



Learning from –omics strategies applied to uncover *Haemophilus influenzae* host-pathogen interactions: Current status and perspectives



Nahikari López-López^{a,1}, Celia Gil-Campillo^{a,1}, Roberto Díez-Martínez^b, Junkal Garmendia^{a,c,*}

^aInstituto de Agrobiotecnología, Consejo Superior de Investigaciones Científicas (IdAB-CSIC)-Gobierno de Navarra, Mutilva, Spain

^bTelum Therapeutics, Noain, Spain

^cCentro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES), Madrid, Spain

ARTICLE INFO

Article history:

Received 10 March 2021

Received in revised form 11 May 2021

Accepted 11 May 2021

Available online 15 May 2021

Keywords:

Haemophilus influenzae

Airway infection

Genome

Transcriptome

Methylome

Proteome

Metabolome

Tn-seq

ABSTRACT

Haemophilus influenzae has contributed to key bacterial genome sequencing hallmarks, as being not only the first bacterium to be genome-sequenced, but also starring the first genome-wide analysis of chromosomes directly transformed with DNA from a divergent genotype, and pioneering Tn-seq methodologies. Over the years, the phenomenal and constantly evolving development of –omic technologies applied to a whole range of biological questions of clinical relevance in the *H. influenzae*-host interplay, has greatly moved forward our understanding of this human-adapted pathogen, responsible for multiple acute and chronic infections of the respiratory tract. In this way, essential genes, virulence factors, pathoadaptive traits, and multi-layer gene expression regulatory networks with both genomic and epigenomic complexity levels are being elucidated. Likewise, the unstoppable increasing whole genome sequencing information underpinning *H. influenzae* great genomic plasticity, mainly when referring to non-encapsulated strains, poses major challenges to understand the genomic basis of clinically relevant phenotypes and even more, to clearly highlight potential targets of clinical interest for diagnostic, therapeutic or vaccine development. We review here how genomic, transcriptomic, proteomic and metabolomic-based approaches are great contributors to our current understanding of the interactions between *H. influenzae* and the human airways, and point possible strategies to maximize their usefulness in the context of biomedical research and clinical needs on this human-adapted bacterial pathogen.

© 2021 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Contents

1. Introduction	3043
2. Constantly evolving information from <i>H. influenzae</i> whole genome sequencing	3043
2.1. <i>H. influenzae</i> population structure	3043
2.2. NTHi isolates from COPD patients: genetic signatures and pathoadaptive traits	3044
2.3. WGS insights into DNA uptake, natural transformation and natural genetic variation by <i>H. influenzae</i>	3045
2.4. Single-molecule real-time (SMRT) sequencing contribution to <i>H. influenzae</i> phase-variable methylome analyses	3045
2.5. Genome-scale microevolution and genetic screening approaches shed light on NTHi adaptation and pathogenesis	3046
3. Unraveling features of the <i>H. influenzae</i> -host interplay by genome-wide gene expression profiling	3046
4. Update on <i>H. influenzae</i> protein and metabolite profiling	3047
5. Future outlook	3048
Funding	3048
Author contributions	3048

* Corresponding author at: Instituto de Agrobiotecnología, Consejo Superior de Investigaciones Científicas (IdAB-CSIC)-Gobierno de Navarra, Mutilva, Spain.

E-mail address: juncal.garmendia@csic.es (J. Garmendia).

¹ These authors contributed equally to this work.

Declaration of Competing Interest 3048
 Acknowledgment 3048
 References 3048

1. Introduction

The Gram negative bacterium *Haemophilus influenzae* is a human-adapted pathogen, commonly associated with human disease in both children and adults. While asymptomatic colonisation begins in the upper airways, it can spread through the respiratory tract, and potentially lead to invasive infections. *H. influenzae* consists of capsulated strains (serotypes a to f), and non-capsulated strains designated as non-typeable (NTHi). The capsulated *H. influenzae* type b (Hib) was the predominating disease-causing serotype, with a high incidence of invasive Hib disease including meningitis and sepsis, until the introduction of the efficient polysaccharide capsule vaccine in the late 1980s/early 1990s, which has driven its almost complete disappearance in countries with established child immunisation programs [1]. In contrast, other *H. influenzae* serotypes and NTHi, which are not targeted by the Hib vaccine, are currently recognised as important causes of infections. In particular, NTHi is a common coloniser of the upper airways in healthy individuals, responsible for multiple acute and chronic infections of the respiratory tract, including otitis media (OM), conjunctivitis, sinusitis and lower respiratory infections in children; exacerbations of chronic obstructive pulmonary disease (COPD) and cystic fibrosis (CF) in adults; and sepsis in neonates, immunocompromised adults, and the elderly [1–5]. Furthermore, *H. influenzae* antimicrobial resistance presents an overall increased trend, in such a way that ampicillin-resistant *H. influenzae* is included in the WHO global priority list of bacteria for which new antibacterial agents are urgently needed [6]. Additionally, fluoroquinolone-resistant *H. influenzae* has quickly increased in recent years and spread worldwide with a variety in epidemiology [7].

The *H. influenzae* strain RdKW20 was the first free-living organism for which the complete genome sequence was established [8], the first naturally transformed organism to be genome-sequenced [9], generated the first genome-scale metabolic model [10,11], and pioneered transposon sequencing methodologies [12]. We further review the contribution of genomic, transcriptomic, proteomic and metabolomic approaches to our understanding of the interactions between *H. influenzae* and the human airways (for a summary, see Fig. 1 and Table 1).

2. Constantly evolving information from *H. influenzae* whole genome sequencing

2.1. *H. influenzae* population structure

On a genomic basis, a whole range of multilocus sequence typing (MLST) and core genome-single nucleotide polymorphisms (core-SNP)-based studies, performed on different *H. influenzae* strain sets encompassing capsulated and non-typeable clinical isolates, commonly share that capsulated strains concentrate into a small number of serotype-specific clusters or clonal populations, while NTHi strains display great genomic diversity [13–18]. About 779 *H. influenzae* genomes are found currently in publicly available databases (May 21st, 2021), and we expect this number to grow

exponentially over time [19], making genomic information dynamic as constantly evolving data requiring regular update. In this context, *H. influenzae* competence in acquiring genomic material throughout the infection process facilitates genomic plasticity (see Section 2.3) leading to an open pan-genome which, for NTHi, was recently updated to 12,249 genes where the core genome accounts for about 7% [20]. The large accessory pan-genome confers diversity and is highly heterogeneous among strains. Many genes are phage-associated, leading to numerous isolates with distinct intact phages integrated in their genome [17,21].

When focusing on the capsule locus, Regions I (containing the *bexABCD* operon) and III (containing the *hcsA* and *hcsB* genes), involved in exporting the capsule polysaccharide, are conserved among all six capsular serotypes, while Region II genes, required for polysaccharide synthesis, are unique to each serotype. Notably, in the frame of transition to whole genome sequencing (WGS)-based surveillance, bioinformatics tools for *in silico* rapid determination of *H. influenzae* serotypes from WGS data have been developed [13,17]. Moreover, capsulated strains genomic closeness has been illustrated for invasive *H. influenzae* type f (Hif) and type a (Hia) isolates [22,23]. Thus, Hif may present a pattern of unique or missing genes maybe increasing the virulence, with the presence of the *sap2* operon, *aef3* fimbriae, and genes for kanamycin nucleotidyltransferase, iron-utilization and putative YadA-like trimeric autotransporters, and the absence of an operon for *de novo* histidine biosynthesis, a *hmg* locus for lipooligosaccharide (LOS) biosynthesis and biofilm formation, and a molybdate transport system [22].

Efforts to define the structure of the NTHi population and to identify groups of related strains led to strain partitioning into six or eight groups or monophyletic clades (clades I-VI / clades I-VIII) based on the core alignment, not predicted on the basis of MLST data, and supported by the composition of the accessory genome [18,20]. The clade I-VI classification allowed proposing clade-specific signatures regarding the exclusive presence/absence of genes encoding surface-associated proteins and LOS components [17,18]. This sorting was further revisited on the clade I-VIII classification by applying pan-genome-wide association studies (GWAS) tools to identify genes enriched in each specific clade [20].

On the other hand, the molecular basis underpinning NTHi transition from commensal to pathogen is not clearly understood. Clade classification did not group strains of common clinical or geographical origin [18,20], and comparative genomics, by analysing paired isogenic strains isolated from the nasopharynx (NP) and bronchoalveolar lavage (BAL) of children with chronic lung disease, did not shed light on this matter as pairs were highly genetically similar [24]. Nevertheless, and despite NTHi heterogeneity, the previously reported association between the NTHi ST14CC-PBP3IIb/A clonal group, increased clinical virulence, antimicrobial resistance and persistence over time, together with NTHi implication in clonal invasive disease, highlights that continuous surveillance and WGS in a clinical laboratory will certainly allow to rapidly address concerns of usually virulent clones or possible outbreaks, particularly in the severe immunocompromised population [25–27].

