Accepted Manuscript

Phylogenetic Relationship of Non-Typeable *Haemophilus influenzae* Isolated in Malaysia

Zaini Mohd-Zain, Nurul H. Kamsani, Norazah Ahmad, Stuart C. Clarke

PII: S1567-1348(15)00393-7
DOI: doi: [10.1016/j.meegid.2015.09.017](https://doi.org/10.1016/j.meegid.2015.09.017)
Reference: MEEGID 2490

To appear in:

Received date: 22 December 2014
Revised date: 9 September 2015
Accepted date: 17 September 2015

Please cite this article as: Mohd-Zain, Zaini, Kamsani, Nurul H., Ahmad, Norazah, Clarke, Stuart C., Phylogenetic Relationship of Non-Typeable *Haemophilus influenzae* Isolated in Malaysia, (2015), doi: [10.1016/j.meegid.2015.09.017](https://doi.org/10.1016/j.meegid.2015.09.017)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Phylogenetic Relationship of Non-Typeable *Haemophilus influenzae* Isolated in Malaysia

Zaini Mohd-Zain*, Nurul H. Kamsani, Norazah Ahmad, Stuart C. Clarke

Zaini Mohd-Zain^{1,2}

¹Core of Health and Wellbeing, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

²Institute of Medical Molecular Biotechnology, Faculty of Medicine, Universiti Teknologi MARA, Sungai Buloh, Selangor, Malaysia

Email: zainimz@salam.uitm.edu.my

Nurul H. Kamsani

Institute of Medical Molecular Biotechnology, Faculty of Medicine, Universiti Teknologi MARA, Sungai Buloh, Selangor, Malaysia

Email: mirakamsani@yahoo.com

Norazah Ahmad

Institute for Medical Research, Kuala Lumpur, Malaysia

Email: norazah@imr.gov.my

Stuart C. Clarke^{1,2,3}

¹Faculty of Medicine and Institute of Life Sciences, University of Southampton, UK

²Southampton NIHR Respiratory Biomedical Research Unit, University Hospital Southampton Foundation NHS Trust, Southampton, UK

³University of Southampton Malaysia Campus, Nusajaya, Johor, Malaysia

Email: S.C.Clarke@soton.ac.uk

*Corresponding author:

Postal address: Faculty of Medicine, Jalan Hospital, 47000 Sungai Buloh, Selangor, Malaysia

Tel.: +6 03 61267650. Fax: +6 03 61267073

E-mail address: zainimz@salam.uitm.edu.my

Abstract

The epidemiology of non-typeable *Haemophilus influenzae* (NTHi) remains poorly understood. We therefore sought to determine the genetic relationship of 25 NTHi isolated from various states in Malaysia using multilocus sequence typing (MLST). The majority of isolates were obtained from sputum. There were 24 novel sequence types (STs). Eight isolates were single-locus variants, the remainder being singletons. Clustering was not based on clinical site of isolation or geographical origin. Despite the limited number of isolates examined in this study, we demonstrate that NTHi isolates in Malaysia are diverse and warrant further investigation.

Keywords: Non-typeable *Haemophilus influenzae*, phylogenetic, Malaysia, MLST

Phylogenetic Relationship of Non-Typeable *Haemophilus influenzae* Isolated in Malaysia

- The phylogenetic relationship of 25 non-typeable *Haemophilus influenzae* (NTHi) isolated in Malaysia was determined by multilocus sequence typing.
- 24 novel sequence types (STs) were obtained.
- Concatenated sequences indicated that the isolates clustered into two major groups.
- Our data shows that NTHi strains in Malaysia are diverse.

