MOLECULAR CHARACTERIZATION OF IS1301 IN AN EMERGENT CLONE OF SEROGROUP C NEISSERIA MENINGITIDIS

by

Ashley M. Conley

BA, Saint Anselm College, 2006

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2008

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Ashley M. Conley

It was defended on

May 27, 2008

and approved by

Thesis Advisor:
Lee H. Harrison, MD
Professor
Department of Medicine, Department of Infectious Diseases and Microbiology
School of Medicine, Graduate School of Public Health
University of Pittsburgh

Committee Member:
Elodie Ghedin, PhD
Assistant Professor
Department of Medicine
School of Medicine
University of Pittsburgh

Committee Member:
Paul Kinchington, PhD
Associate Professor
Department of Ophthalmology
School of Medicine
University of Pittsburgh

Committee Member:
Jeremy Martinson, PhD
Assistant Professor
Department of Infectious Diseases and Microbiology
Graduate School of Public Health
University of Pittsburgh

ACKNOWLEDGEMENTS

I would like to thank Johannes Elias and Ulrich Vogel from the Institute for Hygiene and Microbiology in Germany for providing the reference strain, DE9246. Many thanks go to Dr. Paul Kinchington for discussing various options for locating the integration sites of IS1301 and for the creation of a library to screen the *Neisseria meningitidis* genome for IS1301 integration sites. I would also like to thank Jay DePasse from the laboratory of Dr. Elodie Ghedin and Matthew Keyser from DNASTAR for their assistance in working with the whole genome sequence files. In addition, Dr. Elodie Ghedin's laboratory provided the random hexamers.

MOLECULAR CHARACTERIZATION OF IS1301 IN AN EMERGENT CLONE OF SEROGROUP C NEISSERIA MENINGITIDIS

Ashley M. Conley, M.S.

University of Pittsburgh, 2008

Neisseria meningitidis is a leading cause of invasive meningococcal disease and humans are the only known host. The administration of meningococcal vaccines has reduced the number of meningococcal cases and carriage rates in humans. Current vaccine strategies target important immunological determinants. Insertion sequence 1301 (IS1301) has been shown to facilitate evasion of the host immune response by disrupting antigen expression. The public health importance of this study is in the design of future vaccines against *N. meningitidis* and in understanding the emergence of new clones.

In the 1990s there was an increase in serogroup C meningococcal disease in Maryland that was associated with antigenic shift at the *fetA* gene. The isolates were characterized as either an early clone or late clone based on the outer membrane protein sequence profiles. The 2:P1.5,2:F.1-30 sequence profile is classified as an early clone while the 2:P1.5,2:F.3-6 sequence profile is classified as a late clone. Previous studies determined that the late clone contained IS1301, while the early clone did not. The goal of this present study is to characterize the IS1301 insertion sites in the late clone to determine if this genetic element contributed to clonal emergence. Early and late clone isolates were characterized by DNA sequence analysis of the housekeeping gene, *fumC*. A single nucleotide polymorphism characteristic of the hypervirulent ET-15 clone was identified in the late clone isolates. Southern blot analysis using an IS1301 probe revealed a heterogeneous population with multiple insertion sites, ranging from five to ten

insertions, within the serogroup C late clone genomes. Of note was a high molecular weight triplet banding pattern common to the majority of isolates.

Several different IS1301 specific, PCR-based strategies were performed in an attempt to clone the IS1301 elements corresponding to the bands of the IS1301 triplet. In addition, whole genome sequence analysis was performed on one of the late clone isolates. Initial whole genome analysis demonstrates IS1301 integration within an opacity-associated protein (Opa), which promotes adherence to host cells. Further investigations are necessary to determine the effect of IS1301 insertion on antigenic variation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank Dr. Lee Harrison, my principal investigator, for being an inspirational teacher and mentor. A special thank you goes to Dr. Jane W. Marsh for her constant support, encouragement, advice and guidance while I was doing my research. I want like to thank my committee members, Dr. Elodie Ghedin, Dr. Paul Kinchington, and Dr. Jeremy Martinson, for their support, answering my e-mails, and attending multiple meetings throughout the year. I would like to thank my family, in particular my mom, for always inspiring me to reach my goals and to dream big. A big thank you goes to two special friends, Krisztina Baglyas and Shaylee O'Leary; thank you for becoming wonderful friends and for sharing this journey with me, it would not have been the same without the many study sessions, coffee breaks, and fun times we had together the past two years. Most of all, I want to thank my wonderful husband, Michael, for being there when I needed you, listening to me, and giving me hope; I would not have made it without you. Last but not least, I have met and become friends with many people since the start of this program, especially within the Infectious Diseases and Microbiology Department and the Public Health Infectious Diseases Laboratory; I wish you the best of luck in your futures.

1.0 INTRODUCTION

1.1 EPIDEMIOLOGY OF NEISSERIA MENINGITIDIS

Humans are the only known host to carry *Neisseria meningitidis*, a Gram-negative diplococcus that can colonize the oro-pharyngeal mucosa in a healthy human being without causing disease. Up to 20% of the population carries the organism in the nasopharynx; however, in some cases *N. meningitidis* invades the pharyngeal mucosal epithelium, causing invasive meningococcal disease, commonly characterized by septicemia and meningitis (Davidsen & Tonjum, 2006). *N. meningitidis* is contagious and can be spread when an infected individual exchanges respiratory secretions with another individual. In 2006 in the United States, the average incidence was 0.30/100,000, children less than one year-old had the highest rate of invasive disease and the dominant serogroup was Y (Table 1). Although the incidence of meningococcal disease is relatively low compared to other pathogens, the mortality and morbidity rate is high with 10% fatality. Long term sequelae include seizures, mental retardation, deafness, and amputation ("Active Bacterial Core Surveillance," 2006; "Meningococcal Disease: Frequently Asked Questions," 2008; Pollard, 2004). There are certain high risk groups associated with meningococcal disease including college students living in dormitories, refugees, splenectomy patients, microbiologists that work with isolates of *N. meningitidis* in the laboratory, and

household contacts of case patients ("Active Bacterial Core Surveillance," 2006; "Meningococcal Disease: Frequently Asked Questions," 2008).

Table 1. Rates of invasive meningococcal disease by age and serogroup using Active Bacterial Core Surveillance data gathered in 2006.

	Serogroups							
Age	B		C		Y		Other‡	
(years)	No.	(Rate*)	No.	(Rate*)	No.	(Rate*)	No.	(Rate*)
< 1	13	(2.6)	0	(0.0)	3	(0.61)	0	(0.0)
1	0	(0.0)	0	(0.0)	0	(0.0)	1	(0.2)
2-4	2	(0.13)	0	(0.0)	0	(0.0)	1	(0.07)
5-17	1	(0.02)	8	(0.12)	7	(0.11)	0	(0.0)
18-34	8	(0.09)	11	(0.13)	7	(0.08)	0	(0.0)
35-49	3	(0.04)	8	(0.10)	2	(0.02)	0	(0.0)
50-64	1	(0.02)	2	(0.03)	10	(0.16)	1	(0.02)
≥ 65	2	(0.05)	0	(0.00)	13	(0.32)	4	(0.09)
Total †	30	(0.08)	29	(0.08)	42	(0.12)	7	(0.02)

(*Active Bacterial Core Surveillance Report, Emerging Infections Program Network, Neisseria meningitidis*, 2007)

Note:
Unknown serogroup (n=6) distributed amongst known
*Cases per 100,000 population for ABCs areas excluding Oregon
† All rates excluding Oregon; Rates including Oregon are serogroup B 0.11, serogroup C 0.09, serogroup Y 0.14, and serogroup Other 0.04 cases per 100,000 population for ABCs areas
‡ Other includes serogroup W-135 and non-groupables

Furthermore, *N. meningitidis* is a highly recombinant bacterium, so multiple methods have been developed for classifying the invasive and carrier isolates from individuals. A broad method of characterizing *N. meningitidis* is assigning a serogroup based on the composition of the polysaccharide capsule. There are thirteen serogroups, four of which, A, B, C, and Y, cause a majority of meningococcal cases. *N. meningitidis* strains can be further classified into serosubtype based on their class 1 outer membrane proteins (PorA), serotype based on class 2 or

3 outer membrane proteins (PorB) and immunotype based on the lipooligosaccharides (Rosenstein, Perkins, Stephens, Popovic, & Hughes, 2001).

The outer membrane proteins (OMPs) of *N. meningitidis* are important virulence factors and play a major role in vaccine design and host response. PorA and PorB are two outer membrane proteins and major porins that are recognized by the host immune response and, as outlined above, are important in meningococcal typing. Their main function is to create pores that allow passage of small hydrophilic solutes into the host's cells (Harrison et al., 2006; Rosenstein et al., 2001). In addition, the opacity-associated proteins (Opa) promote adherence to host cells and are a key factor in the pathogenesis of *N. meningitidis*. The Opa proteins vary in their size, antigenicity, and expression in the genome (Rosenstein et al., 2001; Tzeng & Stephens, 2000). Other virulence factors include the pili, which are responsible for adherence to epithelial cells, the meningococcal lipooligosaccharide (LOS), which inhibits serum bactericidal activity, the capsular polysaccharide, which protects the bacteria and enhances survival, and NadA, an adhesin protein and potential candidate for a serogroup B vaccine (Capecchi et al., 2005; Tzeng & Stephens, 2000).

1.2 MOLECULAR TYPING METHODS

Various molecular typing methods are used to understand the genetic characteristics of *N. meningitidis*, including multilocus enzyme electrophoresis (MLEE), multilocus sequence typing (MLST), antigen gene sequence typing of the outer membrane proteins (OMP), and pulse field gel electrophoresis (PFGE). On a smaller, local scale, PFGE is often used to look at outbreaks of *N. meningitidis* to identify related cases. However, it cannot be easily used to analyze global

isolates because the method is highly subjective and the data are difficult to compare between laboratories (Cameron & Tsang, 2007).

Globally, two methods, MLEE and MLST, are frequently used to distinguish genetic differences. MLEE was the first technique used to help map the global spread of meningococcal disease using genotypes and played a role in designing "tailor-made vaccines". In MLEE, fifteen to twenty loci are analyzed and each isolate is assigned an electrophoretic type (ET), such as ET-15. Due to the complex nature of MLEE, and the large number of loci, reference laboratories were unwilling to adopt it for routine typing when it was first developed. MLST has replaced MLEE to determine the genetic lineage of meningococci. MLST uses seven housekeeping genes to assign a sequence type (ST) and is widely used in epidemiological studies to genetically compare multiple isolates (Jolley, Brehony, & Maiden, 2007).

1.3 CURRENT VACCINES

Meningococcal vaccines are an important part of preventing the spread of meningococcal disease and decreasing the morbidity and mortality in high risk groups. Currently, there are two licensed vaccines in the United States, MPSV4, a polysaccharide vaccine, and MCV4, a conjugate vaccine ("Meningococcal Disease: Frequently Asked Questions," 2008; Snape et al., 2008).

MPSV4 (Menomune) protects against serogroups A, C, W-135 and Y, and was licensed in 1982 to be used in individuals $\geq$ 2 years of age. However, the duration of vaccine protection is limited to approximately 3-4 years due to poor immunological memory. In addition, MPSV4 does not provide substantial herd immunity because it does not prevent carriage in individuals. Polysaccharide vaccines use the purified capsular polysaccharides from various *N. meningitidis*

serogroups to provide protection from disease by producing serum antibodies to activate the complement system (Snape et al., 2008).

On the other hand, conjugate polysaccharide vaccines use polysaccharides that are covalently linked to a protein carrier to generate T cell-dependent responses and a better immunological memory than the MPSV4. MCV4 (Menactra) is a conjugate vaccine, licensed in 2005, that covers the same serogroups as MPSV4 and provides herd immunity by preventing carriage. Currently, MCV4 is only licensed for individuals ages 11 to 55, but current research demonstrates a tetravalent conjugate vaccine may be helpful in preventing invasive meningococcal disease in infants (Harrison, 2008; Snape et al., 2008).

At this time, there is no vaccine to prevent disease from *N. meningitidis* serogroup B, due to immunologic reactivity with human neuronal tissue. Unfortunately, serogroup B strains affect a large portion of meningococcal cases in Europe and the Americas and cause most cases in infants. Progress is being made in existing research on developing meningococcal vaccines that include serogroup B and vaccines that can be safely administered to infants (Harrison, 2006).

1.4 SEROGROUP C ANTIGENIC SHIFT

By using horizontal gene transfer, capsular switching, and antigenic variation, *N. meningitidis* can evade the host immune response. In a previous study of isolates from 1992 until 2001, it was observed that the outer membrane profile of the serogroup C isolates was split between an early and late clone OMP profile. The early clone had a profile of 2:P1.5,2:F.1-30 (serotype:serosubtype:immunotype) and was mostly seen in the early 1990s, while the late clone had a profile of 2:P1.5,2:F.3-6 and was seen in the mid-1990s (Figure 1). FetA for the early and

late clones differed greatly and in the variable region of FetA, there was a 28 amino acid difference. The increase in cases in the mid 1990s was partly attributed to antigenic shift, the *porA* deletions and the change in the *fetA* gene (Harrison et al., 2006).

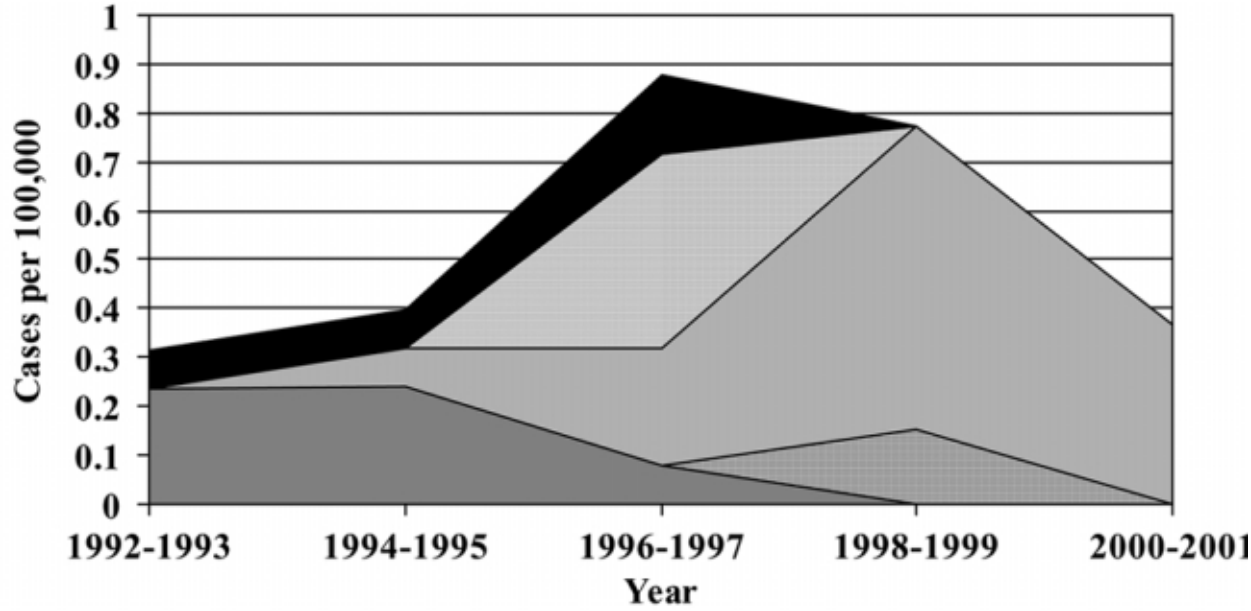

Figure 1: The average annual incidence of serogroup C infection in Maryland from 1992-2001 according to outer membrane sequence profile in 15-24 year olds.
The dark gray area represents infections caused by the early clone (ST-11 complex, 2:P1.5,2:F.1-30), the light gray area represents infections caused by the late clone (ST-11 complex, 2:P1.5,2:F.3-6), and the black area represents infections caused by isolates with other OMP sequence profiles. The dotted areas represent the subset of early and late strains with *por*A deletions (Harrison et al., 2006).

1.5 ET-15 CLONE

The electrophoretic type 15 (ET-15), a clone of ET-37, arose from Canada in the early 1990s, spread worldwide and caused an increase in virulence, rates of mortality and rates of sequelae in comparison to the ET-37 clone. There is a clone specific point mutation in the enzyme for fumerase (*fumC* gene), an A instead of a G, that distinguishes the ET-15 clone from other closely related clones, such as ET-37 (Tsang, Law, Henderson, Blake, & Stoltz, 2006; Tyler & Tsang, 2004; Vogel, Claus, Frosch, & Caugant, 2000).

1.6 INSERTION SEQUENCE 1301

Insertion sequences are mobile elements that are involved in genetic rearrangement and can insert into genomes, cause deletions, knockout gene expression or deregulate downstream genes. This causes a wide variety of genetic characteristics that affects not only the bacteria, but the host the bacterium infects. IS1301 is approximately 800 base pairs (bp), has two overlapping reading frames and is flanked by 19bp inverted repeats. It has been found in serogroups B, C, Y, X, 29E, and W135, but not in other *Neisseria* species such as *Neisseria gonorrheae* or apathogenic *Neisseria* species (Hilse, Hammerschmidt, Bautsch, & Frosch, 1996). Based on similarities and differences in the nucleotide and protein sequences, IS1301 falls into the *IS5*-family, group IS*427,* of insertion sequences (Mahillon & Chandler, 1998).

