Epidemiol. Infect. (2017), **145**, 1727–1736. © Cambridge University Press 2017
doi:10.1017/S0950268817000450

Characterisation of invasive clinical *Haemophilus influenzae* isolates in Queensland, Australia using whole-genome sequencing

M. STAPLES*, R. M. A. GRAHAM AND A. V. JENNISON

Queensland Department of Health, Public Health Microbiology, Forensic and Scientific Services, Brisbane, Queensland, Australia

Received 21 December 2016; Final revision 19 January 2017; Accepted 10 February 2017; first published online 6 March 2017

SUMMARY

Haemophilus influenzae is an important aetiological organism of both adult and child respiratory disease. The number of non-typeable (NTHi) invasive *H. influenzae* isolates referred to the Queensland (QLD) Public Health Microbiology laboratory has increased notably year-by-year. In this study we used whole-genome sequencing to molecularly characterise 100 referred invasive *H. influenzae*, including 74 NTHi isolates over a 15-year period, observing the carriage of capsular and putative virulence genes, including the major adhesins, antimicrobial resistance genes and population diversity. Encapsulated isolates were largely clonal, however NTHi isolates displayed high genetic variability by MLST and single nucleotide polymorphism typing with no dominant clone observed. The only mechanism for β-lactam resistance identified in the QLD isolates was β-lactamase production. No single set of virulence determinants was conclusively associated with invasive QLD NTHi isolates.

Key words: Bacterial infections, emerging infections, *Haemophilus influenzae*, infectious disease, molecular epidemiology.

INTRODUCTION

Haemophilus influenzae is an important cause of both adult and child respiratory disease and severe invasive disease. Strains are characterised largely by the presence or absence of a polysaccharide capsule (CPS). Encapsulated strains (Hi) are further classified serologically, possessing one of six different capsular types, designated a to f. Prior to the inclusion of the *H. influenzae* b (Hib) vaccine in the Australian National Immunisation Program Schedule in 1993, Hib was a common cause of invasive infections, particularly in children. Dramatic reductions in the rates of Hib disease has been noted in countries which have incorporated this vaccine into their vaccination schedules [1, 2].

Un-encapsulated *H. influenzae*, also termed non-typeable (NTHi) do not possess the CPS and as such cannot be classified by serotyping. The CPS is an important virulence factor of encapsulated strains [3], however a number of adhesive factors, IgA-proteases and cell surface glycolipids have been identified in NTHi that promote colonisation and influence host cell interaction and disease [4, 5]. The possession of such genes can contribute to enhanced virulence in some strains. Many recent studies in particular have observed an increase in the frequency of invasive NTHi cases worldwide [6–12].

Hib is a notifiable disease in Queensland (QLD) and the majority of laboratory confirmed invasive

* Author for correspondence: M. Staples, Queensland Department of Health, Public Health Microbiology, Forensic and Scientific Services, Health Support Queensland, PO Box 594, Archerfield 4108, QLD, Australia.
(Email: megan.staples@health.qld.gov.au)

Hi isolated by hospital and private pathology laboratories in the state are forwarded to the QLD Public Health Microbiology (PHM) reference laboratory for serotyping. A large proportion of referred invasive QLD *H. influenzae* isolates are un-encapsulated however there is a distinct lack of published data on the genetic basis and population structure of both QLD and Australian NTHi strains. The aim of this study was to molecularly characterise referred invasive NTHi isolates in QLD, Australia over a 15-year period, observing the carriage of capsular and putative virulence genes including the major adhesins, antimicrobial resistance genes and population diversity.

METHODS

Specimen selection

Between 2001 and 2015, 745 invasive *H. influenzae* isolates from QLD patients (excluding duplicates) were referred to the QLD PHM laboratory. Invasive disease was defined as the isolation of *H. influenzae* from a normally sterile body site in a patient such as blood or cerebrospinal fluid (CSF). As part of reference laboratory surveillance, demographic information including age, gender and postcode was collected, as well as antimicrobial resistance data as recorded by the submitting laboratory using CLSI (Clinical and Laboratory Standards Institute) guidelines [13] for most isolates. Clinical presentation and disease outcome was not routinely provided to the reference laboratory. Isolate identity was confirmed using the RapID™ NH system (Remel, Kansas, USA). Isolates were classified into serotypes a–f (Hia–Hif) or as NTHi using commercially available sera (Phadebact *Haemophilus* Test; MKL Diagnostics AB, Sollentuna, Sweden and Denka Seiken; Tokyo, Japan). β-lactamase production was determined by the nitrocefin biochemical test (Oxoid, Hampshire, UK).

Further characterisation was performed on a subset of 100 isolates chosen to represent varying serotype, year of isolation, patient age and geographical location based on the total isolates received. NTHi (n = 74) formed the majority of isolates with the remaining 26 isolates representing serotypes a (n = 5), b (n = 5), c (n = 4), d (n = 2), e (n = 5) and f (n = 5).

Additionally, 11 publicly available complete genome sequences from encapsulated and un-encapsulated isolates were downloaded from GenBank in June 2016 for use in this study ((Name; GenBank accession no.): Rd_KW20; NC000907, KR494; CP005967, 86-028NP; NC_007146, PittEE; NC_009566, F3031; NC_014920, F3047; NC_014922, 10810; NC_016809, R2866; NC_017451, R2846; NC_017452, 477; NZ_CP007470 and C486; NZ_CP007471).

Whole-genome sequencing of isolates

Genomic DNA was extracted from the isolates using QIAsymphony AP (Qiagen, København, Denmark) as per the manufacturer's instructions. Genomic DNA was sheared to 300 bp using the Covaris S220 ultrasonicator and libraries were prepared using the Ion Plus Fragment Library kit. Sequencing was performed using an Ion Torrent PGM with the Ion PGM IC 200 kit using 316v2 chips (Life Technologies, Carlsbad, USA) according to manufacturer's instructions. FASTQ reads are located in the European Nucleotide Archive in study PRJEB18702.

