Title	Comparative genome analysis identifies novel nucleic acid diagnostic targets for use in the specific detection of Haemophilus influenzae
Author(s)	Coughlan, Helena; Reddington, Kate; Tuite, Nina; Wee Boo, Teck; Cormican, Martin; Barrett, Louise; Smith, Terry; Clancy, Eoin; Barry, Thomas
Publication Date	2015-06-23
Publication Information	Helena Coughlan, Kate Reddington, Nina Tuite, Teck Wee Boo, Martin Cormican, Louise Barrett, Terry J Smith, Eoin Clancy, Thomas Barry (2015) 'Comparative genome analysis identifies novel nucleic acid diagnostic targets for use in the specific detection of Haemophilus influenzae'. Diagnostic Microbiology And Infectious Disease, 83 (2):112-116.
Publisher	Elsevier
Link to publisher's version	http://www.sciencedirect.com/science/article/pii/S0732889315002163
Item record	http://hdl.handle.net/10379/5684
DOI	http://dx.doi.org/10.1016/j.diagmicrobio.2015.06.013

Comparative genome analysis identifies novel nucleic acid diagnostics targets for use in the specific detection of *Haemophilus influenzae*

Running Title: *Haemophilus influenzae* real-time PCR

Abstract Word count: 149

Body of Text Word count: 2, 853

Helena Coughlan[1+2] , Kate Reddington[1], Nina Tuite[1], Teck Wee Boo[3], Martin Cormican[4], Louise Barrett[5], Terry J. Smith[2,6], Eoin Clancy [2, 6], Thomas Barry[1#]

[1] Nucleic Acid Diagnostics Research Laboratory (NADRL), Microbiology, School of Natural Sciences, National University of Ireland, Galway, Ireland

[2]Biomedical Diagnostics Institute Programme, National Centre for Biomedical Engineering Science, National University of Ireland, Galway, Ireland

[3]Department of Clinical Microbiology, University College Hospital, Galway, Ireland

[4]School of Medicine, National University of Ireland, Galway, Ireland

[5]National Centre of Sensor Research, School of Physical Sciences, Biomedical Diagnostics Institute, Dublin City University, Glasnevin, Dublin 9, Ireland

[6] Molecular Diagnostics Research Group (MDRG), School of Natural Sciences, National University of Ireland, Galway, Ireland

Corresponding Author. Mailing Address: Microbiology, School of Natural Sciences, National University of Ireland, Galway, University Road, Co. Galway, Ireland. Telephone: +35391493189, Fax: +35391494598

Abstract

Haemophilus influenzae is recognized as an important human pathogen associated with invasive infections, including blood stream infection and meningitis. Currently used molecular based diagnostics assays lack specificity in correctly detecting and identifying *H. influenzae*. As such, there is a need to develop novel diagnostics assays for the specific identification of *H. influenzae*.

Whole genome comparative analysis was performed to identify putative diagnostics targets which are unique in nucleotide sequence to *H. influenzae*. From this analysis we identified two *H. influenzae* putative diagnostics targets, *phoB* and *pstA*, for use in real-time PCR diagnostics assays. Real-time PCR diagnostics assays using these targets were designed and optimised to specifically detect and identify all 55 *H. influenzae* strains tested. These novel rapid assays can be applied to the specific detection and identification of *H. influenzae,* for use in epidemiological studies and could also enable improved monitoring of invasive disease caused by these bacteria.

Keywords: *H. influenzae*, genome comparison, real-time PCR, *phoB*, *pstA*

1. Introduction

Haemophilus influenzae is the most pathogenic of the 8 *Haemophilus* species which reside as commensals in the human respiratory tract (1). *H. influenzae* strains are categorised based on the presence (encapsulated) or absence (unencapsulated) of a polysaccharide capsule. Encapsulated strains can be further categorised based on their distinct capsular antigens (typeable; serotypes a - f) whilst unencapsulated strains are referred to as Non-Typeable *H. influenzae* (NTHi) (2). In the pre-vaccine era, the encapsulated *H. influenzae* serotype b (Hib), was the primary cause of invasive disease such as meningitis, bacteraemia and pneumonia (3). Upwards of 95% reduction in the number of *H. influenzae* meningitis cases

have been recorded since the implementation of the Hib conjugate vaccine (4). Whilst *H. influenzae* invasive infection in developed countries has declined, NTHi are now the most common cause of *H. influenzae* invasive infections (3, 5-7). In addition, NTHi have replaced type b strains as the most common blood stream isolates (8).

Haemophilus haemolyticus is also a human commensal bacterium that colonizes the respiratory tract and is closely related to *H. influenzae*. Both *H. influenzae* and *H. haemolyticus* require hemin (X factor) and NAD (V factor) for growth. Phenotypic methods of differentiating *H. haemolyticus* from NTHi rely on the ability of *H. haemolyticus* to lyse horse red blood cells (9). However, this haemolysis may be lost after subculture (10, 11) and as a consequence, non-hemolytic *H. haemolyticus* have been misidentified as *H. influenzae* (12). *H. haemolyticus* was considered a rare pathogen (8), however, more recently, a number of cases of invasive disease originally attributed to NTHi have been confirmed as non-haemolytic *H. haemolyticus* (13-15). Phenotypic techniques cannot conclusively differentiate *H. influenzae* from *H. haemolyticus* and other *Haemophilus* species and therefore, alternative molecular techniques must be used to do so (16).

Real-time PCR diagnostics assays for the identification of *H. influenzae* have been developed targeting various genes including *bexA* (17, 18), *ompP2* (19-22), *ompP6* (17, 23, 24), *16S rDNA* (12), *licA* (20), *rnpB* (23), *frdB* (25), *iga* (26) *hpd* (21, 27, 28) and *fucK* (20, 29, 30). Diagnostics assays targeting the *fucK* and *hpd* genes have been established as superior for the specific detection *of H. influenzae* over other routinely used gene targets (20, 31, 32). However, some NTHi strains have been reported as lacking the *fucK* (33-36) or *hpd* genes (37). As such published literature indicates that no one molecular marker can unequivocally differentiate NTHi from *H. haemolyticus* (31, 32).

In this study, we used comparative genome analysis to identify novel diagnostics targets, which could then be used to develop *H. influenzae* specific real-time PCR. Two real-time

PCR assays targeting these novel diagnostics targets were developed and their performance determined. To further validate the real-time PCR diagnostics assays developed, a panel of culture positive clinical isolates, that were identified as *H. Influenzae* using MALDI-TOF MS, were blindly tested. The results of MALI TOF MS and the real-time PCR diagnostics assays were 100% concordant.

