Author version of a manuscript published in FEMS Pathogens and Disease (2019),
DOI: 10.1093/femspd/ftz015

Metabolic analyses reveal common adaptations in two invasive *Haemophilus*
influenzae strains

Noor Marian Muda^a, Marufa Nasreen^a, Rabeb Dhouib^a, Jennifer Hosmer^a, Julian Hill^a,
Manish Mahawar^{a,b}, Horst Joachim Schirra^c, Alastair G. McEwan^a, Ulrike Kappler^{a*}

^a School of Chemistry and Molecular Biosciences, Australian Infectious Disease
Research Centre, Centre for Metals in Biology, The University of Queensland, St. Lucia
4072, Australia

^b Indian Veterinary Research Institute, Izatnagar, India

^c Centre for Advanced Imaging, The University of Queensland, St. Lucia 4072,
Australia

* author for correspondence

E-mail: u.kappler@uq.edu.au

Phone: +61 7 33652978

Abstract

Non-typeable *Haemophilus influenzae* (NTHi) is a major pathogen in upper and lower respiratory tract infections in humans, and is increasingly also associated with invasive disease. We have examined two unrelated NTHi invasive disease isolates, R2866 and C188, in order to identify metabolic and physiological properties that distinguish them from respiratory tract disease isolates such as Hi2019. While the general use of the Hi metabolic network was similar across all three strains, the two invasive isolates secreted increased amounts of succinate which can have anti-inflammatory properties. In addition, they showed a common shift in their carbon source utilization patterns, with strongly enhanced metabolism of nucleoside substrates, glucose and sialic acid. The latter two are major compounds present in blood and CSF. Interestingly, C188 and R2866 also shared a reduced ability to invade or survive intracellularly in 16HBE14 bronchial epithelial cells relative to Hi2019 (4-fold (4h), 25-fold (24h) reduction). Altered metabolic properties such as the ones observed here could arise from genomic adaptations that NTHi undergo during infection. Together these data indicate that shifts in substrate preferences in otherwise conserved metabolic pathways may underlie strain niche specificity and thus have the potential to alter the outcomes of host-NTHi interactions.

Keywords: *Haemophilus influenzae*; invasive disease, metabolism, phenotype, gene expression

Introduction

Haemophilus influenzae (Hi) is a host-adapted human pathogen capable of causing diseases of both the upper and lower respiratory tract (Van Eldere, *et al.*, 2014, Langereis & de Jonge, 2015, Duell, *et al.*, 2016). Hi strains are present in a large proportion of the population as commensals of the nasopharynx, but there is a clear link between Hi dissemination to other areas of the respiratory tract and both childhood and chronic diseases (Murphy, *et al.*, 2009, Essilfie, *et al.*, 2011, King, 2012, Van Eldere, *et al.*, 2014, Duell, *et al.*, 2016). In recent years, a rising incidence of invasive disease caused by non-typeable strains of Hi (NTHi) that lack a polysaccharide capsule has been reported (Tsang, *et al.*, 2014, Ulanova & Tsang, 2014, Centre for Disease Control, 2015, Lm, *et al.*, 2015, Langereis & de Jonge, 2015, Wan Sai Cheong, *et al.*, 2015). This has raised concerns that, following the introduction of a vaccine against the highly virulent Hi serotype b strains, the disease and colonization patterns of this pathogen are changing, favouring NTHi strains and strains from other serotypes and enabling them to disseminate to new body niches.

NTHi strains are known to be genetically variable, with a total of 6453 genes detected across available genomes in a recent study (Moleres, *et al.*, 2018). 1305 core genes were found in over 95% of strains, while 4505 were present in 15% or less of the strains and form the accessory genome (Moleres, *et al.*, 2018). On average, NTHi strains contain 1813 genes (Pettigrew, *et al.*, 2018). This diversity makes it difficult to identify general targets to exploit for the development of new treatment options that would be useful for a wide range of NTHi strains, especially when combined with the fact that many of the known surface proteins and adhesins undergo phase-variation, may not be fully conserved across strains and may also undergo frequent amino acid changes in antigenic surface structures (Hogg, *et al.*, 2007, De Chiara, *et al.*, 2014, Moleres, *et al.*, 2018, Osman, *et al.*, 2018, Pettigrew, *et al.*, 2018). Moreover, the different body niches inhabited by NTHi differ significantly both in nutrient

contents and the availability of oxygen. Mucus layers and biofilm environments as well as the inflamed middle ear would be oxygen limited, while airway epithelial surfaces would be more aerobic (Erwin & Smith, 2007, King, 2012, Othman, *et al.*, 2014). Unlike genes encoding traditional virulence factors, genes involved in central metabolism and energy generation are highly conserved in Hi, and their importance for the virulence of pathogens is being increasingly recognized (Munoz-Elias & McKinney, 2005, Alteri & Mobley, 2012, Bliska & van der Velden, 2012, Othman, *et al.*, 2014). In fact, the concept has been demonstrated in Hi previously for enzymes involved in energy generation (Langereis, *et al.*, 2013, Dhouib, *et al.*, 2016), and there is also evidence for strain specific adaptations in the use of the common central metabolic networks in the reference strain RdKW20 and a COPD isolate, Hi2019 (Othman, *et al.*, 2014). This led us to hypothesise that metabolic adaptation of Hi to suit the nutrient availability of different body niches might be a key driver of infectivity and persistence in these niches (Othman, *et al.*, 2014).

Here we have undertaken a comparative analysis of two invasive isolates of NTHi: the well-studied strain R2866, isolated from a child with meningitis as a complication of otitis media, and strain C188, a blood isolate for which a genome sequence was recently reported (Nizet, *et al.*, 1996, Kappler, *et al.*, 2017). Within the classification proposed by (De Chiara, *et al.*, 2014), R2866 is found in genomic clade VI, while HiC188 is placed in Clade V based on the placement of strains most closely related to C188 at the genomic level (492_HINF; GCU26214; 411; 584; 1104). MLST based analyses placed C188 close to a ST43 strain PittGG), while R2866 represents ST1621 (Pettigrew, *et al.*, 2018). Given the high genetic diversity of Hi strains, it is unclear why some isolates appear to be able to access body niches that are not generally colonized by Hi, and our study aimed to determine whether there were common metabolic and physiological traits associated with these two genetically dissimilar

95 strains that could be associated with niche adaptation and survival of Hi following invasion of
96 the host tissues.

