UAB
Universitat Autònoma de Barcelona

Exploring the Development, Structure, and Role of the Nasal Microbiota in Piglets

Pau Obregon Gutierrez

Exploring the Development, Structure, and Role of the Nasal Microbiota in Piglets

Pau Obregon Gutierrez

PhD Thesis

Exploring the Development, Structure, and Role of the Nasal Microbiota in Piglets

Tesi doctoral presentada per Pau Obregon Gutiérrez per accedir al grau de Doctor en el marc del programa de Doctorat de Medicina i Sanitat Animals de la Universitat Autònoma de Barcelona, sota la direcció de les Dres. Virginia Aragón Fernández i Florencia Correa Fiz.

Bellaterra, Cerdanyola del Vallès, 2025

La Dra. Virginia Aragón Fernández i la Dra. Florencia Correa Fiz investigadores de l'Institut de Recerca i Tecnologia Agroalimentàries – Centre de Recerca en Sanitat Animal (IRTA-CReSA) y el Dr. Francesc Accensi Alemany, del Departament Sanitat i d'Anatomia Animals de la UAB

Informen:

Que els treballs de recerca descrits en aquesta memòria de tesi doctoral "Exploring the Development, Structure, and Role of the Nasal Microbiota in Piglets" presentada per Pau Obregón Gutiérrez per l'obtenció del Grau de Doctor en Medicina i Sanitat Animals s'ha realitzat sota la seva supervisió i tutoria, i n'autoritzen la presentació a fi de ser avaluada per la comissió corresponent.

I perquè així consti i tingui els efectes que corresponguin, signen aquesta declaració.

Directora Directora Tutor

La investigació i la realització d'aquesta tesi han estat finançades pel *Ministerio de Ciencia y Universidades* a través dels projectes PID2019-388106233RB-I00 i PID2022-138657OB-I00/AEI/10.13039/501100011033, i la beca concedida a Pau Obregon Gutierrez FPU19/02126/ AEI/10.13039/501100011033

Abstract

The animal microbiota is closely linked to host health, contributing to metabolic processes, immune regulation, and defense against pathogens. The nasal microbiota has received relatively little attention compared with the gastrointestinal tract, despite its relevant role in respiratory health. Thus, investigating the nasal microbiota represents a promising strategy for improving the prevention and management of respiratory diseases. This thesis explores the establishment of the nasal microbiota in pigs, its detailed composition, and its association with the host immune system.

The first objective of this thesis was to improve the analysis of the nasal microbiota of piglets, resulting in an optimized pipeline, using DNA from nasal swabs and massive sequencing of a 16S rRNA gene fragment. Subsequently, it undergoes various processing steps to assign the taxonomy, calculate the relative abundance of each taxon, and perform different diversity and compositional analyses.

Next, to study the impact of sow-piglet contact in early life, the nasal microbiota of piglets in controlled environmental conditions but with varying durations of contact with their sows, was compared at weaning. Contact with sows proved to be a major factor affecting the nasal microbial composition of their offspring. Piglets with normal contact until weaning developed a nasal microbiota similar to that of healthy farm piglets, while limited or no contact led to altered microbiotas dominated by atypical taxa.

Besides, efforts to characterize the nasal microbiota composition of domestic pigs frequently report the presence of anaerobic bacteria typically found in the gut, such as *Bacteroidales* and *Clostridiales*, although their presence in the nasal cavity remains poorly understood. The findings in this thesis, not only confirmed the fecal origin of these bacteria, but also demonstrated that their detection in the nasal cavity is not artefactual. Furthermore, 16S rRNA cDNA analyses revealed that these taxa are metabolically active.

Among the microbes found in the microbiota, those with the potential of causing disease, called pathobionts, are of special interest. Here, the genomes of *Mycoplasma hyorhinis* strains coming from different clinical backgrounds were compared to identify virulence determinants. Despite most of the genes were shared across strains, we identified a cluster of health-associated strains with possible differential markers.

Finally, in the current scenario of antimicrobial use reduction in animal production, vaccination is critical. Nevertheless, individual variation in antibody response remains poorly understood. While the microbiota has been linked to antibody response, this has not yet been explored in the nasal microbiota of pigs. In this study, piglets with stronger antibody responses had more diverse nasal and rectal microbiotas. Moreover, swine core nasal colonizers, including *Bacteroidales*, *Clostridiales*, *Moraxella*, *Staphylococcus* and *Neisseria* were linked to higher antibody levels. In the gut, *Clostridiales* showed a positive and *Enterobacteriales* a negative association.

Globally, this thesis provides new insights into the development, composition, and immunological relevance of the nasal microbiota in pigs, paving the way for future strategies to promote respiratory health.

Resumen

La microbiota animal está estrechamente ligada a la salud del hospedador, ya que participa en distintos procesos metabólicos, la regulación del sistema inmunitario y la defensa frente a patógenos. Aunque la microbiota nasal desempeña un papel relevante en la salud respiratoria, ha recibido considerablemente menos atención que la gastrointestinal. Por lo tanto, su estudio representa una estrategia prometedora para mejorar la prevención y el control de enfermedades respiratorias. Esta tesis explora el establecimiento de la microbiota nasal en cerdos, su composición detallada y su asociación con el sistema inmunitario del hospedador.

El primer objetivo fue optimizar el análisis de la microbiota nasal en lechones, desarrollando un protocolo mejorado basado en muestras de hisopados nasales y secuenciación masiva de un fragmento del gen 16S rRNA. Este protocolo incluye varios pasos de procesamiento para asignar la taxonomía, calcular la abundancia relativa de cada taxón y realizar análisis de diversidad y composición microbiana.

A continuación, se estudió el impacto del contacto con la madre en la vida temprana de los lechones, comparando su microbiota nasal al destete en condiciones ambientales controladas, pero con diferente duración en el contacto con sus madres. El contacto con la madre demostró ser un factor clave. Aquellos lechones con contacto normal hasta el destete desarrollaron una microbiota similar a la de animales de granja sanos, mientras que el contacto limitado o nulo condujo a microbiotas alteradas dominadas por taxones atípicos.

Asimismo, frecuentemente se observa la presencia de bacterias anaerobias propias del intestino, como *Bacteroidales* y *Clostridiales*, en la cavidad nasal de cerdos domésticos No obstante, su presencia en esta localización es poco comprendida. Los resultados de esta tesis confirmaron no sólo el origen fecal de esas bacterias, sino que su detección en la nariz no es artefactual.

Además, a través del análisis del cDNA del gen rRNA 16S se comprobó que estas baterias eran activas metabólicamente.

Entre los microorganismos detectados, los patobiontes, microorganismos potencialmente patógenos, suscitan especial interés. Se compararon los genomas de cepas de *Mycoplasma hyorhinis* procedentes de diferentes contextos clínicos para identificar determinantes de virulencia. Aunque la mayoría de los genes resultaron ser compartidos, se identificó un grupo de cepas asociadas a animales sanos con marcadores diferenciales prometedores.

Finalmente, en el contexto actual de reducción del uso de antibióticos en producción animal, la vacunación cobra especial relevancia. No obstante, la variabilidad individual en la respuesta de anticuerpos aún no se comprende completamente. Si bien la microbiota se ha vinculado con la respuesta inmunitaria, la participación específica de la microbiota nasal en cerdos no había sido explorada. En este estudio, los lechones con mejores respuestas de anticuerpos mostraron tener microbiotas nasales y rectales más diversas en el momento de la vacunación. Además, varios colonizadores nasales comunes como *Bacteroidales*, *Clostridiales*, *Moraxella*, *Staphylococcus* y *Neisseria* se asociaron con niveles más altos de anticuerpos. En el intestino, *Clostridiales* mostró una asociación positiva y *Enterobacteriales*, negativa.

En conjunto, esta tesis aporta nuevos conocimientos sobre el desarrollo, la composición y la relevancia inmunológica de la microbiota nasal en cerdos, y abre la puerta a futuras estrategias para promover la salud respiratoria.

Resum

La microbiota animal està estretament vinculada a la salut de l'hoste, ja que participa en processos metabòlics, la regulació immunitària i la defensa contra patògens. Tot i que la microbiota nasal juga un paper rellevant en la salut respiratòria, ha rebut menys atenció que la gastrointestinal. És per això, que la seva investigació representa una estratègia prometedora per millorar la prevenció i el control de les malalties respiratòries. Aquesta tesi explora l'establiment de la microbiota nasal en porcs, la seva composició detallada i la seva associació amb el sistema immunitari de l'hoste.

El primer objectiu fou optimitzar l'anàlisi de la microbiota nasal dels garrins mitjançant un protocol basat en hisops nasals i la seqüenciació massiva d'un fragment del gen 16S rRNA. Aquest protocol inclou diversos passos de processament per assignar la taxonomia, calcular l'abundància relativa de cada tàxon i realitzar anàlisis de diversitat i composició microbiana.

Tot seguit, es va estudiar l'impacte del contacte de la truja amb els garrins durant els primers dies de vida, comparant la microbiota nasal dels garrins al deslletament en condicions ambientals controlades però amb diferents durades de contacte amb la truja. Aquest contacte va resultar ser un factor determinant. Aquells garrins amb contacte total van desenvolupar una microbiota nasal semblant a la d'animals sans de granja, mentre que el contacte limitat o nul va conduir a microbiotes alterades dominades per tàxons atípics.

En diversos estudis s'ha observat freqüentment la presència de bacteris anaerobis típics de l'intestí, com els *Bacteroidales* i *Clostridiales*, a la cavitat nasal dels porcs domèstics Tot i això, aquesta presència és poc compresa. Els resultats d'aquesta tesi no només van confirmar l'origen fecal d'aquests bacteris, sinó també la seva que la seva detecció no és artefactual. A més a més, anàlisis de cDNA del rRNA 16S van revelar activitat metabòlica en aquests bacteris.

Entre els microorganismes detectats, els patobionts, microorganismes potencialment patògens, tenen un interès especial. Es van comparar els genomes de soques de *Mycoplasma hyorhinis* procedents de diferents contextos clínics per identificar determinants de virulència. Tot i que la majoria de gens es compartien, es va identificar un grup de soques associades amb animals sans amb marcadors diferencials prometedors.

Finalment, en l'actual context de reducció de l'ús d'antibiòtics en producció animal, la vacunació és clau. No obstant això, la variabilitat individual en la resposta d'anticossos continua sent poc entesa. Encara que s'ha vinculat la microbiota a la resposta immunitària, no s'havia estudiat aquest vincle en la microbiota nasal dels porcs. En aquest treball, els garrins amb respostes d'anticossos més intenses presentaven microbiotes nasals i rectals més diverses. A més a més, colonitzadors nasals habituals com *Bacteroidales*, *Clostridiales*, *Moraxella*, *Staphylococcus* i *Neisseria* s'associaren a nivells més elevats d'anticossos. A l'intestí, *Clostridiales* va mostrar una associació positiva i *Enterobacteriales*, negativa.

En conjunt, aquesta tesi aporta nous coneixements sobre el desenvolupament, la composició i la rellevància immunològica de la microbiota nasal en porcs, i obre la porta a noves estratègies per promoure la salut respiratòria.

Contents

General introduction

1. The animal microbiota

The term “microbiome (or microbiota)” was firstly used in 1988, as “a characteristic microbial community occupying a reasonably well-defined habitat which has distinct physio-chemical properties” (1). Since then, this term has been extensively used in many fields of life sciences. As all biomes, the microbiome can be divided into the living part, the microbiota, plus the microbial structures, components and the environmental conditions that surround it (2). Thus, the animal microbiota can be defined as the ensemble of microorganisms that inhabits the bodies of the animals, including bacteria, archaea, virus, fungi, algae and protozoa (2–4). These microorganisms are primarily located on body surfaces and cavities exposed to the external environment, such as the gastrointestinal and urogenital tracts, the airways, the oral cavity and the skin. The community composition varies markedly across anatomical sites, according to the distinct physicochemical and biological conditions of each niche (2,5).

Traditionally, host-microbe interactions were viewed through a “separation” paradigm, categorizing microorganisms as pathogens, commensals, or symbionts, and interpreting their relationship with the host as either antagonistic or mutualistic. However, this perspective has shifted with the recognition of opportunistic pathogens and pathobionts, giving rise to a more holistic understanding of microbial-host coevolution (2). In this model, the host and its associated microbiota are considered a single evolutionary and functional unit, the holobiont (**Figure 1**), that coevolves as an integrated system. The host provides a stable, nutrient-rich and protective environment, while the microbiota contributes to essential functions, such as metabolic processing, immune modulation and protection against pathogens (2,4,6). A healthy holobiont is typically associated with a state of eubiosis, characterized

by high microbial diversity, resilience, and functional redundancy. This diversity not only enhances ecological stability but also expands the functional repertoire of the holobiont. In contrast, dysbiosis is often characterized by reduced diversity and compositional imbalance, and correlates with the emergence of a pathobiome, a microbial state linked to disease (2).

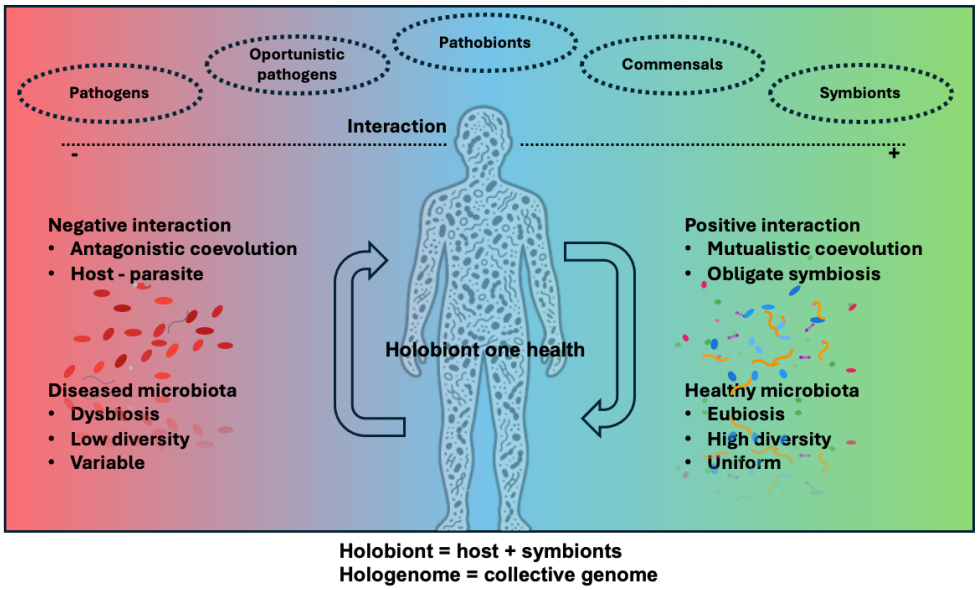


Figure 1. Representation of host-microbes interaction. In the top part, a traditional framework that categorizes microbes based on isolated interactions (e.g., pathogens vs. symbionts) is shown. In the bottom part a holistic view of the holobiont is illustrated. This approach emphasizes that health and disease result not from individual microbes alone, but from the collective behavior, balance, and coevolution of the entire microbial ecosystem in dynamic interaction with the host. Adapted from Berg, *et al*. (2).

Colonization of the host by microbial communities, particularly within the gastrointestinal tract, is associated with enhanced metabolic capacity. This is due to the extensive genetic repertoire contributed by these microorganisms, which collectively encode a gene pool estimated to be approximately 150 times larger than that of the human genome (6). This vast reservoir of

microbial genes significantly expands the host's functional potential. For instance, gut microbiota facilitates the degradation of otherwise indigestible dietary components, including complex polysaccharides, lipids, and proteins, via a diverse array of hydrolytic enzymes. These enzymatic activities contribute to the liberation of bioavailable nutrients and metabolites, playing a critical role in host nutrition, metabolism, and overall physiological homeostasis (2,5,6).

Additionally, the microbiota serves as a primary line of defense against pathogenic microbes, either through direct antagonism or by supporting the stimulation of the host's immune system (4,6), as illustrated in **Figure 2** . Commensal microbes provide direct protection against pathogens through four main mechanisms:

1. Producing antimicrobial compounds like bacteriocins or using type VI secretion systems to kill competitors (4,5).

2.Promoting the formation of a glycoprotein-rich mucus layer that hinders pathogen adhesion and supports beneficial microbes, further enhancing the overall functional capacity of the microbiota (6).

3. Secreting growth inhibitory metabolites, such as short-chain fatty acids or bile acid derivatives (4).

4. Outcompeting effectively pathogens for nutrients and space (environmental niche) (4).

In addition, the microbiota can further contribute to host protection through immune system development, maturation and stimulation through binding of microbe-associated molecular patterns (MAMPs), such as lipopolysaccharide, peptidoglycan flagellin to Toll-like receptors (TLR), or secretion of SCFA (7). These bacterial compounds can contribute to several immune processes, including anti-inflammatory - regulatory cytokines activation, secretion of immunoglobulin A and mucins to the epithelia, downregulation of inflammation pathways (cytokines and macrophages) in absence of infection, sustentation of the production of neutrophils or contribution to T and B cells regulation (4,6,7).

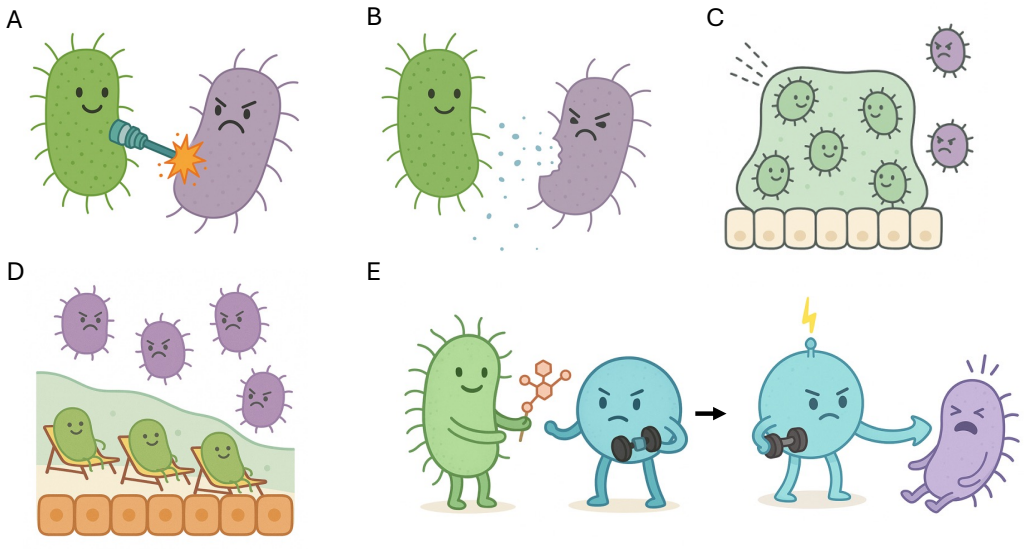


Figure 2. Schematic representation of the beneficial actions exerted by commensal microbiota that contribute to exclusion of harmful pathogens. Here represented: A) Direct killing through type VI secretion system. B) Indirect killing through antimicrobial peptides release. C) Contribution to protective mucus formation and maintenance. D) Niche occupation or outcompetition. E) Immune system development and stimulation through MAMPs.

Furthermore, gut microbial communities are closely linked to the health of various organs through many different compounds in the so-called gut-organ axes. They are involved not only in brain function, affecting mood, social behavior, depression, stress, and neurodegenerative disorders, but also in the health of the respiratory system, liver, kidneys, heart, skin, among others (5,6).

Most of these functions can be compromised when the microbiota is disrupted, generating a multifunctional dysregulation in the host (5,7). In fact, mounting evidence suggests that perturbation of commensal communities in the microbiota is associated with many diseases such as bowel, cardiovascular, respiratory and periodontal diseases, cancer, diabetes, brain and neurodegenerative disorders or systemic diseases affecting specific organs (e.g., kidney, liver), among many others (4,5,7).

2. The pig microbiota

While most microbiota research has historically focused on humans, the study of the animal microbiota offers considerable potential for improving animal health as well. The critical role that the microbiota plays in various aspects of animal physiology, including metabolism, development, immunity, and disease resistance, highlights its potential for applications in veterinary medicine, livestock management, and wildlife conservation. Specifically in pigs, one of the biggest meat industries worldwide (8), microbiota studies are primarily centered around two main goals.: first, to ensure optimal productive, health and welfare parameters in herds, and second, to find feasible alternatives to the use of antibiotics.

As in all animals, the pig microbiota is also shaped by different intrinsic and extrinsic factors acting through lifetime. Piglets initially acquire their microbiota from the vaginal tract during farrowing, and later from the skin, feces and other sow microbiotas, as well as from the farm environment and

early feeding, through maternal milk and colostrum (9,10). Apart from contributing to the establishment of commensal microbes, colostrum is very important in the early immune protection of the piglets (10). During the first life stages, the piglets' microbiota is strongly influenced by the sow until weaning at 3-4 weeks of age (9–11). Weaning is a very critical moment in piglets' life, since they are removed from the sows, mixed with piglets from other litters and changed to solid feeding, causing significant stress in the animals (9,10,12,13). This stress is often accompanied by intestinal alterations (10,14) and microbiota changes (9,13,15), proliferation of pathogens and arise of diseases (10,16,17), and a reduction of growth (9,18). Afterwards, pig microbial communities become more stable and resilient to perturbations(15,19,20). Various external factors influence the microbiota throughout the pig's life, including stress related to piglet processing, antibiotics and other treatments, environmental conditions, aging and interactions with other animals(9,10).

3. The pig respiratory microbiota

3.1 Composition of the pig respiratory microbiota

Among the various microbiotas present in pigs, the respiratory microbiota appears to play an important role in maintaining respiratory health. This is especially relevant given the high prevalence and impact of respiratory diseases in swine populations (21,22). Thus, a well-balanced nasal microbiota may play a significant role in excluding respiratory pathogens, activating and modulating immune responses in the respiratory tract, and maintaining epithelial integrity and function through the stimulation of protective mucus and release of other compounds (5,7,23–25). Conversely, disruptions or imbalances in the microbial communities could impair these vital functions, potentially increasing susceptibility to respiratory infections (26) (**Figure 3**). For these reasons, in-depth investigation of the respiratory

microbiota is essential in research focused on improving swine health and developing effective disease prevention strategies.

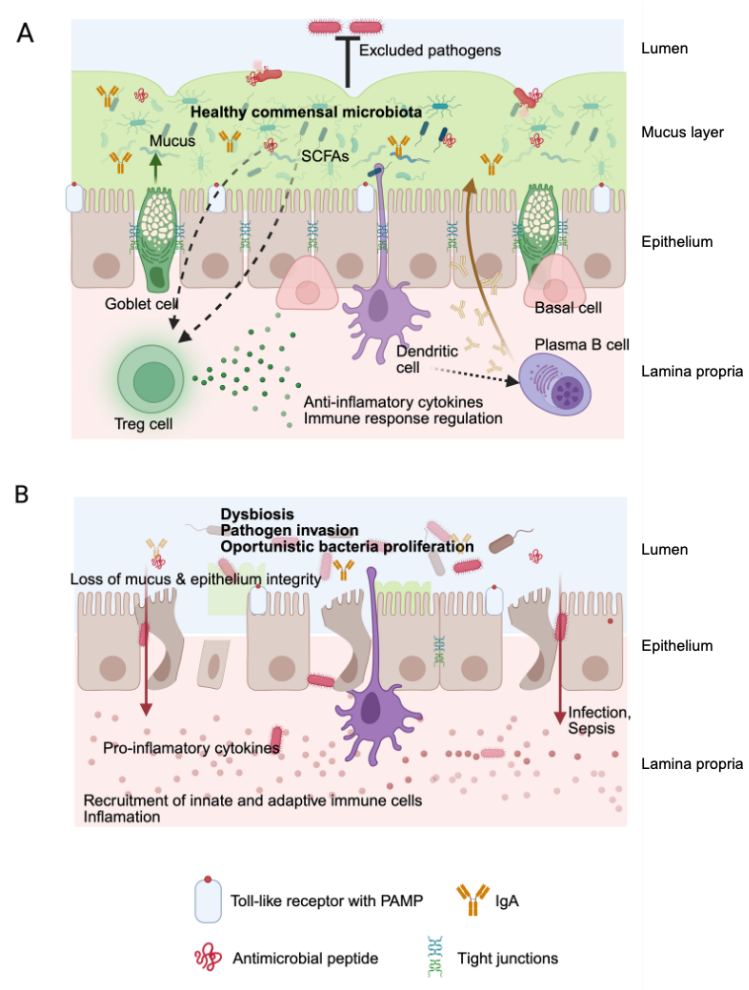


Figure 3. Nasal epithelium conditions in health (A) and disease (B). A) commensal microbiota contributes to pathogen exclusion, immune system stimulation and mucus formation (goblet cells). Dendritic cells sample the lumen and promote regulatory T cells (Immunomodulation/tolerance) and B cells to produce immunoglobulin (Ig)A, which blocks bacterial translocation across the mucosal barrier. Basal cells have the Progenitor/repair role. B) Loss of commensal microbiota and the previous functions. Epithelium loses integrity and pathobiont or opportunistic bacteria are translocated. Innate immune and adaptative (later on) immune responses are triggered producing more inflammation and damage. Adapted from Martel, *et al.* 2022 (7) and Hou, *et al.* 2022 (5) with information from (23–25).