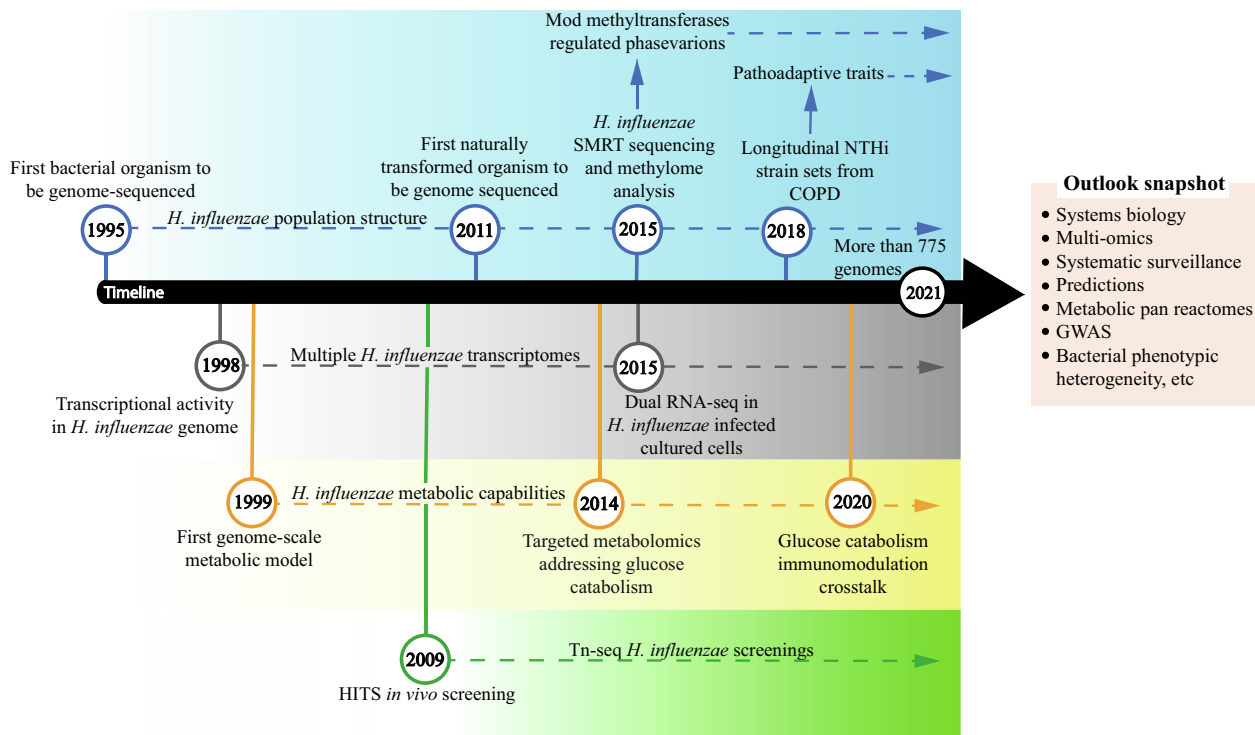


Fig. 1. Summarized landscape of –omics contribution to our understanding of the interactions between *H. influenzae* and the human host. The timeline starts when the first *H. influenzae* complete genome was sequenced, as a major breakthrough on our understanding of the biology, diversity and evolution of bacteria. We highlight key genomic (labeled in blue), transcriptomic (labeled in grey), metabolomic (labeled in yellow), and Tn-seq screening (labeled in green)-based milestones and contributions over time. Information regarding *H. influenzae* population structure, genome-wide expression profiling, metabolic capabilities, and genetic screenings keeps growing along the entire timeline (indicated with dotted arrows). SMRT, single-molecule real-time; NTHi, nontypeable *Haemophilus influenzae*; COPD, chronic obstructive pulmonary disease; dual RNA-seq, dual RNA-sequencing; HITS, high-throughput insertion tracking by deep sequencing; Tn-seq, transposon sequencing; GWAS, genome-wide association studies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Summary of –omics studies contributing to the *H. influenzae*-host interplay.

Type of study	Purpose-outcome of the study	References
A. <i>H. influenzae</i> bacterial whole genome sequencing	A1. Population structure of capsulated and non-capsulated (non-typeable) clinical isolates	[13–18,20–24]
	A2. Within-host pathoadaptive traits on longitudinally collected strain sets from chronic respiratory samples – genome evolution in natural settings	[32–34,36]
	A3. Natural transformation and recombination-driven genetic variation	[9,14,32,38–39,45–46]
	A4. Methylome profiling	[35,54]
	A5. Microevolution in experimental settings	[60–61]
B. Genome-wide genetic screening	B1. HITS* based <i>in vivo</i> genetic screenings	[12,67–68]
	B2. Tn-seq based <i>in vitro</i> genetic screenings	[69–73]
C. Genome-wide gene expression profiling	C1. Bacterial transcriptomic responses to changing environmental conditions	[75–79,82,84,87]
	C2. Bacterial transcriptomic responses to gene inactivation	[35,53–55,59,80–81,83]
	C3. Dual RNA-seq gene expression profiling	[87]
D. Proteomic profiling	D1. Bacterial responses to changing environmental conditions	[89,91–93]
	D2. Bacterial responses to gene inactivation	[54–55,81]
E. Metabolomic profiling	E1. Bacterial responses to changing environmental conditions	[99–101]
	E2. Bacterial responses to gene inactivation	[81,100]

*HITS, high-throughput insertion tracking by deep sequencing; Tn-seq, transposon sequencing; dual RNA-seq, dual RNA-sequencing.

2.2. NTHi isolates from COPD patients: genetic signatures and pathoadaptive traits

COPD is an emerging disease worldwide, and cigarette smoking is its main risk factor. The disease is characterised by dyspnea, chronic cough, sputum production and persistent airflow limitation. COPD exacerbations are often associated to bacterial pathogens, most importantly to NTHi. A main feature of COPD is,

however, a chronic inflammation of the pulmonary parenchyma that is associated with bacterial persistence, and NTHi is a frequently found bacterial species in the airways of these patients [28–31]. When looking for genetic distinctions between NTHi isolates from COPD with respect to other illnesses, the distribution of core genome SNPs does not separate COPD from non-COPD strains [18,20]. However, this distinction is possible when applying discriminant analysis of principal components to the composition of

the accessory genome, further supported by pan-GWAS identification of a repertoire of unique NTHi accessory genes significantly associated with COPD, many of them with predicted roles in virulence, transmembrane transport of metal ions and nutrients, cellular respiration and maintenance of redox homeostasis [20].

Moreover, NTHi also alters its genome during persistence within the lower airways of chronic respiratory patients, as stated by the analyses of longitudinally collected NTHi strains from independent cohorts of COPD patients [32–34]. Pettigrew's group revealed frequent genetic variation due to slipped-strand mispairing in simple sequence repeats (SSR), and diversifying selection of several candidate vaccine antigens during NTHi persistence in the human airways [33]. Follow up studies on this same strain collection showed different proportions of methyltransferase-encoding *modA* alleles compared to those isolated from OM patients [35], and changes in the expression of the IgA protease variants A1, A2, B1 and B2 conferred by changes in genomes during persistent infection which, depending on the variant, could be indels or slipped-strand mispairing in mono- or heptanucleotide repeats [36]. On our part, we addressed molecular genetic changes underlying bacterial pathoadaptation by investigating genetic variants arising from within-patient evolution of NTHi in an independent cohort of COPD patients. Notably, we identified recurrent polymorphisms in several genes and provided experimental evidence for the biological significance of such convergent variation in the *ompP1-fadL* gene, whose recurrent mutations are a likely case of antagonistic pleiotropy during adaptation of NTHi to chronic lung infection associated to COPD [32].

2.3. WGS insights into DNA uptake, natural transformation and natural genetic variation by *H. influenzae*

Natural competence allows bacteria to respond to environmental and nutritional cues by taking up free DNA from their surroundings, gaining both nutrients and genetic information [37]. *H. influenzae* intrinsic transformable nature, self-preferences caused by the uptake machinery bias for uptake signal sequences (USS), and the high rate of homologous recombination events they can undergo, have a great impact on the evolution of the genomes of this pathogen, which seems to be higher in NTHi compared to capsulated strains [38,39].