Materials and Methods

Bacterial Strains & Growth Conditions

H. influenzae strains Hi2019 (Campagnari, *et al.*, 1987), R2866 (Nizet, *et al.*, 1996) and NTHi C188 (Nelson & Smith, 2010) were grown at 37°C with 5% CO₂ for 16h using either brain heart infusion (BHI) agar or broth (Difco) supplemented with hemin (10µg/ml) and NAD (1µg/ml) (sBHI), or chemically defined medium (CDM) (Coleman, *et al.*, 2003). Antibiotic resistance testing used HTM medium inoculated with freshly prepared cell suspension (OD₆₀₀ = 0.05) and antibiotic susceptibility testing discs (Thermo Fisher).

Growth experiments were performed using CDM and aerobic, microaerobic or anaerobic growth conditions with three biological replicates for each strain and condition as in (Othman, *et al.*, 2014). To test self- attachment fully grown cultures (24h, anaerobic) were inverted to resuspend the bacteria and changes in OD₆₀₀ were monitored for a total of 5h. Biofilm formation was investigated using microtiter plate biofilm assay based on the protocol of (Schembri & Klemm, 2001) with modifications as described in (Dhouib, *et al.*, 2016).

Phenotypic Microarray data was collected using PM1 and PM2 carbon substrate arrays (Biolog) and the protocol for fastidious bacteria provided by Biolog. Inoculum (65%T, IF-0a GN with redox dye G and PM1 or 2 supplement for fastidious organisms (12xstock)) was prepared using freshly grown overnight cultures of *H. influenzae* on sBHI plates. Data were recorded for 48h at 37°C using an Omnilog system (Biolog) and analysed using the Omnilog PM software packages. All plates were normalized to the control well (A1) before comparative analyses were carried out. Analyses used both final absorbance and area under the curve for comparisons. Data shown is normalized to the reference well on each plate.

Molecular Biology Methods

Standard methods were used throughout (Ausubel, *et al.*, 2005). Genomic DNA was isolated using the PureLink Genomic DNA Mini Kit (Invitrogen), PCR reactions used GoTaq

Green MasterMix (Promega) according to the manufacturer's instructions. Oligonucleotide primers were designed using Vector NTI 11 (Life Technologies) (Tab.S1). For RNA isolation 1.5 ml of bacterial culture in mid exponential growth phase were treated with RNAProtect Bacteria (Qiagen), RNA was isolated using the Illustra RNASpin Mini Kit (GE Healthcare). Contaminating gDNA was removed with the TurboDNA-free™ kit (Life Technologies) followed by RNA clean up using the RNAspin Kit and concentration determination using the Quant-it RNA Kit (Life Technologies). cDNA was synthesized from 500 ng gDNA-free RNA using Superscript III (Life Technologies) and random hexamer primers (Life Technologies) as per the manufacturer's instructions. cDNA sample were diluted (1:10-1:1000). qRT-PCR was performed at the SCMB qRT-PCR facility using a ViiA or Quantstudio3 system (Applied Biosystems) and SYBRgreen Mastermix (Applied Biosystems). Data analysis used the LinReg software (Ramakers, *et al.*, 2003) to determine PCR efficiencies. All data were normalized to the expression of *gyrA* following analysis of potential reference genes (*gyrA*, *rpoA* and 16S) using Normfinder (Pinto, *et al.*, 2012).

Isolation and Identification of Outer Membrane Proteins

Outer membrane proteins (OMPs) were isolated from stationary growth phase cultures (~18h culture time) on BHI according to the procedure of (Carlone, *et al.*, 1986) and analysed using MS/MS sequencing of tryptic peptides as in (Kappler & Nouwens, 2013). Proteins were identified using Protein Pilot, protein scores of 4.0 and 2.0 were used for stringent (99% confidence) and general (95% confidence) protein identification.

Host Cell Adherence and Invasion Assays

Adherence and invasion assays were performed using 16HBE14 cells as described in (Dhouib, *et al.*, 2016). In brief, confluent layers of 16HBE14 cells were infected with *H. influenzae* (MOI 100:1 - bacteria: tissue cells). Numbers of total adherent (adherent bacteria and any bacteria present inside the cell) and host cell internalized cells were determined after

either 4h or 24h incubation by washing the cell layers with fresh medium, followed by lysis using 1% saponin and dilution plating. For internalized cells the co-cultures were additionally treated with gentamicin (50µg/ml) for 1h prior to the wash steps. The percentage of bacteria located inside the host cells at each timepoint was calculated by dividing CFU/ml of internalized cells by CFU/ml of total adherent bacteria at each timepoint.

Comparative analysis of *H. influenzae* genomes

Proteinortho and Mauve (Darling, *et al.*, 2004, Lechner, *et al.*, 2011) were used for comparative analyses and visualization of published *H. influenzae* genome sequences used data.

Nuclear Magnetic Resonance (NMR) analysis of metabolites

Samples for NMR analysis (1ml) were taken from growing cultures at different timepoints. Cells were removed by centrifugation, and supernatants preserved using azide before storage at -80°C as in (Othman, *et al.*, 2014). For R2866, NMR samples were prepared by mixing 180µl preserved supernatant with 20µl of a solution containing 1mM 2,2-dimethylsilapentane-5-sulfonic acid (DSS) and 1mM 1,1-difluoro-1-trimethylsilanyl phosphonic acid (DFTMP)) in D₂O (DDD solution); for the C188 strain NMR samples were prepared by mixing 165µl preserved supernatant with 15µl of a 2M sodium phosphate buffer, pH 7.4, and with 20µl DDD solution, yielding a total sample volume of 200µl. Samples were transferred into 3mm NMR tubes. ¹H- NMR spectroscopy was performed as in (Schirra, *et al.*, 2008). Proton NMR spectra of the C188 strain were measured on a Bruker AVII+ NMR spectrometer with a frequency of 700.47 MHz, whereas spectra of the R2866 strain were measured on a Bruker AVIII spectrometer with a frequency of 900.13 MHz (Bruker Biospin, Rheinstetten. Germany). Both spectrometers were equipped with a triple resonance cryoprobe and a SampleJet sample changer. 1D Carr-Purcell-Meiboom-Gill (CPMG) spectra were measured at 298 K with the Bruker *cpmgpr1d* pulse sequence, using a spin-lock sequence containing 20 repeats of a (500 ms – 180° pulse – 500 ms) CPMG element. Spectra were acquired with a spectral width

of 14 ppm. The carrier frequency was set to the water signal, which was suppressed with low-power continuous irradiation during the relaxation delay of 3 s. For the C188 samples spectra were acquired with a digital resolution of 64k data points and a receiver gain of 64, whereas for the R2866 sample the digital resolution was 32k data points and the receiver gain 128.