2. Materials and Methods

2.1. Diagnostics Target Identification

Publically available whole genome sequences for *H. influenzae* and *H. haemolyticus* strains were analyzed to identify novel diagnostics targets unique to *H. influenzae* and absent from *H. haemolyticus* and other *Haemophilus* species. *H. influenzae* whole genome sequences are publically available however, *H. haemolyticus* genome sequences (n = 6) are only available as contiguous sequences (contigs) from the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov). For each of the *H. haemolyticus* strains (M19501, M19107, M21127, M21621, M21639 and HK386), contigs were ordered and assembled using the Mauve contig mover tool in the Mauve genome alignment software (http://gel.ahabs.wisc.edu/mauve/). The European Molecular Biology Open Software Suite tool Emboss union (http://emboss.us.es/cgi-bin/emboss/union) was used to concatenate the multifasta Mauve output to create a single FASTA formatted file. Whole genome annotation was then carried out using the online genome annotation service RAST (Rapid Annotation using Subsystem Technology; http://rast.nmpdr.org/) and the resulting data downloaded in EMBL format. Potential diagnostics targets within *H. influenzae* were then identified using WebACT (http://www.webact.org/WebACT/home) by aligning the whole genome sequence of three *H. haemolyticus* strains (M19501, M19107 and HK386) and the whole genome sequence of their most closely related *H. influenzae* strain (R2866: Accession number

NC_017451). Potential diagnostic targets were further analysed using microbial nucleotide BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=MicrobialGenomes) to search for similarities in closely related species.

2.2. Bacterial strains, Culture media and Growth conditions

A panel of *H. influenzae* (n=13), and closely related species (n=33) were obtained from various culture collections (Table 1). All strains were cultured under microaerophillic conditions in haemophilus test media broth, chocolate broth or on Columbia chocolate agar plates at 37^{o}C until sufficient growth was observed.

Table 1: Bacterial species and strains included in this study

Organism	**Strain** [c]/ Source	**Assay**[d]		
		phoB	*pstA*	*fucK*
H. influenzae reference strains				
H. influenzae type a	NCTC 8465	+	+	+
H. influenzae type b	DSMZ 23393[c]	+	+	+
H. influenzae type b	DSMZ 11969[c]	+	+	+
H. influenzae type b	DSMZ 11970[c]	+	+	+
H. influenzae type b	DSMZ 10001[c]	+	+	+
H. influenzae type b	DSMZ 4690[c]	+	+	+
H. influenzae type c	NCTC 8469	+	+	+
H. influenzae type d	DSMZ 11121[c]	+	+	+
H. influenzae type e	NCTC 8472	+	+	+
H. influenzae type f	DSMZ 10000[c]	+	+	+
H. influenzae NTHi	DSMZ 9999[c]	+	+	+
H. influenzae NTHi	CCUG 58365	+	+	-
H. influenzae biogroup *aegyptius*	DSMZ 21187[c]	+	+	+
Clinical isolates (n=6)[a]	Blood	+	+	+
Clinical isolates (n=12)[a]	Ear/Eye swab	+	+	+
Clinical isolates (n=24)[a]	Sputum	+	+	+ (21)[e]
Non- *H. influenzae* strains				
H. haemolyticus	CDC-M19501	-	-	-

H. haemolyticus	CDC-M21127	-	-	-
H. haemolyticus	CDC--M21621	-	-	-
H. haemolyticus	CCUG 24149	-	-	-
H. haemolyticus	CCUG 36015	-	-	-
H. haemolyticus	CCUG 36016	-	-	-
H. haemolyticus	CCUG 15642	-	-	-
H. haemolyticus	CCUG 15312	-	-	-
H. haemolyticus	NCTC 10839	-	-	-
H. haemolyticus	NCTC 10659	-	-	-
H. ducreyi	DSMZ 8925	-	-	/
H. felis	DSMZ 21192	-	-	/
H. haemoglobinophilus	DSMZ 21241	-	-	/
H. paracuniculus	DSMZ 21452	-	-	/
H. parahaemolyticus	DSMZ 21417	-	-	/
H. parainfluenzae	DSMZ 8978	-	-	/
H. paraphrohaemolyticus	DSMZ 21451	-	-	/
H. parasuis	DSM 21448	-	-	/
H. pittmnaie	DSMZ 21203	-	-	/
H. pittmnaie	DSMZ 17420	-	-	/
Haemophilus sp.	CCUG 34110	-	-	-
Actinobacillus suis	DSMZ 22433	-	-	/
Actinobacillus pleuropneumonaie	DSMZ 13472	-	-	/
Aggregatibacter aphrophilus	NCTC 11096	-	-	/
Aggregatibacter aphrophilus	NCTC 10558	-	-	/
Aggregatibacter actinomycetecomitans	DSMZ 11121	-	-	/
Aggregatibacter actinomycetecomitans	DSMZ 8324	-	-	/
Aggregatibacter segnis	NCTC 10977	-	-	/
Avibacterium avium	DSMZ 18557	-	-	/
Avibacterium paragallinarum	DSMZ 18554	-	-	/
Gardnerella vaginalis	DSMZ 4944	-	-	/
Histophilus somni	CCUG 36157	-	-	/
Taylorella equigenitalis	DSMZ 10668	-	-	/

[a] All Clinical isolates were identified as *H. influenzae* using MALDI-TOF MS

[b] NCTC = National Collection of Type Cultures; * DSMZ = The German Collection of Microorganisms; *CCUG = Culture Collection, University of Göteborg, Sweden; *CDC = Centre for Disease Control;

[c] *pstA* and *phoB* gene sequence data was generated for each of these *H. influenzae* strains using primers outlined in Table 2

[d] + ve= Positive; - ve= negative; / = not tested for

[e] 3/24 *H. influenzae* clinical isolates not detected by the *fucK* assay

2.3. DNA isolation and Quantification

Genomic DNA from all species was isolated from 1.5 ml of culture using a DNeasy Blood and Tissue Kit as per manufacturers' instructions (Qiagen, Hilden, Germany). DNA concentrations were determined using the Qubit dsDNA HS Assay and the Qubit 1.0 fluorometer (Life Technologies, Carlsbad, CA). Purified DNA samples were stored at -20°C prior to use.

2.4. Conventional PCR and Nucleic Acid Sequencing

Nucleotide sequence data used for real-time PCR diagnostics assay design was recovered from the National Centre for Biotechnology Information (NCBI-http://www.ncbi.nlm.nih.gov/) or was generated in this study. Sequencing primers were designed to amplify 832 bp of the *pstA* gene and 681 bp of the *phoB* gene of *H. influenzae* to identify optimal diagnostics target regions for primer and probe design.

PCR was carried out using the sequencing primers (Table 2) on an iCycler iQ thermal cycler (Bio-Rad Laboratories Inc., CA) to amplify the *pstA* and *phoB* targets in representative culture collection strains of *H. influenzae*. All reactions were performed using the FastStart PCR Master Kit (Roche Diagnostics, Basel, Switzerland) according to manufacturer's instructions. The thermal cycling parameters used consisted of a denaturation cycle at 95^{o}C for 4 min, followed by 35 cycles at 95 ^{o}C (30 s), 50^{o}C (30 s) and 72^{o}C (30 s), and a final elongation cycle at 72 ^{o}C for 7 min. PCR products were purified using the HighPure PCR product purification kit (Roche Diagnostics) and were sequenced externally (Sequiserve, Vaterstetten, Germany).