In one study, IS1301 inactivated the *siaA* gene, causing a capsule phase variation in *N. meningitidis* serogroup B, which promoted a stronger adherence and increased entry of meningococci into epithelial cells (Hammerschmidt et al., 1996; Hilse et al., 1996). Further research has shown IS1301 to insert into *porA*, *OatWY*, and into meningococcal capsule operons in nongroupable carrier isolates, such as *synA* and *ctrA* (Claus et al., 2004; Dolan-Livengood, Miller, Martin, Urwin, & Stephens, 2003; Jelfs, Munro, Wedege, & Caugant, 2000; Vogel et al., 2000).

Recently, Johannes Elias and Ulrich Vogel published a study where they developed an IS1301-restriction fragment length polymorphism (RFLP) method to be used in epidemiological settings to differentiate ET-15 meningococci. Also, ten genomic insertion sites were characterized in the DE9246 strain to analyze the genetic variability of IS1301. Of note is the insertion of IS1301 into *nadA*, and intragenic insertions in the open reading frames homologous

to MafB and TspB. MafB is an adhesin and TspB is a T-cell and B-cell stimulating antigen (Elias & Vogel, 2007).

2.0 SPECIFIC AIM

During the 1990s, there was an increased incidence of serogroup C *N. meningitidis*, including three outbreaks in Maryland. An early and late clone was identified with the emergence of an ST11 clone that underwent antigenic shift at the *fetA* gene locus. The late clone, not the early clone, is known to carry IS1301. The specific aim of this project is to determine if IS1301 may have contributed to the clonal emergence of the serogroup C late clone through genetic disruption of immunogenic determinants.

3.0 MATERIALS AND METHODS

3.1 STUDY ISOLATES

The isolates used in this study were collected by the Maryland Active Bacterial Core Surveillance (ABCs) from January 1, 1992 to December 31, 2001 as part of the multistate Emerging Infections Program (EIP). Thirty-nine late clone and four early clone isolates were used for this study (Table 2). The isolates were collected from a normally sterile body fluid from Maryland residents, grown on chocolate agar plates and incubated overnight at 37°C and 5% CO_2. The thirty-nine late clone isolates have the subcapsular genotype (2:P1.5,2:F.3-6) and the four early clone isolates have the subcapsular genotype (2:P1.5,2:F1-30) (see section 1.4). In a previous study, the isolates from the ABCs were analyzed by MLST and confirmed to be ST-11 strains (Harrison et al., 2006). The reference isolate, DE9246, was provided by Ulrich Vogel from the Institute for Hygiene and Microbiology at the University of Würzburg, Würzburg, Germany.

Table 2. List of isolates used in this study.

Specimen ID	Culture Date	ST	Clonal Complex	porA VR1	porA VR2	fetA
NM00012	4/20/1992	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00023	10/26/1992	11	ST-11 complex/ET-37 complex	5	2	F1-30
NM00029	2/24/1993	11	ST-11 complex/ET-37 complex	5	2	F1-30
NM00035	4/6/1993	11	ST-11 complex/ET-37 complex	5	2	F1-30
NM00062	2/5/1994	11	ST-11 complex/ET-37 complex	5	2	F1-30
NM00078	1/25/1995	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00084	5/9/1995	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00094	9/30/1995	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00095	10/7/1995	1988	ST-11 complex/ET-37 complex	5	2	F3-6
NM00105	1/26/1996	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00111	3/25/1996	11	ST-11 complex/ET-37 complex	deletion	deletion	F3-6
NM00126	6/19/1996	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00133	7/25/1996	11	ST-11 complex/ET-37 complex	deletion	deletion	F3-6
NM00137	10/30/1996	11	ST-11 complex/ET-37 complex	deletion	deletion	F3-6
NM00144	1/22/1997	11	ST-11 complex/ET-37 complex	deletion	deletion	F3-6
NM00146	1/10/1997	11	ST-11 complex/ET-37 complex	deletion	deletion	F3-6
NM00150	2/13/1997	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00157	3/20/1997	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00162	5/10/1997	11	ST-11 complex/ET-37 complex	deletion	deletion	F3-6
NM00163	5/14/1997	11	ST-11 complex/ET-37 complex	deletion	deletion	F3-6
NM00184	4/13/1998	11	ST-11 complex/ET-37 complex	deletion	deletion	F3-6
NM00199	2/16/1999	11	ST-11 complex/ET-37 complex	deletion	deletion	F3-6
NM00201	2/22/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00202	2/23/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6

Table 2 Continued.

NM00205	3/6/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00208	3/19/1999	2959	ST-11 complex/ET-37 complex	5	2	F3-6
NM00214	4/24/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00216	5/20/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00217	5/22/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00218	5/26/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00222	7/4/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00230	8/20/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00237	11/29/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00239	12/25/1999	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00242	1/5/2000	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00245	3/5/2000	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00259	9/14/2000	50	ST-11 complex/ET-37 complex	5	2	F3-6
NM00260	9/15/2000	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00263	12/12/2000	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00270	1/26/2001	2960	ST-11 complex/ET-37 complex	5	2	F3-6
NM00272	2/1/2001	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00288	5/8/2001	11	ST-11 complex/ET-37 complex	5	2	F3-6
NM00298	8/23/2001	11	ST-11 complex/ET-37 complex	5	2	F3-6

Note: Isolates in red indicate “early clone”.

3.2 SEQUENCING FOR ET-15

All the isolates were subjected to DNA sequencing of the *fumC* gene to reveal the presence of a single nucleotide polymorphism to indicate an ET-15 clone. The primers P1B (5'-ATCCCCGCCGTAAAAGCCCTGC-3') and P2B (5'-TCAACCCGAACGACTGCCCG-3') were used in a polymerase chain reaction (PCR) to amplify the *fumC* gene. The reaction conditions were 94ºC for 2 minutes, followed by forty cycles of amplification (95ºC for 1 minute, 55ºC for 1 minute, and 72ºC for 1 minute). The resulting products were run on a 1% agarose gel containing ethidium bromide to visualize the PCR product. The PCR product was cleaned using EXOSAP-IT (Applied Biosystems, Ipswich, Massachusetts) to remove excess dNTPs and primers. An ABI PRISM® Big Dye Terminator Cycle Sequencing Kit v3.1 (Applied Biosystems) was used with the sequencing primers, P1B and S2 (5'-AGTCGCCAAAACCGCCTACAA-3') and the sequence reaction conditions were as follows: 95ºC for 5 minutes, then 30 cycles of 95ºC for 1 minute, 55ºC for 1 minute, and 72ºC for 1 minute, followed by 72ºC for 7 minutes. The product was cleaned using sodium acetate and ethanol and analyzed by capillary electrophoresis at the University of Pittsburgh Genomics and Proteomics Core Laboratories on an ABI 3730 DNA analyzer. The data was analyzed using SeqMan Pro software (DNASTAR Lasergene 7.2, Madison, Wisconsin).

3.3 IDENTIFICATION OF IS1301 IN LATE CLONE ISOLATES

To confirm the forty late clone isolates had IS1301, a PCR reaction was performed using primers SH42 (5'-TTGAGCTAGGGTCATGG-3') and SH46 (5'-AAATCAGGGTTAGGTTTCTT-3') on the thirty-nine late clone isolates, four early clone isolates, and the reference strain. AmpliTaq Gold® (Applied Biosystems) was used with the following reaction conditions: 95ºC for 10 minutes, followed by forty cycles of amplification (95ºC for 1 minute, 55ºC for 1 minute, and 72ºC for 1 minute) and 7 minutes at 72ºC. The resulting products were run on a 1% agarose gel containing ethidium bromide to visualize the PCR product.

3.4 PREPARATION OF THE IS1301 PROBE

PCR product that was acquired using primers SH42 and SH46 to identify IS1301 in the late clone (see section 3.3) was used to create an IS1301 probe. A QIAquick® PCR Purification Kit (QIAGEN, Valencia, California) was used to purify the PCR product which was then labeled using the DIG High Prime DNA Labeling and Detection Starter Kit II (Roche Diagnostics, Indianapolis, Indiana). The kit utilizes digoxigenin-dUTP (DIG), a steroid hapten, to label the DNA probe for hybridization. The DIG High Prime reagent contains a 5x concentration mixture of random hexamers, dNTPs, and Klenow, to randomly label the DNA probe. Labeling efficiency of the DIG labeled probe was compared to control DIG labeled DNA (provided in the kit) and approximately 25ng/µl of DIG labeled probe was used for the hybridization procedure.

3.5 DNA EXTRACTION AND SOUTHERN HYBRIDIZATION

All the isolates were subject to isolation of genomic bacterial DNA by phenol:chloroform extractions. *N. meningitidis* was harvested from overnight growth on 5% chocolate agar and incubated with 25µl lysozyme (30mg/ml), 5µl proteinase K (20mg/ml), and 10µl RNase (5mg/ml) prior to extraction with phenol:chloroform:isoamyl alcohol (25:24:1 [v/v] saturated with 10mM Tris-HCl, 1mM EDTA, ph 8.0) and chloroform:isoamyl alcohol (24:1 [v/v] saturated with 10mM Tris-HCl, 1mM EDTA, pH 8.0). The DNA was precipitated using 0.1 volumes of 3M sodium acetate (pH 5.3) and 2.5 volumes of 100% ethanol. The extracted genomic DNA was stored at –20°C. Prior to the Southern blot, 2µg genomic DNA was digested with one unit of *Hinc*II in 1X buffer 3 (New England Biolabs, Ipswich, Massachusetts). A 0.8% agarose gel was run with approximately 2µg of restricted *Hinc*II DNA from each isolate. A DIG-labeled DNA Molecular Weight Marker III (Roche Diagnostics) was included in the gel and ranged from 21kb to 125bp. The 0.8% agarose gel was incubated for 30 minutes with 0.25M HCl, followed by 30 minutes in a denaturation solution (1.5M NaCl/0.5 M NaOH), and 30 minutes in a neutralization solution (1.5M NaCl/0.5M Tris-HCl, pH 7.0). A pre-Southern blot image was taken of the agarose gel to visualize the digested DNA.

The DNA was transferred by a standard Southern blotting technique using a positively charged nylon membrane (Roche Diagnostics). The nylon membrane was UV crosslinked with the UV Stratalinker 2400 (Stratagene, La Jolla, California), then air dried. A post-Southern blot transfer image was taken of the 0.8% agarose gel to ensure the DNA adequately transferred to the membrane. The membrane was then pre-hybridized with DIG hybridization granules for 30 minutes at 41ºC and hybridized overnight with the DIG-labeled IS1301 probe. The membrane was put through several stringency washes using 2x and 0.5x sodium citrate, 0.1% SDS, to

remove the unbound DIG-labeled IS1301 probe and then incubated with a blocking solution (included in the kit). Then, the membrane was incubated with anti-digoxigenin-alkaline phosphate conjugate for 30 minutes and subsequently washed using washing buffer (included in kit) to remove unbound antibody. CSPD (Disodium 3-(4-methoxyspiro (1,2-dioxetane-3,2´-(5'-chloro)tricyclo [3.3.1.1^{3,7}]decan)-4-yl)phenyl phosphate), a chemiluminescence substrate that emits light at a wavelength of 477nm, was placed on the membrane and incubated for ten minutes at 37ºC. Lastly, the membrane was exposed to x-ray film for thirty minutes, developed, and analyzed using BioNumerics software (Applied Maths, Austin, Texas).

3.6 EXTRACTION OF HIGH MOLECUAR WEIGHT IS1301 *HINC*II TRIPLET

Thirty-nine out of the forty late clone isolates displayed a high molecular weight IS1301 triplet on the Southern blot analysis. Finding the integration sites of the three IS1301 fragments was of high epidemiological significance. To increase the likelihood of cloning the IS1301 integration sites, the high molecular weight *Hinc*II IS1301 triplet observed by Southern blot was extracted from a 0.8% agarose gel using the QIAquick® Gel Extraction Kit (QIAGEN). Briefly, 10µg of *Hinc*II digested DNA (NM126 or NM263) was run on a 0.8% agarose gel and the high molecular weight IS1301 triplet was extracted using a clean scalpel. The QIAquick® Gel Extraction Kit was used according to manufacturer's protocols and the DNA was eluted in 50µl of elution buffer. The extracted DNA was spot dialyzed against H_2O to remove excess salt and stored at -20ºC for future experiments. To ensure the triplet was adequately extracted from the gel, a Southern blot using the gel extracted DNA was performed to visualize the high molecular weight IS1301 triplet.

3.7 RANDOM HEXAMER STRATEGY

The random hexamer strategy was designed to conduct a large-scale screening process using a random PCR amplification protocol. We used a random hexamer and paired it with a specific primer that was anchored in IS1301 to amplify the flanking regions of the inserted element. Briefly, a random hexamer tagged with a specific primer was incubated with different late clone DNA preparations in the presence of dNTPs and 10x Ecopol buffer at 94ºC for two minutes. After that, the reaction was placed on ice for two minutes followed by the addition of Klenow Fragment (3'->5' exo-) (New England Biolabs) and incubated at 37ºC for one hour. This step was repeated once and then the enzyme was inactivated at 75ºC for 10 minutes. Next, 5µl of this product was used in a PCR reaction with a primer that was specific to the tagged primer but lacked the 6nt random 3' end (Table 3). In this strategy, three forms of *N. meningitidis* DNA were used: late clone whole genomic DNA, *Hinc*II digested genomic DNA, and DNA from the high molecular weight triplet to see how the results would differ, to create a greater specificity for IS1301 and to reduce the amount of background product.

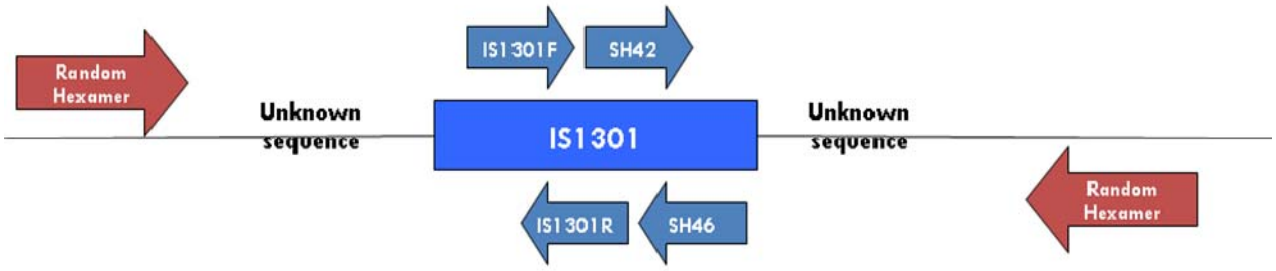

Figure 2: A schematic of the random hexamer strategy (primers not drawn to scale).
This strategy incorporates a random hexamer and known IS1301 specific primers to amplify the unknown sequence, or integration sites.

Table 3. Primer sequences for random hexamer strategy.

Random Hexamer	Primer Sequence
Tag1-6N	5′-ACGAGTGCGTCGTGTCTCTA–NNNNNN-3'
Tag2-6N	5'-ACGCTCGACATAGTATCAGC–NNNNNN-3'
Tag3-6N	5'-AGCACTGTAGTCTCTATGCG–NNNNNN-3'
Tag4-6N	5'-AGACGCACTCATCAGACACG–NNNNNN-3'
Random Hexamer Specific Primer	**Primer Sequence**
Tag1	5'-ACGAGTGCGTCGTGTCTCTA-3'
Tag2	5'-ACGCTCGACATAGTATCAGC-3'
Tag3	5'-AGCACTGTAGTCTCTATGCG-3'
Tag4	5'-AGACGCACTCATCAGACACG-3'
IS1301 Specific Primers	**Primer Sequence**
IS1301F	5′-AGAACCGCAATAACTGACAACATA-3′
IS1301R	5′-CAAGAAGCCCAGTATTTTATCGC-3′
SH42	5′-TTGAGCTAGGGTCATGG-3′
SH46	5′-CAAGAAACCTAACCCTGATTT-3′

3.7.1 Experiment 1 – Late Clone Whole Genomic DNA

Briefly, 1µg of genomic DNA from a late clone isolate was added to 2.5µl of 10x buffer 2 (Applied Biosystems), 1µl of each dNTP (10mM), and 2µl of 10µM random hexamer. Four reactions were set up using the four different tagged primers listed in Table 3. The reaction was incubated at 94ºC for 2 minutes and then placed on ice for 2 minutes. Next, 0.5 units of Klenow Fragment (3'-> 5' exo-) (New England Biolabs) was added and incubated for one hour at 37ºC. This step was repeated once, followed by inactivation of the enzyme by incubation at 75ºC for 10 minutes. Subsequently, 5µl of product was amplified with an IS1301 specific primer, IS1301F and a random hexamer specific primer (Table 3) for round 1 of PCR. The reaction conditions were 94ºC for 10 minutes, followed by forty cycles of amplification (94ºC for 1 minute, 50ºC for 1 minute, and 72ºC for 1.5 minutes) and 7 minutes at 72ºC. The resulting products were run on a 1% agarose gel containing ethidium bromide to visualize the PCR product. A second round of PCR using a nested IS1301 primer, SH42 and the random hexamer specific primer (Table 3) was

performed on 5µl PCR product from the first round of PCR to increase specificity for the IS1301 fragments. The resulting products from round 2 of PCR were run on a 1% agarose gel containing ethidium bromide to analyze the PCR product.