1. Introduction

Haemophilus influenzae remains a major cause of meningitis, sepsis and respiratory tract infection despite the introduction of *Haemophilus influenzae* type b (Hib) vaccination programmes (Watt et al., 2009). Non-Hib and non-typeable *H. influenzae* (NTHi) have traditionally received little attention since they were considered to cause less severe infections. Since Hib vaccination has led to a decreased incidence of infection, interest in non-Hib and NTHi infection has increased (Agrawal and Murphy, 2011; Puig et al., 2014). NTHi are commonly associated with mucosal infections including otitis media, conjunctivitis, chronic bronchitis and community-acquired pneumonia, particularly in children (Murphy et al., 2009). Recent increases in invasive NTHi infections have been noted in Sweden, Norway and the USA over the past few years which could be a cause of concern (Van Eldere et al., 2014). In Malaysia, our understanding of the epidemiology of *Haemophilus* infection is currently limited to a small collection of Hib isolates and another collection of *H. influenzae* which were mostly NTHi (Mohd-Zain et al., 2012).

Typing of NTHi is essential so that the epidemiology of this pathogen and the emergence of strains can be monitored. Multilocus sequence typing (MLST), a method based on the sequencing of house-keeping genes, has been successfully applied to the molecular characterisation of NTHi (Meats et al., 2003). MLST may be used to understand the evolutionary relationships between different bacterial strains, and has the important advantage that data can be compared from different laboratories. We sought to describe the diversity and phylogenetic relationship of historical Malaysian NTHi isolates and examine if they are descendents from strains from other geographical locations.

2. Materials and methods

2.1 Bacterial strains

Twenty-five NTHi isolates were obtained from the Institute for Medical Research (IMR) in Kuala Lumpur, Malaysia (Table 1). IMR is the national laboratory for *H. influenzae* isolated from hospital patients throughout Malaysia (Mohd-Zain et al., 2012). The isolates were from sputum (20), eye swab (1), nasal swab (1), high vaginal swab (2) and one isolate of unknown body site. All isolates were identified as *H. influenzae* using a 16S rRNA polymerase chain reaction (PCR) method (Quentin et al., 1996). They were confirmed to be non-typeable when they failed to react to any of the specific antisera of groups a, b, c, d, e and f antigens used in a standard slide agglutination test (Difco, USA).

2.2 Bacterial growth and preparation of template DNA

All NTHi were grown on chocolate agar at 37°C overnight in the presence of 5% CO₂. A single colony was resuspended in distilled water and heated at 100°C for 10 min. Following centrifugation for 1 min at 14,000 rpm (Eppendorf, Germany), the supernatant was collected and used as DNA template for PCR.

2.3 Multi-locus sequence typing

MLST genes (*atpG*, *frdB*, *pgi*, *adk*, *mdh*, *fucK* and *recA*) were amplified as described previously (Meat et al., 2003). PCR products were purified using MinElute PCR purification kit (Qiagen, Germany) according to the protocol provided by the manufacturer. Purified amplified products were sent to First Base Laboratory, Selangor, Malaysia for DNA sequencing.

2.4 Data analysis

Nucleotide sequences were edited and aligned using BioNumerics 4.0 software (Applied Maths, USA) before they were submitted to the public MLST database website (<http://www.mlst.net>) for assignment of sequence types (ST's). Data from this study were deposited in the *H. influenzae* database. A dendrogram was constructed with MEGA5 software using the unweighted pair group method with arithmetic mean (UPGMA) (Tamura et al., 2011). eBURST version 3 (<http://www.mlst.net>) was used to identify exclusive groups of related genotypes in the population.

3. Results

3.1 Sequence types and phylogenetic relationships of the NTHi isolates

Twenty-five different STs were obtained and, with the exception of isolate H253 (ST388), all were novel due to the occurrence of a new allele or combination of alleles (Table 1). The isolates clustered into two major groups (I and II), twenty in group I and five in group II.

3.2 Analysis of the MLST data using eBURST

eBURST analysis indicated that the Malaysian isolates were diverse by comparing the 25 isolates against the entire *H. influenzae* MLST database (Figure 1). There were eight single-locus variants (SLV's), the remainder being singletons. Isolates H603 and 607 (ST's 1269/1270) and isolates H222 and 223 (ST's 1278/1279) were SLV's of each other, both pairs of isolates having different *frdB* alleles.