2.5. PCR primer and probe design

Nucleotide sequences of the diagnostics targets identified in this study were aligned using ClustalW multiple sequence alignment programme (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Oligonucleotide primers and hydrolysis probes were manually designed in accordance with general recommendations and guidelines (38, 39). Species specific oligonucleotide primers and probe were designed for real-time PCR diagnsotics assays. All primers and probe (Table 2) used in this study were obtained from Eurofins MWG-Operon (Ebersberg, Germany).

Table 2: Nucleotide sequences of primers and probes.

Probe/ Primer	**Function**	**DNA sequence (5' – 3')**	**Nucleotide position**	**Genbank accession no.**	**Concentration (µM)**
pstA F1	Sequencing forward primer	CTAATCAAAACTTGCGT	8-24	CP002277.1	0.4 µM
pstA R1	Sequencing reverse primer	TGTTGAAAGAAGAGTCG	823-839	CP002277.1	0.4 µM
phoB F1	Sequencing forward primer	CTGATAGTTGAAGATG	16-31	CP002277.1	0.4 µM
phoB R1	Sequencing reverse primer	TCATTGTTTATCTCGT	681-696	CP002277.1	0.4 µM
pstA P1	*H. influenzae pstA* forward primer	CGTTTCGCACAAATTACC	310-327	CP002277.1	0.5 µM
pstA P2	*H. influenzae pstA* reverse primer	GTGCGTACCACGATAGG	460-476	CP002277.1	0.5 µM
pstA probe	*H. influenzae pstA* hydrolysis probe	FAM-CTGGAGCATTCGCATTAGCTT-BHQ1	428-448	CP002277.1	0.2 µM
phoB P1	*H. influenzae phoB* forward primer	TTGGATTGGATGCTACC	151-167	CP002277.1	0.5 µM
phoB P2	*H. influenzae phoB* reverse primer	AGTGATGTAGTCATCAGC	292-309	CP002277.1	0.5 µM
phoB probe	*H. influenzae phoB* hydrolysis probe	FAM-AGAAAGCTATGCTGCGATTCC-BHQ1	210-230	CP002277.1	0.2 µM
BSS-1F	IAC forward primer	AACGTAGCATTAGCTGC	111-127	HG519928.1	0.5 µM
BSS-1R	IAC reverse primer	CTCATCTTCTTGCCTGC	260-276	HG519928.1	0.5 µM
BSS-P	IAC – specific hydrolysis probe	Cy5-CACATCCAAGTAGGCTACGCT-BHQ2	179-199	HG519928.1	0.2 µM

158

159 **2.6. IAC Development**

160 A non-competitive internal amplification control (IAC) targeting the *ssrA* gene of *Bacillus*
161 *subtillis* subsp. *spizizenii* previously developed in-house was incorporated into the real-time
162 PCR diagnostics assays (Table 2).

163 **2.7. Real-Time PCR**

164 Duplex real-time PCR diagnostics assays were optimised and the analytical specificity and
165 LOD determined. All real-time PCR diagnostics assays were performed on a LightCycler 480
166 (Roche Diagnostics) using the LightCycler 480 probes master kit (Roche Diagnostics). Real-
167 time PCR reactions were performed in a total volume of 20 µl consisting of 2 x probes
168 master, forward and reverse primers (0.5 µM final concentration), probes (0.2 µM final
169 concentration), template DNA (2 µl, 10^4 genome equivalents GE) and IAC DNA (2 µl, 10^3
170 GE). Nuclease free H_2O was added to make up final volume. A no template control was
171 included in each experiment. The cycling parameters consisted of a 10 min denaturation step
172 followed by 50 cycles of 95^{o}C for 10 s, 60^{o}C for 30 s and a final cooling step at 40^{o}C for 10 s.
173 The LightCycler 480 ramp rate was 4.4^{o}C/s while heating and 2.2^{o}C/s on cooling. To prevent
174 fluorescence signal leakage between detection channels on the LightCycler 480, a colour
175 compensation was generated (40).

176 **2.8. Multiplex PCR assay specificity and LOD**

177 To evaluate the specificity of the diagnostics assays developed, genomic DNA from 55 *H.*
178 *influenzae* and 33 closely related *Haemophilus* species (Table 1), tested in triplicate. The
179 LOD of each assay in multiplex format was determined using probit regression analysis in
180 MiniTab (MiniTab Inc., State College, PA) (41). Purified *H. influenzae* genomic DNA
181 (DSM4690) was diluted in nuclease free dH_2O. Twelve replicates of each dilution 15, 10, 8,
182 6, 4, 2, 1 and 0.1 GE were evaluated.

To determine the efficiency of each assay in both monoplex and multiplex format, serial dilutions of purified *H. influenzae* genomic DNA (DSM4690) at 1 x 10^6, 5 x 10^5, 1 x 10^5, 5 x 10^4, 1 x 10^4, 5 x 10^3, 1 x 10^3, 5 x 10^2, 1 x 10^2 and 5 x10^1 were tested in triplicate. Standard curves of C_T versus GE were then constructed. Efficiency (%) was calculated based on the slope of the line using an online qPCR efficiency calculator (http://www.thermoscientificbio.com/webtools/qpcrefficiency/).

2.9. Clinical Isolate Evaluation

H. influenzae clinical isolates (n= 42) were cultured on Columbia chocolate agar plates and were subjected to MALDI-TOF MS analysis (MALDI biotyper, Bruker Daltonics) and the spectra were compared to the software Version 3.4. A number of clinical isolates were typed using conventional PCR (42)(n = 6, isolated from blood; Table 1) and real-time PCR (18) (n = 3, isolated from sputum; Table 1).

2.10. *fucK* Real-Time PCR

H. influenzae culture confirmed clinical isolates (n = 42), *H. influenzae* culture collection isolates (n = 13), *H. haemolyticus* (n = 10) and 1 *Haemophilus* sp. (previously designated *H. haemolyticus*) were also tested in triplicate using a published *fucK* real-time PCR diagnostics assay (20).

3. Results

3.1. Diagnostics Target Identification

Whole genome sequence comparison of *H. influenzae* R2866 and three *H. haemolyticus* strains (M19502, M19107 and HK386) was carried out using WebACT. *H. influenzae* R2866 was used as a reference genome as it was recognized as being the most closely related strain to *H. haemolyticus* based on the RAST comparison analysis. 325 unique regions, comprising of both genes and intergenic nucleotide sequence, were identified that were present in *H.*

influenzae and absent in *H. haemolyticus*. Each of these regions was BLAST analysed and further evaluated using three criteria to determine their suitability as diagnostics targets. Firstly, regions were refined by ruling out putative diagnostics targets that had significant BLAST hits with *H. haemolyticus.* Secondly, based on *in silico* evaluation, any putative diagnostics targets identified must be present in all 22 publically available whole and draft genome sequences of *H. influenzae*. Finally, diagnostics targets identified with significant inter-strain sequence variation were then also excluded.