Interrogation of assembled sequences for capsular and putative virulence gene detection

Sequences were *de novo* assembled using the Geneious assembler with default settings in Geneious® version 7.1.8 (Biomatters, Auckland, New Zealand) [14]. Assemblies were imported into SeqSphere+ version 3.2.1 (Ridom GmbH, Münster, Germany) and task templates were developed for the detection of targeted putative virulence genes including major proteases and adhesins (*iga, igaB, hmw1A, hmw2A, sodC, hap, hia* and *pepN*), capsule-associated genes (*bexA, bexB, capA, capB, capC, capD, capE* and *capF*), and mobile element IS*1016*. For gene detection, the CDS (complete coding sequences) were downloaded from GenBank using the following reference isolates; 10810 (*bexA, bexB*), Rd_KW20 (*iga, hap, pepN*), R2866 (*hia*), R2846 (*hmw1A, hmw2A*), GenBank accession no. AF549211 (*sodC*), GenBank accession no. NC_022356 (IS*1016* and *capF*), GenBank accession no. Z37516 (*capA*), GenBank accession no. FQ312006 (*capB*), GenBank accession no. HQ651151 (*capC*), GenBank accession no. HQ424464 (*capD*), GenBank accession no. HM053635 (*capE*). Detection of the *hif* gene cluster was achieved using a primer sequence for the conserved *hifBC* region from R2866 [15, 16].

Targets were recorded by SeqSphere+ as present if found in the genome within the parameters of $\geqslant$95% sequence identity and $\geqslant$99% alignment to the reference sequences. For the purposes of this study, only gene detection or non-detection information was recorded and further investigation into the gene functionality

https://doi.org/10.1017/S0950268817000450 Published online by Cambridge University Press

of variants was not performed. The sequences for targets *iga, hap, hia, hmw1A* and *hmw2A* were shown to be highly variable and gave consistently low nucleotide sequence similarity to the references used (<95% identity and <99% alignment) and were recorded as absent by SeqSphere$^+$ and flagged as possible false-negative results. In order to confirm the presence of these targets the reads of all isolates were mapped to the publicly available reference sequences for PittEE, Rd_KW20, R2866 and R2846 using CLC Genomics workbench 8 (Qiagen) which comprised representative nucleotide sequences of the investigated target genes respectively.

Multi locus sequence typing (MLST)

The MLST was determined for each isolate, facilitated using SeqSphere$^+$, which assigned alleles according to the scheme at PubMLST.org. Data were imported into BioNumerics 6.5 software (Applied Maths Inc., Sint-Martens-Latem, Belgium) and a minimum spanning tree (MST) was produced using a categorical coefficient.

Single nucleotide polymorphism (SNP) analysis of isolates

FASTQ files for all isolates from this study, 10 GenBank *Haemophilus influenzae* genome sequences (KR494, 86-028NP, PittEE, F3031, F3047, 10810, R2866, R2846, 477 and C486) and the reference genome, Rd_KW20 were used to produce an alignment of core SNPs using SNIPPY version 3 [17]. A maximum likelihood tree was generated from the SNP alignment, using the Fast Tree and visualised using the Phandango web application (https://github.com/jameshadfield/phandango).

Identification of antimicrobial resistance genes

The presence of acquired resistance genes was investigated for all isolates using the ResFinder tool at CGE (https://cge.cbs.dtu.dk/services/ResFinder-2.1/).

Data analysis

Diversity was calculated using Simpson's index of diversity (SID) [18, 19], with 95% confidence intervals (CI) calculated as previously described [20]. A BioNumerics script available at http://biomath.itqb.unl.pt/ClusterComp [21] was used to facilitate these calculations. Calculation of Yates' χ^2 statistic was calculated as per Preacher [22].

Postcode data were categorised according to QLD Hospital and Health Services (HHS) and was used to assign isolates to one of three regions; North QLD (Torres and Cape, Cairns and Hinterland, North West and Townsville HHS), Central QLD (Central West, Central QLD, Wide Bay and Sunshine Coast HHS) and South East QLD (South West, Darling Downs, West Moreton, Metro North, Metro South and Gold Coast HHS).

RESULTS

Invasive *H. influenzae* isolates referred

A detailed breakdown of data on isolates referred is shown in Table 1. The majority of the 745 invasive Hi isolates referred to the PHM laboratory between 2001 and 2015 were NTHi (72·2%). Of the 207 capsulated Hi, Hib were most common (11·1%), followed by Hif (6·9%), Hia (4·8%) and Hie (4·2%). Of the total referred isolates, 53·4% ($n = 398$) were from female patients and 46·2% ($n = 344$) from males, with three isolates having no gender recorded.

The nitrocefin test for β-lactamase presence was positive for 20·8% of total isolates, a similar rate was observed for NTHi isolates alone (21%). Interestingly, different serotypes exhibited different levels of β-lactamase activity; most notably Hia, Hic and Hid had no β-lactamase-positive isolates, while 30% of Hib were β-lactamase positive.

Location data by postcode were unavailable for 39 isolates. Of the remaining 706 isolates 56·6% were collected from South East QLD, 21·8% from North QLD and 21·6% from Central QLD. When differentiated by serotype, NTHi, Hib, Hie and Hif were predominately from South East QLD, however Hic and Hid were exclusively isolated from North QLD patients, and 65·7% of Hia were also isolated in North QLD.

