Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Niche-specific metabolic signatures in non-typeable Haemophilus influenzae

Keshab Bhattarai
Research Center Borstel, Leibniz Lung Center, Member of German Center for Lung Research (DZL), Airway Research Center North

Bikash Baral
University of Helsinki

Aleksandra Sarnowicz
Research Center Borstel, Leibniz Lung Center, Member of German Center for Lung Research (DZL), Airway Research Center North

Margo Diricks
Research Center Borstel, Leibniz Lung Center, University Hospital Schleswig-Holstein/ Campus Lübeck

Stefan Niemann
Research Center Borstel, Leibniz Lung Center, German Center for Infection Research (DZIF), Hamburg-Lübeck-Borstel-Riems

Jan Rupp
University Hospital Schleswig Holstein/ Campus Lübeck, German Center for Infection Research (DZIF), Hamburg-Lübeck-Borstel-Riems

Katarzyna A. Duda
kduda@fz-borstel.de

Research Center Borstel, Leibniz Lung Center, Member of German Center for Lung Research (DZL), Airway Research Center North

Article

Keywords: Microbiome, Airways, Chronic Lung Diseases, Metabolomics, Niche-association, Genomics

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11523740/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Niche-specific metabolic signatures in non-typeable *Haemophilus influenzae*

Keshab Bhattarai[1†], Bikash Baral[2†], Aleksandra Sarnowicz[1], Margo Diricks[3,5], Stefan Niemann[3,6,7,8], Jan Rupp[4,5,6], Katarzyna A. Duda[1*]

[1]Research Group of Biofunctional Metabolites and Structures, Research Center Borstel, Leibniz Lung Center, Member of German Center for Lung Research (DZL), Airway Research Center North, Borstel, Germany

[2]Organismal and Evolutionary Biology Research Programme, University of Helsinki, FI-00014 Helsinki, Finland

[3]Research Group of Molecular and Experimental Mycobacteriology, Research Center Borstel, Leibniz Lung Center, Borstel, Germany

[4]Institute of Medical Microbiology, University Hospital Schleswig Holstein/ Campus Lübeck, Lübeck, Germany

[5]Infectious Disease Clinic, University Hospital Schleswig-Holstein/ Campus Lübeck, Lübeck, Germany

[6]German Center for Infection Research (DZIF), Hamburg-Lübeck-Borstel-Riems, Germany

[7]Laboratoire Biologie Intégrative des Populations, Evolution Moléculaire, EPHE, PSL University, Paris, France.

[8]Institut de Systématique, Evolution, Biodiversité, ISYEB, Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France.

*** Correspondence:**
Corresponding Author, kduda@fz-borstel.de

[†]These authors contributed equally to this work.

Abstract

Non-typeable *Haemophilus influenzae* (NTHi) is an opportunistic pathogen associated with chronic respiratory diseases. NTHi's metabolic diversity facilitates its survival within a broad spectrum of environmental conditions within the host, making deeper research into NTHi metabolism a promising avenue for novel therapies against NTHi-associated lung diseases. Draft genome sequences of nine NTHi clinical strains from three isolation sites - ear (ear sample, ES), pharynx (pharynx sample, PS), and lower respiratory tract (Lung) - were analyzed. Pathway module coverage was computed and summarized per group for significant annotated metabolites. Metabolites were characterized by mass spectrometry and compared statistically among groups. Gene content across the tested NTHi strains was largely conserved, with limited core-genome SNP variation. Despite minor nucleotide differences, all strains exhibited analogous metabolic genes and capacities. Conversely, metabolomics data revealed differential metabolic profiles among the body-site isolates. In a principal component analysis, the ES group was significantly separated from both the PS and Lung groups, which overlapped considerably. A number of significant differences between the examined groups were detected for inosine, hypoxanthine, niacinamide, isoleucine, phenylalanine, pyridoxine, uracil and valin. This study provides novel insights into the metabolic differences associated with NTHi inhabiting diverse host niches.

Keywords: Microbiome, Airways, Chronic Lung Diseases, Metabolomics, Niche-association, Genomics

1 Introduction

Haemophilus influenzae is a human-restricted, Gram-negative coccobacillus that commonly colonizes the nasopharynx asymptomatically but can also cause disease throughout the respiratory tract. The genome of *H. influenzae* Rd (KW20) was the first complete genome sequenced from a free-living organism, a landmark achievement that ushered in modern bacterial genomics [1]. Strains lacking a polysaccharide capsule are termed non-typeable *H. influenzae* (NTHi) and predominate in mucosal infections, including acute otitis media, pneumonia, sinusitis, conjunctivitis, and exacerbations of chronic lung diseases [2]. NTHi exhibit a genetic strategy distinct from that of encapsulated lineages, which carry a capsule locus defining serotypes a–f [3]. NTHi remain a leading cause of otitis media in children, underscoring their continuing clinical importance [4]. The rise of β-

lactamase-negative ampicillin-resistant (BLNAR) strains has further complicated empirical treatment and patient care due to phenotypic resistance [5].

NTHi strains are genetically diverse, naturally competent and strongly enriched with short "uptake signal sequences" that bias the DNA acquisition and accelerate genome remodelling [6,7]. Analyses of hundreds of genomes show that disease phenotype is only weakly predicted by core-genome multilocus genotyping-based phylogeny or accessory gene content, consistent with frequent microevolution and regulatory switching during niche transitions [8]. Key adhesins (HMW1/HMW2 or the trimeric autotransporter Hia) and IgA1 proteases vary among strains and are dynamically regulated during persistence, shaping adherence, invasion, biofilm development, and immune evasion in airway niches[9,10]. However, the genomic potential underlying primary and secondary metabolic signatures across multiple NTHi strains remains underexplored.

NTHi strains have associated their metabolism to the human airway. They are auxotrophic for heme (X factor) and nicotinamide adenine dinucleotide (NAD, V factor), scavenging these cofactors from host sources through multiple outer-membrane and periplasmic systems [11]. A global transcriptomic and metabolic study reveals that NTHi strains perform "respiration-assisted fermentation," flexibly blending fermentative pathways with electron transport to manage redox stress under microaerobic or anaerobic conditions, which are typical of mucosal biofilms and inflamed airways [12]. During long-term heme-iron restriction, NTHi strain rewire central metabolism and enhance biofilm traits that support survival[13].

Recent mass spectrometry studies have characterized the small-molecule and lipid repertoire of *H. influenzae*, revealing strain-specific phospholipid patterns and previously unannotated metabolites, such as the antimicrobial compound cyclo(Leu-Pro). However, these studies rarely incorporate information on clinical niches [14]. Other studies have identified metabolic profiles involving acetate, formate, succinate, lactate, and hypoxanthine that vary with substrate availability (glucose, inosine, and pyruvate) and aeration conditions in two NTHi strains associated with invasive disease and one associated with chronic lung disease (CLD) [12,15].

Despite extensive genomic research and decades of detailed work on heme/iron acquisition and adhesins, we still lack a clear, systems-level understanding of how association with clinical sites (the ear, pharynx, and lower airways) shapes NTHi metabolism at the level of measurable metabolites. Large comparative genomic studies have shown limited ability to distinguish strains associated with

otitis media, pneumonia, CLD, and other clinical phenotypes based on sequence alone, suggesting that regulatory and metabolic plasticity, rather than fixed gene content, drives niche specialization [8]. NTHi strains have repeatedly been shown to adapt to airway conditions, including hemoglobin limitation, hypoxia, and nitrosative stress, by reprioritizing central metabolism and biofilm physiology. These findings derive largely from transcriptomic studies, single-strain models, multi-isolate metabolome analyses, and niche-aware metabolomics [12,13]. However, comparative datasets linking these metabolite-level features to isolation sites across diverse clinical strains remain lacking [16].

We hypothesized that the association of NTHi with different body sites corresponds to reproducible metabolic "fingerprints" that are not apparent from genome content alone. By integrating whole-genome variation with metabolite profiles, we aimed to identify site-specific metabolic features. This niche-resolved, metabolite-centric perspective complements previous genomic surveys and could reveal actionable metabolic vulnerabilities for diagnosis, prevention, or therapy across the spectrum of NTHi-associated diseases [8,13].

2 Results

2.1 Description of the bacterial strains

Nine clinical NTHi strains previously isolated from patients with otitis media (n=3; H104, H149, and H213) or pneumonia (n=6; H117, H158, H182, H188, H189, and H212) were analyzed [17]. The isolates originated from ear smear (H104 and H213), bronchial secretion (H188 and H212), bronchoalveolar lavage (H117 and H158), tracheal secretion (H182) and pharyngeal smear (H149 and H189). In this study, we categorized all the strains into three major site-specific groups: ear sample (ES), pharyngeal sample (PS), and Lung. The Lung group comprises all the strains isolated from the lower respiratory tract. Importantly, the strains belong to different phylogenetic clades: H149 (PS) belongs to NTHi clade I; H117 (Lung), H158 (Lung) and H212 (Lung) belong to NTHi clade IV; H104 (ES) and H182 (Lung) belong to NTHi clade V; H188 (Lung) and H189 (PS) belong to NTHi clade VIa; and H213 (ES) belongs to NTHi clade VIb (Supplementary Fig. S1 and Supplementary Fig. S2).

2.2 Near-complete genomes with conserved core resistome

Genome assemblies showed high completeness, with BUSCO completeness ranging from 98.2% to 99.4%, fragmented BUSCOs from 0 to 0.8%, and missing BUSCOs from 0 to 1.6% (Supplementary Table S1, Sheet 1). QUAST analysis identified 29–62 contigs per genome, genome sizes of 1.77–1.98 Mb, N50 values of 77,175–359,254 bp, and GC contents of 37.93–38.56% (Supplementary Table S1_Sheet 2). Gene prediction identified 1,750–2,053 genes per genome. Together, these metrics indicated high-quality, near-complete genome assemblies with broadly similar genomic characteristics across the isolates.

Analysis of antimicrobial resistance genes (ARGs) identified a conserved core resistome. The *hmrM* MATE-family efflux pump and the global regulator CRP were detected across all isolates, suggesting a shared basal resistance-associated repertoire (Supplementary Table S1_Sheet 3). FastANI analysis confirmed species-level assignment, with the isolates showing highest similarity to *H. influenzae* NCTC 8143 (maximum ANI, 98.86% for H104; Supplementary Table S1_Sheet 4).