Based on the above criteria, 6 diagnostic targets (*pstS, pstC, pstA, pstB, phoR* and *phoB*) were identified as having the greatest potential for the development of a *H. influenzae* specific real-time PCR diagnostics assays. Nucleotide sequence data for each of the potential diagnostics targets was retrieved from NCBI and *in silico* analysis was performed to further assess their suitability for *H. influenzae* specific detection. Of these, *pstA* and *phoB* were found to contain the most sequence variation when compared with other non- *H. influenzae* bacterial species and contain the least *H. influenzae* inter-strain variation, and were therefore selected for further analysis.

3.2. Real-time PCR assay specificity

The specificity of each real-time PCR diagnostics assay was confirmed in duplex format using the panel of bacteria listed in Table 1. All culture collection strains and clinical isolates of *H. influenzae* were detected by both assays. None of the non-*H. influenzae* microorganisms tested were detected by the assays (Supplementary Fig. 1A-D).

3.3. Real-time PCR assay Efficiency and LOD

The efficiencies of each assay were evaluated in both monoplex and multiplex real-time PCR format. *H. influenzae* genomic DNA (DSMZ 4690) was quantified and serial dilutions were

prepared based on *H. influenzae* genome size (1,830,137 bp) (43) which equates to 1.9 fg DNA per cell (http://www.changbioscience.com/genetics/mw.html). Both assays were highly efficient in monoplex (*phoB* 94.66%, *pstA* 94.51%). In multiplex, efficiency is comparable in the *pstA* assay (93.47%) and is slightly reduced in the *phoB* assay (88.87%). Plots of C_p versus GE were linear over a range of 1×10^6, 5×10^5, 1×10^5, 5×10^4, 1×10^4, 5×10^3, 1×10^3, 5×10^3, 1×10^3, 5×10^2, 1×10^2, and 5×10^1 GEs for the *phoB* and *pstA* monoplex and *pstA* multiplex assays. (Supplementary Fig. 2A-C). Linearity is decreased at lower concentrations of *H. influenzae* DNA in the *phoB* multiplex assay (Supplementary Fig. 2D). In multiplex format, the LOD was determined using probit regression analysis (41). With 95% confidence, the LOD of the *phoB* and *pstA* assays was 6.49 and 10.49 GE respectively.

3.4. *fucK* real-time PCR assay

A *fucK* real-time PCR diagnostics assay was also used to identify all *H. influenzae* culture collection strains and clinical isolates (Table 1) used in this study. From this analysis, 51/55 of the tested *H. influenzae* strains were detected by the *fucK* assay.

3.5. Clinical Isolate Evaluation

All 42 clinical isolates were screened using MALDI-TOF MS. The MALDI-TOF software identified all strains to species level and confirmed that all clinical isolates are *H. influenzae,* all with scores greater than 2.0. All clinical isolates typed (n=9) using published capsular typing methods were confirmed as NTHi.

4. Discussion

H. influenzae is known to cause a wide spectrum of disease, ranging from non-invasive infections including bronchitis and conjunctivitis to more serious infections including meningitis, pneumonia and sepsis. Bacterial culture is the preferred method for the diagnosis of invasive *H. influenzae* (20). However, culture based methods are time-consuming, often

taking up to 36 hr or more (44-46). Also, in many cases, pre-treatment of the infection with broad spectrum antibiotics may produce false negatives (47). Evaluation of a number of biomarker genes for differentiating *H. influenzae* from *H. haemolyticus* and other non- *H. influenzae* species suggests that there is no one diagnostic target identified to date that can unambiguously differentiate the two microorganisms (31).

Recently, draft whole genome sequences of six *H. haemolyticus* strains have been published (15). In the study presented here, comparative analysis between these strains and *H. influenzae,* highlighted a number of variable regions between the genomes of the two species. BLAST and ClustalW analysis identified 6 genes, *phoB, phoR, pstS, pstC, pstA,* and *pstB,* as potential diagnostics targets for the discrimination of the two microorganisms. *In silico* nucleotide sequence analysis revealed that these genes are present in *H. influenzae* but are absent in all other recognised *Haemophilus* species. These genes, which have been studied extensively in *E. coli* , are members of the bacterial *pho* (phosphate) regulon and are involved in the transport and assimilation of inorganic phosphate (Pi) (48, 49).

Two of the putative diagnostic targets identified namely *phoB* and *pstA* were further evaluated for their ability to accurately identify *H. influenzae* and distinguish it from other closely related *Haemophilus* species. *phoB* was chosen as a diagnostics target as it is only present in gram negative bacteria (with a homolog *phoP* in gram positive bacteria) (48). The *pstA* gene was also chosen as a diagnostics target for further evaluation as it was found to have the least inter-strain gene nucleotide sequence heterogeneity and the greatest level of gene nucleotide sequence variation compared with other bacteria (data not shown). Internally controlled real-time duplex PCR assays were developed for each of these selected diagnostics targets. Both assays were determined to be 100% specific for the detection of *H. influenzae.* The established LOD of 6.49 GE and 10.49 GE (95% confidence), for both assays respectively, is comparable to other studies targeting *fucK,* which have LODs of 10 genome

copies (95% confidence) (20) and between 5 (5/8 reactions) and 50 (8/8 reactions) genomes copies (29). We also compared our assays to a previously developed *fucK* real-time PCR assay as this is one of the diagnostics target most commonly used in the literature for the specific detection of *H. influenzae*. This comparative analysis demonstrated that both the *phoB* and *pstA* assays are more specific for the detection of *H. influenzae* compared to the *fucK* assay. Of the 42 clinical isolates identified as *H. influenzae* using MALDI-TOF MS and the novel real-time PCR diagnostics assays developed in this study, 3 (7.1%) were not detected by the *fucK* assay. One of the culture collection strains (CCUG 58365) was also not detected using the *fucK* assay due to a deletion of the fucose operon as previously reported in this strain (35).

When evaluating the two real-time PCR diagnostics assays developed in this study, their performance in terms of specificity and LOD are comparable. However the *pstA* multiplex real-time PCR assay is more efficient than the *phoB* multiplex real-time PCR assay. In addition, IAC amplification is inhibited at a *H. influenzae* DNA concentration of greater than 10^5 GE in the *phoB* multiplex PCR assay, revealing the *pstA* assay is more robust in terms of overall real-time PCR assay performance.

In this study we set out to identify novel diagnostics targets that could be used in real-time PCR assays to rapidly detect and specifically identify *H. influenzae*. The novel real-time PCR diagnostics assays developed could enable improved monitoring of invasive disease caused by this bacteria while also providing the clinician with valuable information about the optimal therapeutic regimen to initiate patient treatment. We propose that the developed assays now merit extensive multicentre evaluation in clinical settings.

Acknowledgements

This work was supported by Science Foundation Ireland as part of the Biomedical Diagnostics Institute Centre for Science Excellence and Technology (10/CE/B1821). We also wish to thank Leonard W. Mayer and Jennifer Dolan Thomas at the Centers for Disease Control for providing some of the *Haemophilus haemolyticus* strains used in this study.