Overall, patient age for QLD invasive Hi cases ranged from 0 to 99 years of age (median age 42·9 years). In total 23·4% of referred isolates belonged to patients aged $\leqslant 5$ years with 9·8% of these being from patients $\leqslant 1$year of age. Patients in the 61–80 year age group made up the second largest percentage of referred Hi (23·1%). When differentiated by type, Hie and Hif demonstrated similar age breakdowns. Interestingly NTHi were less frequently isolated in children $\leqslant 5$ years (17·7%) and had highest frequency from patients 61–80 years of age (25·1%). Other encapsulated serotypes were isolated more frequently from the $\leqslant 5$ years demographic (Hia 63·9%; Hib 37·4%; and Hic 75%).

https://doi.org/10.1017/S0950268817000450 Published online by Cambridge University Press

Table 1. *Data breakdown for Haemophilus influenzae referrals to PHM laboratory 2001–2015*

	Number of isolates							
	Hia	Hib	Hic	Hid	Hie	Hif	NTHi	Total
Total referrals	36	83	4	2	31	51	538	745
Age range								
⩽1	19	21	1	1	8	8	67	125
⩽5	23	31	3	1	10	11	95	174
6–20	1	17	0	0	1	5	49	73
21–40	5	9	0	0	3	3	78	98
41–60	4	13	1	0	7	7	89	121
61–80	3	11	0	0	6	17	135	172
81+	0	2	0	1	4	8	92	107
Median age	1	14	2·5	44	45	59	50·5	42·9
Location								
South QLD	9	33	0	0	17	31	310	400
Central QLD	3	27	0	0	8	9	105	152
North QLD	23	16	4	2	6	8	95	154
Not recorded	1	7	0	0	0	3	28	39
β-lactamase								
Positive	0	24	0	0	7	10	113	154
Negative	35	56	4	2	24	41	425	587
Not recorded	1	3	0	0	0	0	0	4

An increasing trend was observed for the absolute number of NTHi isolates referred each year, with encapsulated strain numbers remaining fairly constant over the time period. A 13% year-on-year increase was observed for NTHi referral numbers from 2001 to 2014, with a 65·2% increase from 2014 to 2015, however this increase in NTHi as a proportion of all referred invasive Hi from 2001 to 2015 was not statistically significant (63% to 83·5%; $\chi^2_{Yates} = 0{\cdot}42$; $P = 0{\cdot}52$).

Detection of capsule-related genes in sequenced isolates

All encapsulated isolates possessed the expected capsule gene for their known serotype, along with the *bexA* and *bexB* genes, and the insertion sequence IS*1016* (absent in one Hid isolate). The *bexA* and/or *bexB* were not detected in any NTHi isolates, although one NTHi isolate (Hi26) possessed the insertion sequence IS*1016*. Further investigation by ordering and aligning the *de novo* assembled contigs for Hi26 to the Hia reference sequence (GenBank CP017811) using Mauve [23] demonstrated that this isolate closely matched the overall structure and content of the Hia reference, but did not contain the capsular region. IS1016 was present in Hi26 at the location of the deleted capsule locus and so is likely to be associated with the deletion of this capsular region.

Detection of putative virulence genes

Presence/absence reporting of putative virulence genes identified 10 separate gene profiles among the 100 isolates (SID = 0·81; 95% CI 0·77–0·85, data not shown). Table 2 shows the proportion of genes detected across each serotype. Genes *iga, hap* and *pepN* were ubiquitous among all isolates, while all other genes were present in a subset of the isolates. All encapsulated isolates lacked *hmw* genes and possessed *hia.* Encapsulated isolates largely clustered together with similar gene profiles, with Hie and Hif grouping exclusively together in one gene profile that also possessed *sodC*.

The *igaB* was detected only in a proportion of NTHi isolates, which all formed one gene profile also having both *hmw* genes. One NTHi isolate possessed *sodC* in addition to both the *hmw* genes. The *hia* was detected in 65% (13/20) of *hmw*-negative isolates. Only seven isolates lacked *hmw* and *hia*, one of these carried *hif*BC, the other six possessed only *iga, hap* and *pepN*.

MLST of sequenced isolates

MLST identified 62 unique sequence types in total among the 100 isolates (SID = 0·99; 95% CI 0·98–0·99). Overall, eight novel alleles were identified (*frdB*, 176–177; *fucK*, 120; *mdh*, 266; *pgi,* 233, 244–245; *recA*, 146) and 18 novel STs (1462, 1463, 1472, 1508–1511, 1589–1599) were assigned.

https://doi.org/10.1017/S0950268817000450 Published online by Cambridge University Press

Table 2. *Putative virulence genes possessed by serotype*

	Gene target % detected								
Serotype	*igaA*	*igaB*	*hmw1A*	*hmw2A*	*sodC*	*hap*	*hia*	*hifBC*	*pepN*
a	100	0	0	0	0	100	100	0	100
b	100	0	0	0	0	100	100	0	100
c	100	0	0	0	0	100	100	100	100
d	100	0	0	0	0	100	100	0	100
e	100	0	0	0	100	100	100	0	100
f	100	0	0	0	100	100	100	0	100
NTHi	100	21·6	71·6	73	1·4	100	17·6	13·5	100

Encapsulated isolates clustered together and were largely separate from NTHi isolates (Fig. 1). Eight clonal complexes (CCs) were identified, with isolates within each CC differing only by one locus. Five CCs (1–5) comprised NTHi isolates exclusively with the largest CC containing 10 isolates, two CCs were comprised only Hib (CC6) and Hie (CC7) respectively and the final cluster (CC8) contained all five Hia isolates and one NTHi (Hi26). No notable association with patient age, gender or location was demonstrated by cluster analysis.

The largest single ST for QLD invasive NTHi was ST103 ($n = 6$), a further 50 STs were associated with the 67 remaining NTHi isolates, which were dispersed widely across the MST. Only one NTHi (Hi26) shared a ST with Hi isolates (ST56). With this exception, Hi fell into exclusive STs: Hia; STs 56 and 1511, Hib; STs 6 and 118, Hic; STs 19, 1508 and 1598, Hid; ST10, Hie; ST18 and 66 and Hif; STs 124 and 16.