3.7.2 Experiment 2 – *Hinc*II Digested Late Clone Whole Genomic DNA

In the second experiment, the same protocol was followed; however, *Hinc*II digested genomic DNA was used instead of whole genomic DNA. Four reactions were set up using the four different tagged primers listed in Table 3. As mentioned above, 5µL of product from each reaction was used in a first round of PCR using IS1301F, IS1301R and the random hexamer specific primers for a total of eight PCR reactions. The resulting products were run on a 1% agarose gel containing ethidium bromide to analyze the PCR product. A second round of PCR using the IS1301 primers, SH42 or SH46, and the random hexamer specific primer, was performed on the PCR product from the first round to increase specificity for the IS1301 fragments. The resulting products from round 2 of PCR were run on a 1% agarose gel containing ethidium bromide to analyze the PCR product.

3.7.3 Experiment 3 – Gel Purified High Molecular Weight IS1301 Triplet

The third experiment used the DNA from the high molecular weight IS1301 triplet that was extracted from the 0.8% agarose gel as template for random priming, instead of genomic DNA or *Hinc*II digested genomic DNA. Out of all three experiments, the extracted triplet should provide the most specificity of amplifying the IS1301 integration sites of the triplet because there is a smaller amount of background DNA. Once again the same protocol was followed, however,

only two reactions using the Tag1-6N, IS1301F, and IS1301R primers were set up due to the limited availability of the IS1301 triplet DNA. 5µL of the random primed product was added to the round 1 PCR reaction using the Tag1 primer and both the IS1301F and IS1301R primers for a total of two PCR reactions. The resulting products were run on a 1% agarose gel containing ethidium bromide to analyze the PCR product. Round 2 PCR used a second IS1301 primer, SH42 or SH46, and the Tag1 primer on 5µl of the PCR product from the first round of PCR. The resulting products were run on a 1% agarose gel containing ethidium bromide to analyze the PCR product.

3.7.4 TOPO TA Cloning® of PCR Product

The TOPO TA Cloning® Kit for Sequencing (Invitrogen, Carlsbad, California) was used to clone the PCR products from both round 1 and round 2 PCR (Table 4) (Note: the PCR product from experiment 1 was not used because the PCR product from the whole genomic DNA is not as specific for IS1301 as the *Hinc*II digested genomic DNA and the IS1301 triplet DNA). In this kit, PCR product can be directly inserted into the plasmid, pCR®4-TOPO® without ligase or PCR specific sequences by using a single 3' thymidine overhang and topoisomerase that is covalently bound to the vector. To each 3' end of the PCR product, a single deoxyadenosine is added using *Taq* polymerase to allow the PCR product to easily ligate to the vector (Invitrogen, 2006). Included in the kit are the necessary tools to make a positive control using control primers and control DNA. Per manufacturer's instructions, the positive control was constructed via PCR and run on a 1% agarose gel to ensure the positive control showed a band at 750bp (data not shown). This positive control was used in the ligation and transformation of One Shot® TOP10 competent cells. A negative control was used in the ligation step, consisting of a "vector-

only" control, to ensure the selection criteria were working correctly. In addition, pUC19 was transformed into the One Shot® TOP10 competent cells to check the competency of the cells.

3.7.4.1 Cloning Experiment 1

For the first cloning experiment, 4µl of PCR product from each of the eight PCR reactions (Table 4) was ligated to 1µl of TOPO® vector (10ng/µl). 2µl of the TOPO Cloning reaction was added to One Shot® TOP10 competent cells (~70µl) and transformed using a BioRad MicroPulser™. After electroporation, 250µl of S.O.C. medium (supplied in kit) was added to the reaction and incubated for one hour at 37ºC with shaking to allow expression of the antibiotic resistance genes. The cells were plated onto 50µg/ml kanamycin (Sigma) Luria-Bertani (LB) agar plates and incubated overnight at 37ºC, 5% CO_2. For the transformation control, 2µl of pUC19 (10pg/µl) was transformed into One Shot® TOP10 competent cells, incubated for one hour at 37ºC with shaking and plated on 100µg/ml ampicillin (Sigma) Luria-Bertani (LB) agar plates and incubated overnight at 37ºC, 5% CO_2.

A random selection of kanamycin resistant colonies from each reaction were sub-cultured and screened in a colony PCR reaction for IS1301 using IS1301 specific primers. PCR product was electrophoresed on a 1% agarose gel to visualize the resulting product. IS1301 positive colonies were then amplified with the M13F (5'-GTAAAACGACGGCCAG-3') and M13R (5'-CAGGAAACAGCTATGAC-3') primers or T3 (5'-ATTAACCCTCACTAAAGGGA-3') and T7 (5'-TAATACGACTCACTATAGGG-3') primers, that are specific to the vector, to ascertain the size of the cloned insert. The PCR product was electrophoresed on a 1% agarose gel to visualize the product. Colonies that were positive for IS1301 and had an insertion size larger than 300bp were suspended in kanamycin (Sigma) LB/broth and incubated overnight on a shaker at 37ºC. Next, the DNA was extracted by plasmid miniprep. Briefly, overnight cultures were

resuspended in plasmid extraction buffer (25mM Trizma base, 50mM glucose, 10mM Na_2EDTA), and added to 0.2ml of alkaline lysis mix (1N NaOH, 10% SDS). Then, 0.15ml of 3M sodium acetate (pH 4.8) was added to form a precipitate. This product was pelleted, 400µl of the supernatant was taken and added to 800µl of 95% ethanol and chilled overnight. Next, the solution was pelleted, the supernatant was discarded, the pellet was rinsed with 95% ethanol, sedimented to remove excess ethanol and resuspended in 50µl of 1x Tris-Borate-EDTA buffer (Sigma). This DNA was then sequenced using the M13F and M13R primers as sequencing primers by cycle sequencing using the BigDye Terminator 3.0 Kit followed by capillary electrophoresis on an AB3730 sequence analyzer available through the Genome and Proteomics Core Laboratory (GPCL) at the University of Pittsburgh. The sequence data was analyzed using SeqMan Pro (DNASTAR Lasergene 7.2).

Table 4. Eight PCR reactions used for cloning.

#	Source of PCR Product	Primers Used	Round of PCR (1 or 2)	# Colonies Screened
1	*Hinc*II gDNA	Tag1/IS1301R	1	~52
2	*Hinc*II gDNA	Tag1/IS1301F	1	~12
3	Gel purified IS1301 Triplet	Tag1/IS1301F	1	~18
4	*Hinc*II gDNA	Tag1/SH46	2	~12
5	Gel purified IS1301 Triplet	Tag1/SH46	2	~12
6	Gel purified IS1301 Triplet	Tag1/SH42	2	~6
7	*Hinc*II gDNA	Tag2/IS1301R	1	~6
8	*Hinc*II gDNA	Tag3/IS1301F	1	~31

Note: The reaction in red was also used in cloning experiment 2 (see section 4.4.5).

3.7.5 pGEM®-T Easy Cloning of PCR Product

The pGEM®-T Easy Vector System (Promega, Madison, Wisconsin) and *E. coli* JM109 High Efficiency Competent Cells (Promega) were used to clone the PCR product for the second cloning experiment. In this system, the vector is cut with *Eco*RV and a 3' terminal thymidine is

added to both ends (Promega, 2007). 'A' overhangs are added to the PCR product to ensure efficient ligation into the vector by taking 7µl of PCR product and incubating at 72C for 15 minutes with 1.5µl 2mM dATP, 1.5µl 10x buffer, 1.5µl 10x MgCl, 0.2µl *Taq* Polymerase (Applied Biosystems) and 4.8µl water. For positive controls, insert DNA (included in the kit) and a plasmid with a known IS1301 insert, LMK140 (Kostelnik, 2006), was chemically transformed into the JM109 competent cells. A negative control was used in the ligation step, consisting of "vector only" to ensure the selection criteria were working correctly. In addition, pUC19 was transformed into the JM109 chemically competent cells to check the competency of the cells.

3.7.5.1 Cloning Experiment 2

The round 2 PCR product generated from the random primed reaction of the triplet DNA (Table 4, reaction #5) was electrophoresed on a 0.8% agarose gel and product >500bp was extracted and purified using the QIAquick® Gel Purification Kit (QIAGEN) to preferentially clone larger product that may include IS1301 and the integration sites. Three different molar ratios of insert:vector of the purified product was used in the ligation step. 100ng, 75ng and 25ng of purified product were ligated into the pGEM®-T Easy Vector (50ng) and 2µl of each ligation reaction was added to ~50µl of JM109 cells. The JM109 cells, with the ligation reactions, were incubated on ice for 20 minutes then heat shocked for 45 seconds at 42ºC, followed by 2 minutes on ice. 250µl of S.O.C. medium (supplied with kit) was added and the cells were shaken for 1.5 hours at 37ºC. The cells were plated onto 100µg/ml ampicillin (Sigma) LB agar plates and incubated overnight at 37ºC, 5% CO_2. The colonies were screened for the presence of IS1301 using colony blots and the DIG High Prime DNA Labeling and Detection Starter Kit II (Roche Diagnostics). Briefly, the agar plates were cooled for 30

minutes then a positively charged nylon 82mm membrane disc (Roche Diagnostics) was laid on the agar plate for 1 minute. The membrane was incubated for 15 minutes in denaturation solution, then 15 minutes in neutralization solution and finally 10 minutes in 2x sodium citrate (SSC). Next, the membranes were UV crosslinked, incubated with proteinase K (14mg/ml diluted 1:10 in 2x SSC) at 37ºC and hybridized with an IS1301 probe (see section 3.4) for 30 minutes as 41ºC. The subsequent steps follow the same procedure outlined in section 3.5 for Southern blot detection of IS1301.

IS1301 positive colonies were selected, subcultured onto LB/ampicillin plates, incubated overnight at 37ºC, 5% CO_2, and screened a second time using colony PCR with IS1301 specific primers. IS1301 PCR positive clones were amplified with M13F and M13R primers. A plasmid miniprep was performed on clones with an insert size larger than 300bp, as previously described. The inserts were sequenced using M13F and M13R as sequencing primers in ABI PRISM® Big Dye Terminator Cycle Sequencing Kit v3.1 (Applied Biosystems). Sequence reaction conditions were as follows: 95ºC for 5 minutes, then 30 cycles of 95ºC for 1 minute, 55ºC for 1 minute, and 72ºC for 1 minute, followed by 72ºC for 7 minutes. The product was cleaned using sodium acetate and ethanol and analyzed by capillary electrophoresis at the University of Pittsburgh Genomics and Proteomics Core Laboratories on an ABI 3730 DNA analyzer. The data was analyzed using SeqMan Pro software (DNASTAR Lasergene 7.2, Madison, Wisconsin).

3.8 LINKER BASED STRATEGY

The overall theory surrounding the linker based strategy is to cut the *N. meningitidis* genome as many times as possible, attach linkers and PCR amplify with linker specific and IS1301 specific primers (Figure 3). In this protocol, genomic DNA from a late clone was digested with *Hha*I, which cuts the genome ~15,000x and also cuts within the IS1301 fragment one time (Figures 3). *Hha*I specific linkers were ligated to the *Hha*I cut *N. meningitidis* late clone DNA. A round of PCR was performed using a primer specific to the *Hha*I linker and an inverse IS1301 specific primer (reference Figure 3, Table 6). Resulting PCR product would then be ligated to *Xba*I/*Eco*RI digested pUC19 and screened for IS1301 inserts. An *Eco*RI site and an *Xba*I site were engineered into the *Hha*I linker primer and the inverse IS1301 specific primers respectively, to facilitate cloning of IS1301 flanking regions.

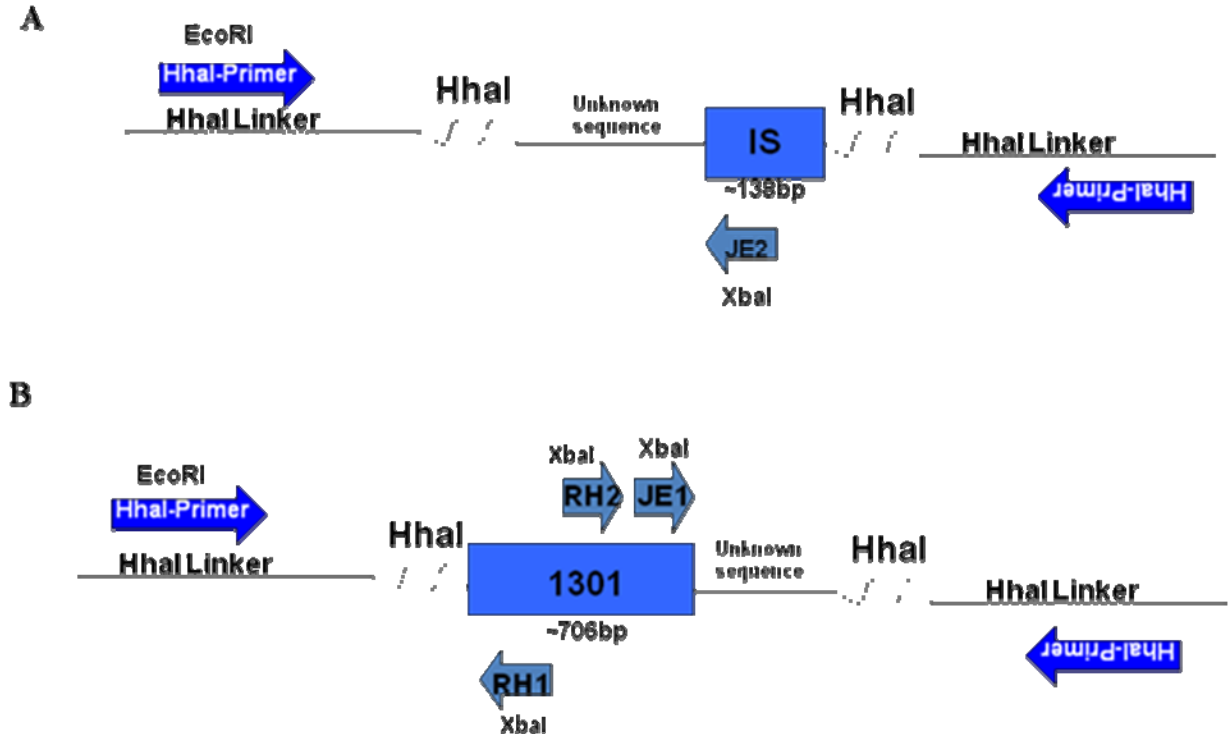

Figure 3: A schematic illustrating the linker based strategy (primers not drawn to scale).

3.8.1 Southern Hybridization – *Hinc*II Digest versus *Hha*I Digest

An IS1301 Southern blot using the same protocol described in section 3.5 was performed to compare the banding pattern of the high molecular weight IS1301 triplet cut with *Hinc*II and *Hha*I to demonstrate the triplet was reduced in size with *Hha*I digestion. Reducing the size of the triplet would help in the cloning procedure since the triplet cut with *Hinc*II is fairly large, approximately 5-10kb. Performing a southern blot on these DNA products allowed visualization of the number and size of the IS1301 fragments in *Hha*I and *Hinc*II cut genomic DNA and the *Hha*I and *Hinc*II cut IS1301 triplet.

3.8.2 Design and Creation of *Hha*I Linkers

The *Hha*I linkers were designed such that annealing of two primers (HhaI-F and HhaI-R) results in the generation the *Hha*I 3' CG overhang. In addition, a HhaI-primer homologous to the 5' end of HhaI-F and containing an *Eco*RI site was designed for subsequent PCR amplification with IS1301 specific primers. To generate the *Hha*I linker, HhaI-F and a HhaI-R primer were combined to a final concentration of 50µM. The primers were boiled for 10 minutes, cooled to room temperature and stored at -20ºC. To visualize the linkers, a 4% high resolution agarose gel was run with 20µl of the HhaI-R primer, HhaI-F primer and the *HhaI* linker to ensure the linker was effectively made (data not shown).

```
HhaI-F                          5′-ATTGACCATCAACTCCCTCGTAGACTGCGTAGCG-3′
HhaI-R                                    3′-TTGAGGGAGCATCTGACGCATC-5′
HhaI-primer       5′-GCGCGAATTCAATTGACCATCAACTCCCTCGT-3′
```

Figure 4: Primer sequences for creation of the *Hha*I linkers.
Primer sequences for HhaI-F, HhaI-R, and the HhaI-primer. The red bases indicate homology to one another, the underlined bases indicate the *Eco*RI restriction site and the green bases indicate the *Hha*I restriction site.