2.3 Conserved core genome and variable accessory genome

Despite their common species identity, the isolates exhibited variation in genome architecture. Protein-coding gene numbers ranged from 1,688 to 1,965, while rRNA and tRNA numbers ranged from 7–19 and 50–58, respectively. Core-genome SNP analysis identified 8–116 pairwise SNP differences, indicating measurable genomic variation among the isolates (Fig. 1a; Supplementary Tables S1 and S2). Comparative genome visualization and phylogenomic analysis further demonstrated the relationships among the isolates (Supplementary Tables S2 and S3). However, given the limited number of isolates and their uneven phylogenetic representation, these data should not be interpreted as capturing the full genomic diversity of NTHi.

Roary-based pangenome analysis identified 4,563 gene clusters across the nine isolates. Of these, 838 were core genes, present in 99-100% of isolates, whereas no soft-core genes were identified. The accessory genome comprised 1,668 shell genes, present in 15–<95% of isolates, and 2,057 cloud genes, present in <15% of isolates (Supplementary Table S1_Sheets 7-9). Thus, the NTHi isolates shared a substantial conserved genomic backbone while retaining a large and variable accessory gene repertoire.

Functional annotation using the RAST SEED framework showed that several major cellular functions were highly conserved, including RNA metabolism, cell division, fatty-acid metabolism, potassium metabolism, and secondary metabolism (Fig. 1b; Supplementary Table S1, Sheet 6). RNA

metabolism, for example, comprised 37–38 genes across isolates, whereas amino-acid metabolism showed greater variation, with 156–181 genes. Variation was also observed in membrane transport, nucleotide metabolism, regulation and signaling, and DNA metabolism. H188 displayed the highest number of membrane transporter genes and also contained comparatively more phage-associated and mobile genetic elements. Virulence-associated genes ranged from 20 to 24 per genome, while iron-acquisition genes were present at similar levels across isolates. These observations indicate conservation of major cellular functions alongside strain-specific differences in accessory functions.

2.4 Deprived biosynthetic potential of secondary metabolites

Given the ecological roles of secondary metabolites in competition, defence, and communication, we first assessed the genomic capacity for secondary metabolite production in the tested NTHi strains by analysing their whole-genome sequences using antiSMASH. Surprisingly, this analysis identified only one secondary metabolite biosynthetic gene cluster, encoding an azole-containing ribosomal peptide, which was conserved across all strains (Supplementary Table S5). This conserved cluster showed no similarity to any known biosynthetic gene clusters or ribosomal peptides.

2.5 Metabolomic profiling revealed diverse primary metabolites

To determine whether the conserved genomic metabolic capacity was reflected in metabolite production, cell-free culture supernatants from the nine NTHi strains grown in chemically defined medium were subjected to metabolomic profiling. MetaboScape analysis initially detected 221 features across 48 samples in positive-ion mode. Following removal of features associated with blanks, early-eluting signals, or the absence of MS2 spectra, 146 features remained. Of these, 38 showed evidence of production or consumption by at least one NTHi strain and were retained for subsequent analysis (Supplementary Table S6).

Eighteen of the 38 features could be annotated using MS^2-spectral library with MS/MS score above 900 (Supplementary Fig. S4), where inosine, hypoxanthine, niacinamide (also known as nicotinamide), uracil and isoleucine were annotated for multiple features.

2.6 Statistically distinct metabolic signatures by isolation site

Thirty-eight samples (comprising nine NTHi strains and media control) were subjected to interactive principal component analysis and linear modelling using MetaboAnalyst 6.0. A synchronized 3D PCA plot (PC1: 44.4%, PC2: 15.0%, PC3: 10.3%) demonstrated that ES samples were separated to

the upper left from the PS and Lung samples while visualized on metadata-overlay (Supplementary Fig. S5). A 2D score plot with 44.4% PC1 (x-axis), 15% PC2 (y-axis), F-value: 10.82, R-squared: 0.48841 and *p*-value (based on 999 permutations): 0.001 indicated that ES still edged to the upper left region as compared to overlapped PS and Lung groups while visualizing on 95% confidence regions (Fig. 2).

Linear modelling with a *P*-value cut-off of 0.05 (False Discovery Rate) identified 28 significant metabolites that varied in site-specific groups: ES, PS, and Lung. The model demonstrated that there were 23, 22, and 2 significant metabolites that very between the ES vs Lung, ES vs PS, and Lung vs PS groups, respectively. The list of significant metabolites varied in different groups as indicated by their respective logFC (fold change) and *P*-values are given in Table 1 and Supplementary Table S7(a-c). Out of these metabolites, eight were annotated against MS^2-spectral libraries: inosine, hypoxanthine, uracil, niacinamide, phenylalanine, isoleucine, pyridoxin and valine (Fig. 3). The most significantly differed metabolite in ES vs Lung and ES vs PS is the unknown compound "1.45_136.06189". This metabolite along with "1.15_136.06189" gave a closer hit with adenine while annotating with reduced threshold of MS/MS score (700-800). Furthermore, inosine, hypoxanthine, niacinamide and uracil have been identified as major, significantly varied metabolites in both the ES vs. Lung and the ES vs. PS. Phenylalanine and isoleucine were only significantly differed metabolites in ES vs Lung, while pyridoxin and valine were only significantly differed metabolites in ES vs PS. Phenylalanine and an unknown metabolite, "1.13_104.10697", were the only compounds that significantly differed between the Lung and PS groups. Box plots of significant metabolites were generated in R (tidyverse package) after averaging repeatedly annotated metabolites to assess abundance across site-specific groups (Fig. 4).

Collectively, NTHi isolates originating from different anatomical sites exhibited distinct extracellular metabolite profiles despite broadly conserved genomic metabolic capacity.

2.7 Conserved metabolic genes despite heterogeneous metabolomes

Given the significant differences observed in metabolites across isolation sites, a subsequent investigation was conducted to ascertain whether these variations were attributable to disparities in metabolic pathways among the isolates. KEGG/KAAS analysis showed that genes associated with these metabolites were highly conserved across all nine NTHi genomes. These pathways included purine metabolism, nicotinate and nicotinamide metabolism, pyrimidine metabolism, phenylalanine

biosynthesis, phenylalanine metabolism, valine-leucine-isoleucine biosynthesis, valine-leucine-isoleucine degradation, and vitamin B6 metabolism. Extracted KEGG orthology (KO) entries which resemble a defined enzyme activity, protein complex, or functional subunit, accompanied by each site-specific group (unified from individual strains) per pathway showed more than 90% identical distribution for purine metabolism, pyrimidine metabolism, and nicotinate and nicotinamide metabolism, while 100% identical distribution for the rest of the pathways (Table 2). Detailed description and complete strain-level KO distributions are provided in Supplementary information and Table S8.

3 Discussion

NTHi strains are clinically significant opportunistic pathogens, particularly in vulnerable patients with respiratory tract infections. Because NTHi lacks capsular polysaccharides, they are not targeted by the Hib vaccine and have emerged as dominant agents in invasive *H. influenzae* disease. Their metabolic versatility is thought to support adaptation to diverse host microenvironments. We therefore hypothesized that isolates from different anatomical sites (ear, pharynx, and lower respiratory tract) exhibit niche-associated metabolic profiles. In this study, we integrated whole-genome sequencing with LC/MS-based metabolomics to assess whether individual as well as niche-specific metabolic phenotypes could be inferred directly from genomic data. We clustered all lower respiratory isolates (tracheal, bronchial, and BAL samples) into one Lung group, as they share similar microbial, metabolic, and abiotic environments compared to ear and pharyngeal isolates[18].

Our pangenome and functional analyses revealed a highly conserved genomic backbone, particularly for essential processes such as RNA metabolism, fatty acid biosynthesis, and cell division [19,20]. Roary identified 4,563 gene clusters comprising 838 core genes, 1,668 shell genes, and 2,057 cloud genes, with no soft-core genes. It is evident from the analysis presented in this study that NTHi strains exhibit conserved genetic capacities while retaining substantial accessory genomic diversity. This diversity has the potential to contribute to strain-level phenotypic variation. As the reference strain 86-028NP was derived from an otitis media isolate, its utilisation as a solitary reference may engender strain- or lineage-specific bias. Consequently, reference-dependent observations were interpreted as comparative rather than representative of the broader NTHi population [21,22]. Variation in amino acid and carbohydrate metabolism and membrane transport further suggested niche-associated adaptation [23]. Notably, the PS isolate H149 and lung isolate H188 showed enhanced glycan degradation capacity, potentially supporting mucosal colonization[24–27]. H188

also contained more mobile elements, consistent with increased genomic plasticity and potential for horizontal gene transfer[28–34].

Subsequently, the analysis was shifted to focus on secondary metabolism. In silico results revealed that all tested NTHi strains lacked the biosynthetic potential to produce secondary metabolites, with the exception of one unknown ribosomal peptide. Secondary metabolites primarily fulfil ecological functions, including competition with other microbes, survival, defence, and microbial communication. This limited secondary-metabolite potential may be indicative of adaptation to an ecological niche in which extensive chemical competition or defence is less critical [54,55]. Consequently, the metabolites detected by LC-MS were predominantly primary metabolites associated with cellular growth and metabolism. The use of biological and technical replicates ensured the robustness of the comparisons of metabolite abundance across groups.

A notable finding was the elevated amount of inosine and hypoxanthine in the samples of the Lung and PS strains as compared to ES strains. As inosine was a constituent of the culture medium, its depletion in ES cultures indicates that ES strains exhibit a more extensive utilisation of inosine compared to PS and Lung strains. Inosine and hypoxanthine are central intermediates of purine metabolism. In humans, inosine has anti-inflammatory effects in allergic lung inflammation and can regulate cytokine generation. It exerts protective effects, reducing the levels of pro-inflammatory cytokines such as tumour necrosis factor-α, interleukin-1β, interleukin-6, and macrophage inflammatory protein-2, reducing leukocyte recruitment into lung tissue [56,57]. Hypoxanthine, a product of inosine metabolism, is a significant biomarker in COPD patients, with elevated levels in serum compared to healthy controls [58]. Elevated hypoxanthine and other adenyl-purine products have also been observed in airway BAL fluid of inflamed lungs, particularly in cystic fibrosis preschoolers with neutrophil inflammation [59,60]. Across human BAL fluid and animal pneumonia models, purine catabolites (inosine and hypoxanthine), NAD-axis metabolites and free amino acids strongly track neutrophilic inflammation and structural lung disease[59–61]. Interestingly, *in vivo* RNA-seq analysis of lungs has revealed up-regulation of NTHi *de novo* purine biosynthesis and fitness dependence as compared to wild type NTHi strains during murine lung infection model. The release of purine breakdown products by PS and Lung strains is consistent with substantial purine salvage in airway-associated strains, contributing to inflammation[62]. Trace hypoxanthine detected in control medium likely resulted from inosine degradation during LC-MS analysis.

