Genetic diversity of *Haemophilus influenzae* isolated from infections and healthy carriers in Iceland 2012

Jana Birta Björnsdóttir

Ritgerð til meistaragráðu
Háskóli Íslands
Læknadeild
Námsbraut í Lífeindafræði
Heilbrigðisvísindasvið

HÁSKÓLI ÍSLANDS

Erfðafræðilegur fjölbreytileiki *Haemophilus influenzae* meðal bera- og sjúkdómsvaldandi stofna á Íslandi 2012

Jana Birta Björnsdóttir

Ritgerð til meistaragráðu í Lífeindafræði

Umsjónarkennari: Helga Erlendsdóttir

Leiðbeinandi: Gunnsteinn Haraldsson

Þriðji maður í MS-nefnd: Karl G. Kristinsson

Læknadeild

Námsbraut í Lífeindafræði

Heilbrigðisvísindasvið Háskóla Íslands

Júní 2014

Genetic diversity of *Haemophilus influenzae* isolated from infections and healthy carriers in Iceland 2012

Jana Birta Björnsdóttir

Thesis for the degree of Master of Science

Supervisor: Helga Erlendsdóttir

Advisor: Gunnsteinn Haraldsson

Third in Masters committee: Karl G. Kristinsson

Faculty of Medicine

Department of Biomedical Science

School of Health Sciences

June 2014

Prentun: Samskipti

Reykjavík, Ísland 2014

Ágrip

Haemophilus influenzae (*Hi*) eru Gram neikvæðir stafir sem finnast í efri loftvegum manna. *Hi* getur bæði dvalið einkennalaust í nefkoki manna (berar), og orsakað sýkingar, þá sérstaklega þegar ónæmisvarnir hýsilsins eru skertar. *Hi* eru flokkaðir eftir tilvist fjölsykruhjúps, sex hjúpgerðum hefur verið lýst; a, b, c, d, e og f. Genamengi hjúpaðra *Hi* sýnir lítinn breytileika innan hverrar hjúpgerðar. Hjúplausir *Hi* (NT*Hi*) hafa engan fjölsykruhjúp og er genamengi þeirra mjög fjölbreytt. *H. haemolyticus* finnst einnig í efri loftvegum manna en er almennt ekki álitinn sýkingarvaldur. Genamengi *H. haemolyticus* er líkt genamengi NT*Hi*. Ein mest notaða aðferðin til að greina á milli þessara tveggja tegunda er myndun *H. haemolyticus* á beta hemólýsu á blóðagar. Erfðafræðilegar greiningar hafa sýnt að *H. haemolyticus* stofnar mynda ekki allir beta hemólýsu og þar af leiðandi hafa óhemólýtískir *H. haemolyticus* verið misgreindir sem NT*Hi* bæði meðal bera og sjúkdómsvaldandi stofna.

Bóluefnið Synflorix gegn 10 hjúpgerðum pneumókokka var innleitt í ungbarnabólusetningar á Íslandi 1. janúar 2011. Prótein D frá *Hi* er notað sem burðarprótein í bóluefninu og hafa rannsóknir sýnt að prótein D hvetji til ónæmissvars gegn *Hi* ásamt hjúpgerðum pneumókokka sem það ber.

Markmið rannsóknarinnar voru að kanna hlutfall *H. haemolyticus* og hjúpaðra *Hi* með PCR ásamt því að skoða genetískan fjölbreytileika *Hi* meðal 511 stofna sem áður höfðu verið greindir sem *Hi* úr klínískum sýnum á Sýklafræðideild Landspítalans 2012. Einnig voru allir 286 *Hi* stofnar sem ræktuðust úr nefkokssýnum leikskólabarna á höfðuborgarsvæðinu í mars 2012 með í rannsókninni. PFGE var gert á alls 303 stofnum, meðal stofna frá innsendum sýnum voru 110 stofnar sem orsökuðu eyrnasýkingar, 39 stofnar sem orsökuðu aðrar sýkingar, 10 stofnar úr nef og nefkokssýnum og meðal stofna frá leikskólabörnum var annar hver stofn greindur, alls 144.

PCR aðferð til aðgreiningar á *Hi* og *H. haemolyticus* var notuð til að leita að tilvist *fucK* og *hpd* genanna sem bæði eru vel varðveitt meðal *Hi* en eiga ekki að vera í *H. haemolyticus*. Sýndi aðferðin að enginn *H. haemolyticus* var meðal stofnanna. Hjúpgreining á *Hi* með PCR sýndi að einn hjúpaður stofn var meðal innsendu sýnanna, var hann af hjúpgerð e (0.2%) og meðal sýna frá leikskólabörnum voru tveir stofnar af hjúpgerð e (0.7%) og þrír af hjúpgerð f (1.1%). Af þeim 303 stofnum sem greindir voru með PFGE mynduðu 254 þeirra 72 PFGE klóna og 49 voru einfarar. PFGE klónar sem innihéldu sjö eða fleiri stofna voru aðeins níu talsins. Stofn af hjúpgerð f sýndi samskonar bandamunstur í PFGE og þrír NT*Hi* stofnar og annar stofn af hjúpgerð f var í sama PFGE klón og tveir stofnar af hjúpgerð e. MLST gaf 10 týpur meðal 12 stofna þar af var ein ný sem ekki hefur verið lýst áður.

Hlutfall *H. haemolyticus* er lægra hér á landi miðað við önnur lönd. Möguleg skýring er mismunur á hefðbundnum greiningaraðferðum sem eru notaðar. Hlutfall hjúpaðra *Hi* er svipað og í erlendum rannsóknum. PFGE greiningin sýndi mikinn fjölbreytileika meðal *Hi* stofna og er það líkt niðurstöðum annarra sambærilegra rannsókna. Í rannsókninni sást að margir PFGE klónanna innihéldu bæði berastofna (frá leikskólabörnum) og sjúkdómsvaldandi stofna frá ýmsum sýnatökustöðum. Er þetta birtingarmynd þess að ónæmiskerfi hýsilsins er mikill áhrifavaldur þess hvaða stofnar sýkja hýsilinn þar sem stofnar af sama PFGE klóni eru ekki bara að valda sýkingum heldur finnast líka í einkennalausum berum. MLST greining sýndi einnig mikinn fjölbreytileika. Frekari rannsóknir á

hjúpgerðum e, f og NT*Hi* stofnunum sem voru annaðhvort með eins bandamunstur eða flokkuðust saman í PFGE klóna eru mikilvægar. Heilgenaraðgreining væri þar besti kosturinn.

Niðurstöður þessarar rannsóknar skapa þekkingagrunn á erfðafræðilegum fjölbreytileika *Hi* áður en möguleg áhrif Synflorix bólusetninganna á *Hi* koma fram.

Abstract

Haemophilus influenzae (*Hi*) is a common colonizer in humans. In carriage *Hi* resides asymptomatically in the human respiratory tract. Because *Hi* can be an opportunistic pathogen, it may cause infections in the respiratory tract and sometimes cause invasive diseases especially in immunosuppressed hosts. *Hi* is either with a polysaccharide capsule (serotype a, b, c, d, e and f) that have a clonal population structure, or without a capsule (nontypable, NT*Hi*) that have a non-clonal population structure. *Haemophilus haemolyticus* also resides in the human nasopharynx and is normally not considered pathogenic. The genome of *H. haemolyticus* is similar to the genome of NT*Hi*. Genotypic methods show that not all *H. haemolyticus* produce hemolysis on blood agar (the most used differentiation method between the species) and that non-hemolytic *H. haemolyticus* have previously been misidentified as NT*Hi*, both among carriage and disease causing isolates.

The Synflorix vaccine against 10 pneumococcus serotypes has been a part of the childhood vaccination program in Iceland since January 1st 2011. Protein D from *Hi* is used as a conjugate in the vaccine. Studies have shown that Protein D as a carrier for pneumococcal polysaccharides also induces some immune response against *Hi*.

The objectives of this study were to examine the percentage of *H. haemolyticus* and capsulated *Hi* and to examine the genetic diversity of *Hi* among 511 isolates isolated from clinical samples sent to the Department of Microbiology, Landspitali in the year 2012, and all 286 *Hi* isolated from children's nasopharyngeal samples collected at day care centers (DCCs) in March the same year. The isolates had previously been identified as *Hi* or *Haemophilus* spp. with phenotypic methods. PFGE was done on 303 isolates, 110 isolates causing ear infections, 39 isolates considered causing other infections, 10 nasopharyngeal and nasal swab isolates and every other isolate from carriage samples, 144 in total.

PCR method identifying the *fuck* and *hpd* genes of *Hi* showed no *H. haemolyticus* among all the isolates (n=797). Results from Multiplex PCR for serotyping showed one capsulated isolate, of serotype e (0.2%), among the clinical isolates, and two serotype e (0.7%) and three serotype f (1.1%) among the isolates from DCCs. Of the 303 isolates typed with PFGE, 254 isolate formed 72 PFGE clones and 49 isolates were singletons. Isolates from two or more different specimen types were found in 45 PFGE clones. Nine PFGE clones contained seven or more isolates. One serotype f isolate had identical PFGE band pattern as three NT*Hi* isolates and another serotype f isolate grouped in the same PFGE clone as two serotype e isolates. A total of 10 different ST's were found among 12 isolates, thereof one new ST that had not been previously recognized.

The percentage *H. haemolyticus* is lower compared to other studies, possibly because of different phenotypic identification methods. The percentage of capsulated *Hi* is consistent to results of other studies. The high genetic diversity apparent in PFGE has been reported in other studies. Many of the PFGE clones contained isolates from different specimen types underlining the fact that the host's immune defenses play a large role in what strains cause disease. MLST results were in accordance to findings of other laboratories. Further analysis of the serotypes e, f and NT*Hi* isolates that were either

identical or in the same PFGE clone are needed, preferably with whole genome sequencing. The results of this study create a baseline on the diversity of *Hi* before possible vaccine effect of the Synflorix vaccine on *Hi* will take place.

Acknowledgements

First and foremost, I thank my supervisors Gunnsteinn, Helga and Karl for the opportunity to work within my field of interest. I would also like to thank them for their time, constructive criticism and encouragement during the course of this study.

I am grateful to all the staff of the Department of Clinical Microbiology for always being willing to share good advices with me and helping me when I needed it. Special thanks to Erla Soffía Björnsdóttir and Freyja Valsdóttir for their PFGE assistance.

I thank Mogens Kilian for sending us a reference strain of serotype d.

I deeply appreciate the support from my fiancé, Hólmsteinn Þór Valdimarsson and my family, thank you for your endless encouragement and for always believing in me, even at times when I didn't believe in myself.

This study was supported by Landspitali University Hospital Research Fund.

Contents

Ágrip 5

Abstract 7

Acknowledgements 9

List of figures 13

List of tables 14

List of abbreviations 15

1 Introduction 16

1.1 *Haemophilus influenzae* – general aspects 16

1.1.1 Polysaccharide capsule 16

1.1.2 Outer membrane 16

1.1.3 Outer membrane proteins 17

1.1.4 Other *Haemophilus* spp. 17

1.2 Virulence of *Hi* 17

1.2.1 Phase variation 18

1.2.2 Antigenic drift 18

1.2.3 Colonization and initiation of *Hi* infection 20

1.2.4 Capsule virulence 22

1.2.4.1 Increased virulence due to partial deletion of IS*1016*-*bexA* 22

1.3 Antibiotic resistance 23

1.4 Epidemiology of *Hi* 23

1.4.1 Importance of surveillance 23

1.4.2 Carriage in children 23

1.4.3 Otitis media 24

1.4.4 Invasive infection caused by capsulated *Hi* 25

1.4.5 Invasive infections caused by NT*Hi* 27

1.5 Vaccine development 28

1.5.1 Vaccine against *Hi*b 28

1.5.1.1 Vaccination against *Hi*b in Iceland 28

1.5.2 Vaccine against NT*Hi* 28

1.6 Phenotypic methods for identification of *Hi* 29

1.6.1 Immunological serotyping methods 30

1.7 Genotypic methods for identification of *Hi* 32

1.7.1 Analysis with PCR 32

1.7.2 Multiplex PCR 32

1.7.3 Differentiation between *Hi* and *H. haemolyticus* with PCR 33

1.7.4 Serotyping of *Hi* with PCR 33

1.7.5 Real time PCR 33
1.8 Methods for determining genetic similarity 34
1.8.1 Pulsed field gel electrophoresis 34
1.8.2 Multi locus sequence typing 34
1.8.2.1 MLST on *Hi* 35
1.8.2.2 PFGE and MLST on *Hi* 35
2 Objectives 37
3 Materials and methods 38
3.1 Isolates 38
3.2 Polymerase chain reaction (PCR) 40
3.2.1 DNA isolation for PCR 40
3.2.2 PCR for differentiation between *Hi* and *H. haemolyticus* 40
3.2.3 Phenotypic analysis of suspected *Haemophilus* isolates 40
3.2.4 PCR assay for serotyping *Hi* 41
3.2.5 PCR assay 41
3.2.6 Electrophoresis 42
3.3 Pulsed-field gel electrophoresis 42
3.4 MLST 43
3.5 Data analysis 44
3.6 Ethics 44
4 Results 45
4.1 Differentiation between *Hi* and *H. haemolyticus* 45
4.2 Serotyping of *Hi* 45
4.3 PFGE 46
4.3.1 Combination of specimen types in PFGE clones 47
4.3.2 Carriage isolates from children attending DCCs 48
4.3.3 Identical PFGE subclones 49
4.3.4 Capsulated *Hi* 51
4.4 MLST 52
5 Discussion 53
5.1 Differentiation between *Hi* and *H. haemolyticus* 53
5.2 Serotyping of *Hi* 54
5.3 Genotyping 56
5.4 Strength, weaknesses and improvement suggestions 58

6 Conclusions 61
References 62
Appendix I 71
Appendix II 72
Appendix III 76

List of figures

Figure 1. Interaction between *Hi* and host in the nasopharynx 21

Figure 2. *Cap* gene locus in *Hi*b where the whole locus is duplicated except a partial deletion of the first *bexA* gene. 22

Figure 3. Classification of results in slide agglutination test.....31

Figure 4. An example of PFGE results. Wells number 1, 14 and 28 contain size references and all other wells contain *HI* isolates from clinical samples. 46

Figure 5. PFGE banding pattern of two similar isolates, isolate 229 from DCC carriage and isolate e1 from blood. 47

Figure 6. PFGE of two ear isolates from the same person taken 49 days apart.....50

Figure 7. PFGE of two ear isolates from the same person taken 9 days apart.....50

Figure 8. PFGE of three tracheal aspirate isolates from the same person taken 4 and 48 days apart. 50

Figure 9. Serotype f isolate from carriage isolate identical in PFGE to three NT*Hi* isolates. 51

Figure 10. Serotype f isolate from carriage isolate in the same PFGE clone as two serotype e isolates. 51

List of tables

Table 1.Virulence factors of *Hi* and their role. 19

Table 2.The results of nasopharyngeal carriage in children attending DCCs in four countries. 24

Table 3. Age of patients with invasive *Hi* infection the year 1983-1989 in Iceland 25

Table 4. Diagnosis, serotype and time interval of infection due to *Hi*b and non *Hi*b in Iceland..... 27

Table 5: Identification of different *Haemophilus* spp. based on factor test and beta-hemolysis 30

Table 6. Isolates used for development of the method to differentiate between *Hi* and *H. haemolyticus*. 38

Table 7. Reference strains used for development of the serotyping method. 39

Table 8. Specimen type of isolates, number of available isolates from 2012, proportion that was analyzed with PFGE and proportion of each specimen type analyzed with PFGE. 39

Table 9. Age of subjects whose isolates were analyzed with PFGE. 40

Table 10. Primers used in the PCR assay for differentiation of *Hi* and *H. haemolyticus* and sizes of their amplicons. 40

Table 11: Primers used in the PCR assay for serotyping *Hi.* 41

Table 12. Primers for MLST. 43

Table 13. Presence of *fucK* and *hpd* genes in the *Hi* isolates from clinical samples and *Hi* isolates from carriage samples. 45

Table 14. Number of capsulated *Hi* and NT*Hi* isolates found among carriage samples and what DCCs they were found in. 46

Table 15. PFGE clones with seven or more isolates, number of subclones, number of isolates of each specimen type and the percentage of total in each PFGE clone. 48

Table 16. PFGE clone and PFGE subclone number of isolates found in two or more children attending DCCs 49

Table 17. Identical PFGE subclones found among PFGE analyzed clinical isolates sent to the Department of Clinical Microbiology the year 2012, specimen type and day between isolation of the isolates. 50

Table 18. Isolates typed using MLST, specimen type, PFGE clone and sequence type. 52

List of abbreviations

adk	Adenylate kinase
AOM	Acute otitis media
ATP	Adenosine triphosphate
atpG	Glyceraldehyde-3-phosphate dehydrogenase
COPD	Chronic obstructive pulmonary disease
frdB	Fumarate reductase iron-sulfur subunit
fucK	Fuculose kinase
Hi	*Haemophilus influenzae*
*Hi*b	*Haemophilus influenzae* serotype b
IgA	Immunoglobulin A
IS	Insertion sequence
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
mdh	Malate dehydrogenase
MLST	Multilocus sequence typing
NAD	Nicotinamide adenine dinuclotide
NT*Hi*	Noncapsulated *Haemophilus influenzae*
PBP	Penicillin binding protein
PCR	Polymerase chain reaction
PFGE	Pulsed field gel electrophoresis
pgi	Glucose-6-phosphate isomerase
PRP	Polyribosyl ribitol phosphate
recA	Recombinase-A
SLPI	Secretory leukocyte protease inhibitor
ST	Sequence type

1 Introduction

1.1 *Haemophilus influenzae* – general aspects

Bacteria of the *Haemophilus* genus are Gram-negative coccobacilli. They share a requirement for blood-derived factors, hence the name *Haemophilus*. Out of all *Haemophilus* spp. infections *Haemophilus influenzae,* (*Hi*) is the most common pathogen and is only found in association with humans (1). *Hi* was first described by Richard Pfeiffer in 1892. While studying an influenza outbreak in America he noticed that the bacteria were present in sputum samples from influenza-infected patients and therefore considered *Hi* to be the cause of influenza. It was not until 1933 that Smith, Andrews and Laidlaw showed that influenza was caused by a virus and *Hi* followed or was present before the viral infections (2, 3).

In 2000 it was estimated that a minimum of eight million cases of serious *Hi* disease occurred worldwide, with 370.000 resulting in death, the majority of fatal cases were in developing countries (4, 5). A conjugated vaccine targeting serotype b has almost eradicated invasive diseases caused by serotype b in countries where the vaccine is widely used. Before vaccination, serotype b caused up to 96% of all *Hi* invasive diseases (6). Nowadays the most common *Hi* are noncapsulated *Hi* (NT*Hi*). Along with being the most common cause of acute otitis media (AOM) after pneumococcus, it can cause invasive disease, especially among immunosuppressed people (7). NT*Hi* are found in the respiratory tract and spread through respiratory droplets from colonized people or through direct contact (1).

1.1.1 Polysaccharide capsule

Hi are classified according to the presence of a polysaccharide capsule. The capsule can protect against phagocytosis because the strong negative charge and chemical composition of the polysaccharides can inhibit complement binding to the surface (8, 9). Immunological differences of the polysaccharide capsule are used to classify *Hi* into serotypes. A total of six serotypes have been described: serotype a, b, c, d, e and f (10).

1.1.2 Outer membrane

The outer membrane of Gram-negative bacilli is made from phospholipids. Many proteins and lipopolysaccharides (LPS) are connected to this membrane (11). LPS are glycolipids consisting of core oligosaccharide, lipid A and O specific polysaccharide (also known as O antigen). The length of the O antigen area correlates with the bacterial cell surface negative charge (12).

A group of Gram-negative bacilli that reside in the mucosal surface have lipooligosaccharides (LOS) instead of LPS and *Hi* is one of them. The main difference between LPS and LOS is that LOS lacks the O specific polysaccharides. This results in a high diversity of LOS, especially in the polysaccharide region. This diversity enables the bacteria to express different isoforms of LOS, which increases the possibility of them being unrecognized by the hosts immune defenses (11, 12).

1.1.3 Outer membrane proteins

Most *Hi* outer membrane proteins increase their virulence, especially by enabling adherence to the nasopharynx. Outer membrane proteins of *Hi* have been studied with the goal of finding a vaccine target. High antigenic variability and variability in preservation have proven to be the main problems in finding a comprehensive vaccine target against an outer membrane protein on *Hi* (13). It is very important that the vaccine target has little antigenic variability to ensure the effectiveness of the vaccine against a wide group of *Hi* strains (14).

One of the most studied outer membrane proteins is protein D. It is coded for by the *hpd* gene and is well-preserved in *Hi* (15). Unlike most outer membrane proteins on *Hi* protein D shows little antigenic variation between strains, making it a candidate for a target in vaccine against NT*Hi* (14).

1.1.4 Other *Haemophilus* spp.

A number of other *Haemophilus* spp. have been described. *Haemophilus parainfluenzae* has been reported to cause pneumonia, endocarditis, meningitis and urinary tract infection both in children and adults (16). Carriage of *H. parainfluenzae* in children is also possible (17). *Haemophilus aegyptius* (also known as *Haemophilus influenzae biogroup aegyptius*) causes conjunctivitis and a more serious infection in children of ages 1-4 years, called Brazilian purpuric fever. The symptoms are meningitis, septicemia and purpuric skin lesions (18). *Haemophilus ducreyi* causes a sexually transmitted infection called chancroid, manifesting as genital lesions (16). *Haemophilus parahaemolyticus* is rarely the cause of infection; however, it is found in some humans nasopharynx and studies have shown it can cause infections in pigs (19).

Haemophilus haemolyticus can reside in the human nasopharynx. Until recently it was considered only to be a human commensal and very rarely the cause of invasive disease. The genome of *H. haemolyticus* is very similar to the genome of *Hi*. When analyzing with phenotypical methods the only factor differentiating the two species is the hemolysis *H. haemolyticus* produces on blood agar. Results from genotypic studies have shown that non-hemolytic *H. haemolyticus* can be misidentified as *Hi* when using only phenotypic methods. Therefore it can be assumed that non-hemolytic *H. haemolytius* is a more common cause of invasive disease than previously reported (20, 21). For example, an American study showed that seven of 374 isolates from invasive diseases previously diagnosed as NT*Hi*, were in fact *H. haemolyticus* when genotypic analysis was conducted. Before the study the literature showed only two previous cases of invasive *H. haemolyticus* (22).

Other *Haemophilus* spp. have been described, such as *Haemophilus segnis*, *Haemophilus aprophilus* and *Haemophilus paraphrohilus* but the rate of carriage and infection is extremely low (1).