The pig respiratory system consists of a series of organs divided into the upper respiratory tract (URT), comprising the nose, pharynx and larynx, and the lower respiratory tract (LRT), formed by the trachea and lungs (26). Since the URT is more exposed to the environment, its colonization starts during and soon after farrowing either from the sows' vaginal, skin and fecal microbiotas but also from the environment (9,27). According to the environmental influence, higher bacterial loads and microbiota diversity were identified significantly in URT compared to LRT, decreasing from the nasal cavity and tonsils to the trachea, bronchi and lungs (28). In humans, several studies reported that bacteria in the URT at birth resembles the mothers' vaginal microbiota in individuals naturally born and the skin microbiota in individuals born by C-section (26), but also that the main nasal microbial source in early life is the maternal nasal microbiota (29). The microbiota of newborn piglets is variable and strongly influenced by the litter of origin (27,30), but later on it stabilizes in a more resilient community after weaning (15).

Regarding the microbiota composition in the URT, it is dominated by *Proteobacteria* and *Firmicutes*, while *Bacteroidetes* and *Actinobacteria* may be found in lower prevalence and abundance (26,28). At genus level, the nasal cavity is dominated by *Moraxella*, together with *Glaesserella*, *Streptococcus*, *Lactobacillus* and *Clostridium* (15,16,26,28). Despite the same phyla (and mostly the same genera) are also found as in high abundance in other URT organs, such as the tonsils and oropharyngeal cavity, the proportions at genus level differ (26,28).

Although the LRT was historically believed sterile in health, mounting evidence has confirmed that microbes are present and shaped by a balance between immigration from the oropharyngeal microbiota and elimination by mucociliary clearance, coughing and the host immune system (28,31). In agreement, Pirolo, *et al*. showed that the microbiota in the posterior part of the URT (choana) resembled the one in the trachea, suggesting a URT

microbial source for the LRT (28). However, it is still a debate whether microbes coming from the URT that are in constant elimination can be considered a true microbial community. In any case, analyses comparing the microbiota of the URT with LRT reported different communities between the two sites (28). Although there is a big variation between studies targeting the microbiota of the LRT, results point towards a dominance of *Firmicutes*/*Proteobacteria* with big variations at genus level, possibly enhanced by the limitations of working with the low biomass in lungs and methodological differences between studies (26,32). A recent LRT microbiota characterization identified *Glaesserella*, *Streptococcus*, *Clostridium* and *Escherichia* (among others in lower abundance), while *Moraxella* would be more URT-associated (28).

3.2 The environment-nasal microbiota-host connection

Interest in studying the pig nasal microbiota as a potential strategy to prevent swine respiratory diseases has grown significantly in recent years. Multiple studies have approached this microbiota from various angles, aiming to better understand its composition, dynamics, and role in health and disease. Taken together, these studies have revealed a connected system in which environmental factors, microbiota composition, and host health are related (**Figure 4**). While some studies have explored how various environmental factors influence the composition and structure of the pig nasal microbiota (together with other respiratory tract organs), others have concentrated on uncovering the functional roles this microbial community plays in host physiology and health.

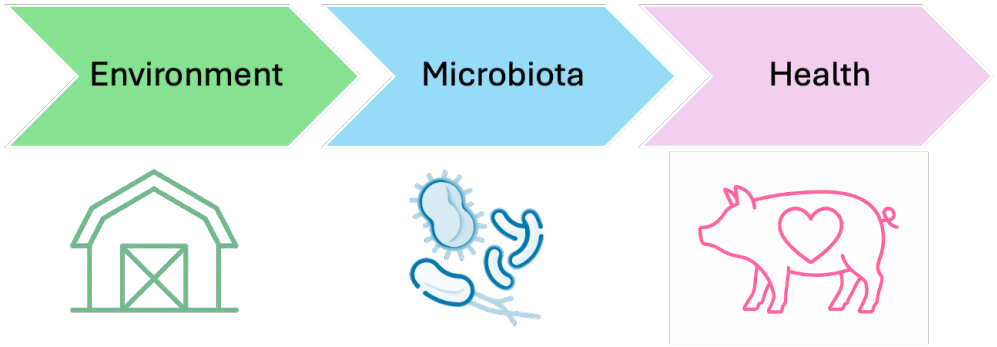


Figure 4. Representation of the environment-microbiota-host connected system where the pig health is directly related with the nasal microbiota and thus, to the environmental agents that may shape its composition.

Regarding the first connection, the most extensively studied environmental factors are those that can disrupt the microbiota composition, such as various treatments commonly used in routine swine industry practices. The antibiotic treatment is maybe the most studied effect on the swine nasal microbiota as environmental perturbation (20,33–36). Collectively, these studies indicate that, although antimicrobials are effective tools for controlling infectious diseases, their use can disrupt the nasal microbiota potentially leading to a range of associated adverse effects. In fact, a previous study demonstrated that removal of metaphylactic antibiotics was a feasible strategy to restore microbial community balance and enhance the health and productivity of pigs (34). Moreover, the use of nasal probiotics has also been suggested as a promising approach to support the restoration of altered microbial communities (20). Vaccination has also been shown to significantly shape the nasal microbiota as well. In this case, piglets born to sows vaccinated against virulent *Glaesserella parasuis* showed a reduction of the microbial diversity as well as in the *G. parasuis* presence (37).

In addition, several studies have shown that respiratory infections in pigs, either viral or bacterial, can induce moderate shifts in the respiratory

microbiota, including changes in community composition, bacterial abundances, and, in some cases, reduced diversity(38–43). Nevertheless, the extent and consistency of these effects vary across studies. Consistent with these findings, microbial differences between healthy and diseased pigs, whether of bacterial or viral origin, have been reported in the URT (44–46), as well as in LRT (32,47,48).

The impact of farming conditions has been studied as well. For instance, the exposure to gaseous ammonia (common in intensive farms) was shown to negatively impact the nasal microbiota by reducing its diversity and some commensals' relative abundance (49). The influence of different animal housing has also been investigated, revealing a significant impact on the microbiota, where environments with increased physical complexity, such as straw-based housing with reduced cleaning frequency, were associated with less optimal microbial profiles (50). Although it was not the main scope of the study, Weese, *et al.* reported that farm management practices, including diet type and antimicrobial exposure, significantly influenced the composition and richness of the nasal microbiota in slaughter-age pigs (51). The impact of the diet was further studied, revealing that dietary additives were associated with alterations in the microbiota, particularly within the nasal tract, and with a reduction in clinical signs (52). In fact, respiratory immunity can also be triggered from the gut through the so-called gut-respiratory axis (52–54) emerging as a promising strategy to improve respiratory health in mammals (55,56).

The intrinsic effect of aging has also been evaluated, and changes and successions in nasal and tonsillar microbial communities have been described in studies collecting longitudinal data (13,15,20,27,30). Generally, these studies show that the porcine respiratory microbiota, both nasal and tonsillar, matures progressively into a more stable and resilient community. Moreover, these microbial communities tend to become more similar across individuals, typically observed after two or three weeks of age (13,15,27).

Nevertheless, this trend has not always been observed (30). Dominant taxa remain relatively stable, but notable shifts in their relative abundances occur over the animal's lifespan (15,20,27). In addition, aging has been associated with a gradual increase in alpha diversity, reflecting a progressive enrichment of the microbiota during the first weeks of life (15,20,30). Weaning is often associated with marked changes in the microbiota, including abrupt shifts in community composition and structure (13,15,27). These variations were linked to concurrent changes, such as dietary transition, environmental change and antimicrobial administration.

Several studies focused on elucidating the role of the pig nasal microbiota on the host. For instance, variations in the diversity and composition of the nasal microbiota at weaning have been associated with the predisposition of developing diseases such as Glässer's disease (16), *Streptococcus suis* (30) and *Mycoplasma hyorhinis* (17) associated disease. Also, the nasal microbiota has been proposed to modulate *Staphylococcus aureus* colonization. A study reported significant associations between carriage status and specific microbial profiles, such as a higher abundance of bacteria with probiotic potential in non-carriers and a predominance of potentially pathogenic taxa in carriers (57), where other did not demonstrate a conclusive impact (51). Furthermore, the influence of the nasal microbiota in a viral disease outcome was revealed (58), showing that the nasal microbiota composition can determine the disease outcome of the animal in a herd naturally infected with a highly virulent PRRSV-1 variant. With a different focus, the oropharyngeal microbiota was associated with pig productive performance, since its composition may also influence the animal weight gain (59).

All environmental factors and host-effects studied for the respiratory microbiota of pigs are summarized in **Figure 5.**

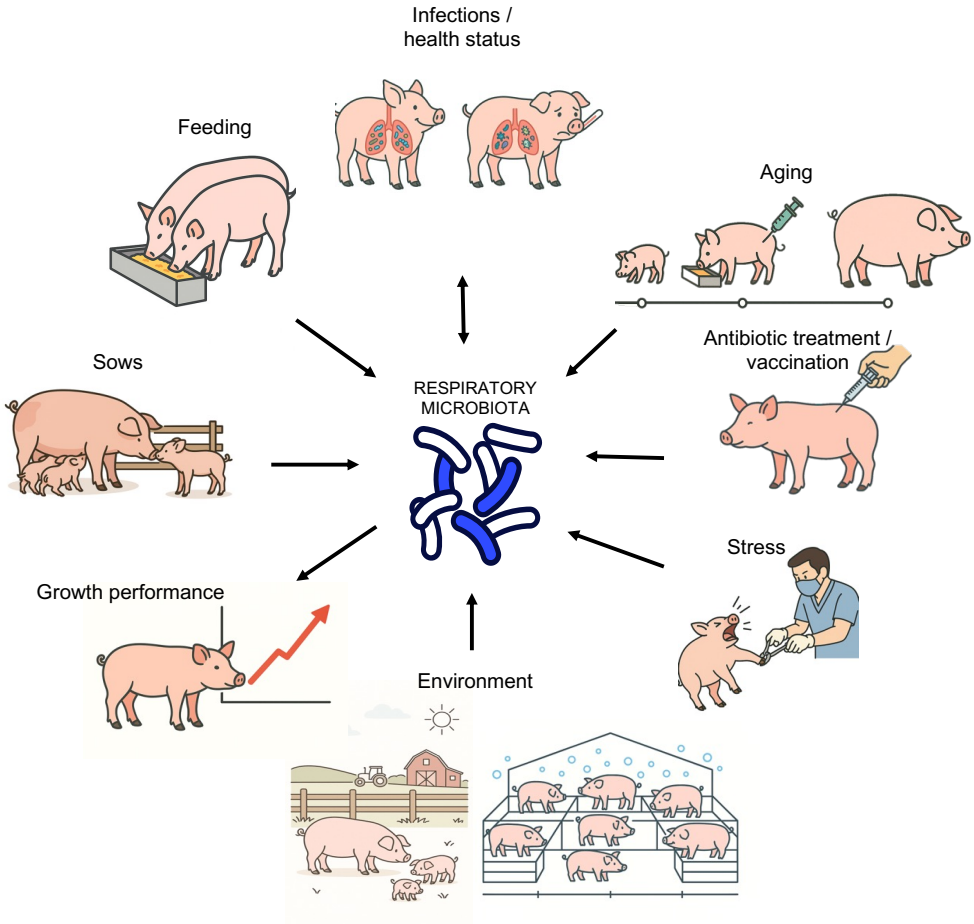


Figure 5. Overview of the findings yielded by studies examining the swine respiratory microbiota, including both factors reported to modulate its composition and microbiota-mediated effects on the host. Arrows represent the directionality of associations described in the literature.

3.3 Main respiratory bacterial pathogens of piglets

All microbiotas harbor a fraction of pathobionts, bacterial species normally present but with the potential to cause systemic infections and disease under certain circumstances (4). In the case of the pig nasal microbiota, the most common pathobionts associated with early-life respiratory infections are *Glaesserella parasuis*, *Streptococcus suis* and *Mycoplasma hyorhinis*, which

are particularly recurrent and prevalent in the swine industry (21,60). These bacteria can spread beyond the respiratory tract, leading to systemic lesions such as polyserositis (8,21,22,60).

3.3.1 *Streptococcus suis*

Streptococcus suis is a Gram-positive bacterium capable of colonizing almost all anatomical sites in swine, transmitted vertically at birth and subsequently through horizontal routes, mainly via direct contact and aerosols (61). Although the palatine tonsil is considered its main niche (30), it is commonly detected in other sites of the upper respiratory tract, as well as in the intestinal and urogenital tracts, showing high prevalence across all age groups (21,62).

Currently, 35 *S. suis* serotypes were initially identified, although some were later reclassified as other species (63,64). Serotype is 2 the most abundant worldwide as well as the most associated with pig and human diseases (65). Although there is no universal agreement on the determinants of virulence (66), several factors have been described, including capsular polysaccharides (CPS), surface components and secreted effectors. These contribute to evasion of phagocytosis, systemic dissemination, adhesion, cytotoxicity and induction of an exaggerated inflammatory response (21,62,65).

It is frequent to find multiple strains of *S. suis*, even from different serotypes, within a single animal. However, clinical outbreaks in a herd are typically caused by a single strain (21,62), unless co-infections with other pathogens compromise the host's immune system (67) and allow multiple strains to establish infection. In diseased animals, *S. suis* initially colonizes the respiratory mucosa, then breaches the epithelium and disseminates systemically through the bloodstream or lymphatics. The bacterium is also capable of crossing the blood-brain barrier (61), causing meningitis. Alternatively, systemic dissemination may originate from the gastrointestinal tract (21,62).

Clinical manifestations of *S. suis* infection are nonspecific and overlap with other infections, requiring isolation and identification of the pathogen for definitive diagnosis (21,62). The most common signs include fever, depression, inappetence, and lameness due to arthritis. Less frequently, nervous signs, anorexia, abortions and lesions (e.g. endocarditis, pneumonia, polyserositis and vaginitis) may occur (21,62,65) .

Treatments with penicillin accompained by other antibiotics are widely used for therapeutic intervention (65). In addition, *S. suis* is susceptible to common disinfectants.

3.3.2 *Glaesserella parasuis*

Glaesserella parasuis, formerly *Haemophilus parasuis*, is a Gram-negative bacterium that colonizes the nasal cavity of piglets shortly after birth (28,68). It is transmitted vertically from sows to piglets and horizontally through direct contact (69). *G. parasuis* is considered a common pathobiont within the porcine respiratory microbiota. During early life, piglets are protected by maternal immunity, which declines as the piglets' own immune system matures. Factors, such as early weaning, suboptimal management, stress, co-infection with other pathogens and alterations in the nasal microbiota composition, have been associated with disease onset (69).

To date, 15 serovars have been described (70), with serovars containing virulent and avirulent strains (69). Virulent and commensal strains of *G. parasuis* can coexist within the same herd, or even within a single animal (69). Nevertheless, outbreaks are usually related with one strain (71). PCR targeting *vtaA* genes (virulence-associated trimeric autotransporters) allows discrimination between virulent and avirulent strains (72).

Avirulent strains typically remain restricted to the nasal cavity and are effectively cleared by host defenses, particularly porcine alveolar macrophages (PAMs), upon reaching the lung (73,74). These strains are also incapable of surviving in the bloodstream. In contrast, virulent strains evade

pulmonary and serum innate immune responses, delaying PAM activation and surviving in the blood to invade systemically (75). The resulting inflammatory response leads to polyserositis lesions, typical of Glässer's disease (69). Although the role of *G. parasuis* as a primary respiratory pathogen remains unclear, it is frequently isolated from lungs of animals with respiratory disease, particularly in the context of secondary infections involving other viral or bacterial agents (69,74). Clinical manifestations are nonspecific and may include fever, coughing, abdominal breathing, swollen joints, lameness and neurological signs in acute cases. In chronic presentations, reduced growth and tissue lesions may be observed (8,21,69). Confirmation of clinical and pathological diagnosis needs to be performed in the laboratory.

While commercial vaccines are available, *G. parasuis* infections are typically treated with antibiotics such as synthetic penicillin, ceftiofur, ampicillin and enrofloxacin (21,69). Nevertheless, antimicrobial resistance is increasingly reported, highlighting the importance of alternative strategies such as vaccination (8,69). Also, sow vaccination is being explored to enhance passive immunity, delay colonization in piglets, or selectively eliminate virulent strains from herds while maintaining commensal ones (68,76).

3.3.3 *Mycoplasma hyorhinis*

Mycoplasma hyorhinis is a small wall-less bacterium that inhabits the porcine respiratory microbiota, primarily colonizing the nasal cavity and tonsils. Transmission occurs both vertically and horizontally via direct contact or aerosols, contributing to the pathobiont's endemic persistence within herds (21,77,78).

Although no definitive associations have been established between genotype and virulence (79), strains exhibit variability in dissemination potential and pathogenicity (80–82). It remains unclear whether these differences stem

from genetic variation or from expression changes driven by environmental factors.

M. hyorhinis has been isolated from the lungs of pigs with respiratory disease (78). Nevertheless, its role as a primary respiratory pathogen remains controversial, as it is more commonly involved in secondary infections in the presence of agents such as PRRSV or PCV2 (21,78). It is also frequently recovered from asymptomatic animals, supporting its opportunistic nature (79).

Under specific, yet poorly understood conditions, certain strains are capable of systemic dissemination, causing various forms of serositis, including polyarthritis, pericarditis, pleuritis, peritonitis and meningitis (77,78), as well as otitis (83), conjunctivitis (84,85) and abortion (86). Associated clinical signs include fever, depression, reduced appetite and movement, respiratory distress, abdominal tenderness, joint swelling, and lameness (21).

Due to the nonspecific nature of these symptoms, etiological diagnosis relies on tissue sampling followed by culture or PCR confirmation (78). To treat *M. hyorhinis*, antimicrobial therapy, typically involving fluoroquinolones, macrolides, pleuromutilins or tetracyclines, can be employed; however, resistance is rising. Promising alternatives include autogenous vaccines derived from at least three local strains (21,78).

3.4. Control of swine respiratory pathogens

Swine respiratory infections can be managed through two main strategies: prevention and treatment. Preventive measures focus on maintaining hygienic husbandry practices, preventing primary infections and disrupting pathogen transmission within herds, both of which are strongly associated with effective control of most respiratory pathogens (21,69). Although perinatal metaphylactic antibiotic treatments have been used to prevent

bacterial infections and promote growth (although this use was banned in Europe in 2006) (87–89), their use is increasingly discouraged due to associated risks and limited long-term benefits (87–89). Despite other preventive treatments such as vaccination with autogenous and commercial vaccines have been developed and are especially important to control viral infections (21,65,69,78), its use is still minor regarding bacterial infections in the swine industry (8). Treatment is mainly focused on using different antibiotics. Nevertheless, although antibiotics remain essential in modern medicine, their use is closely associated with the emergence of resistant bacterial strains and the disruption of beneficial microbial communities (20,33–36,90). Therefore, the development of alternative strategies for managing bacterial infections has become a critical research priority, as reflected in institutional calls (91).

Host-targeted strategies stand as a promising option, especially those focused on protecting and supporting the beneficial microbiota, which deserve more attention (36). Beneficial microbiota can be modulated to support host health by preventing harmful factors from disturbing its composition or through different complementary strategies such as probiotics, prebiotics and postbiotics. Probiotics are live commensal microorganisms that provide beneficial functions to the host. Prebiotics, on the other hand, are defined as 'substrates selectively utilized by host microorganisms which in turn confer a health benefit. Postbiotics refer to bioactive compounds resulting from bacterial activity, including microbial components, which exert positive effects on the host (92). In addition, other microbiota modulation strategies, such as microbiota transplantation and phage therapy, can be used to modulate the microbiota composition (5,6,92,93).

Despite such alternatives have been studied and even applied in human medicine for both gut and nasal microbiota (92–94), research on swine microbiota, particularly beyond the gut (9,95), remains limited (52). Up to now, few studies have evaluated the effect of probiotics in the swine respiratory

tract, showing the immunomodulatory capacity of *Bacillus subtillis* and a *Lactococcus spp.* cocktail in the nasal mucosa (96,97), or the capacity to modulate the microbiota composition with five porcine nasal colonizers and ameliorate changes introduced by antibiotic administration (20). Therefore, advancing research on microbiota modulation represents a promising alternative to reduce antibiotic use. By supporting the development of a stable and beneficial microbial community, these strategies may significantly enhance disease resistance and offer substantial health benefits for the swine industry.

4. Study of the microbiota

Metagenomic studies targeting the animal microbiota are mainly focused to understand the microbial community composition (i.e. “who’s there?”) and the functionality that these microbes bring (i.e. “what can they do?”) (98). The microbiota can be approached using culture-dependent and culture-independent techniques (6). Culture-dependent microbiota studies are based on culturing, isolating and identifying microorganisms from a microbiota sample, but the number of bacteria isolable by culture-dependent methods is limited in comparison to the true size of the microbiota (6,99). The way to overcome this issue is through culture-independent techniques. Initially methods based on PCR like denaturing gradient gel electrophoresis, fragment polymorphisms, Sanger library sequencing or techniques based on DNA hybridization allowed to investigate bacterial diversity (6). Nevertheless, culture-independent techniques for the study of microbial populations quickly evolved with the appearance of next-generation sequencing (99,100), which allowed massive sequencing more efficiently in terms of effort, cost and time (6).

The sequencing of DNA can be targeted (i.e. amplicon sequencing) or untargeted (i.e. whole-genome shotgun sequencing), with different pros and cons that are summarized in **Table 1** (98,99).

With shotgun metagenomics virtually all the DNA extracted from a sample is sequenced, enabling comprehensive taxonomic resolution down to the strain level (99,101). It also allows the recovery of all functional genes present, offering direct insights into the functional potential of the microbial community (99). Moreover, it captures the entire community, including bacteria, archaea, viruses, and eukaryotes (99) . Despite these advantages, shotgun approaches require significantly higher sequencing coverage, which increases the cost, data complexity, and computational requirements (98,99). In addition, non-desired contaminant DNA can become a problem, especially for low-biomass samples (102).

On the contrary, amplicon sequencing is directed to a specific marker, mainly to the 16S rRNA gene, taking advantage of its nine hypervariable regions to distinguish between bacteria (100,101), using one of the available broad databases of 16S genes: Greengenes, Silva or Ribosomal Database Project (103–105). In marker gene-based approaches, sequenced amplicons are processed to identify distinct sequence variants and determine their abundances, which are then used to infer the taxonomic composition and diversity of the microbiota (100,101). These methods are cost-effective, computationally efficient, and well-suited for comparative studies. However, they present the limitation of having a resolution generally restricted to the genus or species level, which can hinder the differentiation of closely related strains (98,99,101). Additionally, they do not directly provide functional information, as only a marker gene is sequenced. Although functional inference tools such as PICRUSt or Tax4Fun are available (106,107), these approaches rely on indirect predictions and may be subject to inaccuracies. Furthermore, PCR-based amplification in this technique introduces bias,

primarily due to primer / region selection, but also because of differential efficiency, and chimera formation (99,100).