NMR spectra were processed in Topspin version 3.1 Software (Bruker Biospin) by multiplying with a sine window function shifted by $\pi/2$, followed by Fourier-transformation, manual baseline correction, and calibration of the chemical shift axis on the DSS signal at 0 ppm. For targeted metabolite analysis, the NMR signals of 26 metabolites were integrated in AssureNMR (Bruker Biospin) with an automated integration and quantification method, optimised for the 700 and 900 MHz spectra respectively. Metabolite concentrations were obtained from the raw integrals via the formula:

$$c_x = \frac{nH_{ref}}{nH_x} \frac{I_x}{I_{ref}} c_{ref}$$

with c_x being the concentration of a metabolite of interest, c_{ref} being the known concentration of a reference substance, nH_x and nH_{ref} being the number of protons causing the NMR signals of the metabolite of interest and the reference metabolite, respectively, and I_x and I_{ref} being the raw integral of these two signals. The signal of ethylene glycol served as a concentration reference, which was present in the C188 samples at a concentration of 25.770 mM, and at a concentration of 28.113 mM in the R2866 samples. The raw intensities I_{ref} of the ethylene glycol signal were averaged for the C188 and R2866 spectra respectively, and the corresponding average $\langle I_{ref} \rangle$ used for I_{ref} in the above formula. Concentrations were plotted with an in-house Matlab script.

Results

Despite being both associated with invasive disease, genomic and physiological properties of C188 and R1866 are distinct

The average genome size for NTHi strains that has been reported as 1.841Mb with 1813 encoded proteins (Pettigrew, *et al.*, 2018), and this agrees well with the 1.932 Mb and 1803 protein coding genes for *H. influenzae* R2866, while Hi C188 (proteins:1720, genome size: 1.848 Mb) encodes 100 proteins less than the average strain. Although the two strains belong to only distantly related STs, 1529 protein coding genes were common to the two strains, ~ 200 more than form the Hi core genome (1305 genes) (Moleres, *et al.*, 2018), with 265 (R2866) and 186 (C188) unique genes, respectively. When the comparison was expanded to include the third strains studied here, Hi2019 and the HiRDKW20 reference strain, only 1348 protein coding genes were common to all four strains (Fig.1A, Tab.S5). Alignment of the R2866, C188, Hi2019 and RDKW20 genomes using MAUVE showed overall synteny across the genome sequences, with some inversions and insertions distinguishing the strains (Fig.S1A).

In both R2866 and C188 the unique genes were dominated by hypothetical proteins (48% and 53%, R2866 & C188), followed by about 15% of virulence associated genes that included outer membrane and LPS synthesis proteins as well as antibiotic resistance genes (Tab.S2, Fig.S1B). General functions/housekeeping, metabolism, transport and regulators were categories present at <8% each. The unique proteins included urease in strain C188 for which R2866 is known to be negative, as well as a beta lactamase in strain R2866 which correlated with an increased ampicillin resistance (Amp10 inhibition zone R2866: 26 mm, C188: 38 mm). Interestingly, unique phage and transposon associated genes were more abundant in strain R2866 (13%) than in strain C188 (7%). Analyses of the genes/proteins shared by all strains revealed that genes involved in central metabolism and respiration (analysed for Hi2019 and HiRDKW20 in (Othman, *et al.*, 2014))(Tab.S3, S5) were also conserved in Hi strains R2866, C188 (amino acid

similarity: 96-100%, Tab.S3) with the only exception of a second copy of a xanthine phosphoribosyltransferase encoding gene (*gtp2*) that was absent in HiC188, confirming that central metabolic functions are conserved in these strains Hi.

Despite their common origin from invasive diseases, the physiological properties of Hi strains R2866 and C188 varied, likely reflecting their genetic dissimilarity. On glucose-containing CDM, growth of R2866 exceeded that of C188 (aerobic: 1.28 x, microaerobic: 1.17 x) and resembled our published data for Hi2019 (Dhouib, *et al.*, 2016), while anaerobic growth was similar (Fig.S2A). Strain R2866 showed strong self-adhesion with rapid settling occurring following resuspension of cell in the growth medium (OD₆₀₀ reduction: 2h: >50%, 5h: >90%), while for HiC188 ODs were only reduced by ~50% after 5h (Fig.S2B). Similarly, *in vitro* biofilm formation was approximately 3-fold enhanced for R2866 relative to C188 under all condition tested, while Hi2019 showed an intermediate phenotype (Fig.1B). Strong biofilm formation has been observed previously for strain R2866 (Marti, *et al.*, 2017), and has also been reported as a general feature of many invasive and otitis media NTHi isolates (Puig, *et al.*, 2014).

Hi outer membrane proteins are essential for virulence, with a number of OMPs known to undergo phase variation and to be fast evolving (Moleres, *et al.*, 2018). Outer membrane protein preparations of C188 and R2866 (57(40) and 58(36) proteins, 95%(99%) confidence) showed extensive similarities, with 39 proteins present in both samples (Tab.S4, S5). The top five identified proteins included BamA (D15 surface antigen), OmpP2 and OmpP5 in both strains, and Hia, which was the most abundant OMP in R2866. Hi2019 samples also had overall similar protein profiles, and additionally contained the HMW-type adhesin for which C188 & R2866 are negative. Other OMPs identified included transport proteins, LPS modification & synthesis proteins and outer membrane protein assembly factors (Tab.S4), all of which are

typical components of outer membrane protein and outer membrane vesicle preparations (Sharpe, *et al.*, 2011, Roier, *et al.*, 2012).