1.2 Virulence of *Hi*

The virulence factors of *Hi* make the bacteria fit to attach to host cells or the host's surfaces for colonization, evade the host's innate and adaptive immune response and obtain nutrients in its surroundings essential for replication (23). *Hi* is a successful colonizer of the nasopharynx and a pathogen because it does not rely on any single mechanism to attach to the host and, furthermore has a quick response to the host's immune defense mechanism with phase variation and antigenic drift

(13). A universal virulence gene or characteristic has not been found to be associated with all the disease-causing isolates, supporting the fact that NT*Hi* is genetically diverse (24).

1.2.1 Phase variation

Phase variation helps the bacteria adapt to the host's environment by turning on and off genes that encode for factors in its outer membrane. Phase variation is caused by slipped-strand mispairing due to simple sequence repeats (SSR) in the coding or upstream promoter region of genes. These sequence repeats are unstable and can cause an error in the translational frameshift or change the distance that the promoter covers, thus altering the gene expression (23). Many genes of *Hi* have been found to be phase variable, for further information see Table 1.

1.2.2 Antigenic drift

Antigenic drift enables *Hi* to evade the host's immune response. The antigenic drifts are due to the switch, deletion or addition of amino acids and can also occur after horizontal gene transfer from a nearby strain. These changes can result in structural changes in the immunodominant area and therefore affect the binding of previously produced antibody to the antigen. Virulence factors of *Hi* have been reported to have antigenic drift (Table 1) (11).

Table 1.Virulence factors of *Hi* and their role.

Virulence factor		Role	Citation
Pili	Protein complex	Helps with adherence to epithelial cells.	(25)
Hap*	Outer membrane protein	Increases adherence to epithelial cells and other bacteria. When close to secretory leukocyte protease inhibitor (SLPI) degrading of Hap is inhibited, causing increased adherence, aggregation of bacteria and formation of microcolonies.	(26)
Hia	Outer membrane protein	Promotes adherence to epithelial cells through binding domains HiaBD1 and HiaBD2.	(27)
HMW1 and HMW2	Outer membrane protein	Increases adherence to epithelial cells.	(8, 11)
IgA1 protease	Proteolytic enzyme	Inactivates Immunoglobulin A by cleavage.	(28, 29)
Lipid A*	Glucosamine-based saccharolipid	Endotoxin activates cells through Toll-like receptor 4 (TLR4) and induces inflammation response of the innate immunity. High concentration of Lipid A in the body can cause septic shock.	(12, 30)
LOS	Glycolipid	Helps adherence to epithelial cells, weakens ciliary function in the nasopharynx.	(8, 31)
OapA	Outer membrane protein	Increases adherence to epithelial cells and possibly other eukaryotic cells.	(11, 32)
P2	Outer membrane protein	Serves as a porin, lets molecules up to 1400 Da in size through the membrane. Can also bind to mucin in the nasopharynx.	(13, 33)
P5	Outer membrane protein	Increases adherence to epithelial cells. Binds to CEACAM-1 that has increased expression during viral infections.	(11, 34)
Protein D*	Outer membrane protein	Indicators that it serves a role in impairing the ciliary function in the nasopharynx.	(11, 14)

*No reports found on the factor being susceptible to phase variation or antigen drift.

1.2.3 Colonization and initiation of *Hi* infection

For *Hi* to colonize the nasopharynx and cause disease, the interplay of virulence factors and host's factors is required in many different steps (Figure 1). When the host is exposed to allergic disease, viral infection or other types of immune suppression, *Hi* can go from its colonization state in the nasopharynx and spread to other sites in the respiratory tract and cause disease (11).

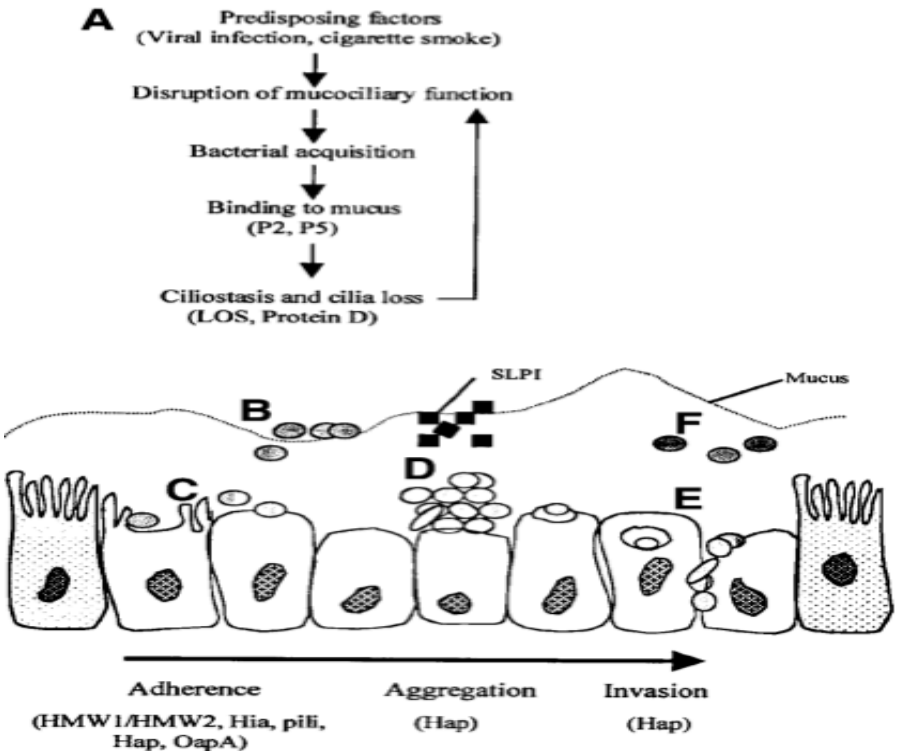


Figure 1. Interaction between *Hi* and host in the nasopharynx (11)**.**

A, B: *Hi* enters the nasopharynx and meets the mucociliary escalator that mostly clears it out. Viral infection, cigarette smoke and immune suppression of some kind reduces the clearance ability of the mucociliary escalator. *Hi* that is not cleared out binds to mucin through the adherence factors P2 and P5. LOS, protein D and possibly other factors damage the cilia, which further impairs the mucociliary escalator. *Hi* can then continue downwards.

C: *Hi* comes in contact with the respiratory epithelium via adherence factors like pili, HMW1, HMW2, Hia, Hap and OapA.

D: In the presence of SLPI, a protective factor that is secreted by the respiratory epithelium, Hap promotes adherence and mediates bacterial aggregation. SLPI then encourages infection instead of preventing it (11, 26).

E: Hap and possibly other factors mediate bacterial aggregation into epithelial cells and subepithelial space, providing the bacteria protection against the host's immune response. *Hi* can take up iron and heme, which allows *Hi* to stay alive in the respiratory mucosa despite nutrient deficiency.

F: When the concentration of SLPI is low, Hap is autoproteolysed, which causes *Hi* to spread throughout the respiratory tract (11). IgA1 protease, phase variation and antigen drift can make *Hi* invisible to the host's immune response. In colonization, *Hi* gets past IgA1 that is secreted by the respiratory epithelium to prevent adherence and invasion of bacteria with IgA1 protease that inactivates IgA. It has been shown that Hap can induce microcolony formation if point mutation is in the gene. This is possibly a contributor to *Hi* ability to stay in the nasopharynx and survive the attack of protective factors secreted by the respiratory epithelium (11, 35).

1.2.4 Capsule virulence

The capsule of *Hi* serves as an important virulence factor. Capsule material can act as interference to binding of opsonizing antibodies and complements to the bacterial surface. Polysaccharides in the capsule provide it with strong negative charge that can be electrostatically repelling for phagocytic cells (8).

The *cap* gene locus codes for capsule formation in *Hi*. The locus is made up of three regions, region I and region III are similar between serotypes, they are necessary for processing of the capsule content and exporting to the surface. Region II contains genes that take part in the biosynthesis of polysaccharides and is variable between serotypes (11).

1.2.4.1 Increased virulence due to partial deletion of IS*1016-bexA*

A group of capsulated *Hi* have IS*1016* insertion sequences inside the *cap* gene locus (36). Insertion sequences are small DNA fragments capable of inserting at various sites in a chromosome, making certain events more likely to occur in the bacterial chromosome (37).

The *cap* gene locus is commonly duplicated in *Hi*. Tandem duplication of the *cap* locus in *Hi*b is characterized by a partial deletion of the other *bexA* gene that contains an IS*1016* insertion sequence; IS*1016-bexA* (Figure 2). In homologous recombination on these loci, two events can occur. Homologous recombination causes one of the *cap* gene loci to be deleted because of specific recombination events. The *cap* gene locus that remains contains a partial deletion of IS*1016-bexA* and cannot produce a capsule. These strains are named with a minus sign after their serotype, for example *Hi*b-. The other event that can occur is that the presence of IS*1016* in homologous recombination can lead to additional *cap* gene loci being expressed in the cell. Additional *cap* gene loci mean the bacterium produces more capsule, leading to even more resistance to the hosts immune defenses (11, 36, 38).

Partial deletion of IS*1016-bexA* has also been noted in invasive strains of serotype a, e and f (38, 39). There are indications that these strains are more virulent than other strains of the same serotype. The epidemiology of these strains needs to be studied because the eradication of *Hi*b following vaccination can possibly influence the frequency of these strains (39).

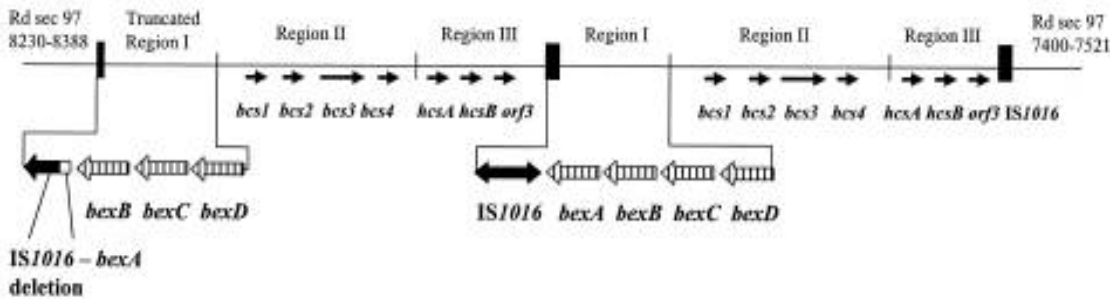


Figure 2. *Cap* gene locus in *Hi*b where the whole locus is duplicated except a partial deletion of the first *bexA* gene (36).

1.3 Antibiotic resistance

Penicillin-binding proteins (PBPs) are crucial for bacterial cell wall synthesis, beta-lactam antibiotics bind to the PBPs resulting in cell death (40). The frequent use of beta-lactam antibiotics has resulted in *Hi* developing resistance to beta-lactam agents (41).

Resistance to beta-lactam antibiotics can occur for one or more of the following reason: 1) A change in PBPs causing an altered target with less or no binding to the antibiotic; 2) A change in outer membrane porins resulting in limited ability of the drug to penetrate across the membrane and to the PBPs; 3) Bacteria produce one or more beta-lactamases enzymes that inactivate the beta-lactam molecules with hydrolysis. Plasmid encoded beta-lactamase like TEM-1 and ROB-1 are usually the reason for beta-lactam resistance in *Hi*. Those isolates are called beta-lactamase positive ampicillin resistant (BLPAR) (41, 42). These strains are normally resistant to ampicillin and amoxicillin but are susceptible to beta-lactamase inhibitor combined with beta-lactam antibiotic such as amoxicillin with clavulanic acid (41). The most common resistance mechanism in the beta-lactam negative ampicillin resistant (BLNAR) category is due to a mutation in the *fstI* gene that leads to a change in the PBP3, making *Hi* less susceptible to penicillins and cephalosporins (42). Some beta-lactamase producing *Hi* are resistant to amoxicixillin/clavulanic acid (BLPACR), which is most likely due to a combination of the mechanisms behind the BLPAR and BLNAR (40, 43).

Though resistance to beta-lactam antibiotics is the most common type of resistance among *Hi*, resistance or lower susceptibility to azalide, ketolide, macrolide, tetracycline and quinolone has been described in several cases. Epidemiological changes of *Hi* call for continued surveillance of susceptibility to antibiotics (44).

1.4 Epidemiology of *Hi*

1.4.1 Importance of surveillance

Epidemiologic surveillance of *Hi* is very important, especially to monitor if serotype replacement occurs as a result of the *Hi*b vaccination (45). Knowledge of which serotypes cause diseases and their prevalence makes it possible to assess if new preventive measures are necessary (46).

1.4.2 Carriage in children

Hi is a member of the commensal flora of the human nasopharynx in both children and adults (11, 47). Prevalence of *Hi* colonization varies, according to different studies (Table 2). Carriage of *Hi* is highest in children and much lower or non-existent in adults. The time of carriage of one genotypically distinct strain varies from two weeks to five months (48).

In Table 2 the results of nasopharyngeal carriage in children attending daycare centers (DCCs) from four different studies are listed. The *Hi* isolates in the Japanese study were not serotyped and none of the children had been vaccinated against *Hi*b (49). A recent Icelandic carriage rate study showed a high carriage rate, 62% of children attending DCCs carried *Haemophilus* spp. in the year 2012 (50) (Table 2), results of further analyses of these isolates will be presented in this thesis. Carriage rates might possibly vary between similar studies due to frequent antibiotic use, children's

surroundings and seasonal changes (51). Capsulated isolates are in general uncommon in nasopharyngeal carriage. Serotypes e and f are the most frequent capsulated strains found in carriages according to several studies, (51, 52).

Table 2. The results of nasopharyngeal carriage in children attending DCCs in four countries.

Country	**% of NT*Hi***	**% of capsulated**	**Citation**
Brazil	23	other than *Hi*b: 8	(51)
France	40	other than *Hi*b: <1	(52)
Japan	47 of both NT*Hi* and capsulated *Hi*		(49)
Iceland	62 of suspected *Hi*		(50)

The effect of 7-valent pneumococcal vaccination on *Hi* nasopharyngeal carriage has been examined by van Gils *et al.* in the Netherlands. Their findings showed no difference between the nasopharyngeal carriage of *Hi* in vaccinated and unvaccinated children during the first two years of life (53).

1.4.3 Otitis media

There are three main classifications of otitis media. Acute otitis media (AOM) is defined as the presence of middle ear effusion along with the rapid onset of one or more symptoms of inflammation like otalgia, otorrhea, irritability and fever. Otitis media with effusion (OME) is when there are no signs of acute infection in the ear but fluid present in the ear. Chronic suppurative otitis media (CSOM) is when there is chronic inflammation (six weeks to three months) in the middle ear and discharge (54).

AOM is a common childhood infection and a frequent cause for pediatric visits and antibiotic use. Children are at the most risk for AOM because of their immature specific immune system and because their Eustachian tubes are smaller and more level than those of adults. Around 80% of children in USA have experienced at least one episode of AOM by three years of age (54). NT*Hi* are the leading cause of AOM in children after pneumococcus and account for 25-35% of all AOM in children. Nasopharyngeal colonization is the precursor of otitis media and children who are frequently colonized with NT*Hi* are more likely to get AOM. Along with colonization, other risk factors for AOM, include, DCC attendance, male gender, passive smoke, having an otitis prone sibling and a previous viral infection. When comparing clinical manifestations of NT*Hi* to pneumococcus it has been noted that they are different. NT*Hi* are less likely to cause fever and discharge from the ear; however, the symptoms are insufficient to distinguish between the bacteria. Therefore obtaining a culture of middle ear fluid is the only way to determine what bacterium is causing AOM. When a child has more than four episodes of AOM in one year or has middle ear effusion eight months of the year it is classified as otitis prone (7, 35). Studies have shown that NT*Hi* is a frequent cause of repeated AOM. Insertion of tympanostomy tubes is the most effective treatment for repeated AOM (55).

Almost all *Hi* found in AOM are NT*Hi*. Therefore the start of vaccination against *Hi*b had no visible effect on the prevalence of AOM due to *Hi* (54). Since 2000, most infants in USA received a 7-valent pneumococcal vaccine. Studies comparing the pathogen distribution before and after the initiation of

the vaccine effect showed an increase in the proportion of AOM caused by NT*Hi* while AOM due to pneumococcus was significantly lower, as was expected (56, 57).

1.4.4 Invasive infection caused by capsulated *Hi*

Before vaccination against *Hi*b began in the late 1980s, infections due to *Hi*b were as high as 93-96% of all *Hi* infections in children in Scandinavia, UK and USA (58-60). In Iceland before the initiation of the *Hi*b vaccination in 1990, the overall incidence of invasive *Hi*b infection was 6.4/100.000/year (60). Children under five years of age are at a greater risk of getting *Hi*b infection because of their immature immune response to polysaccharides. This can clearly be seen when the rate of invasive *Hi*b disease is viewed by age (Table 3) (60). In Iceland the age specific incidence of invasive *Hi*b disease for the years before the initiation of the *Hi*b vaccination, 1983-1989 the incidence of *Hi*b disease among children under the age of one was 109/100.000/year, but only 41/100.000/year among children from one to five years of age (Table 3) (60).

Table 3. Age of patients with invasive *Hi* infection the year 1983-1989 in Iceland (60).

Age	***Hi*b**	**Cases per 100.000 population per year**	**Non *Hi*b***	**Cases per 100.000 population per year**
	n		n	
<1	32	109,0	2	6,8
1 to 5	6	41,1	1	0,7
6 to 20	5	1,1	2	0,5
21-40	1	0,2	1	0,2
41-60	4	1,3	2	0,6
>60	5	2,2	6	2,7

* Non *Hi*b includes NT*Hi* and all serotypes except b

In a summary on worldwide data from 1970-1990 on *Hi*b infection that affected the neurological system and normally sterile sites, the age-specific incidence among children under the age of five was 71/100.000/year. The age-specific incidence was much lower among adults, only 3/100.000/year (61).

Aboriginal populations tend to have increased susceptibility to invasive disease caused by capsulated bacteria and also a lower immunization effect. The reasons are not well understood, but both genetic and environmental factors have been speculated in previous studies. In the pre-*Hi*b vaccine era, *Hi*b infections in Navajo and White Mountain Apache children were three to five times more common compared to rates in other children in the US population. Post-vaccination the rates among Navajo and White Mountain Apache children under the age of two decreased but still remain up to twenty times higher than in other children in the US (59).

After the initiation of the childhood *Hi*b vaccination, *Hi*b carriage and infection has almost been eliminated in countries where the vaccine is in routine use. By the year 2006, children in approximately 108 countries, mainly in the developed countries were vaccinated, representing 55% of all the children in the world (4).

International cooperation to monitor the effect of *Hi*b vaccination on other invasive *Hi* disease was started in 1996 (6). Their findings were published in a large European study on invasive *Hi* disease from the years 1996-2006. All 28 countries included in the study had vaccination programs against *Hi*b. Invasive disease due to other encapsulated *Hi* than *Hi*b were very low, with an annual incidence of only 0.036/100.000. Serotype e caused 21% and serotype f 72% of these infections throughout the study period. A total of 44% of these cases occurred in people over 65 years of age (6).

The majority of serotype a invasive infections occur in young children. Similar to *Hi*b, invasive isolates of serotype a with the IS*1016-bexA* partial deletion have been documented (39, 62). Incidence of serotype a invasive infections is low in Europe. The highest incidence is seen in children under the age of five and the incidence has been reported as 0.012/100.000/year (6). However, age specific incidence for the years 1991-2003 among Apache and Navajo children less than five years was 25/100.000/year. *Hi*b vaccination was widespread in Alaska since 1992, and no increase of invasive infections caused by serotype a was found in this population and similar ones after the introduction of *Hi*b vaccine (39, 62).

Serotype c and d are the serotypes of *Hi* that most seldom cause disease in Europe and results showed that these serotypes mostly effect the adult population (6). This is in accordance with an American study showing that three of seven serotype c and four of six serotype d invasive *Hi* infections were in persons 18-64 years of age (63). According to the previously mentioned European study, the mortality of invasive *Hi* infections caused by c and d is the lowest among all the serotypes, suggesting that these serotypes are not as virulent as the others. The combined case fatality rate of invasive infections due to other serotypes than b is 9%. This number is possibly affected by the fact that serotypes other than b often infect people with immune suppression (6).

An Icelandic study on invasive infections due to *Hi* spanning 25 years showed that serotype f was the most common capsulated *Hi* after *Hi*b (Table 4). The most common type of infection before the initiation of the *Hi*b routine childhood vaccination (1989) was meningitis, mostly caused by *Hi*b (Table 4) (60).

The current rate of non-*Hi*b infections is much lower than the rate of *Hi*b disease was before the introduction of the *Hi*b vaccine, and infections caused by non-*Hi*b did not increase after the vaccination against *Hi*b (58).

Table 4. Diagnosis, serotype and time interval of infection due to *Hi*b and non *Hi*b in Iceland (60).

Diagnosis	Serotypes (1983-2008)										1983-1989		1990-2008	
	a	b	c	d	e	f	NT*Hi*	*H. para-influenzae*	Non b	Un-certain	*Hi*b	Non *Hi*b	*Hi*b	Non *Hi*b
Pneumonia	1	12	1			6	16			4	9	4	3	24
Meningitis		58				3				1	55	2	3	2
Bacteremia	2	16			1	2	16	3			14	5	2	19
Epiglottitis		4				1					4			1
Cellulitis		10				2	1	1		1	10	1		4
Osteomyelitis or septic arthritis		10					1	1			9		1	2
Other		5		1			6	2	3		1	2	4	10
Diagnosis not available		8	1							2	7		1	3
Total	**3**	**123**	**2**	**1**	**1**	**14**	**40**	**7**	**3**	**8**	**109**	**14**	**14**	**65**

1.4.5 Invasive infections caused by NT*Hi*

NT*Hi* was not a common cause of invasive disease before widespread use of the vaccination against *Hi*b (59). The Advisory Committee for Immunization Practices in USA has recommended *Hi*b vaccinations since the beginning of the year 1991 (64).

After the adaptation of routine vaccination for *Hi*b, NT*Hi* became the most common of all *Hi* to cause invasive infections (59). Out of all invasive *Hi* infections in the US reported to the Active Bacterial Core surveillance (ABCs) in the year 1989, only 17% were caused by NT*Hi* but by 2008 they had increased to 68%. It is important to bear in mind that the annual incidence of *Hi* invasive infection decreased by 65% during the course of the study (4.4/100.000 down to 1.5/100.000) (65). The annual mean incidence of NT*Hi* invasive infection in Europe 1996-2006 was 0.15/100.000 (6). An increase in NT*Hi* invasive infection among people over 65 years of age has been reported but the rate of all *Hi* infections has not reached rates like those seen with *Hi*b before vaccination (6, 65).