Additionally, studies inferring the microbiota composition from retrotranscribed rRNA have arisen with the aim to identify active populations inside a whole microbial community. While this technique offers valuable insights, it also has its limitations, as rRNA levels do not always perfectly correlate with microbial growth or activity, although a general association is often observed (108).

Table 1. Overview of the two main microbiota profiling approaches based on second generation sequencing tools.

	16S rRNA	**Whole-genome shotgun**
Target	Specific regions of the 16S rRNA gene (bacteria / archaea)	All DNA in the sample (bacteria, archaea and eukaryotes)
Taxonomic resolution	Genus level (sometimes species)	Strain level
Functional profiling	Indirect (functional prediction tools)	Directly from sequenced genes
Coverage required	Low (targeted amplification)	High (in order to capture low abundant microbes / genes)
Cost	Lower	Higher
Contaminant DNA	Lower (targeted amplification reduces noise)	Higher (non-target DNA introduces noise, especially host-DNA)

Best suited for	Microbial community profiling, especially in low biomass samples	Comprehensive microbiome analysis, including functional potential
Main limitations	• Limited to bacteria/archaea • Confident to genus level • Lacks functional insights • Depends on PCR	• Higher cost • Complex data analysis • More susceptible to contamination (non-desired DNA)

Sequencing methods are susceptible to other types of biases that can be introduced during microbiota sampling, sequencing and analysis, including differences in laboratory protocols, parameter choices, and computational pipelines (26,100). For instance, results can vary according to the sample type, sampling methodology, sample processing, sequencing strategy (including library preparation and chosen platform), marker gene region (in the case of 16S rRNA sequencing) and bioinformatic tools / databases used, among others (26,99,100). This methodological heterogeneity across studies can further amplify variability and hinder direct comparisons. Hence, reporting all methodological steps and addressing standardization remains a major challenge in microbiota profiling.

Recent advances in third-generation sequencing technologies, like PacBio SMRT and Oxford Nanopore, offer key advantages over conventional 16S rRNA amplicon sequencing. Unlike short-read methods (~250–300 bp), they can sequence the full 16S gene (read lengths up to 10,000 bp), improving taxonomic resolution down to species and strain levels (6,98). These platforms also support PCR-free metagenomic sequencing, avoiding amplification bias (6). Furthermore, their long reads enable better genome or functional genes assembly in complex communities (Scholz). Moreover, they

can detect epigenetic modifications like DNA methylation, providing insights into host–microbe interactions through gene regulation (6). These improvements make third-generation sequencing a promising addition to the current microbiome research toolbox.

Motivation, Hypothesis and Objectives

The nasal microbiota has a crucial role in pig health. Proper establishment of these microbial communities and perturbations in early life can significantly influence the host's health throughout life. However, many aspects remain poorly understood and require further study. This thesis aims to understand the establishment, the detailed composition and the role of the nasal microbiota to identify beneficial members that may help developing future strategies for the stabilization of the microbiota through rational manipulation. Specifically, this thesis will investigate: the role of sows in shaping the early microbiota of piglets, the occurrence of anaerobic bacteria and virulent pathobionts in the nasal cavity, and the contribution of the nasal microbiota to the host immune response.

The **hypothesis** of this thesis is that the swine nasal microbiota consists of a relatively defined community acquired in early life that is connected to the health of the host through the interaction with the immune system.

The **main objective** of this thesis is to investigate the swine nasal microbiota composition, the factors driving its development, and its functional role on the host health.

The **specific objectives** are:

1. Establish an optimized and standardized pipeline for the study of the nasal microbiota of pigs.
2. To investigate to which extent sows are important for the early establishment of the nasal microbiota in piglets.
3. To expand the characterization of bacterial members of the swine nasal microbiota, including anaerobes.
4. To study the role of the nasal microbiota in the antibody response to vaccination.
5. To explore genomic variations among *M. hyorhinis* strains from different clinical origins with the goal of identifying virulence markers.

CHAPTER 1

Analysis of the Nasal Microbiota in Healthy and Diseased Pigs

(Adaptation of a 16S rRNA gene microbiota analysis pipeline for swine nasal microbiota studies)

Pau Obregon-Gutierrez, Virginia Aragon, Florencia Correa-Fiz. *Methods Mol Biol*. 2024; 2815:93-113. doi:10.1007/978-1-0716-3898-9_8

Abstract

Massive sequencing of a fragment of 16S rRNA gene allows the characterization of bacterial communities in different body sites: the microbiota. Nasal microbiota can be analysed by DNA extraction from nasal swabs, amplification of the specific fragment of interest, and posterior sequencing. The raw sequences obtained need to go through a computational process to check their quality and then assign the taxonomy. Here, we will describe the complete process from sampling to get the microbial diversity of nasal microbiota in health and disease.

Introduction

The microbiota has attracted significant attention in the last years, especially with the increased availability and cost-effective methodologies for massive sequencing. Most of the microbiota studies have focused on the human gut microbiota, but the microbiota from other mammalian hosts and other body locations is also of major interest to understand the health and physiology of the host and the interactions with pathogens. The microbiota found in the upper respiratory tract is considered one of the first lines of defense of the host against respiratory pathogens. These pathogens need to overcome the microbiota and the immune response of the host (which is in turn modulated by the microbiota) to colonize and infect the lower respiratory tract. Once there, some pathogens such as *Streptococcus suis* can invade and cause systemic disease. The main habitat of *S. suis* is the upper respiratory tract, particularly the tonsils, but this bacterium can also be detected in the nasal cavity as well. Little is known about the role of the microbiota in the tonsil or the nose of piglets in controlling the spread of *S. suis* disease.

To study the different bacterial species composing a sample, traditional bacterial culture techniques were used to allow the biochemical and phenotypical characterization. However, this method is known to have several limitations related to the difficulty to culture obligate anaerobes or to control their survival during the transport or storage of the samples under standard procedure (109). The molecular approaches that appeared late in the twentieth century can overcome these restraints and rapidly became the gold standard to characterize these complex communities. Shotgun metagenomics or 16S rRNA amplicon sequencing are among the methods of choice to perform this kind of analysis with different limitations (110). Metagenome sequencing includes the massive sequencing of fragmented DNA that needs to be assembled in full or partial genomes through a complex and computing-demanding process. This sequencing method does not depend on amplification; hence it avoids any bias related to this, but it has its

own caveats. Computational limitations together with biases related to host sequencing or other non-target DNA sequencing are among the most common problems found. It is also significantly more expensive and challenging from the bioinformatics analysis point of view. On the other hand, amplicon sequencing, the most used method in the field of microbiome research (111), involves the amplification of a particular region of interest. Sequencing the 16S rRNA gene is typically used due to its high conservation throughout the whole *Bacteria* domain allowing the amplification of all bacterial members using degenerate primers designed to target this region. But also, within the V3–V4 regions typically targeted in this gene, there are highly divergent regions that allow distinguishing between different bacterial genera. The obtained reads can be taxonomically classified to infer the bacterial composition of the sample by comparison to a reference database. This method, apart from having a much lower economic cost, outputs lower complexity sequence data more computationally manageable. However, the classification obtained from this short region is not reliable at species or strain levels, since individual bacteria can share identical amplicon sequences making impossible the differentiation. Additionally, the amplification step can lead to more opportunities of error directly impacting in bacterial identification from sequences (112). The functional capacities can be inferred from these data, although the shotgun metagenomic approach can give a better resolution of the genes and pathways present in the dataset. Both methods share some technical issues due to DNA extraction, sample manipulation, and in-silico processing analyses. Controlling the methodology becomes crucial to deliver accurate results able to be reproduced and compared between experiments. Therefore, the relevance of defining a workflow from sampling to unravelling the communities' composition and diversity becomes evident.

Materials

Study Design

The study design will depend on the objective of the study, but commonly different conditions of interest are selected to be compared; in this case, the analysis is focused in comparing animals showing clinical signs compatible with *S. suis* disease and healthy pigs. Several factors that can influence the nasal microbiota composition have been determined, including antibiotic treatment (34), contact with the sows (113), diet and food additives (52), or the age of the animals (9). Therefore, they must be controlled among the groups of interest. The type of sample will also depend on the objective. Using nasal swabs is usually enough for sampling the nasal microbiota; however, some authors have performed nasal washes, albeit detecting lower diversity (35). To study the tonsil microbiota, swabs may be considered too superficial, and the whole tissue can be used. It is evident that sometimes the availability of the samples will determine the sample to be used, since tissue samples normally require euthanasia of the animals, and this needs to be justified. In this chapter, we will focus on nasal swab samples.

Laboratory Materials and Equipment

- Sample Collection and Storage

    1.Thin aluminum swabs.

    2.1X phosphate-buffered saline.

    3.Microtubes (“Eppendorf tubes”).

    4.Stabilization solution for nucleic acids, such as DNA/RNA shield.

- DNA Extraction

    1.DNA extraction kits.

    2.Vortex.

37.Centrifuge.

48.Spectrophotometer (e.g., BioDrop).

Methods

Sample Collection and Storage

Before the insertion of the swab in the nasal cavity, the snout needs to be cleaned to avoid introduction of external contamination in the samples. For young piglets, we recommend using thin aluminum swabs due to their narrow nasal passages. The two nares can be sampled with a unique swab (see **Note 1**). Eight to ten animals are recommended to be sampled per group to obtain valid results, ensuring at least five samples per group for the bioinformatics analysis. Since sows are a major source of variability in young piglets, when possible, animals from different litters should be included in each group. Importantly, samples should always be taken in the same standardized way to guarantee unbiased sampling (see **Note 2**).

Samples can be stored in 500 µL of PBS at −80 °C after vigorous vortexing. If transportation of the samples is necessary or freezing is not possible, samples should be placed in DNA/RNA shield, or similar stabilization solution for nucleic acids, immediately after collection (see **Note 3**). In parallel, a clean swab should be conserved in the same fashion to be processed as a negative control.

All the information about the origin of the samples (age of animals, farming system, treatments, clinical status, sow origin, sow parity, sex, feed, etc.), which constitute the metadata, needs to be compiled and organized to enable the subsequent analysis.

DNA Extraction

For DNA extraction, 200 µL of the swab suspensions can be used as a starting material for commercial kits. We have successfully used the Nucleospin Blood kit following the manufacturer's instructions, but using 50 µL of elution buffer to elute the DNA from the column resulted in better yields. After extraction, the quality and quantity of DNA must be assessed with a spectrophotometer to be within the optimal range. For Illumina sequencing, the minimum required DNA concentration is set to 5 ng/µL per sample; however, in our experience, only when DNA extracted from nasal swabs is in the range of 30–200 ng/µL and with a ratio A260/280 higher than 1.8 is the sequencing yield adequate for analysis. Other kits, such as MagMAX Pathogen RNA/DNA kit or ZymoBIOMICS DNA/RNA Miniprep kits, can be used, especially with samples in DNA/RNA shield. Preliminary tests to optimize the methods for DNA extraction and storage conditions may be required with the specific samples and available equipment.

All the samples should be prepared following the same protocol to avoid differences due to the method used. It is always a good practice to include a negative control (a clean swab in the same storage solution), especially if you think that the bacterial load could be low in the sampled animals (such as after an intensive antimicrobial treatment).

Sequencing

DNA samples will follow a standard protocol to produce a library for sequencing (see **Note 4**). This library can be prepared following the manufacturer's recommendations. During this process, primers for sequencing will be incorporated into the DNA fragments, as well as specific barcodes for each sample since several samples are normally included in a single run. This barcoding allows the correct assignation of the reads from each sample (demultiplexing steps). Alternatively, the sequencing library may be performed by a specialized sequencing service. In such a case, it is recommendable to send the DNA samples within 24h post-extraction under

refrigeration (4 °C). If the samples are not processed within this time, storage should be done at −20 °C or below.

There are several different platforms that can be used for amplicon sequencing, but Illumina is the most frequently used. Amplification of the V3–V4 region of the 16S rRNA gene [10] allows the identification of typical nasal taxa using Illumina MiSeq pair-end 2 × 250 bp or 2 × 300 bp technology (MS-102-2003 MiSeq® Reagent Kit v2, 500-cycle or MS-102-3003 MiSeq Reagent Kit v3, 600-cycle). For nasal samples, recommended primers are those used in the standard Illumina protocol (114):

16S amplicon PCR forward primer:

5’TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG

16S amplicon PCR reverse primer:

5’GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC

The cycling conditions include initial denaturation at 95 °C for 3 min followed by 30 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s and a final extension at 72 °C for 5 min. The PCR product must be checked to verify its size after purification (Bioanalyzer DNA 1000 chip).

A number of 30–40 samples per run would output a total of 80,000 high-quality raw sequences per sample, where ~40% reads are filtered out in the following steps. This will allow the analysis to be done at a sampling depth of ~30,000 sequences per sample. The rarefaction curves must be inspected in the downstream analysis to estimate the species richness. These accumulation curves represent the number of species as a function of the number of sequences taken, hence they grow rapidly (as the most common species are found), but they plateau when only rare species are left to be sampled, meaning that the expected number of species is reached. If more than 40

samples are to be sequenced, multiple runs should be performed. In this case, runs should be organized to contain samples from all the groups to be compared. Ideally, one of the samples should be sequenced in different runs as a control, especially in the case the runs are not performed at the same time. All samples (including those from different runs) should reach the plateau at the maximum depth to be included in the analysis, confirming the sample procedure was adequate. Occasionally, a sample fails to yield enough reads and should be excluded from the analysis. Reasons for this problem can be traced to the failure in obtaining enough bacterial material during the sampling procedure, in a deficient DNA extraction, or in problems when generating the sequencing library.

Analysis

The pipeline described below is intended to be used with the newest Quantitative Insights into microbial Ecology 2 (QIIME2) release (115); for this example Version 2023.2 was used. All the documentation is available at https://docs.qiime2.org/2023.2/. The step-by-step pipeline, schematically represented in **Figure 1**, is designed to be performed with Illumina paired-end output raw reads to get the final description of the microbiota composition in different groups (e.g., health and disease) and the diversity (within each sample and between the different groups) of the communities. Not all the parameters or the parameters available for each command are explained, but every command applied is described with the objective that an unexperienced user can go through it.

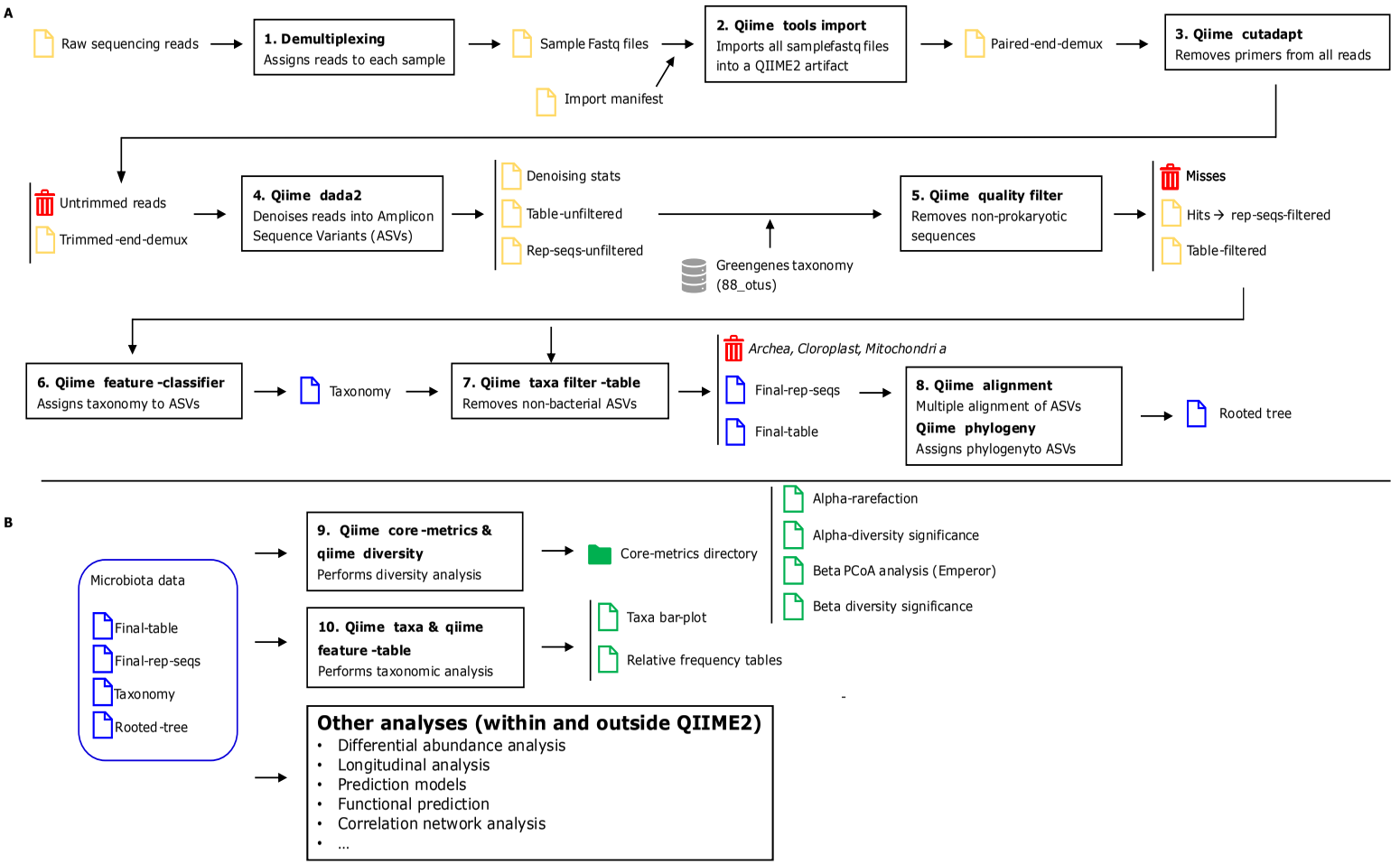

A
Raw sequencing reads
1. Demultiplexing
Assigns reads to each sample
Sample Fastq files
Import manifest
2. Qiime tools import
Imports all samplefastq files into a QIIME2 artifact
Paired-end-demux
3. Qiime cutadapt
Removes primers from all reads
Untrimmed reads
Trimmed-end-demux
4. Qiime dada2
Denoises reads into Amplicon Sequence Variants (ASVs)
Denoising stats
Table-unfiltered
Rep-seqs-unfiltered
Greengenes taxonomy (88_otus)
5. Qiime quality filter
Removes non-prokaryotic sequences
Misses
Hits → rep-seqs-filtered
Table-filtered
6. Qiime feature-classifier
Assigns taxonomy to ASVs
Taxonomy
7. Qiime taxa filter-table
Removes non-bacterial ASVs
Archea, Cloroplast, Mitochondria
Final-rep-seqs
Final-table
8. Qiime alignment
Multiple alignment of ASVs
Qiime phylogeny
Assigns phylogenyto ASVs
Rooted tree
B
Microbiota data
Final-table
Final-rep-seqs
Taxonomy
Rooted-tree
9. Qiime core-metrics & qiime diversity
Performs diversity analysis
Core-metrics directory
Alpha-rarefaction
Alpha-diversity significance
Beta PCoA analysis (Emperor)
Beta diversity significance
10. Qiime taxa & qiime feature-table
Performs taxonomic analysis
Taxa bar-plot
Relative frequency tables
Other analyses (within and outside QIIME2)
• Differential abundance analysis
• Longitudinal analysis
• Prediction models
• Functional prediction
• Correlation network analysis
• ...

Figure 1. Workflow chart illustrating the steps for the in-silico analysis pipeline. A) Preprocessing steps from raw sequencing reads to feature table and intermediate files needed to perform downstream analyses. B) Result-oriented steps from microbiota intermediate files previously obtained to taxonomy and/or diversity analysis. Each box represents a set of commands detailed in the main text and linked by the number. For clarification, preprocessed files are shown in yellow, removed files in red, the external data base in gray, the processed microbiota data in blue, and output analysis files in green.

As in any in-silico analysis, the initial step is to ensure that your computer fulfils the specific requirements to run QIIME2 software, where a minimum of 32-Gb RAM would be enough for most users, but faster CPUs will allow shorter run times. The installation of a software is not a trivial step since, many times, a large list of dependencies is needed to complete the full installation. Fortunately, for QIIME2 core distribution, these steps are well described in the developers' webpage (https://docs.qiime2.org/2023.2/install/) where its native installation is recommended as a conda environment. A virtual machine can be used alternatively but only if the conda installation is not available. The installation and correct activation of the QIIME2 environment can be verified by typing "qiime" in the command line, which should list all the available plugins. The QIIME2 Core 2023.2 distribution includes plugins and interfaces developed, maintained, tested, and distributed by the QIIME2 development team, but additional plugins are available to be installed. All the analyses detailed below are included in the core distribution and can be run in either Mac or Linux operating systems, where the latest versions are always recommended. Many steps in the pipeline can be computationally demanding, so they can be efficiently accelerated by parallelized computing. In the *- -p-n-threads* option, the number of concurrent tasks to be performed at the same time should be set (as included in the examples below, where this option is allowed).

The sample metadata file (named as metadata.tsv) is one of the essential files in this kind of analyses, since it contains all the description of the samples

and will be used to define the study groups, so it is important to include all the data that can be useful in the downstream analysis (see **Note 5**). This file can be created using any spreadsheet program (e.g., Excel, LibreOffice or Google sheet) provided it is saved as a tab-separated-values (tsv) file, where all the collected parameters are located in different columns. Technical details, such as time point, body site, or sequencing run, should be included together with all the information specific for the microbiome study. As an example, to compare health and disease, a column indicating the clinical status of the animals must be included to split the animals into these two groups. If the process was correctly followed, the metadata.tsv file should be ready to proceed with the analyses; however, the file format can be previously validated to avoid errors when loaded in the following step (see **Note 6**).

Once the software is correctly installed and the metadata file validated, the first step is to demultiplex the reads, i.e., sort the reads by sample. QIIME2 provides tools to perform this step (Figure 1, step 1); however, most of the times, the sequencing services provide already demultiplexed samples, and this step is skipped. To import demultiplexed fastq files into a QIIME2 artifact (see **Note 7**), a manifest file must be prepared (Figure 1, step 2). In this comma-separated file (manifest), the unique sample id, the full path of the reads, and their direction (forward or reverse) must be specified (see **Note 8**). Also, it is important to specify in the same command the type of reads (“single” or “paired-end” reads). Raw reads are imported into an executable QIIME2 artifact (qza) which can be converted into a viewable format (qzv) after using *demux summarize* command (see **Note 9**). This qzv file contains a report including the number, the quality, and the length of the reads in the samples. Secondly, primers can be removed with *Cutadapt plugin* (Figure 1, step 3). The primer sequences from the sequencing step should be included to be detected and trimmed (see **Note 10**). The *–p-discard-untrimmed* option allows to filter out those reads without primers, which can be contaminants, since they did not result from the amplification during the sequencing library

preparation. Finally, we must convert the qza format to a viewable qzv artifact as mentioned above. The commands to perform steps 2 and 3 as depicted in Figure 1, are detailed below.

```
qiime tools import \
 --type 'SampleData[PairedEndSequencesWithQuality]' \
 --input-path nasal-PE-manifest \
 --output-path paired-end-demux.qza \
 --input-format PairedEndFastqManifestPhred33

qiime demux summarize \
 --i-data paired-end-demux.qza \
 --p-n 100000 \
 --o-visualization paired-end-demux.qzv

qiime cutadapt trim-paired \
 --i-demultiplexed-sequences paired-end-demux.qza \
 --p-cores 24 \
 --p-front-f CCTACGGGNGGCWGCAG \
 --p-front-r NACTACHVGGGTATCTAATCC \
 --p-match-adapter-wildcards \
 --p-match-read-wildcards \
 --p-discard-untrimmed \
 --o-trimmed-sequences trimmed-demux.qza

qiime demux summarize \
 --i-data trimmed-demux.qza \
 --p-n 100000 \
 --o-visualization trimmed-demux.qzv
```

The intermediate and final visualization outputs (paired-end-demux.qzv and trimmed-demux.qzv) should be checked in QIIME2 view (https://view.qiime2.org/). If the sequencing has been successful, these two

steps should only output few reads of difference (untrimmed reads not containing the primers). After, the quality of the reads at each nucleotide position should be inspected by visualizing the trimmed-demux.qzv file generated, to decide the parameters to be used in the next step: denoising. An example on how to decide dada2 truncation parameters is shown and explained in **Figure 2** and **Note 11**, respectively.