References

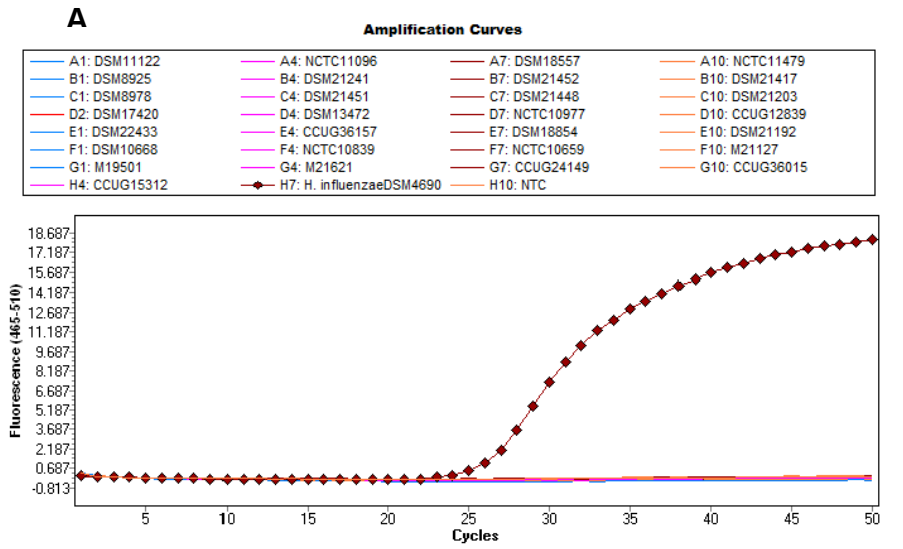
1. **McCrea Kirk W., Xie Jingping, LaCross Nathan, Patel Mayurika, Mukundan Deepa, Murphy Timothy F., Marrs Carl F., Gilsdorf Janet R. 2008. Relationships of nontypeable Haemophilus influenzae strains to hemolytic and nonhemolytic Haemophilus haemolyticus strains. Journal of clinical microbiology 46:406-416.**
2. **Pittman M. 1931. VARIATION AND TYPE SPECIFICITY IN THE BACTERIAL SPECIES HEMOPHILUS INFLUENZAE. The Journal of experimental medicine 53:471-492.**
3. **van Wessel Karen, Rodenburg Gerwin D., Veenhoven Reinier H., Spanjaard Lodewijk, van der Ende Arie, Sanders Elisabeth A. M. 2011. Nontypeable Haemophilus influenzae Invasive Disease in the Netherlands: A Retrospective Surveillance Study 2001-2008. Clinical Infectious Diseases 53:E1-E7.**
4. **Brouwer Matthijs C., Tunkel Allan R., van de Beek Diederik. 2010. Epidemiology, Diagnosis, and Antimicrobial Treatment of Acute Bacterial Meningitis. Clinical Microbiology Reviews 23:467-+.**
5. **MacNeil Jessica R., Cohn Amanda C., Farley Monica, Mair Raydel, Baumbach Joan, Bennett Nancy, Gershman Ken, Harrison Lee H., Lynfield Ruth, Petit Susan, Reingold Arthur, Schaffner William, Thomas Ann, Coronado Fatima, Zell Elizabeth R., Mayer Leonard W., Clark Thomas A., Messonnier Nancy E. 2011. Current Epidemiology and Trends in Invasive Haemophilus influenzae Disease-United States, 1989-2008. Clinical Infectious Diseases 53:1230-1236.**
6. **Dworkin Mark S., Park Lee, Borchardt Stephanie M. 2007. The changing epidemiology of invasive Haemophilus influenzae disease, especially in persons >= 65 years old. Clinical Infectious Diseases 44:810-816.**
7. **Resman F., Ristovski M., Ahl J., Forsgren A., Gilsdorf J. R., Jasir A., Kaijser B., Kronvall G., Riesbeck K. 2011. Invasive disease caused by Haemophilus influenzae in Sweden 1997-2009; evidence of increasing incidence and clinical burden of non-type b strains. Clinical Microbiology and Infection 17:1638-1645.**
8. **Agrawal Aarti, Murphy Timothy F. 2011. Haemophilus influenzae Infections in the H. influenzae Type b Conjugate Vaccine Era. Journal of Clinical Microbiology 49:3728-3732.**
9. **Sandstedt S. A., Zhang L., Patel M., McCrea K. W., Qin Z., Marrs C. F., Gilsdorf J. R. 2008. Comparison of laboratory-based and phylogenetic methods to distinguish between Haemophilus influenzae and H. haemolyticus. J Microbiol Methods 75:369-371.**
10. **Mukundan Deepa, Ecevit Zafer, Patel Mayuri, Marrs Carl F., Gilsdorf Janet R. 2007. Pharyngeal colonization dynamics of Haemophilus influenzae and Haemophilus haemolyticus in healthy adult carriers. Journal of Clinical Microbiology 45:3207-3217.**
11. **Kilian M. 1976. The haemolytic activity of Haemophilus species. Acta pathologica et microbiologica Scandinavica Section B, Microbiology 84B:339-341.**
12. **Murphy Timothy F., Brauer Aimee L., Sethi Sanjay, Kilian Mogens, Cai Xueya, Lesse Alan J. 2007. Haemophilus haemolyticus: A human respiratory tract commensal to be distinguished from Haemophilus influenzae. Journal of Infectious Diseases 195:81-89.**