In Iceland the annual incidence of invasive infection due to other than *Hi*b before vaccination against *Hi*b was 0.9/100.000. After the vaccination started (1990-2008) the annual incidence increased to 1.2/100.000, this increase is not statistically significant. In 40 out of 59 cases, NT*Hi* was the cause of the invasive infection (60). NT*Hi* tends to affect different age groups than *Hi*b. The incidence of NT*Hi* invasive infections is the highest among newborns and people over 65 years of age. Immune suppression or chronic conditions are often a preexisting condition in populations that develop NT*Hi* infections (6, 60). Pregnant women are significantly more likely to have NT*Hi* bacteremia than other women in the same age group, most likely due to the immunosuppressive effects of pregnancy (6, 66). The Icelandic study showed the incidence rate of invasive *Hi* disease among pregnant women to be 9.2/100.000/year while the rate among non-pregnant women was only 0.4/100.000/year (60).

1.5 Vaccine development

1.5.1 Vaccine against *Hi*b

In 1931 Avery and Goebel described how linking polysaccharides from bacteria to a protein resulted in an increased immune response against the bacteria (67). Antigens made in response to protein carrying the polysaccharide induce the helper T cell. This results in T cells signaling to carbohydrate-specific B cells which leads to a good immune response and immunological memory against the polysaccharide part of the vaccine (58, 68). Because young children have poorly developed B cell immunity, this discovery by Avery and Goebel has made it possible to develop many effective polysaccharide-conjugated vaccines, providing protection to vulnerable populations (58).

In the 1970s the first vaccine against *Hi*b was licensed (58). It was a polysaccharide vaccine without conjugation to a carrier protein. Immune response in children under 18 months showed lack of protection, leaving the age group with highest *Hi*b disease rate unprotected. This vaccine was in overall 50%-60% effective. The first licensed conjugated vaccine against *Hi*b had polyribosyl ribitol phosphate (PRP) conjugated to diphtheria toxoid (PRP-D). This vaccine proved to be 87% effective after 3 doses in a childs first year of life (69). Today there are three licensed vaccines against *Hi*b: PRP-T, which has PRP conjugated to a tetanus toxoid; HbOC, which has PRP conjugated to a mutated toxin (CRM197) from the *Corynebacterium diphtheria*; and the PRP-OMP, which has PRP conjugated to an outer membrane protein of *Neisseria meningitidis* group b. Comparison of the antibody responses after three doses of the vaccines showed PRP-T had the highest response (58).

1.5.1.1 Vaccination against *Hi*b in Iceland

Results of a Finnish study published in 1987 (70) were the main rationale for vaccination against *Hi*b in Iceland. The results showed 87% protection against *Hi*b disease in children less than 12 months of age after they had been vaccinated with three doses of PRP-D (ProHIBIT, Connaught, Ltd). Vaccination in Iceland was initiated in 1989, and other nearby European countries, aside from Finland initiated their vaccine program 2-4 years later (68, 69).

It is possible to administer *Hi*b vaccine along with other vaccines. Currently, children in Iceland are vaccinated with the Pentavac® vaccine which contains vaccines against pertussis, diphtheria, tetanus, *Hi*b and polio. Administration occurs at 3, 5 and 12 months of age and each dose contains 10 µg of PRP conjugated to a tetanus toxoid (71).

1.5.2 Vaccine against NT*Hi*

There is a basis for vaccination against NT*Hi*. Up to 35% of all AOM is caused by NT*Hi*. They are also commonly found in purulent secretions in people who suffer from chronic obstructive pulmonary disease (COPD) and cystic fibrosis. These patient groups would benefit most from vaccination against NT*Hi* because their quality of life is greatly affected by NT*Hi* infections. Genetic diversity and antigen drift of NT*Hi* has made it difficult to find a common antigen to use against the majority or all NT*Hi* in a vaccine (14). Protein D from *Hi* was chosen to be a conjugate for 10 pneumococcus serotypes

(Synflorix) because it is located on the surface, is well-preserved, has limited antigenic drift and is a virulence factor. Other important reason was the results of Prymula *et al.* when studying the effect of 11-valent pneumococcal vaccine that contained protein D showed the vaccine efficacy against *Hi* AOM to be 35% (14, 72). van der Berg *et al.* recently published results on the effect of 10-valent pneumococcal vaccine with Protein D from *Hi* conjugate on nasopharyngeal colonization of NT*Hi*. Their results showed no significant difference in NT*Hi* nasopharyngeal colonization compared to carriage of children vaccinated with 7-valent pneumococcal vaccine (73). All children born in Iceland after January 1st the year 2011 are vaccinated against 10 pneumococcal serotypes with the Synflorix vaccine (71). Synflorix also contains the protein D conjugate, it will be interesting to follow the effects of this vaccination on NT*Hi* in the future (74).

1.6 Phenotypic methods for identification of *Hi*

Phenotypic methods to identify *Hi* are used in clinical microbiology laboratories around the world. The process of identification takes 2-3 days, starting with culturing the sample on chocolate agar, which is enriched with nutrients vital for *Haemophilus* spp. growth. Optimum temperature for growing *Haemophilus* spp. is 37°C and growth can be induced with CO_2. Under these conditions, growth appears in 24 hours (1).

In Gram staining *Haemophilus* are identified as Gram-negative coccobacilli. When Gram staining is done directly on a patient specimen, the pink background color can make it difficult to identify *Haemophilus* because they fall into the background of the sample.

Haemophilus colonies on chocolate agar are small, soft and translucent. Capsulated isolates produce bigger and slimier colonies. If colonies only grow on chocolate agar, smell and look like *Haemophilus* colonies, the next step is to determine which *Haemophilus* species is present in the sample.

Factor test is used to examine isolates requirement of hemin, Nicotinamide adenine dinucleotide (NAD) or both for growth. The growth in a solution of 0.5 McFarland is plated on plain agar that is without nutrients. Factor disks, one containing hemin, one containing NAD and one containing both hemin and NAD, are placed on the agar and incubated overnight at 35-37°C in O_2. Growth around the factor disks is examined (Table 5) (1). If no growth appears, the test is repeated in CO_2 because some *Haemophilus* spp. require CO_2 for growth (75).

The majority of *H. haemolyticus* and *H. parahaemolyticus* produce beta-hemolysis. To identify beta-hemolysis, the isolate is plated a 5% blood agar beside *Stapylococcus aureus*. The isolate gets hemin from the blood agar and NAD from *S. aureus*. After one night in 35-37°C CO_2 the hemolysis is examined. A clear zone around the *Haemophilus* spp. colonies on the blood agar means that the isolate produces beta-hemolysis (1).

Table 5: Identification of different *Haemophilus* spp. based on factor test and beta-hemolysis (1).

Species	Hemin (X-factor)	NAD (V-factor)	Beta-hemolysis
H. influenzae	+	+	-
H. haemolyticus	+	+	+
H. parainfluenzae	-	+	-
H. parahaemolyticus	-	+	+

In Iceland, susceptibility tests on *Haemophilus* spp. isolates are done according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) standard (76). Isolates resistant to penicillin are checked for beta-lactamase with the Nitrocefin test (77). Nitrocefin is a Cephalosporin substrate containing a beta-lactam-ring. A drop of Nitrocefin is placed on filter paper with a loopfill that contains *Haemophilus* spp. growth. If the *Haemophilus* spp. isolate contains beta-lactamase, the Nitrocefins beta-lactam-ring disrupts resulting in a color change from yellow to red (78).

1.6.1 Immunological serotyping methods

Serotyping of *Hi* is a simple and highly important tool for identifying which serotypes are present in specific geographic areas and to evaluate the effect of *Hi*b vaccination around the world (58). Serotyping is commonly done with antisera in slide agglutination tests. The antisera can be polyvalent or specific to one particular serotype. Suspension of *Hi* growth and antisera are mixed together on a slide. When an antibody present in the antisera matches the serotype of *Hi* it binds to the capsule and causes agglutination of the *Hi* cells. Agglutination is classified into six categories depending on the appearance of agglutination on the glass slide (Figure 3) (79).

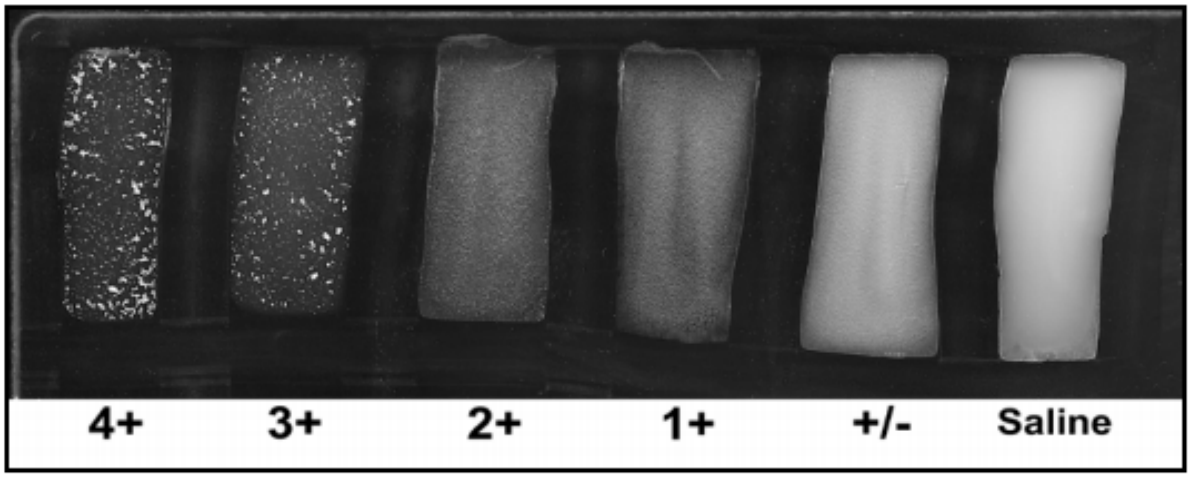


Figure 3. Classification of results in slide agglutination test (79).

4+: all cells agglutinate and the background appears clear.

3+: around 75% of the cells agglutinate and the background appears slightly cloudy.

2+: around 50% of the cells agglutinate and the background appears cloudy.

1+: around 25% of the cells agglutinate and the background appears cloudy.

+/-: 25% or less of the cells agglutinate and the background appears fine-granulated

0 (saline): no visible agglutination. The solution remains unchanged (79).

Isolates are classified as capsulated when the agglutination appears 3+ or 4+ in two minutes. Isolates are classified as non-capsulated when the agglutination appears 0, +/-, 1+ or 2+ with antisera specific for serotypes a, b, c, d, e and f (79).

Limitations of the slide agglutination test do raise concerns. Despite the use of control isolates and user manual, the test has been proven to give inconsistent results. Laclaire *et al*. found that 56 (40%) of *Hi* isolates sent to the Center for Disease Control (CDC) were serotyped incorrectly when they compared results from slide agglutination test to PCR results (80). The decline in *Hi*b disease resulted in laboratories performing the slide agglutination test less frequently, possibly one of the reasons for misidentification along with the fact that different categories of agglutination make the test very subjective (81). Troubles with non-specific agglutination or cross-reactions of some test kits have been reported (82, 83). Bokerman *et al.* compared two different slide agglutination tests and one polyvalent antiserum to PCR serotyping. The polyvalent antiserum proved least reliable, with 66% sensitivity and 92% specificity (83).

1.7 Genotypic methods for identification of *Hi*

1.7.1 Analysis with PCR

Polymerase chain reaction (PCR) is a method that allows for the detection of specific DNA or RNA. Substrate solution for the reaction contains isolated DNA, nucleotides, primers specific for the target gene and a DNA polymerase with temperature optimum around 70°C. The substrate solution is placed in a device that regulates the temperature in repeated cycles, resulting in amplification of the target gene if the primers are correctly located (84, 85). The cycle is repeated until there is enough of the amplified product (amplicon) if the target gene is present in the substrate solution to make it visible with electrophoresis. With electrophoresis it is possible to see if the amplicon is present and compare its size to DNA size markers. If the size of the amplicon and target gene match, the amplification was successful (84).

1.7.2 Multiplex PCR

Having two set of primer in a substrate solution is called a duplex. More than two set of primer in a substrate solution is known as a multiplex PCR. The advantages of a duplex or multiplex PCR are that it enables the use of an inner control to prevent false negative results due to a PCR obstruction in the substrate solution. Multiplex PCR is also less expensive and takes less time than using several monoplex PCR (85).

Certain factors need to be considered when designing a multiplex PCR. Primers need to be specific to the nucleotide sequence that is being examined so only the target sequence will be amplified. The amplicons cannot be of similar size if differentiation between the products is done with electrophoresis. Primers that are planned to be placed together in a multiplex PCR need to have a similar design regarding their length, G/C ratio and melting temperature. Each primer pair is first tested in a monoplex PCR to ensure that the amplification takes place at the selected temperature and the amplicon is of the correct size (85, 86).

The possibility of non-specific binding of the primers and artifacts increases with each added primer. Therefore a set of primer that work well in monoplex do not necessarily work with other primers in multiplex. Common problems in multiplex PCR are primer dimers. This happens when primers pair with each other and are amplified instead of the nucleotide sequence of the target gene. This also causes competition between the correctly placed primers and the primer dimers for nucleotides and use of the polymerase. This can sometimes be fixed with a Hot-start method, change of primers concentration and adjustments of cycles and anneal temperature (85, 86).

If tests show non-specific amplification, there is a reason to try Hot-start PCR. Non-specific amplification can happen when primers pair with each other and make primer dimers or the primers pair non-specifically with an incorrect nucleotide sequence in the substrate solution before the first cycle begins. To prevent non-specific annealing and activation of the polymerase in the beginning of the PCR reaction it is possible to use a special *Taq* polymerase in a Hot-start. The polymerase does not activate until the temperature has reached 95°C, so if the non-specific amplification is because of

premature annealing of the primers the Hot-start should fix it. If it does not work then the primer sequence needs to be reselected (85, 86).

1.7.3 Differentiation between *Hi* and *H. haemolyticus* with PCR

Several PCR methods to differentiate between *Hi* and *H. haemolyticus* have been developed. Among genes that have been tried as a target for differentiation are IgA protease gene (*iga) (87)*, lipo-oligosaccharide gene (*igtC)* (88) and pilus gene (*pilA)* (74). Detections of the Protein D gene (*hpd*) and fuculose kinase gene (*fucK*) have proven to be effective to distinguish between the two species because *H. haemolyticus* should not have these genes in its genome (22, 89).

When both genes are present in the genome it is a strong indicator that the bacterium is *Hi* because the genes are highly preserved in the *Hi* genome. In cases where only one of the two genes is present, the bacterium is also considered *Hi*. Because of high horizontal gene transfer it is very difficult to develop a PCR method that is 100% sensitive, but by using *hpd* and *fucK* together a more reliable identification than a method with detection of only one of the gene is possible (21).

1.7.4 Serotyping of *Hi* with PCR

Falla *et al.* developed a PCR serotyping method for *Hi*. To identify if *Hi* had a capsule they designed a primer pair targeting the *bexA* gene in region I of the *cap* gene locus that is present in all capsulated *Hi*. All six serotype specific primers were designed from region II of the *cap* gene locus, this region is unique for each serotype (a-f). First the strains were tested in a duplex PCR with primers for the *bexA* gene, serving as an inner control for a capsule and serotype b. This enabled identification of NT*Hi*, b- and b strains. The remaining strains were then tested with serotype-specific primers other than b in monoplex PCR (90).

1.7.5 Real time PCR

Real time PCR is when the amplification of the target gene and analysis is done in the same machine and at the same time. This method has an advantage over conventional PCR because it does not require electrophoresis of the PCR amplicon, making it less time consuming and minimizing the risk of contamination (85).

Real time PCR uses fluorescent probes or DNA dyes to detect amplification as it happens. The sooner the machine detects color or fluorescence, the more of the target DNA is in the sample.

Wroblewski *et al.* recently designed a two-step real time PCR method for serotyping *Hi*. The *bexA* primer they used was designed with regard to previously discovered deleted regions of the *bexA* gene (81, 91). The primers were divided into two reactions. The first reaction had primers for the *bexA* gene and serotype b. If the isolate was negative for both primers, it was determined as NT*Hi*. Isolates positive for *bexA* and serotype b were determined as serotype b. Isolates positive for *bexA* but negative for serotype b were tested in the second substrate solution with primers for serotype a,c-f (81).

1.8 Methods for determining genetic similarity

Because NT*Hi* needs to survive the environment of the human respiratory tract, it relies on selection and clonal expansion of certain strains instead of regulatory systems that are commonly used by other bacterial species (92, 93). This characteristic of NT*Hi* calls for exploration if certain genotypes are more prevalent than others. Frequent recombination of bacterial lineages in pneumococcus has previously been associated with vaccine escape and resistance to antibiotics; therefore it is possible that the same goes for NT*Hi,* especially because of its ability to take up DNA from the environment and integrate it into its chromosome (94). By comparing genetic information of different isolates, it is possible to discover if a certain DNA fingerprint or type is more prevalent in certain groups of patients, carriages or in a specific geographic area (94).

1.8.1 Pulsed field gel electrophoresis

Genetic relatedness can be assessed with Pulsed field gel electrophoresis (PFGE) where restriction enzymes cut the bacterial genome into variously sized DNA molecules (95). Conventional electrophoresis cannot precisely resolve DNA molecules larger than 40-50 kilobases because of their size independent co-migration (96). For electrophoresis of various sized DNA molecules, contour clamped homogeneous electric field electrophoresis was developed. It is based on periodically changing the direction of the electric field and therefore the direction DNA molecules move in the gel. Because small DNA molecules need less time to reorient than big DNA molecules, a gradual time increase of the electrophoresis from each direction (ramping) enables the separation of DNA molecules that have a wide variety of sizes (kilobase to megabase).

When preparing DNA for PFGE the aim is to get an intact chromosome with the right concentration and quality so the band patterns are well identifiable after electrophoresis. The protocol for DNA preparation is different depending on bacteria species. For instance, conditions for lysis are not the same for Gram-negative and Gram-positive bacteria (85). Restriction enzymes are selected with regard to suitable frequency of the restriction enzymes recognition sites in the chromosome of the bacteria. *SmaI* is the most commonly used restriction enzyme for *Hi*. The choice of the changing electrophoresis condition depends on the size of DNA molecules that are being separated. It is important to place a size ladder in the agarose gel to make it possible to estimate the size of DNA molecules. It is also common to use well-known bacterial genomes as a size standard, making the size calculation more accurate if the fragments have the same or a larger range of sizes as the bacteria being tested (85, 97).

Interpreting the band patterns and similarity between strains is usually done with the help of computer programs. Strains that have the same band patterns are classified as the same strain. The more differences there are between band patterns of strains, the less related they are (97).

1.8.2 Multi locus sequence typing

Multi locus sequence typing (MLST) is a method based on the sequencing of well-preserved “housekeeping” genes (98). For *Hi* these genes are seven, adenylate kinase (*adk),* glyceraldehyde-3-phosphate dehydrogenase *(atpG),* fumarate reductase iron-sulfur subunit *(frdB),* fuculose kinase

(fucK), malate dehydrogenase *(mdh),* glucose-6-phosphate isomerase *(pgi)* and recombinase-A *(recA).* Each unique sequence of these genes is assigned an allele number in order of discovery. Every isolate has a seven-figure number (allelic profile) and from that is given a sequence type (ST). Very closely related isolates have the same ST or ST that have little difference between their allelic profile hence; isolates that are unrelated have unrelated ST. The main advantage of MLST compared to PFGE is that it provides unambiguous results so comparison between study facilities around the world is possible through an online database (99).

1.8.2.1 MLST on *Hi*

A number of studies have used MLST to examine the genetic relationship between *Hi* strains. Samples were collected from children in New York, US at 9, 12, 15, 18, 24 and 30 months old and if a child was diagnosed with AOM a sample was collected. MLST was conducted on 165 NT*Hi* isolates from 73 children, and 70 different ST were found. The same ST was found in the ear of a child with AOM (37 children) and their nasopharynx in 84% of cases. When viewing STs from different time points, 85% of the NT*Hi* isolates had disappeared after 3 months (47).

Erwin *et al.* examined all the MLST of *Hi* that were in the online database in the year 2006 (a total of 195 different ST of NT*Hi*). Their results indicated that NT*Hi* did not show a clustering of genetic content by clinical isolation or geography, however, the capsulated *Hi* did show serotype-specific phylogenetic groups (24).

LaCross *et al*. found 109 different ST in 170 NT*Hi* isolates from children with AOM and isolates from the nasopharynx of healthy children from three countries (Finland, Israel and US). The ST found were divided into clonal complexes where the ST had to be identical in at least six of the seven MLST loci to be in the same clonal complex. The biggest clonal complex only had 5 ST, containing both isolates from carriages and otitis media from various geographic areas. All available NT*Hi* ST in the online MLST database (537 at the time) were analyzed into clonal complexes. The largest clonal complex that formed only consisted of 19 different ST, further validating previous reports on the great genetic diversity of NT*Hi* (100).

1.8.2.2 PFGE and MLST on *Hi*

There is no agreement between studies on which is better, PFGE or MLST. There are studies that reveal similar resolution given with MLST and PFGE, some studies reveal that MLST has less discriminatory power than PFGE and others reveal that MLST has more discriminatory power. This can depend on which genes are used in the MLST scheme and which restriction enzyme is used in PFGE; but because MLST does not catch the whole genome, PFGE can be considered to have more discriminatory power than MLST in those species that have a great amount insertion sequences, gene islands and other mobile elements (101).

There is limited amount of studies using both PFGE and MLST on the same *Hi* isolates. A total of 95 *Hi* isolates from Spanish patients with community acquired pneumonia were divided into 47 PFGE patterns, 21 PFGE clones were formed containing 2-15 isolates and 26 isolates were genotypically unique. MLST was done on all 95 isolates giving 67 different ST. Identical PFGE patterns had the same ST. Isolates classified in the same PFGE clone did not always have the same ST (102). A

similar study was conducted on invasive isolates from Manitoba, Canada. Out of 20 NT*Hi* isolates they found three PFGE clones, containing two isolates each. MLST of the 20 isolates revealed 18 different ST (103). Among nine NT*Hi* isolates that caused meningitis in Italy from 1997-2006, there were no clonal formations even though two isolates came from the same hospital within one month. Each of the strains had a different ST (104). A total of 28 invasive NT*Hi* isolates from Arkansas, US in 1993-2001 were analyzed with MLST. Interestingly, four separate pairs of isolates did share alleles in five out of seven genes. PFGE was performed on these eight isolates, confirming their close genetic relationship. Two pairs were identical and the other two had ≤3 bands difference. No relationship between the infected patients was found that could explain the similarities (105).

2 Objectives

The Synforix 10-valent pneumococcal vaccine uses protein D from *H. influenzae* as a conjugate. It has been a part of the childhood vaccination program in Iceland since the beginning of the year 2011. The effect of this vaccine on the genetic diversity of *H. influenzae* is unknown. This thesis will create a baseline on the diversity of *H. influenzae* before the vaccine effects take place.

The percentage of *H. haemolyticus,* capsulated and noncapsulated *H. influenzae,* in children attending DCCs is unknown in Iceland. Furthermore, genetic diversity of *H. influenzae* has also not been studied in Iceland.