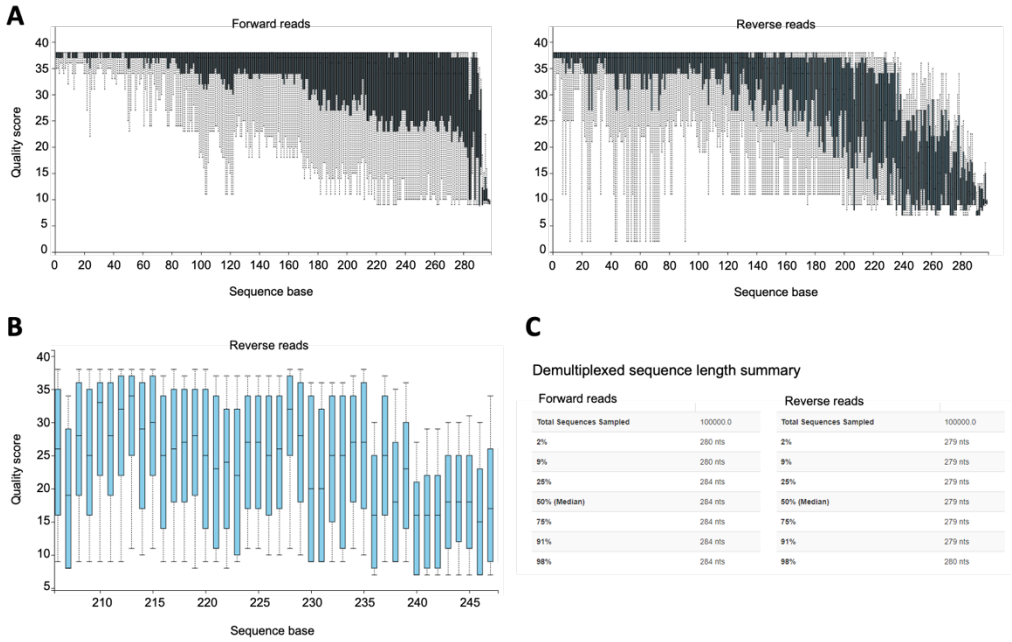


Figure 2. Plots evaluated to decide truncation and trimming parameters (a) Quality plot per base pair of forward and reverse reads. (b) Quality plot of reverse reads (zoomed in). The line in boxplots indicates the mean of the quality at each position. (c) Forward and reverse read length summary.

Dada2 (116) is an algorithm developed to perform the denoising steps, i.e., filtering reads by quality, joining paired-end reads, removing chimeras, and sorting the reads into amplicon sequence variants (ASVs) implemented in QIIME2 in *qiime dada2* command (Figure 1, step 4). Many parameters can be set manually (or used by default) in this step, but it is relevant to pay

especial attention to two of them: *--p-trim-left* and and *–p-trunc-len*. These options are thought to remove low-quality positions of the reads in the 5′ and 3′ ends, respectively. When denoising reads with dada2, it is important to remove low-quality nucleotide positions that would add no valuable information but noise to the analysis (as stated in **Note 11**). Dada2 provides three output files: 1, denoising stats, allows to assess how reads were processed and deleted at each step; 2, the representative sequences file, which contains all the different ASVs; and 3, the feature table, where ASVs' abundances across samples can be found. Next steps are performed to convert these three files for visualization (taking as input each of the generated "qza" files and ouputting the corresponding "qzv" file). The commands to perform step 4 as depicted in Figure 1 are shown below.

```
qiime dada2 denoise-paired \
 --i-demultiplexed-seqs trimmed-demux.qza \
 --p-trunc-len-f 230 \
 --p-trunc-len-r 216 \
 --o-table table-unfiltered.qza \
 --o-representative-sequences rep-seqs-unfiltered.qza \
 --o-denoising-stats denoising-stats.qza \
 --p-n-threads 24

qiime feature-table summarize \
 --i-table table-unfiltered.qza \
 --m-sample-metadata-file metadata.tsv \
 --o-visualization table-unfiltered.qzv

qiime metadata tabulate \
 --m-input-file denoising-stats.qza \
 --o-visualization denoising-stats.qzv
qiime feature-table tabulate-seqs \
 --i-data rep-seqs-unfiltered.qza \
 --o-visualization rep-seqs-unfiltered.qzv
```

Denoising with dada2 algorithm corrects reads using an error model specifically trained for the specific dataset. Therefore, all the stated steps above need to be individually done for each sequencing run (in the example commands below, run A, B, and C) for a correction model suitable for each batch to be used. Afterwards, feature tables and representative sequences obtained from different runs should be joined to be further analyzed together. This merging step can be done using *qiime feature-table merge* as described below. As this step is optional, it has not been included in Figure 1.

```
qiime feature-table merge \
 --i-tables table-unfiltered-runA.qza \
 --i-tables table-unfiltered-runB.qza \
 --i- tables table-unfiltered-runC.qza \
 --o-merged-table table-merged.qza

qiime feature-table summarize \
 --i-table table-merged.qza \
 --m-sample-metadata-file metadata.tsv \
 --o-visualization table-merged.qzv

qiime feature-table merge-seqs \
 --i-data rep-seqs-unfiltered-runA.qza \
 --i-data rep-seqs-unfiltered-runB.qza \
 --i-data rep-seqs-unfiltered-runC.qza \
 --o-merged-data rep-seqs-merged.qza

qiime feature-table tabulate-seqs \
 --i-data rep-seqs-merged.qza \
 --o-visualization rep-seqs-merged.qzv
```

In order to ensure that all amplicons (ASVs) are not artifacts generated in sequencing, it is important to filter them out using *qiime2 quality control* plugin (117). In this step (Figure 1, step 5), all ASVs that do not match at 65% identity

and 50% query cover against the latest Greengenes bacterial 16S rRNA gene database (118) (clustered at 88%, available at https://docs.qiime2.org/2020.8/data-resources/) using vsearch (119) are discarded (see **Note 12**). Those ASVs kept after this filtering step are stored as "hits" (hits.qza), while those not matching the selected parameters are stored as "misses" (misses.qza). Below are found the commands to perform step 5 as depicted in Figure 1.

```
qiime quality-control exclude-seqs \
 --i-query-sequences rep-seqs-merged.qza \
 --i-reference-sequences 88_otus.qza \
 --p-method vsearch \
 --p-perc-identity 0.65 \
 --p-perc-query-aligned 0.5 \
 --o-sequence-hits hits.qza \
 --o-sequence-misses misses.qza

qiime feature-table tabulate-seqs \
 --i-data hits.qza \
 --o-visualization hits.qzv

qiime feature-table tabulate-seqs \
 --i-data misses.qza \
 --o-visualization misses.qzv
```

Finally, this "hits.qza" file is renamed for clarification as rep-seqs-filtered.qza, and it is then used to filter the feature table (table-merged.qza or table.qza) to retain only the ASVs included in it (using *qiime feature-table filter-features*).

```
mv hits.qza rep-seqs-filtered.qza

qiime feature-table filter-features \
 --i-table table-merged.qza \
 --m-metadata-file rep-seqs-filtered.qza \
```

```
 --o-filtered-table table-filtered.qza

qiime feature-table summarize \
 --i-table table-filtered.qza \
 --m-sample-metadata-file metadata.tsv \
 --o-visualization table-filtered.qzv
```

The obtained table contains the ASVs and their relative abundance in each sample. To unravel the taxonomical composition in the samples, these ASVs must be compared to a database to infer their classification (Figure 1, step 6). Although QIIME2 software offers several options to taxonomically classify the ASVs, it is recommended to use a pretrained Naïve-Bayes classifier (120) method (available at https://docs.qiime2.org/2023.2/data-resources/) using *feature-classifier* plugin (see **Note 13**). The output taxonomy.qza file contains the taxonomical assignation to each ASV with a percentage of confidence and can be inspected when converted to qzv file. The required commands to perform step 6 (Figure 1) are described below.

```
qiime feature-classifier classify-sklearn \
 --i-classifier classifier.qza \
 --i-reads rep-seqs-filtered.qza \
 --o-classification taxonomy.qza

qiime metadata tabulate \
 --m-input-file taxonomy.qza \
 --o-visualization taxonomy.qzv
```

Once the sequences have been classified, the obtained taxonomy.qza file should be used to perform a last filtering step to the feature table to remove nonbacterial 16S rRNA gene ASVs (Figure 1, step 7). Since this gene can also be found in *Archaea*, *chloroplasts*, and *mitochondria*, these ASVs must

be removed from the feature table and, accordingly, from the sequences. The files obtained until here will serve as inputs for all the downstream analyses.

```
qiime taxa filter-table \
 --i-table table-filtered.qza \
 --i-taxonomy taxonomy.qza \
 --p-exclude Archaea,Chloroplast,mitochondria \
 --o-filtered-table final-table.qza

qiime feature-table summarize \
 --i-table final-table.qza \
 --m-sample-metadata-file metadata.tsv \
 --o-visualization final-table.qzv

qiime feature-table filter-seqs \
 --i-data rep-seqs-filtered.qza \
 --i-table final-table.qza \
 --p-no-exclude-ids \
 --o-filtered-data final-rep-seqs.qza

qiime feature-table tabulate-seqs \
 --i-data final-rep-seqs.qza \
 --o-visualization final-rep-seqs.qzv
```

One of the main goals in microbiome analysis consists in characterizing the diversity within the samples (alpha diversity) and between (beta diversity) the groups under study. There are several metrics that can be used to perform these estimations. In this chapter, the estimation of those metrics more commonly used is described (see **Note 14**).

Several phylogenetic diversity metrics require a rooted phylogenetic tree to be computed to classify the ASVs (Figure 1, step 8). First, a multiple sequence alignment is done with MAFFT (*qiime alignment mafft*, (121)). The

hypervariable regions are generally considered to add noise to a resulting phylogenetic tree so they should be masked (122)- (*qiime alignment mask*). Finally, the tree is built from the masked alignment using Fasttree (123) and rooted (*qiime phylogeny midpoint-root*). The required commands to perform step 8 in Figure 1 are the following:

```
qiime alignment mafft \
 --i-sequences final-rep-seqs.qza \
 --o-alignment aligned-rep-seqs.qza \
 --p-n-threads 8

qiime alignment mask \
 --i-alignment aligned-rep-seqs.qza \
 --o-masked-alignment masked-aligned-rep-seqs.qza

qiime phylogeny fasttree \
 --i-alignment masked-aligned-rep-seqs.qza \
 --o-tree unrooted-tree.qza \
 --p-n-threads 8

qiime phylogeny midpoint-root \
 --i-tree unrooted-tree.qza \
 --o-rooted-tree rooted-tree.qza
```

All the downstream analyses will be performed using the final filtered representative sequences (rep-seqs.qza), the feature table (table.qza), and the phylogenetic tree (unrooted-tree.qza and rooted-tree.qza) generated throughout these previous steps.

Next steps will permit the computation of both alpha and beta diversity metrics which are essential to understand the variability of the samples (alpha diversity) and the diversity between (beta diversity) the different groups under study (Figure 1, step 9). To make these analyses comparable among samples

with different number of reads, it is mandatory to set a sampling depth common to all the samples to be analyzed. This value, selected by inspecting the table.qzv, is normally set to the lowest sampling depth at which all the diversity was captured in the dataset (10,000 in this example), with the objective of preserving the maximum number of samples and retaining enough number of sequences. The total counts from each sample will be randomly subsampled using the value provided when set in the *- -p-sampling-depth* parameter using the *qiime diversity core-metrics* command (which is a workflow itself calculating different metrics for alpha and beta diversity). The output for this step will be saved in the newly created core-metrics directory (defined in the *- - output-dir* option). The "qza" files inside this folder will allow further diversity analysis or plot generation, and "qzv" files will explain the main diversity characteristics of the samples. The diversity of each sample can be investigated with the alpha-rarefaction-full.qzv file (rarefaction curves of the computed indexes per sample groups through different sampling depths), while beta diversity can be visualized using the emperor PcoA plots (124,125). When opening the emperor files for each beta diversity metric in *qiime view*, several options become available interactively that help the user to understand the differences between groups (such modifying color or changing the shape by different categories).

```
qiime diversity core-metrics-phylogenetic \
 --i-phylogeny rooted-tree.qza \
 --i-table final-table.qza \
 --p-sampling-depth 10000 \
 --m-metadata-file metadata.tsv \
 --output-dir core-metrics \
 --p-n-jobs-or-threads 4
qiime diversity alpha-rarefaction \
 --i-table final-table.qza \
 --i-phylogeny rooted-tree.qza \
 --p-max-depth 10000 \
```

```
 --m-metadata-file metadata.tsv \
 --p-metrics chao1 \
 --p-metrics shannon \
 --p-metrics observed_features \
 --p-metrics simpson \
 --p-metrics faith_pd \
 --o-visualization core-metrics/alpha-
rarefaction_full.qzv

qiime diversity alpha \
 --i-table final-table.qza \
 --p-metric chao1 \
 --o-alpha-diversity core-metrics/chao1_vector.qza

qiime diversity alpha \
 --i-table final-table.qza \
 --p-metric simpson \
 --o-alpha-diversity core-metrics/simpson_vector.qza
```

The diversity analyses within QIIME2 provide the computation of alpha and beta diversity metrics and output interactive visualization, but also they can provide the statistical tests to find out the variable (included as a column in the metadata) that is most strongly associated (and statistically significant) to the differences observed in the microbial diversity. These questions can be answered executing the *qiime diversity alpha* or *beta-group significance* scripts (126,127). The results from alpha significance can be visualized as boxplots of the computed alpha diversity indexes per category at the maximum sampling depth. The statistical significance of alpha diversity differences is calculated with nonparametric t tests (with 999 random permutations). On the other hand, the significance of the beta diversity comparisons (suspected through the visualization of emperor plots) can be calculated using PERMANOVA, and the estimation of *P-values* is done

through 999 random permutations of the dataset. Moreover, the percentage of variation between grouped samples (size effect) can be measured by R^2 using the Adonis function (128), implemented in *qiime diversity adonis* (see **Note 15**). All these steps also belong to step 9 in Figure 1.

```
qiime diversity alpha-group-significance \
 --i-alpha-diversity core-
metrics/observed_features_vector.qza \
 --m-metadata-file metadata.tsv \
 --o-visualization core-metrics/observed-features-
group-significance.qzv

qiime diversity alpha-group-significance \
 --i-alpha-diversity core-metrics/shannon_vector.qza \
 --m-metadata-file metadata.tsv \
 --o-visualization core-metrics/shannon-group-
significance.qzv

qiime diversity alpha-group-significance \
 --i-alpha-diversity core-metrics/chao1_vector.qza \
 --m-metadata-file metadata.tsv \
 --o-visualization core-metrics/chao1-group-
significance.qzv

qiime diversity alpha-group-significance \
 --i-alpha-diversity core-metrics/simpson_vector.qza \
 --m-metadata-file metadata.tsv \
 --o-visualization core-metrics/simpson-group-
significance.qzv

qiime diversity beta-group-significance \
 --i-distance-matrix core-
metrics/jaccard_distance_matrix.qza \
 --m-metadata-file metadata.tsv \
```

```
 --m-metadata-column variable \
 --o-visualization core-metrics/jaccard_variable-
significance.qzv \
 --p-pairwise

qiime diversity beta-group-significance \
 --i-distance-matrix core-
metrics/bray_curtis_distance_matrix.qza \
 --m-metadata-file metadata.tsv \
 --m-metadata-column variable \
 --o-visualization core-metrics/bray_curtis_variable-
significance.qzv \
 --p-pairwise

qiime diversity adonis \
 --i-distance-matrix core-
metrics/jaccard_distance_matrix.qza \
 --m-metadata-file metadata.tsv \
 --p-formula variable_or_formula_of_variables \
 --o-visualization core-metrics/jaccard_ADONIS-
variable.qzv

qiime diversity adonis \
 --i-distance-matrix core-
metrics/bray_curtis_distance_matrix.qza \
 --m-metadata-file metadata.tsv \
 --p-formula variable_or_formula_of_variables \
 --o-visualization core-metrics/bray_curtis_ADONIS-
variable.qzv
```

The visualization of the taxonomical composition of the samples at different taxonomical levels, while grouping samples into the study groups, can be done using bar plots (Figure 1, step 10). The whole

relative frequency table of the taxa found in the samples can be extracted by grouping the ASVs from the feature table at a desired taxonomic level in the command *qiime taxa collapse* (from domain to species, 1–7). In this case, to find out the abundance of the *Streptococcus* genus in the samples, the relative frequency of all genera detected can be extracted (level 6), and then *Streptococcus* string can be used to filter by this genus (see **Note 16**). Sometimes, it may be interesting to have the information of the relative frequency table at ASV level, then the first step intended to collapse at higher taxonomical levels needs to be omitted.

```
qiime taxa barplot \
 --i-table final-table.qza \
 --i-taxonomy taxonomy.qza \
 --m-metadata-file metadata.tsv \
 --o-visualization taxa-bar-plots.qzv

qiime taxa collapse \
 --i-table final-table.qza \
 --i-taxonomy taxonomy.qza \
 --p-level 6 \
 --o-collapsed-table table-l6.qza

qiime feature-table relative-frequency \
 --i-table table-l6.qza \
 --o-relative-frequency-table table-rel-freq-l6.qza

qiime tools export \
 --input-path table-rel-freq-l6.qza \
 --output-path l6-rel-freq-table
biom convert \
-i l6-rel-freq-table/feature-table.biom \
-o l6-rel-freq-table.tsv –to-tsv
```

This pipeline is intended to give the user an overview on the diversity, the taxonomical composition, and the significant differences between and within the groups under study. However, the feature table and the representative sequences can be used as inputs for many downstream analyses within QIIME2 by using several plugins (Figure 1, step 11). Moreover, other bioinformatics tools (outside QIIME2) will accept QIIME2-exported files as input and may be useful to expand or focus the analyses.

Notes

1. The swab needs to go deep enough into the nose to obtain a representative sample.

2. It is recommended that sampling is performed by the same person to avoid differences in sample collection.

3. Note that RNA later is not a good option for storage of swabs.

4. Note that also RNA may be used as starting material for sequencing after a step of retro-transcription.

5. The sample metadata file is a tab-separated text which contains sample ids in the first column, while in the rest of columns include all the information for each sample. The first two rows include the names of the categories and a label defining whether these categories are numeric or categorical. An example sample metadata file can be downloaded from a QIIME2 tutorial (https://docs.qiime2.org/2023.2/tutorials/moving-pictures-usage/). It is important to check that everything, including the words “numeric” and “categorical,” are well spelled (they are case sensitive, e.g., within the category “clinical status,” “health,” and “Health” would be treated as samples from different groups).

6. Although the generated file as metadata can be directly loaded in Qiime2, if errors are present, they will be detected one at a time making the identification and resolving process long. Instead, it is recommended to perform the validation

of the metadata file before using Keemei (129) that will make a report with all the errors and warnings at a time.

7. QIIME2 software works with artifacts, which are compressed files (zip) that contain both data and the information about the data (metadata). There are two types of artifacts: the executable (.qza) and the visualizations (.qzv), viewable in QIIME2 view interface (available at https://view.qiime2.org/).

8. A typical manifest file is a plain file which contains all the information described tab separated, with each sample in a different row separated by a newline.

9. The *qiime demux summarize* command allows to visualize the demultiplexed sequence artifact. In the case of fastq reads, this option outputs a plot describing the quality for each nucleotide position of the original reads using a subsample of a defined number of reads (this parameter can be set with the *- -p-n flag*).

10. In this example command, *cutadapt* plugin is used to remove the primers from paired-end reads. The primers targeted to amplify the V3–V4 region can be set in the corresponding parameters *(--p-front-f, --p-front-r*). This plugin also offers options to trim the reads (i.e., trim the first n bases) or to demultiplex barcodes in case the sequencing facility outputs undemultiplexed reads.

11. Example on deciding dada2 truncation and trimming parameters. In the example (Figure 2), it is clear how quality drops along the read (Figure 2a), especially in reverse reads. Thus, if no truncation is applied (*--p-trunc-len,* in 3′ end), all reads with low quality will be discarded. Zooming in qiime2 view (Figure 2b), a truncation position where quality is acceptable can be selected, but this should be always lower than the reads' length (Figure 2c). If a truncation position higher than the reads' length is selected (Figure 2c), all reads smaller than this selection will be discarded. If no truncation position is specified, all reads will be used, and most of them will be discarded when low-quality positions are detected. But on the contrary, if truncation is too extreme, the merging of the paired-end reads performed in the next step will not be successful. Therefore, a compromise between eliminating bad-quality positions allowing the merging of the reads must be achieved. Similarly, the low-quality positions in the 5′ end can be trimmed (--

p-trim-left); however, quality is usually high at these positions. It may be required to perform this step several times to get the best parameters for the analysis.

12. Quality filtering parameters used herein may seem quite permissive, but the purpose of this step is only to remove sequences not coming from 16S rRNA gene amplification. This filtering can be applied directly on dada2 output sequences (in the example shown there is only one run) or on the merged sequences (if different runs were merged). After the filtering, it is important to check that the majority of the ASVs were kept in the hits.qza file, with the expected amplicon lengths in the sequence length summary. On the contrary, few sequences should be discarded as "misses" (and with shorter lengths than expected).

13. There are other options to classify our features, as well as training the classifiers (see https://docs.qiime2.org/2023.2/plugins/available/feature-classifier/ for further info).

14. In this diversity analysis example, four alpha diversity measures assessing the richness and evenness of our samples are calculated (observed features, Chao (130), Shannon (131), and Simpson (132) indexes). The beta diversity of the samples is estimated with the non-phylogenetic dissimilarity indexes Bray and Curtis (133) and Jaccard (134) (quantitative and qualitative, respectively) and weighted (135) and unweighted (136) UniFrac phylogenetic distances (quantitative and qualitative, respectively).

15. The beta diversity significance commands (PERMANOVA and Adonis) are shown for the two non-phylogenetic indexes, but the same commands can be applied for the two UniFrac phylogenetic analyses changing the qza files as input.

16. In order to find out the relative abundance across samples of the taxa, the relative frequency table can be manually inspected. However, simple commands can help to automatically extract the abundances of desired taxa. As an example, it is provided a way to perform this step using the *grep* command (directly applied in the terminal, bash language). Three options are provided with the *-w -F -f* flags. First, the string we want to search which is the taxon name (an asterisk in the

beginning, replaces any character found before the taxon name); second, the file where the search has to be done, i.e., the relative frequency table; and third, the output file name has to be defined (in the example, streptococcus-abundances.tsv). The command applied for *Streptococcus* would look like this:

*grep -w -F -f *Streptococcus l6-rel-freq-table.tsv > streptococcus-abundances.tsv.*

If the relative frequency of a bunch of taxa is desired, a plain file with the taxa names separated by comma (csv file, my_taxa.csv in the example below) should be prepared in any spreadsheet program and then use it in the -f option when applying this command:

grep -w -F -f my_taxa.csv l6-rel-freq-table.tsv > taxa-abundances.tsv

CHAPTER 2

Sow Contact Is a Major Driver in the Development of the Nasal Microbiota of Piglets

Pau Obregon-Gutierrez, Virginia Aragon, Florencia Correa-Fiz. *Pathogens*. 2021;10(6):697. doi:10.3390/pathogens10060697

Abstract

The nasal microbiota composition is associated with the health status of piglets. Sow-contact in early life is one of the factors influencing the microbial composition in piglets; however, its impact has never been assessed in the nasal microbiota of piglets reared in controlled environmental conditions. Nasal microbiota of weaning piglets in high-biosecurity facilities (BSL3) with different time of contact with their sows (no contact after farrowing, contact limited to few hours or normal contact until weaning at three weeks) was unveiled by 16S rRNA gene sequencing. Contact with sows demonstrated to be a major factor affecting the nasal microbial composition of the piglets. The nasal microbiota of piglets that had contact with sows until weaning, but were reared in high biosecurity facilities, was richer and more similar to the previously described healthy nasal microbiota from conventional farm piglets. On the other hand, the nasal communities inhabiting piglets with no or limited contact with sows was different and dominated by bacteria not commonly abundant in this body site. Furthermore, the length of sow–piglet contact was also an important variable. In addition, the piglets raised in BSL3 conditions showed an in-creased richness of low-abundant species in the nasal microbiota. Artificially rearing in high biosecurity facilities without the contact of sows as a source of nasal colonizers had dramatic im-pacts on the nasal microbiota of weaning piglets and may introduce significant bias into animal research under these conditions.