SNP typing

The classification of isolates by SNP groupings correlated well with serotype, MLST ST and the distribution of putative virulence genes (Fig. 2). No notable correlation was observed between SNP groupings and year, age, location or gender (data not shown).

Overall, SNP typing resolved the isolates into three major groups (A, B and C). Group C was a distinct group comprising entirely of Hie and Hif isolates which were more distantly related to the other encapsulated and NTHi isolates. All remaining isolates (with the exception of Hi42) clustered into either group A or B. Group A comprised entirely of NTHi ($n = 33$) while group B included NTHi ($n = 40$) and encapsulated isolates belonging to Hia, Hib, Hic and Hid.

Comparisons between NTHi in groups A and B demonstrated a notable difference; isolates possessing *igaB* belonged exclusively within group A and comprised 48·5% ($n = 16$) of these isolates. Conversely, the proportion of isolates which carried the *hif* gene cluster was higher in group A (18·2%, $n = 6$) than in group B (10%, $n = 4$).

Within these major groups, NTHi isolates were largely heterogeneous by SNP typing, however isolates within the same serotype and MLST ST clustered together. Isolates within these clusters were also observed to contain similar virulence gene profiles. Hi26 was the only NTHi to cluster with encapsulated isolates by SNP analysis, grouping with the Hia ST56 isolates.

SNP typing identified one isolate (Hi42) which failed to cluster with any of the three main SNP groupings. This isolate was characterised by MLST as ST1595, a newly assigned ST with two novel alleles (*frdB* = 177 and *fucK* = 120). Similarly to the other isolates, Hi42 possessed *hap, pepN* and *iga*, however was negative for all other putative virulence genes.

Antimicrobial resistance

Ampicillin-related sensitivity data were available for 92 of the 100 isolates and of these, 23 (25%) demonstrated resistance, all due to β-lactamase production. Only one of these demonstrated resistance to amoxicillin/clavulanic acid. Of these 23 isolates, 87% were NTHi ($n = 20$) with one isolate each of Hib, Hie and Hif. The $bla_{\text{TEM-1}}$ gene was detected exclusively in each of the β-lactamase producers. No β-lactamase-negative ampicillin resistant (BLNAR) isolates were identified. No change in antimicrobial resistance was observed over time.

Additionally, 18·7% of isolates with phenotypic data ($n = 14/75$), all NTHi, recorded resistance to cotrimoxazole. No other significant resistances to antibiotics were observed and ResFinder did not detect the presence of any other acquired resistance genes in addition to $bla_{\text{TEM-1}}$.

https://doi.org/10.1017/S0950268817000450 Published online by Cambridge University Press