3.8.3 Ligation of *Hha*I Linker to *Hha*I Cut Genomic DNA

2µg of genomic DNA from a late clone (NM272) was digested with one unit of *Hha*I and 1X buffer 4 (New England Biolabs). An aliquot of the digest was run on a 0.8% agarose gel to ensure the genome was adequately digested by *Hha*I (data not shown). To remove excess *Hha*I enzyme from the digest, a phenol:chloroform extraction was performed, the DNA was resuspended in 10µl of deionized water and quantified. In subsequent experiments, the DNA from the *Hha*I digest was purified using a QIAquick® Gel Extraction Kit (QIAGEN) to obtain higher DNA yields. In addition, the cloning vector, pUC19 (0.5µg), was prepared by performing a double digest with *Eco*RI and *Xba*I (New England Biolabs) and run on a 1% agarose gel to ensure the plasmid was cut sufficiently by the enzymes (data not shown).

3.8.4 Experiment 1 - Linker Strategy – Ligation & PCR

Five different linker-ligation reactions were performed using varying amounts of *Hha*I cut genomic DNA, 1ng, 10ng, 25ng, and 50ng. A 50ng control was performed without ligase. The *Hha*I cut genomic DNA was diluted to 5.0ng/µl, 2.5ng/µl and 0.25ng/µl working concentrations to make a 1ng, 10ng, 25ng, and 50ng concentrations of DNA in the ligation reactions. Table 5 depicts the five *Hha*I linker ligation reactions:

Table 5. Linker ligation reagents and DNA concentrations.

	***Hha*I Cut gDNA Concentrations**				
Reagents	**1ng**	**10ng**	**25ng**	**50ng**	**50ng (control)**
***Hha*I Linker (5pmol/µl)**	2	x	x	x	x
***Hha*I Linker (50pmol/µl)**	x	2 (50pmol)	5 (250pmol)	10 (500pmol)	10 (500pmol)
10x Ligase Buffer (µl)	2	2	2	2	2
T4 Ligase (µl)	1	1	1	1	x
DNA (µl)	4 (0.25ng/µl)	4 (2.5ng/µl)	10 (2.5ng/µl)	10 (5ng/µl)	10 (5ng/µl)
dH_20 (µl)	9	10	2	x	x

The reagents were vortexed and incubated at room temperature for 10 minutes. The ligations were purified using the QIAquick® PCR Purification Kit (QIAGEN) (per manufacturer's instructions) to remove excess linker and ligase.

The HhaI-primer with an *Eco*RI site and an IS1301 specific primer, either RH1XBAI or RH2XBAI (Table 6), was used to PCR amplify the *Hha*I linker ligation reactions. 5µl from each of the five purified ligation reactions shown in table 5 were PCR amplified. The following controls were performed:

- Uncut genomic DNA with RH1XBAI and RH2XBAI
- *Hha*I cut genomic DNA (no linker) with RH1XBAI and RH2XBAI
- Uncut genomic DNA with HhaI-primer
- *Hha*I cut genomic DNA (no linker) with HhaI-primer
- *Hha*I cut genomic DNA (ligated linker) with HhaI-primer
- *Hha*I linkers (no gDNA) with either RH1XBAI or RH2XBA1
- *Hha*I linkers (no gDNA, no ligase) with either RH1XBAI or RH2XBA1
- *Hha*I linkers (no gDNA, with ligase) with either RH1XBAI or RH2XBA1

AmpliTaq Gold® (Applied Biosystems) was used with the following reaction conditions: 95ºC for 10 minutes, followed by forty cycles of amplification (95ºC for 1 minute, 55ºC for 1 minute, and 72ºC for 1 minute) and 7 minutes at 72ºC. The resulting products were run on a 1% agarose gel containing ethidium bromide to visualize the PCR product.

Table 6. IS1301inverse primer sequences for use in the linker strategy.

Primer Name	Primer Sequence
RH1XBAI	5′-GTACTCTAGATCTTCCATTGATGACAGCCG-3′
RH2XBAI	5′-GTACTCTAGAATTACGCCACTTAGTCGAGA-3′
JE1XBAI	5′-GCGCTCTAGACGATGCTTTACTTGGCTTGCT-3′
JE2XBAI	5′-GCGCTCTAGACTCGCCATTGTTTGTGTTTGC-3′

Note: Red bases indicate an *Xba*I site.

3.8.5 Experiment 2 Linker Strategy – Ligation & PCR

To reduce the amount of excess linker in the linker ligation PCR reaction, the linker ligation reactions were gel purified. A master mix consisting of a 10 molar excess of *Hha*I linker (620pmol), 10µl 10x ligation buffer, 2µl T4 ligase, and 3µg of gel purified *Hha*I cut genomic DNA was prepared (QIAquick® PCR Purification Kit, QIAGEN). A negative control, without the addition of ligase, was also included. The reactions were incubated overnight at 14ºC. After incubation, the linker ligation was run on a 0.8% agarose gel. Ligation products from 300bp-700bp and >700bp were extracted from the gel using the QIAquick® Gel Extraction Kit (QIAGEN) to remove excess linker. PCR amplification of the 2 sized purified linker ligation reactions was performed as described in section 3.8.4 using linker and IS1301 specific primers.

3.9 INVERSE PCR

The theory behind inverse PCR is depicted in Figure 5 where genomic DNA from a late clone (NM263) was PCR amplified using the inverse RH1XBAI and RH2XBAI primers or the JE1XBAI and JE2XBAI primers using the following conditions: 95ºC for 10 minutes, followed by forty cycles of amplification (95ºC for 1 minute, 55ºC for 1 minute, and 72ºC for 1 minute)

and 7 minutes at 72ºC. PCR product was then electrophoresed on a 0.8% agarose gel containing ethidium bromide to visualize the PCR product. Each band from the PCR product (Figure 27, lane 1) was extracted from the gel and purified using the QIAquick® Gel Extraction Kit (QIAGEN). The PCR product was cleaned using EXOSAP-IT (Applied Biosystems) to remove excess dNTPs and primers. Cylce sequencing was performed as previously described using primers JE1XBAI and JE2XBAI. The product was cleaned using sodium acetate and ethanol and analyzed by capillary electrophoresis at the University of Pittsburgh Genomics and Proteomics Core Laboratories on an ABI 3730 DNA analyzer. The data was analyzed using SeqMan Pro software (DNASTAR Lasergene 7.2, Madison, Wisconsin).

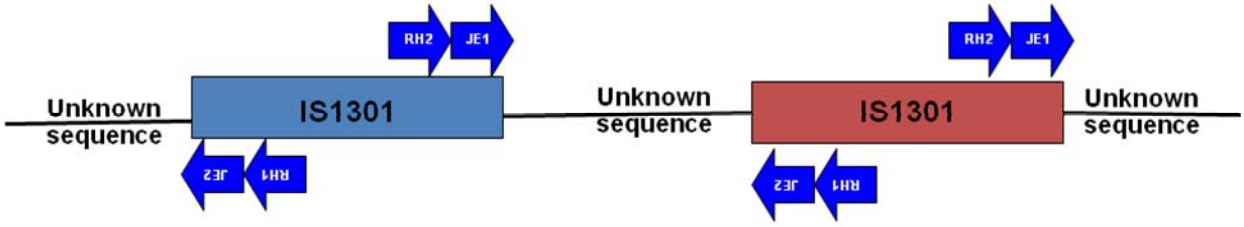

Figure 5: A schematic illustrating the inverse PCR strategy.
Inverse IS1301 primers were used to amplify unknown sequences between two IS1301 fragments that were close to one another on the *N. meningitidis* genome.

3.10 WHOLE GENOME SEQUENCING

3.10.1 Whole Genome Assembly

The DNA from isolate NM216 was purified by phenol:chloroform extraction. The quality of the DNA was checked by running 2µg of DNA on a 0.8% agarose gel. Whole genome sequencing was performed by the GPCL on a Roche Genome Sequencer FLX system. The resulting sequence data was assembled using the SeqMan Genome Assembler by DNAstar (DNAstar,

Madison WI). Two different genome assemblies were performed - *de novo* and a combination, FAM18C/IS1301 templated and *de novo* assembly with junkyard sequences. For the combination assembly, a templated assembly to the sequenced FAM18C reference genome was performed with the IS1301 sequence included. This templated assembly was broken up such that all areas of coverage became separate contigs. A second *de novo* assembly was then performed with the resulting contigs and the junkyard sequences. The Lasergene v8.1 suite of sequence analysis software was used to search for IS1301 integration sites within resulting contigs. BLAST analysis was performed on IS1301 containing sequences to determine the identity of potential IS1301 insertion sites.

3.10.2 *opa* PCR and Sequencing

Based upon BLAST homologies of the 386 kb contig (558435-601117) and the 40 kb contig (902277-935662) from the combination FAM18C/IS1301 templated - junkyard *de novo* assembly, primers that flank the potential IS1301 insertion in the *opaD* gene of NM216 were designed. PCR amplification of NM216, DE9246 and the early clone strain, NM23 was performed with opaD-F (5'-TATCGGGGAATCAGAAGCGGTAGC-3') and opaD-R (5'-GGCAGGCGGCAAGTGAAGACAGC-3') at a final concentration of 10μM. In addition, the opaD primers were used in combination with the inverse facing IS1301 specific primers, JE1XBAI and JE2XBAI to look for IS1301 specific amplification of *opaD* flanking regions. PCR amplification with IS1301F and IS1301R was included as positive control. The inverse PCR products resulting from the combinations of either JE1XBAI and opaD-F or JE2XBAI and opaD-R from NM216 and DE9246 were sequenced using these primer pairs.

Briefly, 5µl of PCR product was cleaned with 2µl of EXOSAP-IT (Applied Biosystems) and incubated for 37ºC for 15 minutes and heat inactivated for 15 minutes at 80ºC to remove excess dNTPs and primers. An ABI PRISM® Big Dye Terminator Cycle Sequencing Kit v3.1 (Applied Biosystems) was used and the sequence reaction were performed with 1µl of the cleaned PCR product and 3.2µM primer at the GPCL facility. The sequencing conditions were as follows: 95ºC for 5 minutes, then 30 cycles of 95ºC for 1 minute, 55ºC for 1 minute, and 72ºC for 1 minute, followed by 72ºC for 7 minutes. The samples were cleaned and analyzed by capillary electrophoresis at the University of Pittsburgh Genomics and Proteomics Core Laboratories on an ABI 3730 DNA analyzer. Contigs of the resulting sequence data were generated using SeqMan Pro v8.0 (DNAstar).

4.0 RESULTS

4.1 SEQUENCING FOR ET-15

The 39 late clones and four early clones were sequenced using primers for the *fumC* gene to determine if the isolates were ET-15 clones. The electropherogram of the early clone showed a guanine, while the late clone possessed an adenine (Figure 8), which results in an amino acid change from lysine to glutamic acid. The 39 late clone isolates all possessed the single nucleotide polymorphism that distinguishes them as belonging to the ET-15 clone while the four early clone isolates did not.

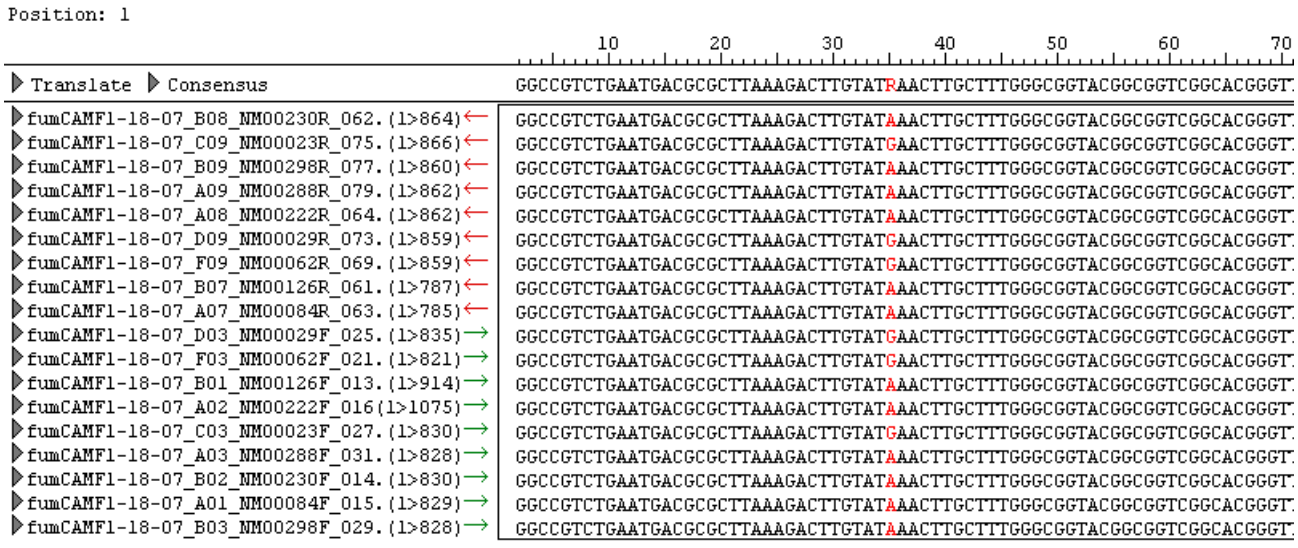

Figure 6: Representation of *fumC* sequence analysis to distinguish ET-15 clone.
SeqMan Pro (Lasergene 7.2) analysis showing the single nucleotide polymorphism.

4.2 IDENTIFICATION OF IS1031 IN LATE CLONE ISOLATES

The thirty-nine late clone isolates, reference strain DE9246, and the four early clone isolates, underwent PCR analysis using IS1301 specific primers, SH42 and SH46, to determine if the isolates contained IS1301. As shown in figure 7, a representative gel of 18 of the late clone isolates showing one band at approximately 450bp (orange arrow) corresponding to IS1301, while the early clones and the negative PCR control do not display a PCR product.

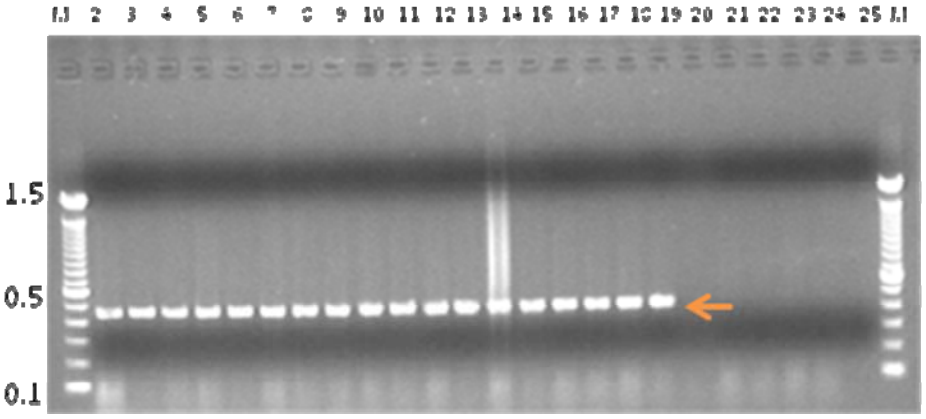

Figure 7: PCR using the SH42 and SH46 primers with late and early clone DNA.
Molecular weight markers (M), 100bp ladder with sizes in kilobase pairs, present in lanes 1 and 26. A representation of the late clone (lanes 2-19) and early clone (lanes 20-24) isolates used in this study amplified with SH42 and SH46 primers. Negative PCR control (lane 25).

4.3 SOUTHERN BLOT ANALYSIS

Characterization of IS1301 insertion sites was performed by Southern blot analysis of *Hinc*II genomic digests of the late clones. Figure 8 displays a representation of a pre-Southern blot image of a 0.8% agarose gel with 2µg of *Hinc*II digested genomic DNA. After Southern blot transfer, hybridization, and immunodetection, the membrane was exposed to x-ray film to visualize the IS1301 positive fragments (Figure 9). The reference strain, DE9246 is present in

lane 11 and ten of the IS1301 insertions sites in this strain have been previously characterized (Elias & Vogel, 2007). The uncharacterized high molecular weight IS1301 triplet is indicated by a red arrow in Figure 10. The late clone IS1301 banding pattern obtained by Southern blot was analyzed for genetic similarity using BioNumerics v.5.10 (Applied Maths) software. A distance tree using clustering with the Unweighted Pair Group Method with Arithmatic Mean (UPGMA) and the Pearson coefficient was performed to normalize the banding patterns of the IS1301 fragments. The analysis shows multiple copies of IS1301 in all late clone isolates, ranging from five to ten insertions. The high molecular weight IS1301 triplet was seen in the majority of late clone isolates with the exception of isolate NM222 (Figure 9, lane 4). NM23, an early clone, did not show IS1301 fragments when hybridized with an IS1301 probe (data not shown). Furthermore, three outbreaks in Maryland during the study period displayed the same banding pattern within each outbreak (Figure 10, colored isolate names). These data are consistent with previous work by Vogel *et al*, to develop IS1301Southern blots as a typing method for the ET-15 clone (Elias & Vogel, 2007).