Introduction

The dynamic ensemble of microbiological communities inhabiting animal body sites is called microbiota (4) and comprises many different types of microorganisms that provide the host with an important set of functions (137–139). Studies unraveling the many positive aspects of the microbiota on different animal hosts (such as metabolic benefits, immune system maturation, and protection against pathogens) (4,9,10,140), together with the fact that the microbiota comprises potential pathogens (4), have put this area in the focus of microbiological research during recent years. Many studies have shown that a stable, specific and diverse community of commensal microbes is associated with health in humans (3), but also in different animals (9,10,137). On the other hand, alterations on beneficial bacteria can result in overgrowth of opportunistic pathogenic bacteria and predispose to disease (4). However, most of these studies have been carried out on humans, leaving animal microbiota less assessed (3). Animal microbiota composition varies among individuals due to the action of many factors, such as the environment, host-genetics, the diet, the use of antimicrobials and the maternal transmission in early life, among others (4,9,10,141).

Among animal farming, pig industry is one of the most important worldwide, since pork is one of the world's most consumed meats (142). Antimicrobials have been widely used for many years in swine production with a great success controlling bacterial diseases, but antimicrobial resistance has generated a strong institutional call for the reduction of antimicrobial use in farm animals (91,143). In addition, antimicrobials have the potential to disrupt the beneficial microbial communities of the microbiota (10,14,34). Therefore, alternative tools to control diseases while ensuring balanced costs and low economic loses are of great interest (26), including those directed to the development and maintenance of healthy microbial communities (10,137).

The largest animal microbial ecosystem is the gastrointestinal tract, which has been the focus of most microbiota studies, because of its impact on the host (10,14,139,144,145). Nevertheless, other microbial niches (nasal, oral, skin, vagina, etc.) may also play important roles (16,145), and need more consideration. The upper respiratory tract, in particular the nasal cavity, is the entry way of many respiratory pathogens. Therefore, understanding the nasal microbiota and the factors that modulate its composition can be essential for improving the control of respiratory infections (26). Furthermore, early life conditions can determine the development and subsequent composition of the microbiota and affect the health status later in life (16).

Recently, some studies have correlated the pig microbiota composition with health or disease (10,12,14,16). This is especially important at weaning, which is a critical moment due to the changes in feeding, antimicrobial treatments, and physical and social environment (9,10,12,14,26). The microbial composition at this moment is still unstable and can play a key role in the subsequent health status of the animals (10,12,16,17,140,146,147). Later, the microbiota constitutes a stable and resilient community that is able to resist some perturbations (26). Hence, during the first period of life, when the microbiota can undergo many fluctuations, it may be important to favor its stabilization with beneficial communities in order to wean healthier and more resilient piglets.

Several studies have established that correlations between the microbiota of sows and their piglets depend on different aspects such as the delivery mode, feeding and direct contact between the litter and sow (9–11,14,27,148). Initial colonization of piglets is mainly derived from vaginal microbiota in natural delivery, and later from skin, food or the environment (10,26,27,140,149). However, the nasal microbiota has not been the focus of these studies. On the other hand, some disease models for endemic pathogens require piglets with no contact with the sows and are therefore artificially reared; the so-called snatch-farrowed colostrum-deprived piglets or the caesarean-derived

colostrum-deprived models (150–152). Isolation facilities and high biosafety level facilities, which ensure a particularly clean environment, are sometimes used to house these highly susceptible pig models. It is clear that depriving the newborn piglets from colostrum, their initial source of immunity, has a big impact on the piglet health (10). Nevertheless, the impact of this artificial environment and rearing conditions on the microbiota of the piglets has never been evaluated, especially focusing on nasal microbiota. Therefore, the aim of this study was to assess how a reduced contact with sows and the high biosecurity conditions where piglets are raised affect the nasal microbiota composition of piglets at weaning.

Results

Sample Collection and Sequencing

Nasal samples were taken at three to four weeks of age (common age of weaning) from piglets raised either in Biosafety Level 3 (BSL3) facilities or in conventional farms, as depicted in **Table 1**. To examine the effect of the sow in the nasal microbiota of the offspring, piglets were allowed to have different degree of contact with their sows. Thus, piglets grown in BSL3 conditions (L3) belonged to three different groups: one group of snatch-farrowed colostrum-deprived piglets, taken at delivery in farms with no contact with the dams except for the birth canal (no-contact, L3-NC); a second group with less than 12 h of contact with the sows after birth in the farm, before they were transferred to the BSL3 facilities for housing (limited-contact, L3-LC); and a third group of piglets that were born and raised in the BSL3 facilities and stayed with their sows until sampling (full-contact, L3-FC). In addition, piglets in group L3-LC belonged to two subgroups depending on the treatment at birth with crystalline ceftiofur (treated or control). Total DNA was extracted from nasal swabs and the V3-V4 region of the 16S gene was amplified and sequenced. The sequences from the nasal swabs of a group of conventional

piglets at weaning, born and raised at the farm (farm born-farm raised, FB-FR) from three farms with confirmed good health status (no clinical signs in maternity and nursery units) from Catalonia (GM, PT and VL) from a previous study (16), were also included in the analysis. A total of 31,245,358 paired-end reads were obtained from five different runs. Paired-end reads were joined, quality filtered and sorted into amplicon sequence variants (ASVs). An amount of 29,794 ASVs were obtained, with a total frequency of 3,569,216 sequences (mean frequency of 71,384.32 per sample). The rarefaction curves were built (**Figure 1A** and **Figure S1A**) to determine the optimal depth where the plateau was reached and, hence, the microbial communities were well represented. In order to normalize data (rarefaction), one sample from L3-LC group was removed due to its low coverage (5,320 sequences in comparison with 73,555 for the rest of the samples) to mitigate possible biases related to uneven sampling depths (153).

Table 1. Sample group distribution with main characteristics and dataset study reference.

Group	Housing	Birthplace	Sow-Contact	Perinatal Treatment	Reference
L3-NC	BSL3	Farm (snatch farrowed)	No	Colistin and crystalline ceftiofur (N = 6)	This study
L3-LC	BSL3	Farm	Less than 12 h	Crystalline ceftiofur (N = 7) Untreated (N = 6)	This study
L3-FC	BSL3	BSL3	Until weaning	Untreated (N = 11)	This study

FB-FR	Farm GM	Farm GM	Until weaning	Amoxicillin (N = 10)	[18]
	Farm PT	Farm PT	Until weaning	Not available (N = 5)	
	Farm VL	Farm VL	Until weaning	Ceftiofur and tulathromycin (N = 5)	

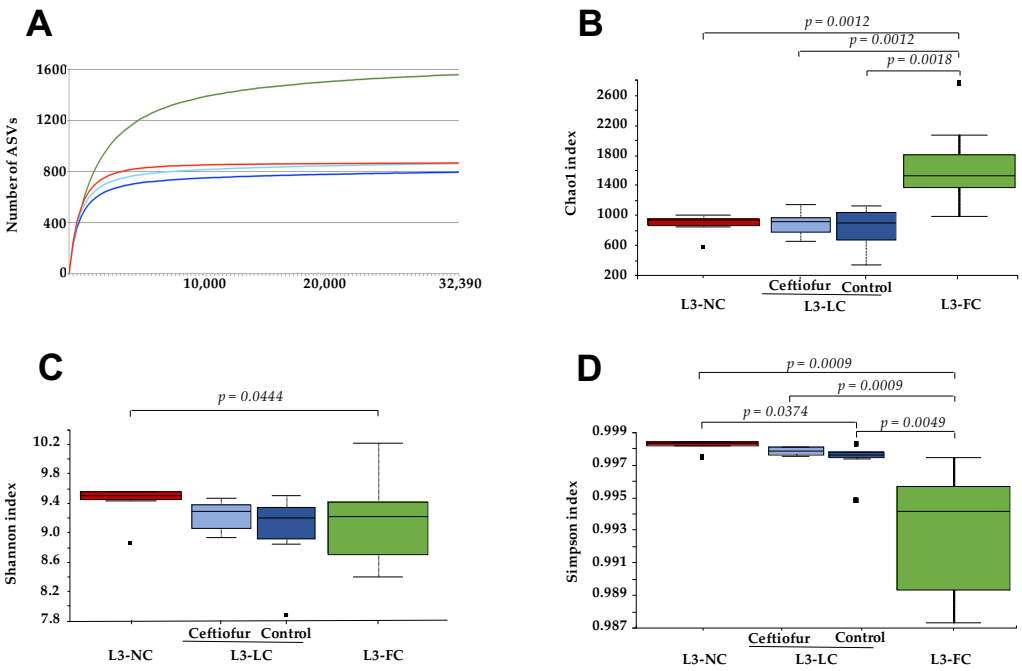


Figure 1. Alpha diversity of the nasal microbiota of 3–4-week-old piglets raised under BSL3 conditions with different degree of contact with their sows. Alpha diversity was measured at a depth of 32,392 by observed ASVs (A), Chao1 (B), Shannon–Weaver (C) and Simpson (D) indices for the different groups under study: piglets with no contact with the sow except for the birth canal (L3-NC, in red); piglets with limited contact of less than 12 h (L3-LC, in blue) and those that had full normal contact with their sows (L3-FC, in green). Group L3-LC included animals treated with ceftiofur ('Ceftiofur', in light blue) or non-treated ('Control', in dark blue). Significant p values from Kruskal–Wallis pairwise tests are depicted.

Alpha Diversity of the Nasal Microbiota of Piglets with Different Degree of Sow Contact

To assess the effect that sows have on the nasal microbiota of their piglets, samples from piglets raised under the same BSL3 conditions but with different sow–contact time were compared. Samples were rarefied to an even depth of 32,392 sequences/sample and alpha diversity was evaluated through different metrics. The mean number of ASVs at this depth was 860 for L3-NC, 803 for L3-LC and 1470 for the L3-FC group (**Figure 1A**). The highest richness of L3-FC group was confirmed when estimated through Chao1 index (Kruskal–Wallis test, 999 permutations, $p < 0.05$, **Figure 1B**). However, when evenness was taken into consideration, the group showing significantly higher diversity was L3-NC, either by Shannon (**Figure 1C**) or Simpson (**Figure 1D**) indices. The alpha diversity variability within each group increased with the length of sow–piglet contact, especially when including evenness (Figure 1C,D).

The comparison of ceftiofur-treated and non-treated piglets within the L3-LC group did not show significant differences in alpha diversity ($p = 0.87$, observed features; $p = 0.75$, Chao1; $p = 0.63$, Shannon; $p = 0.42$, Simpson).

Nasal Microbiota Composition of Piglets Differed Depending on the Degree of Contact with Sows

To unravel compositional differences between groups, beta diversity was estimated with unweighted and weighted UniFrac phylogenetic distances. Principal coordinate analysis (PCoA) was performed to assess the degree of clustering among groups and the percentage of variation explained by each clustering was estimated (Adonis function from the Vegan package). Sample clustering of the three groups under study was statistically significant when evaluated with unweighted and weighted UniFrac distance matrices (PERMANOVA, $p = 0.001$ in all cases). Sow–piglet contact explained 43%

and 82% of the clustering in the unweighted and weighted analysis, respectively (Adonis, $p = 0.001$). A PCoA biplot was built to represent the most relevant genera that describe the separation in the same PCoA space together with the samples (**Figure 2A**). In the biplot analysis, *Enhydrobacter*, *Moraxella* and *Rothia* vectors were directed to the PCoA space occupied by samples from L3-FC group, indicating that these genera might be associated to this group. The vectors of two genera from *Firmicutes* (from *Clostridiales* order and *Ruminococcaceae* family) were pointing to the L3-NC space and a genus from *Bacteroidetes* (from family *S24-7*, also known as *Muribaculaceae*) to the L3-LC space. Interestingly, L3-NC and L3-LC groups were qualitative and quantitatively more similar between them than to L3-FC, as indicated by the lower distances between them (**Figure 2B,C**).

Ceftiofur treatment within L3-LC did not influence the nasal microbiota composition under these conditions (Figure 2A). This was confirmed by comparing only these two subgroups, ceftiofur treated and non-treated, from L3-LC under beta diversity metrics (PERMANOVA, $p = 0.445$ and $p = 0.146$ for unweighted and weighted UniFrac, respectively).

The effect of contact between sow and piglets was also evaluated by comparing the clustering between the group of 'positive contact' (L3-LC and L3-FC together) and 'negative contact' (L3-NC). Moreover, we calculated the effect of 'normal sow–contact' (L3-FC) versus 'altered contact' (L3-NC and L3-LC together). The percentage of explanation of these clusters were lower in positive or negative contact (12% and 18% for unweighted and weighted Unifrac, $p < 0.01$) than for normal vs. altered contact (31% and 68%, $p = 0.001$), indicating that the time of contact with the sows is a stronger driver of microbiota composition than the mere existence of this contact.

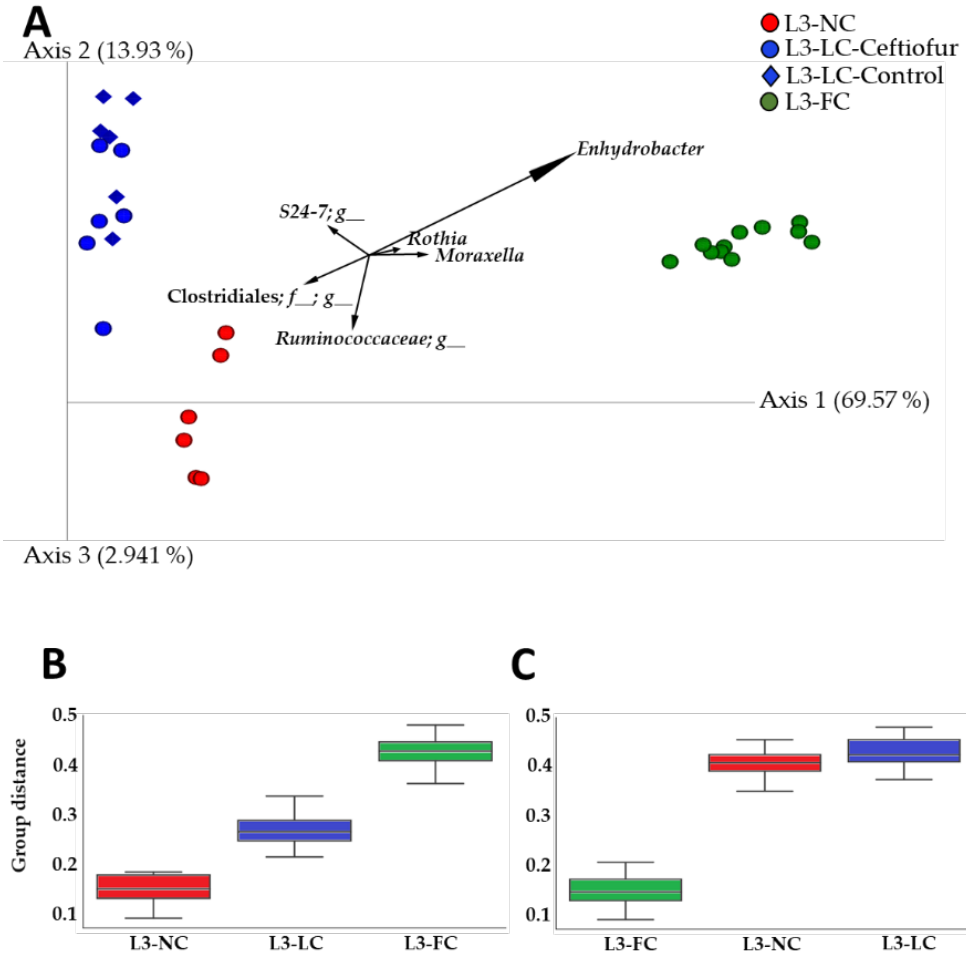


Figure 2. Beta diversity of the nasal microbiota of 3–4-week-old piglets raised under BSL3 conditions with different degree of contact with their sows. A) Principal coordinate analysis representing beta diversity of the nasal microbiota of piglets raised in BSL3 conditions with no contact with sows except the birth canal (L3-NC, in red), limited contact of less than 12 h (L3-LC, in blue) and full normal contact (L3-FC, in green) was computed through weighted UniFrac analysis. The six most relevant genera explaining the differences among groups are plotted in the PCoA space. The length of each of the taxonomic vectors approximates the variance of each taxon throughout the samples. Group clustering distance to L3-NC (B) and to L3-FC (C) computed by PERMANOVA pairwise test with weighted UniFrac distance matrix (p = 0.001 in all cases).

Comparison of the Nasal Microbiota Core from Animals with Variable Sow Contact

To unravel the composition of the different groups with variable sow contact, the ASVs were taxonomically classified. Three phyla comprised 88% of the global relative abundance: *Firmicutes* (49%), *Bacteroidetes* (23%) and *Proteobacteria* (16%). Within the *Firmicutes*, and *Clostridiales* orders it was the most relatively abundant (41% of the global relative abundance), comprising of mainly two families: *Ruminococcaceae* and *Lachnospiraceae*. Among the classified genera, the most relatively abundant were *Oscillospira*, *Ruminococcus* and *Coprococcus*. *Firmicutes* were similarly represented by *Bacillales* and *Lactobacillales* orders, such as *Streptococcus*, *Lactobacillus* and *Staphylococcus*. The most relatively abundant families among *Bacteroidetes* phylum belonged to *Bacteroidales* order (22% of the global relative abundance), where *Prevotellaceae* (mainly represented by *Prevotella* genus) and S24-7 were the most relatively abundant. Most *Proteobacteria* were classified as *Gammaproteobacteria*, where *Moraxellaceae* was the most relatively abundant family within this phylum and the second globally, mainly composed of *Enhydrobacter* and *Moraxella* genera. Regarding other phyla, some relatively abundant families were *Micrococcaceae* (*Actinobacteria*), mainly represented by the genus *Rothia*; and *Spirochaetaceae* (*Spirochaetes*), mainly represented by *Treponema*. Taxonomical composition at genus level of the most relatively abundant taxa globally (>1% relative abundance in one group at least) is depicted in **Figure 3**. The full list of the taxonomical assignment of the most relatively abundant taxa from the different groups is included in **Table S1**, and the corresponding plots at phyla and family levels can be found in **Figure S1**.

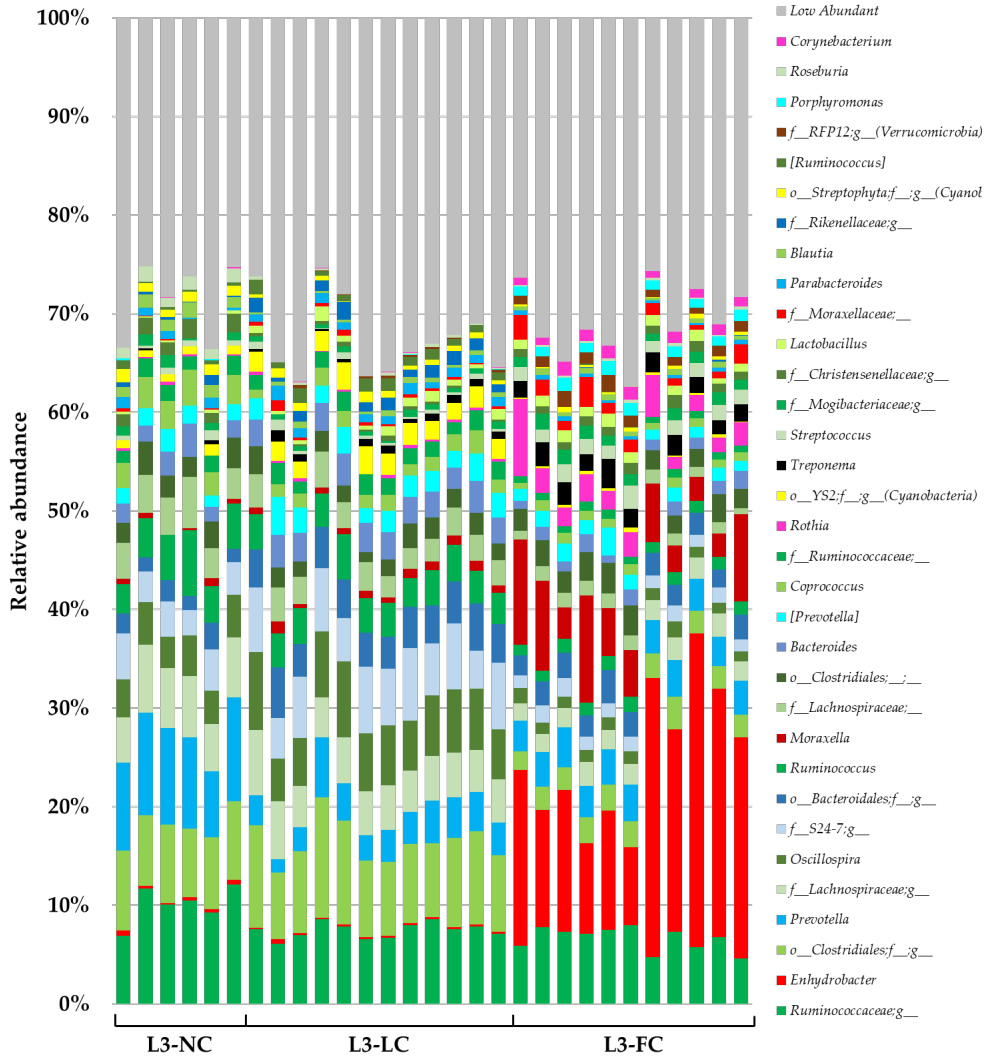


Figure 3. Taxonomical composition at genus level of the nasal microbiota from piglets raised in BSL3 conditions with variable contact with sows. Only taxa with global relative abundance higher than 1% in at least one group is color-coded. All genera with less relative abundance than 1% have been grouped and are shown in grey. L3-NC, no contact with the sows except the birth canal; L3-LC, limited contact of less than 12 h; and L3-FC, full normal contact until weaning at 3 weeks of age. Genera belonging to most relatively abundant phyla have been colored in common color-scheme to simplify its visualization; green for *Firmicutes*, blue for *Bacteroidetes*, and red for *Proteobacteria*. The legend shows the genera ordered by global relative abundance from bottom to top.

The core microbiota was defined as the taxa represented in more than 80% of samples in each group at two different levels: family and genus. The number of core taxa found was 65, 75 and 102 families and 86, 109 and 185 genera for L3-NC, L3-LC and L3-FC, respectively. The number of shared and exclusive taxa among groups is represented in **Figure 4**, while the full list of taxa is available in **Table S2**. A common core for the three groups was composed of 43 families and 53 genera, which included the majority of most relatively abundant families and genera (26 out of 33 of families, and 26 out of 34 genera), previously defined as those taxa present in over 1% relative abundance in at least one group (included in Table S1). The two groups with sow-contact, L3-LC and L3-FC, had more taxa in common than with L3-NC. Finally, L3-FC held the highest number of exclusive taxa.

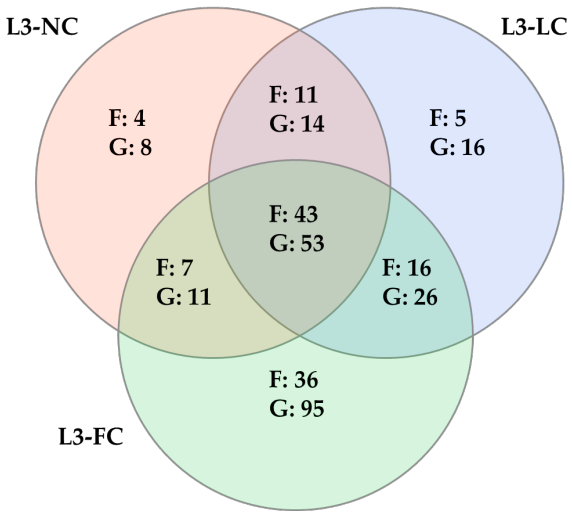


Figure 4. Venn diagram of the shared and specific taxa from the nasal microbiota of piglets raised under BSL3 conditions with different degree of contact with their sows: no contact except the birth canal (L3-NC, in red); limited contact of less than 12 h (L3-LC, in blue) and full normal contact until weaning at 3 weeks of age (L3-FC, in green). Taxa was analyzed at family (F) and genus (G) levels. Only taxa present in more than 80% of samples in a group were included.