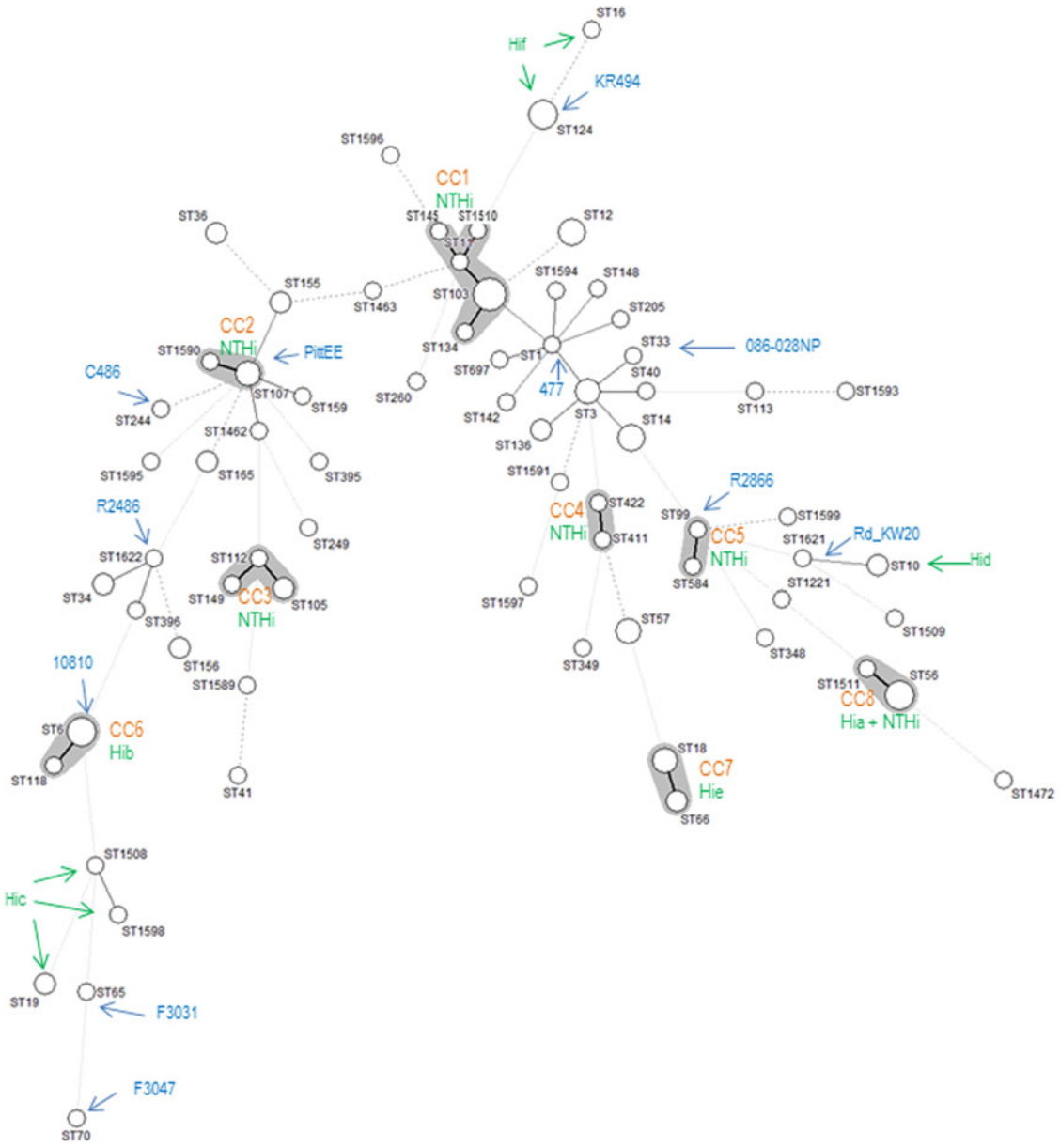

Fig. 1. Minimum spanning tree based on allelic profiles obtained by MLST using a categorical coefficient. The 11 publicly available reference sequences have also been included (labelled in blue). Additionally, serotype data of significant clusters and encapsulated isolates are labelled. Each node represents a different ST, with the size of nodes proportional to the number of included isolates. Thick dark lines separate types differing by a single locus, while thin continuous lines connect double-locus variants. Faint lines represent types differing by three or more loci. Shading around nodes demonstrates partitioning of isolates differing by only one locus, and indicates designated CCs (1–8).

DISCUSSION

Rates of NTHi among invasive Hi isolates

NTHi were the most frequent cause of invasive Hi disease in QLD (72·2%), followed by Hib (11·1%) and Hif (6·9%). Similar observations of invasive Hi proportions have also been made in Spain (NTHi 86·6%) and Portugal (NTHi 77·1%) [6, 24]. Additionally, the yearly increase in the referred numbers of invasive NTHi observed for QLD has also been reported worldwide [7–12]. The highest NTHi rates from QLD patients were observed in the older age categories (82·3% in patients over 5 years of age), which is consistent with some observations [7, 24] though different from a study performed in Arkansas which saw a majority of NTHi invasive cases in children younger than 4 years of age [4]. This

https://doi.org/10.1017/S0950268817000450 Published online by Cambridge University Press

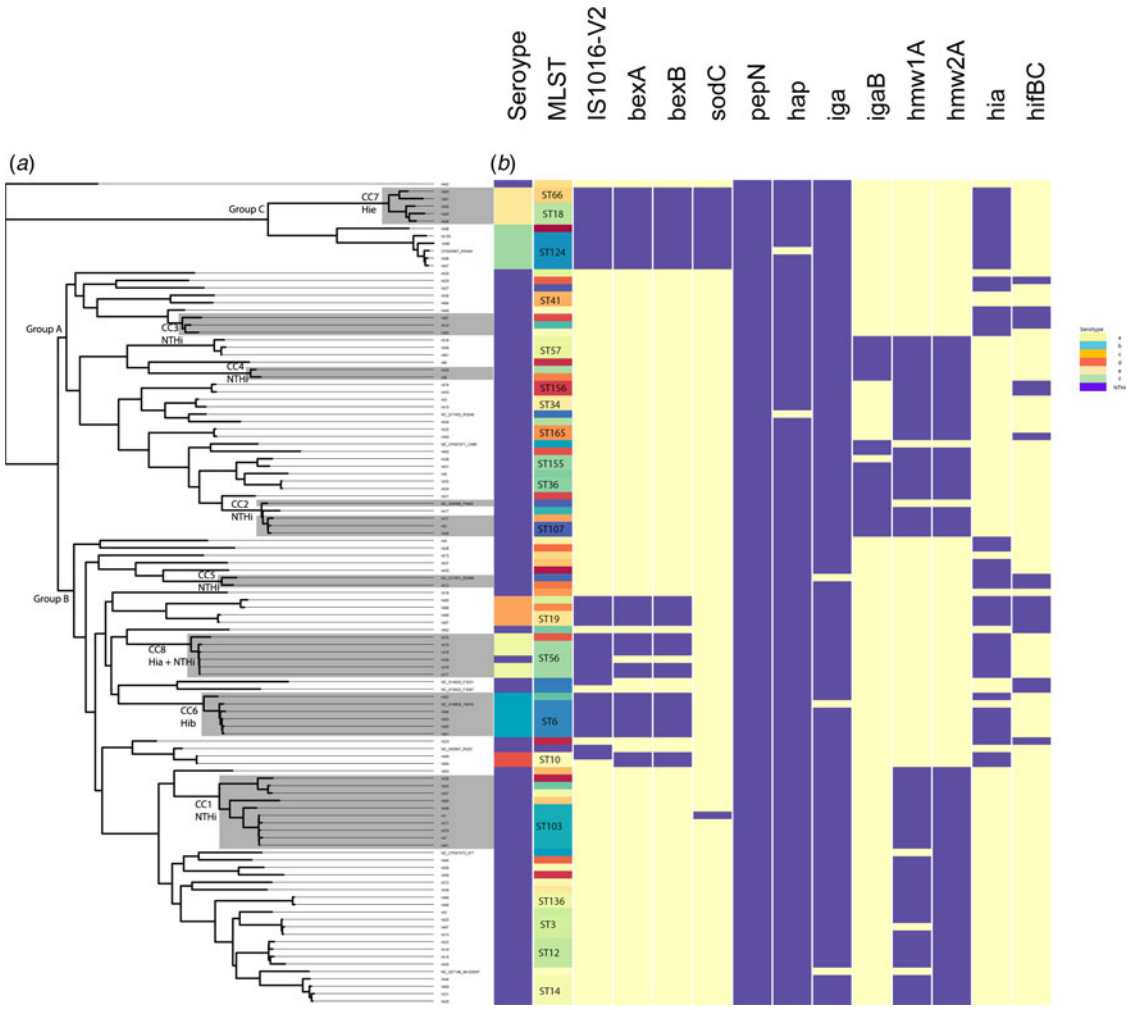

Fig. 2. Genetic relatedness, characteristics and putative virulence gene profile of invasive QLD *H. influenzae* isolates. (*a*) Maximum-likelihood phylogeny based on core-genome SNPs of the 100 isolates, 10 GenBank *Haemophilus influenzae* genome sequences and the reference genome, Rd_KW20. SNP groups are labelled (a–c) and CCs as defined by MLST are shaded. (*b*) Serotype, MLST ST and putative virulence gene profile for each isolate. Serotype and STs are designated individual colours, while for the binary virulence profiles, purple represents present, yellow absent.