Such frequent genetic exchange also occurs among *Haemophilus* spp., (*H. influenzae*, *H. aegyptius* and *H. haemolyticus* belong to the cluster “*Haemophilus sensu stricto*” [40,41]), causing frequent misidentifications of *H. influenzae* from clinical specimens when using selective culture methods, a limitation overcome by using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry [42,43]. On this matter, phylogeny studies may also allow not only seeking for *H. influenzae*-specific signatures of diagnostic potential such as the *fucP* locus [16], but also speciating *Haemophilus* isolates from COPD, further supporting that the heterogeneity of NTHi may indeed provide a genetic continuum between NTHi and *H. haemolyticus* [44].

When focusing on bacterial genome shaping by natural transformation, it is noticeable that the first WGS characterization of recombination events was performed by using DNA from a clinical isolate of NTHi (strain 86-028NP) to transform competent cells of a laboratory strain (strain RdKW20), leading to the first genome-wide analysis of chromosomes directly transformed with DNA from a divergent genotype [9,45]. Such observations were independently supported by transformation of the same laboratory strain with donor DNA from a heterologous Hib strain [14]. Detailed genome-wide analyses of not only *H. influenzae* bacteria experimentally transformed with donor DNA from diverged clinical isolates, but also DNA uptake across the outer membrane of naturally competent *H. influenzae* bacteria, turned out to be extre-

mely informative to understand the relevance and evolution of uptake sequences in the genome, and even to predict the outcome of competition between self- and non-self DNA in the respiratory tract environment [39,45]. Moreover, transformation-based experimental evolution studies showed that transformation of single competent cells is extensive, suggesting that it could be used as a tool to map traits that vary between clinical isolates. This notion was applied to develop “transformed recombinant enrichment profiling” (TREP), in which natural transformation is used to generate complex pools of recombinants, phenotypic selection is used to enrich for specific recombinants, and deep sequencing is used to survey for the genetic variation responsible. TREP was applied to investigate the genetic architecture of intracellular epithelial invasion by *H. influenzae*, and identified specific allelic variants of the HMW1 adhesin as a key factor on this matter [46]. Connecting such *in vitro* transformation studies with the high levels of natural genetic variation found in NTHi, variation within closely related genomes from clinical strains collected from COPD respiratory samples shows numerous closely spaced variants which resemble natural transformation events seen in the laboratory, i.e. recombination tracts distinguishing the clinical isolates [32].

2.4. Single-molecule real-time (SMRT) sequencing contribution to *H. influenzae* phase-variable methylome analyses

Formation of C-/N-methyl-cytosine and N-methyl-adenine in bacterial genomes is a postreplicative phenomenon that occurs at specific targets, has multiple roles in bacterial physiology, and can be genome-wide deciphered by using SMRT or nanopore sequencing [47–49]. DNA methyltransferases can exist as part of restriction-modification (R-M) systems and be subject to phase variation, i.e. the random and reversible switching of gene expression due to the presence of inverted repeats (IR) or SSRs in their coding regions. Phase-variable methyltransferases control the expression of multiple genes via epigenetic mechanisms [50–52]. Such systems, called phasevarions for phase-variable regulons, were originally identified in *H. influenzae*, where the DNA methyltransferase ModA is a type III R-M system component containing a variable number of SSRs within its coding sequence [53]. ModA phase variation generates ON/OFF protein variants resulting in changes in methylation patterns, and therefore in phasevarion expression profiles and phenotypic heterogeneity. Besides phase variation, ModA also undergoes allelic variation among *H. influenzae* strains contributing to a wide range of predicted DNA methylation target sequences [50–52]. By analysing *modA* allelic variation across NTHi isolates from healthy individuals, OM and COPD patients, ModA1 to ModA21 variants have been identified, whose distribution and prevalence vary depending on the strain origin, suggesting that different *modA* alleles may provide distinct advantages in the differing human body niches and even have a marker potential [35,54].

Several studies tackled phenotypic heterogeneity due to NTHi ModA2 phase variation, showing that (i) under alkaline conditions such as those in the middle ear during chronic disease, NTHi strains expressing ModA2 (*modA2* ON) form biofilms with greater biomass and less distinct architecture than those formed by a ModA2-deficient population (*modA2* OFF) [55]; (ii) ModA2 ON are more sensitive to oxidative stress than ModA2 OFF variants [56]. Regarding *in vivo* selection during infection of the middle ear in a chinchilla model of OM, *modA2* ON was preferentially selected compared to *modA2* OFF [54], further supported by the fact that a shift from OFF to ON within the middle ear results in increased disease severity [57], and that a unique host immune response may be mounted against each discrete *modA* subpopulation [58].

In addition, coupling of SMRT sequencing with gene and/or protein expression profiling allows combining knowledge of methylation specificity and gene expression changes proportional with methyltransferase phase variation. Thus, genome-wide gene expression profiling showed (i) differences when comparing RdKW20 WT and *modA* mutant strains for genes encoding, among others, surface-exposed proteins, transporters and heat-shock proteins [53]; (ii) differences when comparing R2866 WT and *modA10* mutant strains, including upregulation of membrane protein encoding genes such as *olpA2*, which is involved in host cell adhesion and invasion [59]; (iii) the effects of *modA2* phase variation, including those within biofilms by using an ON/OFF strain pair [54,55]; (iv) the effects of *modA15* and *modA18* phase variation, by using a single ON/OFF pair of strains for each *modA* allele. In the case of the *ModA15* phase variation, the DNA binding transcriptional regulator Fis was up-regulated in ON relative to OFF, as a likely case of regulation of a regulator [35]. Conversely, proteomic analyses (i) determined relative protein abundances within the biofilms formed by each *modA2* subpopulation, with the greatest differences occurring in alkaline conditions including a significant decrease in abundance of the DNABII DNA-binding protein HU within biofilms formed by *modA2* ON compared to those formed by *modA2* OFF bacteria [55], and (ii) showed differences related to changes in the expression of outer membrane proteins when comparing *modA* ON/OFF strain pairs [54].

2.5. Genome-scale microevolution and genetic screening approaches shed light on NTHi adaptation and pathogenesis

Unravelling pathoadaptive genome evolution related to chronicity in the natural setting is not only thrilling but also challenging, as it requires longitudinal and preferably prospective clinical isolate sampling [32,33]. This notion, when applied to well-suited experimental settings, also allows unravelling microevolution traits contributing to the pathogen persistence. This is the case of the mutation of the *icc* gene, encoding a 3',5'-cyclic adenosine monophosphate phosphodiesterase, identified as a *H. influenzae* microevolution trait in response to transient nutrient limitation, which associates with increased development of intracellular bacterial communities (IBC) [60]. Likewise, when using a pre-clinical model of OM to assess NTHi pathoadaptation during sequential episodes of disease, microevolution of haemoglobin binding and LOS biosynthesis genes was observed in such OM-adapted strains, which in turn promoted increased biofilm formation, inflammation, stromal fibrosis, and an increased propensity to form IBCs [61].

On the other hand, sequencing and data analysis tools have also revolutionized genome-wide gene discovery aimed to determine gene disruptions differentially represented in a mutant population upon screening. Historical limitations by the labour involved in mapping the locations of the mutations, are overcome by using random transposition insertion mutagenesis followed by deep sequencing. This is achievable by high-throughput insertion tracking by deep sequencing (HITS), insertion sequencing (IN-seq), transposon sequencing (Tn-seq), or transposon-directed insertion sequencing (TraDIS), collectively referred to as Tn-seq [62–66]. *H. influenzae* pioneered genome-wide gene discovery by starring the original HITS proof of concept, developed to analyse genes required by *H. influenzae* to resist clearance from the lung in a murine pulmonary model, and identified a whole range of genes with reported or potential roles in survival during nutrient limitation, oxidative stress, and exposure to antimicrobial membrane perturbations [12]. HITS was also useful to identify *H. influenzae* genes required to colonize an influenza A virus (IAV) co-infection murine model, showing genes both exclusive or commonly shared by bacterial and IAV co-infection model systems [67], or to screen

genome-wide genetic interactions to unravel epistatic interaction accounting for *H. influenzae* survival in the murine lung [68]. Tn-seq suitability has also been reported for *H. influenzae* *in vitro* genome-wide screening studies including the identification of (i) essential genes [69]; (ii) the carbonic anhydrase involvement in NTHi adaptation to changes in environmental CO₂ levels and intracellular survival [70]; (iii) a role for the outer membrane protein P5 in NTHi survival to complement-mediated killing [71], and for galactose-containing oligosaccharide structures in NTHi survival to complement-dependent neutrophil-mediated killing [72]; (iv) the contribution of the *mltC* and *lppB* genes to biofilm formation [73].