NTHi strains R2866 and C188 have a reduced level of survival during interactions with epithelial tissue cells compared to the Hi2019 COPD isolate strain Being able to adhere to and invade human epithelial cells is essential for Hi being able to establish infection. In co-culture with 16HBE14 bronchial epithelial cells, clear differences were observed between the three strains. For planktonic growth, CFU/ml for all strains were similar after 4h of incubation ($1.2\text{--}1.6 \times 10^8$ CFU/ml), however, after 24h cell numbers of Hi2019 were 145% of the 4h value, but were only 2.5% and 16%, respectively, for R2866 and C188 (Fig.1C, S3).

For Hi2019, cell counts for total tissue cell adherent bacteria increased (2-fold, $p < 0.0001$) from 4h to 24h of incubation, but decreased by 50% for R2866 ($p = 0.0013$) and essentially remained the same for C188 (20% decrease, not significant)(Fig.1C, S3). Levels of intracellular bacteria were 0.04% (4h) and 0.23% (24h) of total adherent cells for C188 and R2866, compared to 0.14% (4h) and 1.8% (24h) for Hi2019 (Fig.1C, S3). Thus, both invasive disease strains showed reduced adherence and a reduced ability to invade and/or replicate following entry into the tissue cells when compared to Hi2019 which was used as the reference strains in these experiments.

Metabolic and metabolomic analyses revealed common substrate utilization capabilities across NTHi strains but key differences in the pattern of substrate utilization General metabolism, including energy generation and the ability to utilize substrates available in the respiratory tract or other body niches is a relatively unexplored aspect of NTHi physiology. The Biolog Phenotypic Microarray PM1 & 2 plates were used to investigate carbon substrate utilization for R2866, C188 and Hi2019. At a general level, the carbon source

metabolic patterns were very similar, with R2866, C188 and Hi2019 utilizing 50, 54 and 48 of 196 substrates, respectively. 38 substrates supported growth of all three strains (Tab.S6), however, while some substrates such as L-lyxose, D- & L- arabinose, L- lactic acid and D- ribose were used strongly by all strains, both R2866 and C188 had an increased ability to metabolize nucleosides (e.g. inosine, adenosine, thymidine) as well as pyruvate, glucose and N-acetyl neuraminic (sialic) acid when compared to Hi2019 (Fig.2). This difference was present in both final absorbance values reached and utilization kinetics, with the invasive NTHi strains showing higher substrate utilization rates, reflected in a strong, linear increases in redox dye absorbance. In contrast, for substrates such as ribose utilization rates and final absorbances reached were essentially the same for all three strains (Fig.2).

We then used $^1\text{H-NMR}$ to analyse metabolite production by strains C188 and R2866 as a function of time and oxygen tension in CDM medium containing glucose, inosine and pyruvate as major carbon sources (Fig.3, Tab.1). Pyruvate was metabolized completely within the first few hours of growth, while inosine and glucose were consumed more slowly, with concentrations approaching zero only after 24h. Utilization rates for glucose and inosine during exponential growth of the cultures decreased by 30-70% depending on the strain and substrate with decreasing oxygen concentrations (Fig.3).

For all conditions, acetate was the major metabolite produced, with highest concentrations being produced during aerobic growth, where the only other major product was hypoxanthine, a by-product of inosine breakdown. Product formation rates matched the respective changes of substrate concentrations for glucose and inosine (Figs.3&4). With decreasing oxygen tension, acetate production decreased, and additional products such as formate and succinate were present in increasing amounts (Fig.4). The highest succinate concentrations were observed under anaerobic conditions after 24h. A similar endproduct spectrum although with lower final succinate and higher formate concentrations was previously

reported for Hi2019 and HiRDKW20, and is typical for bacterial fermentative metabolism (Othman, *et al.*, 2014). Acetate and succinate have both been shown to have immunomodulatory properties, with the effects of acetate having been particularly well defined for host-bacteria interactions in the human gut (Louis, *et al.*, 2014, Peruzzotti-Jametti, *et al.*, 2018).

Expression of metabolic genes of NTHi R2866 and NTHi C188 in response to changing growth conditions matches the metabolic patterns To confirm some of the above observations, we analysed the expression of conserved genes in central metabolism as a function of oxygen availability (Fig.5). Throughout, both R2866 and C188 expressed key genes for the Embden-Meyerhof Pathway (glycolysis) (*pfkA*), the hexose monophosphate shunt (*zwf*), which feeds into the pentose phosphate pathway, and *ackA*, involved in the production of acetate, which matches the metabolite profile. The increased expression of pyruvate:formate lyase (*pflA*) at lower oxygen tension (microaerobic: 2.5/2.8 x increase, anaerobic: 4.5/8.9 x, $p < 0.0001$ (t-test)) corresponds well with the increased production of formate during anaerobic conditions. We also observed high expression of the *deoD* gene, which encodes a purine:nucleoside phosphorylase that converts inosine to hypoxanthine and ribose-1-phosphate, while other enzymes in purine salvage metabolism such as *hpt* (inosine to IMP conversion), *guaB* (IMP to XMP) and *guaA* (XMP to GMP) were only expressed at 30% (R2866) to 50% (C188) of the *deoD* expression levels. These gene expression patterns are in keeping with the observed formation of hypoxanthine as a product of inosine consumption during growth, with the ribose-1-phosphate likely being used as a supplementary carbon source via the pentose phosphate pathway.

Although the metabolic endproducts we identified indicated a mostly fermentative metabolism, components of the Hi respiratory chain were also highly expressed. Of the Hi

respiratory chain dehydrogenases the genes encoding the sodium-translocating NADH dehydrogenase, *nqrB*, and the L-lactate dehydrogenase (*lldD*) were highly expressed, followed by fumarate reductase (*frdA*) involved in the production of succinate (Fig.5). For the respiratory chain reductases expression levels of the cytochrome *bd* oxidase were relatively constant throughout, while the alternative reductases (*napA*, *nrfA*, *mtsZ*, *dmsA*, converting nitrate, nitrite and S-/N-oxides, respectively) increased in expression under oxygen limiting conditions ($p < 0.005$ for all genes, t-test). Genes encoding nitrite and nitrate reductases (*nrfA* and *napA*) were the most highly expressed respiratory enzymes under microaerobic and anaerobic conditions. Together with data from previous work (Othman, *et al.*, 2014) these results indicate that all three Hi strains use the same basic mechanisms of energy generation, but show variations in their affinities for specific substrates, where despite belonging to different phylogenetic clades and sequence types, the two invasive strains showed a common, enhanced ability to metabolize nucleosides and sugar derivatives.