Niacinamide, a precursor and product associated with NAD^+, plays a crucial role in nicotinate and nicotinamide metabolism. Notably, all strains produced niacinamide, but the ES group showed significantly higher levels. Functionally, niacinamide is involved in energy processes, DNA repair, redox reactions, and cellular stress mechanisms. Previous studies have demonstrated that niacinamide acts as an antimicrobial compound against bacteria and fungi by arresting the microbial cell cycle[63]. Furthermore, elevated niacinamide and NAD^+ levels have been observed in BAL fluid from early cystic fibrosis patients during neutrophilic inflammation[60]. Niacinamide has also been shown to suppress lung injury and proinflammatory mediator accumulation by inhibiting MAPK and AKT/NF-κB signaling pathways[64]. *H. influenzae* is unable to synthesize NAD and relies on NAD/nicotinamide riboside/nicotinamide mononucleotide salvage, the detection of niacinamide in the supernatant is compatible with NAD turnover by these pathways[65–68]. However, the biological significance of the observed site-associated niacinamide differences remains unclear and requires validation under *in vivo* conditions.

Given the presence of isoleucine, phenylalanine, pyridoxine, and uracil in the culture medium allowed the assessment of its differential uptake by tested NTHi strains.

Phenylalanine, an essential aromatic amino acid, is produced by many bacteria and is one of the nine essential amino acids in the human diet [75]. In cystic fibrosis, free amino acids, including phenylalanine, valine, and isoleucine, increase in BAL fluid with neutrophil abundance, consistent with host proteolysis[60]. Furthermore, elevated phenylalanine levels in lung and blood samples have been shown to induce inflammation by enhancing the innate immune response and releasing proinflammatory cytokines. This can also promote pyroptosis of alveolar macrophages, exacerbating lung inflammation and acute respiratory distress syndrome lethality[76]. However, in contrast to published data, our study found that all tested strains metabolised phenylalanine, with most strains isolated from Lung.

Pyridoxine (vitamin B6) was utilized by all strains, but the PS strains showed the highest utilization compared to ES and Lung strains. In bacteria, pyridoxine plays a crucial role as a coenzyme in various reactions, primarily involved in amino acid metabolism. Additionally, pyridoxine exhibits potent antioxidant activity, effectively neutralizing reactive oxygen species and contributing to cellular well-being[69]. In some pathogenic bacteria such as *Streptococcus pneumoniae*, *Helicobacter pylori,* and *Mycobacterium tuberculosis*, vitamin B6 biosynthesis is essential for their survival, motility, and virulence[70].

Uracil, a product of bacterial metabolism, can act as a mucosal signal that stimulates host dual oxidase (DUOX)-dependent antimicrobial reactive oxygen species (ROS) production, thereby contributing to host-cross-talk[59,71–73]. Uracil is generated during uridine catabolism, together with ribose, and nucleotide-metabolism genes are more frequently associated with pathogenic than commensal bacteria. Ribose, another product of nucleoside catabolism, can also promote bacterial quorum sensing (QS) and virulence-gene expression, potentially contributing to a commensal-to-pathogen transition[74]. In our study, uracil abundance was highest in ES isolates compared with PS and lung isolates, suggesting site-associated differences in uracil production or utilization. However, the biological significance of this difference for host-microbe interactions remains to be determined.

Biosynthesis of branched-chain amino acids, such as valine, leucine, and isoleucine, is crucial for bacterial survival and growth[77]. The study of Begona et al. found that NTHi up-regulates non-aromatic amino acid (leucine, isoleucine, and valine) biosynthesis during lung infection[62]. Furthermore, research suggests that a decrease in valine can increase competence in NTHi, promoting extracellular DNA uptake for nutritional purposes[78,79]. Our study demonstrated that Lung strains utilized isoleucine to the greatest extent compared to ES and PS strains. In contrast, valine was produced only by ES strains.

Our findings demonstrate that there are site-associated differences in the extracellular metabolic profiles of NTHi isolates, suggesting that strains from different anatomical niches may exhibit distinct metabolic phenotypes. Nevertheless, the present study does not resolve the underlying mechanisms for these differences. Additionally, the small and unequal sample size (2 ear smear strains, 2 pharyngeal smear strains, and 5 lower respiratory tract strains) limits the generalizability of our findings to site-specific NTHi strains. Further *in vitro* and *in vivo* studies utilising larger, phylogenetically diverse isolate collections are required to ascertain the physiological significance and reproducibility of these observations. Nevertheless, the findings of this study provide a foundation for future research on the development of targeted, site-specific therapeutic approaches for NTHi-associated CLD.

Some significant compounds remained unannotated against the available mass spectral libraries under the given annotation parameters[65,80]. Application of further methods, such as gas chromatography coupled with mass spectrometry, should be used for further identification of the significant signals. Some compounds, such as inosine, hypoxanthine, niacinamide, and uracil appeared as multiple features in our analysis. Retention time shifts over a long sequence of

acquisition samples, broad peak elution from the LC column, accurate mass deviation beyond the threshold limit and isomerism can trigger such identical identification of multiple features [81].

The observed divergence in metabolite abundance despite broadly conserved metabolic gene content suggests that genomic potential alone does not fully determine the metabolic state of NTHi isolates. These differences may reflect differential gene regulation, pathway activity, substrate availability, or metabolic flux within a conserved genomic framework. However, the present study does not resolve the regulatory mechanisms underlying these metabolic states, as gene expression and regulatory activity were not directly assessed. One possible explanation for this phenotypic divergence could be the presence of phase-variable regulons (phasevarions) in human-associated bacterial pathogens, including NTHi. Phasevarions can alter the expression of multiple genes through phase-variable DNA methylation, resulting in ON-OFF switching of regulons involved in bacterial adaptation and pathobiology[16]. Whether these metabolic states influence bacterial fitness or reflect selection by host-derived factors also remains unresolved and will require targeted functional studies.

Given that core-genome SNP analysis revealed limited genomic divergence (8-116 SNPs) among the strains, the observed metabolic differences are more likely attributable to regulatory or epigenetic mechanisms rather than major gene-content variation. To validate this assumption, further analyses such as transcriptomics and epigenetics profiling would be necessary. Furthermore, as a limitation of this study, post-translational biosynthetic regulation (such as enzyme activity modulation through phosphorylation and acetylation), and differences in growth phases (in our study, the growth was adjusted to the incubation period) and the resulting metabolic states could have influenced the genomic-metabolic disconnection in this study [82,83].

In summary, this study underscores the significance of metabolomic investigations from a bacterial perspective. A deeper understanding of pathogen-specific metabolic profiles provides insight into survival strategies within the host and informs about site/disease specific microbial communities. Our findings support the hypothesis that NTHi strains associate to distinct infection sites through metabolic reprogramming, despite possessing a largely homogeneous genomic capacity for metabolite synthesis (Fig. 5). The observed metabolic variations suggest that NTHi may modulate nutrient uptake and physiological processes in a site-specific manner. Future research should focus on validating these pathways functionally and assessing their relevance *in vivo*.

4 Materials and methods

4.1 Ethics approval and consent to participate

The retrospective collection of defined clinical and patient-related parameters from patients with detection of *Haemophilus* sp. in routine clinical samples did not require direct informed consent. The full name of ethical committee which waived the need for consent is as follows: Ethics Committee at the University of Lübeck, Ratzeburger Allee 160, Building 70, 23562 Lübeck. All data were obtained in a blind manner. The study protocol was approved by the local ethics committee of the University of Lübeck (Az 13-072). The waiver of consent by patients in the retrospective analysis of the study was specified in the study protocol and approved by Ethics Committee Resolution No. 13-072. It should be noted that, in today's ethics applications, a separate application is required for the waiver of consent, whereas at that time it was still addressed within the study protocols. Research and data collection was conducted in compliance with the principles of the Helsinki Declaration as described by Diricks et al.[17]. The collected data was accessed for research purposes since 01/06/2013.

2.2. Genome assembly, quality assessment, and comparative genomic analysis

Previously generated Illumina MiSeq 251-bp paired-end reads were quality-checked, trimmed, assembled, and polished using Shovill with SPAdes v3.14.0 [17,84]. Assembly completeness and quality were assessed using BUSCO v5.8.2 with the bacteria_odb10 dataset and QUAST v5.3.0, respectively[85,86]. Antimicrobial resistance determinants were screened using ABRicate v1.0.1 against the NCBI AMR database, and species identity was confirmed by FastANI using the *H. influenzae* CP000057.2 reference genome[21,22]. The complete, high-quality genome of the well-characterized NTHi strain 86-028NP (CP000057.2) was selected as a comparative genomic framework for annotation, gene-content assessment, and functional characterization. As an otitis media isolate, 86-028NP was not considered representative or ancestral to the study strains; therefore, potential strain- or lineage-specific reference bias was acknowledged, and comparative findings were interpreted within the genomic diversity of the study isolates. The complete genomes of all isolated strains are deposited in the European Nucleotide Archive database (Bioproject; PRJEB43356). High-quality assemblies (BUSCO completeness 98.2-99.4%) were retained for comparative analyses.

Genomes were annotated using Prokka v1.14.6, with RAST used as an independent resource for functional and pathway annotation[87]. Prokka annotations provided the standardized gene set for comparative and pangenome analyses, whereas RAST was used to cross-validate functional

assignments. The complete NTHi strain 86-028NP genome (CP000057.2) was used as a comparative framework for gene-content and functional analyses, but not as an ancestral or representative reference for the clinical isolates. Accordingly, strain-specific and lineage-specific differences were interpreted within the genomic diversity of the isolates. Core- and accessory-genome variation was assessed using Roary based on Prokka-generated GFF3 files. RAST and Prokka were used as complementary annotation approaches, with Prokka providing the standardized genome annotation and protein-coding sequences for downstream comparative and pangenome analyses, while RAST served as an independent resource for cross-validation of gene functions and pathway assignments; the two annotation outputs were not combined as independent gene sets.Core-genome SNPs were extracted from the core-gene alignment using snp-sites, and pairwise SNP distances were calculated using snp-dists. Orthologous genes were functionally categorized using COG assignments to identify conserved and variable functional capacities among isolates. Together, these analyses defined the extent of genomic conservation and variation across the strains and provided the genomic framework for subsequent analysis of metabolic potential.