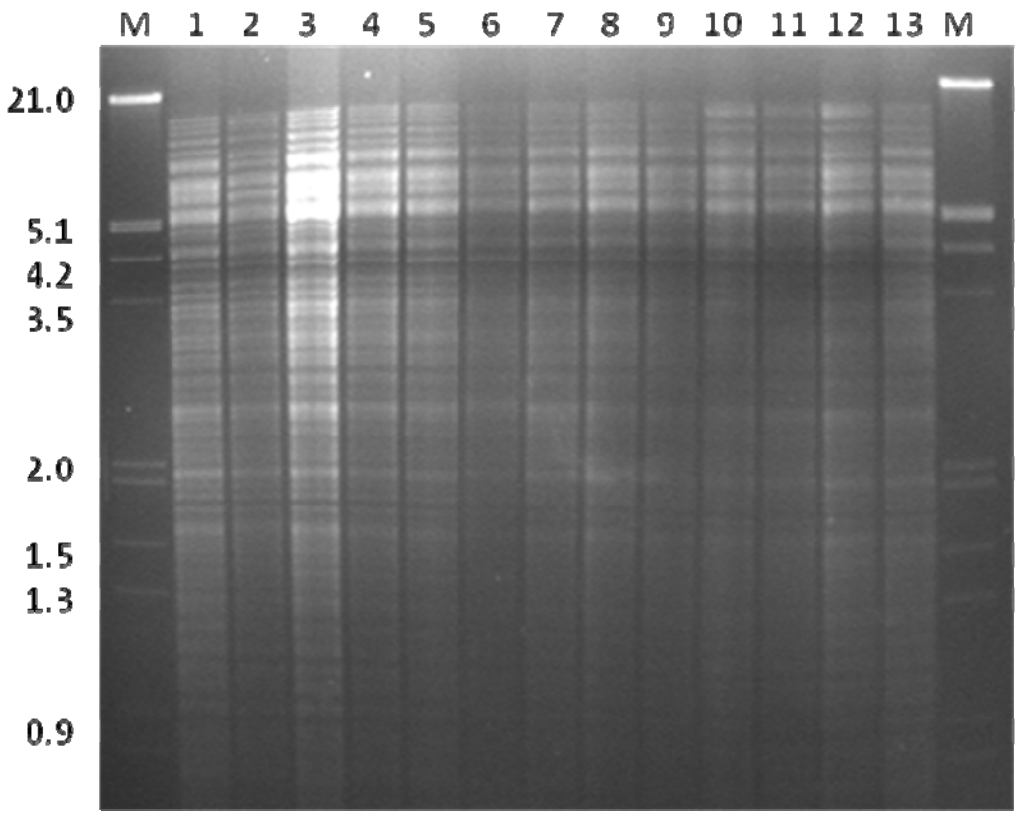

Figure 8: Restriction digest of *N. meningitidis* using *Hinc*II.
A representational 0.8% agarose gel with *Hinc*II digested DNA from 13 late clone isolates (lanes 1-13). Lanes (M) DNA Molecular Weight Marker III, digoxigenin-labeled, marked in kilobase pairs (Roche Diagnostics).

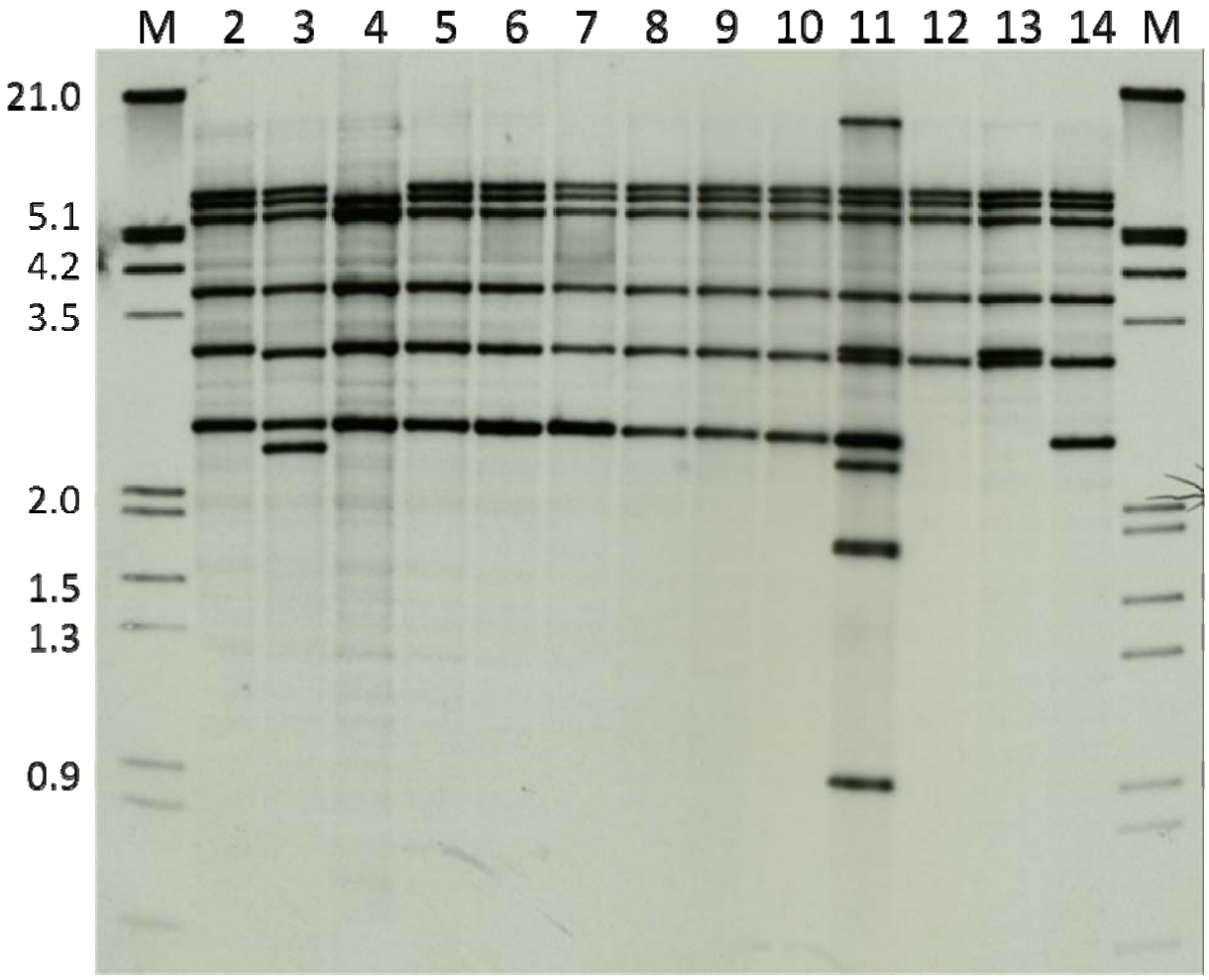

Figure 9: IS1301 Southern blot of *Hinc*II gDNA from 13 late clones.
Lanes (M) DNA Molecular Weight Marker III, digoxigenin-labeled, marked in kilobase pairs (Roche Diagnostics). Lanes 2-14, *Hinc*II gDNA from 13 late clones.

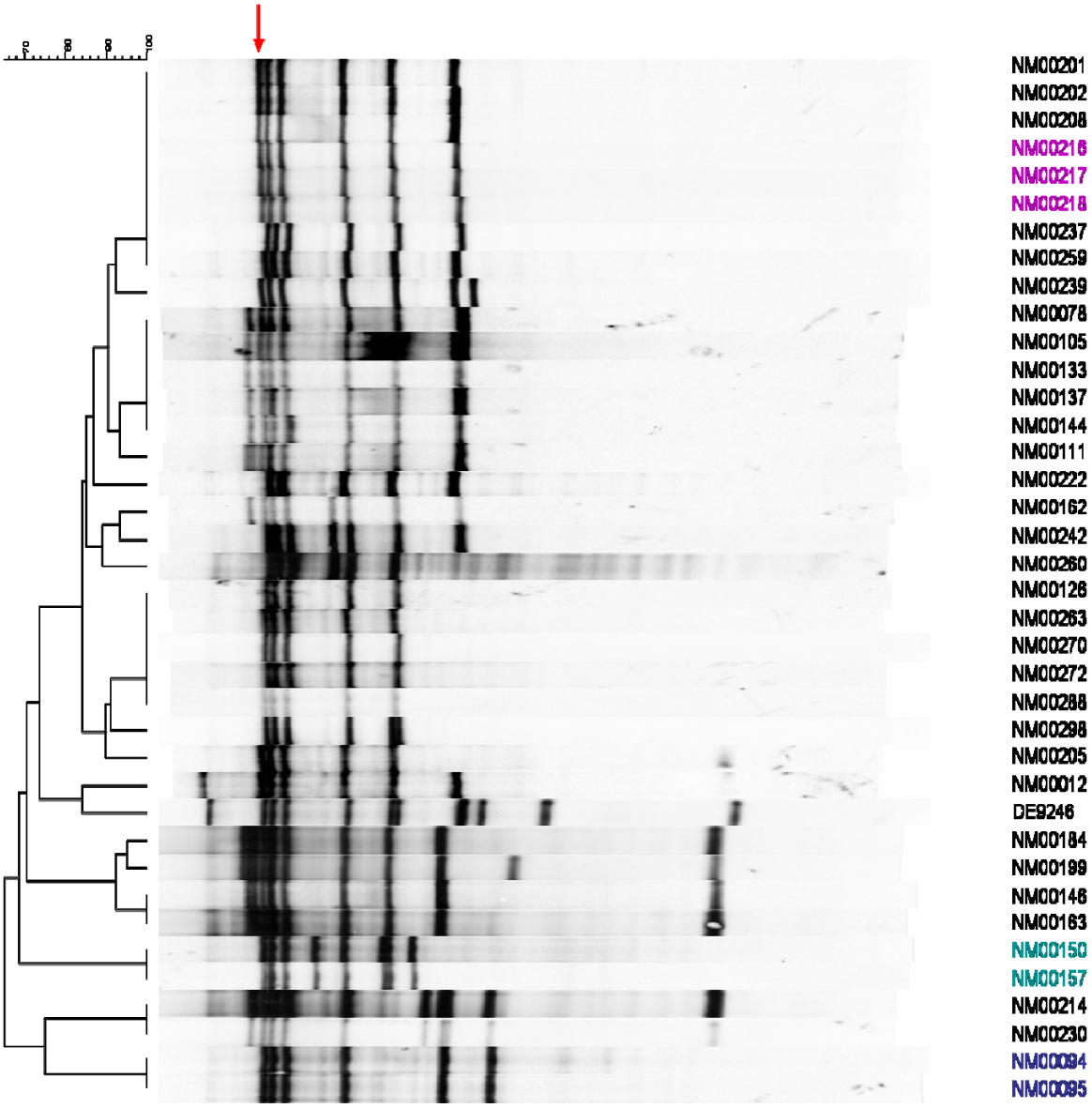

Figure 10: Dendogram of IS1301 Southern blot of 37 late clone isolates.
UPGMA of IS1301 banding pattern. The colored isolates indicate three outbreaks in Maryland that have the same IS1301 banding pattern, red arrow indicating the IS1301 triplet.

4.4 RANDOM HEXAMER STRATEGY

4.4.1 Experiment 1 – Late Clone Whole Genomic DNA

The first round of PCR using either random hexamer specific primer, Tag1, Tag2, Tag3, or Tag4 and IS1301F generated a smear of products in all four reactions (Figure 11, lanes 2-5). A smear of product was expected because the random hexamer binds sporadically throughout the genome, generating various sizes of product. The positive control using late clone genomic DNA and IS1301 specific primers produced a band at ~450bp, characteristic of IS1301 (Figure 11, lane 6). The second round of PCR using the SH42 and SH46 primers produced a band at approximately 100bp (Figure 12, lanes 2-5, orange arrow). A 100bp product was unexpected; PCR products of various sizes were expected when using the random hexamer specific primer and the nested IS1301 primer to amplify template from the first round of PCR. The positive control using late clone genomic DNA and IS1301 specific primers, once again produced a band at ~450bp (Figure 12, lane 6).

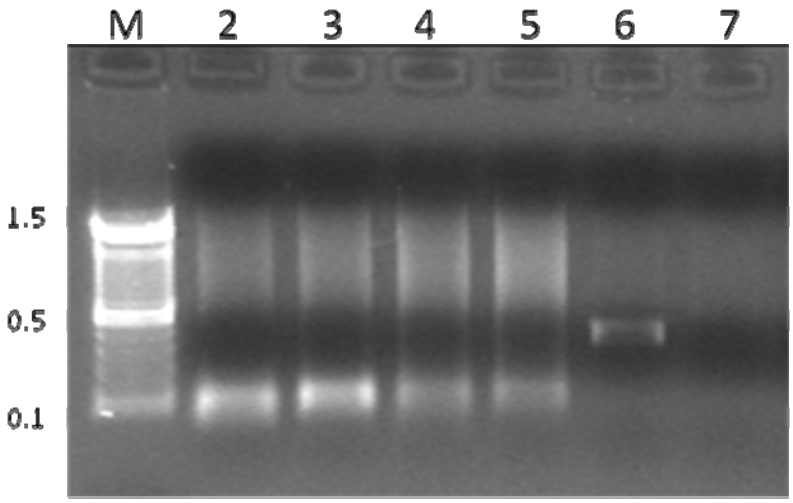

Figure 11: Random hexamer PCR of *N. meningitidis* genomic DNA.
Molecular weight marker (M), a 100bp ladder, is in lane 1 marked with kilobase pairs. Lanes 2-5 PCR amplification Tag1, Tag2, Tag3, or Tag4, respectively and IS1301F. Positive PCR control using SH42 and SH46 primers with genomic DNA (lane 6). Negative PCR control (lane 7).

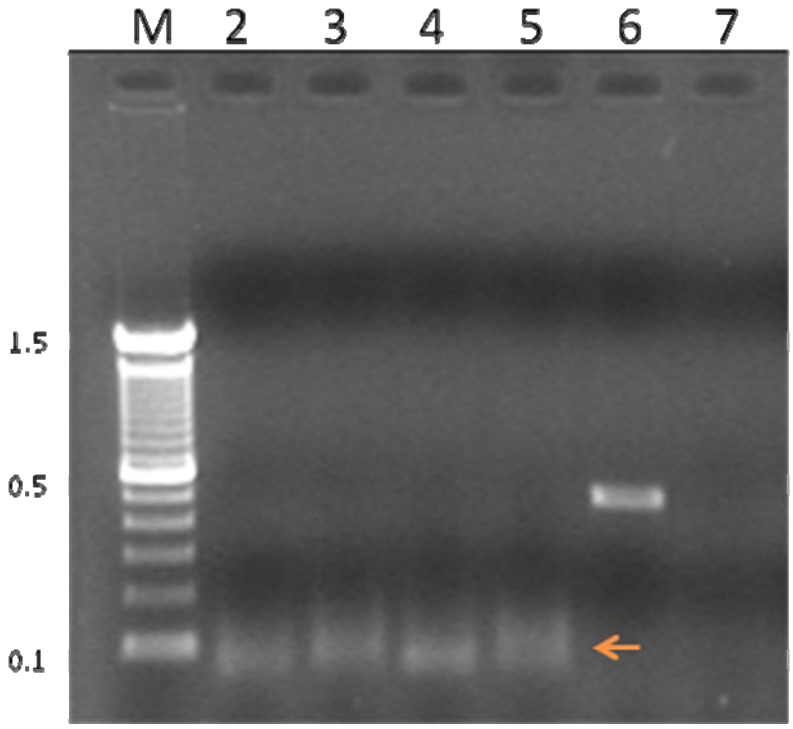

Figure 12: Nested PCR of round 1 PCR products.
Molecular weight marker (M), 100bp ladder, sizes shown in kilobase pairs (lane 1). Lanes 2-5 used either Tag1, Tag2, Tag3, or Tag4 and SH42. Positive PCR control using SH42 and SH46 primers with genomic DNA (lane 6). Negative PCR control (lane 7).

4.4.2 Experiment 2 – *Hinc*II Digested Late Clone Genomic DNA

Instead of using genomic DNA, *Hinc*II digested late clone genomic DNA was used to reduce the amount of background by creating smaller fragments of DNA for the random hexamer to bind. The DNA was amplified with random hexamer primers, an IS1301 specific primer, and Klenow. Round 1 PCR, using random specific primers and IS1301F or IS1301R, produced a smear of PCR product ranging in size from 100bp to 1kb. The product also showed distinct bands at various sizes within each reaction (Figure 13, lanes 2-9, yellow arrows). Amplification of PCR product from round 1 with random hexamer specific primers and a second IS1301 primer, either SH42 or SH46 in a subsequent round of PCR produced a smear of product ranging from 100bp to >1kb with no distinct bands (Figure 14, lanes 2-9). The PCR products in red boxes were used in later cloning experiments (see sections 4.4.4 and 4.4.5).