Therefore the objectives of this thesis are to:

- Compare the percentage of *H. haemolyticus* in children attending DCCs in the capital region of Iceland in 2012 to results from disease causing isolates identified in clinical samples sent to the Department of Clinical Microbiology in 2012.
- Compare the percentage of capsulated *H. influenzae,* within the same groups.
- Develop a PFGE method and use it to examine the genetic diversity of *H. influenzae* in selected isolates from children attending DCCs in 2012 and disease causing isolates identified in clinical samples sent to the Department of Clinical Microbiology in 2012.
- Use MLST on selected isolates that represent clusters of genetically related isolates found with PFGE and compare the MLST results to results from laboratories around the world.

3 Materials and methods

3.1 Isolates

Hi isolates were obtained from clinical samples sent to the Department of Clinical Microbiology in 2012. All *Hi* isolates that were considered disease causing, a total of 521 *Hi* isolates were stored. A total of five were dead and five could not be found. All remaining 511 isolates were included in this study.

Nasopharyngeal samples from 465 children attending 15 DCCs in the Reykjavík capital region were collected in March of 2012. Samples were collected from five DCCs in Hafnarfjörður, five in Kópavogur and five in Reykjavík. The DCCs were marked in alphabetic order from A to O. Suspected *Haemophilus* spp. were found in 62% of the samples. Two isolates were found in 10 samples giving a total of 301 isolates. A total of five were dead or not found, leaving 296 isolates to be identified, 10 turned out not being *Hi*, giving a total of 286 *Hi* isolates.

Reference strains used for method development and control are listed in Table 6 and 7. Information on the number of isolates from each specimen type and the age of the subjects whose isolates were analyzed with PFGE are listed in Table 8 and 9. All isolates were frozen in tryptose-glycerol freezing medium (Difco Laboratories, USA) at -80°C.

Table 6. Isolates used for development of the method to differentiate between *Hi* and *H. haemolyticus*.

Species	**Isolates number**
Haemophilus parainfluenzae	070322-0075[1]
Haemophilus influenzae serotype b	070322-0082[1]
Haemophilus influenzae serotype f	090414-0119[1]
Haemophilus haemolyticus	120605-0009[2]

[1]Isolates from the Department of Clinical Microbiology collection.
[2]Isolate grown from patient throat sample.

Table 7. Reference strains used for development of the serotyping method.

Serotype	Reference strain
a	ATCC 9006[1]
b	ATCC 10211[1]
c	ATCC 9007[1]
d	HK-644[2]
e	ATCC 8142[1]
f	ATCC 9833[1]
b-	Fb 7132[1]
NT*Hi*	NCTC 11315[1]
NT*Hi*	ATCC 49247[1]

[1] Strains from the Department of Clinical Microbiology collection.
[2] Reference strain from the Aarhus University.

Table 8. Specimen type of isolates, number of available isolates from 2012, proportion that was analyzed with PFGE and proportion of each specimen type analyzed with PFGE.

Specimen type	Number of available isolates	Number of isolates analyzed with PFGE	% of PFGE analyzed isolates from the same specimen type	% of all PFGE isolates (n=303)
Abscess	1	1	100	0.3
Blood	2	2	100	0.7
Bronchial aspirate	25	2	8	0.7
Ear	257	110	43	36
Nasal smear	12	1	8	0.3
Nasopharyngeal	46	9	20	2.9
Nasopharyngeal DCC	287	144	50	48
Sinus	10	4	40	1.3
Sputum	129	23	18	7.6
Tissue	1	1	100	0.3
Tracheal aspirate	22	5	36	1.6
Women genital	2	1	50	0.3
Total	**794**	**303**	**38**	**100**

Table 9. Age of subjects whose isolates were analyzed with PFGE.

Agegroup of subject and specimen type	Number of isolates	Age of subjects		
		age median	average	range
Isolates from clinical samples adults >18	36	61	59	21-94
Isolates from clinical samples children under <18	123	1.9	2.2	0.3-12
Isolates from clinical samples children age 1.5-6.5	44	2.5	3	2-5
Isolates from children attending DCCs	144	2.6	4	1.5-6.2

3.2 Polymerase chain reaction (PCR)

3.2.1 DNA isolation for PCR

Isolates were grown on chocolate agar plates (Oxoid, Hampshire, UK) and incubated at 37°C in 5% CO_2 overnight. Two to five colonies were suspended in 500 µl of 5% Chelex 100® (Bio-Rad Laboratories, Hercules, CA, USA). The suspension was heated for 10 minutes at 100°C and then centrifuged at 14000 rpm for 10 minutes at 4°C. Around 200-300 µl of the supernatant containing the bacteria DNA was stored at -80°C until used in PCR.

3.2.2 PCR for differentiation between *Hi* and *H. haemolyticus*

The author had previously developed the method used for differentiation of *Hi* and *H. haemolyticus* (106). Primer specific for the *hpd* gene coding for protein D in *Hi* (107) and primer specific for the *fucK* gene coding for fuculose kinase in *Hi* (107, 108) were used together in a duplex PCR (Table 10). The method was tested on isolates (Table 6) and done on all 511 *Hi* isolates obtained from clinical samples and all 296 suspected *Haemophilus* spp. carriage isolates.

Table 10. Primers used in the PCR assay for differentiation of *Hi* and *H. haemolyticus* and sizes of their amplicons.

Primers	Sequence (5'-3')	Amplicon (bp)	Citation
*fucK*2-f	ATG GCG GGA ACA TCA ATG A	266	(107)
*fucK*1-*r*	AAG ATT TCC CAG GTG CCA GA		(108)
hpd#1-f	GAT TGG AAA GAA ACA CAA GAA AAA G	112	(108)
hpd#1-r	CAC CAT CGG CAT ATT TAAC CA		

3.2.3 Phenotypic analysis of suspected *Haemophilus* isolates

Carriage isolates negative in PCR for differentiation between *Hi* and *H. haemolyticus* were tested with phenotypic methods. Isolates were grown on chocolate agar plates (Oxoid) and incubated overnight at 37°C in 5% CO_2. Factor requirements for growth were examined with V, X and XV factor disks (BD BBL™ Taxo™, MD, USA) The disks and culture (0.5 MacFarland in saltwater) were placed on plain agar (Oxoid) and grown overnight in at 37°C in 5% CO_2. To establish if the isolate produced hemolysis, the isolate was grown overnight at 37°C in 5% CO_2 on blood agar (Oxoid) next to

Staphylococcus aureus. In two isolates the suspected *Haemophilus* colonies were hard to isolate from other bacteria of the nasopharynx. The growth was replated on chocolate agar (Oxoid) and Bacitracin (BD BBL™ Taxo™) disks were placed next to the growth.

3.2.4 PCR assay for serotyping *Hi*

The author had previously designed a multiplex PCR method for serotyping *Hi* (Table 11) with reference strains (Table 7) (106). This method was used on *Hi* isolates obtained from clinical samples. Isolates from the nasopharynx and nasal swabs (n=58) were excluded from serotyping because of uncertainty if they were disease-causing isolates.

When serotyping the carriage isolates (n=286) primers specific for the *bexA* gene (Table 11) were used in duplex PCR with either *fucK* or *hpd* primers as internal control (Table 10). The *bexA* gene is present in all capsulated strains (90). When a strain showed the presence of the *bexA* gene the isolate was further analyzed with serotype specific primers (Table 11).

Table 11: Primers used in the PCR assay for serotyping *Hi.*

Reaction	Serotype	Primers	Sequence (5'-3')	Amplicon (bp)	Citation
I	a	AcsA-f/r	GCA ACC ATC TTA CAA CTT AG	77	(81)
			CGG TGT CCT GTG TTT AG		
	b	b1/b2	GCG AAA GTG AAC TCT TAT CTC	481	(90)
			CTT ACG CTT CTA TCT CGG T		
	d	d1/d2	TGA TGA CCG ATA CAA CCT G	166	(90)
			TCC ACT CTT CAA ACC ATT C		
	bexA inner control	Hi-1/Hi-2	CGT TTG TAT GAT GTT GAT CC	340	(90)
			CCA TGT CTT CAA AAT GAT G		
II	c	c1/c4	TCT GTG TAG ATG ATG GTT CA	146	(90)
			TCA ATG AAA GTA ACC CAT TC		
	e	EcsC2-F/EcsC-R	CAC ACT ACC TTT TGA GAA GAG	209	*
			TAG TTT GAA AGA ACC CTC TG		(81)
	f	FcsA-f/r	CCT GAA ATT TGC TAT TAC TTT A	94	(81)
			GTG GTC TAT TTC CAT TCT CTT		
	bexA inner control	Hi-1/Hi-2	CGT TTG TAT GAT GTT GAT CC	340	(90)
			CCA TGT CTT CAA AAT GAT G		

*New primer designed for this study.

3.2.5 PCR assay

Quick Load® Taq 2X Master Mix (New England Biolabs Inc., Ipswich, MA, USA) was used for all PCR assays. Primers (TAG Copenhagen A/S, Copenhagen, Denmark) were diluted with Sigma water (Sigma-Aldrich Co, St. Louis, MI, USA) to the final concentration of 0.5 µmol. Each 25 µl PCR reaction contained 12.5 µl Master Mix, 5 µl of DNA extraction, 1 µl of each primer and 6.5 µl Sigma water (Sigma-Aldrich Co). When primer pairs were more than one, water was reduced so the total amount in

each reaction was always 25 µl. PCR reaction was conducted in a 2720 Thermal cycler (Applied Biosystems, Foster City, CA, USA). Denaturation temperature was 94°C for 4 minutes, then 30 cycles of denaturation at 94°C for 30 seconds, annealing at 52°C for 30 seconds and elongation at 72°C for 30 seconds, after 30 cycles the final elongation temperature was 72°C for 5 minutes, then cooled to 4°C storage temperature.

3.2.6 Electrophoresis

PCR products were visualized by electrophoresis in 1.5% agarose gel (Sigma-Aldrich Co) in a sub cell horizontal electrophoresis system (Bio-rad Laboratories, Inc., CA, USA) containing 0.5 x TBE buffer (Tris-borat-EDTA). Each well was loaded with 6 µl of PCR product and 4 µl of 100 bp size ladder (Quick-load 100 bp DNA Ladder, New England Biolabs Inc) was used as size marker.

Electrophoresis was at 120 V for 30-70 minutes, depending on the gel size. The gel was stained with Ethidium bromide for 15 minutes destained for 30-40 minutes in distilled water and photographed (Syngene, Cambridge, UK) under UV light and the photograph was then saved as a TIF file.

3.3 Pulsed-field gel electrophoresis

PFGE was done on 303 selected isolates from clinical and carriage samples, 38% of all available *Hi* isolates from 2012. Almost every other *Hi* isolated from ears was analyzed with PFGE. Every other nasopharyngeal *Hi* isolate available from children attending DCCs in 2012 were analyzed. *Hi* isolates from other specimen types were systematically picked and all isolates from abscess, blood and tissue were analyzed (Table 8).

The method for *Hi* PFGE was based on previously designed methods for *Streptococcus pneumoniae* (109) and *Legionella* (110).

Briefly, *Hi* isolates were grown on chocolate agar overnight (Oxoid, Hampshire, UK). A single colony was picked, re-plated on chocolate agar and incubated overnight. Bacterial growth was suspended in PIV buffer and cell suspension adjusted to absorbance of 0.400-0.600 at 650 nm in a spectrometer. After centrifugation the bacterial pellet was re-suspended in 200 µl TEN buffer, 150 µl of the re-suspended bacterial pellet was then gently mixed with 150 µl of 1.5% Seaplaque® GTG® agarose (Lonza, Rockland Inc., Rockland, ME, USA). Agarose disks (20 µl) were cast and incubated in 1 ml EC solution with lysozyme (10 mg/ml) (Sigma-Aldric Co) and RN-ase (10 mg/ml) (New England Biolabs Inc) at 37°C overnight. Then the agarose disks were incubated in 1 ml ES solution with proteinase K (1 mg/ml) (Sigma-Aldric Co) at 37°C overnight and then washed four times with TE buffer. The DNA in the disks was then digested with *Sma*I restriction enzyme (10 U/µl) (Thermo Fisher Scientific Inc., MA, USA) overnight at 30°C. The restriction enzyme was stopped with Gel Loading Dye Blue 6X (New England BioLabs Inc).

Electrophoresis was performed in a 1.5% Seakem LE agarose (Lonza) for 23 hours at 6 V/cm with pulsed time of 5 to 35 seconds. Lambda ladder PFGE marker (New England BioLabs, Inc) was placed at each end and in the middle well of the gel for size references.

The gel was stained with Ethidium bromide for 15 minutes and destained for 3-5 days in distilled water. The gel was photographed (Syngene) under UV light and the image saved as a TIF file.

Band patterns were analyzed with BioNumerics (Applied Maths NC, Sint-Martens-Latem, Belgium) and the similarity matrix calculated using the Dice coefficient with 1.5% band position tolerance and 0.5% pattern optimization. Dendrogram was drawn using UPGMA clustering.

The PFGE pattern of each isolate consisted of a mean of 11 bands. Because three band difference can be caused by a single genetic event according to Tenover's criteria (97) it was decided to group isolates with 75% (8/11=72%, rounded up to 75%) similarity according to BioNumerics and ≤3 band difference together in a PFGE clone. The PFGE subclones were given numbers according to band difference and visual analysis so each PFGE subclone number represents a unique band pattern.

3.4 MLST

Isolates analyzed with MLST were chosen from the PFGE results. Isolates from large PFGE clones containing isolates from different specimen types were chosen rather than isolates from smaller PFGE clones containing one specimen type. Both isolates from blood were analyzed with MLST. The MLST primers for the PCR and sequencing of the seven housekeeping genes were received from http://haemophilus.mlst.net (Table 12). PCR assay for amplification of the seven housekeeping genes for MLST was conducted with denaturation temperature of 95°C for 4 minutes, then 30 cycles of denaturation at 95°C for 30 seconds, annealing at 55°C for 30 seconds and elongation at 72°C for 60 seconds, after 30 cycles the final elongation temperature was 72°C for 10 minutes, then cooled to 4°C storage temperature.

Gene sequencing was performed at Matís ohf in an ABI3730XL (Applied Biosystems), the data was then analyzed in BioNumerics.

Table 12. Primers for MLST.

Primers	Sequence 5'-3'
adk up/dn	GGT GCA CCG GGT GCA GGT AA
	CCT AAG ATT TTA TCT AAC TC
atpG up/dn	ATG GCA GGT GCA AAA GAG AT
	TTG TAC AAC AGG CTT TTG CG
frdB up/dn	CTT ATC GTT GGT CTT GCC GT
	TTG GCA CTT TCC ACT TTT CC
fucK up/dn	ACC ACT TTC GGC GTG GAT GG
	AAG ATT TCC CAG GTG CCA GA
mdh up/dn	TCA TTG TAT GAT ATT GCC CC
	ACT TCT GTA CCT GCA TTT TG
pgi up/dn	GGT GAA AAA ATC AAT CGT AC
	ATT GAA AGA CCA ATA GCT GA
recA up/dn	ATG GCA ACT CAA GAA GAA AA
	TTA CCA AAC ATC ACG CCT AT

3.5 Data analysis

Data was analyzed using Excel (Microsoft, Redmond, WA, USA). For statistical analysis, Fisher exact was used and the p value of <0.05 was considered significant. Similarity between isolates was calculated in BioNumerics.

3.6 Ethics

This study is a part of the VIce study – Effects of pneumococcal vaccination in Iceland. License number from the Icelandic National Bioethics Committee: 13-010-S1.

4 Results

4.1 Differentiation between *Hi* and *H. haemolyticus*

All 511 available *Hi* isolates from clinical samples were tested with this method as a part of the author's diploma project. All the isolates were positive for *fucK*, *hpd* or both genes (Table 13).

Haemophilus spp. were suspected in 296 isolates from carriage samples. All isolates were identified with the method for differentiating between *Hi* and *H. haemolyticus*. The genes *fucK* and *hpd* were present in 286 isolates (Table 13). No amplification of either *fucK* or *hpd* was found in 10 of the carriage isolates. The isolates were identified with phenotypic methods, one was *H. parahaemolyticus*, one was *H. parainfluenzae* and the rest of the isolates were not *Haemophilus* spp.

Table 13. Presence of *fucK* and *hpd* genes in the *Hi* isolates from clinical samples and *Hi* isolates from carriage samples**.**

Gene	***Hi* isolates from clinical samples**		***Hi* isolated from carriage samples**	
	n	**%**	**n**	**%**
fucK	32	6.3	17	5.9
hpd	12	2.3	3	1.1
fucK and *hpd*	467	91.4	266	93
total	511	100	286	100

4.2 Serotyping of *Hi*

Part of the author's diploma project was developing a serotyping method using seven reference strains (Table 7). The method (Table 11) was then used to serotype isolates from clinical samples. Nasal and nasopharynx isolates among the clinical samples were excluded, leaving 453 isolates serotyped. All isolates were NT*Hi,* except one serotype e isolate (0.2%).

PCR serotyping method was done on all 286 isolates from carriage samples positive for either *fucK* or *hpd* genes.

All but five isolates from carriage samples were NT*Hi,* two isolates were serotype e and three isolates were serotype f. All three serotype f isolates were found in the same DCC but the two serotype e isolates were found in two different DCCs (Table 14).

There was a significant difference (p=0.03) between the number of capsulated isolates from carriage samples and isolates from clinical samples.

Table 14. Number of capsulated *Hi* and NT*Hi* isolates found among carriage samples and what DCCs they were found in.

Serotype	n	DCCs	%
e	2	1 from DCC B	0.7
		1 from DCC J	
f	3	DCC A	1.1
NT*Hi*	281	All DCCs	98.2
Total	286	All DCCs	100

4.3 PFGE

Isolates with 75% similarity, according to BioNumerics and ≤3 band difference, were classified together in one PFGE clone. A PFGE subclone number was assigned to identical isolates inside each PFGE clones so that each unique band pattern got an individual number.

Of the 303 isolates analyzed with PFGE, 254 formed 72 PFGE clones and 49 were singletons or 16% of all the isolates. The full dendrogram can be found in Appendix I and II, and PFGE clone details are in Appendix III. Example of PFGE results are shown in Figure 4.

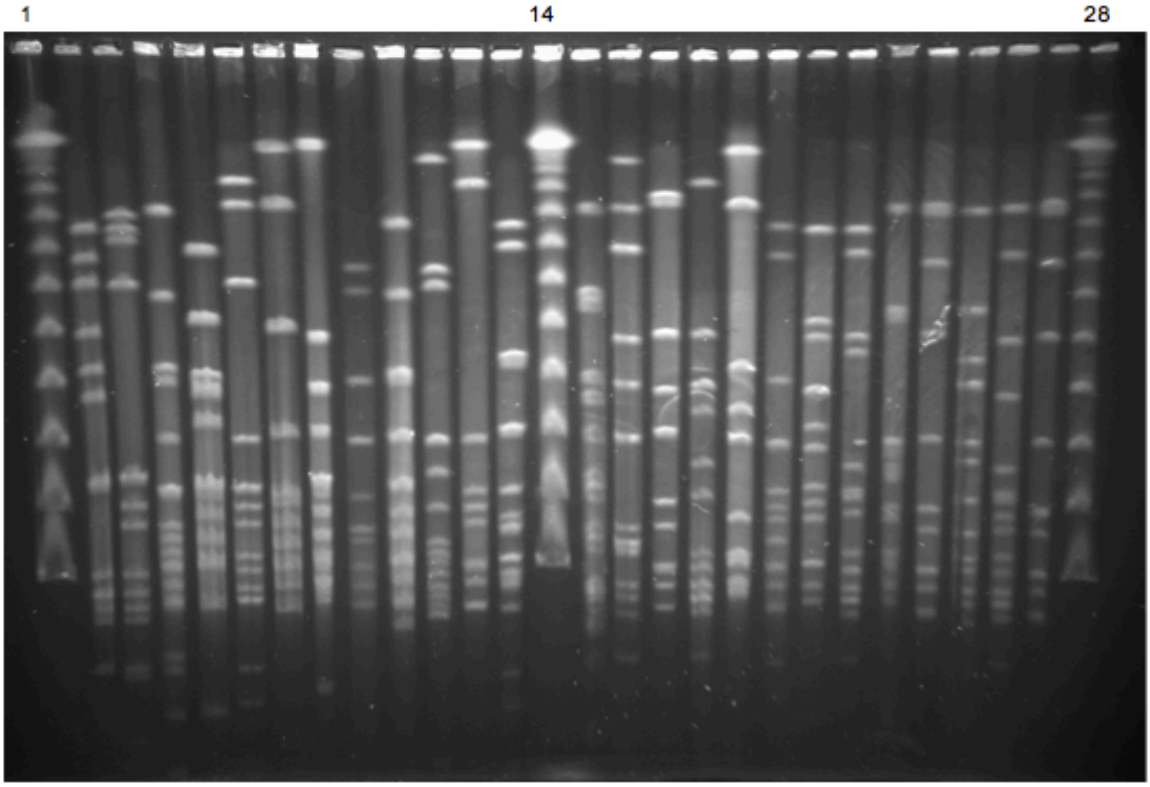


Figure 4. An example of PFGE results. Wells number 1, 14 and 28 contain size references and all other wells contain *HI* isolates from clinical samples.

4.3.1 Combination of specimen types in PFGE clones

Isolates from two different specimen types could be found in 35 PFGE clones and 10 PFGE clones included isolates from three or more specimen types (Appendix III).

The PFGE clone with the highest number of isolates was PFGE clone 16; it contained 10 isolates from carriage samples and four isolates from ear (Table 15). PFGE clone 16 was also the PFGE clone that had the highest number isolates from the same specimen type, 10 isolates from carriage samples. The PFGE clones with the highest number of ear isolates were PFGE clones 2, 5, 11, and 16, all with four isolates each (Appendix III).

Only two *Hi* blood isolates were available from 2012; one of them was similar to a DCC carriage isolate (78% similarity) forming PFGE clone 52 (Figure 5). The other blood isolate was a singleton.

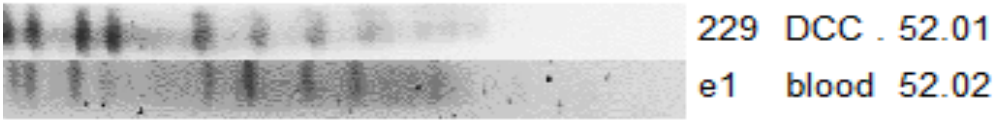


Figure 5. PFGE banding pattern of two similar isolates, isolate 229 from DCC carriage and isolate e1 from blood.

Of the 303 isolates analyzed with PFGE 144 (48%) were isolates from carriages (nasopharyngeal *Hi* isolates from children attending DCCs), 110 (36%) isolates from ear infections and 39 (13%) isolates from other infections. 10 nasopharyngeal and nasal swab isolates sent to the Department of Clinical Microbiology were also analyzed with PFGE 10 (3%) (Table 8).