3.3 Analysis of housekeeping genes from the NTHi isolates

Multiple alignment of each of the seven housekeeping genes indicated that, with exception of the *adk* gene, the nucleotide sequences had 90-99% similarity. For the *adk* gene, nucleotide sequences were more divergent although 22 isolates fell into one major branch of the tree (Figure 2). The *adk* sequence of H209 had 94% similarity to the other 22 isolates whereas those of H220 and H226 had only 86.6-88.1% similarity.

4. Discussion

In an effort to control infections caused by NTHi, improving our understanding of the epidemiology of the bacterium is essential. Since the decline of Hib in numerous countries worldwide, NTHi is now considered a major cause of respiratory infections. MLST is one of the main molecular methods used for epidemiological studies of infectious diseases. In this study, allelic profiles of the NTHi isolates produced 24 novel STs. However, concatenation of the seven housekeeping gene sequences revealed two phylogenetic groups (I and II), suggesting that they are genetically divergent, although the majority of isolates fell into group I. Our data supports previous findings that NTHi isolates are diverse and show less evidence of clonality compared to encapsulated Hi (Meats et al., 2003; Musser et al., 1986).

As also observed by Erwin et al. (2008), there was no correlation between the clinical source and phylogenetic clustering. We were unable to distinguish other characteristics that may have divided them into two groups as demographic details were not available. It has been previously reported that asymptomatic NTHi that colonise the nasopharynx are not phylogenetically distinct from those strains that caused acute otitis media although they have different STs (Kaur et al., 2011). Conversely, LaCross et al. (2013) examined 170 NTHi and showed that commensal NTHi were genetically separated from NTHi associated with otitis media. Our study was limited to only clinical isolates but our observations are consistent with the recent finding that NTHi occur as distinct evolutionary clades (De Chiara et al., 2014). Clustering of the strains is likely to be due to small numbers of mutations that have occurred in the housekeeping genes. Our data indicates that *adk* was most diverse compared to the other six housekeeping genes. Some *H. influenzae* strains lack the fucose operon (*fucK*) which may make them untypeable by MLST (Ridderberg et al., 2010, Shuel et al., 2011). All isolates in our study possessed the *fucK* gene which enabled us to type them by MLST.

The present study provides further evidence on the utility of MLST in assessing diversity and the relationships between NTHi. Our study is limited by the small number of samples that were isolated in the 1990s in some Malaysian states. Sampling was also biased towards non-invasive bacterial isolates, mostly from sputum, although these are also underrepresented in the MLST database. However, this is the first report of sequence typing of NTHi in Malaysia. As it highlights the diversity of NTHi even in a limited collection of isolates, further studies are required to further support our understanding of NTHi infections in Malaysia and elsewhere.

Acknowledgements

This work was supported by the Excellence Fund of Universiti Teknologi MARA, Malaysia (600-RMI/ST/Dana5/3 DST 332/2011). This publication made use of the *Haemophilus influenzae* MLST website (<http://pubmlst.org/hinfluenzae/>) sited at the University of Oxford (Jolley & Maiden 2010, *BMC Bioinformatics*, **11**:595). The development of this site has been funded by the Wellcome Trust.