high proportion of invasive NTHi infections suggests that the lack of capsule genes does not appear to affect the ability of *H. influenzae* to cause invasive disease and there must be other factors which contribute to invasiveness.

Detection of putative virulence genes

All invasive QLD Hi isolates possessed *pepN* in addition to homologues of the *iga* and *hap* genes. These results are consistent with previous reports [4, 5, 25–27]. Vitovski *et al.* [27] postulated that the presence of Iga1 protease contributes significantly to the pathogenic potential of Hi due to increased prevalence in symptomatic strains over asymptomatic strains, however as only invasive isolates were investigated in this study no further conclusions can be drawn.

The *igaB* gene was observed in 21·6% of invasive NTHi isolates and occurred in 12 different STs (including CCs 2 and 4), all of which shared a consistent gene profile (possession of *iga, igaB, hmw* genes, *hap* and *pepN*). All isolates with *igaB* were clustered within SNP group A. Possession of this second copy of the immunoglobulin A1 protease gene is thought to contribute to enhanced virulence [28].

We saw a high rate of *hmw1A* and *hmw2A* prevalence among NTHi isolates with 71·6% and 73% of isolates possessing these genes respectively, proportions which remained consistent within SNP groups A and B. This is higher than rates observed by Ecevit *et al.*

https://doi.org/10.1017/S0950268817000450 Published online by Cambridge University Press

and O'Neill *et al.* [4, 15] but similar to those in Kostyanev and Sechanova [25], and St. Geme *et al.* [29]. Consistent with other studies, no *hmw* genes were observed in encapsulated isolates, all of which possessed *hia* [15, 29]. Both HMW proteins and the *H. influenzae* adhesin (encoded by *hia*) mediate the attachment of *H. influenzae* to human epithelial cells and they are seldom observed together in *H. influenzae* strains. It postulated that the possession of a particular adhesin could determine genetic origin [29], with possession of the *hia* gene by an NTHi strain being suggestive of evolution from an encapsulated ancestor [30]. In this study, all NTHi possessing *hia* isolates clustered in SNP group C with the encapsulated isolates. A homologue of *hia* was observed in all invasive encapsulated isolates as expected, and in 17·6% of NTHi (65%, $n = 13/20$ of the *hmw*-negative isolates), lower than those observed by Ecevit *et al.* [15] and O'Neill *et al.* [4] but again similar to rates observed in Kostyanev and Sechanova [25] and St. Geme *et al.* [29].

The *hifBC* target representing the pilus gene cluster was detected in 13·5% of NTHi isolates, a similar rate as observed by Ecevit *et al.* [15] but much less than the almost 50% rate as reported by Kostyanev and Sechanova [25]. The pilus cluster was not observed in any of the QLD encapsulated isolates except for Hic, although other studies have also observed these genes in both Hib and Hif [4, 15].

The *sodC* gene was detected in all Hie and Hif isolates and one NTHi (Hi1). This isolate was not associated with Hie or Hif by SNP or MLST; rather it was ST103, part of the largest NTHi CC1, with five other ST103 NTHi and four SLV (single-locus variants) which all additionally possessed *hmw* targets. Additionally, as no capsule related or IS*1016* targets were detected in this isolate; it does not appear to be a capsule-deficient mutant of an encapsulated strain. It is possible that NTHi such as this isolate possess *sodC* through interspecies recombination with *Haemophilus haemolyticus* [16].

In total, six invasive NTHi isolates did not possess *hmw* genes, *hia* or *hifBC*, and we postulate that either Hap or other additional adhesive structures may have facilitated adherence and invasion for these isolates. These isolates did not demonstrate notable genetic relatedness by SNP or MLST.

Detection of capsule-related genes in sequenced isolates

All encapsulated isolates possessed the expected capsule gene for their known serotype, and both *bexA* and *bexB* genes in addition to the insertion sequence IS*1016* (absent in one Hid isolate). No evidence of capsule-specific sequences or *bexA* and/or *bexB* were found within any of the NTHi isolates although Hi26 did possess DNA homologous to the insertion sequence IS*1016* and was ST56, with a sequence type shared with Hia isolates. This isolate also possessed *hia* and was absent for *hmw* genes, similar to other IS*1016*-positive invasive NTHi from previous studies [29, 31]. IS*1016* has been linked with the capacity to cause severe infections [32] and Satola *et al.* [31] also observed a statistically significant association between younger age and the presence of IS*1016* in invasive NTHi. Hi26 was isolated from a 3-year-old patient. Within Hi26, the capsule region has been deleted, possibly by recombination between duplicated IS*1016* copies flanking the capsule locus, but now retains a single IS*1016* copy and is non-encapsulated, similar to the laboratory strain Rd_KW20. Interestingly, a previous study also demonstrated evidence of a capsule-deficient Hia variant isolated from blood at the same time of an encapsulated Hia from CSF in a meningitis patient [33]. In this case it was postulated that the loss of the capsule may have been a result of antibody-driven selective pressure, which may also have been the case for this isolate.