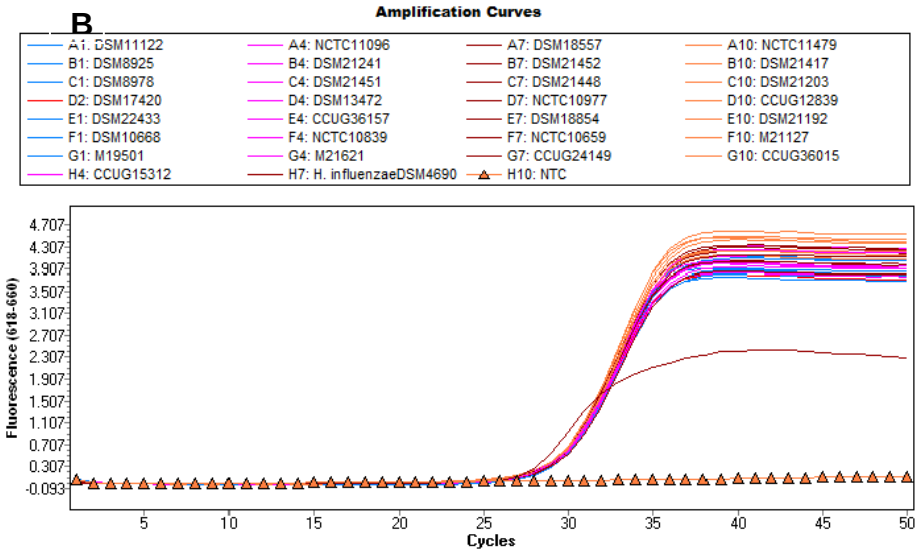
13. Morton Daniel J., Hempel Randy J., Whitby Paul W., Seale Thomas W., Stull Terrence L. 2012. An Invasive Haemophilus haemolyticus Isolate. Journal of Clinical Microbiology 50:1502-1503.
14. Anderson Raydel, Wang Xin, Briere Elizabeth C., Katz Lee S., Cohn Amanda C., Clark Thomas A., Messonnier Nancy E., Mayer Leonard W. 2012. Haemophilus haemolyticus Isolates Causing Clinical Disease. Journal of Clinical Microbiology 50:2462-2465.
15. Jordan I. King, Conley Andrew B., Antonov Ivan V., Arthur Robert A., Cook Erin D., Cooper Guy P., Jones Bernard L., Knipe Kristen M., Lee Kevin J., Liu Xing, Mitchell Gabriel J., Pande Pushkar R., Petit Robert A., Qin Shaopu, Rajan Vani N., Sarda Shruti, Sebastian Aswathy, Tang Shiyuyun, Thapliyal Racchit, Varghese Neha J., Ye Tianjun, Katz Lee S., Wang Xin, Rowe Lori, Frace Michael, Mayer Leonard W. 2011. Genome Sequences for Five Strains of the Emerging Pathogen Haemophilus haemolyticus. Journal of Bacteriology 193:5879-5880.
16. Theodore M. Jordan, Anderson Raydel D., Wang Xin, Katz Lee S., Vuong Jeni T., Bell Melissa E., Juni Billie A., Lowther Sara A., Lynfield Ruth, MacNeil Jessica R., Mayer Leonard W. 2012. Evaluation of new biomarker genes for differentiating Haemophilus influenzae from Haemophilus haemolyticus. Journal of clinical microbiology 50:1422-1424.
17. Vanketel R. J., Dewever B., Vanalphen L. 1990. DETECTION OF HAEMOPHILUS-INFLUENZAE IN CEREBROSPINAL FLUIDS BY POLYMERASE CHAIN-REACTION DNA AMPLIFICATION. Journal of Medical Microbiology 33:271-276.
18. Wroblewski Danielle, Halse Tanya A., Hayes Jill, Kohlerschmidt Donna, Musser Kimberlee A. 2013. Utilization of a real-time PCR approach for Haemophilus influenzae serotype determination as an alternative to the slide agglutination test. Molecular and Cellular Probes 27:86-89.
19. Favaro Marco, Savini Vincenzo, Favalli Cartesio, Fontana Carla. 2013. A Multi-Target Real-Time PCR Assay for Rapid Identification of Meningitis-Associated Microorganisms. Molecular Biotechnology 53:74-79.
20. Meyler Kenneth L., Meehan Mary, Bennett Desiree, Cunney Robert, Cafferkey Mary. 2012. Development of a diagnostic real-time polymerase chain reaction assay for the detection of invasive Haemophilus influenzae in clinical samples. Diagnostic Microbiology and Infectious Disease 74:356-362.
21. Wang Xin, Mair Raydel, Hatcher Cynthia, Theodore M. Jordan, Edmond Karen, Wu Henry M., Harcourt Brian H., Carvalho Maria da Gloria S., Pimenta Fabiana, Nymadawa Pagbajab, Altantsetseg Dorjpurev, Kirsch Mariah, Satola Sarah W., Cohn Amanda, Messonnier Nancy E., Mayer Leonard W. 2011. Detection of bacterial pathogens in Mongolia meningitis surveillance with a new real-time PCR assay to detect Haemophilus influenzae. International Journal of Medical Microbiology 301:303-309.
22. Hobson R. P., Williams A., Rawal K., Pennington T. H., Forbes K. J. 1995. INCIDENCE AND SPREAD OF HAEMOPHILUS-INFLUENZAE ON AN ANTARCTIC BASE DETERMINED USING THE POLYMERASE CHAIN-REACTION. Epidemiology and Infection 114:93-103.
23. Abdeldaim Guma M. K., Stralin Kristoffer, Kirsebom Leif A., Olcen Per, Blomberg Jonas, Hermann Bjorn. 2009. Detection of Haemophilus influenzae in respiratory secretions from pneumonia patients by quantitative real-time polymerase chain reaction. Diagnostic Microbiology and Infectious Disease 64:366-373.
24. Hedberg Sara Thulin, Olcen Per, Fredlund Hans, Molling Paula. 2009. Real-time PCR detection of five prevalent bacteria causing acute meningitis. Apmis 117:856-860.
25. Kunthalert Duangkamol, Henghiranyawong Kritsada, Sistayanarain Anchalee, Khoothiam Krissana. 2013. A single-step polymerase chain reaction for simultaneous detection and differentiation of nontypeable and serotypeable Haemophilus influenzae, Moraxella catarrhalis and Streptococcus pneumoniae. International Journal of Pediatric Otorhinolaryngology 77:275-280.

26. Vitovski S., Dunkin K. T., Howard A. J., Sayers J. R. 2002. Nontypeable Haemophilus influenzae in carriage and disease - A difference in IgA1 protease activity levels. Jama-Journal of the American Medical Association 287:1699-1705.
27. Pickering Janessa, Binks Michael J., Beissbarth Jemima, Hare Kim M., Kirkham Lea-Ann S., Smith-Vaughan Heidi. 2014. A PCR-High-Resolution Melt Assay for Rapid Differentiation of Nontypeable Haemophilus influenzae and Haemophilus haemolyticus. Journal of Clinical Microbiology 52:663-667.
28. Hare K. M., Binks M. J., Grimwood K., Chang A. B., Leach A. J., Smith-Vaughan H. 2012. Culture and PCR Detection of Haemophilus influenzae and Haemophilus haemolyticus in Australian Indigenous Children with Bronchiectasis. Journal of Clinical Microbiology 50:2444-2445.
29. Abdeldaim Guma M. K., Stralin Kristoffer, Olcen Per, Blomberg Jonas, Molling Paula, Herrmann Bjorn. 2013. Quantitative fucK gene polymerase chain reaction on sputum and nasopharyngeal secretions to detect Haemophilus influenzae pneumonia. Diagnostic Microbiology and Infectious Disease 76:141-146.
30. Meats E., Feil E. J., Stringer S., Cody A. J., Goldstein R., Kroll J. S., Popovic T. J., Spratt B. G. 2003. Characterization of encapsulated and noncapsulated Haemophilus influenzae and determination of phylogenetic relationships by multilocus sequence typing. Journal of Clinical Microbiology 41:1623-1636.
31. Binks Michael J., Temple Beth, Kirkham Lea-Ann, Wiertsema Selma P., Dunne Eileen M., Richmond Peter C., Marsh Robyn L., Leach Amanda J., Smith-Vaughan Heidi C. 2012. Molecular Surveillance of True Nontypeable Haemophilus influenzae: An Evaluation of PCR Screening Assays. Plos One 7.
32. Theodore M. Jordan, Anderson Raydel D., Wang Xin, Katz Lee S., Vuong Jeni T., Bell Melissa E., Juni Billie A., Lowther Sara A., Lynfield Ruth, MacNeil Jessica R., Mayer Leonard W. 2012. Evaluation of New Biomarker Genes for Differentiating Haemophilus influenzae from Haemophilus haemolyticus. Journal of Clinical Microbiology 50:1422-1424.
33. Norskov-Lauritsen Niels. 2009. Detection of Cryptic Genospecies Misidentified as Haemophilus influenzae in Routine Clinical Samples by Assessment of Marker Genes fucK, hap, and sodC. Journal of Clinical Microbiology 47:2590-2592.
34. Fenger Mette G., Ridderberg Winnie, Olesen Hanne V., Norskov-Lauritsen Niels. 2012. Low occurrence of 'non-haemolytic Haemophilus haemolyticus' misidentified as Haemophilus influenzae in cystic fibrosis respiratory specimens, and frequent recurrence of persistent H. influenzae clones despite antimicrobial treatment. International Journal of Medical Microbiology 302:315-319.
35. Ridderberg Winnie, Fenger Mette G., Norskov-Lauritsen Niels. 2010. Haemophilus influenzae may be untypable by the multilocus sequence typing scheme due to a complete deletion of the fucose operon. Journal of Medical Microbiology 59:740-742.
36. Shuel Michelle L., Karlowsky Kathleen E., Law Dennis K. S., Tsang Raymond S. W. 2011. Nonencapsulated or nontypeable Haemophilus influenzae are more likely than their encapsulated or serotypeable counterparts to have mutations in their fucose operon. Canadian Journal of Microbiology 57:982-986.
37. Smith-Vaughan Heidi C., Chang Anne B., Sarovich Derek S., Marsh Robyn L., Grimwood Keith, Leach Amanda J., Morris Peter S., Price Erin P. 2014. Absence of an Important Vaccine and Diagnostic Target in Carriage- and Disease-Related Nontypeable Haemophilus influenzae. Clinical and Vaccine Immunology 21:250-252.
38. Dorak M. Tevfik. 2006. Real-Time PCR. http://wwwdorakinfo/genetics/realtimehtml.
39. Robertson JamesM, Walsh-Weller Jennifer. 1998. An Introduction to PCR Primer Design and Optimization of Amplification Reactions, p 121-154. *In* Lincoln P, Thomson J (ed), Forensic DNA Profiling Protocols, vol 98. Humana Press.