References

- Agrawal, A., Murphy, T.F., 2011. *Haemophilus influenzae* infections in the *H. influenzae* type b conjugate vaccine era. *J. Clin. Microbiol.* **49**, 3728-3732.
- De Chiara, M., Hood, D., Muzzi, A., Pickard, D.J., Perkins, T., Pizza, M., Dougan, G., Rappuoli, R., Moxon, E.R., Soriani, M., Donati, C., 2014. Genome sequencing of

- disease and carriage isolates of nontypeable *Haemophilus influenzae* identifies discrete population structure. Proc. Nat. Acad. Sci. www.pnas.org/cgi/doi/10.1073/pnas.14033-53111. Pg 1-6.
- Erwin, A., Sandstedt, S.A., Bonthuis, P.J., Geelhood, J.L., Nelson, K.L., Unrath, W.C.T., Diggle, M.A., Theodore, M.J., Pleatman, C.R., Mothershed, E.A., Sacchi, C.T., Mayer, L.W., Gilsdorf, J.R., Smith, A.L., 2008. Analysis of genetic relatedness of *Haemophilus influenzae* isolated by multilocus sequence typing. J. Bact. 190, 1473-1483.
- Kaur, R., Chang, A., Xu, Q., Casey, J.R., Pichichero, M.E., 2011. Phylogenetic relatedness and diversity of non-typeable *Haemophilus influenzae* in the nasopharynx and middle ear fluid of children with acute otitis media. Pediatr. Infect. Dis. J. 31, 145-9.
- LaCross, N.C., Marrs, C.F., Patel, M., Sandstedt, S.A., Gilsdorf, J.R., 2008. High genetic diversity of nontypeable *Haemophilus influenzae* isolated from two children attending a day care centre. J. Clin. Microbiol. 46, 3817-3821.
- LaCross, N.C., Marrs, C.F., Gilsdorf, J.R., 2013. Population structure in nontypeable *Haemophilus influenzae*. Infect. Genet. Evol. 14, 125-136.
- Meats, E., Feil, E.J., Stringer, S., Cody, A.J., Golstein, R., Kroll, J.S., Popovic, T., Spratt, B.G., 2003. Characterization of encapsulated and noncapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing. J. Clinical Microbiol. 41, 1623-1636.
- Musser, J.M., Barenkamp, S.J., Granoff, D.M., Selander, R.K., 1986. Genetic relationships of serologically non-typeable and serotype b strains of *Haemophilus influenzae*. Infect. Immun. 52, 183-191.
- Mohd-Zain, Z., Kamsani, N.H., Ismail, I.S., Ahmad, N., 2012. Antibiotic susceptibility profile of *Haemophilus influenzae* and transfer of co-trimoxazole resistance determinants. Trop. Biomed. 29, 372-380.
- Murphy, T.F., Faden, H., Bakaletz, L.O., Kyd, J.M., Forsgren, A., Campos, J., Virji, M., Pelton, S.I., 2009. Nontypeable *Haemophilus influenzae* as a Pathogen in Children. Pediatr. Infect. Dis. J. 28, 43-48.
- Puig, C., Grau, I., Marti, S., Tubau, F., Calatayud, K., Pallares, R., Liñares, J., Ardanuv, C., 2014. Clinical and molecular epidemiology of *Haemophilus influenzae* causing invasive disease in adult patients. PLOS One. 9, e112711. doi:10.1371/journal.pne.0112711.
- Quentin, R., Ruimy, R., Rosenau, A., Musser, J.M., Christen, R., 1996. Genetic identification of cryptic genospecies of *Haemophilus* causing urogenital and neonatal infections by PCR using specific primers targeting genes coding for 16S rRNA. J. Clin. Microbiol. 34, 1380-1385.
- Ridderberg, W., Fenger, M.G., Nerskov-Lauritsen, N., 2010. *Haemophilus influenzae* may be untypeable by the multilocus sequence typing scheme due to a complete deletion of the fucose operon. J. Med. Microbiol. 59, 740-742.
- Shuel, M.L., Karlowsky, K.E., Law, D.K., Tsang, R.S., 2011. Non-capsulated or nontypeable *Haemophilus influenzae* are more likely than their encapsulated or serotypeable counterparts to have mutations in their fucose operon. Can. J. Microbiol. 57, 982-986.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol. Biol. Evol. 28, 2731-2739.
- Van Eldere, J., Slack, M.P.E., Ladhani, S., Cripps, A.W., 2014. Non-typeable *Haemophilus influenzae*, an under-recognised pathogen. The Lancet Infectious Diseases. Lancet. 14, 1281-1292.
- Watt, J.P., Wolfson, L.J., O'Brien, K.I., Henkle, E., Deloria-Knoll, M., McCall, N., Lee, E., Levine, O.S., Hajjeh, R., Mulholland, K., Cherian, T., 2009. Burden of disease caused

by Haemophilus type b in children younger than 5 years: global estimates. Lancet. 394, 903-911.