Genetic diversity

Analysis by MLST and SNP typing demonstrated high genetic variability among the QLD NTHi isolates, a trait that has been observed by many other studies [4, 6, 34–36]. No dominant NTHi clone was observed among the QLD isolates, although the most common STs were ST103 ($n = 6$), ST14 ($n = 4$) and ST12 ($n = 4$), with 38 unique STs, many of which have been seen in other NTHi studies [10, 36].

In contrast, the encapsulated isolates largely demonstrated genetic relatedness. The STs associated with our encapsulated isolates have also been reported in the same serotypes disseminated globally [6, 10, 36–39]. No capsular switching events were evident among our isolates; the only NTHi to cluster with the encapsulated strains by MLST was Hi26 (ST56) as discussed previously.

In particular, NTHi isolates belonging within SNP group A appear to have a distinct genotype from the encapsulated isolates, suggesting a more separate population of *H. influenzae* than the group B NTHi isolates, which may be derived from encapsulated strains by the loss of the capsule locus through more recent evolutionary events.

https://doi.org/10.1017/S0950268817000450 Published online by Cambridge University Press

Antimicrobial susceptibility testing

The only mechanism for β-lactam resistance identified in the QLD isolates was β-lactamase production; with resistance rates at 25%, within the 10–25% reported in most regions although well below rates identified in areas of Asia [11]. Consistent with a previous global prevalence report [40] the β-lactamase identified in all QLD isolates was TEM-1. Similarly to the low rates of BLNAR seen in USA and most of Europe, no BLNAR were identified from these invasive QLD isolates [11].

Limitations of study

One limitation to this study is the lack of clinical data available to the laboratory, preventing extrapolation of disease severity to genotype. As only invasive isolates were investigated; further comments cannot be made regarding the sharing of virulence determinants between QLD invasive and non-invasive infections and future studies are warranted regarding this.

CONCLUSION

The total numbers of NTHi identified in invasive Hi disease in QLD is increasing yearly. Analysis of virulence genes could not identify any virulence determinants conclusively associated with invasive NTHi QLD isolates. Furthermore, the heterogeneity of isolates demonstrated that the QLD invasive NTHi population is genetically diverse, indicating that clonal expansion is not the reason for the increase in NTHi invasive infections in QLD. Genotyping did not identify a close relationship with encapsulated *H. influenzae* isolates, with only one NTHi showing evidence of a recent loss of encapsulation. In conclusion, the highly genetically diverse NTHi population identified in QLD offers a challenge to vaccine programmes and demonstrates the importance of maintaining genetic surveillance of *H. influenzae* isolates.

ACKNOWLEDGEMENTS

This work was supported by the Health Support Queensland Research and Development Funding. The authors would like to acknowledge Vicki Hicks and Lawrence Ariotti for their work in the identification and serotyping of referred isolates, and Helen Smith for supervision of the Clinical Reference area and for input into the study design. This publication made use of the *Haemophilus influenzae* MLST website (http://pubmlst.org/hinfluenzae/) sited at the University of Oxford [41]. Special thanks to Roisin Ure for collation of this database and assigning of new alleles and STs. The development of this site has been funded by the Wellcome Trust.

DECLARATION OF INTEREST

None.

REFERENCES

1. **Falla TJ, *et al.*** Population-based study of non-typable *Haemophilus influenzae* invasive disease in children and neonates. *Lancet* 1993; **341**: 851–854.
2. **Jordens JZ, Slack MPE.** *Haemophilus influenzae*: then and now. *European Journal of Clinical Microbiology and Infectious Diseases* 1995; **14**: 935–948.
3. **Gilsdorf JR.** What the pediatrician should know about non-typeable *Haemophilus influenzae*. *Journal of Infection* 2015; **71**: S10–S14.
4. **O'Neill JM, *et al.*** Invasive disease due to nontypeable *Haemophilus influenzae* among children in Arkansas. *Journal of Clinical Microbiology* 2003; **41i**: 3064–3069.
5. **De Chiara M, *et al.*** Genome sequencing of disease and carriage isolates of nontypeable *Haemophilus influenzae* identifies discrete population structure. *Proceedings of the National Academy of Sciences of the United States of America* 2014; **111i**: 5439–5444.
6. **Bajanca-Lavado MP, *et al.*** Characteristics of *Haemophilus influenzae* invasive isolates from Portugal following routine childhood vaccination against *H. influenzae* serotype b (2002–2010). *European Journal of Clinical Microbiology and Infectious Diseases* 2014; **33i**: 603–610.
7. **Dworkin MS, Park L, Borchardt SM.** The changing epidemiology of invasive *Haemophilus influenzae* disease, especially in persons $\geqslant$65 years old. *Clinical Infectious Diseases* 2007; **44i**: 810–816.
8. **Giufre M, *et al.*** Ten years of Hib vaccination in Italy: prevalence of non-encapsulated *Haemophilus influenzae* among invasive isolates and the possible impact on antibiotic resistance. *Vaccine* 2011; **29i**: 3857–3862.
9. **Sarangi J, *et al.*** Invasive *Haemophilus influenzae* disease in adults. *Epidemiology and Infection* 2000; **124**: 441–447.
10. **Sill ML, *et al.*** Population genetics and antibiotic susceptibility of invasive *Haemophilus influenzae* in Manitoba, Canada, from 2000 to 2006. *FEMS Immunology and Medical Microbiology* 2007; **51**: 270–276.
11. **Van Eldere J, *et al.*** Non-typeable *Haemophilus influenzae*, an under-recognised pathogen. *Lancet Infectious Diseases* 2014; **14**: 1281–1292.
12. **Wan Sai Cheong J, *et al.*** Trends in the epidemiology of invasive *Haemophilus influenzae* disease in Queensland, Australia from 2000 to 2013: what is the impact of an increase in invasive non-typable *H. influenzae* (NTHi)? *Epidemiology and Infection* 2015; **12**: 1–8.
13. **Clinical and Laboratory Standards Institute.** *Performance Standards for Antimicrobial Susceptibility Testing: 26th*