In Table 15, PFGE clones with seven or more isolates are listed for comparison of the composition of the specimen type in each PFGE clones and the composition of each specimen type of all PFGE analyzed isolates.

Table 15. PFGE clones with seven or more isolates, number of subclones, number of isolates of each specimen type and the percentage of total in each PFGE clone.

PFGE clone (n)	Number of subclones	Specimen type and number of isolates		% in PFGE clone
1 (7)	2	Carriage	7	100
5 (11)	1	Carriage	6	55
		Ear	4	36
		Other	1	9
11 (11)	7	Carriage	7	64
		Ear	4	36
14 (7)	4	Carriage	3	42
		Ear	2	29
		Other	2	29
16 (14)	4	Carriage	10	71
		Ear	4	29
17 (7)	1	Carriage	5	71
		Ear	2	29
19 (10)	4	Carriage	5	50
		Ear	3	30
		Nasopharyngeal	1	10
		Other	1	10
23 (7)	6	Carriage	1	14
		Ear	3	43
		Nasopharyngeal	1	14
		Other	2	28
72 (7)	6	Carriage	3	43
		Ear	3	43
		Other	1	14

Statistic analysis was done on all clones containing seven or more isolates. PFGE clone 1 was the only clone where all the isolates were from the same specimen type, it was only found in carriage. This differed significantly (P=0.005) from the composition of all analyzed isolates (48% of the isolates analyzed with PFGE were from carriage).

4.3.2 Carriage isolates from children attending DCCs

A total of 52 PFGE clones were found among the carriage isolates (Appendix III). Of the 52 PFGE clones, 20 could be found at different DCCs. A total of 14 different PFGE clones could be found in only one DCC. The PFGE clones containing two or more isolates that were only found in one DCC, were: PFGE clone 1, 4, 17, 24, 29, 31, 37, 54, 60, 72, 73, 74, 76, and 78. PFGE clone 1 contained seven isolates from DCC D, PFGE clone 4 contained six isolates from DCC F, PFGE clone 17 contained five isolates from DCC J and the rest of the PFGE clones that were found in only one DCC contained either two or three isolates (Table 16). Out of all the DCCs carriage isolates that were analyzed with PFGE there were 18 singletons (Appendix III).

Table 16. PFGE clone and PFGE subclone number of isolates found in two or more children attending DCCs

PFGE clone	DCC
1.01	D
1.01	D
1.01	D
1.01	D
1.01	D
1.02	D
1.02	D
4.01	F
4.01	F
4.01	F
4.01	F
4.01	F
4.01	F
5.01	E
5.01	F
5.01	F
5.01	G
5.01	G
5.01	H
11.05	A
11.05	A
11.05	A
11.01	E
11.07	E
11.03	H
11.02	J
13.04	G
13.01	K
14.04	E
14.03	G
14.02	J
16.01	B
16.01	B
16.01	F
16.01	F
16.02	G

PFGE clone	DCC
16.01	H
16.03	I
16.01	L
16.01	L
17.01	J
17.01	J
17.01	J
17.01	J
17.01	J
19.01	E
19.01	E
19.01	G
19.04	M
19.04	N
21.02	I
21.01	O
21.01	O
22.02	E
22.01	H
24.01	C
24.01	C
28.01	A
28.03	B
29.01	H
29.02	H
30.02	H
30.01	M
30.01	N
31.01	M
31.02	M
33.01	C
33.01	O
37.01	D
37.02	D
40.01	L
40.02	O

PFGE clone	DCC
43.01	N
43.02	O
47.01	E
47.02	L
49.02	F
49.04	L
50.01	G
50.02	L
53.01	G
53.03	O
54.01	A
54.01	A
54.01	A
55.01	L
55.01	G
60.01	C
60.01	C
60.01	C
63.01	E
63.01	N
63.02	M
64.02	E
64.01	O
64.01	O
72.01	I
72.03	I
72.03	I
73.01	O
73.02	O
73.02	O
74.01	N
74.01	N
76.01	I
76.01	I
78.01	N
78.01	N

4.3.3 Identical PFGE subclones

Personal identifications were available from all the isolates sent to the Department of Clinical Microbiology in 2012. The isolates forming PFGE subclone 3.01 were isolated from the ear of the same person, with 49 days apart (Figure 6). The isolates forming PFGE subclone 26.01 were also

found in the ear of the same person (Figure 7), with 9 days apart. The three isolates forming PFGE subclone 77.01 were found in the same person, isolated from tracheal aspirate with 52 days apart from the first and last isolation (Figure 8) (Table 17).

Table 17. Identical PFGE subclones found among PFGE analyzed clinical isolates sent to the Department of Clinical Microbiology the year 2012, specimen type and day between isolation of the isolates.

PFGE subclone	Specimen type	Number of isolates	Time between (days)
3.01	Ear	2	49
26.01	Ear	2	9
77.01	Tracheal aspirate	3	1st and 2nd: 4 2nd and 3rd: 48

Figure 6. PFGE of two ear isolates from the same person taken 49 days apart.

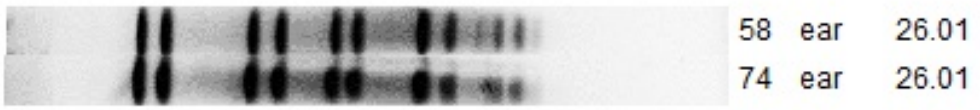


Figure 7. PFGE of two ear isolates from the same person taken 9 days apart.

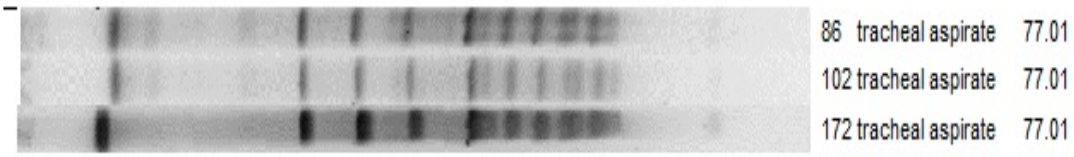


Figure 8. PFGE of three tracheal aspirate isolates from the same person taken 4 and 48 days apart.

4.3.4 Capsulated *Hi*

A total of four capsulated *Hi* were analyzed with PFGE, two were serotype e and two were serotype f according PCR. Serotype f isolate from carriage sample was identical (same subclone) to three NT*Hi* isolates (Figure 9). Another serotype f isolate from carriage sample was classified in a clone with two serotype e isolates (Figure 10).

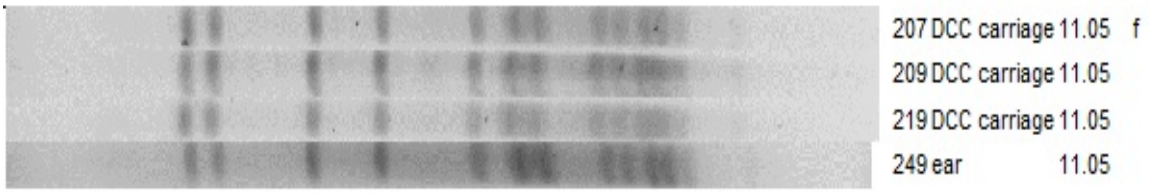


Figure 9. Serotype f isolate from carriage isolate identical in PFGE to three NT*Hi* isolates.

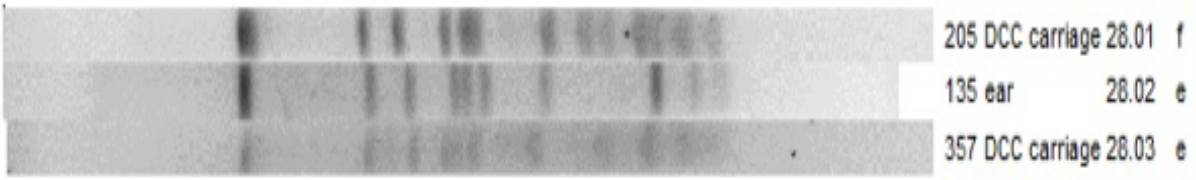


Figure 10. Serotype f isolate from carriage isolate in the same PFGE clone as two serotype e isolates.

4.4 MLST

MLST was done on 21 isolates. Sequencing of all seven housekeeping genes gave results for 12 isolates, in nine isolates sequencing errors in one or two genes circumvented results. Results can be seen in Table 18.

Table 18. Isolates typed using MLST, specimen type, PFGE clone and sequence type.

Isolate code	Specimen type	PFGE clone	Sequence type
13	Carriage	5.01	ST474
37	Carriage	4.01	
54	Carriage	1.01	ST840
160	Carriage	72.03	
190	Carriage	54.01	ST2
219	Carriage	11.05	
239	Carriage	16.01	ST1030
284	Carriage	63.02	New ST
301	Carriage	18.01	
337	Carriage	17.01	ST946
390	Carriage	19.01	
42	ear	6.01	
102	tracheal aspirate	77.01	ST925
167	ear	62.03	ST12
215	ear	16.01	ST1030
249	ear	11.05	ST544
292	ear	14.02	
297	ear	10.01	ST266
394	ear	2.01	ST925
e1	blood	52.01	
e2	blood	singleton	

5 Discussion

The results from the differentiation of *Hi* and *H. haemolyticus* showed that no *H. haemolyticus* had been identified as *Hi* among the isolates from clinical samples and no *H. haemolyticus* was among the isolates from carriage samples.

Only six capsulated *Hi* were found among the isolates, one was isolated from an ear sample and the others from carriage samples. The rest of the isolates were all NT*Hi*.

The isolates analyzed with PFGE showed high genetic diversity. A total of 72 PFGE clones were found among the 303 isolates and 49 isolates were singletons.

5.1 Differentiation between *Hi* and *H. haemolyticus*

Hemolysis is the most used test for differentiating *Hi* and *H. haemolyticus* when identifying with phenotypic methods. Because the biomedical scientists of the Department of Clinical Microbiology had already excluded all hemolytic isolates we conclude that no non-hemolytic *H. haemolyticus* was identified as *Hi* among the clinical samples.

The identification of isolates from nasopharyngeal carriage samples from children attending DCCs in Iceland's capital region in 2012 gave the similar results, no *H. haemolyticus* was found among those isolates. The percentage of isolates containing both *fucK* and *hpd* genes that are generally present in *Hi* were over 90% of both among the isolates from clinical samples and carriage sample (Table 13).

Correct identification between the two species can be important for patients. For example *Hi* is one of the most common bacteria that cause exacerbations of symptoms in COPD patients. When *H. haemolyticus* is in respiratory tract isolates of COPD patients and is correctly identified as such, it spares the patient from unnecessary antibiotic use (111). The differentiation between *Hi* and *H. haemolyticus* in invasive isolates is important for epidemiologic knowledge. *H. haemolyticus* had rarely been reported to cause invasive disease until Anderson *et al.* repeated analysis on 374 isolates previously identified as NT*Hi*. A total of seven isolates were *H. haemolyticus* according to 16S rRNA gene sequencing (22).

Non-hemolytic *H. haemolyticus* have been found in previous studies (19). Murphy *et al.* did a study on various strains isolated from invasive diseases, sputum of COPD patients, children's nasopharynx and middle ear fluid, all previously identified as *Hi*. The isolates were analyzed with 16S rRNA gene sequencing, MLST, DNA-DNA hybridization and sequencing of the gene coding for *P6* outer membrane protein. The results showed 27% of the nasopharyngeal isolates and 40% of the sputum isolates were in fact *H. haemolyticus* according to the genotypic methods. The authors suggest that it is possible that the rate of *H. haemolyticus* is higher among respiratory tract isolates because it is a respiratory tract commensal and not an as common pathogen like *Hi*. This suggestion is also supported by the fact that none of the 130 middle ear fluid isolates contained *H. haemolyticus* and only four (all cryptic genospecies) out of 54 invasive isolates were *H. haemolyticus*. The *H. haemolyticus* isolates were checked for hemolysis on blood agar, all of the nasopharyngeal isolates and almost half of the sputum isolates proved to be hemolytic (111).

Kirkham *et al.* did a similar study on 266 nasopharyngeal *Hi* isolates from Australian children either with or without reoccurring AOM. They analyzed the isolates with PCR for *hpd* and *P*6 genes along with 16S rRNA gene sequencing. Among the isolates from children without reoccurring AOM 12% proved to be *H. haemolyticus* and among the isolates from children with reoccurring AOM, 9% were *H. haemolyticus*. The *H. haemolyticus* found in the study were not checked for hemolysis. They did not identify any *H. haemolyticus* among isolates from ear effusion samples from children with reoccurring AOM, which is in accordance with the results of Murphy *et al*. (111, 112).

Norskov-Lauritsen identified 480 clinical isolates from a Danish hospital with hybridization for *fucK, hap* and *sodC* with gene specific-probes and 16S rRNA gene sequencing. A total of two isolates were determined as non-hemolytic *H. haemolyticus* (113).

The best way to differentiate between *Hi* and *H. haemolyticus* is by doing whole genome sequencing (20). However, whole genome sequencing is expensive and might not be available for laboratories on a low budget. Theodore *et al.* did a study comparing 16S rRNA gene sequencing to the *fucK* and *hpd* PCR method. The *hpd* and *fucK* genes were not found in 46 isolates. The results of 16S rRNA gene sequencing showed that 40 out of the 46 were *H. haemolyticus*. All these 40 isolates were non-hemolytic. The six remaining isolates were all classified with *Hi*. All 78 isolates in the study that contained *fucK* and/or *hpd* were classified as *Hi* in the 16S rRNA gene sequencing. The sensitivity and specificity of the PCR method compared to 16S rRNA gene sequencing was the highest when both *fucK* and *hpd* were used, with 92% sensitivity and 98% specificity. It is important to keep in mind that because of the high frequency of horizontal gene transfer of these two species, no single PCR method can be 100% sensitive. Based on the latest knowledge available, using *fucK* and *hpd* is a reliable method for differentiating between the two species (21). The fact that sequencing of the 16S rRNA gene sequencing is most commonly used as a standard for comparing the differentiation ability of other methods like PCR, does raise the question of if the nucleotide information of only one gene, 16S rRNA, is specific enough (88).

The rate of both *H. haemolyticus* and non-haemolytic *H. haemolyticus* among previously identified *Hi* isolates differs from specimen types and possibly varies between different countries. It is possible that not all studies done on the misidentification of *H. haemolyticus* as *Hi* are compatible because of their different identification methods. Differences in the blood used in agar production might have something to do with the beta-hemolysis manifestations of *H. haemolyticus*. For example, in Iceland we use horse blood from Icelandic horses but sheep blood is mostly used in other countries (79).

5.2 Serotyping of *Hi*

Only six of 739 isolates were found to have a capsule. Among the serotyped isolates from clinical samples (n=453), one isolate was serotype e (0.2%), and 452 were NT*Hi*. Among isolates from carriage samples (n=286), two isolates were serotype e (0.7%) and three were serotype f (1.1%), the remaining 281 were NT*Hi* (Table 14).

Among invasive *Hi* isolates diagnosed at the Department of Clinical Microbiology in the years 1983-2008, only one isolate of serotype e was identified among the 79 non-*Hi*b isolates (Table 4) (60). It is

possible that serotype e has been more common among *Hi* isolates other than invasive, because generally isolates are not serotyped unless they were the cause of invasive infections.

The isolate of serotype e that was found among the clinical samples was from the middle ear of a 16-month old child. In this study a total of 257 isolates from ears were serotyped, therefore the percentage of serotype e from ear isolates was 0.4%. These results are in accordance to other studies reporting that when *Hi* is found in ear samples it is almost always NT*Hi* (7, 55).

NT*Hi* was the cause of two invasive infections in 2012, both isolates were found in blood. Ever since the *Hi*b vaccine became widely used, NT*Hi* has been the most common of *Hi* to cause infections, both in Iceland and in other developed countries around the world (6, 7, 60). Laupland *et al*. showed the annual incidence rate of NT*Hi* bacteremia in Australia, Canada and Denmark from 2000-2008 was 0.98/100.000. Serotypes a, c, e, d and f had the annual incidence rate of 0.22/100.000 and *Hi*b only had 0.08/100.000 (46). This low rate seen in bacteremia caused by non-*Hi*b underlines that serotype replacement among *Hi* has not occurred after the *Hi*b vaccinations (114).

It is common for no *Hi*b to be found among populations where the majority of the people are vaccinated, similar to this study. The vaccinated population creates a herd effect so that unvaccinated people are at a lower risk for getting *Hi*b. Like Watts *et al.* showed in a global estimate on the burden of *Hi*b disease, not everyone has the privilege of either getting vaccinated or being touched by the vaccine herd effect. They estimated that in 2000, *Hi*b caused over eight million of serious infections among children under the age of five around the world. Developing countries in Africa and Asia did largely contribute to this high number. Deaths in children under the age of five in 2000 were estimated to be 371.000 and 61% of these deaths occurred in 10 African and Asian countries. The vast majority of these serious infections and deaths could have been prevented by vaccine use (4).

There was a significant difference between the numbers of capsulated isolates between carriage samples and clinical samples. No compatible studies comparing rates of capsulated *Hi* between carriage and disease isolates were found in the literature. However, infections due to serotype e and f are very rare and they do effect the same age group as NT*Hi* (6).

In 2012, *Haemophilus* spp. were found in 62% of children attending DCCs in Iceland (50). Nasopharyngeal carriage of *Hi*b in children was approximately 3-5% in infants and approximately 8-12% in children attending DCCs before *Hi*b vaccinations (59). In *Hi*b vaccinated populations serotyping of *Hi* found among DCCs carriage is a good way to follow the vaccine effect. Carvalho *et al*. found that among carriage isolates from 1192 Brazilian children attending 62 DCCs (taken from August to December 2005), serotype f was the most frequent (4.6%) after NT*Hi* (23.3%) among DCCs carriage. Serotype a accounted for 2% of the *Hi* isolates and *Hi*b were found in 0.7% of the *Hi* isolates, demonstrating the decrease in carriage of *Hi*b after the vaccination (51). Two unrelated French studies were conducted at a similar time in a different geographic area. The first study was on the carriage rate among 20 DCCs (February to April 1997) and reported 26% NT*Hi* carriage among the DCCs children and two capsulated isolates, serotype e and f each 0.6% (115). The other study, conducted two years later in three different regions of France (65 DCCs), also reported the only capsulated strains to be serotype e (0.6%) and f (0.4%). The carriage rate of *Hi* was analyzed during two different periods; the rate for May to June, 1999 was 41.8% and November to December of the same year was

39.8% (52). An even higher NT*Hi* carriage rate of 64% was reported among 198 children from 16 different DCCs from February to April in Michigan, USA in 2001.

The carriage rate of NT*Hi* in Iceland (around 60%) can be considered high compared to the similarly conducted studies listed above. The reasons are most likely due to a combination of many factors; for example, differences in hygiene and antibiotic use are likely to cause differences in carriage rates. The carriage samples from DCCs in Reykjavík capital region were taken in March. Seasonal changes are also known to affect the carriage rate of *Hi* (51, 116). Differences in the phenotypic identification of the isolates can possibly affect the rate of *Hi* that are identified. In this study we used chocolate agar and bacitracin sensitivity disks to make *Hi* more visible from the other bacteria in the nasopharynx.

5.3 Genotyping

PFGE was done on 303 isolates. Of these 110 (36%) were isolates causing ear infections, 39 (13%) isolates considered causing other infections, 10 (3%) nasopharyngeal and nasal swab isolates and every other isolate from carriage samples was included, a total of 144 isolates (48%) (Table 8).

The PFGE pattern of each isolate consisted of a mean number of 11 bands. As three band difference can be caused by a single genetic event according to Tenover's criteria (97), all isolates representing 75% (8/11) similarity were grouped together in the same PFGE clone.

Of the 303 isolates analyzed with PFGE, 254 were in 72 PFGE clones and the remaining 49 isolates were not similar to another isolate (Appendix III).

We noticed one noteworthy PFGE clone when viewing the composition of isolates in each of the larger clones, compared to the composition of all isolates analyzed with PFGE (Table 15). PFGE clone 1 was the only clone with seven or more isolates where all the isolates were from the same specimen type, it was only found in carriage samples. This differed significantly (P=0.005) from the composition of all analyzed isolates (48% of the isolates analyzed with PFGE were from carriage samples). This might indicate that isolates from PFGE clone 1 exhibit less virulence than isolates from the other PFGE clones. The composition of other PFGE clones containing seven or more isolates did not differ significantly from the composition of all analyzed isolates.

Isolates from other specimen types were grouped in PFGE clones with ear and isolates from carriage samples and there was no obvious trend of isolates from subjects of a certain age grouped together in clones. Only two invasive *Hi* isolates (from blood) were found in 2012 one of them was similar to an isolate from a carriage sample (78% similarity), together they formed PFGE clone 52 (Figure 5) and the other one was a singleton. PFGE clone 77 only contained three tracheal aspirate isolates from the same person where 52 days passed from the first to the last isolation of the specimens. Two other persons had identical subclones at a different time of isolation, both from ears (Table 17).

All isolates from carriage samples analyzed with PFGE were classified into 52 different PFGE clones and 18 singletons (Table 16 and Appendix III). Of the 52 PFGE clones, 20 could be found at different DCCs and 18 different PFGE clones contained only one carriage isolate. A total of 14

different PFGE clones could be found in only one DCC. PFGE clones including five or more isolates at the same DCC were PFGE clone 1 that contained seven isolates from DCC D, PFGE clone 4 that contained six isolates from DCC F and PFGE clone 17 that contained five isolates from DCC J (Table 16). The rest of the local PFGE clones only included two or three isolates from a specific DCC. These results show that the same PFGE clone can be found in many children at the same DCC because of the close surroundings and the same clone can be spread among DCCs, as has previously been described (47, 52). The high diversity among *Hi* isolates from children in close surroundings can also been caused by new strains (PFGE clones) brought in by the children and the staff from outside the DCCs. High turnover rate of *Hi* has been documented, for example Kaur *et al.* studied carriage in 33 children with a three month interval, 75% of the children had acquired a new NT*Hi* strain after three months (47).