2.3. Biosynthetic and metabolic potential

To determine whether genomic conservation was associated with conserved metabolic capacity, biosynthetic gene clusters (BGCs) were predicted using antiSMASH v8.0.1 from Prokka-annotated GenBank files[88,89]. Comparative BGC profiling was performed across the nine pathogenic NTHi, with conserved and strain-enriched clusters identified for subsequent consideration in relation to metabolite production. Metabolic functions were further characterized using KEGG and RAST pathway annotations[90–93]. Particular emphasis was placed on carbohydrate metabolism because of the observed inter-strain differences in carbohydrate utilization and its potential relevance to host adaptation and pathogenicity. Genes involved in central carbohydrate metabolism, polysaccharide and monosaccharide utilization, amino-sugar and organic-acid metabolism, fermentation, glycoside hydrolysis, and related pathways were identified and compared across strains. KEGG orthologies were assigned using KAAS with the bi-directional best-hit method, and pathway completeness and gene presence/absence were evaluated using KEGG Mapper[94]. Conserved genes and pathways across the isolates were considered evidence of a shared metabolic framework, whereas differences in accessory genes, pathway completeness, or biosynthetic gene clusters highlighted potential sources of strain-specific metabolic phenotypes. This genomic conservation provided the rationale for

investigating whether differences in metabolite profiles arise primarily from differences in metabolic gene content or instead from differential regulation of otherwise conserved metabolic pathways.

4.4 Cultivation and extraction

Each strain of NTHi strains was streaked on chocolate agar (BD) and incubated at 37°C for 24 h. A pure colony grown after 24 h was inoculated into 6 mL chemically defined medium as mentioned by Coleman et al. (2003) and incubated for 24 h at 37°C at an agitation of 160 rpm[95]. The cultivation medium consisted of 194 mL of RPMI 1640 with L-glutamine and phenol red but without 25 mM HEPES, 2 mL of a 100 mM MEM sodium pyruvate solution (Sigma-Aldrich), 2 mL of β-NAD^+ stock, 4 mL of heme-L-histidine stock, 10 mL of a 2 mg/mL uracil solution dissolved in 0.1 M NaOH, and 20 mL of a 20 mg/mL inosine solution prepared by dissolving inosine (Sigma-Aldrich) in deionized water and filter sterilized (0.2 µm pore size). The β-NAD^+ stock was prepared by dissolving 100 mg of β-NAD^+ (Sigma-Aldrich) in 100 mL of deionized water and filter sterilized (0.2 µm pore size). The heme-L-histidine stock was prepared by dissolving 0.2 g of L-histidine (Sigma-Aldrich) in 200 mL of deionized water and adding 0.2 g of hemin HCl (Sigma-Aldrich) and 4 mL of 1 M NaOH and steaming over a water bath for 10 min to solubilize the mixture, and filter sterilized (0.2 µm pore size). After measuring optical density at 600nm (OD_{600}), 1 mL of each seed culture was further inoculated into 18 mL of fresh medium (main culture) and incubated for 24 h at 37°C and 160 rpm. Depending on the growth (OD_{600}) of different strains, to obtain the $OD_{600} > 0.50$, the incubation period varied up to 28 h. The main culture of each strain was centrifuged (Avanti J-25 XP) for 30 min at 4°C and 30074 × *g*. The supernatant was filter sterilized (0.2 µm pore size) and stored at -20°C until n-butanol extraction, while the cells were washed with sterile water, liquid N_2-killed, and stored at -80°C.

For solvent extraction, the supernatant of each culture was mixed with 20 mL n-butanol and shaken vigorously for 20 min. Afterwards, the mixture was centrifuged (Thermo Scientific Sorvall RC 6+ Centrifuge) for 10 min at 10°C and 2290 × *g*. The upper solvent phase was collected, evaporated under a rotary evaporator, redissolved in 1 mL methanol (Millipore), and N_2-dried to yield the final dry extract ready for LC/MS measurements.

4.5 LC-MS measurement and metabolomics

Samples for LC-MS measurement were prepared at a concentration of 0.1 mg/mL in methanol and analysed by ultrahigh-performance liquid chromatography-electrospray ionization-high-resolution

tandem mass spectrometry (UHPLC-ESI-HRMS/MS) using an Elute UHPLC system coupled to a Compact QTOF mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with a reversed-phase Intensity Solo 2 C18 column (100 mm × 2.1 mm, 2.2 µm; Bruker Daltonics, Bremen, Germany). The UHPLC acquisition method was set with column temperature: 35°C, autosampler temperature: 4°C, sample injection volume: 10 µL with flush volume of 45 µL via partial loop injection. Water with 0.1% formic acid (buffer A) and acetonitrile with 0.1% formic acid (buffer B) were used for mobile phases with the elution gradient of 10% B isocratically for 2 min, then to 100% B in 17 min, washing the column with 100% B for 5 min, and finally recalibration of the column with 10% B for 5 min. The MS acquisition was set up with following parameters; mass spectra peak detection: use maximum intensity, absolute threshold (per 1000 sum.): 25 cts., absolute threshold: 20 cts., peak summation width: 5 pts., source: ESI, end plate offset: 500 V, capillary: 4500 V, nebulizer: 2.2 Bar, dry gas: 10L/min, dry temperature: 220°C, Funnel 1RF and Funnel 2RF: 200 Vpp, isCID energy: 0 eV, hexapole: 50 Vpp, quadrupole Ion energy: 5 eV, collision cell energy: 5 eV, pre pulse storage 5 µs, stepping mode: basic, collision RF: 200-700 Vpp, transfer time: 20-80µs, MS/MS only collision energy: 100-250%, precursor ion cycle time: 0.3 sec, MS/MS absolute threshold (per 1000 sum.): 400 cts., MS/MS absolute threshold: 313, active exclusion: after 3 spectra and 0.3 min, calibration mode: HPC, calibration reference compound: sodium formate, ion polarity: positive, scan mode: MS, scan range: 20-1300 *m/z*. Each sample was measured twice accompanied by external quality control (QC, 0.0125 mg/mL biotin, injection volume of 2 µL) and methanol blank. The raw LC-MS data of 48 samples, including 4 external QC samples, 2 n-butanol blank samples, 4 methanol blank samples, 2 media control samples, and 36 bacterial-culture extracts (encompassing 2 biological- and 2 technical- replicates) were processed through MetaboScape 2022b Version 9.0.1. In MetaboScape, following parameters were set: minimum number of features for extraction: 4, presence of features in minimum number of analyses: 4, processing method: T-ReX 3D Processing, intensity threshold: 8000 counts, feature signal: area, minimum peak length: 10 spectra, retention time range: 0-22 min and mass range (*m/z*): 50-1300. Furthermore, ion deconvolution was set with EIC correlation: 0.8, primary ion: $[M+H]^+$, seed ions: $[M+Na]^+$, $[M+K]^+$, $[M+2H]^{2+}$ and $[M+2Na]^{2+}$, and common ions: $[M+H-H_2O]^+$ and $[M-H_2O+2H]^{2+}$.

In MetaboScape, the feature table of metabolites was annotated against different mass spectral libraries including Bruker MetaboBASE Personal Library 3.0 (100679 compounds), MassBank_NIST (21804 compounds), MSMS_Public_Pos_VS17 (21127 compounds) and MoNA (17748 compounds). During annotation against libraries, the parameters were set as *m/z* tolerance of

5 ppm, mSigma value up to 60, and a minimum MS/MS score of 800. The feature table of metabolites was exported as an excel file (.csv) from MetaboScape after filtering out the features eluted before 0.5 min, features lacking MS/MS spectra, features present extensively in n-butanol.

4.6 Statistical analysis

The feature table generated from nine NTHi strains was considered for statistical analysis. It was further cleaned by filtering out all the features present in 4 methanol blank samples. Additionally, features present in only media-control samples were also removed. At the final cleaning step, all those features were kept in the table that were explicitly and comparatively present in at least one set of 4 values representing 2 biological- and 2 technical- replicates while the rest were completely filtered out from the feature table. Additionally, QC, n-butanol, and methanol blank samples were also removed from the table. In the table, the samples were attributed to four groups: ES, PS, Lung and not applicable ("na"). Lung group consisted of samples from the bacteria isolated from tracheal secretion, bronchial secretion and bronchoalveolar lavage. Media control sample was assigned to "na" group.

The final feature table was split into data table consisting of sample names versus feature names with their peak area and a metadata table consisting of two columns: sample names and sampling sites (ES, PS, Lung, na). In the data table, samples were kept in rows and features were kept in columns.

In MetaboAnalyst 6.0, statistical analysis module is used for the comparative analysis of metabolites produced by NTHi strains isolated from different body sites[96–98]. The data and metadata files were uploaded by considering peak intensities as data type. Data filtering step was skipped while manual data cleaning was done before processing data in MetaboAnalyst. Afterwards, the data were normalized by log transformation (base 10). Finally, principal component analysis (PCA) and linear model were performed by considering each group as reference and the other group as contrast with FDR-adjusted *p*-value cutoff of 0.05 without covariate adjustment. The significant metabolites were tabulated with their logFC, *p*-value, and adj *p*-value. Box plots of each significant metabolite in different comparison groups were generated using the *tidyverse* package in R (version 4.5.1).

5 References

1. Fleischmann, R. D. *et al.* Whole-Genome Random Sequencing and Assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512 (1995).

2. Van Eldere, J., Slack, M. P. E., Ladhani, S. & Cripps, A. W. Non-typeable *Haemophilus influenzae*, an under-recognised pathogen. *Lancet Infect. Dis.* **14**, 1281–1292 (2014).

3. Davis Gregg S., Sandstedt Sara A., Patel Mayuri, Marrs Carl F., & Gilsdorf Janet R. Use of bexB To Detect the Capsule Locus in *Haemophilus influenzae*. *J. Clin. Microbiol.* **49**, 2594–2601 (2011).

4. Seppanen, E. J. *et al. Haemophilus influenzae* remains the predominant otitis media pathogen in Australian children undergoing ventilation tube insertion in the PCV13 era. *J. Infect.* **90**, 106478 (2025).

5. Yamada, S. *et al.* β-Lactamase-non-producing ampicillin-resistant *Haemophilus influenzae* is acquiring multidrug resistance. *J. Infect. Public Health* **13**, 497–501 (2020).