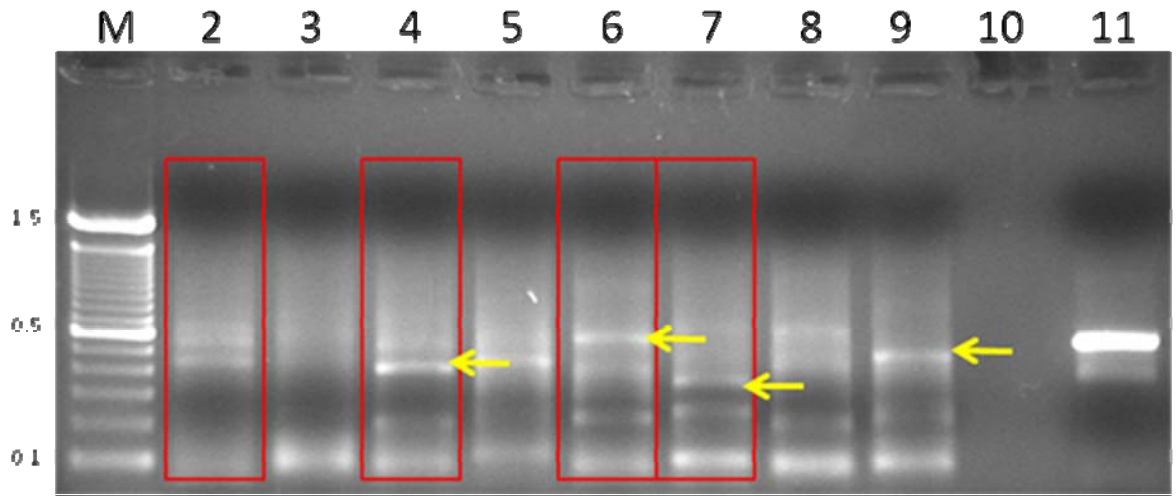

Figure 13: Random hexamer PCR *Hinc*II digested *N. meningitidis* DNA.
Molecular weight marker (M), 100bp ladder, sizes shown in kilobase pairs (lane 1). Lanes 2-5 used Tag1, Tag2, Tag3, or Tag4 and IS1301F. Lanes 6-9 used Tag1, Tag2, Tag3, or Tag4 and IS1301R. Negative PCR control (lane 10), positive PCR control using gDNA and SH42 and SH46 primers (lane 11).

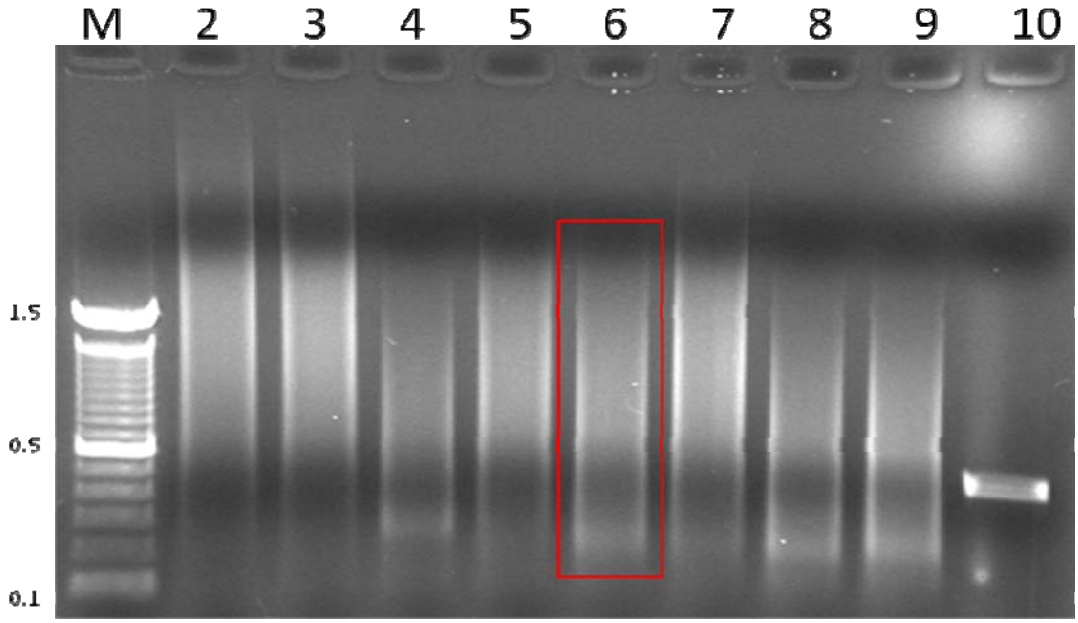

Figure 14: PCR of round 1 PCR products (*Hinc*II digested gDNA).
The molecular weight marker (M), a 100bp ladder, is in lane 1 with the sizes in kilobase. Lanes 2-5 used Tag1, Tag2, Tag3, or Tag4 and SH42. Lanes 6-9 used Tag1, Tag2, Tag3, or Tag4 and SH46. Positive PCR control (lane 10), which used SH42 and SH46 primers with gDNA.

4.4.3 Experiment 3 – Gel Extracted High Molecular Weight IS1301 Triplet

In experiment 3, the gel purified IS1301 triplet was used as template in the random hexamer PCR to increase the likelihood of amplifying these DNA fragments. Round 1 PCR using the random hexamer primer, Tag1 and IS1301F or IS1301R produced a smear ranging in size from 100bp to 900bp. In addition, several distinct bands were visible within the smear of PCR products (Figure 15, lanes 2-3). The second round of PCR amplification of these products generated a smear of DNA ranging in size from 100bp to 2kb, without specific bands (Figure 16). PCR product in red boxes was used in subsequent cloning procedures.

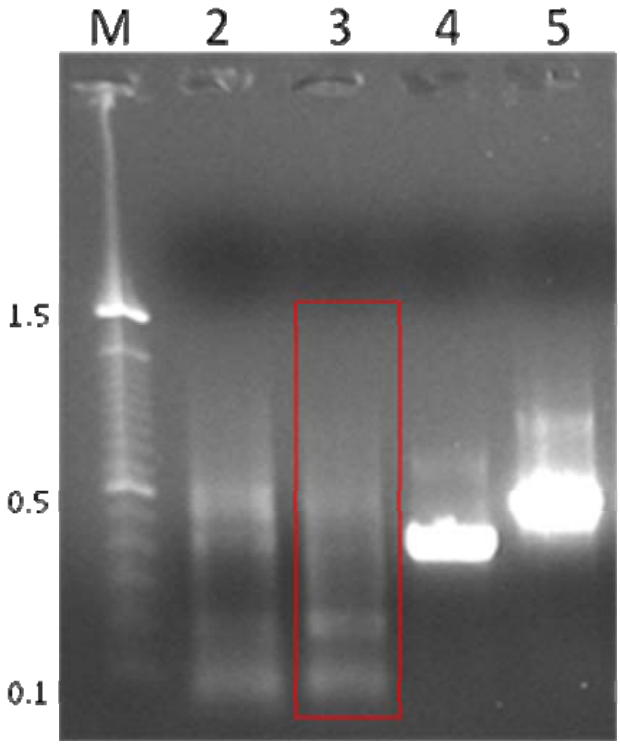

Figure 15: Random hexamer PCR of purified *N. meningitidis* IS1301 triplet.
Molecular weight marker (M), a 100bp ladder with sizes in kilobase pairs (lane 1). Lanes 2-3 used Tag1 and either IS1301F (lane 2) or IS1301R (lane 3). Two positive controls, SH42 and SH46 primers and IS1301F and IS1301R primers with gDNA (lanes 4-5).

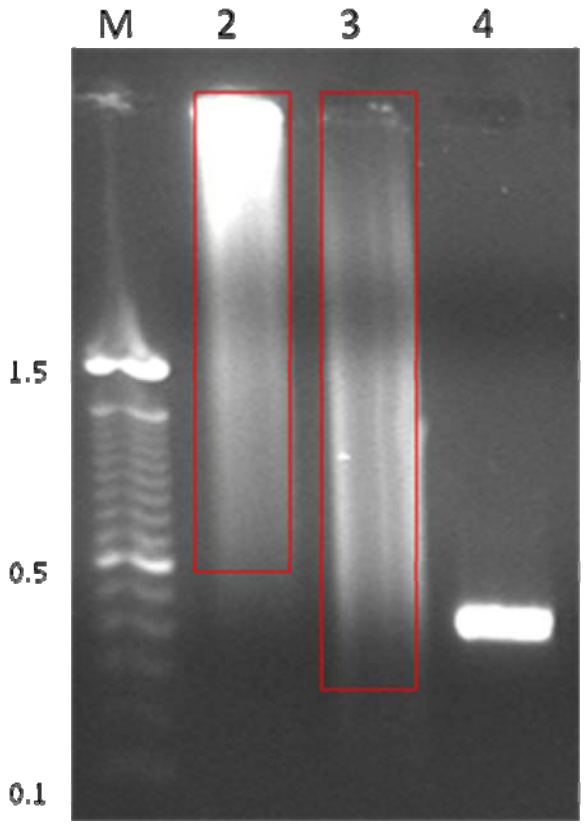

Figure 16: Nested PCR of round 1 PCR products (IS1301 triplet).
Molecular weight marker (M), a 100bp ladder, is in lane 1 with sizes in kilobase pairs. Tag1 and either SH42 (lane 2) or SH46 (lane 3). Positive control using SH42 and SH46 primers with gDNA (lane 4).

4.4.4 Cloning Experiment 1

Eight of the various combinations of round 1 and round 2 PCR reactions described in sections 4.4.2-3 above were cloned (Table 4) to search for inserts containing IS1301 and the flanking regions. Colony PCR with IS1301 and plasmid specific primers on 149 of the resulting transformants identified 13 clones with IS1301 sequences that were >300bp in size (Figure 17). Sequence analysis of the plasmid inserts demonstrated that while all of the clones possessed a portion of IS1301 sequence, none of the cloned sequence extended beyond IS1301 into the

flanking regions, which suggests the random hexamer is binding close to or within the IS1301 sequence.

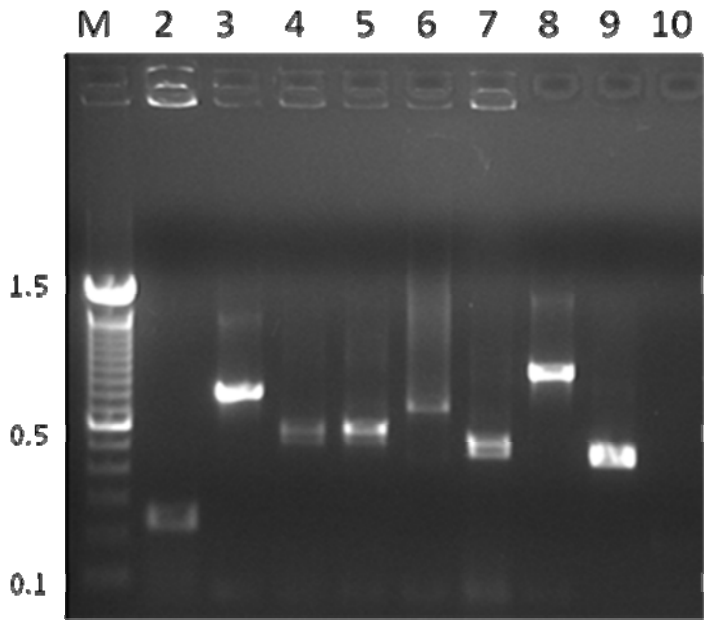

Figure 17: Approximate insertion size of IS1301 positive colonies.
Molecular weight marker (M), a 100bp ladder with sizes in kilobase pairs (lane 1). A representational image of the colonies that were positive for IS1301 in the colony PCR using M13F and M13R primers (lanes 2-8). Positive PCR control using SH42 and SH46 with gDNA (lane 9), and negative PCR control (lane 10).

4.4.5 Cloning Experiment 2

A second cloning experiment was performed using the round 1 random primer PCR and nested PCR product from the purified high molecular weight IS1301 triplet (Table 4). Resulting transformants were screened by colony blots with an IS1301 probe. Figure 18 shows JM109 transformed colonies (A), while (B) shows the colonies positive for IS1301 insertion by colony blot using an IS1301 probe (see section 3.4). The colonies circled red on the agar plate, were IS1301 positive by colony blot, however, the colonies circled in yellow on the agar plate were not positive for IS1301 by colony blot (Figure 18). This suggested that the colony blot was an effective screen for IS1301 containing clones.

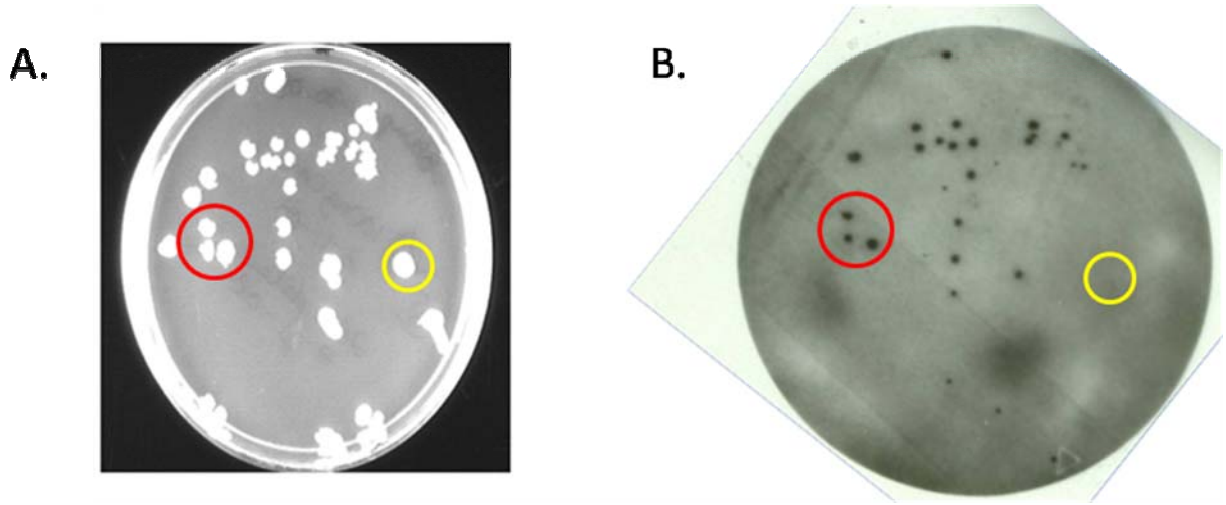

Figure 18: Representational image of transformed colonies and corresponding colony blot.
The transformed JM109 cells were plated (A) and screened using a colony blot with an IS1301 probe. The positive colonies are depicted on x-ray film (B).

Approximately 230 colonies were positive for IS1301 by colony blot. The positive clones were screened by colony PCR with primers specific for IS1301. Of the 230 IS1301 positive colonies identified by colony blotting, only 43 were PCR positive for IS1301. Figure 19 illustrates 7 IS1301 positive clones (lanes 3, 6, 12, 13, 14, 17, 18, red boxes). The band size (~450bp) is consistent with the expected IS1301 product. Subsequent PCR using vector primers (Figure 20), show the insert size to be ~300bp (green lines), with the exception of plasmid #B.3.6 which shows an insert size of ~700bp (red arrow). Once again, sequence analysis of plasmid #B.3.6 demonstrated that while the clone possessed a portion of IS1301 sequence, the cloned sequence did not extended beyond IS1301 into the flanking regions. The vector only control (lane 11) shows a band at ~300bp and the control DNA (included in kit) shows a characteristic band at 750bp (lane 13).

The second cloning experiment did not produce clones with IS1301 and its flanking regions. Possible reasons include failure of the selection criteria, causing a large number of "vector-only" plasmids, poor probe specificity with IS1301 and poor ligation of the PCR product

to the plasmid. To further investigate the specificity of the IS1301 probe, a plasmid known to contain IS1301, LMK140 (Kostelnik, 2006), was hybridized with and without the IS1301 probe. In addition, the colonies from the control DNA (supplied with the Promega kit) was blotted and hybridized with the IS1301 probe. The LMK140 IS1301 hybridization gave positive colonies as well as some nonspecific background (data not shown). In addition, the colony blot of control DNA, showed a few IS1301 positive colonies and some nonspecific background (data not shown). These data suggest that the IS1301 probe was binding non-specifically to non-IS1301 DNA. In future experiments, a freshly labeled or more specific probe using radioactivity are needed to more adequately screen a large number of colonies for IS1301.

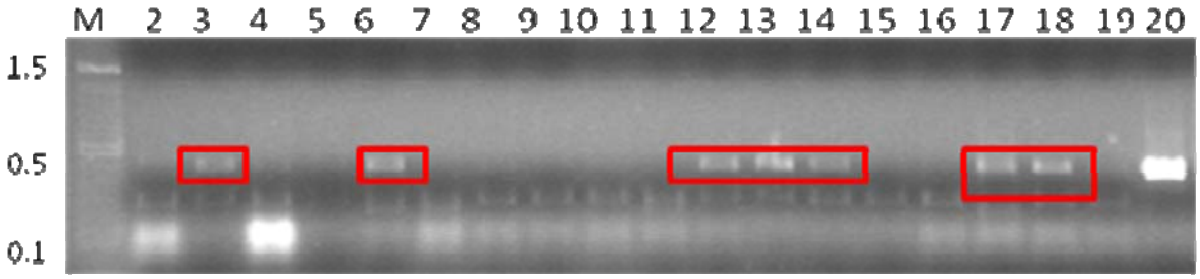

Figure 19: Representational image of colony PCR.
Molecular weight marker (M), a 100bp ladder, with sized in kilobase pairs (lane 1). Colony PCR using primers SH42 and SH46 (lanes 3, 6, 12, 13, 14, 17, 18, red boxes). Positive control using gDNA and SH42 and SH46 primers (lane 20).