The four capsulated isolates that were analyzed with PFGE were of particular interest. A serotype f isolate was identical to two other NT*Hi* isolates from carriage samples and one NT*Hi* isolate from ear (Figure 9). There is very limited literature about similar cases; However, Ohkusu *et al.* published an interesting study on two different *Hi* isolates that were found in a child with meningitis. The isolate found in cerebrospinal fluid was serotype a, and the isolates found in the patient's blood was NT*Hi*. When the restriction enzyme *Nae*I was used in PFGE, the two isolates were identical. The difference between the band patterns of the two isolates when analyzed with *Sma*I, *Apa*I and *Sac*II in PFGE were only the size of one fragment. They confirmed with southern blot to screen that the *cap* locus was not intact in the NT*Hi* isolate and that was the reason for the differences in PFGE band patterns. Results from PCR analysis confirmed that the NT*Hi* isolate did possess the insertion sequence IS*1016* (commonly found in capsulated strains because they flank the *cap* locus). The serotype a isolate did, however, possess the IS*1016*, *bexA* and cap part II. The IS*1016* from the serotype a and NT*Hi* isolate were sequenced and they were identical. Therefore it was confirmed that the NT*Hi* isolate was a capsule-deprived variant from the serotype a isolate (117). The *Hi*b- variants have been studied more extensively, those variants do contain the *cap* locus but with an IS*1016*-*bexA* partial deletion and cannot produce capsule. The case described by Ohkusu *et al.* is different because the NT*Hi* strain only contained IS*1016* and can therefore not be called a-minus variant (*Hi*a-). Invasive NT*Hi* isolates from Portugal did show more similarity (not identical) to *Hi*b than other NT*Hi* strains according to PFGE with *Sma*I. These NT*Hi* isolates did not contain part of the *cap* locus according to PCR (118). The most famous capsule deficient variant of *Hi* is the *Rd* strain, it was the first free living organism to have its complete genome fully sequenced (119). The *Rd* strain lost its capsule through a recombination event, leaving only one copy of IS*1016* in section 97 of the genome. Satola *et al*. searched for IS*1016* among 201 NT*Hi* isolates, 19 (9.5%) showed evidence of the IS*1016.* but they could not confirm that IS*1016* was in section 97 and therefore could not prove if the NT*Hi* isolates had lost their capsule like *Rd* (36).

In PFGE clone 28 there were three capsulated isolates, one was serotype f from a carriage sample and the other two were serotype e, one from ear and one from carriage sample (Figure 10). This did raise questions because the homogeneity of capsulated *Hi* is well documented (120-122). For example PFGE analysis on capsulated strains almost always group together in a dendrogram (118,

123) An exception was found in a study by Omikunle *et al.* in which one invasive serotype f isolate was less similar than the other serotype f (n=8) when digested with *Sma*I, but when digested with *Apa*I restriction enzyme the isolate appeared to be as similar to serotype a isolates as to the other serotype f isolates (123). Clustering according to serotype has also been reported with MLST analysis (121). Interestingly, when all available isolates in the MLST database at the time were studied (359 different ST), the dendrogram showed almost all the NT*Hi* strains were classified into eight different clades. In addition there was one clade made up of serotype a, e, f and NT*Hi* strains and another clade that consisted of serotype b and NT*Hi* strains. This does suggest that these serotypes and NT*Hi* do have a common ancestor (24, 121).

A total of 10 different ST's were obtained from 12 isolates. We identified ST2 from carriage isolate and ST12 from ear isolate, similar to the results of Kaur *et al.*, who found ST2 in the nasopharynx and ear of a child with AOM and ST12 was found in another child, both in the nasopharynx and ear isolate. ST1030 was found in ear and carriage isolate from the same PFGE subclone (16.01), according to the online database, ST1030 was first identified from cerebrospinal fluid in 2010 (http://haemophilus.mlst.net). ST925 was found in two isolates that belonged to a different PFGE clone, one isolate was from tracheal aspirate and the other from ear. The clones were viewed together in BioNumerics and it was apparent that the isolates in these two PFGE clones had a similar band pattern (72% similarity). It is possible that the lambda ladder used as a size reference did not work as expected in this case and therefore the variations between gels created this inconsistency. We also need to keep in mind that MLST results are not always consistent to PFGE results and vice versa (101).

A new ST was identified, in an isolate from a carriage sample, it had a single base pair change from *pgi* allele number 40 and is therefore a single locus variant from ST1012. Another isolate was found that had a single base pair change from *mdh* allele 87 but the sequence quality of three other genes was not good enough to assign an allele number. The new ST will be added to the online database and assigned an ST number.

The genetic diversity of *Hi* in Iceland in 2012 is similar to what has been documented in previous studies (47, 100, 121). The same PFGE clone can be seen in carriages, and can also cause AOM or other diseases. This underlines the fact that the characteristics of the hosts immune defenses play a large role in what strains cause disease rather than certain strains being extremely virulent (23, 47, 92, 122).

5.4 Strength, weaknesses and improvement suggestions

If only the *hpd* primers had been used for differentiation of *Hi* and *H. haemolyticus*, 6.3% of the isolates from clinical samples and 5.9% of the *Hi* isolates from carriage samples would have been considered *H. haemolyticus* until further identification. Therefore, including two primer pairs does increase the differentiation method and also reduces the possibility of false negative results because of an obstruction in the PCR reaction.

A primer pair detecting a gene only present in *H. haemolyticus* would have been ideal. According to the available literature, a gene only present in *H. haemolyticus* has yet to be identified (20, 107).

Despite *fucK* and *hpd* genes being well-preserved in *Hi*, horizontal gene transfer can possibly cause *Hi* to lose the both genes or *H. haemolyticus* to gain them, therefore to unambiguously differentiate between the two species, whole genome sequencing is required (21). Because we found no non-hemolytic *H. haemolyticus* among the clinical isolates it would be preferable to examine hemolysis of previously identified non-hemolytic *H. haemolytcus* from other laboratories. This would determine if the horse blood agar does show hemolysis differently.

The multiplex serotyping PCR method developed in the author's diploma project did show some slight unspecific amplification. Therefore, a new duplex was developed to search for capsulated *Hi* among isolates from carriage samples. The new duplex method included primers detecting for capsule and *fucK* or *hpd* (depending on which gene was present in the differentiation between *Hi* and *H. haemolyticus* method) serving as an inner control. Both the old and new method had strengths, they are cheaper than real time PCR and more objective than the agglutination tests (81). What the new method has over the old one is an inner control that confirms that a PCR reaction did take place in each well, because in the old method amplification of no band meant the isolate was NT*Hi*. In general, there is little chance of an obstruction of amplification (86), but it is better to ensure a reaction in each well, if possible. There was significant difference between the number of capsulated isolates in clinical samples and the isolates from carriage samples. We need to take into consideration that the serotyping was not done with the same method. There is a slight possibility that the number of capsulated isolates was higher in the disease causing isolates, because no amplification of the genes in the multiplex PCR meant that the isolates were NT*Hi*. To make sure that the difference in the number of capsulated isolates is not due to the new method we would have to repeat all disease-causing isolates with the new duplex method.

Because of how time-consuming genotyping with PFGE is, we were only able to genotype 38% of all available isolates from 2012. Having more isolates genotyped would have given an even clearer picture of the genetic diversity of *Hi* in Iceland for 2012 and possibly higher statistical power. We first tried to develop a RAPD PCR method to genotype the isolates, but it only amplified 3-6 bands in each isolates and the amplicons from the reference strains were of similar size so the differentiation ability of the RAPD method was low. If the RAPD PCR method would have had more discriminatory power we would have been able to analyze all the *Hi* isolates in the given timeframe. The PFGE method did work well; this was seen in particular when an isolate from the same patient at a different time was identified as the same PFGE subclone. We only got ST for 12 of the 21 isolates, because of poor sequence quality. The sequence typed isolates did show that two isolates from the same PFGE subclone were both ST1030, which tells us that in this case the discriminatory power of the PFGE is good when compared to MLST. There are other studies that report identical clones from PFGE have the same results as MLST (102). In contradiction, we identified ST925 in two isolates from two different PFGE clones. Variation between gels might have created this inconsistency because PFGE clones 2 and 77 had similar band patterns although the similarity was only 72%. Placing these isolates next to each other on a gel, might have had revealed higher similarity in BioNumerics. This is a weakness of the PFGE method in general, especially when genotyping a highly diverse bacteria like *Hi*. The other main weakness of the PFGE method is that comparison of the results between

laboratories is difficult, MLST does, however, give unambiguous results that can prove to be important in epidemiological studies.

The next steps are to conduct a similar study on isolates from the year 2015 to monitor the effect of the Synflorix vaccine on the genetic diversity of *Hi*. Further analysis of the serotypes e, f and NT*Hi* isolates that were either identical or in the same PFGE clone are needed, preferably with whole genome sequencing. We also plan to analyze all 120 invasive *Hi* isolates that have been stored by the Department of Clinical Microbiology.

6 Conclusions

To our knowledge, the effect of Synforix 10-valent pneumococcal vaccine on genetic diversity of *Hi* has not been examined. It is a possibility that the genetic diversity of *Hi* will change after the possible effects of the Synforix 10-valent pneumococcal vaccine. To enable comparison of the era before vaccination and after, it is important to have a baseline like the one provided in our study.

The following conclusions are drawn from this study:

- *H. haemolyticus* was not found in isolates from clinical samples sent to the Department of Clinical Microbiology and isolates from carriage samples of children attending DCCs in 2012.
- If only the *hpd* primers had been used for differentiation between *Hi* and *H. haemolyticus,* around 6% of all the isolates would have been considered *H. haemolyticus* until further identification.
- Among 453 serotyped isolates from clinical samples sent to the clinical Department of Clinical Microbiology in 2012, one capsulated isolate was found (0.2%) which is significantly lower than the five capsulated *Hi* isolates found among 286 serotyped carriage isolates (1.8%)
- The *Hi* isolates showed high genetic diversity in PFGE. The 303 isolates analyzed with PFGE formed 72 PFGE clones and 49 isolates were singletons.
- PFGE clones contained *Hi* isolates both from carriage and clinical samples.
- A serotype f isolate and two serotype e isolates were classified together in one PFGE clone.
- A serotype f isolate was identical in PFGE to three NT*Hi* isolates.
- We identified one new ST, the other ST were in accordance to previous findings.

References

1. Forbes BA, Sahm DF, Weissfeld AS. Baily & Scott's diagnostic microbiology. 12 ed: Mosby; 2007.

2. Turk DC. The pathogenicity of *Haemophilus influenzae*. J Med Microbiol. 1984;18(1):1-16.

3. Barry JM. The Great Influenza: The story of the deadliest pandemic in history: Penguin Books; 2004.

4. Watt JP, Wolfson LJ, O'Brien KL, Henkle E, Deloria-Knoll M, McCall N, Lee E, Levine OS, Hajjeh R, Mulholland K, Cherian T. Burden of disease caused by *Haemophilus influenzae* type b in children younger than 5 years: global estimates. Lancet. 2009;374(9693):903-11.

5. WHO position paper on *Haemophilus influenzae* type b conjugate vaccines. (Replaces WHO position paper on *Hi*b vaccines previously published in the Weekly Epidemiological Record). Wkly Epidemiol Rec. 2006;81(47):445-52.

6. Ladhani S, Slack MP, Heath PT, von Gottberg A, Chandra M, Ramsay ME. Invasive *Haemophilus influenzae* disease, Europe, 1996-2006. Emerg Infect Dis. 2010;16(3):455-63.

7. Murphy TF, Faden H, Bakaletz LO, Kyd JM, Forsgren A, Campos J, Virji M, Pelton SI. Nontypeable *Haemophilus influenzae* as a pathogen in children. Pediatr Infect Dis J. 2009;28(1).

8. Marrs CF, Krasan GP, McCrea KW, Clemans DL, Gilsdorf JR. *Haemophilus influenzae*—human specific bacteria. Front Biosci. 2001;6:E41-E60.

9. Hallstrom T, Riesbeck K. *Haemophilus influenzae* and the complement system. Trends Microbiol. 2010;18(6):258-65.

10. Pittman M. Variation and type specificity in the bacterial species *Haemophilus influenzae*. J Exp Med. 1931;53(4):471-92.

11. Herbert MA, Hood DW, Moxon ER. *Haemophilus influenzae* protocols: Springer; 2003.

12. Swords WE, Jones PA, Apicella MA. The lipo-oligosaccharides of *Haemophilus influenzae:* an interesting array of characters. J Endotoxin Res. 2003;9(3):131-44.

13. Foxwell AR, Kyd JM, Cripps AW. Nontypeable *Haemophilus influenzae*: pathogenesis and prevention. Microbiol Mol Biol Rev. 1998;62(2):294-308.

14. Forsgren A, Riesbeck K, Janson H. Protein D of *Haemophilus influenzae*: a protective nontypeable *H. influenzae* antigen and a carrier for pneumococcal conjugate vaccines. Clin Infect Dis. 2008;46(5):726-31.

15. Song XM, Forsgren A, Janson H. The gene encoding protein D (*hpd*) is highly conserved among *Haemophilus influenzae* type b and nontypeable strains. Infect Immun. 1995;63(2):696-9.

16. Musher DM. *Haemophilus Species*. In: Baron S, editor. Medical Microbiology. Galveston TX: The University of Texas Medical Branch at Galveston; 1996.

17. Tian GZ, Zhang LJ, Wang XL, Zhang L, Li SF, Gu CM, Sun J, Cui BY. Rapid detection of *Haemophilus influenzae* and *Haemophilus parainfluenzae* in nasopharyngeal swabs by multiplex PCR. Biomed Environ Sci. 2012;25(3):367-71.

18. Harrison LH, da Silva GA, Pittman M, Fleming DW, Vranjac A, Broome CV. Epidemiology and clinical spectrum of Brazilian purpuric fever. Brazilian Purpuric Fever Study Group. J Clin Microbiol. 1989;27(4):599-604.

19. Kilian M. A taxonomic study of the genus *Haemophilus*, with the proposal of a new species. J Gen Microbiol. 1976;93(1):9 - 62.

20. Jordan IK, Conley AB, Antonov IV, Arthur RA, Cook ED, Cooper GP. Genome sequences for five strains of the emerging pathogen *Haemophilus haemolyticus*. J Bacteriol. 2011;193(20):5879-80.

21. Theodore MJ, Anderson RD, Wang X, Katz LS, Vuong JT, Bell ME, Juni BA, Lowther SA, Lynfield R, MacNeil JR, Mayer LW. Evaluation of new biomarker genes for differentiating *Haemophilus influenzae* from *Haemophilus haemolyticus*. J Clin Microbiol. 2012;50(4):1422-4.

22. Anderson R, Wang X, Briere EC, Katz LS, Cohn AC, Clark TA, Messonnier NE, Mayer LW. *Haemophilus haemolyticus* isolates causing clinical disease. J Clin Microbiol. 2012;50(7):2462-5.

23. Garmendia J, Martí-Lliteras P, Moleres J, Puig C, Bengoechea JA. Genotypic and phenotypic diversity in the noncapsulated *Haemophilus influenzae*: adaptation and pathogenesis in the human airways. Int Microbiol. 2013;15(4):157-70.

24. Erwin AL, Sandstedt SA, Bonthuis PJ, Geelhood JL, Nelson KL, Unrath WC, Diggle MA, Theodore MJ, Pleatman CR, Mothershed EA, Sacchi CT, Mayer LW, Gilsdorf JR, Smith AL. Analysis of genetic relatedness of *Haemophilus influenzae* isolates by multilocus sequence typing. J Bacteriol. 2008;190(4):1473-83.

25. Mhlanga-Mutangadura T, Morlin G, Smith AL, Eisenstark A, Golomb M. Evolution of the major pilus gene cluster of *Haemophilus influenzae*. J Bacteriol. 1998;180(17):4693-703.

26. Hendrixson DR, St. Geme JW 3rd. The *Haemophilus influenzae* hap serine protease promotes adherence and microcolony formation, potentiated by a soluble host protein. Mol Cell. 1998;2(6):841-50.

27. Yeo HJ, Cotter SE, Laarmann S, Juehne T, St. Geme JW 3rd, Waksman G. Structural basis for host recognition by the *Haemophilus influenzae* Hia autotransporter. EMBO J. 2004;23(6):1245-56.

28. Reinholdt J, Kilian M. Comparative analysis of immunoglobulin A1 protease activity among bacteria representing different genera, species, and strains. Infect Immun. 1997;65(11):4452-9.

29. Lomholt H. Molecular biology and vaccine aspects of bacterial immunoglobulin A1 proteases. APMIS Suppl. 1996;62:5-28.

30. Raetz CR, Guan Z, Ingram BO, Six DA, Song F, Wang X, Zhao J. Discovery of new biosynthetic pathways: the lipid A story. J Lipid Res. 2009;50(Supplement):S103-S8.

31. Lysenko E, Richards JC, Cox AD, Stewart A, Martin A, Kapoor M, Weiser JN. The position of phosphorylcholine on the lipopolysaccharide of *Haemophilus influenzae* affects binding and sensitivity to C-reactive protein-mediated killing. Mol Microbiol. 2000;35(1):234-45.

32. Prasadarao NV, Lysenko E, Wass CA, Kim KS, Weiser JN. Opacity-associated protein A contributes to the binding of *Haemophilus influenzae* to chang epithelial cells. Infect Immun. 1999;67(8):4153-60.

33. St. Geme JW 3rd. Molecular and cellular determinants of non-typeable *Haemophilus influenzae* adherence and invasion. Cell Microbiol. 2002;4(4):191-200.

34. Hill DJ, Toleman MA, Evans DJ, Villullas S, Van Alphen L, Virji M. The variable P5 proteins of typeable and non-typeable *Haemophilus influenzae* target human CEACAM1. Mol Microbiol. 2001;39(4):850-62.

35. St. Geme JW 3rd. The pathogenesis of nontypable *Haemophilus influenzae* otitis media. Vaccine. 2000;19, Supplement 1(0):S41-S50.

36. Satola SW, Schirmer PL, Farley MM. Complete sequence of the cap locus of *Haemophilus influenzae* serotype b and nonencapsulated b capsule-negative variants. Infect Immun. 2003;71(6):3639-44.

37. Mahillon J, Chandler M. Insertion sequences. Microbiol Mol Biol Rev. 1998;62(3):725-74.

38. Ogilvie C, Omikunle A, Wang Y, St. Geme JW 3rd, Rodriguez CA, Adderson EE. Capsulation loci of non-serotype b encapsulated *Haemophilus influenzae*. J Infect Dis. 2001;184(2):144-9.

39. Kapogiannis BG, Satola S, Keyserling HL, Farley MM. Invasive infections with *Haemophilus influenzae* serotype a containing an IS1016-bexA partial deletion: possible association with virulence. Clin Infect Dis. 2005;41(11):e97-103.

40. Kostyanev TS, Sechanova LP. Virulence factors and mechanisms of antibiotic resistance of *Haemophilus influenzae*. Folia Med (Plovdiv). 2012;54(1):19-23.

41. Pitout JD, Sanders CC, Sanders WEJ. Antimicrobial resistance with focus on beta-lactam resistance in gram-negative bacilli. Am J Med. 1997;103(1):51-9.

42. Harrison OB, Brueggemann AB, Caugant DA, van der Ende A, Frosch M, Gray S, Heuberger S, Krizova P, Olcen P, Slack M, Taha MK, Maiden MC. Molecular typing methods for outbreak detection and surveillance of invasive disease caused by *Neisseria meningitidis, Haemophilus influenzae* and *Streptococcus pneumoniae*, a review. Microbiology. 2011;157(Pt 8):2181-95.

43. Matic V, Bozdogan B, Jacobs MR, Ubukata K, Appelbaum PC. Contribution of beta-lactamase and PBP amino acid substitutions to amoxicillin/clavulanate resistance in beta-lactamase-positive, amoxicillin/clavulanate-resistant *Haemophilus influenzae*. J Antimicrob Chemother. 2003;52(6):1018-21.

44. Tristram S, Jacobs MR, Appelbaum PC. Antimicrobial resistance in *Haemophilus influenzae*. Clin Microbiol Rev. 2007;20(2):368-89.

45. Lipsitch M. Bacterial vaccines and serotype replacement: lessons from *Haemophilus influenzae* and prospects for *Streptococcus pneumoniae*. Emerg Infect Dis. 1999;5(3):336-45.

46. Laupland KB, Schønheyder HC, Østergaard C, Knudsen JD, Valiquette L, Galbraith J, Kennedy KJ, Gradel KO. Epidemiology of *Haemophilus influenzae* bacteremia: A multi-national population-based assessment. J Infect. 2011;62(2):142-8.

47. Kaur R, Chang A, Xu Q, Casey JR, Pichichero ME. Phylogenetic relatedness and diversity of non-typable *Haemophilus influenzae* in the nasopharynx and middle ear fluid of children with acute otitis media. J Med Microbiol. 2011;60(Pt 12):1841-8.

48. García-Rodríguez JÁ, Fresnadillo Martínez MJ. Dynamics of nasopharyngeal colonization by potential respiratory pathogens. J Antimicrob Chemother. 2002;50(suppl 3):59-74.

49. Hashida K, Shiomori T, Hohchi N, Muratani T, Mori T, Udaka T, Suzuki H. Nasopharyngeal *Haemophilus influenzae* carriage in Japanese children attending day-care centers. J Clin Microbiol. 2008;46(3):876-81.

50. Andrésdóttir BDI. Faraldsfræði *Streptococcus pneumoniae*, *S. pyogenes* og *Haemophilus sp.* í leikskólabörnum árið 2012. 2012.

51. de Carvalho CX, Kipnis A, Thorn L, de Andrade JG, Pimenta F, Brandileone MC, Zanella RC, Flannery B, Sgambatti S, Andrade AL. Carriage of *Haemophilus influenzae* among Brazilian children attending day care centers in the era of widespread *Hi*b vaccination. Vaccine. 2011;29(7):1438-42.

52. Dabernat H, Plisson-Saune MA, Delmas C, Seguy M, Faucon G, Pelissier R, Carsenti H, Pradier C, Roussel-Delvallez M, Leroy J, Dupont MJ, De Bels F, Dellamonica P. *Haemophilus influenzae* carriage in children attending French day care centers: a molecular epidemiological study. J Clin Microbiol. 2003;41(4):1664-72.

53. van Gils EJ, Veenhoven RH, Rodenburg GD, Hak E, Sanders EA. Effect of 7-valent pneumococcal conjugate vaccine on nasopharyngeal carriage with *Haemophilus influenzae* and *Moraxella catarrhalis* in a randomized controlled trial. Vaccine. 2011;29(44):7595-8.

54. Cripps AW, Otczyk DC, Kyd JM. Bacterial otitis media: a vaccine preventable disease? Vaccine. 2005;23(17-18):2304-10.

55. Leibovitz E, Jacobs MR, Dagan R. *Haemophilus influenzae*: a significant pathogen in acute otitis media. Pediatr Infect Dis J. 2004;23(12):1142-52.

56. Casey JR, Pichichero ME. Changes in frequency and pathogens causing acute otitis media in 1995-2003. Pediatr Infect Dis J. 2004;23(9):824-8.

57. Block SL, Hedrick J, Harrison CJ, Tyler R, Smith A, Findlay R, Keegan E. Community-wide vaccination with the heptavalent pneumococcal conjugate significantly alters the microbiology of acute otitis media. Pediatr Infect Dis J. 2004;23(9):829-33.

58. WHO. *Haemophilus influenzae* type b vaccines. (Immunological basis for immunization series ; module 9). Switzerland: Department of Immunization, Vaccines and Biologicals; 2007.

59. Ulanova M, Tsang RSW. Invasive *Haemophilus influenzae* disease: Changing epidemiology and host–parasite interactions in the 21st century. Infect Genet Evol. 2009;9(4):594-605.