6. Mell, J. C., Hall, I. M. & Redfield, R. J. Defining the DNA uptake specificity of naturally competent *Haemophilus influenzae* cells. *Nucleic Acids Res.* **40**, 8536–8549 (2012).

7. Bakkali, M., Chen, T.-Y., Lee, H. C. & Redfield, R. J. Evolutionary stability of DNA uptake signal sequences in the Pasteurellaceae. *Proc. Natl. Acad. Sci.* **101**, 4513–4518 (2004).

8. KC, R. *et al.* Whole-genome analyses reveal gene content differences between nontypeable *Haemophilus influenzae* isolates from chronic obstructive pulmonary disease compared to other clinical phenotypes. *Microbial Genomics* vol. 6 (2020).

9. Rempe Katherine A., Porsch Eric A., Wilson Jolaine M., & St. Geme Joseph W. The HMW1 and HMW2 Adhesins Enhance the Ability of Nontypeable *Haemophilus influenzae* To Colonize the Upper Respiratory Tract of Rhesus Macaques. *Infect. Immun.* **84**, 2771–2778 (2016).

10. Meng, G., St. Geme, J. W. & Waksman, G. Repetitive Architecture of the *Haemophilus influenzae* Hia Trimeric Autotransporter. *J. Mol. Biol.* **384**, 824–836 (2008).

11. Seale Thomas W. *et al.* Complex Role of Hemoglobin and Hemoglobin-Haptoglobin Binding Proteins in *Haemophilus influenzae* Virulence in the Infant Rat Model of Invasive Infection. *Infect. Immun.* **74**, 6213–6225 (2006).

12. Othman, D. S. M. P., Schirra, H. J., McEwan, A. G. & Kappler, U. Metabolic versatility in *Haemophilus influenzae*: a metabolomic and genomic analysis. *Front. Microbiol.* **Volume 5-2014**, (2014).

13. Harrison, A. *et al.* Reprioritization of biofilm metabolism is associated with nutrient adaptation and long-term survival of *Haemophilus influenzae*. *Npj Biofilms Microbiomes* **5**, 33 (2019).

14. Fernández-García, M. *et al.* Multiplatform Metabolomics Characterization Reveals Novel Metabolites and Phospholipid Compositional Rules of *Haemophilus influenzae* Rd KW20. *Int. J. Mol. Sci.* **24**, (2023).

15. Muda, N. M. *et al.* Metabolic analyses reveal common adaptations in two invasive *Haemophilus influenzae* strains. *Pathog. Dis.* **77**, ftz015 (2019).

16. Atack, J. M., Murphy, T. F., Pettigrew, M. M., Seib, K. L. & Jennings, M. P. Non-typeable *Haemophilus influenzae* isolates from patients with chronic obstructive pulmonary disease contain new phase-variable modA methyltransferase alleles controlling phasevarions. *Sci. Rep.* **9**, 15963 (2019).

17. Diricks, M. *et al.* Whole genome sequencing-based classification of human-related *Haemophilus* species and detection of antimicrobial resistance genes. *Genome Med.* **14**, 13 (2022).

18. Natalini, J. G., Singh, S. & Segal, L. N. The dynamic lung microbiome in health and disease. *Nat. Rev. Microbiol.* **21**, 222–235 (2023).

19. McCubbin, T. *et al.* A Pan-Genome Guided Metabolic Network Reconstruction of Five Propionibacterium Species Reveals Extensive Metabolic Diversity. *Genes* **11**, (2020).

20. Norsigian, C. J. *et al.* Systems biology approach to functionally assess the Clostridioides difficile pangenome reveals genetic diversity with discriminatory power. *Proc. Natl. Acad. Sci.* **119**, e2119396119 (2022).

21. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).

22. Hou, T. *et al.* Whole-genome and pan-genome analyses reveal genomic differences among nontypeable *Haemophilus influenzae* isolates from bronchiectasis, community-acquired pneumonia, and chronic obstructive pulmonary disease. *Front. Cell. Infect. Microbiol.* **Volume 16-2026**, (2026).

23. Zhu, B. *et al.* Multi-omics analysis of niche specificity provides new insights into ecological adaptation in bacteria. *ISME J.* **10**, 2072–2075 (2016).

24. Yamaguchi, M. & Yamamoto, K. Mucin glycans and their degradation by gut microbiota. *Glycoconj. J.* **40**, 493–512 (2023).

25. Bell, A. & Juge, N. Mucosal glycan degradation of the host by the gut microbiota. *Glycobiology* **31**, 691–696 (2021).

26. Luis, A. S. & Hansson, G. C. Intestinal mucus and their glycans: A habitat for thriving microbiota. *Cell Host Microbe* **31**, 1087–1100 (2023).

27. Belzer, C. Nutritional strategies for mucosal health: the interplay between microbes and mucin glycans. *Trends Microbiol.* **30**, 13–21 (2022).

28. Consentino, L., Rejasse, A., Crapart, N., Bevilacqua, C. & Nielsen-LeRoux, C. Laser capture microdissection to study Bacillus cereus iron homeostasis gene expression during *Galleria mellonella* in vivo gut colonization. *Virulence* **12**, 2104–2121 (2021).

29. Weisberg, A. J. & Chang, J. H. Mobile Genetic Element Flexibility as an Underlying Principle to Bacterial Evolution. *Annual Review of Microbiology* vol. 77 603–624 (2023).

30. Reams, A. B. & Neidle, E. L. Selection for Gene Clustering by Tandem Duplication. *Annual Review of Microbiology* vol. 58 119–142 (2004).

31. Mbelle, N. M. *et al.* The Resistome, Mobilome, Virulome and Phylogenomics of Multidrug-Resistant *Escherichia coli* Clinical Isolates from Pretoria, South Africa. *Sci. Rep.* **9**, 16457 (2019).

32. Karczmarczyk Maria, Abbott Yvonne, Walsh Ciara, Leonard Nola, & Fanning Séamus. Characterization of Multidrug-Resistant *Escherichia coli* Isolates from Animals Presenting at a University Veterinary Hospital. *Appl. Environ. Microbiol.* **77**, 7104–7112 (2011).

33. Mbelle, N. M. *et al.* Pathogenomics and Evolutionary Epidemiology of Multi-Drug Resistant Clinical *Klebsiella pneumoniae* Isolated from Pretoria, South Africa. *Sci. Rep.* **10**, 1232 (2020).

34. Adolf, L. A. & Heilbronner, S. Nutritional Interactions between Bacterial Species Colonising the Human Nasal Cavity: Current Knowledge and Future Prospects. *Metabolites* **12**, (2022).

35. Seo, H., Kim, S., Beck, S. & Song, H.-Y. Perspectives on Microbiome Therapeutics in Infectious Diseases: A Comprehensive Approach Beyond Immunology and Microbiology. *Cells* **13**, (2024).

36. Müller Miklós *et al.* Biochemistry and Evolution of Anaerobic Energy Metabolism in Eukaryotes. *Microbiol. Mol. Biol. Rev.* **76**, 444–495 (2012).

37. Cavalcanti, J. H. F. *et al.* Evolution and Functional Implications of the Tricarboxylic Acid Cycle as Revealed by Phylogenetic Analysis. *Genome Biol. Evol.* **6**, 2830–2848 (2014).

38. Nogales, J., Gudmundsson, S., Duque, E., Ramos, J. L. & Palsson, B. O. Expanding the computable reactome in *Pseudomonas putida* reveals metabolic cycles providing robustness. *bioRxiv* 139121 (2017) doi:10.1101/139121.

596 39. Jardón Raquel, Gancedo Carlos, & Flores Carmen-Lisset. The Gluconeogenic Enzyme Fructose-
597 1,6-Bisphosphatase Is Dispensable for Growth of the Yeast Yarrowia lipolytica in Gluconeogenic
598 Substrates. *Eukaryot. Cell* **7**, 1742–1749 (2008).

599 40. Paixão, L. *et al.* Transcriptional and metabolic effects of glucose on Streptococcus pneumoniae
600 sugar metabolism. *Front. Microbiol.* **Volume 6-2015**, (2015).

601 41. Awad, K., Maghraby, A. S., Abd-Elshafy, D. N. & Bahgat, M. M. Carbohydrates Metabolic
602 Signatures in Immune Cells: Response to Infection. *Front. Immunol.* **Volume 13-2022**, (2022).

603 42. Paixão, A. L. The role of pneumococcal carbon metabolism in colonisation and invasive disease.
604 in (2015).

605 43. Ravcheev, D. A. & Thiele, I. Comparative Genomic Analysis of the Human Gut Microbiome
606 Reveals a Broad Distribution of Metabolic Pathways for the Degradation of Host-Synthetized
607 Mucin Glycans and Utilization of Mucin-Derived Monosaccharides. *Front. Genet.* **Volume 8-**
608 **2017**, (2017).

609 44. Lal, M. A. & Bhatla, S. C. Metabolism of Storage Carbohydrates. in *Plant Physiology,*
610 *Development and Metabolism* (eds Bhatla, S. C. & Lal, M. A.) 231–259 (Springer Nature
611 Singapore, Singapore, 2023). doi:10.1007/978-981-99-5736-1_9.

612 45. Jo, M.-S., Yang, H.-W., Park, J.-H., Shin, J.-M. & Park, I.-H. Glycolytic reprogramming is
613 involved in tissue remodeling on chronic rhinosinusitis. *PLOS ONE* **18**, e0281640 (2023).

614 46. Tuchscherr, L. *et al.* Reduced Glycolysis and Cytotoxicity in *Staphylococcus aureus* Isolates
615 from Chronic Rhinosinusitis as Strategies for Host Adaptation. *Int. J. Mol. Sci.* **25**, (2024).

616 47. Ruan, Y. *et al.* Comparative genome analysis of Prevotella intermedia strain isolated from
617 infected root canal reveals features related to pathogenicity and adaptation. *BMC Genomics* **16**,
618 122 (2015).

48. Chen, F. *et al.* Age-Associated Changes of Nasal Bacterial Microbiome in Patients With Chronic Rhinosinusitis. *Front. Cell. Infect. Microbiol.* **Volume 12-2022**, (2022).

49. Carvalho, S. M., de Jong, A., Kloosterman, T. G., Kuipers, O. P. & Saraiva, L. M. The *Staphylococcus aureus* α-Acetolactate Synthase ALS Confers Resistance to Nitrosative Stress. *Front. Microbiol.* **Volume 8-2017**, (2017).