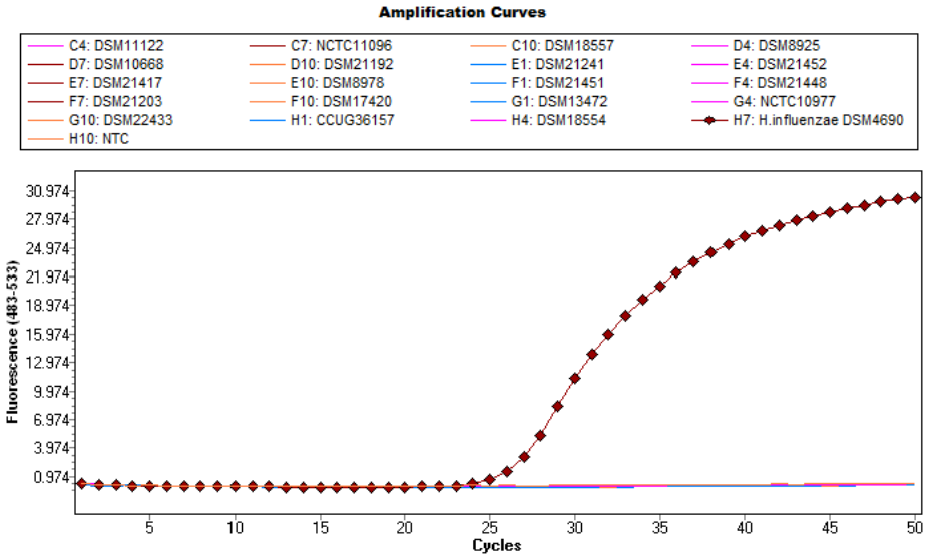
40. Anonymous. 2008. LightCycler 480 instrument operator's manual, p 248 - 256. Roche Diagnostics, Basel, Switzerland.
41. Burd Eileen M. 2010. Validation of Laboratory-Developed Molecular Assays for Infectious Diseases. Clinical Microbiology Reviews 23:550-+.
42. Falla TJ, Crook DW, Brophy LN, Maskell D, Kroll JS, Moxon ER. 1994. PCR for capsular typing of Haemophilus influenzae. Journal of Clinical Microbiology 32:2382-2386.
43. Fleischmann R. D., Adams M. D., White O., Clayton R. A., Kirkness E. F., Kerlavage A. R., Bult C. J., Tomb J. F., Dougherty B. A., Merrick J. M., McKenney K., Sutton G., Fitzhugh W., Fields C., Gocayne J. D., Scott J., Shirley R., Liu L. I., Glodek A., Kelley J. M., Weidman J. F., Phillips C. A., Spriggs T., Hedblom E., Cotton M. D., Utterback T. R., Hanna M. C., Nguyen D. T., Saudek D. M., Brandon R. C., Fine L. D., Fritchman J. L., Fuhrmann J. L., Geoghagen N. S. M., Gnehm C. L., McDonald L. A., Small K. V., Fraser C. M., Smith H. O., Venter J. C. 1995. WHOLE-GENOME RANDOM SEQUENCING AND ASSEMBLY OF HAEMOPHILUS-INFLUENZAE RD. Science 269:496-512.
44. Corless C. E., Guiver M., Borrow R., Edwards-Jones V., Fox A. J., Kaczmarski E. B. 2001. Simultaneous detection of Neisseria meningitidis, Haemophilus influenzae, and Streptococcus pneumoniae in suspected cases of meningitis and septicemia using real-time PCR. Journal of Clinical Microbiology 39:1553-1558.
45. Greenberg-Kushnir Noa, Haskin Orly, Yarden-Bilavsky Havatzelet, Amir Jacob, Bilavsky Efraim. 2012. Haemophilus influenzae Type b Meningitis in the Short Period after Vaccination: A Reminder of the Phenomenon of Apparent Vaccine Failure. Case reports in infectious diseases 2012:950107-950107.
46. NAKHJAVANI F.A., BONAKDAR HASHEMI F., KALANi M.T., KAZEMI B., NOURI K., AZADI N., HAGHI ASHTIANI M., ALIGHOLI M., ERFANI U., ABEDINI M. 2005. DETECTION OF HAEMOPHILUS INFLUENZAE TYPE B IN CEREBROSPINAL FLUID OF SUSPECTED CHILDREN WITH MENINGITIS BYPCR. Medical Journal Of The Islamic Republic Of Iran 19:181-184.
47. Dalton H. P., Allison M. J. 1968. Modification of laboratory results by partial treatment of bacterial meningitis. American Journal of Clinical Pathology 49:410-&.
48. Vershinina O. A., Znamenskaia L. V. 2002. The Pho regulons of bacteria. Mikrobiologiia 71:581-595.
49. Crepin Sebastien, Chekabab Samuel-Mohammed, Le Bihan Guillaume, Bertrand Nicolas, Dozois Charles M., Harel Josee. 2011. The Pho regulon and the pathogenesis of Escherichia coli. Veterinary microbiology 153:82-88.