60. Berndsen MR, Erlendsdottir H, Gottfredsson M. Evolving epidemiology of invasive *Haemophilus* infections in the post-vaccination era: results from a long-term population-based study. Clin Microbiol Infect. 2012;18(9):918-23.

61. Peltola H. Worldwide *Haemophilus influenzae* type b disease at the beginning of the 21st century: global analysis of the disease burden 25 years after the use of the polysaccharide vaccine and a decade after the advent of conjugates. Clin Microbiol Rev. 2000;13(2):302-17.

62. Jin Z, Romero-Steiner S, Carlone GM, Robbins JB, Schneerson R. *Haemophilus influenzae* type a infection and its prevention. Infect Immun. 2007;75(6):2650-4.

63. Dworkin MS, Park L, Borchardt SM. The changing epidemiology of invasive *Haemophilus influenzae* disease, especially in persons ≥65 years old. Clin Infect Dis. 2007;44(6):810-6.

64. White KE, Pabst LJ, Cullen KA. Up-to-date *Haemophilus influenzae* type b vaccination coverage during a vaccine shortage. Pediatrics. 2011;127(3):e707-e12.

65. MacNeil JR, Cohn AC, Farley M, Mair R, Baumbach J, Bennett N, Gershman K, Harrison LH, Lynfield R, Petit S, Reingold A, Schaffner W, Thomas A, Coronado F, Zell ER, Mayer LW, Clark TA, Messonnier NE. Current epidemiology and trends in invasive *Haemophilus influenzae* disease--United States, 1989-2008. Clin Infect Dis. 2011;53(12):1230-6.

66. Cherpes TL, Kusne S, Hillier SL. *Haemophilus influenzae* septic abortion. Infect Dis Obstet Gynecol. 2002;10(3):161-4.

67. Avery OT, Goebel WF. Chemo-immunological studies on conjugated carbohydrate-proteins: V. The immunological specifity of an antigen prepared by combining the capsular polysaccharide of type III pneumococcus with foreign protein. J Exp Med. 1931;54(3):437-47.

68. Kelly DF, Moxon ER, Pollard AJ. *Haemophilus influenzae* type b conjugate vaccines. Immunology. 2004;113(2):163-74.

69. Jónsdóttir KE, Hansen HJ, Arnórsson VH, Laxdal Þ, Stefánsson ML. Ungbarnabólusetning á Íslandi gegn *Haemophilus influenzae* af hjúpgerð b: árangur eftir sex ára notkun PRP-D (ProHIBiT). Læknablaðið. 1996;82(1):32-8.

70. Eskola J, Peltola H, Takala AK, Kayhty H, Hakulinen M, Karanko V, Kela E, Rekola P, Ronnberg PR, Samuelson JS, et al. Efficacy of *Haemophilus influenzae* type b polysaccharide-diphtheria toxoid conjugate vaccine in infancy. N Engl J Med. 1987;317(12):717-22.

71. Landlæknir. Almennar bólusetningar barna á Íslandi frá janúar 2013. 2013 [cited 2013 March 20th]; Available from: http://www.landlaeknir.is/servlet/file/store93/item19484/Almennar_bolusetningar_barna_eftir_jan%C3%BAar 2013.pdf.

72. Prymula R, Peeters P, Chrobok V, Kriz P, Novakova E, Kaliskova E, Kohl I, Lommel P, Poolman J, Prieels J-P, Schuerman L. Pneumococcal capsular polysaccharides conjugated to protein D for prevention of acute otitis media caused by both *Streptococcus pneumoniae* and non-typable *Haemophilus influenzae*: a randomised double-blind efficacy study. Lancet.367(9512):740-8.

73. van den Bergh MR, Spijkerman J, Swinnen KM, Francois NA, Pascal TG, Borys D, Schuerman L, Ijzerman EP, Bruin JP, van der Ende A, Veenhoven RH, Sanders EA. Effects of the 10-valent pneumococcal nontypeable *Haemophilus influenzae* protein D-conjugate vaccine on nasopharyngeal bacterial colonization in young children: a randomized controlled trial. Clin Infect Dis. 2013;56(3):e30-9.

74. Prymula R, Kriz P, Kaliskova E, Pascal T, Poolman J, Schuerman L. Effect of vaccination with pneumococcal capsular polysaccharides conjugated to *Haemophilus influenzae*-derived protein D on nasopharyngeal carriage of *Streptococcus pneumoniae* and *H. influenzae* in children under 2 years of age. Vaccine. 2009;28(1):71-8.

75. Zinnemann K, Rogers KB, Frazer J, Devaraj SK. A haemolytic V-dependent CO2-preferring *Haemophilus* species (*Haemophilus paraphrohaemolyticus* Nov. Spec.). J Med Microbiol. 1971;4(1):139-43.

76. Leclercq R, Cantón R, Brown DFJ, Giske CG, Heisig P, MacGowan AP, Mouton JW, Nordmann P, Rodloff AC, Rossolini GM, Soussy CJ, Steinbakk M, Winstanley TG, Kahlmeter G. EUCAST expert rules in antimicrobial susceptibility testing. Clin Microbiol Infect. 2013;19(2):141-60.

77. EUCAST. Frequently asked questions. 2012 [cited 2013 april 10th]; Available from: http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Disk_test_documents/120210_Frequently_Asked_Questions_EUCAST.pdf.

78. O'Callaghan CH, Morris A, Kirby SM, Shingler AH. Novel method for detection of beta-lactamases by using a chromogenic cephalosporin substrate. Antimicrob Agents Chemother. 1972;1(4):283-8.

79. CDC. Laboratory methods for the diagnosis of meningitis. 2011 [cited 2013 March 20th]; Available from: http://www.cdc.gov/meningitis/lab-manual/chpt09-id-characterization-hi.pdf.

80. LaClaire LL, Tondella MLC, Beall DS, Noble CA, Raghunathan PL, Rosenstein NE, Popovic T. Identification of *Haemophilus influenzae* serotypes by standard slide agglutination serotyping and PCR-based capsule typing. J Clin Microbiol. 2003;41(1):393-6.

81. Wroblewski D, Halse TA, Hayes J, Kohlerschmidt D, Musser KA. Utilization of a real-time PCR approach for *Haemophilus influenzae* serotype determination as an alternative to the slide agglutination test. Mol Cell Probes. 2013;27(2):86-9.

82. Shively RG, Shigei JT, Peterson EM, de la Maza LM. Typing of *Haemophilus influenzae* by coagglutination and conventional slide agglutination. J Clin Microbiol. 1981;14(6):706-8.

83. Bokermann S, Zanella RC, Lemos AP, de Andrade AL, Brandileone MC. Evaluation of methodology for serotyping invasive and nasopharyngeal isolates of *Haemophilus influenzae* in the ongoing surveillance in Brazil. J Clin Microbiol. 2003;41(12):5546-50.

84. Markham A. The polymerase chain reaction: a tool for molecular medicine. BMJ. 1993;306(6875):441-6.

85. Persing DH, Tenover FC, Versalovic J, YiWei T, Unger ER, Relman DA, White T. Molecular microbiology: diagnostic principles and practice: ASM Press; 2004.

86. Edwards MC, Gibbs RA. Multiplex PCR: advantages, development, and applications. PCR Methods Appl. 1994;3(4):S65-75.

87. Norskov-Lauritsen N, Overballe MD, Kilian M. Delineation of the species *Haemophilus influenzae* by phenotype, multilocus sequence phylogeny, and detection of marker genes. J Bacteriol. 2009;191(3):822-31.

88. McCrea KW, Xie J, LaCross N, Patel M, Mukundan D, Murphy TF, Marrs CF, Gilsdorf JR. Relationships of nontypeable *Haemophilus influenzae* strains to hemolytic and nonhemolytic *Haemophilus haemolyticus* strains. J Clin Microbiol. 2008;46(2):406-16.

89. Ridderberg W, Fenger MG, Norskov-Lauritsen N. *Haemophilus influenzae* may be untypable by the multilocus sequence typing scheme due to a complete deletion of the fucose operon. J Med Microbiol. 2010;59(Pt 6):740-2.

90. Falla TJ, Crook DW, Brophy LN, Maskell D, Kroll JS, Moxon ER. PCR for capsular typing of *Haemophilus influenzae*. J Clin Microbiol. 1994;32(10):2382-6.

91. Davis GS, Sandstedt SA, Patel M, Marrs CF, Gilsdorf JR. Use of *bexB* to detect the capsule locus in *Haemophilus influenzae*. J Clin Microbiol. 2011;49(7):2594-601.

92. Murphy TF, Lesse AJ, Kirkham C, Zhong H, Sethi S, Munson RSJ. A clonal group of nontypeable *Haemophilus influenzae* with two IgA proteases is adapted to infection in chronic obstructive pulmonary disease. PLoS One. 2011;6(10):e25923.

93. Gilsdorf JR. Antigenic diversity and gene polymorphisms in *Haemophilus influenzae*. Infect Immun. 1998;66(11):5053-9.

94. Connor TR, Corander J, Hanage WP. Population subdivision and the detection of recombination in non-typable *Haemophilus influenzae*. Microbiology. 2012;158(Pt 12):2958-64.

95. Goering RV. Pulsed field gel electrophoresis: a review of application and interpretation in the molecular epidemiology of infectious disease. Infect Genet Evol. 2010;10(7):866-75.

96. Chu G. Bag model for DNA migration during pulsed-field electrophoresis. Proc Natl Acad Sci USA. 1991;88(24):11071-5.

97. Tenover FC, Arbeit RD, Goering RV, Mickelsen PA, Murray BE, Persing DH, Swaminathan B. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol. 1995;33(9):2233-9.

98. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci USA. 1998;95(6):3140-5.

99. Urwin R, Maiden MC. Multi-locus sequence typing: a tool for global epidemiology. Trends Microbiol. 2003;11(10):479-87.

100. LaCross NC, Marrs CF, Gilsdorf JR. Population structure in nontypeable *Haemophilus influenzae*. Infect Genet Evol. 2013;14:125-36.

101. Cooper JE, Feil EJ. Multilocus sequence typing–what is resolved? Trends Microbiol. 2004;12(8):373-7.

102. Puig C, Calatayud L, Marti S, Tubau F, Garcia-Vidal C, Carratala J, Linares J, Ardanuy C. Molecular epidemiology of nontypeable *Haemophilus influenzae* causing community-acquired pneumonia in adults. PLoS One. 2013;8(12):e82515.

103. Tsang RSW, Mubareka S, Sill ML, Wylie J, Skinner S, Law DKS. Invasive *Haemophilus influenzae* in Manitoba, Canada, in the postvaccination era. J Clin Microbiol. 2006;44(4):1530-5.

104. Cardines R, Giufre M, Mastrantonio P, Ciofi degli Atti ML, Cerquetti M. Nontypeable *Haemophilus influenzae* meningitis in children: phenotypic and genotypic characterization of isolates. Pediatr Infect Dis J. 2007;26(7):577-82.

105. O'Neill JM, St. Geme JW 3rd, Cutter D, Adderson EE, Anyanwu J, Jacobs RF, Schutze GE. Invasive disease due to nontypeable *Haemophilus influenzae* among children in Arkansas. J Clin Microbiol. 2003;41(7):3064-9.

106. Björnsdóttir JB. *Haemophilus influenzae* - Aðgreining frá *Haemophilus haemolyticus* og hjúpgreining með multiplex PCR aðferð. Reykjavík: Háskóli Íslands; 2012.

107. Binks MJ, Temple B, Kirkham LA, Wiertsema SP, Dunne EM, Richmond PC, Marsh RL, Leach AJ, Smith-Vaughan HC. Molecular surveillance of true nontypeable *Haemophilus influenzae*: an evaluation of PCR screening assays. PLoS One. 2012;7(3):e34083.

108. Meyler KL, Meehan M, Bennett D, Cunney R, Cafferkey M. Development of a diagnostic real-time polymerase chain reaction assay for the detection of invasive *Haemophilus influenzae* in clinical samples. Diagn Microbiol Infect Dis. 2012;74(4):356-62.

109. Sa-Leao R, Tomasz A, Sanches IS, Brito-Avo A, Vilhelmsson SE, Kristinsson KG, de Lencastre H. Carriage of internationally spread clones of *Streptococcus pneumoniae* with unusual drug resistance patterns in children attending day care centers in Lisbon, Portugal. J Infect Dis. 2000;182(4):1153-60.

110. Sabrià MMDP, García-Núñez MB, Pedro-Botet MLMDP, Sopena NMD, Gimeno JMMD, Reynaga EMD, Morera JMDP, Rey-Joly CMDP. Presence and chromosomal subtyping of *Legionella* species in potable water systems in 20 hospitals of Catalonia, Spain. Infect Control Hosp Epidemiol. 2001;22(11):673-6.

111. Murphy TF, Brauer AL, Sethi S, Kilian M, Cai X, Lesse AJ. *Haemophilus haemolyticus*: A human respiratory tract commensal to be distinguished from *Haemophilus influenzae*. J Clin Microbiol. 2007;195(1):81-9.

112. Kirkham LAS, Wiertsema SP, Mowe EN, Bowman JM, Riley TV, Richmond PC. Nasopharyngeal carriage of *Haemophilus haemolyticus* in otitis-prone and healthy children. J Clin Microbiol. 2010;48(7):2557-9.

113. Nørskov-Lauritsen N. Detection of cryptic genospecies misidentified as Haemophilus influenzae in routine clinical samples by assessment of marker genes *fucK*, *hap*, and *sodC*. J Clin Microbiol. 2009;47(8):2590-2.

114. Ladhani S, Ramsay ME, Chandra M, Slack MP. No evidence for *Haemophilus influenzae* serotype replacement in Europe after introduction of the *Hi*b conjugate vaccine. Lancet. 2008;8(5):275-6.

115. Talon D, Leroy J, Dupont MJ, Bertrand X, Mermet F, Thouverez M, Estavoyer JM. Antibiotic susceptibility and genotypic characterization of *Haemophilus influenzae* strains isolated from nasopharyngeal specimens from children in day-care centers in eastern France. Clin Microbiol Infect. 2000;6(10):519-24.

116. Marchisio P, Gironi S, Esposito S, Schito GC, Mannelli S, Principi N. Seasonal variations in nasopharyngeal carriage of respiratory pathogens in healthy Italian children attending day-care centres or schools. J Med Microbiol. 2001;50(12):1095-9.

117. Ohkusu K, Nash KA, Inderlied CB. Molecular characterisation of *Haemophilus influenzae* type a and untypeable strains isolated simultaneously from cerebrospinal fluid and blood: novel use of quantitative real-time PCR based on the cap copy number to determine virulence. Clin Microbiol Infect. 2005;11(8):637-43.

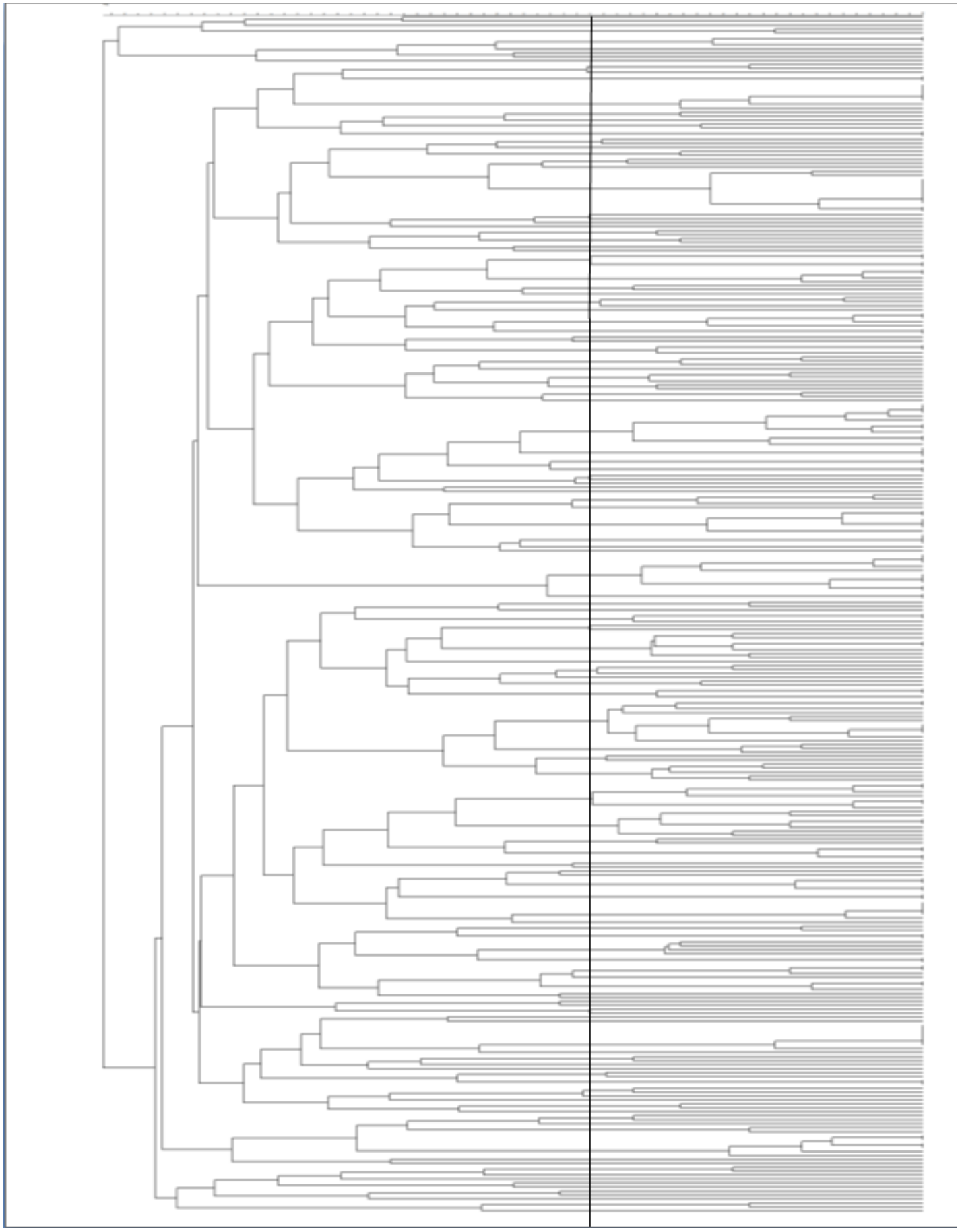
118. Bajanca P, Canica M. Emergence of nonencapsulated and encapsulated non-b-type invasive *Haemophilus influenzae* isolates in Portugal (1989-2001). J Clin Microbiol. 2004;42(2):807-10.

119. Fleischmann R, Adams M, White O, Clayton R, Kirkness E, Kerlavage A, Bult C, Tomb J, Dougherty B, Merrick J. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science. 1995;269(5223):496 - 512.

120. Skoczyńska A, Lewandowska M, Klarowicz A, Hryniewicz W. Prevalence and serotype distribution of encapsulated *Haemophilus influenzae* isolates from patients with lower respiratory tract infections in Poland. J Clin Microbiol. 2005;43(2):938-41.
121. Meats E, Feil E, Stringer S, Cody A, Goldstein R, Kroll J, Popovic T, Spratt B. Characterization of encapsulated and noncapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing. J Clin Microbiol. 2003;41(4):1623 - 36.
122. De Chiara M, Hood D, Muzzi A, Pickard DJ, Perkins T, Pizza M, Dougan G, Rappuoli R, Moxon ER, Soriani M, Donati C. Genome sequencing of disease and carriage isolates of nontypeable *Haemophilus influenzae* identifies discrete population structure. Proc Natl Acad Sci USA. 2014;111(14):5439-44.
123. Omikunle A, Takahashi S, Ogilvie CL, Wang Y, Rodriguez CA, St. Geme JW 3rd, Adderson EE. Limited genetic diversity of recent invasive isolates of non-serotype b encapsulated *Haemophilus influenzae*. J Clin Microbiol. 2002;40(4):1264-70.