50. Weber Michael & Fuchs Thilo M. Metabolism in the Niche: a Large-Scale Genome-Based Survey Reveals Inositol Utilization To Be Widespread among Soil, Commensal, and Pathogenic Bacteria. *Microbiol. Spectr.* **10**, e02013-22 (2022).

51. Michell, R. H. Inositol and its derivatives: Their evolution and functions. *Adv. Enzyme Regul.* **51**, 84–90 (2011).

52. Michell, R. H. Inositol derivatives: evolution and functions. *Nat. Rev. Mol. Cell Biol.* **9**, 151–161 (2008).

53. An, D., Na, C., Bielawski, J., Hannun, Y. A. & Kasper, D. L. Membrane sphingolipids as essential molecular signals for Bacteroides survival in the intestine. *Proc. Natl. Acad. Sci.* **108**, 4666–4671 (2011).

54. Dinglasan, J. L. N., Otani, H., Doering, D. T., Udwary, D. & Mouncey, N. J. Microbial secondary metabolites: advancements to accelerate discovery towards application. *Nat. Rev. Microbiol.* **23**, 338–354 (2025).

55. Zhang Zhiguo *et al.* Diversity and distribution of biosynthetic gene clusters in agricultural soil microbiomes. *mSystems* **9**, e01263-23 (2024).

56. da Rocha Lapa, F. *et al.* Anti-inflammatory effects of inosine in allergic lung inflammation in mice: evidence for the participation of adenosine A2A and A3 receptors. *Purinergic Signal.* **9**, 325–336 (2013).

57. Yu, M. *et al.* Aberrant purine metabolism in allergic asthma revealed by plasma metabolomics. *J. Pharm. Biomed. Anal.* **120**, 181–189 (2016).

58. Ma, C. *et al.* Hypoxanthine as an Important Biomarker Participate in Intestinal Tissue Injury in COPD Patient Via 'Lung-Gut Axis'. *SSRN Electron. J.* https://doi.org/10.2139/ssrn.4160039 (2022) doi:10.2139/ssrn.4160039.

59. Cui, L. *et al.* Metabolomics Investigation Reveals Metabolite Mediators Associated with Acute Lung Injury and Repair in a Murine Model of Influenza Pneumonia. *Sci. Rep.* **6**, 26076 (2016).

60. Esther, C. R. *et al.* Metabolomic biomarkers predictive of early structural lung disease in cystic fibrosis. *Eur. Respir. J.* **48**, 1612–1621 (2016).

61. Evans, C. R. *et al.* Untargeted LC–MS Metabolomics of Bronchoalveolar Lavage Fluid Differentiates Acute Respiratory Distress Syndrome from Health. *J. Proteome Res.* **13**, 640–649 (2014).

62. Euba Begoña *et al.* In Vivo Genome-Wide Gene Expression Profiling Reveals That *Haemophilus influenzae* Purine Synthesis Pathway Benefits Its Infectivity within the Airways. *Microbiol. Spectr.* **11**, e00823-23 (2023).

63. Ziklo, N., Bibi, M., Sinai, L. & Salama, P. Niacinamide Antimicrobial Efficacy and Its Mode of Action via Microbial Cell Cycle Arrest. *Microorganisms* **12**, (2024).

64. Zhang, Q., Li, J., Zhong, H. & Xu, Y. The mechanism of nicotinamide on reducing acute lung injury by inhibiting MAPK and NF-κB signal pathway. *Mol. Med.* **27**, 115 (2021).

65. Reidl, J. *et al.* NADP and NAD utilization in *Haemophilus influenzae*. *Mol. Microbiol.* **35**, 1573–1581 (2000).

66. Kemmer Gabriele *et al.* NadN and e (P4) Are Essential for Utilization of NAD and Nicotinamide Mononucleotide but Not Nicotinamide Riboside in *Haemophilus influenzae*. *J. Bacteriol.* **183**, 3974–3981 (2001).

67. Herbert Mark *et al.* Nicotinamide Ribosyl Uptake Mutants in *Haemophilus influenzae*. *Infect. Immun.* **71**, 5398–5401 (2003).

68. Sauer Elizabeta, Merdanovic Melisa, Price Mortimer Anne, Bringmann Gerhard, & Reidl Joachim. PnuC and the Utilization of the Nicotinamide Riboside Analog 3-Aminopyridine in *Haemophilus influenzae*. *Antimicrob. Agents Chemother.* **48**, 4532–4541 (2004).

69. Tramonti Angela *et al.* Knowns and Unknowns of Vitamin B6 Metabolism in *Escherichia coli*. *EcoSal Plus* **9**, eESP-0004-2021 (2021).

70. Grubman Alexandra *et al.* Vitamin B6 Is Required for Full Motility and Virulence in Helicobacter pylori. *mBio* **1**, 10.1128/mbio.00112-10 (2010).

71. Lee, K.-A. *et al.* Bacterial-Derived Uracil as a Modulator of Mucosal Immunity and Gut-Microbe Homeostasis in Drosophila. *Cell* **153**, 797–811 (2013).

72. Lee, W.-J. & Kim, S.-H. Role of DUOX in gut inflammation: lessons from Drosophila model of gut-microbiota interactions. *Front. Cell. Infect. Microbiol.* **Volume 3-2013**, (2014).

73. Sarr, D., Tóth, E., Gingerich, A. & Rada, B. Antimicrobial actions of dual oxidases and lactoperoxidase. *J. Microbiol.* **56**, 373–386 (2018).

74. Kim, E.-K. *et al.* Bacterial Nucleoside Catabolism Controls Quorum Sensing and Commensal-to-Pathogen Transition in the Drosophila Gut. *Cell Host Microbe* **27**, 345-357.e6 (2020).

75. Sun, W. *et al.* Transcriptomics and metabolomics analysis of L-phenylalanine overproduction in *Escherichia coli*. *Microb. Cell Factories* **22**, 65 (2023).

685 76. Tang, Y. *et al.* Phenylalanine promotes alveolar macrophage pyroptosis via the activation of
686 CaSR in ARDS. *Front. Immunol.* **Volume 14-2023**, (2023).

687 77. Amorim Franco, T. M. & Blanchard, J. S. Bacterial Branched-Chain Amino Acid Biosynthesis:
688 Structures, Mechanisms, and Drugability. *Biochemistry* **56**, 5849–5865 (2017).

689 78. Harrison, A. *et al.* Comprehensive Proteomic and Metabolomic Signatures of Nontypeable
690 *Haemophilus influenzae*-Induced Acute Otitis Media Reveal Bacterial Aerobic Respiration in an
691 Immunosuppressed Environment*. *Mol. Cell. Proteomics* **15**, 1117–1138 (2016).

692 79. Marsh, R. L. *et al.* Panel 4: Recent advances in understanding the natural history of the otitis
693 media microbiome and its response to environmental pressures. *Res. Panel Publ. Int. Soc. Otitis*
694 *Media Hollywood CA June 2019* **130**, 109836 (2020).

695 80. Sinha, S., Mell, J. & Redfield, R. The availability of purine nucleotides regulates natural
696 competence by controlling translation of the competence activator Sxy. *Mol. Microbiol.* **88**,
697 1106–1119 (2013).

698 81. Habra, H. *et al.* metabCombiner 2.0: Disparate Multi-Dataset Feature Alignment for LC-MS
699 Metabolomics. *Metabolites* **14**, (2024).

700 82. Albertin, W. *et al.* Linking Post-Translational Modifications and Variation of Phenotypic Traits*.
701 *Mol. Cell. Proteomics* **12**, 720–735 (2013).

702 83. Hosmer, J. *et al.* Access to highly specialized growth substrates and production of epithelial
703 immunomodulatory metabolites determine survival of *Haemophilus influenzae* in human airway
704 epithelial cells. *PLOS Pathog.* **18**, e1010209 (2022).

705 84. Bankevich, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications to
706 Single-Cell Sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).

85. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).

86. Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A. & Zdobnov, E. M. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).

87. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).

88. Medema, M. H. *et al.* antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **39**, W339–W346 (2011).

89. Blin, K. *et al.* antiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. *Nucleic Acids Res.* **53**, W32–W38 (2025).

90. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

91. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951 (2019).

92. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res.* **53**, D672–D677 (2025).

93. Aziz, R. K. *et al.* The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9**, 75 (2008).

94. Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C. & Kanehisa, M. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* **35**, W182–W185 (2007).

95. Coleman Hannah N., Daines Dayle A., Jarisch Justin, & Smith Arnold L. Chemically Defined Media for Growth of *Haemophilus influenzae* Strains. *J. Clin. Microbiol.* **41**, 4408–4410 (2003).

96. Pang, Z. *et al.* MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. *Nucleic Acids Res.* **52**, W398–W406 (2024).

97. Xia, J., Psychogios, N., Young, N. & Wishart, D. S. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **37**, W652–W660 (2009).

98. Ewald, J. D. *et al.* Web-based multi-omics integration using the Analyst software suite. *Nat. Protoc.* **19**, 1467–1497 (2024).

Funding

This work was supported by German Center for Lung Research (DZL), project: DZL4AA22D (to KB and KAD).

Author Contributions

KB conceived and initiated the project, designed and performed metabolomics experiments, and prepared the manuscript.

BB designed and performed genomic experiments, and prepared the manuscript.

AS prepared the manuscript.

MD prepared the manuscript.

SN prepared the manuscript.

JR contributed with the bacterial isolates and prepared the manuscript.

KAD initiated the project, got funding, designed and supervised experiments, and prepared the final version of the manuscript.

Competing interests

The authors declare no competing interests.