Discussion

While *H. influenzae* is known to be able to survive and cause infection in a variety of body niches, specific factors leading to niche adaptation of isolates have so far not been identified. We began this work with the hypothesis that despite the genetic diversity inherent in NTHi strains, strains from similar niches of isolation might share some characteristics that would enhance their ability to preferentially colonize a given body niche and have used a comparative study of two invasive disease isolates, strains *H. influenzae* R 2866 and C188, and the COPD isolate Hi2019 to investigate this, with a particular focus on metabolic adaptations that may promote enhanced survival in additional body niches. The two strains differ clearly in their genomic content as well as basic physiological parameters which is typical for unrelated Hi isolates.

We did, however, observed striking and unexpected changes to the metabolic abilities of the two invasive strains: Compared to Hi2019, the two invasive isolates showed an increased ability to metabolize a variety of purine and pyrimidine nucleosides, as well as glucose, pyruvate and N-acetyl neuraminic acid (Neu5Ac, sialic acid) (Fig.2). Glucose is abundant in blood (conc. >4 mM), and increases in serum concentrations of nucleosides have been shown to be associated with injury (Chaleckis, *et al.*, 2016). In addition, extracellular DNA, present, e.g., as a result of NET (Neutrophil Extracellular Trap) formation, can be a source of nucleosides. Sialic acid is known to be acquired by *H. influenzae* and can be used either for incorporation into the LPS or as a carbon source that is metabolized via glucose-6 phosphate (Apicella, 2012). In blood, levels of sialic acid for healthy individuals have been reported to be in the mM range (1.58-2.22 mM (Sillanauke, *et al.*, 1999)) making an increased ability to use sialic acid advantageous for invasive NTHi isolates.

Despite these documented changes in the substrate utilization profile, all three NTHi strains carried out a form of mixed-acid fermentative metabolism irrespective of oxygen

availability, which is supported by the action of a respiratory chain. Under aerobic conditions acetate is the major carbon-endproduct from glucose, suggesting that a significant amount of ATP synthesis in Hi derives from substrate level phosphorylation. This incomplete oxidation of glucose, even under aerobic conditions is consistent with the absence of a complete oxidative branch of the TCA cycle in this bacterium (Tatusov, *et al.*, 1996, Othman, *et al.*, 2014). As oxygen availability decreases, succinate and formate are produced as additional carbon metabolites, consistent with standard patterns of mixed acid fermentation where the ability to re-oxidise NADH becomes restricted, and e.g. expression of pyruvate-formate lyase (*pflA*) is elevated since this enzyme does not generate NADH. Expression of the fumarate reductase was relatively constant regardless of oxygen availability, suggesting that in the presence of oxygen, succinate produced by this enzyme was either metabolized further, or that, as in *E.coli*, this enzyme could be reversible (Maklashina, *et al.*, 2013) and could also act as a succinate dehydrogenase. In the current study production of succinate by the two invasive isolates as a major metabolite under anaerobic conditions was more pronounced than previously observed (Othman, *et al.*, 2014).

Unlike what happens in purely fermentative metabolism, in NTHi NADH produced during carbon metabolite oxidation, would mostly be re-oxidised by the respiratory chain with oxygen as the terminal electron acceptor. The Nqr NADH dehydrogenase may have a major role in this process by transferring electrons to the menaquinone pool while the Nqr sodium motive capability may play a role in sodium homeostasis and in providing an ion-gradient to drive substrate uptake. Transcripts for the second NADH dehydrogenase (*ndh*) were present at a lower level, and this enzyme is not linked to ion-translocation. Formate dehydrogenase (*fdxG*), and D-lactate dehydrogenase (*dld*) were expressed at low levels and would be expected not to contribute significantly to Hi respiration under the chosen experimental conditions.

The oxygen-reducing cytochrome *bd* oxidase is likely to play a major role under aerobic and microaerobic growth conditions and has a high affinity for oxygen but very poor proton-translocating capability (Borisov, *et al.*, 2011). It is therefore reasonable to assume that the primary function of the aerobic respiratory chain in Hi is in redox balance rather than proton-motive force generation to drive ATP synthesis, although a proton-translocating ATP synthase is present (Othman, *et al.*, 2014).

Expression of NTHi respiratory chain enzymes also varied when oxygen concentrations became limited with several enzymes supporting anaerobic respiration (e.g. on nitrate and nitrite) increasing in expression despite the absence of added electron acceptors. In contrast, the *bd* cytochrome oxidase continued to be expressed at similar levels under microaerobic and anaerobic conditions. This could also indicate that as described for other bacteria, the Hi *bd* oxidase could have additional roles e.g. in preventing nitrosative stress, which would be an oxygen-independent function (Giuffrè, *et al.*, 2014).

Together, these data show that metabolic adaptation in *H. influenzae* could be a possible determinant of niche preferences and may be based on differences in the level of expression and regulation of a common repertoire of metabolic enzymes. During infection, Hi strains undergo genomic adaptation, but genes shown to be under strong evolutionary pressure generally encode outer membrane proteins that undergo strong variation in surface loops (De Chiara, *et al.*, 2014, Garmendia, *et al.*, 2014, Osman, *et al.*, 2018, Pettigrew, *et al.*, 2018). In some cases a selection for the loss of a protein is also observed as in the case of the FadL surface protein where a loss of the protein was shown to be advantageous post-infection (Moleres, *et al.*, 2018). So far none of the genes involved in core energy and carbon metabolism have been implicated in this type of evolutionary adaptation, but our data indicate that subtle changes in the ability to utilize specific growth substrates may enable the colonization of additional body niches by Hi strains. A possible scenario for this could be changes in e.g. the promoter regions

of the affected genes or in regulator proteins that control expression of metabolic functions, or potentially also amino acid substitutions that change the properties of key enzymes in the degradation or uptake of specific substrates and this should be investigated in the future.

Interestingly, in addition to the variations in growth substrate utilization, when compared to the Hi2019 COPD isolate, the two invasive HI isolates showed a common, reduced ability to establish themselves in during co-culture with host cells. This affected not only adherent growth, but led to a reduction in cell invasion/ intracellular survival compared to t Hi2019, suggesting that the changes in the metabolic profile seen in the invasive isolates may not be advantageous for intracellular survival. Low levels of host cell invasion have previously been observed for strain R2866 during co-culture with immortalized A549 lung cells, or Chang conjunctival cells (Tsao, *et al.*, 2012), however, cell invasion by strain R2866 can vary depending on the cell line used, with higher rates observed for normal human bronchial epithelia cells (NHBE) (Watson, *et al.*, 2004, VanWagoner, *et al.*, 2016).