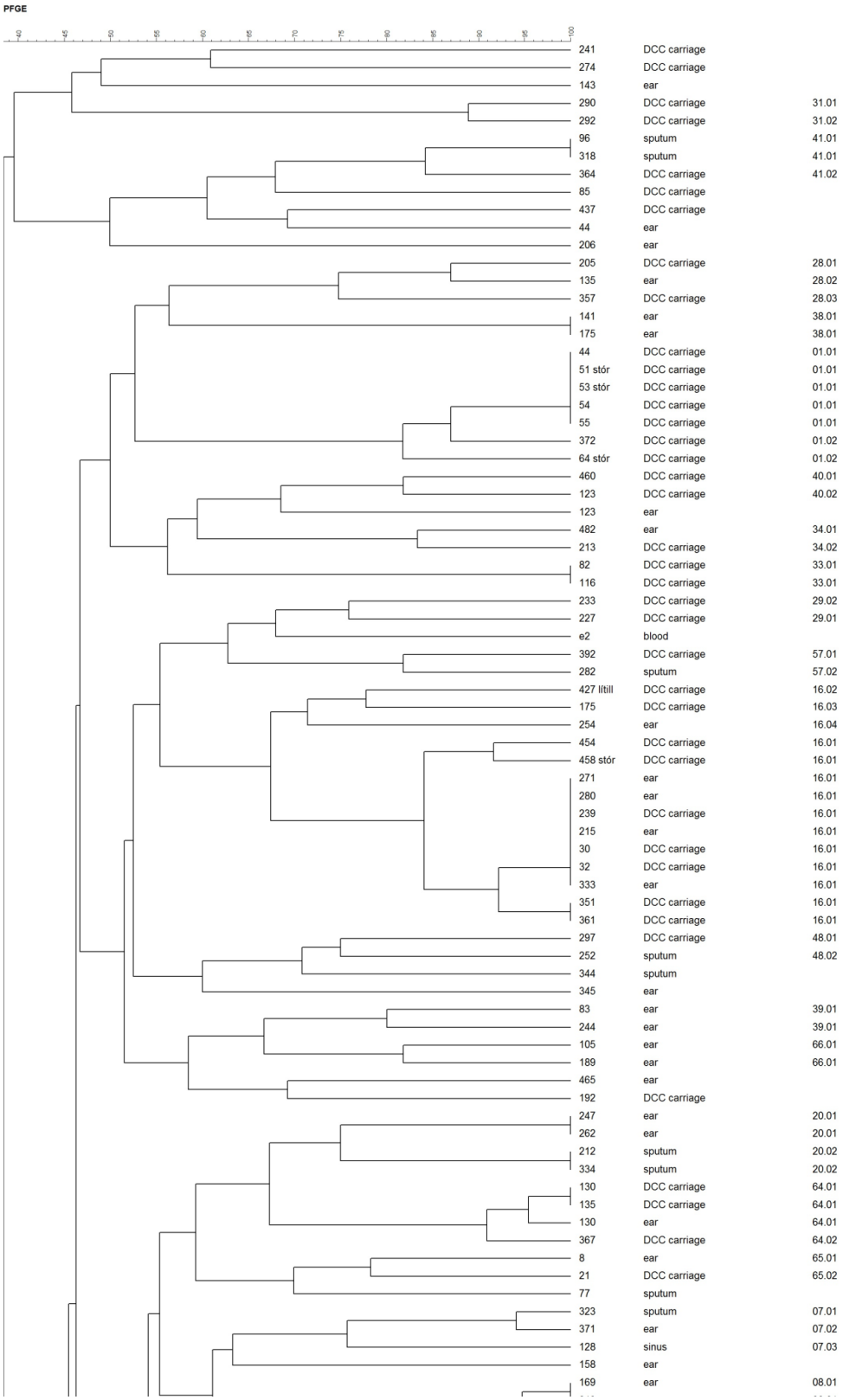
Appendix I

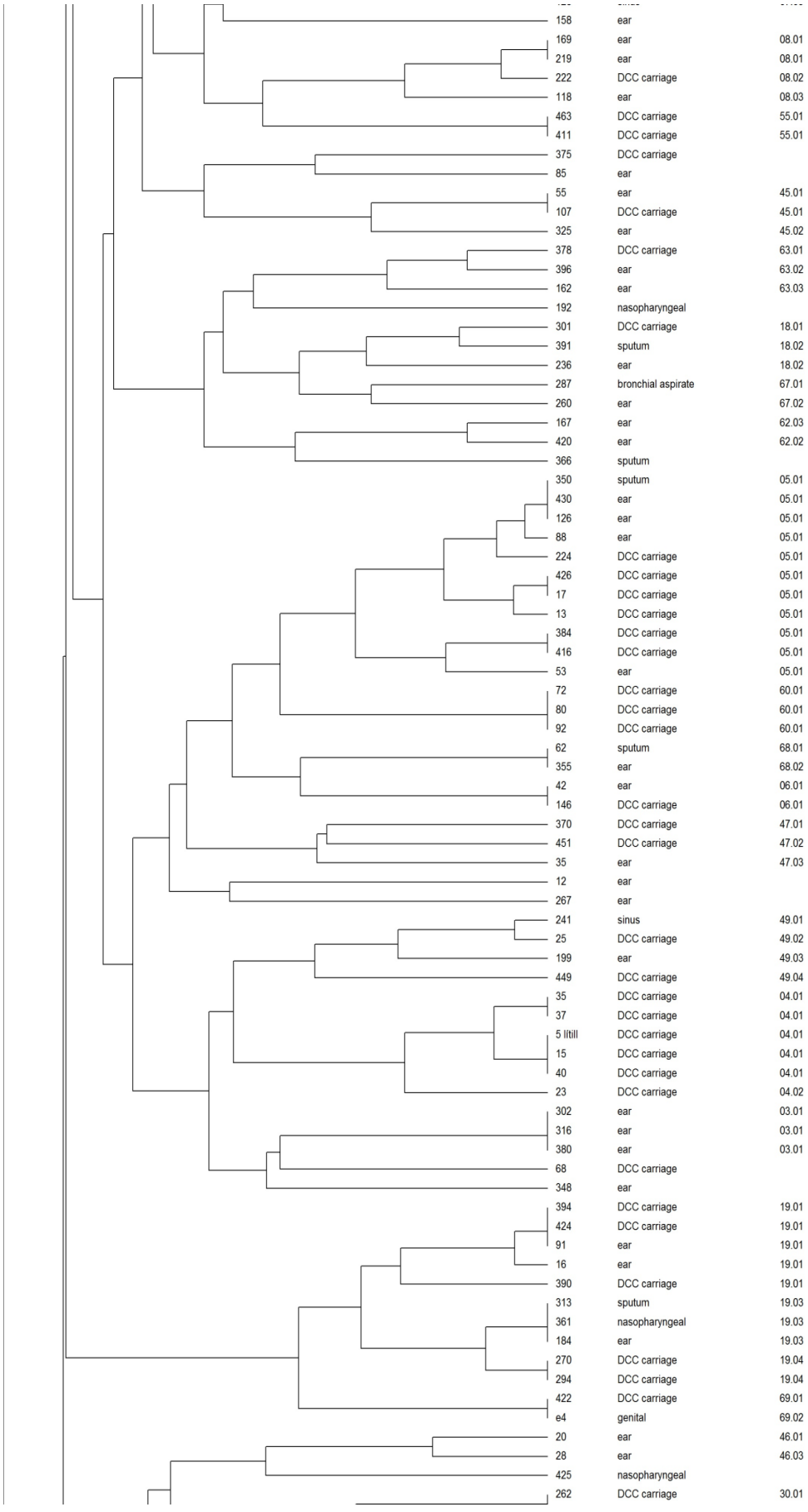
The dendrogram of all *Hi* isolates analyzed with PFGE, the vertical line represents 75% similarity.

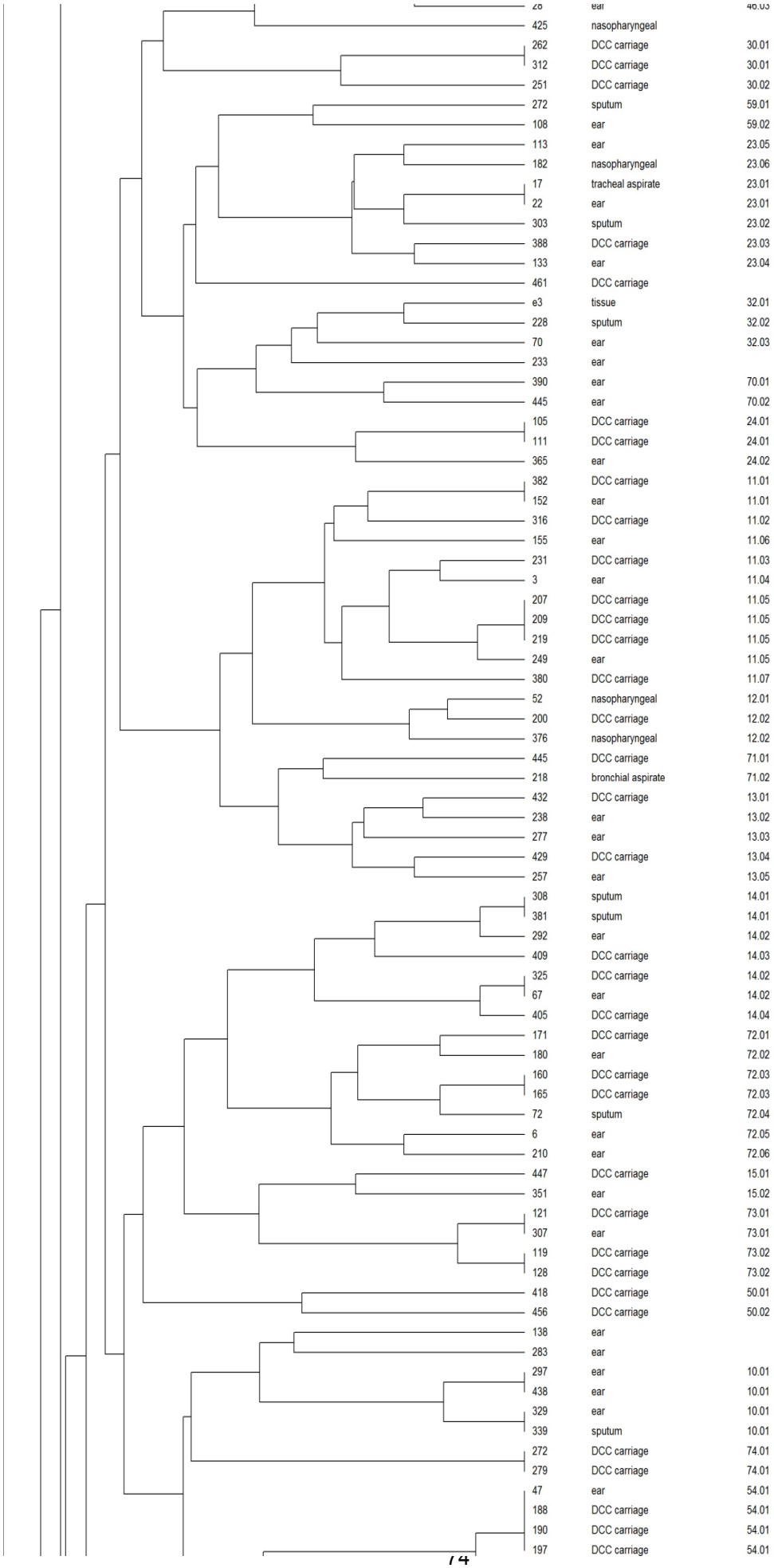


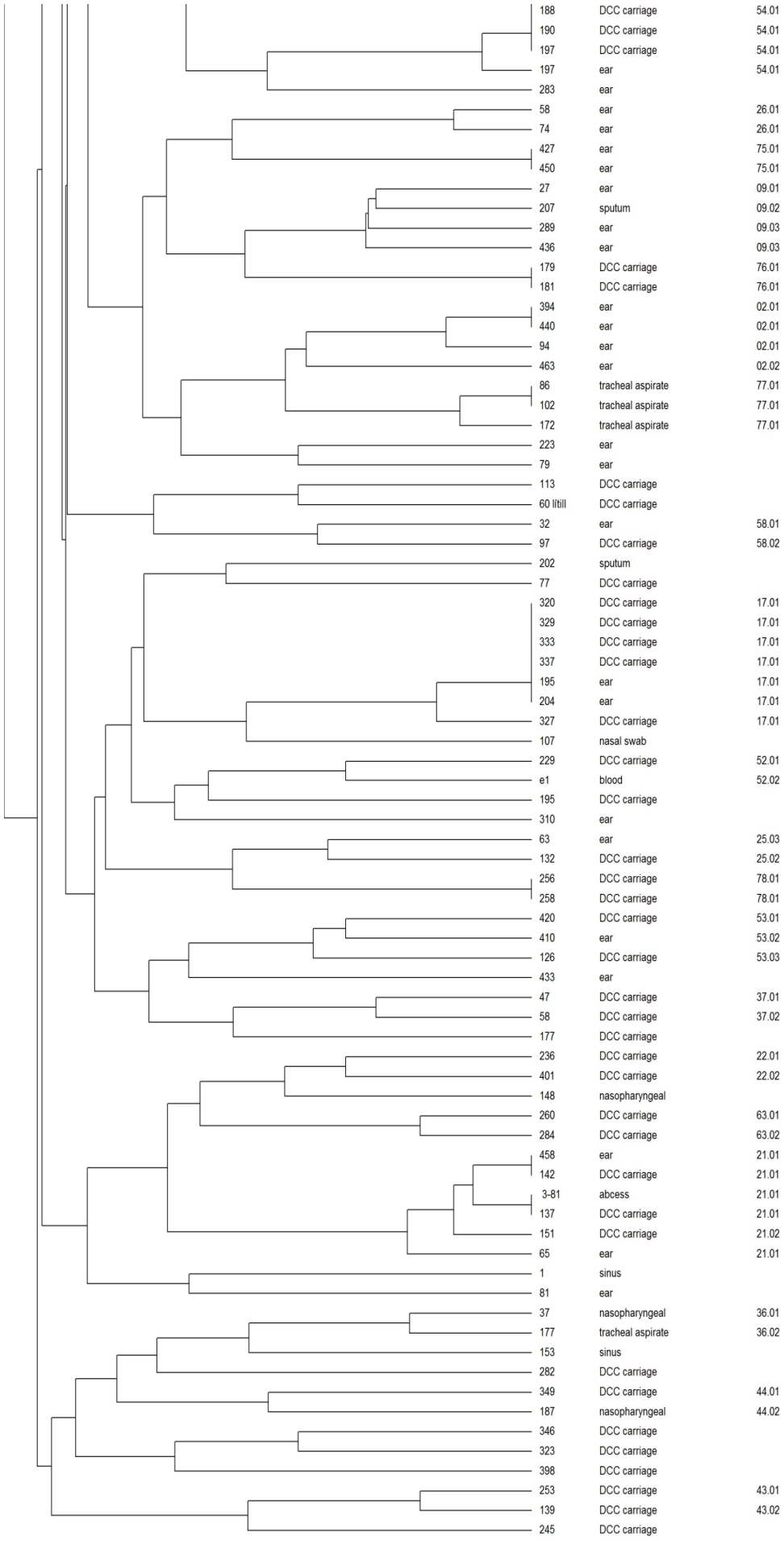
Appendix II

The dendrogram with information on the isolate specimen type and PFGE clone number.




158 ear
169 ear 08.01
219 ear 08.01
222 DCC carriage 08.02
118 ear 08.03
463 DCC carriage 55.01
411 DCC carriage 55.01
375 DCC carriage
85 ear
55 ear 45.01
107 DCC carriage 45.01
325 ear 45.02
378 DCC carriage 63.01
396 ear 63.02
162 ear 63.03
192 nasopharyngeal
301 DCC carriage 18.01
391 sputum 18.02
236 ear 18.02
287 bronchial aspirate 67.01
260 ear 67.02
167 ear 62.03
420 ear 62.02
366 sputum
350 sputum 05.01
430 ear 05.01
126 ear 05.01
88 ear 05.01
224 DCC carriage 05.01
426 DCC carriage 05.01
17 DCC carriage 05.01
13 DCC carriage 05.01
384 DCC carriage 05.01
416 DCC carriage 05.01
53 ear 05.01
72 DCC carriage 60.01
80 DCC carriage 60.01
92 DCC carriage 60.01
62 sputum 68.01
355 ear 68.02
42 ear 06.01
146 DCC carriage 06.01
370 DCC carriage 47.01
451 DCC carriage 47.02
35 ear 47.03
12 ear
267 ear
241 sinus 49.01
25 DCC carriage 49.02
199 ear 49.03
449 DCC carriage 49.04
35 DCC carriage 04.01
37 DCC carriage 04.01
5 lltill DCC carriage 04.01
15 DCC carriage 04.01
40 DCC carriage 04.01
23 DCC carriage 04.02
302 ear 03.01
316 ear 03.01
380 ear 03.01
68 DCC carriage
348 ear
394 DCC carriage 19.01
424 DCC carriage 19.01
91 ear 19.01
16 ear 19.01
390 DCC carriage 19.01
313 sputum 19.03
361 nasopharyngeal 19.03
184 ear 19.03
270 DCC carriage 19.04
294 DCC carriage 19.04
422 DCC carriage 69.01
e4 genital 69.02
20 ear 46.01
28 ear 46.03
425 nasopharyngeal
262 DCC carriage 30.01


26 ear 46.03
425 nasopharyngeal
262 DCC carriage 30.01
312 DCC carriage 30.01
251 DCC carriage 30.02
272 sputum 59.01
108 ear 59.02
113 ear 23.05
182 nasopharyngeal 23.06
17 tracheal aspirate 23.01
22 ear 23.01
303 sputum 23.02
388 DCC carriage 23.03
133 ear 23.04
461 DCC carriage
e3 tissue 32.01
228 sputum 32.02
70 ear 32.03
233 ear
390 ear 70.01
445 ear 70.02
105 DCC carriage 24.01
111 DCC carriage 24.01
365 ear 24.02
382 DCC carriage 11.01
152 ear 11.01
316 DCC carriage 11.02
155 ear 11.06
231 DCC carriage 11.03
3 ear 11.04
207 DCC carriage 11.05
209 DCC carriage 11.05
219 DCC carriage 11.05
249 ear 11.05
380 DCC carriage 11.07
52 nasopharyngeal 12.01
200 DCC carriage 12.02
376 nasopharyngeal 12.02
445 DCC carriage 71.01
218 bronchial aspirate 71.02
432 DCC carriage 13.01
238 ear 13.02
277 ear 13.03
429 DCC carriage 13.04
257 ear 13.05
308 sputum 14.01
381 sputum 14.01
292 ear 14.02
409 DCC carriage 14.03
325 DCC carriage 14.02
67 ear 14.02
405 DCC carriage 14.04
171 DCC carriage 72.01
180 ear 72.02
160 DCC carriage 72.03
165 DCC carriage 72.03
72 sputum 72.04
6 ear 72.05
210 ear 72.06
447 DCC carriage 15.01
351 ear 15.02
121 DCC carriage 73.01
307 ear 73.01
119 DCC carriage 73.02
128 DCC carriage 73.02
418 DCC carriage 50.01
456 DCC carriage 50.02
138 ear
283 ear
297 ear 10.01
438 ear 10.01
329 ear 10.01
339 sputum 10.01
272 DCC carriage 74.01
279 DCC carriage 74.01
47 ear 54.01
188 DCC carriage 54.01
190 DCC carriage 54.01
197 DCC carriage 54.01


188 DCC carriage 54.01
190 DCC carriage 54.01
197 DCC carriage 54.01
197 ear 54.01
283 ear
58 ear 26.01
74 ear 26.01
427 ear 75.01
450 ear 75.01
27 ear 09.01
207 sputum 09.02
289 ear 09.03
436 ear 09.03
179 DCC carriage 76.01
181 DCC carriage 76.01
394 ear 02.01
440 ear 02.01
94 ear 02.01
463 ear 02.02
86 tracheal aspirate 77.01
102 tracheal aspirate 77.01
172 tracheal aspirate 77.01
223 ear
79 ear
113 DCC carriage
60 l/till DCC carriage
32 ear 58.01
97 DCC carriage 58.02
202 sputum
77 DCC carriage
320 DCC carriage 17.01
329 DCC carriage 17.01
333 DCC carriage 17.01
337 DCC carriage 17.01
195 ear 17.01
204 ear 17.01
327 DCC carriage 17.01
107 nasal swab
229 DCC carriage 52.01
e1 blood 52.02
195 DCC carriage
310 ear
63 ear 25.03
132 DCC carriage 25.02
256 DCC carriage 78.01
258 DCC carriage 78.01
420 DCC carriage 53.01
410 ear 53.02
126 DCC carriage 53.03
433 ear
47 DCC carriage 37.01
58 DCC carriage 37.02
177 DCC carriage
236 DCC carriage 22.01
401 DCC carriage 22.02
148 nasopharyngeal
260 DCC carriage 63.01
284 DCC carriage 63.02
458 ear 21.01
142 DCC carriage 21.01
3-81 abcess 21.01
137 DCC carriage 21.01
151 DCC carriage 21.02
65 ear 21.01
1 sinus
81 ear
37 nasopharyngeal 36.01
177 tracheal aspirate 36.02
153 sinus
282 DCC carriage
349 DCC carriage 44.01
187 nasopharyngeal 44.02
346 DCC carriage
323 DCC carriage
398 DCC carriage
253 DCC carriage 43.01
139 DCC carriage 43.02
245 DCC carriage

Appendix III

Information on each PFGE clone, number of isolates of each specimen type and the percentage of total in each PFGE clone.

clone	number of sub clones	Specimen type and number of isolates		% in clone
1	2	DCCs carriage	7	100
2	2	ear	4	100
3	1	ear	3	100
4	2	DCCs carriage	6	100
5	1	sputum	1	9
		ear	4	36
		DCCs carriage	6	55
6	1	ear	1	50
		DCCs carriage	1	50
7	3	sputum	1	33
		ear	1	33
		sinus	1	33
8	3	ear	3	75
		DCCs carriage	1	25
9	3	ear	3	75
		sputum	1	25
10	1	ear	3	75
		sputum	1	25
11	7	DCCs carriage	7	64
		ear	4	36
12	2	nasopharynx	2	67
		DCCs carriage	1	33
13	5	DCCs carriage	2	40
		ear	3	60
14	4	sputum	2	29
		ear	2	29
		DCCs carriage	3	42
15	2	ear	1	50
		DCCs carriage	1	50
16	4	DCCs carriage	10	71
		ear	4	29
17	1	DCCs carriage	5	71
		ear	2	29
18	2	DCCs carriage	1	33
		sputum	1	33

clone	number of sub clones	Specimen type and number of isolates		% in clone
		ear	1	33
19	4	DCCs carriage	5	50
		ear	3	30
		sputum	1	10
		nasopharynx	1	10
20	2	ear	2	50
		sputum	2	50
21	2	DCCs carriage	3	50
		ear	2	33
		abcess	1	17
22	2	DCCs carriage	2	100
23	6	ear	3	43
		tracheal aspirate	1	14
		nasopharynx	1	14
		DCCs carriage	1	14
		sputum	1	14
24	2	DCCs carriage	2	67
		ear	1	33
25	2	ear	1	50
		DCCs carriage	1	50
26	1	ear	2	100
28	3	DCCs carriage	2	67
		ear	1	33
29	2	DCCs carriage	2	100
30	2	DCCs carriage	3	100
31	2	DCCs carriage	2	100
32	3	tissue	1	33
		sputum	1	33
		ear	1	33
33	1	DCCs carriage	2	100
34	2	DCCs carriage	1	50
		ear	1	50
36	2	nasopharynx	1	50
		tracheal aspirate	1	50
37	2	DCCs carriage	2	100
38	1	ear	2	100
39	1	ear	2	100
40	2	DCCs carriage	2	100
41	2	sputum	2	67
		DCCs carriage	1	33
43	2	DCCs carriage	2	100

clone	number of sub clones	Specimen type and number of isolates		% in clone
44	2	DCCs carriage	1	50
		nasopharynx	1	50
45	2	ear	2	67
		DCCs carriage	1	33
46	2	ear	2	100
47	3	DCCs carriage	2	67
		ear	1	33
48	2	DCCs carriage	1	50
		sputum	1	50
49	4	sinus	1	25
		DCCs carriage	2	50
		ear	1	25
50	2	DCCs carriage	2	100
52	2	DCCs carriage	1	50
		blood	1	50
53	3	DCCs carriage	2	67
		ear	1	33
54	1	ear	2	40
		DCCs carriage	3	60
55	1	DCCs carriage	2	100
57	2	DCCs carriage	1	50
		sputum	1	50
58	2	DCCs carriage	1	50
		ear	1	50
59	2	sputum	1	50
		ear	1	50
60	1	DCCs carriage	3	100
62	2	ear	2	100
63	3	ear	2	40
		DCCs carriage	3	60
64	2	DCCs carriage	3	75
		ear	1	25
65	2	DCCs carriage	1	50
		ear	1	50
66	1	ear	2	100
67	2	bronchial aspirate	1	50
		ear	1	50
68	2	sputum	1	50
		ear	1	50
69	2	genital	1	50
		DCCs carriage	1	50

clone	number of sub clones	Specimen type and number of isolates		% in clone
70	2	ear	2	100
71	2	DCCs carriage	1	50
		bronchial aspirate	1	50
72	6	ear	3	43
		DCCs carriage	3	43
		sputum	1	14
73	2	DCCs carriage	3	75
		ear	1	25
74	1	DCCs carriage	2	100
75	1	ear	2	100
76	1	DCCs carriage	2	100
77	1	tracheal aspirate	3	100
78	1	DCCs carriage	2	100
singletons	49	blood	1	2
		DCCs carriage	18	37
		ear	20	41
		sputum	4	8
		nasopharynx	3	6
		nasal	1	2
		sinus	2	4