3. Unraveling features of the *H. influenzae*-host interplay by genome-wide gene expression profiling

Since the first high-density oligonucleotide probe array containing probes representing 106 *H. influenzae* genes [74], microarray and later on RNA sequencing (RNA-seq) technologies have greatly contributed to broaden our understanding of NTHi physiology and host interplay. Please note that profiling of gene expression changes regulated by methyltransferases and their phase variation is reviewed in Section 2.4. Moreover, *H. influenzae* transits between niches within the host that differ in oxygen levels, and the ArcAB two-component regulatory system controls gene expression in response to different respiratory conditions of growth. Gene expression profiling contributed not only to identify the *H. influenzae* ArcA regulon upon anaerobic growth [75], but also to decipher changes in *H. influenzae* transcriptional responses upon changes in oxygen availability, showing that, in high oxygen, arginine uptake, arginine/aspartate metabolism, and glutamate pathways leading into the Krebs cycle are commonly up-regulated [76].

Genome-wide expression profiling also identified a putative core of genes responsive to iron and heme (FeHm) availability in several NTHi strain backgrounds including RdKW20, 10810, R2866, 86-028NP or R2846. Included in the core iron/heme module were genes preferentially expressed under iron/heme limitation, most of which are directly involved with iron and/or heme acquisition [77–79]. The ferric uptake regulator (Fur) regulon has also been studied, which identified the first small RNA described in any *Haemophilus* species, HrrF, overexpressed in the absence of Fur and responsive to iron levels, in turn regulating the expression of molybdate uptake, and deoxyribonucleotide and amino acid synthesis genes [80]. As mentioned above, transient nutrient limitation, in particular FeHm restriction, increases the longevity of NTHi survival *in vitro* and promotes bacterial microevolution by mutation of the *icc* gene [60]. Such adaptive trait, besides increased development of IBCs [60], also relates to increased biofilm formation and gene expression changes within the *in vitro* biofilms, some of them associated to competence and therefore likely linked to the observed increased transformation efficiency of those biofilms [81].

Regarding natural competence, changes in gene expression during competence development confirmed the existence of a competence regulon characterized by a promoter-associated competence regulatory element (CRE) closely related to the cAMP receptor protein (CRP) binding consensus, and where the essential competence gene *sxy* is induced early in competence development [82]. Such observations were further extended by differential gene expression profiling of *H. influenzae* wild-type, *crp* and *sxy* mutant strains, which also shed light on the competence-regulated toxin-antitoxin ToxTA system [83]. Moreover, analysis of global transcriptional changes has also been of use when assessing *H. influenzae* response to antibiotic pressure, showing increased *ponB* (encoding PBP1b) and *acrR* (negative regulator of AcrAB-TolC efflux

pump) gene expression during heat stress, which may relate to the observed decrease in bacterial viability after incubation with imipenem at 42 °C as compared to 37 °C [84]. An effort to identify convergent molecular signatures related to lung adaptation led to comparative transcriptomic analysis of paired, isogenic NTHi strains, isolated from the NP and BAL of children with chronic lung disease, showing no convergence at the gene level, but a possible trend in terms of functional enrichment among genetically unrelated NTHi strains [24].

Although most studies have analyzed *H. influenzae* gene expression, host response to bacterial infection also needs obvious attention. Gene expression profiling identified host functions differentially expressed in NTHi-infected human type II pneumocytes, which allowed identifying a repertoire of host target candidates for pharmacological modulation [85]. In an independent study, the transcriptome of a complete episode of acute OM in mice identified sets of genes involved in regulation of immune responses, changes in epithelial and stromal cell markers, and the recruitment/function of neutrophils and macrophages [86]. Notably, a dual RNA-seq approach was undertaken for pathogen and host genome-wide expression profiling during *H. influenzae* infection of ciliated human bronchial epithelial cells. Temporal profiling of host mRNA signatures revealed significant dysregulation of the target cell cytoskeleton elicited by bacterial infection, with a profound effect on the intermediate filament network and junctional complexes. In turn, NTHi downregulated its central metabolism, and increased the expression of transporters, suggesting a change in the metabolic regime due to the availability of host substrates (see Section 4), and of stress-induced defense mechanisms, including the transport of exogenous glutathione and activation of toxin-antitoxin family components [87]. Currently lacking, *in vivo* dual RNA-seq would likely be of great help to gather a more comprehensive view of the *H. influenzae*-host interplay.

4. Update on *H. influenzae* protein and metabolite profiling

Proteomic studies on the *H. influenzae*-host interaction are relatively scattered. From a diagnostic perspective, *H. influenzae* species-unique peptide biomarkers discovered by tandem mass spectrometry may be a valuable tool for clinical sample proteotyping [88]. Protein profiling regulated by methyltransferase phase variation is reviewed in Section 2.4. Independently, proteomic expression profiling of *H. influenzae* grown in pooled human sputum revealed increased expression of antioxidant, stress-response proteins, and cofactor and nutrient uptake systems compared to media grown cells [89]. Definition of the protein content of NTHi lysates identified numerous unique *H. influenzae* proteins contributing to biofilm formation, immune evasion, or epithelial inflammation, classified as outer membrane/cell surface associated, metabolic, biosynthesis mediators, proteases, chaperones/DNA binding proteins, transporters, reductases or hydrolases [90]. Proteomic profiling was also used to demonstrate differential expression of the NTHi biofilm to planktonic samples, where ArcA showed a high level of downregulation in the biofilm [91], and whose inactivation led to ArcA-regulated proteomic changes, some of them maybe involved in serum susceptibility [92]. Proteomic profiling was also useful to prove that the two bacterial populations released from a NTHi biofilm by using antibodies directed against the type IV pilus or, alternatively, against a DNABII DNA-binding protein, are different not only from planktonically grown NTHi, but also from each other despite genetic identity. Even more, each newly released population had a distinct, increased susceptibility to antibiotic killing compared to planktonic bacteria, highlighting new clues for anti-biofilm therapeutics [93].

From the host side, proteomic and metabolomic signatures of NTHi-induced acute OM in a chinchilla model were delineated in infected middle ear tissue lysates, revealing that establishment of disease coincides with actin morphogenesis, suppression of inflammatory mediators, and bacterial aerobic respiration. As a first step toward identification of clinically meaningful metabolic biomarkers, decreased biogenic amines and sphingomyelin molecules, increased taurine, glutamine and ornithine were observed upon infection [94].

Elucidating the metabolic determinants in the lung during respiratory infection is likely to be key to develop therapeutics modulating symptom and/or disease severity [95]. Lung metabolomics have not been performed on the *H. influenzae*-human airways interplay, but currently available insights on bacterial metabolism may guide us when assessing this infectious process. The *H. influenzae* RdKW20 strain drove the first genome-scale metabolic model. Besides assessing basic structural features of the *H. influenzae* metabolic network, such model facilitated addressing minimal substrates requirements for the network to allow biomass production. Minimal requirements included fructose, a likely preferred carbon source for which a phosphotransferase system (PTS) exists in *H. influenzae* [96,97], but fructose could also be replaced by other carbon sources including glucose [10,11]. *H. influenzae* possesses complete glycolysis and pentose phosphate pathways for glucose catabolism, lacks most enzymes of the oxidative branch of the Krebs cycle, and holds a respiratory chain with several dehydrogenases transferring electrons into the menaquinone pool, and terminal reductases transferring the electrons to a variety of electron acceptors, altogether driving a so-called glucose respiration-assisted fermentation, where acetate is the main end-product under aerobic growth. When oxygen is limiting, acetate is still produced but other products arise including formate and succinate, at variable rates depending on the tested strains [98–101]. End-product excretion profiling is informative not only from the bacterial perspective, but also due to their possible immunomodulatory roles. Although lung nutritional immunity keeps low glucose levels in the airway surface liquid of healthy individuals, respiratory disease generates a glucose-rich niche favorable for pathogens able to use glucose as carbon source. As this being the case for *H. influenzae*, glucose catabolism enhances bacterial growth, but also promotes the release of end-products acting as pro-inflammatory metabolites, more specifically acetate, which may contribute to lung colonization and inflammation in chronic respiratory patients [100]. Besides the above described genomic flexibility (see Section 2), *H. influenzae* metabolic versatility among strains has also been reported, in terms of substrate utilization, histidine synthesis or urease activity, suggesting metabolic adaptive traits favoring differential access to throat, ear, lower airways or blood niches [101–104].