Edition. Approved Standard M100-S26. Pennsylvania, USA: Clinical and Laboratory Standards Institute, 2016.
14. **Kearse M, *et al.*** Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* (*Oxford, England*) 2012; **28**: 1647–1649.
15. **Ecevit IZ, *et al.*** Prevalence of the *hifBC, hmw1A, hmw2A, hmwC*, and *hia* Genes in *Haemophilus influenzae* isolates. *Journal of Clinical Microbiology* 2004; **42**: 3065–3072.
16. **Mccrea KW, *et al.*** Prevalence of the *sodC* gene in nontypeable *Haemophilus influenzae* and *Haemophilus haemolyticus* by microarray-based hybridization. *Journal of Clinical Microbiology* 2010; **48**: 714–719.
17. **Seemann T.** Rapid haploid variant calling and core SNP phylogeny software. In. 3 ed, 2016 (https://github.com/tseemann/snippy).
18. **Simpson EH.** Measurement of diversity. *Nature* 1949; **163**: 688–688.
19. **Hunter PR, Gaston MA.** Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *Journal of Clinical Microbiology* 1988; **26**: 2465–2466.
20. **Grundmann H, Hori S, Tanner G.** Determining confidence intervals when measuring genetic diversity and the discriminatory abilities of typing methods for microorganisms. *Journal of Clinical Microbiology* 2001; **39**: 4190–4192.
21. **Carriço JA, *et al.*** Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. *Journal of Clinical Microbiology* 2006; **44**: 2524–2532.
22. **Preacher KJ.** Calculation for the chi-square test: an interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. 2001 (http://quantpsy.org). Last accessed April 2016.
23. **Rissman AI, *et al.*** Reordering contigs of draft genomes using the Mauve Aligner. *Bioinformatics* (*Oxford, England*) 2009; **25**: 2071–2073.
24. **Garcia-Cobos S, *et al.*** Isolates of beta-lactamase-negative ampicillin-resistant *Haemophilus influenzae* causing invasive infections in Spain remain susceptible to cefotaxime and imipenem. *Journal of Antimicrobial Chemotherapy* 2014; **69**: 111–116.
25. **Kostyanev TS, Sechanova LP.** Virulence factors and mechanisms of antibiotic resistance of *Haemophilus influenzae*. *Folia Medica Plovdiv* 2012; **54**: 19–23.
26. **Rodriguez CA, *et al.*** Prevalence and distribution of adhesins in invasive non-type b encapsulated *Haemophilus influenzae*. *Infection and Immunity* 2003; **71**: 1635–1642.
27. **Vitovski S, *et al.*** Nontypeable *Haemophilus influenzae* in carriage and disease: a difference in Iga1 protease activity levels. *JAMA* 2002; **287**: 1699–1705.
28. **Fernaays MM, *et al.*** Characterization of *igaB*, a second immunoglobulin A1 protease gene in nontypeable *Haemophilus influenzae*. *Infection and Immunity* 2006; **74**: 5860–5870.
29. **St. Geme JW, *et al.*** Prevalence and distribution of the *hmw* and *hia* genes and the HMW and Hia adhesins among genetically diverse strains of nontypeable *Haemophilus influenzae*. *Infection and Immunity* 1998; **66**: 364–368.
30. **Krasan GP, *et al.*** Adhesin expression in matched nasopharyngeal and middle ear isolates of nontypeable *Haemophilus influenzae* from children with acute otitis media. *Infection and Immunity* 1999; **67**: 449–454.
31. **Satola SW, Napier B, Farley MM.** Association of IS1016 with the *hia* adhesin gene and biotypes V and I in invasive nontypeable *Haemophilus influenzae*. *Infection and Immunity* 2008; **76**: 5221–5227.
32. **Karlsson EVA, Melhus Å.** Nontypeable *Haemophilus influenzae* strains with the capsule-associated insertion element IS1016 may mimic encapsulated strains. *Acta, Pathologica, Microbiologica, et Immunologica Scandinavica* 2006; **114**: 633–640.
33. **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. *Clinical Microbiology and Infection* 2005; **11**: 637–643.
34. **Musser JM, *et al.*** Genetic relationships of serologically nontypable and serotype b strains of *Haemophilus influenzae*. *Infection and Immunity* 1986; **52**: 183–191.
35. **Shuel M, *et al.*** Invasive *Haemophilus influenzae* in British Columbia: non-Hib and non-typeable strains causing disease in children and adults. *International Journal of Infectious Diseases* 2011; **15**: 167–173.
36. **Puig C, *et al.*** Clinical and molecular epidemiology of *Haemophilus influenzae* causing invasive disease in adult patients. *PLoS ONE* 2014; **9**: e112711.
37. **Ladhani SN, *et al.*** Invasive *Haemophilus influenzae* serotype e and f disease, England and Wales. *Emerging Infectious Diseases* 2012; **18**: 725–732.
38. **Vickers AA, *et al.*** Characterisation of invasive non-type B *Haemophilus influenzae* disease in children in England and Wales, 2003–2010. *Clinical Microbiology and Infection* 2012; **18**: 368–369.
39. **Ulanova M, Tsang RSW.** *Haemophilus influenzae* serotype a as a cause of serious invasive infections. *Lancet Infectious Diseases* 2013; **14**: 70–82.
40. **Farrell DJ, *et al.*** Global distribution of TEM-1 and ROB-1 beta-lactamases in *Haemophilus influenzae*. *Journal of Antimicrobial Chemotherapy* 2005; **56**: 773–776.
41. **Jolley KA, Maiden MCJ.** BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 2010; **11**: 595–595.