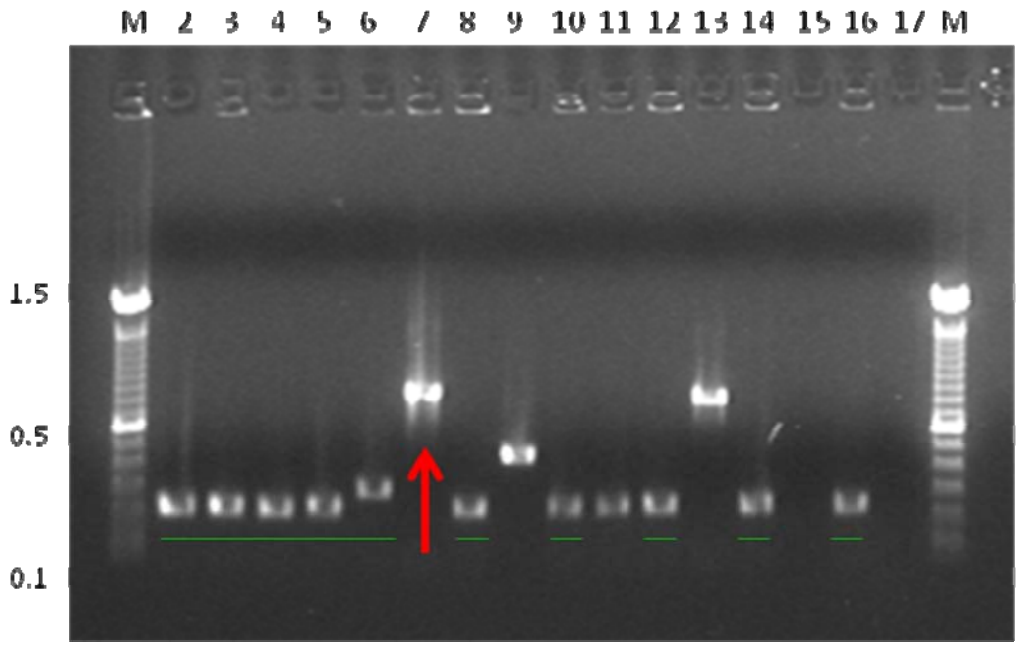

Figure 20: Representational image of PCR using M13F/R primers.
Molecular weight marker (M), 100bp ladder in lanes 1 and 18 marked with sizes in kilobase pairs. Lanes 2-6, 8, 10, 12, 14, and 16 shows an insert size of ~300bp. Lane 7 shows colony #B.3.6 (red arrow), lane 9 is the positive control using SH42 and SH46 primers with genomic DNA, lane 13 is the positive control using control DNA (included in kit). Lane 11 is a vector only control (no insert) and lanes 15 and 17 are negative PCR controls.

4.5 LINKER BASED STRATEGY

4.5.1 IS1301 Southern Blot – *Hinc*II Digest & *Hha*I Digest

An IS1301 Southern blot was done to determine the size of the IS1301 fragments from *Hha*I cut genomic *N. meningitidis* DNA (Figure 21). The gel purified IS130 triplet was only supposed to contain the three IS1301 fragments, but Southern blot analysis revealed a fourth IS1301 fragment that was also extracted from the gel (Figure 21, lane 6). The sizes of the IS1301 fragments from both the gDNA and the IS1301 triplet preparations were reduced in size when cut with *Hha*I

(Figure 21, lanes 3 and 4) from 5-10kb to 0.7-1.5kb. The smaller *Hha*I fragments would facilitate cloning of the IS1301 elements from the triplet DNA.

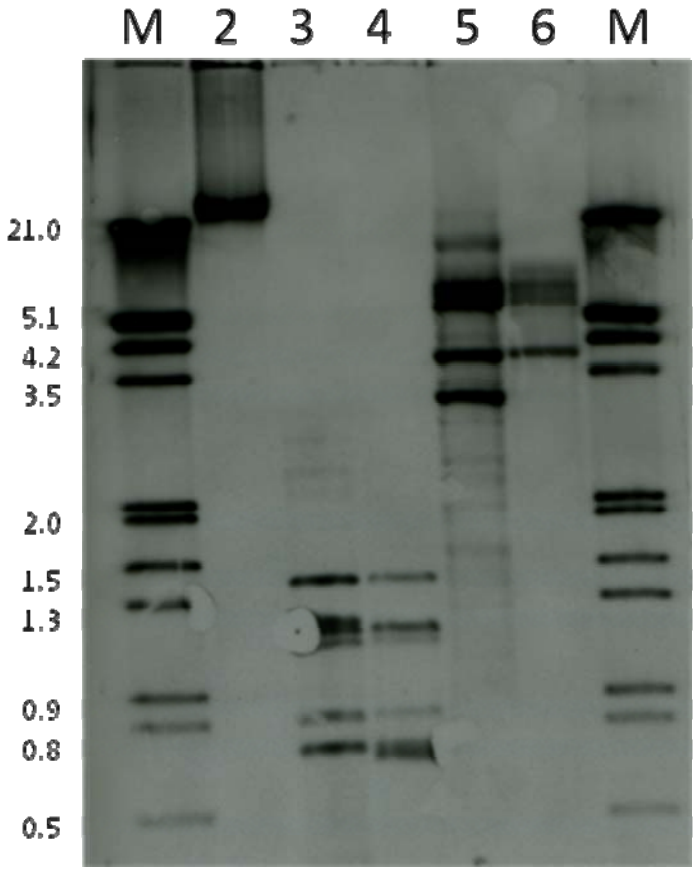

Figure 21: Southern blot of *Hinc*II or *Hha*I digest of *N. meningitidis* DNA.
Lanes (M) DNA Molecular Weight Marker III, digoxigenin-labeled, marked in kilobase pairs (Roche Diagnostics). Uncut *N. meningitidis* DNA, NM126 (lane 2), *Hha*I cut *N. meningitidis* DNA, NM126 (lane 3), IS1301 triplet cut with *Hha*I, NM272 (lane 4), *Hinc*II cut *N. meningitidis*, NM126 (lane 5), gel purified IS1301 triplet, NM272 (lane 6).

4.5.2 Experiment 1 Linker Strategy - PCR

The five linker ligation reactions (Table 5) were amplified in a PCR reaction with the HhaI-primer and an inverse IS1301 primer. Product amplified with RH1XBAI and the HhaI-primer generated a 100bp band (Figure 22, lanes 3-6). A 100bp product was expected because *Hha*I

cuts within the IS1301 fragment and the RH1XBAI and HhaI-primer does not amplify unknown sequence (Figure 3). The 100bp band is a result of the *Hha*I linkers, as PCR with the linkers only (no gDNA) also produced a band at 100bp (Figure 22, lanes 11-14, 16-17). However, PCR amplification of the linker ligated gDNA with the HhaI-primer and the RH2XBAI primers generated two products of 100bp and 300bp (Figure 22, lanes 7-9, red arrow). The 300bp product may contain approximately 100bp of the flanking region of an IS1301 fragment. It was expected that this PCR reaction would have generated more than one PCR product because there are multiple IS1301 fragments in the isolates used; however, it is possible that the smallest fragment was preferentially amplified in the PCR reaction, leaving out the larger fragments.

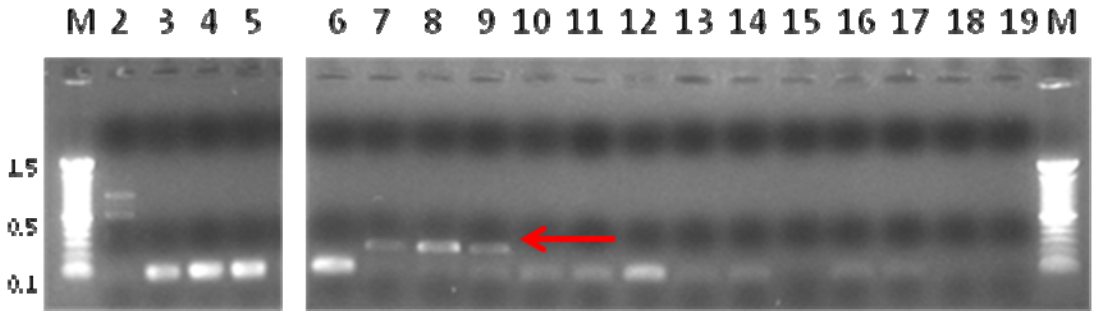

Figure 22: Experiment 1 Linker Strategy PCR Results
Lanes 1 and 20, molecular weight marker (M), 100bp ladder with sizes in kilobase pairs. Uncut genomic DNA with RH1XBAI and RH2XBAI (lane1), *Hha*I cut genomic DNA (attached linker) with HhaI-primer and RH1XBAI with 1ng, 10ng, 25ng, 50ng of DNA (lanes 3-6), *Hha*I cut genomic DNA (attached linker) with HhaI-primer and RH2XBAI with 1ng, 10ng, 25ng, 50ng of DNA (lanes 7-10), *Hha*I linkers (with ligase) with RH1XBAI and RH2XBA1 (lane 11), *Hha*I linkers (no ligase) with RH2XBAI and HhaI-primer (lane 12), *Hha*I linkers (no ligase) with RH1XBA1 and HhaI-primer (lane 13), *Hha*I linkers (no ligase) with RH2XBA1 and RH2XBAI (lane 14) uncut genomic DNA with HhaI-primer (lane 15), *Hha*I linkers (no ligase) with HhaI-primer (lane 16), *Hha*I linkers (with ligase) with HhaI-primer (lane 17), negative PCR controls (lanes 18-19).

4.5.3 Experiment 2 Linker Strategy - PCR

Since the first experiment showed a large amount of excess linker, the *Hha*I cut DNA with the attached *Hha*I linkers was run on a gel and the region from 300bp-700bp and >700bp was extracted (Figure 23, yellow and red boxes). Next, the two gel purified products were amplified

in a PCR reaction with the HhaI-primer and inverse IS1301 primers. The *Hha*I cut genomic DNA (attached linker, 300-700bp) amplified with HhaI-primer and RH2XBAI or RH2XBAI produced a band at ~100bp (Figure 24, lanes 6, 7). *Hha*I cut genomic DNA (attached linker, >700bp) amplified with HhaI-primer and RH1XBAI also produced a band at 100bp, however, the *Hha*I cut genomic DNA (attached linker, >700bp) amplified with HhaI-primer and RH2XBAI produced two bands at 100bp and ~300bp (Figure 24, lanes 8, 9). This 300bp (red arrow) product could potentially have approximately 100bp of unknown sequence, similar to experiment 1. The linker controls without the DNA template also produced a band at 100bp (data not shown) similar to experiment 1 and uncut genomic DNA and *Hha*I cut genomic DNA with the ligated linkers amplified with the HhaI-primer did not produce a band (Figure 24, lanes 4, 5).

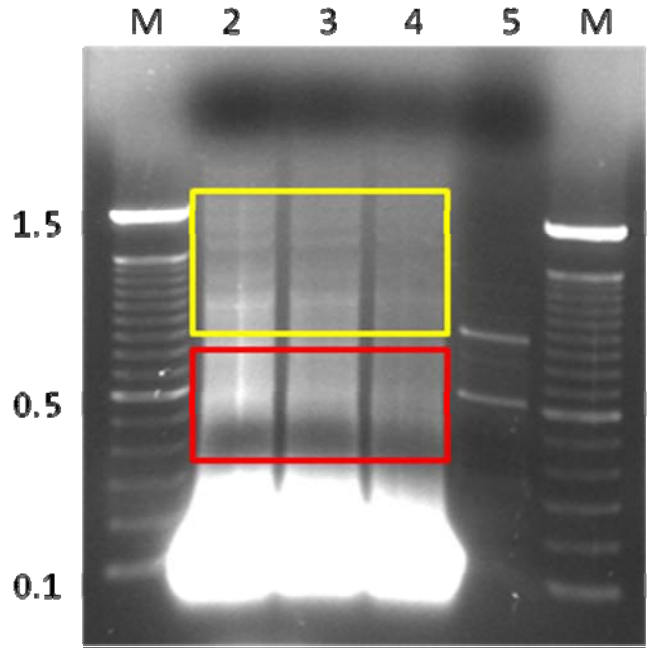

Figure 23: Gel extraction of *Hha*I cut DNA (with linkers).
Molecular weight marker (M), 100bp ladder with sizes in kilobase pairs (lanes 1 and 6). *Hha*I cut DNA with the attached *Hha*I linkers (lanes 2-4), genomic *N. meningitidis* DNA with RH1XBAI and RH2XBAI primers (lane 5).

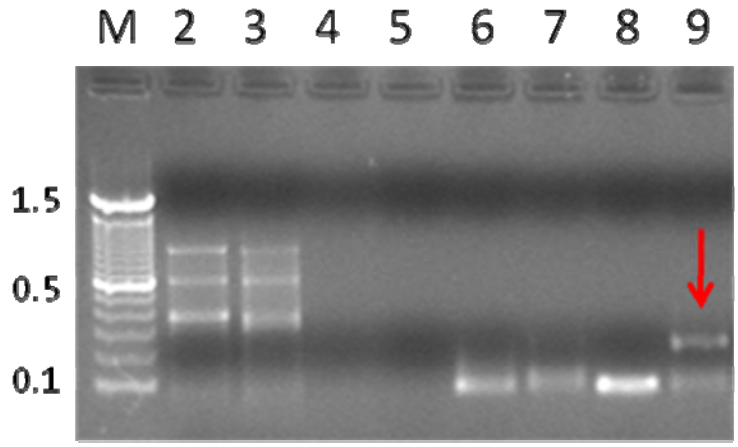

Figure 24: Experiment 2 Linker Strategy PCR Results.
Molecular weight markers (M), 100bp ladder with sizes in kilobase pairs (lane 1). Genomic *N. meningitidis* DNA with RH1XBAI and RH2XBAI primers (lanes 2-3), genomic *N. meningitidis* DNA with HhaI-primer (lane 4), *Hha*I cut genomic DNA (attached linker, 300-700bp) with HhaI-primer (lane 5), *Hha*I cut genomic DNA (attached linker, 300-700bp) with HhaI-primer and RH1XBAI (lane 6), *Hha*I cut genomic DNA (attached linker, 300-700bp) with HhaI-primer and RH2XBAI (lane 7), *Hha*I cut genomic DNA (attached linker, >700bp) with HhaI-primer and RH1XBAI (lane 8), *Hha*I cut genomic DNA (attached linker, >700bp) with HhaI-primer and RH2XBAI (lane 9).

4.6 INVERSE PCR

While performing PCR for the linker strategy, the inverse primers (Table 6) were used on genomic late clone DNA and produced multiple PCR products. In this experiment, an inverse PCR was performed using the primers RH1XBAI and RH2XBAI which produced four distinct bands and two faint bands (Figure 25, lane 1). PCR with primers JE1XBAI and JE2XBAI produced five distinct bands and two faint bands (Figure 25, lane 3). The four distinct bands from lane 1 were gel purified and sequenced with the nested JE1XBAI and JE2XBAI primers. Sequencing of the bands failed, possibly due to random amplification of the genome by the RH1XBAI and RH2XBAI primers or poor gel extraction. In addition, DNA from the IS1301 triplet and DNA from an early clone were amplified with the inverse primers to check for PCR product, but the primers did not amplify product in either reaction (data not shown). A better

method of extracting the bands from the gel or cloning of the PCR product and using vector primer for sequencing could be done in future experiments in order to fully ascertain the sequence of the PCR product.

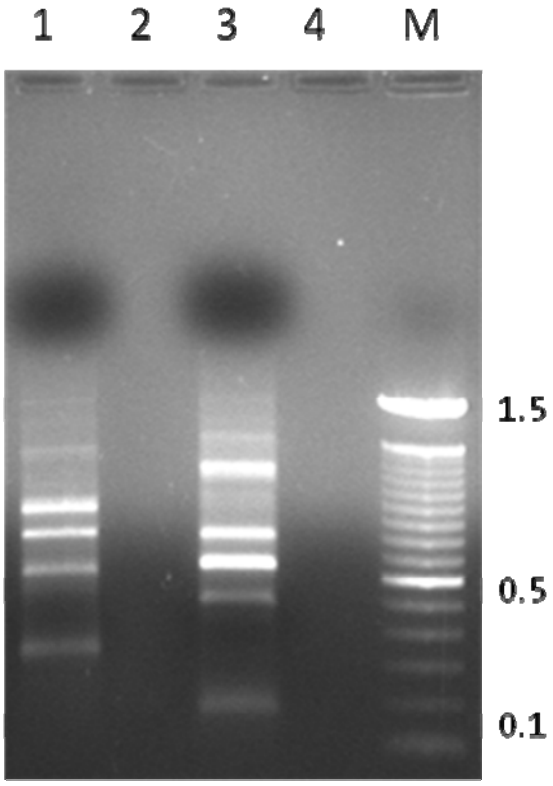

Figure 25: Inverse PCR of genomic *N. meningitidis* DNA.
Molecular weight marker (M), 100bp ladder with sizes in kilobases, genomic *N. meningitidis* DNA amplified with RH1XBAI and RH2XBAI (lane 1), or JE1XBAI and JE2XBAI (lane 3), lanes 2 and 4 are empty.