Finally, and considering that *H. influenzae* metabolic machinery adapts to the host-imposed milieu, transcriptome signatures upon cell infection showed that the genes encoding the bacterial biosynthesis pathways of carbohydrates, lipids, amino acids, nucleotides, and energy metabolism were mostly downregulated, although the shikimate pathway for precursor chorismate production and the consequent tryptophan biosynthesis pathway was upregulated after initial bacterial-cell contact [87]. Increased tryptophan biosynthetic pathway was also observed in biofilms formed by strain variants adapted to transient FeHm restriction where, moreover, amino acids and enzymes associated with central metabolism, and metabolites associated with the urea cycle were also increased [81]. Linked to the metabolic machinery, attention should also be paid to bacterial metabolite transport systems. In fact, lactate, arginine, tryptophan, cysteine/glutathione, thiamine, thiamine pyrophosphate, iron and polyamines uptake and transport systems were highly expressed during *H. influenzae* cell infec-

tion [87]. Likewise, increased production of serine, sialic acid and tryptophan transporters was observed in biofilms formed by strain variants adapted to transient FeHm restriction [81], together suggesting that changes in central metabolism combined with increased nutrient stores may help counteracting nutritional immunity at the colonizing niches.

5. Future outlook

The last twenty five years, since the *H. influenzae* RdKW20 strain was genome sequenced, have gone a long way, greatly propelled by the unstoppable –omics technologies. As a human-adapted pathogen, integrative systems biology approaches to the interactions between *H. influenzae* and the human airways, are instrumental in the discovery of specific bacterial recognition, host signal transduction or immune tolerance, together with multi-omics approaches to decipher the genetic, immunologic, (post)transcriptional, (post)translational, and metabolic mechanisms underlying the outcome of such interactions. Those approaches will be enormously enriched by bacterial single-cell resolution methodologies such as microfluidics, or single-cell sorting, microscopy and genome/RNA sequencing, to allow overlooking the remarkable bacterial phenotypic heterogeneity, and shed light on complex phenotypes such as antibiotic persistence or metabolic specialization. Guidance of such integrative work by the human specificity of the infection will provide meaningful information on *H. influenzae* patho-adaptive evolution within the human host. Furthermore, we are collecting a wealth of genomic and epigenomic data, and this will continue to grow with the introduction of routine sequencing for disease surveillance. These data will be of great use to predict how changes in genome sequence lead to different host-pathogen interaction outcomes by applying machine learning approaches, or to build genome-scale metabolic network reconstructions assessing core and pan reactomes. Also, combination of deep bacterial phenotyping and pan-GWAS methodologies will allow not only predicting bacterial genotype-phenotype associations, but also deciphering genetic determinants of resistance to antimicrobials, virulence or niche adaptation. Altogether, therapeutic/vaccination target and biomarker discovery will be moved forward, to be eventually translated to the clinical settings and reach end users.

Funding

N.L.-L. is funded by a PhD studentship from Regional Navarra Govern, Spain, reference 0011-1408-2017-00000. C.G.-C. is funded by a PhD studentship from Agencia Española de Investigación (AEI), Spain, reference PRE2019-088382. This work has been funded by grants from MINECO RTI2018-096369-B-I00, 875/2019 from SEPAR, and PC150-151–152 from Gobierno de Navarra to J. G. CIBER is an initiative from Instituto de Salud Carlos III (ISCIII), Madrid, Spain.

Author contributions

Conceptualization: J.G.; writing (original draft preparation): N. L.-L. C.G.-C. and J.G.; writing (review and editing): all authors; supervision: J.G.; funding acquisition: J.G.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We are grateful to the Garmendia's lab for manuscript critical reading.

References

- [1] Jalalvand F, Riesbeck K. Update on non-typeable *Haemophilus influenzae*-mediated disease and vaccine development. *Expert Rev Vaccines* 2018;17:503–12. <https://doi.org/10.1080/14760584.2018.1484286>.
- [2] Ladhani S, Slack MPE, Heath PT, von Gottberg A, Chandra M, Ramsay ME. Invasive *Haemophilus influenzae* disease, Europe, 1996–2006. *Emerg Infect Dis* 2010;16:455–63. <https://doi.org/10.3201/eid1603.090290>.
- [3] Ahearn CP, Gallo MC, Murphy TF. Insights on persistent airway infection by non-typeable *Haemophilus influenzae* in chronic obstructive pulmonary disease. *Pathog Dis* 2017;75:1–18. <https://doi.org/10.1093/femsdp/ftx042>.
- [4] Duell BL, Su YC, Riesbeck K. Host–pathogen interactions of nontypeable *Haemophilus influenzae*: from commensal to pathogen. *FEBS Lett* 2016;590:3840–53. <https://doi.org/10.1002/1873-3468.12351>.
- [5] Su YC, Jalalvand F, Thegerström J, Riesbeck K. The interplay between immune response and bacterial infection in COPD: focus upon non-typeable *Haemophilus influenzae*. *Front Immunol* 2018;9:1–26. <https://doi.org/10.3389/fimmu.2018.02530>.
- [6] Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect Dis* 2018;18:318–27. [https://doi.org/10.1016/S1473-3099\(17\)30753-3](https://doi.org/10.1016/S1473-3099(17)30753-3).
- [7] Wen S, Feng D, Chen D, Yang L, Xu Z. Molecular epidemiology and evolution of *Haemophilus influenzae*. *Infect Genet Evol* 2020;80. <https://doi.org/10.1016/j.meegid.2020.104205>.
- [8] Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* (80–) 1995;269:496–512. doi:10.1126/science.7542800.
- [9] Mell JC, Shumilina S, Hall IM, Redfield RJ. Transformation of natural genetic variation into *Haemophilus influenzae* genomes. *PLoS Pathog* 2011;7:32–4. <https://doi.org/10.1371/journal.ppat.1002151>.
- [10] Edwards JS, Pálsson BO. Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *J Biol Chem* 1999;274:17410–6. <https://doi.org/10.1074/jbc.274.25.17410>.
- [11] Schilling CH, Pálsson B. Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis. *J Theor Biol* 2000;203:249–83. <https://doi.org/10.1006/jtbi.2000.1088>.
- [12] Gawronski JD, Wong SMS, Giannoukos G, Ward DV, Akerley BJ. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung. *Proc Natl Acad Sci U S A* 2009;106:16422–7. <https://doi.org/10.1073/pnas.0906627106>.
- [13] Potts CC, Topaz N, Rodríguez-Rivera LD, Hu F, Chang HY, Whaley MJ, et al. Genomic characterization of *Haemophilus influenzae*: a focus on the capsule locus. *BMC Genomics* 2019;20:1–9. <https://doi.org/10.1186/s12864-019-6145-8>.
- [14] Power PM, Bentley SD, Parkhill J, Moxon ER, Hood DW. Investigations into genome diversity of *Haemophilus influenzae* using whole genome sequencing of clinical isolates and laboratory transformants. *BMC Microbiol* 2012;12:1. <https://doi.org/10.1186/1471-2180-12-273>.
- [15] Staples M, Graham RMA, Jennison AV. Characterisation of invasive clinical *Haemophilus influenzae* isolates in Queensland, Australia using whole-genome sequencing. *Epidemiol Infect* 2017;145:1727–36. <https://doi.org/10.1017/S0950268817000450>.
- [16] Price EP, Sarovich DS, Nosworthy E, Beissbarth J, Marsh RL, Pickering J, et al. *Haemophilus influenzae*: using comparative genomics to accurately identify a highly recombinogenic human pathogen. *BMC Genomics* 2015;16:1–10. <https://doi.org/10.1186/s12864-015-1857-x>.
- [17] Pinto M, González-Díaz A, Machado MP, Duarte S, Vieira L, Carriço JA, et al. Insights into the population structure and pan-genome of *Haemophilus influenzae*. *Infect Genet Evol* 2019;67:126–35. <https://doi.org/10.1016/j.meegid.2018.10.025>.
- [18] De Chiara M, Hood D, Muzzi A, Pickard DJ, Perkins T, Pizza M, et al. Genome sequencing of disease and carriage isolates of nontypeable *Haemophilus influenzae* identifies discrete population structure. *Proc Natl Acad Sci U S A* 2014;111:5439–44. <https://doi.org/10.1073/pnas.1403353111>.
- [19] Koonin EV, Makarova KS, Wolf YI. Evolution of microbial genomics: conceptual shifts over a quarter century. *Trends Microbiol* 2021;1–11. <https://doi.org/10.1016/j.tim.2021.01.005>.
- [20] Rajendra KC, Leong KWC, Harkness NM, Lachowicz J, Gautam SS, Cooley LA, et al. Whole-genome analyses reveal gene content differences between nontypeable *Haemophilus influenzae* isolates from chronic obstructive pulmonary disease compared to other clinical phenotypes. *Microb Genomics* 2020;6:1–12. <https://doi.org/10.1099/mgen.0.000405>.
- [21] Tanaka E, Wajima T, Nakaminami H, Noguchi N. Whole-genome sequence of *Haemophilus influenzae* ST422 outbreak clone strain 2018–Y40 with low quinolone susceptibility isolated from a paediatric patient. *J Glob Antimicrob Resist* 2020;22:759–61. <https://doi.org/10.1016/j.jgar.2020.06.024>.