ACCEPTED MANUSCRIPT

Fig. 1. Analysis of the NTHi isolates allelic profile to a list of all isolates profile in the database using eBURST. STs differing from another ST by only one of seven loci are connected by lines and form a group. All the 25 NTHi isolates examined in this work are shown as magenta spots.

Fig. 2. Dendrogram of nucleotide sequences of the *adk* gene of 25 NTHi isolated in Malaysia

Table 1

Properties and Sequence Type of NTHi Isolates Examined by MLST

Id	Isolates	Specimen	Region	Year	Allelic Profile							ST*	Group
					<i>adk</i>	<i>atpG</i>	<i>frdB</i>	<i>fucK</i>	<i>mdh</i>	<i>pgi</i>	<i>recA</i>		
2117	H151	Eye swab	Sabah	1995	14	3	154	14	238	14	21	1259	I
2142	H152	Sputum	Kuala Lumpur	1995	175	1	156	102	13	1	135	1275	II
2118	H153	Sputum	Kuala Lumpur	1995	52	1	156	14	15	1	3	1260	I
2119	H154	Sputum	Sg. Buloh	1995	63	32	7	18	46	1	29	1253	I
2120	H155	Nasal	Sabah	1995	63	50	22	15	30	1	3	1254	I
2122	H159	Sputum	Sabah	1995	173	33	7	101	11	1	31	1261	I
2123	H209	HVS**	Kuala Lumpur	1995	50	12	32	54	147	1	125	1252	I
2144	H219	Sputum	Penang	1995	172	120	8	2	14	14	136	1277	II
2124	H220	Sputum	Johor Bahru	1995	170	50	152	15	30	1	31	1262	I
2145	H222	Sputum	Penang	1995	171	33	1	99	238	14	31	1278	II
2146	H223	Sputum	Penang	1995	171	33	4	99	238	14	31	1279	II
2125	H224	Sputum	Penang	1995	14	7	13	7	17	14	17	1255	I
2126	H225	Sputum	Penang	1995	28	33	7	1	236	1	48	1263	I
2127	H226	Sputum	Penang	1995	170	50	155	15	30	14	134	1264	I
2128	H228	Sputum	Kota Bharu	1995	33	33	22	15	51	14	3	1258	I
2129	H252	Sputum	Penang	1995	1	115	71	1	31	1	5	1257	I
2130	H253	Sputum	Penang	1995	60	51	16	48	15	1	31	388	I
2131	H254	Unknown	Unknown	1995	60	51	16	48	15	14	37	1256	I
2132	H255	Sputum	Penang	1995	33	1	22	100	51	14	125	1265	I
2133	H256	Sputum	Penang	1995	5	115	7	15	237	1	29	1266	I
2134	H300	Sputum	Penang	1995	1	122	1	14	9	14	31	1267	I
2135	H597	Sputum	Kuala Lipis	1996	33	8	155	14	239	14	134	1268	I
2136	H603	Sputum	Kuala Lumpur	1995	174	121	1	14	240	1	8	1269	I
2137	H607	Sputum	Alor Setar	1998	174	121	8	14	240	1	8	1270	I
2143	H210	HVS**	Kuala Lumpur	1995	171	33	153	99	238	1	29	1276	II

* ST = sequence type, **HVS = High vaginal swab. New alleles are indicated in bold typeface.