In order to find differentially abundant taxa in the different groups with different sow contact, analysis of composition of microbiomes (ANCOM) and a discrete false-discovery-rate (DS-FDR) analysis were performed. The taxa with more than 1% of relative abundance in at least one group and found differentially abundant in either test, are shown in **Table 2**. Among the highest relative abundance taxa, *Firmicutes* was detected in higher abundance in altered sow–piglet contact groups, attributed mainly to *Lachnospiraceae* and *Ruminococcaceae* families (*Clostridiales* order, with some exceptions). *Bacteroidales* order (*Bacteroidetes*) was also increased in these groups, with higher abundance of family *S24-7*, or *Prevotella* in L3-NC and p-2534-18B5 in L3-LC. Contrary, the order *Flavobacteriales* or the genus *Porphyromonas* were more represented in L3-FC. On other hand, an increase in the relative abundance of *Actinobacteria* and *Proteobacteria* (mainly *Moraxella* and *Enhydrobacter*) was observed in L3-FC group, most of these observations validated with both approaches. Similarly, *Bacillales* and *Lactobacillales* orders (*Firmicutes*) were increased when the time of sow–piglet contact was longer (detected only with DS-FDR). These analyses were done at each taxa level, whilst ANCOM was also performed at ASVs level, which confirmed the results obtained at higher taxonomic levels (data not shown).

Table 2. Differently abundant taxa at phylum, order, family and genus levels found by DS-FDR Kruskal–Wallis test and ANCOM (when indicated) with the mean relative abundance for nasal microbiota of piglets raised under BSL3 conditions with different degree of contact with their sows: no contact except the birth canal (L3-NC); limited contact of less than 12 h (L3-LC) and full normal contact until weaning at 3 weeks of age (L3-FC).

Taxonomy	**Relative Abundance (%)**			**Statistics**	
Taxa*	**L3-NC**	**L3-LC**	**L3-FC**	**DS-FDR (*p*)**	**ANCOM* (W)**

Acidobacteria	1.19	0.68	0.02	0.001	
Actinobacteria	1.72	0.71	6.21	0.001	30
Actinomycetales	0.97	0.38	5.02	0.001	122
Corynebacteriaceae	0.03	0.04	1.03	0.001	211
Corynebacterium	0.03	0.04	1.03	0.001	
Micrococcaceae	0.34	0.24	2.92	0.001	
Rothia	0.20	0.24	2.76	0.001	
Coriobacteriales	0.26	0.05	1.11	0.001	131
Coriobacteriaceae	0.26	0.05	1.11	0.001	206
Bacteroidetes	22.53	28.78	16.54	0.001	25
Bacteroidales	21.46	27.84	14.62	0.001	
Bacteroidales;f__	1.85	4.00	2.36	0.001	
Bacteroidales;f__;g__	1.85	4.00	2.36	0.001	
Bacteroidaceae	2.11	2.85	1.29	0.001	
Bacteroides	2.11	2.85	1.29	0.001	
Porphyromonadaceae	0.81	0.95	1.70	0.001	
Porphyromonas	0.06	0.01	1.09	0.001	448
Prevotellaceae	9.28	3.46	3.46	0.002	
Prevotella	9.28	3.46	3.46	0.002	
Rikenellaceae	0.56	1.61	0.13	0.001	216
Rikenellaceae;g__	0.38	1.14	0.03	0.001	496
S24-7	3.56	5.85	1.47	0.001	

S24-7;g__	3.56	5.85	1.47	0.001	
Paraprevotellaceae	1.82	2.87	2.72	0.006	
(*Prevotella*)	1.69	2.34	1.46	0.001	
p-2534-18B5	0.48	5.53	0.10	0.001	248
p-2534-18B5;g__	0.48	5.53	0.10	0.001	505
Flavobacteriales	0.43	0.76	1.76	0.002	
Flavobacteriaceae	0.16	0.19	1.19	0.001	
Cyanobacteria	2.77	3.71	0.59	0.001	29
Firmicutes	58.43	50.69	37.69	0.027	
Bacillales	0.36	0.56	1.53	0.001	
Staphylococcaceae	0.26	0.55	1.12	0.001	
Lactobacillales	1.23	1.68	5.00	0.001	
Aerococcaceae	0.06	0.01	1.31	0.001	224
Lactobacillaceae	0.20	0.76	1.01	0.001	
Lactobacillus	0.20	0.76	1.00	0.001	
Streptococcaceae	0.73	0.49	1.62	0.001	
Streptococcus	0.59	0.42	1.52	0.001	
Clostridiales	54.65	45.83	28.82	0.001	
Clostridiales;f__	7.58	8.77	2.46	0.001	
Clostridiales;f__;g__	7.58	8.77	2.46	0.001	
Christensenellaceae	1.51	0.26	0.92	0.001	
Christensenellaceae;g__	1.51	0.26	0.92	0.001	

Clostridiaceae	1.14	0.67	1.53	0.001	
Lachnospiraceae	17.25	11.67	5.94	0.001	
Lachnospiraceae;__	3.77	2.79	1.23	0.001	
Lachnospiraceae;g__	5.76	4.68	2.07	0.001	
Blautia	1.23	0.41	0.50	0.001	
Coprococcus	2.93	1.48	0.91	0.001	
Roseburia	1.22	0.17	0.23	0.001	
Ruminococcaceae	21.77	19.88	10.86	0.001	
Ruminococcaceae;__	1.61	1.71	0.75	0.001	
Oscillospira	3.87	6.03	1.24	0.001	
Ruminococcus	4.41	3.52	1.27	0.001	
(*Mogibacteriaceae*)	0.89	0.52	1.20	0.001	
(*Mogibacteriaceae*);g__	0.87	0.52	1.19	0.001	
(*Tissierellaceae*)	0.14	0.05	1.15	0.007	
Proteobacteria	8.21	9.46	31.06	0.001	27
Rhizobiales	0.95	1.18	0.16	0.001	123
Neisseriales	0.05	0.37	1.03	0.001	143
Neisseriaceae	0.05	0.37	1.03	0.005	241
Pseudomonadales	1.38	2.19	26.17	0.001	135
Moraxellaceae	1.15	2.01	26.01	0.002	216
Moraxellaceae;__	0.12	0.37	1.41	0.001	
Enhydrobacter	0.34	0.23	18.32	0.001	450

Moraxella	0.44	0.77	5.98	0.001	
Spirochaetes	0.38	0.77	2.56	0.001	29
Spirochaetales	0.09	0.64	2.01	0.001	145
Spirochaetaceae	0.09	0.64	2.01	0.001	249
Treponema	0.09	0.64	2.01	0.001	493
Tenericutes	1.38	1.66	0.66	0.001	27
Verrucomicrobia	0.68	0.59	1.55	0.001	
WCHB1-41	0.00	0.16	1.18	0.001	144
RFP12	0.00	0.15	1.08	0.001	249
RFP12;g__	0.00	0.15	1.08	0.001	522

* The brackets indicate that the taxonomic name is contested.

Sow–Piglet Contact Had a Stronger Impact on the Nasal Microbiota of BSL3 Piglets than the Environment

To assess the relevance of raising piglets under different environment early in life, we compared the nasal microbiota composition of BSL3-raised piglets to same-aged animals raised in farms (FB-FR). Alpha diversity estimated at a maximum even depth of 17,360 showed higher richness in L3-FC group compared to FB-FR groups, measured by Observed features and Chao1 indices ($p < 0.01$; **Figure S2 A,B**). L3-NC and L3-LC groups had similar richness than the whole FB-FR group, only statistically higher than one of the farms (PT, $p < 0.05$), but higher diversity when measured by Shannon's and Simpson's indices (**Figure S2 C,D**). Particularly L3-NC was statistically higher than other groups in both indices ($p < 0.05$). L3-FC diversity was statistically higher than FB-FR when computed by Shannon's index ($p = 0.026$) that includes both richness and evenness.

To understand compositional similarities and dissimilarities, beta diversity was estimated for FB-FR and the three BSL3 groups. All groups were statistically different in both quantitative and qualitative analyses (weighted and unweighted Unifrac, $p < 0.01$). The biplot analysis showed that *Pasteurellaceae* and *Bergeyella* shared the space with the FB-FR group, and again, *Moraxella* and *Enhydrobacter* were associated with normal sow–piglet contact (**Figure 5**). PCoA plots showed clustering of L3-NC and L3-LC groups when farms were included in the qualitative and quantitative analysis (Figure 5A,B), especially in the latter. This was confirmed by measuring the group distance with both metrics, where the distance among L3-NC and L3-LC groups to L3-FC was similar to the distance to FB-FR, even though the animals were all raised in BSL3.

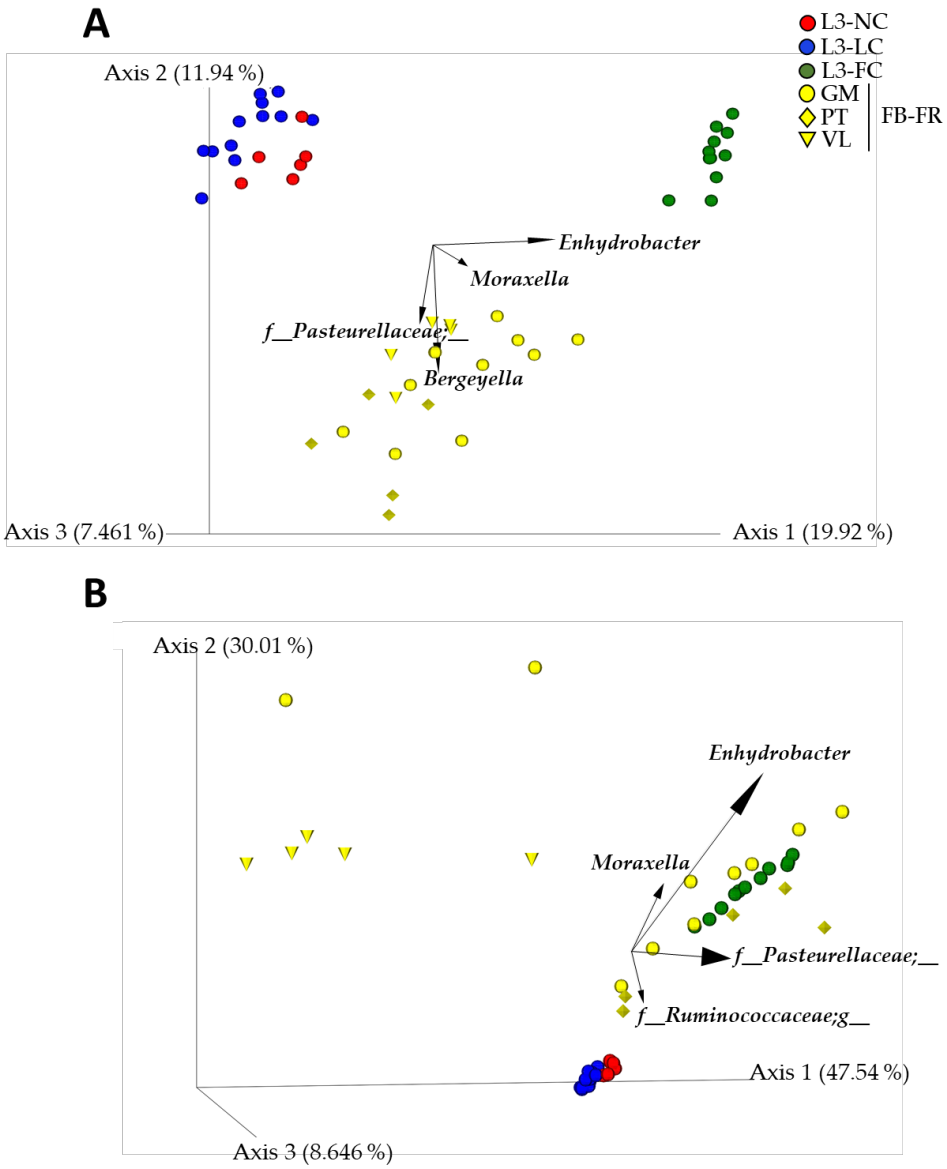


Figure 5. Principal component plot of the nasal microbiota of piglets raised in farms (FB-FR, in yellow) or under BSL3 conditions with different degree of contact with their sows (no contact except the birth canal (L3-NC, in red); limited contact of less than 12 h (L3-LC, in blue) and full normal contact until weaning at 3 weeks of age (L3-FC, in green)) computed through unweighted (A) and weighted (B) UniFrac analysis. The four most relevant genera explaining the differences among groups are plotted in the PCoA space. The length of each of the taxonomic vectors approximates the variance of each taxon throughout the samples. FB-FR samples have different shapes depending on the farm.

To estimate the percentage explained by grouping animals based on the environment, they were raised and compare it to the effect of sow–piglet contact, we performed PERMANOVA tests with Adonis function with all groups. The percentage explained by grouping the animals according to the environment (BLS3 vs. FB-FR groups) was 11.29% and 17.08% in the unweighted or weighted analysis, respectively (p = 0.001). In addition, we evaluated the effect of different sow–piglet contact, after considering the environment (nested Adonis), where L3-FC and FB-FR were computed as a single group. The percentage of explanation by sow–piglet contact was 25.67% and 24.44% in unweighted and weighted analysis, respectively (p = 0.001). A normal sow–piglet contact (grouping L3-NC and L3-LC compared to L3-FC and farms) was responsible for most of the sows' effect variation, while the mere presence of this contact (L3-NC versus the other groups) accounted for a smaller part of this effect.

In order to qualitatively analyze the similarities among BSL3 and FB-FR groups, the core taxa of each group (represented in more than 80% of samples) were compared at family and genus levels. The number of families were 65, 75, 102 and 55; while the number of different genera were 86, 109, 185 and 75 in L3-NC, L3-LC, L3-FC and FB-FR groups, respectively. Remarkably, the longer the sow contact in the BSL3 groups, the more taxa were shared with the FB-FR group. The number of shared and exclusive taxa from each group is shown in **Figure 6**, while the list of common and exclusive taxa is presented in **Table S3**.

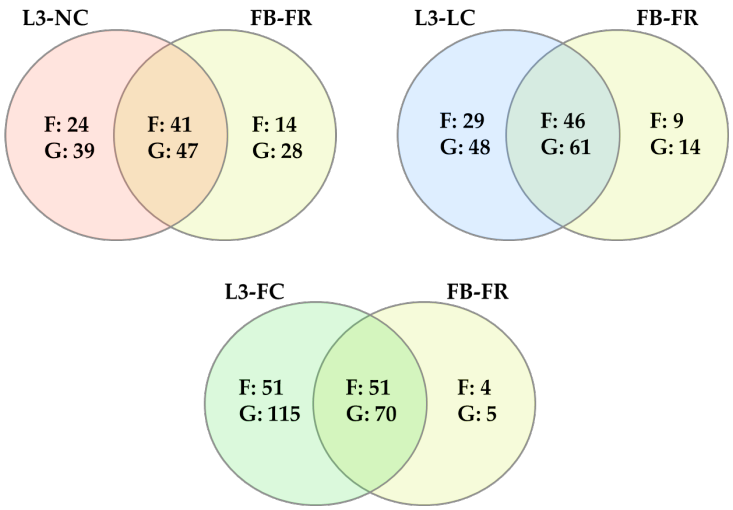


Figure 6. Venn diagrams of the number of taxa at genus (G) and family levels (F) found in the nasal microbiota of piglets raised in farms (FB-FR, in yellow) and under BSL3 conditions with different degree of contact with their sows. No-contact except for the birth canal (L3-NC) in red; limited contact of less than 12 h (L3-LC), in blue and full normal contact until weaning (L3-FC), in green. Only taxa present in more than 80% of samples per group were considered.

After confirming that L3-FC and FB-FR groups were more similar compared to other BSL3 groups (as a result of a normal contact with sows), we assessed the differential abundant taxa between these two groups, in order to explore the effect of raising piglets under BSL3 conditions. Few differences were found and most of them were detected in low abundant taxa, except for *Bergeyella*, *Rikenellaceae* and *Mycoplasma*, which were in higher relative abundance in farms, and families from *Actinobacteria*, generally more represented in L3-FC. The complete list with ANCOM results is presented in **Table S4**.

Discussion

The swine microbiota composition plays an important role in the physiology and immunity of the host (9,10,16,26,34,154). Pigs are normally raised in the farms with their sows until weaning, a critical period when piglets undergo drastic changes in life conditions, and many complex diseases can arise (9,10,12,14,16,26). Microbial colonization in early life can promote short- and long-term health benefits leading to different susceptibility to disease (16). The sow–piglet contact and the environment have emerged as main factors influencing the microbiota composition in piglets (9,10), but in particular, the effect of growing piglets in biosecurity facilities with or without the presence of sows had not been previously assessed. Here, we found that sow–contact has more impact than the facilities environment in shaping the nasal microbiota of the piglets. The BSL3 environment increased the richness of the microbiota, probably by the colonization of transient species, which were detected in low abundance.

The nasal microbiota composition of the animals raised in BSL3 facilities was different depending on the time the piglets spent with their sows. The piglets raised with sows in the facilities showed to be dominated by *Firmicutes* and *Proteobacteria*, and in a lower relative abundance by *Bacteroidetes* and *Actinobacteria* following the tendencies observed in the healthy porcine respiratory microbiome (26). For instance, families and genera related with swine respiratory or oral tract and sow skin microbiotas, such as *Staphylococcaceae*, *Aerococcaceae*, *Lactobacillus* or *Streptococcus* (*Firmicutes*), were more relatively abundant in piglets with normal sow contact (26,27,148,154,155). Similarly, *Proteobacteria* was more relatively abundant in L3-FC, especially *Enhydrobacter* and *Moraxella*, both from *Moraxellaceae* family and present in the respiratory microbiota in healthy pigs (16,26,27). In agreement, some genera such as *Moraxella*, *Rothia* and *Staphylococcus*, have been identified in teat skin and tonsils of piglets (27). Other common members of swine microbiota, such as *Porphyromonas* and

Flavobacteriaceae (*Bacteroidetes*) (154,155), *Corynebacterium* (*Actinobacteria*) (26,27), *Neisseriaceae* (*Proteobacteria*) (27) and *Treponema* (*Spirochaetes*) (154) or *Bergeyella* (156), *Haemophilus* (16,27) and *Mycoplasma* (17,157), were also identified as sow-derived in BSL3 samples, as they were found in the core microbiota of piglets that had contact with their sows. In contrast, an increase in *Firmicutes* and *Bacteroidetes* was found in the groups with no or limited contact with their sows (L3-NC and L3-LC), mainly due to an increase of taxa commonly found in the gastrointestinal tract of healthy pigs, such as *Lachnospiraceae*, *Ruminococcaceae* and other taxa from the *Clostridiales* order (154,158,159); or *Prevotella* (also found in tonsils) (154,155,160), S24-7 (161) and *Bacteroides* (154,155,159,162) from *Bacteroidales*. The richness observed in the nasal microbiota of these groups could be explained by the unusual abundance of these taxa. We hypothesize that the increased abundance of these fecal bacteria could be due to the reduced presence of sow-derived natural nasal colonizers, which would otherwise compete for the colonization, causing a decrease of community evenness, as seen in the pigs with normal contact with their sows.

The importance of sows in the development of the nasal microbiota early in life was also confirmed for BSL3 animals when the three groups were compositionally compared to normal farm samples. The normal contact with sows was more important than the environment. Importantly, we also found that the length of this contact shaped the nasal microbiota composition. Qualitative analysis showed that the longer the sow–piglet contact time was, the more common taxa were found between BSL3 and farm piglets, including swine nasal colonizers, such as *Bergeyella* (156), *Glaesserella* (*Haemophilus*) (16,27), and members from *Moraxellaceae* (16,26,27), or others found in the pig microbiota relatively abundant as *Corynebacterium* (26,27), *Treponema* (154) and *Porphyromonas* (154,155). However, the reduced abundance of some taxa containing potential pathogens in L3-FC compared to FB-FR, such as *Pasteurellaceae* (16,26,27), *Bergeyella* (156),

or *Mycoplasma* (17,157), may reflect the effect of the high biosecurity facilities as compared to the farm group. Noteworthy, *Enhydrobacter* was found in higher relative abundance in L3-FC compared to the nasal core microbiota from healthy farms (16). Some ASVs classified as *Enhydrobacter* could really belong to Moraxella (as classified in the NCBI database using BLASTn; results not shown) in agreement with the current information about swine nasal microbiome (26); however, there might also be some displacement of *Moraxella* by *Enhydrobacter* in BSL3 conditions. The inherent variability observed in farms complicates the achievement of a simple conclusion, especially in quantitative analysis where each farm is dominated by few specific taxa (possibly due to each farm's treatment, environment, etc.). Actually, the use of antimicrobials might be also affecting the results. However, when we compared the microbial composition in animals treated with ceftiofur in both groups with altered contact (L3-LC-ceftiofur and L3-NC) and the microbiota from non-treated animals with different degree of sow–piglet contact (L3-FC and L3-LC-control), the results supported the role of the sow–piglet contact as a major driver in shaping the piglets' microbiota.

In spite of the differences found between BSL3 groups, a common core was identified and included many taxa identified in previous studies from farm herds. Pena Cortes, *et al*., identified bacteria present in the vagina and teat skin of sows in strong correlation with piglet tonsillar microbiome (27) which were also identified in our nasal core microbiota of BSL3 piglets, regardless of the contact with sows. Those taxa included the *Firmicutes* families *Streptococcaceae*, *Staphylococcaceae* and *Lactobacillaceae*, as well as *Pasteurellaceae* and *Moraxellaceae* (from *Proteobacteria*), and *Micrococcaceae* (from *Actinobacteria*), and some of these were also identified in tonsils by Lowe et al. (155). Similarly, Murase *et al*. (148) found *Bacteroides* in swine vaginal mucus or *Streptococcus*, *Moraxella* and *Rothia* in saliva, which we found in increased relative abundance in the piglets with normal contact with sows. Correa Fiz *et al*. also identified most of these species

among the nasal respiratory core microbiota in healthy farms from Spain and the UK (16). The vertical transmission of most of these taxa has been also described, at least, in humans (163). Among this core, some fecal associated taxa (previously described) and other taxa that may come from food (*Cyanobacteria*) were found as well.

In this study, the nasal microbiota was assessed at the moment of weaning, which is one of the most critical moments in the development of microbial communities in piglets' lives (9,26). Important changes on the nasal microbiota were reported at this time-point depending on the variable contact of piglets with sows in artificially reared piglets. Nevertheless, the implications of the nasal microbiota later in life have not been assessed. Moreover, the main fact of having unbalanced microbial communities and especially the reduced contact with swine pathogens at the moment of weaning, could impair the control respiratory infections later in life (16). Nevertheless, the health status of these animals outside this controlled environment has not been assessed. We have shown that the nasal microbiota composition changes due to the highly-controlled rearing conditions in BSL3 facilities, but more importantly, due to the reduced contact of the piglets with their mothers. The sows provide relevant colonizers to the piglets, however, some studies require to avoid sow–piglet contact and clean facilities to reproduce endemic diseases or to avoid the interaction with other farm or environmental pathogens (150–152). More research is needed to unveil the implications that the distinct microbiota composition acquired in this particular environment may have on the conclusions arisen with these animal models.

Materials and Methods

Samples Included in the Study

The 50 samples included in the present study were taken from the nasal cavity of piglets at 3–4 weeks of age, at weaning. All piglets were normally delivered, except for L3-NC that were snatch farrowed. However, their

housing and the extent of contact with their mother differed, as depicted in Table 1. Samples were from experiments previously performed with other purposes under institutional authorization and the approval of the Ethics Commission in Animal Experimentation of the Generalitat de Catalunya (Approved Protocol Numbers 9211 and 9485).