- [22] Su YC, Resman F, Hörhold F, Riesbeck K. Comparative genomic analysis reveals distinct genotypic features of the emerging pathogen *Haemophilus influenzae* type f. *BMC Genomics* 2014;15. <https://doi.org/10.1186/1471-2164-15-38>.
- [23] Tsang RSW, Shuel M, Ahmad T, Hayden K, Knox N, Van Domselaar G, et al. Whole genome sequencing to study the phylogenetic structure of serotype a *Haemophilus influenzae* recovered from patients in Canada. *Can J Microbiol* 2020;66:99–110. <https://doi.org/10.1139/cjm-2019-0406>.
- [24] Aziz A, Sarovich DS, Nosworthy E, Beissbarth J, Chang AB, Smith-Vaughan H, et al. Molecular signatures of nontypeable *Haemophilus influenzae* lung adaptation in pediatric chronic lung disease. *Front Microbiol* 2019;10. <https://doi.org/10.3389/fmicb.2019.01622>.
- [25] Månsson V, Skaare D, Riesbeck K, Resman F. The spread and clinical impact of ST14CC-PBP3 type IIb/A, a clonal group of non-typeable *Haemophilus influenzae* with chromosomally mediated β -lactam resistance—a prospective observational study. *Clin Microbiol Infect* 2017;23:209.e1–7. <https://doi.org/10.1016/j.cmi.2016.11.006>.
- [26] Andersson M, Resman F, Eitrem R, Drobni P, Riesbeck K, Kahlmeter G, et al. Outbreak of a β -lactam resistant non-typeable *Haemophilus influenzae* sequence type 14 associated with severe clinical outcomes. *BMC Infect Dis* 2015;15. <https://doi.org/10.1186/s12879-015-1319-8>.
- [27] John MA, Beukers A, Chan R, van Hal S. Whole genome sequencing identifies opportunistic non-typeable *Haemophilus influenzae* rather than a hypervirulent clone. *Pathology* 2020;10–1. <https://doi.org/10.1016/j.pathol.2020.08.011>.
- [28] Ritchie AI, Wedzicha JA. Definition, causes, pathogenesis, and consequences of chronic obstructive pulmonary disease exacerbations. *Clin Chest Med* 2020;41:421–38. <https://doi.org/10.1016/j.ccm.2020.06.007>.
- [29] Celli BR, Wedzicha JA. Update on clinical aspects of chronic obstructive pulmonary disease. *N Engl J Med* 2019;381:1257–66. <https://doi.org/10.1056/NEJMr1900500>.
- [30] Hurst JR, Vestbo J, Anzueto A, Locantore N, Müllerova H, Tal-Singer R, et al. Susceptibility to exacerbation in chronic obstructive pulmonary disease. *N Engl J Med* 2010;363:1128–38. <https://doi.org/10.1056/NEJMoa0909883>.
- [31] Wedzicha JA, Seemungal TAR. COPD exacerbations: defining their cause and prevention. *Lancet* (London, England) 2007;370:786–96. [https://doi.org/10.1016/S0140-6736\(07\)161382-8](https://doi.org/10.1016/S0140-6736(07)161382-8).
- [32] Moleres J, Fernandez-Calvet A, Ehrlich RL, Marti S, Perez-Regidor L, Euba B, et al. Antagonistic pleiotropy in the bifunctional surface protein FadL (OmpP1) during adaptation of *Haemophilus influenzae* to chronic lung infection associated with chronic obstructive pulmonary disease. *MBio* 2018;9. <https://doi.org/10.1128/mBio.01176-18>.
- [33] Pettigrew MM, Ahearn CP, Gent JF, Kong Y, Gallo MC, Munro JB, et al. *Haemophilus influenzae* genome evolution during persistence in the human airways in chronic obstructive pulmonary disease. *Proc Natl Acad Sci U S A* 2018;115:E3256–65. <https://doi.org/10.1073/pnas.1719654115>.
- [34] Garmendia J, Viadas C, Calatayud L, Mell JC, Marti-Llitas P, Euba B, et al. Characterization of nontypeable *Haemophilus influenzae* isolates recovered from adult patients with underlying chronic lung disease reveals genotypic and phenotypic traits associated with persistent infection. *PLoS ONE* 2014;9. <https://doi.org/10.1371/journal.pone.0097020>.
- [35] Attack JM, Murphy TF, Pettigrew MM, Seib KL, Jennings MP. Non-typeable *Haemophilus influenzae* isolates from patients with chronic obstructive pulmonary disease contain new phase-variable *modA* methyltransferase alleles controlling phasevariations. *Sci Rep* 2019;9:1–10. <https://doi.org/10.1038/s41598-019-52429-6>.
- [36] Gallo MC, Kirkham C, Eng S, Bewawee RS, Kong Y, Pettigrew MM, et al. Changes in IgA protease expression are conferred by changes in genomes during persistent infection by nontypeable *Haemophilus influenzae* in chronic obstructive pulmonary disease. *Infect Immun* 2018;86:1–11. <https://doi.org/10.1128/IAI.00313-18>.
- [37] Mell JC, Redfield RJ. Natural competence and the evolution of DNA uptake specificity. *J Bacteriol* 2014;196:1471–83. <https://doi.org/10.1128/JB.01293-13>.
- [38] Connor TR, Corander J, Hanage WP. Population subdivision and the detection of recombination in non-typeable *Haemophilus influenzae*. *Microbiol (United Kingdom)* 2012;158:2958–64. <https://doi.org/10.1099/mic.0.063073-0>.
- [39] Mora M, Mell JC, Ehrlich GD, Ehrlich RL, Redfield RJ. Genome-wide analysis of DNA uptake across the outer membrane of naturally competent *Haemophilus influenzae*. *iScience* 2021;24. <https://doi.org/10.1016/j.isci.2020.102007>.
- [40] Nørskov-Lauritsen N, Bruun B, Kilian M. Multilocus sequence phylogenetic study of the genus *Haemophilus* with description of *Haemophilus pittmaniae* sp. nov. *Int J Syst Evol Microbiol* 2005;55:449–56. <https://doi.org/10.1099/ijso.0.63325-0>.
- [41] Hedegaard J, Okkels H, Bruun B, Kilian M, Mortensen KK, Nørskov-Lauritsen N. Phylogeny of the genus *Haemophilus* as determined by comparison of partial *infB* sequences. *Microbiology* 2001;147:2599–609. <https://doi.org/10.1099/00221287-147-9-2599>.
- [42] Marti S, Puig C, de la Campa AG, Tirado-Velez JM, Tubau F, Domenech A, et al. Identification of *Haemophilus haemolyticus* in clinical samples and characterization of their mechanisms of antimicrobial resistance. *J Antimicrob Chemother* 2016;71:80–4. <https://doi.org/10.1093/jac/dkv307>.