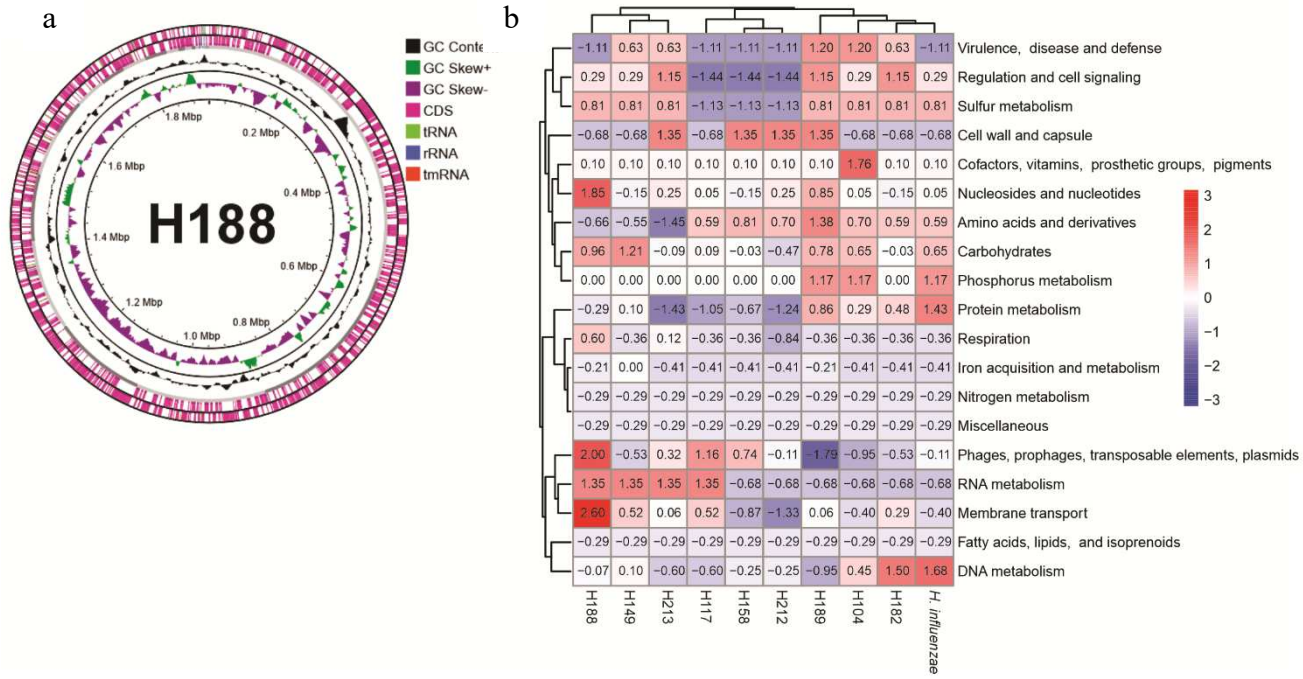
Data Availability

The datasets (raw spectral data) for this study can be found in the MassIVE at doi:10.25345/C58S4K27R, reference number MassIVE MSV000099632.

Supplementary Material

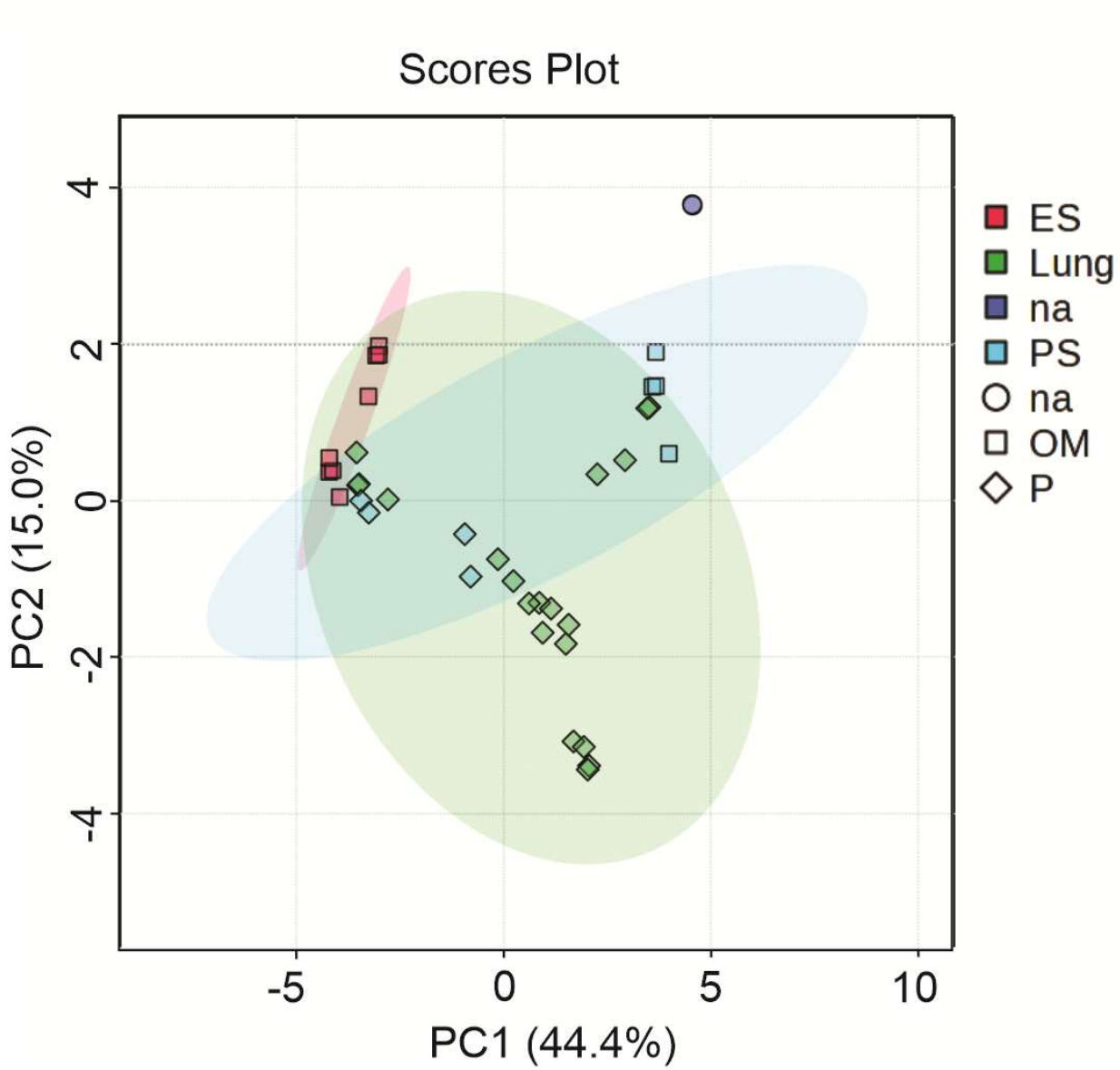
Please see separate supplementary files.

Figures & figures legends:



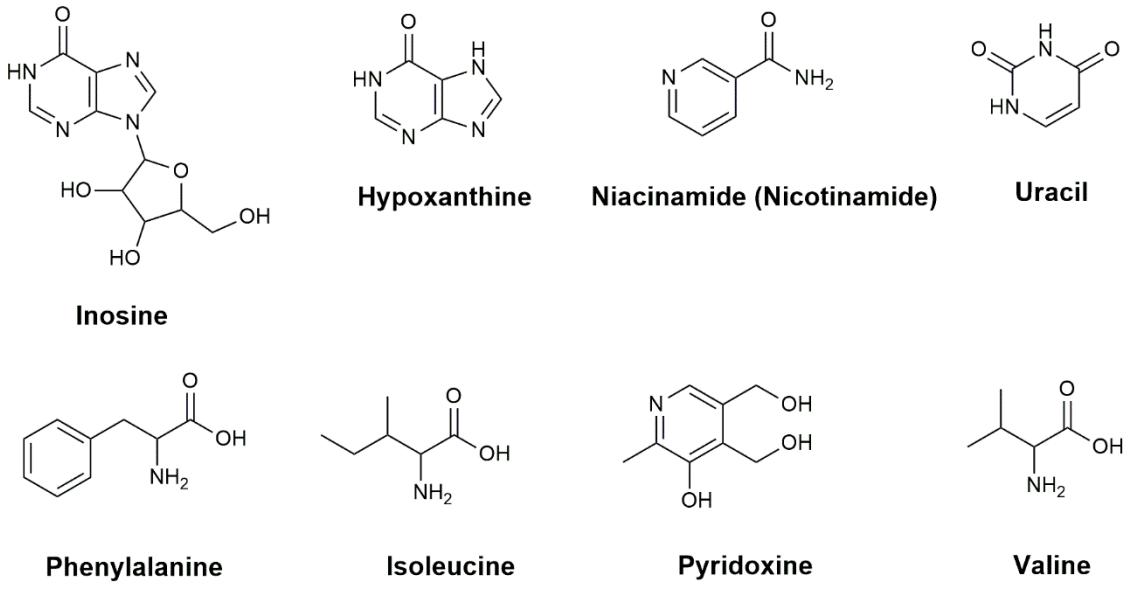
762

763 **Fig. 1.** Genomic overview and accompanied metabolic extents of 9 different NTHi strains. Circular
764 map and total genome information of a representative strain (H188) used in this experiment, showing
765 the GC content, GC skew, CDS, tRNA, rRNA, and tmRNA (**a**). The circular map was created using
766 the fasta sequence of H188 strain in Proksee (www.proksee.ca). Subsystem distribution of different
767 isolated strains based on the data obtained through the RAST server (**b**) (Color codes are based on
768 their numbers obtained from the subsystems of the RAST output). Detailed data for individual
769 cellular processes is provided in Supplementary Table S1_Sheet 6. Comparative analysis of different
770 subsystem distributions in different isolated strains is shown. *Haemophilus influenzae* strain
771 CP000057.2 was used as a reference.