Supplemental Figures



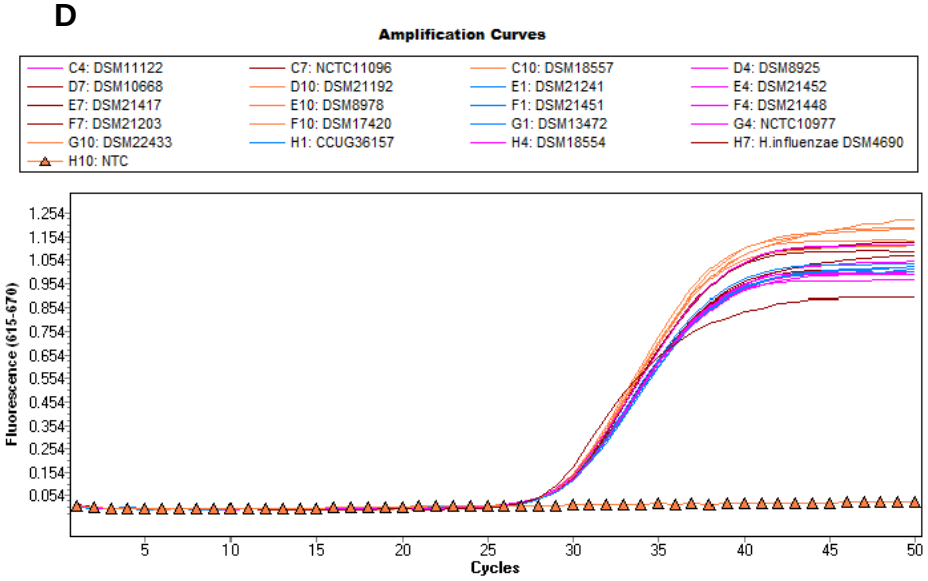


C



508

D



509

510 **Figure 1: Representative amplification curves for multiplex real-time PCR diagnostics**
511 **assay.** Fig. 1A: Real-time amplification curves for 1 strain of *H. influenzae* (DSM4690)
512 (Diamonds) targeting the *phoB* gene in the FAM channel (465 -510 nm). No non-*H.*
513 *influenzae* species were detected by the assay; Fig. 1B: Amplification curves for IAC in Cy5
514 channel (618-660nm) with NTC highlighted with triangles. All samples detected except the
515 no template control (NTC). Fig. 1C: Real-time amplification curves for 1 strain of *H.*
516 *influenzae* (DSM4690) (Diamonds) targeting the *pstA* gene in the FAM channel (483 -533
517 nm). No amplification occurred in non-*H. influenzae* species; Fig. 1D: Real-time
518 amplification curves for IAC in Cy5 channel (615-670nm) with NTC highlighted with
519 triangles. All samples detected except the no template control (NTC).
520

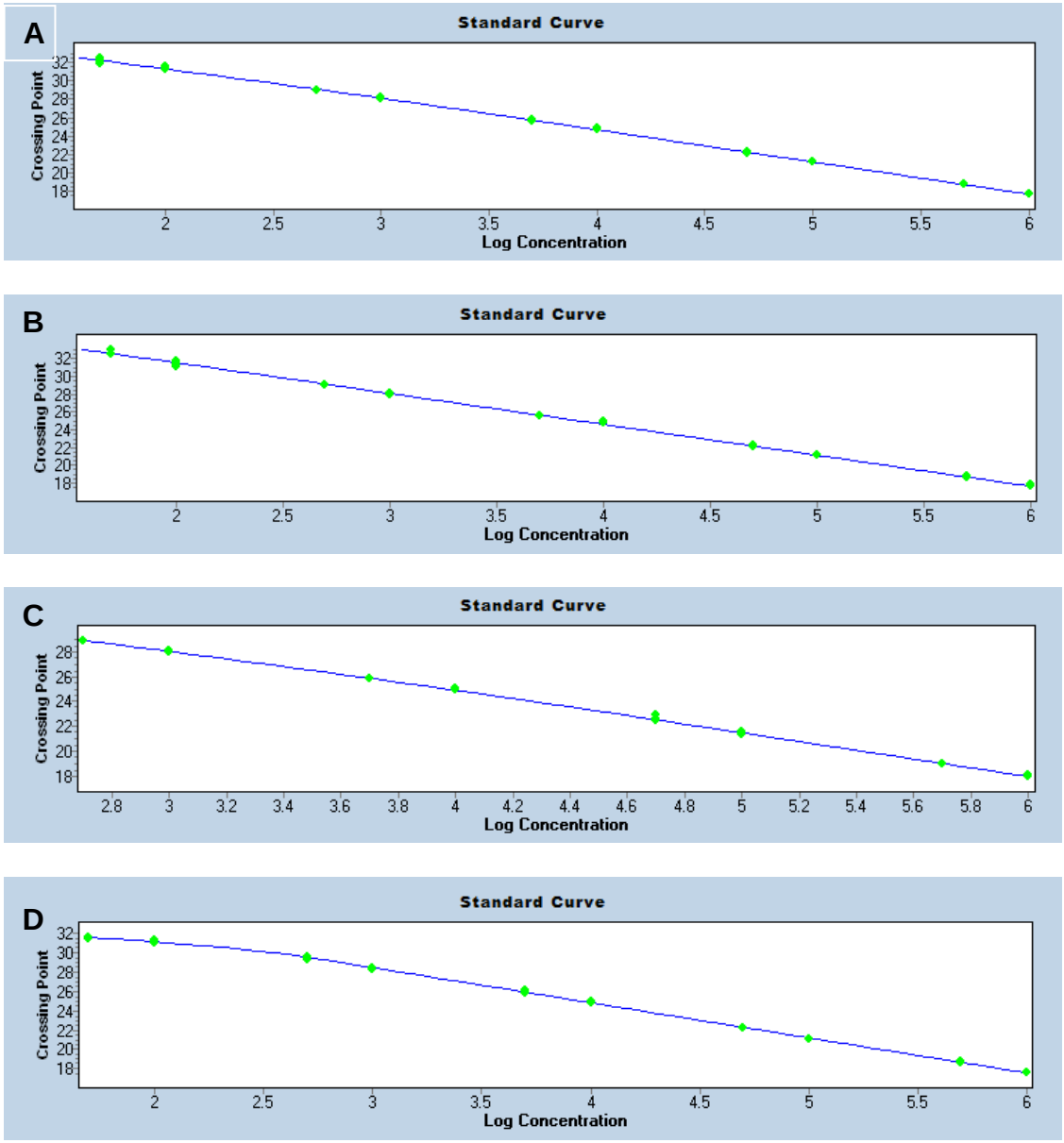


Figure 2: Standard curve of Log10 GE *H. influenzae* DSM4690 DNA versus crossing point (Cp) values obtained for monoplex *pstA* (A) and *phoB* (B) assays and multiplex *pstA* (C) and *phoB* (D) assays.