- [43] Frickmann H, Christner M, Donat M, Berger A, Essig A, Podbielski A, et al. Rapid discrimination of *Haemophilus influenzae*, *H. parainfluenzae*, and *H. haemolyticus* by fluorescence *in situ* hybridization (FISH) and two matrix-assisted laser-desorption-ionization time-of-flight mass spectrometry (MALDI-TOF-MS) Platforms. *PLoS ONE* 2013;8. <https://doi.org/10.1371/journal.pone.0063222>.
- [44] Osman KL, Jefferies JMC, Woelk CH, Devos N, Pascal TG, Mortier MC, et al. Patients with chronic obstructive pulmonary disease harbour a variation of *Haemophilus* species. *Sci Rep* 2018;8:1–11. <https://doi.org/10.1038/s41598-018-32973-3>.
- [45] Mell JC, Lee JY, Firme M, Sinha S, Redfie RJ. Extensive cotransformation of natural variation into chromosomes of naturally competent *Haemophilus influenzae*. *G3 Genes Genomes, Genet* 2014;4:717–31. <https://doi.org/10.1534/g3.113.009597>.
- [46] Mell JC, Viadas C, Moleres J, Sinha S, Fernández-Calvet A, Porsch EA, et al. Transformed recombinant enrichment profiling rapidly identifies HMW1 as an intracellular invasion locus in *Haemophilus influenzae*. *PLoS Pathog* 2016;12. <https://doi.org/10.1371/journal.ppat.1005576>.
- [47] Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods* 2010;7:461–5. <https://doi.org/10.1038/nmeth.1459>.
- [48] Sánchez-Romero MA, Cota I, Casadesús J. DNA methylation in bacteria: from the methyl group to the methylome. *Curr Opin Microbiol* 2015;25:9–16. <https://doi.org/10.1016/j.mib.2015.03.004>.
- [49] Sánchez-Romero MA, Casadesús J. The bacterial epigenome. *Nat Rev Microbiol* 2020;18:7–20. <https://doi.org/10.1038/s41579-019-0286-2>.
- [50] Attack JM, Tan A, Bakaletz LO, Jennings MP, Seib KL. Phasevariations of bacterial pathogens: methylomics sheds new light on old enemies. *Trends Microbiol* 2018;26:715–26. <https://doi.org/10.1016/j.tim.2018.01.008>.
- [51] Phillips ZN, Husna AU, Jennings MP, Seib KL, Attack JM. Phasevariations of bacterial pathogens—phase-variable epigenetic regulators evolving from restriction-modification systems. *Microbiol (United Kingdom)* 2019;165:917–28. <https://doi.org/10.1099/mic.0.000805>.
- [52] Seib KL, Srihanta YN, Attack JM, Jennings MP. Epigenetic regulation of virulence and immunoevasion by phase-variable restriction-modification systems in bacterial pathogens. *Annu Rev Microbiol* 2020;74:655–71. <https://doi.org/10.1146/annurev-micro-090817-062346>.
- [53] Srihanta YN, Maguire TL, Stacey KJ, Grimmond SM, Jennings MP. The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. *Proc Natl Acad Sci U S A* 2005;102:5547–51. <https://doi.org/10.1073/pnas.0501169102>.
- [54] Attack JM, Srihanta YN, Fox KL, Jurcisek JA, Brockman KL, Clark TA, et al. A biphasic epigenetic switch controls immunoevasion, virulence and niche adaptation in non-typeable *Haemophilus influenzae*. *Nat Commun* 2015;6. <https://doi.org/10.1038/ncomms58828>.
- [55] Brockman KL, Azzari PN, Branstool MT, Attack JM, Schulz BL, Jen FEC, et al. Epigenetic regulation alters biofilm architecture and composition in multiple clinical isolates of nontypeable *Haemophilus influenzae*. *MBio* 2018;9:1–15. <https://doi.org/10.1128/mBio.01682-18>.
- [56] Brockman KL, Branstool MT, Attack JM, Robledo-Avila F, Partida-Sanchez S, Jennings MP, et al. The *ModA2* phasevarion of nontypeable *Haemophilus influenzae* regulates resistance to oxidative stress and killing by human neutrophils. *Sci Rep* 2017;7:1–11. <https://doi.org/10.1038/s41598-017-03552-9>.
- [57] Brockman KL, Jurcisek JA, Attack JM, Srihanta YN, Jennings MP, Bakaletz LO. *ModA2* phasevarion switching in nontypeable *Haemophilus influenzae* increases the severity of experimental otitis media. *J Infect Dis* 2016;214:817–24. <https://doi.org/10.1093/infdis/jiw243>.
- [58] Robledo-Avila FH, Ruiz-Rosadode JD, Partida-Sanchez S, Brockman KL. A Bacterial epigenetic switch in non-typeable *Haemophilus influenzae* modifies host immune response during otitis media. *Front Cell Infect Microbiol* 2020;10:512743. <https://doi.org/10.3389/fcimb.2020.512743>.
- [59] VanWagoner TM, Attack JM, Nelson KL, Smith HK, Fox KL, Jennings MP, et al. The *modA10* phasevarion of nontypeable *Haemophilus influenzae* R2866 regulates multiple virulence-associated traits. *Microb Pathog* 2016;92:60–7. <https://doi.org/10.1016/j.micpath.2015.12.006>.
- [60] Hardison RL, Harrison A, Wallace RM, Heimlich DR, O'Bryan ME, Sebra RP, et al. Microevolution in response to transient heme-iron restriction enhances intracellular bacterial community development and persistence. *PLoS Pathog* 2018;14:1–25. <https://doi.org/10.1371/journal.ppat.1007355>.
- [61] Harrison A, Hardison RL, Fullen AR, Wallace RM, Gordon DM, White P, et al. Continuous microevolution accelerates disease progression during sequential episodes of infection. *Cell Rep* 2020;30:2978–2988.e3. <https://doi.org/10.1016/j.celrep.2020.02.019>.
- [62] van Opijnen T, Bodi KL, Camilli A. Tn-seq: High-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods* 2009;6:767–72. <https://doi.org/10.1038/nmeth.1377>.
- [63] Van Opijnen T, Camilli A. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nat Rev Microbiol* 2013;11:435–42. <https://doi.org/10.1038/nrmicro3033>.
- [64] Cain AK, Barquist L, Goodman AL, Paulsen IT, Parkhill J, van Opijnen T. A decade of advances in transposon-insertion sequencing. *Nat Rev Genet* 2020;21:526–40. <https://doi.org/10.1038/s41576-020-0244-x>.
- [65] Burby PE, Nye TM, Schroeder JW, Simmons LA. Implementation and data analysis of Tn-seq, whole-genome resequencing, and single-molecule real-time sequencing for bacterial genetics. *J Bacteriol* 2017;199:1–11. <https://doi.org/10.1128/JB.00560-16>.