Piglets were either born and raised in farms (Farm Born-Farm Raised, FB-FR groups), born in farms but raised under BSL3 conditions (L3-NC and L3-LC), or born and raised entirely in BSL3 (L3-FC). Samples from FB-FR group came from a previous study (16) and belong to 3 farms from Catalonia (farms IDs: GM, PT and VL), which use a farrow-to-finish or multi-site production system. All farms were PRRSV stable with no clinical signs in maternity and nursery units, and had a confirmed good health status. All BSL3-raised piglets were housed in IRTA-CReSA BSL3 facilities until sampling. The time of contact with their sows (sow-contact) was an especially important variable for BSL3-raised piglets, dividing them into 3 groups. One group included piglets taken at delivery in farms (snatch farrowed), with no contact with their mothers except in the birth canal, colostrum-deprived and transferred to the BSL3 facilities for housing immediately after birth (no-contact, L3-NC). This group was entirely treated with colistin and ceftiofur to reduce mortality, since piglets were suffering from diarrhea, as they were deprived of colostrum (no maternal-derived immunity). Another group, consisted of piglets (from four different sows) that were born in farms and stayed several hours in contact with sows (less than 12h), before they were transferred to the BSL3 facilities for housing (altered-contact, L3-LC). Half of the L3-LC group was treated with ceftiofur while the other half remained untreated. The last BSL3 group consisted of piglets born from two different sows and raised in BSL3 facilities. These piglets stayed with the sows during the whole study (full-contact, L3-FC) and were untreated since it was not necessary. All sows and piglets from BSL3L groups came from the same farm. The housing in BSL3 facilities was performed according to the study groups, i.e., L3-NC piglets were housed in

the same pen, L3-LC piglets were divided in two pens inside the same unit according to their treatment (ceftiofur and control), while L3-FC group was divided in two pens (one for each sow and piglets) inside the same unit. Finally, FB-FR piglets were born and grown under conventional industrial pig farming process where piglets remained with the sows for 3–4 weeks, when they were weaned.

DNA Extraction and 16S rRNA Gene Sequencing

Nasal samples were taken with swabs made of aluminium with a cotton tip, not flocked. The sampling procedure included the insertion of each swab into both nares. Nasal swabs were placed into sterile tube and transported under refrigeration to the laboratory, where they were resuspended (vortexed) in 500 µL of PBS and stored at −80 °C until DNA extraction. Genomic DNA was extracted with Nucleospin Blood kit (Machinery Nagel, GmbH & Co, Düren, Germany). To assess DNA quantity and quality, BioDrop DUO (BioDrop Ltd., Cambridge, UK) was used.

DNA from nasal swabs was used for 16S rRNA gene library preparation and sequencing performed at Servei de Genòmica, Universitat Autònoma de Barcelona, using Illumina pair-end 2 X 250 bp sequencing with MiSeq following the manufacturer instructions (MS-102-2003 MiSeq® Re-agent Kit v2, 500 cycle). The target region for amplification was the variable regions V3 and V4 of the 16S rRNA gene (~460 bp), using Illumina recommended primers. The PCR product was purified and checked to verify its size on a Bioanalyzer DNA 1000 chip (Agilent). Sequences were bioinformatically sorted in samples and downstream in silico analysis was performed. The sequences analyzed in this study belong to five different runs.

All the sequences included in this study are available at the NCBI database, BioProject ID PRJNA717747.

Microbiota in silico Analysis

Sequenced reads were processed using quantitative insights into microbial ecology (QIIME) 2 software version 2020.8 (115), and microbiota composition was inferred through an in-house pipeline, briefly described. First, sequenced reads were imported in QIIME2 (*Qiime Import Tools*) and quality was assessed with the *Qiime Demux* plugin. Secondly, the DADA2 algorithm (116) was used as qiime2 plugin under the default parameters to quality filter, denoise and trim low-quality reads, remove chimeras, establishing a sequence limit length of 250bp and 240 bp for the forward and reverse reads respectively (based on quality plots), join paired-ends and sort sequences into amplicon sequence variants (ASVs). An additional quality filtering was applied using VSEARCH (119) in order to remove unspecific contaminants, where ASVs not matching latest Greengenes database (13.8) (118), provided by QIIME2 project, available at qiime2 software repository (https://docs.qiime2.org/2020.8/data-resources/, accessed on February 2021); at 88% identity clustered with 65% identity and 50% query cover were filtered out with the *Qiime Quality Control* plugin (117). Since the obtained reads for the samples included in this study came from five different runs, several actions were considered to deal with possible batch effects. All these steps were done individually per run in order to train the error model of this algorithm specifically based on each run's characteristics but ran under the same parameters so as to be comparable. After this denoising step was completed, all data was merged to proceed with the downstream analysis. Sample depths were evaluated with the *Qiime Diversity Alpha-Rarefaction* plugin and normalized to an even depth in order to avoid the methodological issues [30]. Finally, after multiple aligning of sequences using MAFFT (121) and masking hypervariable positions (122) (*Qiime Alignment* plugin), a phylogenetic tree was built with Fasttree (123) in order to perform further alpha and beta diversity analysis with *Qiime Diversity* plugin.

Alpha diversity metrics were estimated with observed-features, Chao (130), Shannon (131), and Simpson (132) indices, and used to compute alpha group

significance by pairwise non-parametric t-tests (999 random permutations) in *Qiime Diversity Alpha-Group-Significance* plugin (126). Beta diversity metrics were computed with *Core-Metrics-Phylogenetic* plugin, and used to build distance matrices, in order to perform principal coordinate (PCoA) and PCoA-biplot analysis (125,164), which were visualized using Emperor (124). Beta diversity metrics used were Unweighted UniFrac (136) ad Weighted UniFrac (135),qualitative and quantitative, respectively. The percentage of variation in study groups was assessed with the Adonis Test Function from Vegan Package in R software (128). The sample group variation was calculated by PERMANOVA pairwise test, with 999 random permutations in *q2 Diversity Beta-Group-Significance* plugin (127).

Taxonomical assignment to representative sequences was obtained with the machine learning Python library Scikit-Learn using the pre-trained naïve Bayes classifier (120) trained against V3-V4 regions from 16S rRNA gene Greengenes (13.8 version) pre-clustered at 99% sequence identity, as suggested by Werner, *et al*. (165). Finally, ANCOM (166) and DS-FDR (167) approaches were used to identify taxa that were differently abundant in the studied groups. The significance in the DS-FDR test was computed with a p value, while in the ANCOM test it was represented by the W score, which is a measure of how many ratio pairs of a feature (with other features) are significantly different at a fixed taxonomic level.

Conclusions

Sows are one of the most important sources of colonizers of the upper respiratory tract of piglets in early life. Artificial rearing of piglets without the presence of sows and in highly controlled environments can have a big impact on the nasal microbiota of weaning piglets and may introduce bias into research. The nasal microbial composition of piglets that had normal contact with sows is more similar to the known healthy swine nasal microbiota, while

animals with altered sow–piglet contact were dominated by bacteria not commonly abundant in this body site. Understanding how sows influence the developing microbiota of piglets is a key point in understanding the swine microbiome and all the implications on piglets' health.

Supplementary materials

Supplementary figures

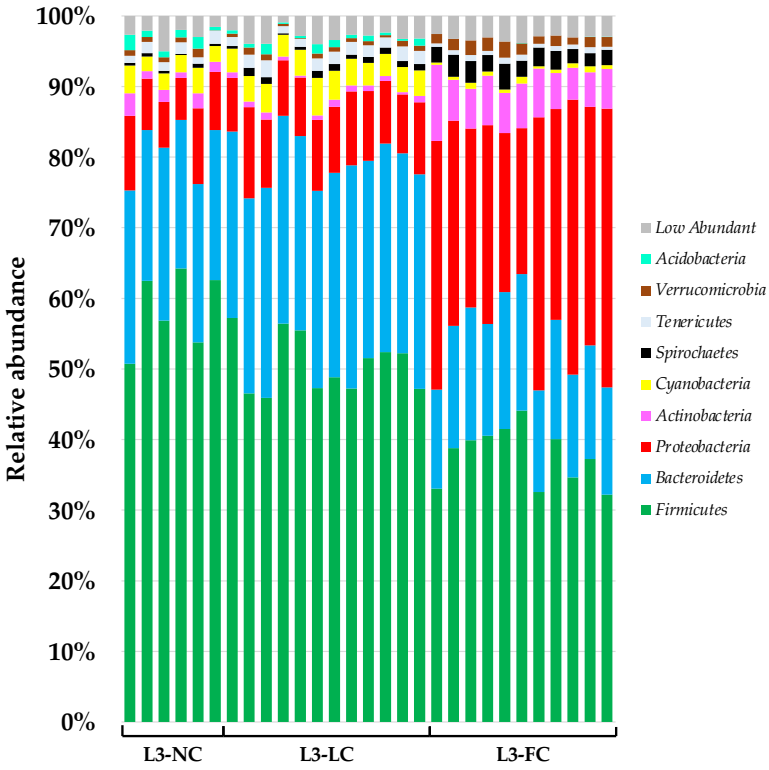


Figure S1A: Taxonomical composition at phylum level of the nasal microbiota of piglets raised in BSL3 conditions with variable contact with sows: L3-NC, no contact except the birth canal; L3-LC, limited contact of less than 12 h; and L3-FC, full normal contact until weaning at 3 weeks of age. Only taxa > 1% global relative abundance in at least one group is represented. The legend shows the taxa ordered by global relative abundance from bottom to top.

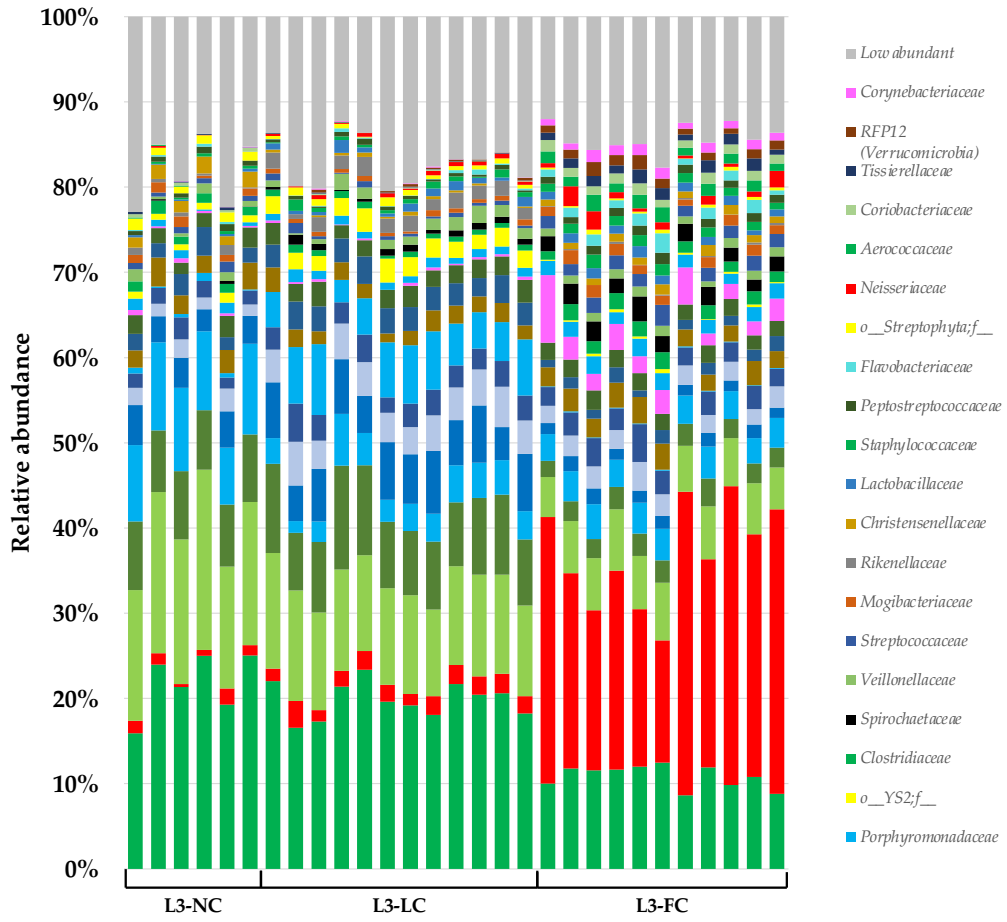
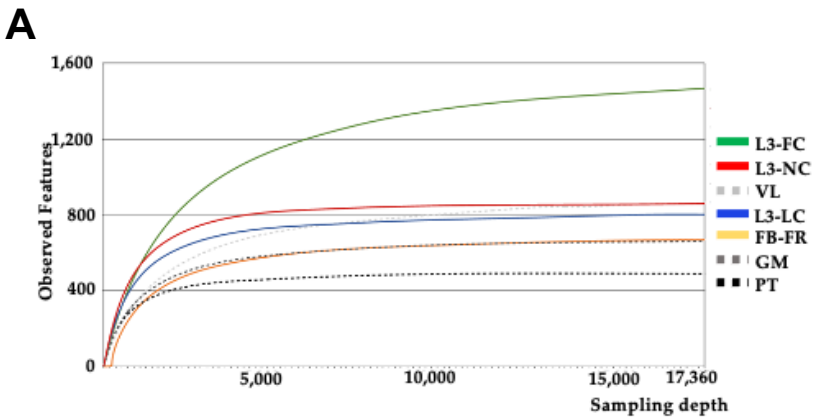
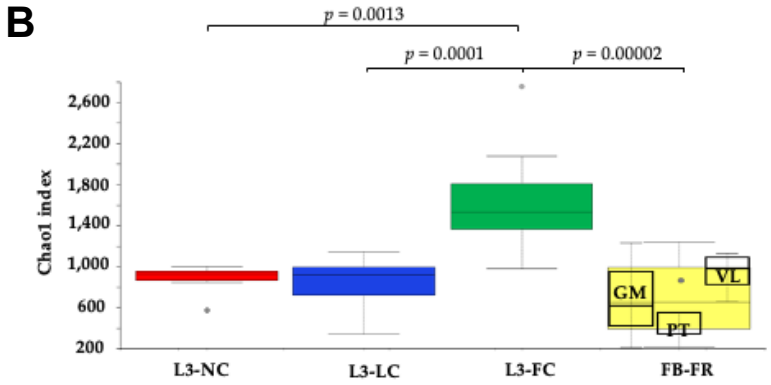
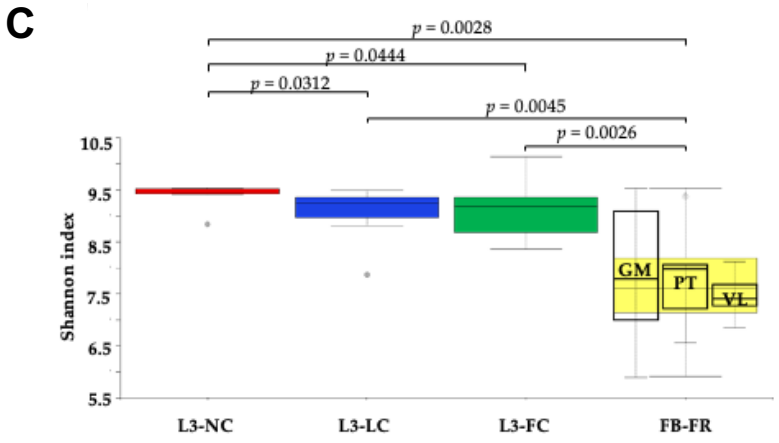


Figure S1B: Taxonomical composition at) family level of the nasal microbiota of piglets raised in BSL3 conditions with variable contact with sows: L3-NC, no contact except the birth canal; L3-LC, limited contact of less than 12 h; and L3-FC, full normal contact until weaning at 3 weeks of age. Only taxa > 1% global relative abundance in at least one group is represented. Families belonging most relatively abundant Phyla have been colored in common color-scheme to simplify its visualization; green for Firmicutes, blue for Bacteroidetes, and red for Proteobacteria. The legend shows the taxa ordered by global relative abundance from bottom to top.


A
1,600
1,200
800
400
0
Observed Features
5,000
10,000
15,000
17,360
Sampling depth
L3-FC
L3-NC
VL
L3-LC
FB-FR
GM
PT


B
p = 0.0013
p = 0.0001
p = 0.00002
2,600
2,200
1,800
1,400
1,000
600
200
Chao1 index
GM
PT
VL
L3-NC
L3-LC
L3-FC
FB-FR


C
p = 0.0028
p = 0.0444
p = 0.0312
p = 0.0045
p = 0.0026
10.5
9.5
8.5
7.5
6.5
5.5
Shannon index
GM
PT
VL
L3-NC
L3-LC
L3-FC
FB-FR

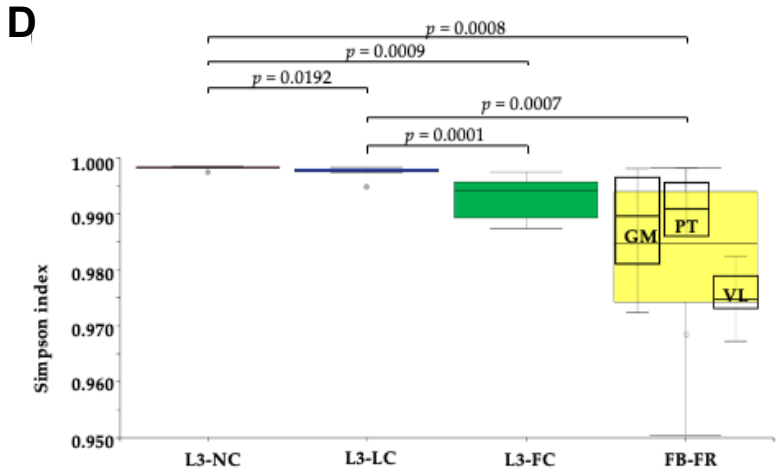


Figure S2: Alpha diversity of the nasal microbiota of piglets raised under BSL3 conditions with different degree of contact with their sows (no contact except the birth canal (L3-NC, in red); limited contact of less than 12 h (L3-LC, in blue) and full normal contact until weaning at 3 weeks of age (L3-FC, in green)); and piglets from conventional farm industry (FB-FR, in yellow). Group FB-FR included piglets from three farms (GM, PT and VL). Alpha diversity was measured at a depth of 17,360 by observed ASVs (A), Chao1 (B), Shannon–Weaver (C) and Simpson (D) indices. Significant Kruskal–Wallis pairwise tests are depicted in the top bars.

Supplementary Tables

Supplementary tables are available within the publication in:

https://www.mdpi.com/2076-0817/10/6/697

CHAPTER 3

Gut-associated microbes are present and active in the pig nasal cavity

Pau Obregon-Gutierrez[#], Laura Bonillo-Lopez[#], Florencia Correa-Fiz, Marina Sibila, Joaquim Segalés, Karl Kochanowski, Virginia Aragon. *Sci Rep.* 2024;14(1):8470. doi:10.1038/s41598-024-58681-9

[#]Equal contribution

Abstract

The nasal microbiota is a key contributor to animal health, and characterizing the nasal microbiota composition is an important step towards elucidating the role of its different members. Efforts to characterize the nasal microbiota composition of domestic pigs and other farm animals frequently report the presence of bacteria that are typically found in the gut, including many anaerobes from the *Bacteroidales* and *Clostridiales* orders. However, the in vivo role of these gut-microbiota associated taxa is currently unclear. Here, we tackled this issue by examining the prevalence, origin, and activity of these taxa in the nasal microbiota of piglets. First, analysis of the nasal microbiota of farm piglets sampled in this study, as well as various publicly available data sets, revealed that gut-microbiota associated taxa indeed constitute a substantial fraction of the pig nasal microbiota that is highly variable across individual animals. Second, comparison of herd-matched nasal and rectal samples at amplicon sequencing variant (ASV) level showed that these taxa are largely shared in the nasal and rectal microbiota, suggesting a common origin driven presumably by the transfer of fecal matter. Third, surgical sampling of the inner nasal tract showed that gut-microbiota associated taxa are found throughout the nasal cavity, indicating that these taxa do not stem from contaminations introduced during sampling with conventional nasal swabs. Finally, analysis of cDNA from the 16S rRNA gene in these nasal samples indicated that gut-microbiota associated taxa are indeed active in the pig nasal cavity. This study shows that gut-microbiota associated taxa are not only present, but also active, in the nasal cavity of domestic pigs, and paves the way for future efforts to elucidate the function of these taxa within the nasal microbiota.

Introduction

The network of microorganisms inhabiting the bodies of animals is known as the microbiota (1,3,4). The microbiota has been shown to play a pivotal role in various aspects of host health, for example by providing critical support in immune system maturation, nutrient utilization (3,4), and defense against pathogen invasion (4,168). In the case of respiratory pathogens, one of the first lines of defense is the nasal microbiota (169), and many recent studies have focused on characterizing the commensal nasal microbiota and its relationship with respiratory pathogens in humans (169,170) and various animal species (16,26,171–173). These studies have identified a variety of taxa that are frequently found in the nasal microbiota of different host species, including members from different genera such as *Moraxella*, *Lactobacillus*, *Streptococcus*, *Haemophilus*/*Glaesserella*, and *Staphylococcus* (16,26,170–173) .

Surprisingly, these studies also frequently detected microorganisms in the nasal cavity that are normally associated with the gut microbiota, in particular many anaerobic bacteria from the *Bacteroidales* and *Clostridiales* orders. These gut-microbiota associated taxa can be found in human nasal microbiota, where anaerobic Gram-negative bacteria such as *Prevotella* and *Veillonella* are frequently detected (170), but are particularly prevalent in the nasal microbiota of pigs (26). Considering that the respiratory tract is unlikely to have anaerobic niches (13), there is ongoing discussion about the *in vivo* role of these gut-microbiota associated taxa in the nasal microbiota. Recent *in vivo* studies in piglets have shown that these taxa are differently abundant under different health-status scenarios (16,17,40,46,47,50) and are variable through age stages (13,15,27), pointing towards a functional role. In contrast, other studies have attributed the presence of these gut-microbiota associated taxa to contamination from fecal material and/or soil (13,57), which could be explained by the rooting behavior of pigs (13,174,175).

In this study, we investigate the presence of these gut-microbiota associated anaerobic taxa in the nasal microbiota of piglets. Specifically, we focus on three questions: **1)** how prevalent are these gut-microbiota associated taxa in the nasal microbiota of domestic pigs? **2)** what is their source (i.e. do these microorganisms truly originate from the gut)? and **3)** are these taxa active in the aerobic nasal environment? We tackle these questions using a combination of 16SrRNA amplicon sequencing of DNA (total communities) and cDNA retrotranscribed from RNA (active communities) in matched *in vivo* samples obtained from individual animals. We confirm that gut-microbiota associated taxa indeed represent a substantial fraction of the pig nasal microbiota across a wide range of samples from this study and literature. Comparison of Amplicon Sequencing Variants (ASVs) in matched rectal/nasal samples suggests a shared pool of these taxa in both body sites, pointing to a common source. Moreover, surgical sampling of the inner nasal tract indicates that these gut-microbiota associated taxa are not introduced during sampling but are truly located in the pig nose. Finally, comparison of total and active microbial communities suggests that these taxa are active throughout the nasal cavity of pigs. Overall, this work sheds light on the role of gut-microbiota associated taxa in the nasal microbiota of pigs and supports the notion that these taxa are not only present, but also active.