4.7 WHOLE GENOME SEQUENCING

4.7.1 Whole Genome Sequence Analysis

The *de novo* assembly of the NM216 whole genome sequence data using the SeqMan Genome Assembler (DNAstar) generated 341 contigs from 225,568 sequences with an N50 of 26 kb. The

N50 represents the contig size at which 50% of the sequence data is represented. Therefore, the larger the N50, the better the assembly. The largest contig from the *de novo* assembly was 56 kb. IS1301 sequences were identified in 6 of contigs (Table X). All of these corresponded to bands previously described by Vogel (Elias & Vogel, 2007). The BLAST homologies of contigs 00015 and 00025 were highest to GenBank accession number AM412561, an intergenic fragment from DE9246. No reference to this GenBank accession number is made in the Vogel paper and these contig sequences showed no homology to the FAM18 reference genome (Elias & Vogel, 2007).

Table 7. IS1301 inserts identified from de novo assembly of NM216 whole genome sequence.

Contig #	Size (kb)	IS1301	Coverage at IS1301	BLAST	Vogel Insert No.
00015	12.1	5’	20	AM412561	AM412561
00025	57.6	3’	5-10	AM412561	AM412561
00070	1.5	3’	16	TspB	7
00075	4.3	3’	10	Intergenic	5a
00084	56.9	3’		Intergenic	5b
00096	7.1	5’	8	Intergenic	5a

Thus, the whole genome *de novo* assembly of NM216 identified 4 IS1301 insertions. Three of these correspond to insertion described by Vogel. The fourth insertion, while referenced by Vogel, is not indicated in their paper. We speculate that this sequence corresponds to the Vogel insertion no. 6 as this band appears to be shared between DE9246 and NM216 (Figure 10). Further investigation is required to confirm the identity of this IS1301 insertion in NM216.

The combination FAM18C/IS1301 templated and *de novo* assembly with junkyard sequences generated 249 contigs from 224,885 sequences with an N50 of 152 kb. This assembly generated 2 very large contigs of 388 kb in length. The 3’ end of IS1301 was identified in one of these large contigs (558435-601117). The IS1301 insertion site has 20X sequence coverage and a nr BLAST homology search of the first 500 nucleotides upstream of the IS1301 sequence showed 99.8% sequence identity to chromosome location 901462←901004 of the FAM18C

genome. This location in the FAM18C genome corresponds to NMC0903, a pseudogene of the opacity protein, OpaD.

A search of the other contigs in the assembly identified the 5' end of IS1301 in a 40.4 kb contig (902277-935662). This IS1301 insertion site has 13X sequence coverage and a BLAST homology search of 200 nucleotides downstream of the IS1301 insertion showed 99% sequence identitiy to chromosome location 901505 – 901645 of the FAM18C genome corresponding to NMC0903 the *opaD* pseudogene described above. Based on these data, primers that flank the *opaD* insertion were designed for PCR amplification.

4.7.2 *opa* PCR and Sequencing

Based on the contig from the boneyard SMGA, primers were designed to amplify from the 5' and 3' ends of IS1301 and the *opa* gene. When NM216 was amplified with the OpaD-F and the JE1XBAI inverse IS1301 primer, a 500bp band was produced (lane 7) which was expected based on the contig. When the DNA was amplified with OpaD-R and JE2XBAI inverse IS1301 primer a 400bp band was generated (lane 8) which was also expected based on the sequence of the contig. When the NM216 DNA was amplified with OpaD-F and OpaD-R a 700bp product was produced (lane 6) which was also seen when early clone (NM23) DNA was used (lane 2). It was expected that the early clone would have a smaller product than the late clone because the IS1301 fragment is not found in the early clone. In addition, when the early clone was amplified with the OpaD-R and JE2XBAI primers, a 600bp product was produced (lane 4) which was not expected because it was thought the JE2XBAI primer was specific to IS1301. The early clone and NM216 were amplified with IS1301 specific primers and NM216 produced a band at ~450bp characteristic of IS1301 (lane 9) but the early clone did not (lane 5). The reference strain

was also amplified with the *opa* primers and inverse IS1301 primers and had similar results to the NM216 late clone (Figure 26, lanes 10-13). This PCR confirms the whole genome sequence but still leaves questions as to how IS1301 fits within the *opa* gene.

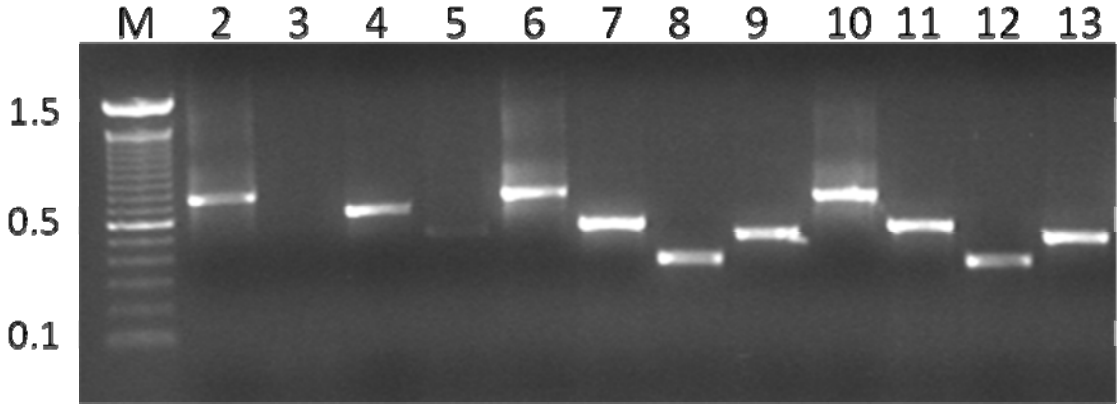

Figure 26. *opa* PCR with inverse IS1301 primers.
Molecular weight marker (M), 100bp ladder with sizes in kilobase pairs. NM23 (early clone) amplified with OpaD-F and OpaD-R (lane 2), OpaD-F and JE1XBAI (lane 3), OpaD-R and JE2XBAI (lane 4), IS1301F and IS1301R (lane 5). NM216 (late clone, sequenced genome) amplified with OpaD-F and OpaD-R (lane 6), OpaD-F and JE1XBAI (lane 7), OpaD-R and JE2XBAI (lane 8), IS1301F and IS1301R (lane 9). DE9246 (reference strain) amplified with OpaD-F and OpaD-R (lane 10), OpaD-F and JE1XBAI (lane 11), OpaD-R and JE2XBAI (lane 12), IS1301F and IS1301R (lane 13).

5.0 DISCUSSION

The increase in invasive meningococcal disease in Maryland in the mid-1990s from serogroup C *N. meningitidis* was characterized by the emergence of a "late clone" with an outer membrane protein profile (2:P1.5,2:F.3-6), *porA* deletions and variations in FetA. This study demonstrated that the late clone isolates belong to the ET-15 clonal complex, a hypervirulent clone. Previous research using representational difference analysis showed the late clone to have IS1301 while the early clone did not. Future studies could be conducted to ascertain whether or not the triplet is found across the United States and around the globe in serogroup C *N. meningitidis* isolates to understand the molecular epidemiology of the IS1301 globally.

Elias and Vogel also described the ET-15 clone in a grouping of isolates from Germany that contained multiple inserts of IS1301. Several of the integration sites were characterized in their study, one IS1301 fragment was specifically found within *nadA*, a candidate for a serogroup B vaccine (Elias & Vogel, 2007). The high molecular weight IS1301 triplet was also observed in their isolates but was not characterized in their study. This triplet is of epidemiological interest because it was found in the majority of their isolates, and in the majority of late clone isolates described in this study. This high molecular weight IS1301 triplet may represent genetic disruption of immunogenic determinants. IS1301 southern blot analysis of the late clone isolates showed IS1301 insertion as few as five times and as many as ten times within the genome. During the study period, there were three outbreaks in the state of Maryland and by IS1301

analysis the outbreaks had the same IS1301 banding pattern, consistent with the IS1301-RFLP protocol designed by Vogel (Elias & Vogel, 2007).

To characterize the IS1301 triplet, various PCR-based strategies: random hexamer strategy, linker based strategy, and inverse PCR, were designed and implemented in an attempt to clone the IS1301 triplet and identify the flanking regions. One additional method to isolate the flanking regions of IS1301 that was not a part of this study but is currently being explored is the creation of a library using the DNA from a late clone and screening the plasmids with multiple IS1301 probes. Initial data from this library screen show IS1301 integration within an *opa* gene.

The random primer strategy used a random primer and an IS1301 specific primer to amplify the IS1301 fragments and the flanking sequences. The cloning of the PCR product successfully cloned IS1301 but failed to include the flanking regions of the IS1301 fragments. This may be due to preferential cloning of smaller PCR products. To circumvent this problem a second cloning strategy was attempted using gel purified PCR product larger than 500bp to preferentially clone larger fragments. This protocol also failed to recover any IS1301 specific fragments with significant flanking sequence. Possibly due to the high background obtained with the digoxigenin labeled IS1301 probe used for the colony blot screenings. In future experiments using the random hexamer strategy, PCR products larger than 700bp should be targeted and enriched by nested PCR. In addition, a specific and accurate screening method, such as a radioactive probe, for locating IS1301 positive colonies is necessary.

The second PCR based method, the linker strategy, generated minimal data with the exception of one PCR product ~300bp in length having IS1301 flanking sequence. As demonstrated in the results section, the PCR product was mainly left over linker and specific

amplification of the linker ligated DNA did not succeed. The design of the linker or the ligation of the linker to the *Hha*I cut DNA could have been a problem, although positive and negative controls with the ligation of the linkers was successful. Another dilemma encountered with this strategy was not getting a high DNA yield after doing a phenol:chloroform extraction on the *Hha*I cut DNA which was performed to remove excess enzyme. In future experiments, a new linker should be designed and additional measures should be taken to reduce the amount of excess linker that is put into the PCR reactions.

Moreover, the inverse PCR strategy did not isolate the IS1301 fragments with the flanking insertion sites. The PCR product produced when two inverse primers were used was possibly due to the close proximity of two IS1301 fragments within the genome. The amplified PCR products could not be extracted with sufficient yield to permit sequencing. It is possible the primers bound non-specifically on the *N. meningitidis* genome and produced PCR product. If this strategy were to be continued, a better method of extracting the bands could be used. In addition, a Southern blot of the gel with the amplified product could be done to see if the PCR products contain IS1301. Another possible strategy would be to TA clone the PCR product and screen using an IS1301 probe to determine if the product is IS1301 specific and then sequence to identify flanking regions in positive clones.

Initial experiments with the whole genome sequencing show possible IS1301 insertion within *opaD*, encoding an outer membrane protein. Current experiments using PCR and sequence analysis are being performed to confirm the insertion of IS1301 in *opaD* and to determine if this insertion corresponds to one of the high molecular weight IS1301 triplet bands. The opacity proteins are outer membrane proteins which mediate meningococcal adhesion to the nasopharyngeal epithelium. Expression of the *opa* genes is known to undergo phase variation

enabling the bacterium to evade host immune responses. Thus, IS1301 may play a role in mediating *opa* gene phase variation, permitting clonal emergence by rendering naïve populations susceptible to infection. Future experiments are needed to fully ascertain the role IS1301 plays in *opa* expression and host immune response. Consideration should be given to IS1301 in relation to future vaccine designs against serogroup C *N. meningitidis*.

Overall, IS1301 is fascinating and the search for the flanking regions of the IS1301 triplet should be continued because it is epidemiologically interesting and may be pertinent to understanding the emergence of new clones and provide useful information in the design of new vaccines against *Neisseria meningitidis*.

BIBLIOGRAPHY

Active Bacterial Core Surveillance. (2006). Retrieved April 15, 2008, from http://www.cdc.gov/ncidod/dbmd/abcs/

Active Bacterial Core Surveillance Report, Emerging Infections Program Network, Neisseria meningitidis. (2007). Centers for Disease Control and Preventiono. Document Number)

Cameron, M. L., & Tsang, R. S. (2007). Analysis of phenotypic variants of the serogroup C ET-15 clone of *Neisseria meningitidis* by pulsed-field gel electrophoresis. *J Clin Microbiol, 45*(7), 2351-2352.

Capecchi, B., Adu-Bobie, J., Di Marcello, F., Ciucchi, L., Masignani, V., Taddei, A., et al. (2005). *Neisseria meningitidis* NadA is a new invasin which promotes bacterial adhesion to and penetration into human epithelial cells. *Mol Microbiol, 55*(3), 687-698.

Claus, H., Borrow, R., Achtman, M., Morelli, G., Kantelberg, C., Longworth, E., et al. (2004). Genetics of capsule O-acetylation in serogroup C, W-135 and Y meningococci. *Mol Microbiol, 51*(1), 227-239.

Davidsen, T., & Tonjum, T. (2006). Meningococcal genome dynamics. *Nat Rev Microbiol, 4*(1), 11-22.

Dolan-Livengood, J. M., Miller, Y. K., Martin, L. E., Urwin, R., & Stephens, D. S. (2003). Genetic basis for nongroupable *Neisseria meningitidis*. *J Infect Dis, 187*(10), 1616-1628.

Elias, J., & Vogel, U. (2007). IS1301 fingerprint analysis of *Neisseria meningitidis* strains belonging to the ET-15 clone. *J Clin Microbiol, 45*(1), 159-167.

Hammerschmidt, S., Hilse, R., van Putten, J. P., Gerardy-Schahn, R., Unkmeir, A., & Frosch, M. (1996). Modulation of cell surface sialic acid expression in *Neisseria meningitidis* via a transposable genetic element. *EMBO J, 15*(1), 192-198.

Harrison, L. H. (2006). Vaccine prevention of meningococcal disease: making slow progress. *Clin Infect Dis, 43*(11), 1395-1397.

Harrison, L. H. (2008). A multivalent conjugate vaccine for prevention of meningococcal disease in infants. *JAMA, 299*(2), 217-219.

Harrison, L. H., Jolley, K. A., Shutt, K. A., Marsh, J. W., O'Leary, M., Sanza, L. T., et al. (2006). Antigenic shift and increased incidence of meningococcal disease. *J Infect Dis, 193*(9), 1266-1274.

Hilse, R., Hammerschmidt, S., Bautsch, W., & Frosch, M. (1996). Site-specific insertion of IS1301 and distribution in *Neisseria meningitidis* strains. *J Bacteriol, 178*(9), 2527-2532.

Invitrogen. (2006). *TOPO TA Cloning Kit for Sequencing*. Carlsbad, California: Invitrogen Corporation.

Jelfs, J., Munro, R., Wedege, E., & Caugant, D. A. (2000). Sequence variation in the porA gene of a clone of *Neisseria meningitidis* during epidemic spread. *Clin Diagn Lab Immunol, 7*(3), 390-395.

Jolley, K. A., Brehony, C., & Maiden, M. C. (2007). Molecular typing of meningococci: recommendations for target choice and nomenclature. *FEMS Microbiol Rev, 31*(1), 89-96.

Kostelnik, L. M. (2006). *Representational Difference Analysis (RDA) for the Detection of Genetic Elements Associated with Increased Incidence of Serogroup C Neisseria meningitidis Infection.* Univeristy of Pittsburgh, Pittsburgh.

Mahillon, J., & Chandler, M. (1998). Insertion sequences. *Microbiol Mol Biol Rev, 62*(3), 725-774.

Meningococcal Disease: Frequently Asked Questions. (2008). Retrieved June 5, 2008, from http://www.cdc.gov/meningitis/bacterial/faqs.htm

Pollard, A. J. (2004). Global epidemiology of meningococcal disease and vaccine efficacy. *Pediatr Infect Dis J, 23*(12 Suppl), S274-279.

Promega. (2007). *Technical Manual: pGEM-T and pGEM-T Easy Vector Systems*. Madison, Wisconsin: Progema Corporations.

Rosenstein, N. E., Perkins, B. A., Stephens, D. S., Popovic, T., & Hughes, J. M. (2001). Meningococcal disease. *N Engl J Med, 344*(18), 1378-1388.

Snape, M. D., Perrett, K. P., Ford, K. J., John, T. M., Pace, D., Yu, L. M., et al. (2008). Immunogenicity of a tetravalent meningococcal glycoconjugate vaccine in infants: a randomized controlled trial. *JAMA, 299*(2), 173-184.

Tsang, R. S., Law, D. K., Henderson, A. M., Blake, M. L., & Stoltz, J. (2006). Increase in serogroup C meningococcal disease in Canada is associated with antigenic changes in the protein antigens of the ET-15 clone of *Neisseria meningitidis*. *J Infect Dis, 194*(12), 1791-1792; author reply 1792-1793.

Tyler, S., & Tsang, R. (2004). Genetic analysis of Canadian isolates of C:2a:P1.2,5 and B:2a:P1.2,5 *Neisseria meningitidis* strains belonging to the hypervirulent clone of ET-15. *Can J Microbiol, 50*(6), 433-443.

Tzeng, Y. L., & Stephens, D. S. (2000). Epidemiology and pathogenesis of *Neisseria meningitidis*. *Microbes Infect, 2*(6), 687-700.

Vogel, U., Claus, H., Frosch, M., & Caugant, D. A. (2000). Molecular basis for distinction of the ET-15 clone within the ET-37 complex of *Neisseria meningitidis*. *J Clin Microbiol, 38*(2), 941-942.