In summary our analyses provide first evidence for changes in the *H. influenzae* metabolic profile as a possible determinant of niche adaptation in *H. influenzae* strains, with changes in host cell interactions being a second parameter that should be evaluated in this regard. Further studies of metabolic properties of NTHi isolates from different body niches/disease types should be undertaken to provide further insights into the role of substrate utilization and host cell interactions as potential niche-specific adaptations.

Acknowledgements

This work was funded by the National Health & Medical Research Council of Australia (GNT1043532) to UK and AGM. NMM was supported by a scholarship from the Malaysian government (MARA), MN was supported by a Faculty for the Future Scholarship (Schlumberger Foundation), JH is supported by a Research Training Program Scholarship (Department of Education and Training, Commonwealth of Australia). The authors do not declare a conflict of interest.

References

- Alteri CJ & Mobley HLT (2012) *Escherichia coli* physiology and metabolism dictates adaptation to diverse host microenvironments. *Current Opinion in Microbiology* **15**: 3-9.
- Apicella MA (2012) Nontypeable *Haemophilus influenzae*: the role of N-acetyl-5-neuraminic acid in biology. *Frontiers in Cellular and Infection Microbiology* **2**: 19.
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA & Struhl K (2005) Current Protocols in Molecular Biology. *Current Protocols in Molecular Biology*, Vol. 1 (Janssen K, eds.), pp. John Wiley & Sons Inc., Hoboken, NJ.
- Bliska JB & van der Velden AWM (2012) *Salmonella* ,Sops' Up a Preferred Electron Receptor in the Inflamed Intestine. *mBio* **3**: 10.1128/mBio.00226-00212.
- Borisov VB, Gennis RB, Hemp J & Verkhovsky MI (2011) The cytochrome bd respiratory oxygen reductases. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1807**: 1398-1413.
- Campagnari AA, Gupta MR, Dudas KC, Murphy TF & Apicella MA (1987) Antigenic diversity of lipopolysaccharides on nontypable *Haemophilus influenzae*. *Infection and Immunity* **55**: 882-887.
- Carlone GM, Thomas ML, Rumschlag HS & Sottnek FO (1986) Rapid microprocedure for isolating detergent-insoluble outer membrane proteins from *Haemophilus* species. *Journal of Clinical Microbiology* **24**: 330-332.
- Centre for Disease Control (2015) *Haemophilus influenzae* invasive disease. <http://wwwn.cdc.gov/nndss/conditions/haemophilus-influenzae-invasive-disease/case-definition/2015/>.
- Chaleckis R, Murakami I, Takada J, Kondoh H & Yanagida M (2016) Individual variability in human blood metabolites identifies age-related differences. *Proceedings of the National Academy of Sciences* **113**: 4252-4259.
- Coleman HN, Daines DA, Jarisch J & Smith AL (2003) Chemically defined media for growth of *Haemophilus influenzae* strains. *Journal of Clinical Microbiology* **41**: 4408-4410.

- 469 Darling AC, Mau B, Blattner FR & Perna NT (2004) Mauve: multiple alignment of conserved
470 genomic sequence with rearrangements. *Genome Res* **14**.
- 471
- 472 De Chiara M, Hood D, Muzzi A, *et al.* (2014) Genome sequencing of disease and carriage
473 isolates of nontypeable *Haemophilus influenzae* identifies discrete population structure.
474 *Proceedings of the National Academy of Sciences of the United States of America* **111**: 5439-
475 5444.
- 476
- 477 Dhouib R, Othman DSMP, Lin V, *et al.* (2016) A Novel, Molybdenum-Containing Methionine
478 Sulfoxide Reductase Supports Survival of *Haemophilus influenzae* in an In vivo Model of
479 Infection. *Frontiers in Microbiology* **7**: e1743.
- 480
- 481 Duell BL, Su Y-C & Riesbeck K (2016) Host–pathogen interactions of nontypeable
482 *Haemophilus influenzae*: from commensal to pathogen. *FEBS Letters* **590**: 3840-3853.
- 483
- 484 Erwin AL & Smith AL (2007) Nontypeable *Haemophilus influenzae*: understanding virulence
485 and commensal behavior. *Trends in Microbiology* **15**: 355-362.
- 486
- 487 Essilfie A-T, Simpson JL, Horvat JC, *et al.* (2011) *Haemophilus influenzae* Infection Drives
488 IL-17-Mediated Neutrophilic Allergic Airways Disease. *PLoS Pathog* **7**: e1002244.
- 489
- 490 Garmendia J, Viadas C, Calatayud L, *et al.* (2014) Characterization of Nontypable
491 *Haemophilus influenzae* Isolates Recovered from Adult Patients with Underlying Chronic Lung
492 Disease Reveals Genotypic and Phenotypic Traits Associated with Persistent Infection. *PLoS*
493 *ONE* **9**: e97020.
- 494
- 495 Giuffrè A, Borisov VB, Arese M, Sarti P & Forte E (2014) Cytochrome bd oxidase and bacterial
496 tolerance to oxidative and nitrosative stress. *Biochimica et Biophysica Acta (BBA) -*
497 *Bioenergetics* **1837**: 1178-1187.
- 498
- 499 Hogg JS, Hu FZ, Janto B, *et al.* (2007) Characterization and modeling of the *Haemophilus*
500 *influenzae* core and supragenomes based on the complete genomic sequences of Rd and 12
501 clinical nontypeable strains. *Genome Biology* **8**: R103-R103.
- 502
- 503 Kappler U & Nouwens AS (2013) The molybdoproteome of *Starkeya novella* - insights into
504 the diversity and functions of molybdenum containing proteins in response to changing growth
505 conditions. *Metallomics* **5**: 325-334.