- [66] Wong SMS, Akerley BJ. Genome-scale approaches to identify genes essential for *Haemophilus influenzae* pathogenesis. *Front Cell Infect Microbiol* 2012;2:23. <https://doi.org/10.3389/fcimb.2012.00023>.
- [67] Wong SM, Bernui M, Shen H, Akerley BJ. Genome-wide fitness profiling reveals adaptations required by *Haemophilus* in coinfection with influenza A virus in the murine lung. *Proc Natl Acad Sci U S A* 2013;110:15413–8. <https://doi.org/10.1073/pnas.1311217110>.
- [68] Wong SM, Jackson MD, Akerley BJ. Suppression of alternative lipooligosaccharide glycosyltransferase activity by UDP-galactose epimerase enhances murine lung infection and evasion of serum IgM. *Front Cell Infect Microbiol* 2019;9:160. <https://doi.org/10.3389/fcimb.2019.00160>.
- [69] Mobegi FM, van Hijum SAFT, Burghout P, Bootsma HJ, de Vries SPW, van der Gaast-de Jongh CE, et al. From microbial gene essentiality to novel antimicrobial drug targets. *BMC Genomics* 2014;15. 10.1186/1471-2164-15-958.
- [70] Langereis JD, Zomer A, Stunnenberg HG, Burghout P, Hermans PWM. Nontypeable *Haemophilus influenzae* carbonic anhydrase is important for environmental and intracellular survival. *J Bacteriol* 2013;195:2737–46. <https://doi.org/10.1128/JB.01870-12>.
- [71] Langereis JD, de Jonge MI, Weiser JN. Binding of human factor H to outer membrane protein P5 of non-typeable *Haemophilus influenzae* contributes to complement resistance. *Mol Microbiol* 2014;94:89–106. <https://doi.org/10.1111/mmi.12741>.
- [72] Langereis JD, Weiser JN. Shielding of a lipooligosaccharide IgM epitope allows evasion of neutrophil-mediated killing of an invasive strain of nontypeable *Haemophilus influenzae*. *MBio* 2014;5:1–10. <https://doi.org/10.1128/mBio.01478-14>.
- [73] Marti S, Puig C, Merlos A, Viñas M, de Jonge MI, Linares J, et al. Bacterial lysis through interference with peptidoglycan synthesis increases biofilm formation by nontypeable *Haemophilus influenzae*. *MSphere* 2017;2:1–12. <https://doi.org/10.1128/msphere.00329-16>.
- [74] De Saizieu A, Certa U, Warrington J, Gray C, Keck W, Mous J. Bacterial transcript imaging by hybridization of total RNA to oligonucleotide arrays. *Nat Biotechnol* 1998;16:45–8. <https://doi.org/10.1038/nbt0198-45>.
- [75] Wong SMS, Alugupalli KR, Ram S, Akerley BJ. The ArcA regulon and oxidative stress resistance in *Haemophilus influenzae*. *Mol Microbiol* 2007;64:1375–90. <https://doi.org/10.1111/j.1365-2958.2007.05747.x>.
- [76] Jiang D, Tikhomirova A, Kidd SP. *Haemophilus influenzae* strains possess variations in the global transcriptional profile in response to oxygen levels and this influences sensitivity to environmental stresses. *Res Microbiol* 2016;167:13–9. <https://doi.org/10.1016/j.resmic.2015.08.004>.
- [77] Whitby PW, VanWagoner TM, Seale TW, Morton DJ, Stull TL. Transcriptional profile of *Haemophilus influenzae*: effects of iron and heme. *J Bacteriol* 2006;188:5640–5. <https://doi.org/10.1128/JB.00417-06>.
- [78] Whitby PW, Seale TW, VanWagoner TM, Morton DJ, Stull TL. The iron/heme regulated genes of *Haemophilus influenzae*: comparative transcriptional profiling as a tool to define the species core modulon. *BMC Genomics* 2009;10:1–19. <https://doi.org/10.1186/1471-2164-10-6>.
- [79] Whitby PW, VanWagoner TM, Seale TW, Morton DJ, Stull TL. Comparison of transcription of the *Haemophilus influenzae* iron/heme modulon genes *in vitro* and *in vivo* in the chinchilla middle ear. *BMC Genomics* 2013;14. <https://doi.org/10.1186/1471-2164-14-925>.
- [80] Santana EA, Harrison A, Zhang X, Baker BD, Kelly BJ, White P, et al. HrrF is the Fur-regulated small RNA in nontypeable *Haemophilus influenzae*. *PLoS One* 2014;9. <https://doi.org/10.1371/journal.pone.0105644>.
- [81] Harrison A, Hardison RL, Wallace RM, Fitch J, Heimlich DR, Bryan MO, et al. Reprioritization of biofilm metabolism is associated with nutrient adaptation and long-term survival of *Haemophilus influenzae*. *Npj Biofilms Microbiomes* 2019;5. <https://doi.org/10.1038/s41522-019-0105-6>.
- [82] Redfield RJ, Cameron ADS, Qian Q, Hinds J, Ali TR, Kroll JS, et al. A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenzae*. *J Mol Biol* 2005;347:735–47. <https://doi.org/10.1016/j.jmb.2005.01.012>.
- [83] Black HF, Mastromatteo S, Sinha S, Ehrlich RL, Nislow C, Mell JC, et al. A competence-regulated toxin-antitoxin system in *Haemophilus influenzae*. *PLoS ONE* 2020;15:1–23. <https://doi.org/10.1371/journal.pone.0217255>.
- [84] Cherkaoui A, Diene SM, Fischer A, Leo S, François P, Schrenzel J. Transcriptional modulation of penicillin-binding protein 1b, outer membrane protein P2 and efflux pump (AcrAB-TolC) during heat stress is correlated to enhanced bactericidal action of imipenem on non-typeable *Haemophilus influenzae*. *Front Microbiol* 2018;8:1–12. <https://doi.org/10.3389/fmicb.2017.02676>.
- [85] Euba B, Moleres J, Segura V, Vidas C, Morey P, Moranta D, et al. Genome expression profiling-based identification and administration efficacy of host-directed antimicrobial drugs against respiratory infection by nontypeable *Haemophilus influenzae*. *Antimicrob Agents Chemother* 2015;59:7581–92. <https://doi.org/10.1128/AAC.01278-15>.
- [86] Hernandez M, Leichtle A, Pak K, Webster NJ, Wasserman SI, Ryan AF. The transcriptome of a complete episode of acute otitis media. *BMC Genomics* 2015;16:1–16. <https://doi.org/10.1186/s12864-015-1475-7>.
- [87] Baddal B, Muzzi A, Censini S, Calogero RA, Torricelli G, Guidotti S, et al. Dual RNA-seq of nontypeable *Haemophilus influenzae* and host cell transcriptomes reveals novel insights into host-pathogen cross talk. *MBio* 2015;6:1–14. <https://doi.org/10.1128/mBio.01765-15>.
- [88] Karlsson R, Thorsell A, Gomila M, Salvà-Serra F, Jakobsson HE, Gonzales-Siles L, et al. Discovery of species-unique peptide biomarkers of bacterial pathogens by tandem mass spectrometry-based proteotyping. *Mol Cell Proteomics* 2020;19:518–28. <https://doi.org/10.1074/mcp.RA119.001667>.
- [89] Qu J, Lesse AJ, Brauer AL, Cao J, Gill SR, Murphy TF. Proteomic expression profiling of *Haemophilus influenzae* grown in pooled human sputum from adults with chronic obstructive pulmonary disease reveal antioxidant and stress responses. *BMC Microbiol* 2010;10. <https://doi.org/10.1186/1471-2180-10-162>.
- [90] Preciado D, Poley M, Tsai S, Tomney A, Brown K, Val S. A proteomic characterization of NTHi lysates. *Int J Pediatr Otorhinolaryngol* 2016;80:8–16. <https://doi.org/10.1016/j.ijporl.2015.11.016>.
- [91] Post DMB, Held JM, Ketterer MR, Phillips NJ, Sahu A, Apicella MA, et al. Comparative analyses of proteins from *Haemophilus influenzae* biofilm and planktonic populations using metabolic labeling and mass spectrometry. *BMC Microbiol* 2014;14:329. <https://doi.org/10.1186/s12866-014-0329-9>.
- [92] De Souza-Hart JA, Blackstock W, Di Modugno V, Holland IB, Kok M. Two-component systems in *Haemophilus influenzae*: a regulatory role for ArcA in serum resistance. *Infect Immun* 2003;71:163–72. <https://doi.org/10.1128/iai.71.1.163-172.2003>.
- [93] Mokran EM, Ahearn CP, Buzzo JR, Novotny LA, Zhang Y, Goodman SD, et al. Nontypeable *Haemophilus influenzae* newly released (NReI) from biofilms by antibody-mediated dispersal versus antibody-mediated disruption are phenotypically distinct. *Biofilm* 2020;2. <https://doi.org/10.1016/j.biofilm.2020.100039>.
- [94] Harrison A, Dubois LG, John-Williams LS, Moseley MA, Hardison RL, Heimlich DR, et al. Comprehensive proteomic and metabolomic signatures of nontypeable *Haemophilus influenzae*-induced acute otitis media reveal bacterial aerobic respiration in an immunosuppressed environment. *Mol Cell Proteomics* 2016;15:1117–38. <https://doi.org/10.1074/mcp.M115.052498>.
- [95] Bernatchez JA, McCall LI. Insights gained into respiratory infection pathogenesis using lung tissue metabolomics. *PLoS Pathog* 2020;16:1–8. <https://doi.org/10.1371/journal.ppat.1008662>.
- [96] Macfadyen LP, Redfield RJ. Life in mucus: Sugar metabolism in *Haemophilus influenzae*. *Res Microbiol* 1996;147:541–51. [https://doi.org/10.1016/0923-2508\(96\)84010-1](https://doi.org/10.1016/0923-2508(96)84010-1).
- [97] Macfadyen LP, Dorocicz IR, Reizer J, Saler MH, Redfield RJ. Regulation of competence development and sugar utilization in *Haemophilus influenzae* Rd by a phosphoenolpyruvate:fructose phosphotransferase system. *Mol Microbiol* 1996;21:941–52. <https://doi.org/10.1046/j.1365-2958.1996.441420.x>.
- [98] Raghunathan A, Price ND, Galperin MY, Makarova KS, Purvine S, Picone AF, et al. *In silico* metabolic model and protein expression of *Haemophilus influenzae* strain Rd KW20 in rich medium. *Omi A J Integr Biol* 2004;8:25–41. <https://doi.org/10.1089/153623104773547471>.
- [99] Othman DSMP, Schirra H, McEwan AG, Kappeler U. Metabolic versatility in *Haemophilus influenzae*: a metabolomic and genomic analysis. *Front Microbiol* 2014;5:1–10. <https://doi.org/10.3389/fmicb.2014.00069>.
- [100] López-López N, Euba B, Hill J, Dhoui R, Caballero L, Leiva J, et al. *Haemophilus influenzae* glucose catabolism leading to production of the immunometabolite acetate has a key contribution to the host airway-pathogen interplay. *ACS Infect Dis* 2020;6:406–21. <https://doi.org/10.1021/acscinfdis.9b00359>.
- [101] Muda NM, Nasreen M, Dhoui R, Hosmer J, Hill J, Mahawar M, et al. Metabolic analyses reveal common adaptations in two invasive *Haemophilus influenzae* strains. *Pathog Dis* 2019;77:1–12. <https://doi.org/10.1093/femspd/ftz015>.
- [102] Juliao PC, Marrs CF, Xie J, Gilsdorf JR. Histidine auxotrophy in commensal and disease-causing nontypeable *Haemophilus influenzae*. *J Bacteriol* 2007;189:4994–5001. <https://doi.org/10.1128/JB.00146-07>.
- [103] Zhang L, Patel M, Xie J, Davis GS, Marrs CF, Gilsdorf JR. Urease operon and urease activity in commensal and disease-causing nontypeable *Haemophilus influenzae*. *J Clin Microbiol* 2013;51:653–5. <https://doi.org/10.1128/JCM.03145-12>.
- [104] Murphy TF, Brauer AL. Expression of urease by *Haemophilus influenzae* during human respiratory tract infection and role in survival in an acid environment. *BMC Microbiol* 2011;11. <https://doi.org/10.1186/1471-2180-11-183>.