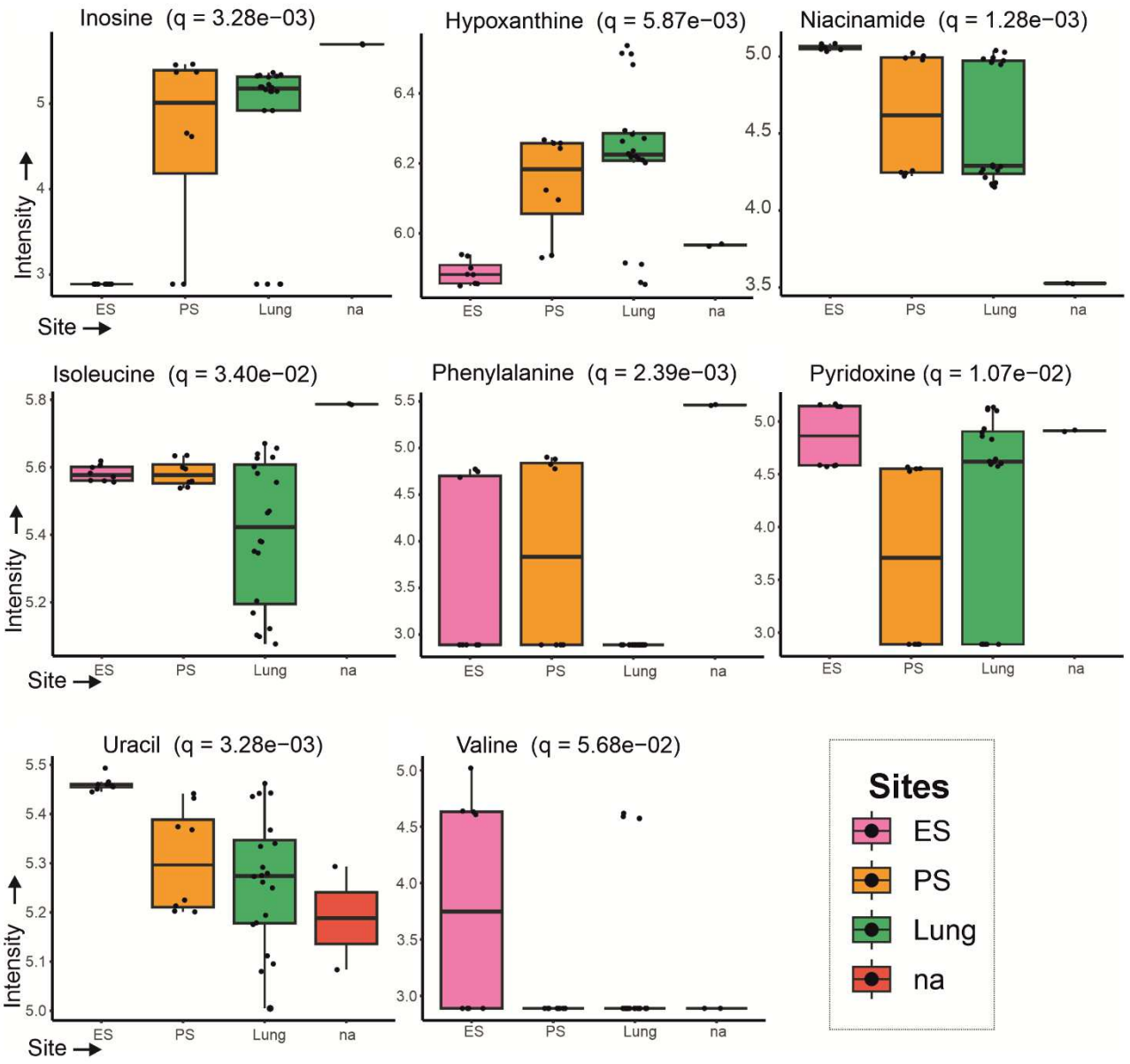
772

773 **Fig. 2.** PCA score plot of isolation site-specific samples (PC1 vs. PC2) based on metabolite
774 intensities. (ES = ear samples, PS = pharyngeal sample, OM = otitis media, P = pneumonia, na =
775 medium control).



776

777 **Fig. 3.** Chemical structures of metabolic signatures from site specific samples.



778

779 **Fig. 4.** Box plots of significantly varied metabolites in site-specific samples. The q values presented
780 in each box plot represent the Kruskal-Wallis false discovery rates, FDR. (na = medium control, ES =
781 ear sample, PS = pharyngeal sample).

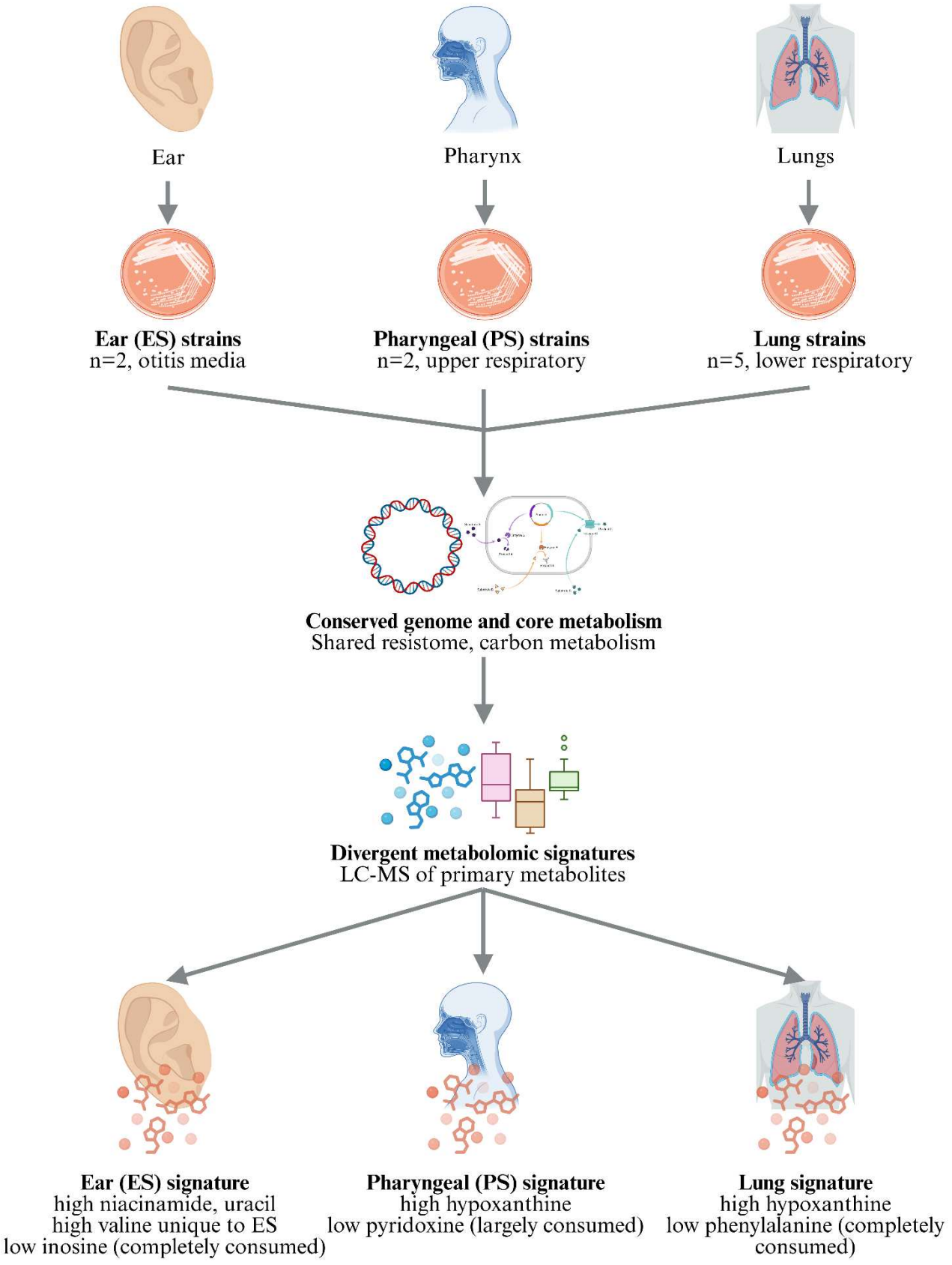


Fig. 5. A graphical and conceptual model summarizing the proposed framework of this study.

784

785

786

787

788

789

Tables:

Table 1. Summary of significantly produced or consumed metabolites in ES vs Lung, ES vs PS, and Lung vs PS ranked by abundance. (S = significant, NS = non-significant, ES = ear sample, PS = pharyngeal sample).

SN	Compounds	ES vs Lung	ES vs PS	Lung vs PS	Abundance ranking
1	1.45_136.06189	S	S	NS	ES > PS > Lung
2	Inosine	S	S	NS	Lung > PS > ES
3	1.15_291.07034	S	S	NS	Lung > PS > ES
4	9.4_190.0866	S	S	NS	ES > PS = Lung
5	Hypoxanthine	S	S	NS	Lung > PS > ES
6	20.03_408.30839	S	NS	NS	ES > PS > Lung
7	Niacinamide	S	S	NS	ES > PS > Lung
8	Uracil	S	S	NS	ES > PS > Lung
9	1.18_427.10863	S	S	NS	Lung > PS > ES=
10	1.56_159.02792	S	S	NS	Lung > PS > ES >
11	1.12_216.92314	S	S	NS	ES > PS > Lung
12	1.13_104.10697	S	NS	S	ES > PS > Lung
13	1.78_86.09661	S	NS	NS	ES > PS > Lung
14	Phenylalanine	S	NS	S	PS > ES > Lung
15	1.15_136.06189	S	S	NS	ES > Lung > PS
16	Isoleucine	S	NS	NS	PS > ES > Lung
17	Pyridoxine	NS	S	NS	ES > Lung > PS
18	Valine	NS	S	NS	ES > Lung > PS
19	1.2_294.15499	NS	S	NS	PS > Lung > ES
20	1.78_294.15506	NS	S	NS	PS > Lung > ES
21	1.19_537.1687	NS	S	NS	PS > Lung > ES

797 **Table 2.** Retrieved number of KO entries of NTHi strains by group for selected KEGG pathways.
798 (ES = ear sample, PS = pharyngeal sample).

Purine metabolism	**KO entries**	**Distribution (%)**	**Vitamin B6 metabolism**	**KO entries**	**Distribution (%)**
All Unified	45		All Unified	6	
ES	44	97.78	ES	6	100
PS	42	93.33	PS	6	100
Lung	44	97.78	Lung	6	100
Pyrimidine metabolism	**KO entries**	**Distribution (%)**	**Phenylalanine biosynthesis**	**KO entries**	**Distribution (%)**
All Unified	27		All Unified	17	
ES	26	96.3	ES	17	100
PS	25	92.59	PS	17	100
Lung	27	100	Lung	17	100
Valine, leucine, isoleucine biosynthesis	**KO entries**	**Distribution (%)**	**Phenylalanine metabolism**	**KO entries**	**Distribution (%)**
All Unified	11		All Unified	2	
ES	11	100	ES	2	100
PS	11	100	PS	2	100
Lung	11	100	Lung	2	100
Valine, leucine, isoleucine degradation	**KO entries**	**Distribution (%)**	**Nicotinate and nicotinamide metabolism**	**KO entries**	**Distribution (%)**
All Unified	3		All Unified	11	
ES	3	100	ES	11	100
PS	3	100	PS	10	90.91
Lung	3	100	Lung	11	100

799

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS7.docx
- SupplementaryTableS2.xlsx
- SupplementaryTableS4.xlsx
- SupplementaryTableS8.xlsx
- SupplementaryTableS6.xlsx
- SupplementaryTableS5.docx
- SupplementaryTableS3.xlsx
- SupplementaryTableS1.xlsx
- SupplementaryfileincludingsupplementaryfiguresBhattaraietal.forsubmission.pdf