- 506
- 507 Kappler U, Dhouib R, Nair RP & McEwan AG (2017) Draft genome sequences of three non-
 508 typeable strains of *Haemophilus influenzae*, C188, R535 and 1200 isolated from different types
 509 of disease. *Genome Announcements* **5**: e00035-00017.
- 510
- 511 King P (2012) *Haemophilus influenzae* and the lung (*Haemophilus* and the lung). *Clinical and*
 512 *Translational Medicine* **1**: 10.
- 513
- 514 Lâm T-T, Claus H, Elias J, Frosch M & Vogel U (2015) Ampicillin resistance of invasive
 515 *Haemophilus influenzae* isolates in Germany 2009–2012. *International Journal of Medical*
 516 *Microbiology* **305**: 748-755.
- 517
- 518 Langereis JD & de Jonge MI (2015) Invasive Disease Caused by Nontypeable *Haemophilus*
 519 *influenzae*. *Emerging Infectious Diseases* **21**: 1711-1718.
- 520
- 521 Langereis JD, Zomer A, Stunnenberg HG, Burghout P & Hermans PWM (2013) Nontypeable
 522 *Haemophilus influenzae* Carbonic Anhydrase Is Important for Environmental and Intracellular
 523 Survival. *Journal of Bacteriology* **195**: 2737-2746.
- 524
- 525 Lechner M, Findeiß S, Steiner L, Marz M, Stadler PF & Prohaska SJ (2011) Proteinortho:
 526 Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics* **12**: 1-9.
- 527
- 528 Louis P, Hold GL & Flint HJ (2014) The gut microbiota, bacterial metabolites and colorectal
 529 cancer. *Nat Rev Micro* **12**: 661-672.
- 530
- 531 Maklashina E, Cecchini G & Dikanov S (2013) Defining a direction: electron transfer and
 532 catalysis in *Escherichia coli* complex II enzymes. *Biochimica et biophysica acta* **1827**: 668-
 533 678.
- 534
- 535 Marti S, Puig C, Merlos A, *et al.* (2017) Bacterial Lysis through Interference with
 536 Peptidoglycan Synthesis Increases Biofilm Formation by Nontypeable *Haemophilus*
 537 *influenzae*. *mSphere* **2**: e00329-00316.
- 538
- 539 Moleres J, Fernández-Calvet A, Ehrlich RL, *et al.* (2018) Antagonistic Pleiotropy in the
 540 Bifunctional Surface Protein FadL (OmpP1) during Adaptation of *Haemophilus influenzae* to

- 541 Chronic Lung Infection Associated with Chronic Obstructive Pulmonary Disease. *mBio* **9**:
542 e01176-01118.
- 543
- 544 Munoz-Elias EJ & McKinney JD (2005) *Mycobacterium tuberculosis* isocitrate lyases 1 and 2
545 are jointly required for in vivo growth and virulence. *Nat Med* **11**: 638-644.
- 546
- 547 Murphy TF, Faden H, Bakaletz LO, *et al.* (2009) Nontypeable *Haemophilus influenzae* as a
548 Pathogen in Children. *The Pediatric Infectious Disease Journal* **28**: 43-48.
- 549
- 550 Nelson KL & Smith AL (2010) Determination of capsulation status in *Haemophilus influenzae*
551 by multiplex polymerase chain reaction. *Diagn Microbiol Infect Dis* **66**: 235-240.
- 552
- 553 Nizet V, Colina KF, Almquist JR, Rubens CE & Smith AL (1996) A Virulent Nonencapsulated
554 *Haemophilus influenzae*. *Journal of Infectious Diseases* **173**: 180-186.
- 555
- 556 Osman KL, Jefferies JMC, Woelk CH, *et al.* (2018) Patients with Chronic Obstructive
557 Pulmonary Disease harbour a variation of *Haemophilus* species. *Scientific Reports* **8**: 14734.
- 558
- 559 Othman DSMP, Schirra HJ, McEwan AG & Kappler U (2014) Metabolic versatility in
560 *Haemophilus influenzae*: a metabolomic and genomic analysis. *Frontiers in Microbiology* **doi:**
561 **10.3389/fmicb.2014.00069**.
- 562
- 563 Peruzzotti-Jametti L, Bernstock JD, Vicario N, *et al.* (2018) Macrophage-Derived Extracellular
564 Succinate Licenses Neural Stem Cells to Suppress Chronic Neuroinflammation. *Cell Stem Cell*
565 **22**: 355-368.e313.
- 566
- 567 Pettigrew MM, Ahearn CP, Gent JF, *et al.* (2018) *Haemophilus influenzae* genome evolution
568 during persistence in the human airways in chronic obstructive pulmonary disease. *Proceedings*
569 *of the National Academy of Sciences* **115**: E3256-E3265.
- 570
- 571 Pinto F, Pacheco CC, Ferreira D, Moradas-Ferreira P & Tamagnini P (2012) Selection of
572 Suitable Reference Genes for RT-qPCR Analyses in Cyanobacteria. *PLoS ONE* **7**: e34983.
- 573
- 574 Puig C, Domenech A, Garmendia J, *et al.* (2014) Increased Biofilm Formation by Nontypeable
575 *Haemophilus influenzae* Isolates from Patients with Invasive Disease or Otitis Media versus

- 576 Strains Recovered from Cases of Respiratory Infections. *Applied and Environmental*
 577 *Microbiology* **80**: 7088-7095.
- 578
- 579 Ramakers C, Ruijter JM, Deprez RHL & Moorman AFM (2003) Assumption-free analysis of
 580 quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters* **339**: 62-66.
- 581
- 582 Roier S, Leitner DR, Iwashkiw J, *et al.* (2012) Intranasal Immunization with Nontypeable
 583 *Haemophilus influenzae* Outer Membrane Vesicles Induces Cross-Protective Immunity in
 584 Mice. *PLoS ONE* **7**: e42664.
- 585
- 586 Schembri MA & Klemm P (2001) Biofilm formation in a hydrodynamic environment by novel
 587 fimH variants and ramifications for virulence. *Infection and Immunity* **69**: 1322-1328.
- 588
- 589 Schirra HJ, Anderson CG, Wilson WJ, Kerr L, Craik DJ, Waters MJ & Lichanska AM (2008)
 590 Altered metabolism of growth hormone receptor mutant mice: a combined NMR metabonomics
 591 and microarray study. *PLoS One* **3**: e2764.
- 592
- 593 Sharpe SW, Kuehn MJ & Mason KM (2011) Elicitation of Epithelial Cell-Derived Immune
 594 Effectors by Outer Membrane Vesicles of Nontypeable *Haemophilus influenzae*. *Infection and*
 595 *Immunity* **79**: 4361-4369.
- 596
- 597 Sillanauke, Pönniö & Jääskeläinen (1999) Occurrence of sialic acids in healthy humans and
 598 different disorders. *European Journal of Clinical Investigation* **29**: 413-425.
- 599
- 600 Tatusov RL, Mushegian AR, Bork P, *et al.* (1996) Metabolism and evolution of *Haemophilus*
 601 *influenzae* deduced from a whole-genome comparison with *Escherichia coli*. *Current Biology*
 602 **6**: 279-291.
- 603
- 604 Tsang RSW, Bruce MG, Lem M, Barreto L & Ulanova M (2014) A review of invasive
 605 *Haemophilus influenzae* disease in the Indigenous populations of North America.
 606 *Epidemiology and Infection* **142**: 1344-1354.
- 607
- 608 Tsao D, Nelson KL, Kim D & Smith AL (2012) Infant rat infection modifies phenotypic
 609 properties of an invasive nontypeable *Haemophilus influenzae*. *Microbes and Infection / Institut*
 610 *Pasteur* **14**: 509-516.
- 611

- 612 Ulanova M & Tsang RSW (2014) *Haemophilus influenzae* serotype a as a cause of serious
613 invasive infections. *The Lancet Infectious Diseases* **14**: 70-82.
- 614
- 615 Van Eldere J, Slack MPE, Ladhani S & Cripps AW (2014) Non-typeable *Haemophilus*
616 *influenzae*, an under-recognised pathogen. *The Lancet Infectious Diseases* **14**: 1281-1292.
- 617
- 618 VanWagoner TM, Attack JM, Nelson KL, *et al.* (2016) The modA10 phasevarion of
619 Nontypeable *H. influenzae* R2866 regulates multiple virulence-associated traits. *Microbial*
620 *pathogenesis* **92**: 60-67.
- 621
- 622 Wan Sai Cheong J, Smith H, Heney C, Robson J, Schlebusch S, Fu J & Nourse C (2015) Trends
623 in the epidemiology of invasive *Haemophilus influenzae* disease in Queensland, Australia from
624 2000 to 2013: what is the impact of an increase in invasive non-typable *H. influenzae* (NTHi)?
625 *Epidemiology and Infection* **143**: 2993-3000.
- 626
- 627 Watson ME, Jarisch J & Smith AL (2004) Inactivation of deoxyadenosine methyltransferase
628 (dam) attenuates *Haemophilus influenzae* virulence. *Molecular Microbiology* **53**: 651-664.
- 629
- 630
- 631

Legends to Figures

Figure 1 Genomic and physiological comparison of the invasive *H. influenzae* strains R2866 and C188 with the COPD isolate strain Hi2019. **Panel A:** Venn diagram detailing the number of protein coding genes shared between strains R2866, C188, 2019 and the reference strain RDKW20 **Panel B:** Comparison of *in vitro* biofilm formation in strains in response to changing oxygen tensions for Aerobic (AE) Microaerobic (MA) or Anaerobic (AN) growth. Adjusted p-values were <0.0001 (****, 2way ANOVA) for all comparisons within one growth condition except C188AE-Hi2019AE, adjusted p=0.003,***; not shown in graph for clarity). **Panel C:** Co-infection of 16HBE14 tissue cells with *H. influenzae* for NTHi strains 2019, C188 and R2866. Only 24h data points are shown, strain 2019 was used as a reference strain. Adjusted p-values (2-way ANOVA) :*** p=0.0002 & 0.0006; **** p<0.0001.

Figure 2 Carbon substrate utilization patterns for selected substrates included in the Omnilog PM1 & PM2 phenotypic microarray plates. **A:** Comparison of absorbance endpoints after growth on various substrates for strains Hi2019, R2866 and C188 **B:** Comparison of substrate utilization kinetics of strains R2866 and C188 (grey) vs. Hi2019 (black) demonstrating the stronger substrate utilization in the two blood isolates.

Figure 3 Consumption of major metabolites in NTHi R2866 and C188 in the presence of different oxygen concentrations detected by ¹H-NMR. Concentrations are given in mM.

Figure 4 Production of major metabolites in NTHi R2866 and C188 in the presence of different oxygen concentrations detected by ¹H-NMR. Concentrations are given in mM.

Figure 5 Expression of metabolic genes in NTHi R2866 and C188 in response to changing oxygen tensions. Gene expression data was normalized to expression of *gyrA*, mean expression values and standard deviation of the mean are shown.

Colour scheme: White: genes encoding proteins involved in glucose and pyruvate conversions, light grey: genes encoding respiratory chain enzymes, dark grey: genes encoding enzymes involved in purine and pyrimidine metabolism.

Tables

Table 1 –Major metabolites detected in cultures of *Haemophilus influenzae* strains

R2866 and HiC188 following growth on CDM with glucose under different aeration regimes

Growth Condition		Aerobic		Microaerobic		Anaerobic	
Strain/Medium	CDM	R2866	C188	R2866	C188	R2866	C188
Substrates [mM]							
<i>Glucose</i>	9.71±0.25	0	0.24±0.2	0	0.08±0.08	0	1.26±0.26
<i>Inosine</i>	6.59±0.42	0.06±0.02	0.11±0.17	0.00	0.01±0.007	0.94±0.054	0.002±0.004
<i>Pyruvate</i>	0.6±0.03	0.01±0.001	0.06±0.04	0.02±0.002	0.03±0.01	0.04±0.004	0.05±0.02
Products [mM]							
<i>Acetate</i>	0.08±0.003	16.15±0.58	22.95±0.22	13.59±0.058	19.28±1.65	7.81±0.11	11.04±0.4
<i>Formate</i>	0.08±0.01	0.29±0.22	0.16±0.03	1.97±0.18	4.13±0.18	1.4±0.05	3.15±0.34
<i>Succinate</i>	0	0	0.03±0.01	5.84±0.79	6.78±1.80	15.8±0.36	20.00±0.82
<i>Hypoxanthine</i>	0.58±0.07	5.58±0.14	7.49±0.32	4.03±0.46	6.83±0.24	4.06±0.05	4.57±0.48

Figures

Figure 1

Figure 2

Figure 3

Figure 4

695

696

697

698

699

700

701