Results

Characterizing the gut-microbiota associated fraction in the pig nasal microbiota

As the starting point of this study, we aimed to characterize the fraction of gut-microbiota associated taxa found in the nasal microbial communities of piglets. Towards this end, we sampled the nasal cavity and rectum of 24 healthy animals from three different commercial farms located in Spain without a history of respiratory disease outbreaks and used 16S rRNA gene sequencing to determine the microbiota composition (**Figure 1A**). We obtained a total number of 9,103 different ASVs (mean of 132,051.44 read counts per sample) after processing the raw reads. To determine the fraction of gut-microbiota associated microorganisms in these samples, we focused on two orders (*Bacteroidales*/*Clostridiales*) which constitute the most abundant taxa in the gut microbiota and defined these as "gut-microbiota associated taxa". We found that *Bacteroidales* and *Clostridiales* represented a substantial fraction of the pig nasal microbiota in most animals, accounting for 9.43 ± 3.8 % and 20 ± 7.7 % of the total composition, respectively (mean ± SD across samples) (**Figure 1B, left**). As expected, this fraction was higher in the respective rectal swabs (**Figure 1B, right**). We also determined this gut-microbiota associated fraction in the pig nasal microbiota in an alternative way by identifying taxa that are prevalent in a reference data set of gut microbiota (obtained from about 300 animals (176)) and obtained very similar results (**supplementary Text 1 and supplementary Figure 1**). Consistent with previous reports (13,16,26), in the nasal samples, *Prevotella* and *Bacteroides* together with *Ruminococcaceae* and *Lachnospiraceae,* were the most prevalent taxa within *Bacteroidales* and *Clostridiales*, respectively. Reassuringly, other reported nasal colonizers such as *Moraxella*, *Acinetobacter*, and *Enhydrobacter* genera from *Pseudomonadales*, *Lactobacillus* and *Streptococcus* genera from *Lactobacillales* and

Glaesserella from *Pasteurellales* were also highly abundant in the nasal samples (**supplementary Figure 2**).

To assess whether this substantial fraction of gut-microbiota associated taxa is a unique feature of these particular samples, we quantified the fraction of *Bacteroidales* and *Clostridiales* in 11 publicly available data sets of pig nasal microbiota samples ((16,34,35,37,38,46,52,57,177,178), see **supplementary Table 1**). We found that across different countries of origin, pig ages, and sequenced regions, *Bacteroidales* and *Clostridiales* represent a substantial (albeit variable between individual animals) fraction of the nasal microbiota (**supplementary Figure 3, supplementary Table 2**). Taken together, these results confirm that the pig nasal microbiota has a substantial fraction of commonly gut-microbiota associated taxa.

Identifying the source of gut-microbiota associated taxa in the pig nasal microbiota

The most probable source of gut-microbiota associated taxa in the pig is the gut microbiota itself, and in support of this hypothesis we found that the most abundant genera belonging to either *Bacteroidales* or *Clostridiales* in rectal samples also tended to be highly abundant in matched nasal samples (**Figure 1C**). On the other hand, most genera previously reported as nasal colonizers exhibited low abundances in rectal samples (**Figure 1C)**. However, it is conceivable that although the gut-microbiota associated taxa in these two body sites belong to the same genera, pig nose and gut may nevertheless be inhabited by distinct strains with different niche preferences.

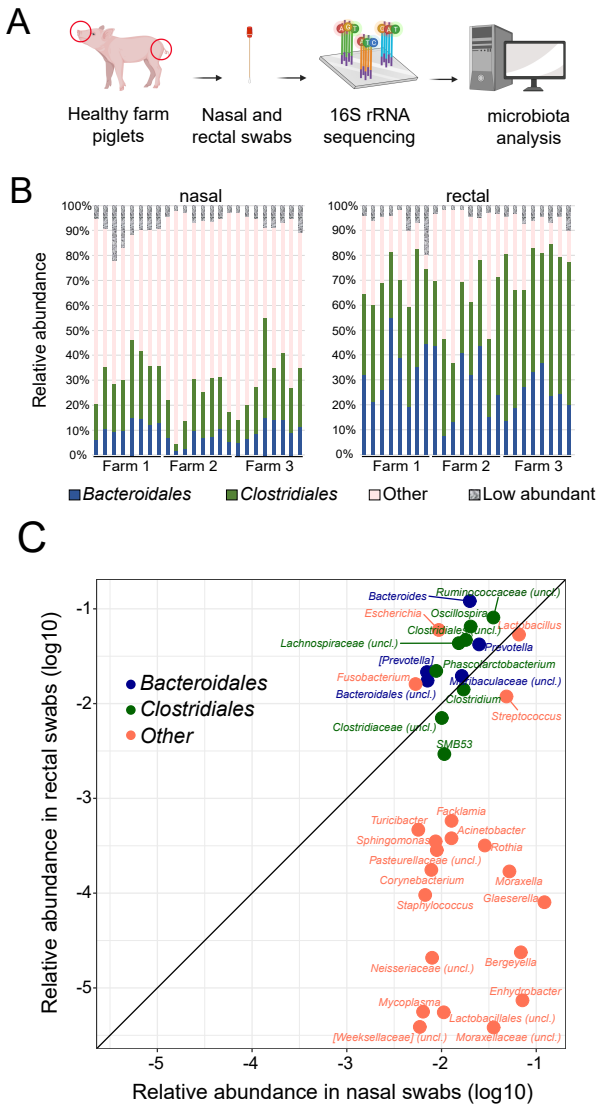
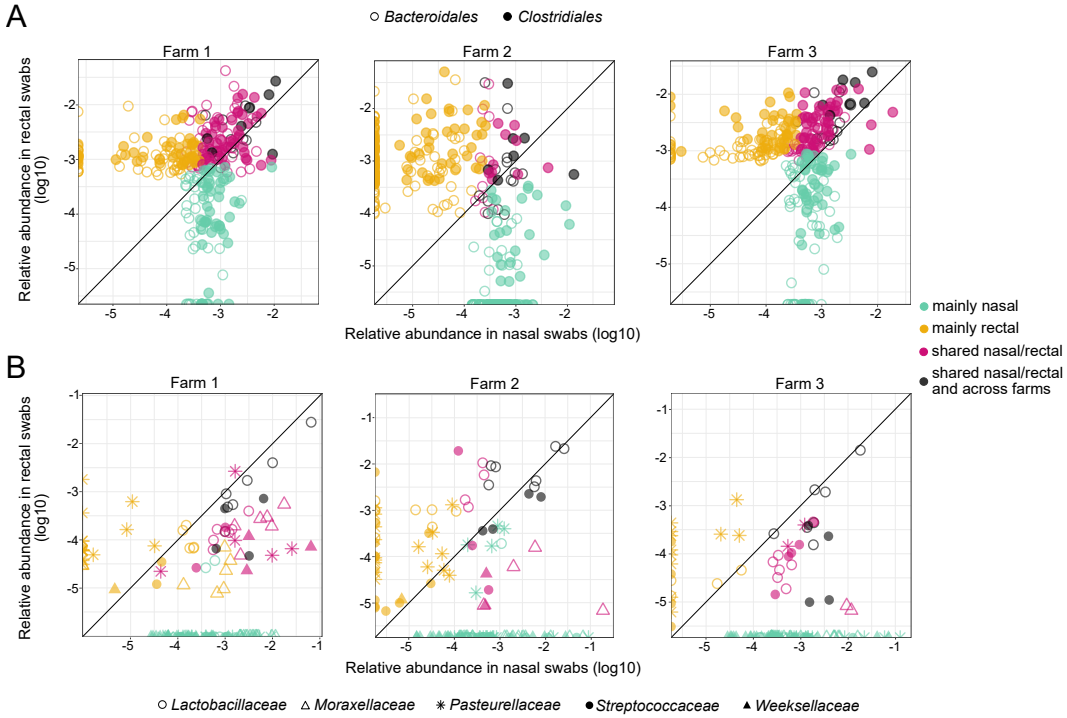


Figure 1. Detection of gut-microbiota associated taxa in nasal and rectal swabs from healthy 2–3 week-old piglets from three farms. (A) Schematic of microbiota sampling and sequencing approaches (created with BioRender.com). (B) Summed relative abundance of *Bacteroidales* (blue) and *Clostridiales* (green) taxa in nasal (left) and rectal (right) swabs of 24 individual animals. Summed abundance of other taxa (order level) with > 1% mean relative abundance across samples are shown as “Other” (pink). Orders with <1% relative abundance are summed in the category “Low abundant”. (C) Relative abundances in nasal and rectal samples (24 individuals mean) of genera in nasal microbiota with >0.5% mean relative abundance across animals. Square brackets in taxonomical assignations indicate contested names in the reference Greengenes database used (13.8).

To address this question, we examined the individual ASVs (as a proxy for strain identity (179)) found in nasal and rectal swabs within each farm. Specifically, for each farm we identified the most abundant *Bacteroidales* and *Clostridiales* ASVs in nasal and rectal samples and determined their overlap (the 100 most relatively abundant ASVs in nasal and rectal samples were selected, **Figure 2A, supplementary Table 3**). We found that 39%, 14% and 30% of ASVs belonging to *Bacteroidales,* were shared between nose and rectal samples in farms 1, 2 and 3, respectively. In the case of *Clostridiales,* these proportions were 36%, 9% and 39%. Importantly, we found that the ASVs shared between both body sites were often more highly abundant than the site-specific ones (especially in farms 1 and 3). Notably, only few of these shared ASVs were also found across farms (dark gray circles in Figure 2A), suggesting that our observations are not simply the result of biases due to high sequence conservation in these taxa. Moreover, we did not observe any biases for certain bacterial families or genera within the body-site specific ASVs. For example, we found ASVs classified as *Prevotella*, *Bacteroides*, *Veillonella* and *Oscillospira* among the site-specific ASVs in both nasal and rectal samples (**supplementary Table 3)**, and phylogenetic analysis of these most abundant *Bacteroidales* and *Clostridiales* sequences, showed little clustering across body sites and farms (**supplementary figure 4**). In contrast to *Bacteroidales* and *Clostridiales*, ASVs belonging to families of known nasal colonizers such as *Moraxellaceae*, *Pasteurellaceae* and *Weeksellaceae* showed a much lower degree of overlap between body sites (**Figure 2B**). Among the exceptions we found *Lactobacillaceae* and *Streptococcaceae*, whose ASVs still exhibited substantial overlap between the two body sites. Thus, the analysis of nasal and rectal microbiota at ASV level suggests a shared origin of the gut-microbiota associated taxa (presumably

environmental e.g. through transfer of fecal matter) found in the pig nasal and rectal microbiota.

Figure 2. Nasal and rectal overlap of gut-microbiota associated taxa (A) and nasal colonizers (B) most prevalent ASVs per sampled farm. For gut-microbiota



associated taxa the 100 most abundant ASVs in nasal and rectal samples were selected. For other nasal colonizers' families (**B**), the top 20 ASVs were considered. Each dot corresponds to one ASV (mean abundance of 8 animals from each herd). Turquoise: ASV is among the most abundant ASVs in nasal but not rectal swabs. Yellow: ASV is among the most abundant ASVs in rectal but not nasal swabs. Pink: ASV is among the most abundant ASVs in both sites. Dark gray: ASV is among the most abundant ASVs in both sites, and also shared across the three sampled farms.

Surgical deep sampling of the pig nasal cavity to identify gut-microbiota associated taxa at different depths

The shared gut-microbiota associated taxa in nasal and rectal swabs described above are consistent with the hypothesis that these taxa enter the pig nasal cavity through the transfer of fecal matter driven by the animals' rooting behavior. To further validate that these taxa were indeed resident in the pig nasal cavity and did not stem from contaminations introduced by the sampling through the nostrils, we performed additional in-depth sampling (**Figure 3A**). Specifically, we surgically opened the nasal cavity of five animals post-mortem to enable sampling of its (normally inaccessible) deep and middle parts (as detailed in the Methods section), avoiding a possible swab contamination from the skin surrounding the nostril openings. Additionally, standard nasal swabs, swabs of the external nasal area, and rectal swabs of each animal were taken for comparison. We observed a gradient in microbial load (highest to lowest, as determined by 16S rRNA gene qPCR) from external to deep internal samples, which was consistent with the number of total reads and abundance of contaminant sequences from negative control samples (**supplementary Figure 5**). Moreover, the deep nasal microbiota exhibited lower richness compared to samples from external nose (Chao1 index $p < 0.05$, **supplementary Figure 6A**), and was identified as a distinct community (as determined by beta diversity analysis) compared to the external nares (Jaccard and Bray-Curtis PERMANOVA $p < 0.05$, **supplementary Figure 6A, supplementary Table 4**), when deep nasal samples were compared pairwise with standard and external swab samples. Comparison of the composition at different nasal cavity locations within the same animal revealed that those taxa which were found at all sampling sites (of which there were only few in any given animal) tended to constitute the bulk of the observed microbiota (**supplementary Figure 7, see supplementary text 2 for detailed compositional analysis**). This dominance of few taxa was found both at the genus level, as well as at the

level of individual ASVs, suggesting that the nasal microbiota of individual animals is consistently composed of few dominating strains across the whole nasal cavity. Importantly, nasal swabs as well as matched surgical nasal samples showed a variable, but substantial, fraction of gut-microbiota associated taxa (**Figure 3B and supplementary Figure 6B**), many of which were found consistently throughout the nasal cavity in each animal (**Figure 3C**). Thus, these findings suggest that gut-microbiota associated taxa do reside in the pig nasal cavity and do not stem from contaminations introduced by the sampling procedure (i.e. sampling through the nostril using swabs) itself.

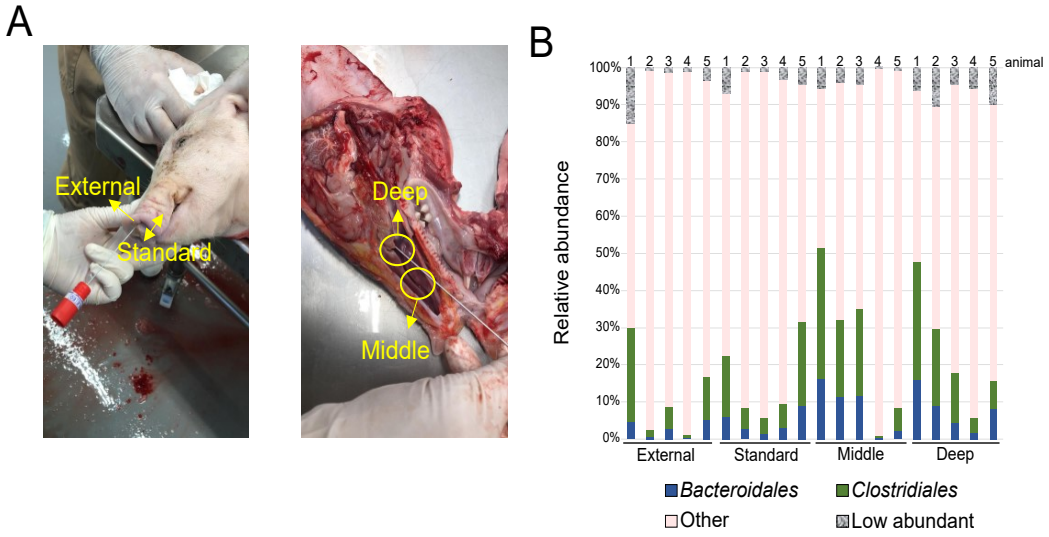


Figure 3. Characterization of surgical microbiota samples. (A) Collection of nasal samples at four sites. External and standard nasal sampling; deep and middle nose samples after longitudinal surgical cuts of piglet heads (see methods). (B) Summed relative abundance of *Bacteroidales* (blue) and *Clostridiales* (green) taxa in the different types of nasal swabs of the 5 individual animals. Summed abundance of orders with > 1% mean relative abundance across samples are shown as “Other” (pink). Taxa with <1% relative abundance are summed in the category “Low abundant”.

C

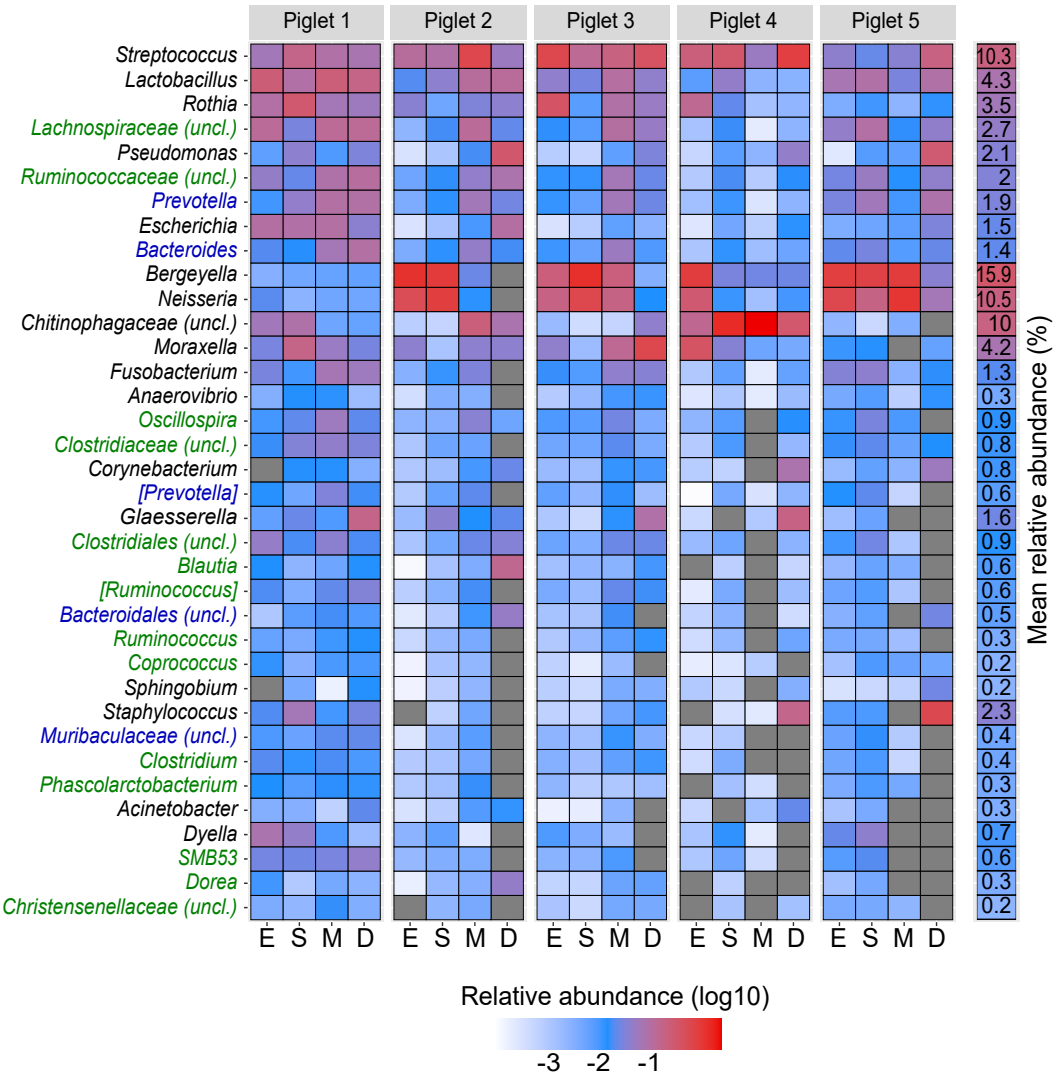


Figure 3. Characterization of surgical microbiota samples. (C) Most prevalent genera across the nose of the 5 sampled piglets. Genera are ordered from top to bottom by prevalence (present in most samples) and global relative abundance. Gut-microbiota associated genera are labelled in blue (*Bacteroidales*) and green (*Clostridiales*). Numbered column on the right: mean relative abundance across all samples. Sample sites are marked as E(xternal), S(tandard), M(iddle), D(eep).

Assessing the activity of gut-microbiota associated taxa by quantifying the 16S rRNA transcripts

The results described above suggested that gut-microbiota associated taxa are present throughout the whole pig nasal cavity. Next, we wanted to determine whether these taxa (which include many obligate anaerobic species) are also active in the aerobic environment of the pig nose. Towards this end, we quantified 16S rRNA gene transcripts (a proxy for protein synthesis potential and thus indirect measure of cellular activity, (108,180,181)) in the nasal microbiota samples described above. We found that gut-microbiota associated taxa constituted a similar portion in these RNA derived samples as in the respective DNA samples described above. Specifically, in nasal samples taken from 24 animals across 3 farms, *Bacteroidales* and *Clostridiales* orders accounted for a mean 7.1% ± 6.2 and 22.4% ± 14.1, respectively (**Figure 4A and Supplementary Figure 8A**). In the surgical samples taken at different sites in the nasal cavity, *Bacteroidales* accounted for 10.3% ± 8.8 (Deep), 8.2% ± 10.4 (Middle), 1.6% ± 3.2 (Standard) and 0.1% ± 0.1 (External); and *Clostridiales* for 3.6% ± 2.8 (Deep), 3.9% ± 3.6 (Middle), 3.2% ± 5.6 (Standard) and 2.9% ± 4.6 (External) (**Figure 4A and Supplementary Figure 8B**). Moreover, we found that gut-microbiota associated taxa had similar RNA to DNA ratios as the ratios for reported nasal colonizers when examined globally (summing across different taxa) in individual animals (**Figure 4B**), or when examining different genera (**Figure 4C**) and families (**supplementary Figure 9**). To identify individual families belonging to gut-microbiota associated taxa that deviate from this general trend, we finally compared their RNA/DNA ratio distributions with those of nasal colonizers in the farm animal samples (**supplementary Figure 10**). With some exceptions (e.g. lower RNA/DNA ratios for *Bacteroidales (unclass.)* and higher ratios for *Clostridiaceae*), RNA/DNA ratio distributions were largely not significantly different for most gut-microbiota associated taxa compared to reported nasal colonizers. Taken together, these results suggest

that gut-microbiota associated taxa are not only present in the pig nasal environment, but also active.

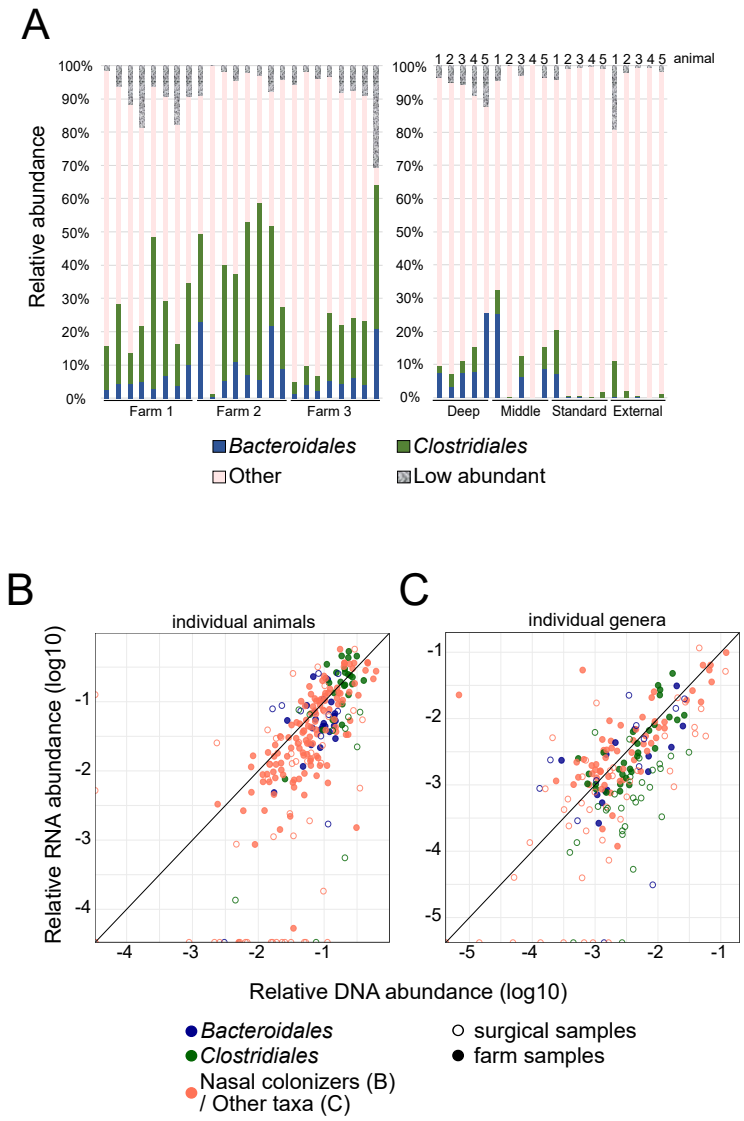


Figure 4. RNA-based quantification of nasal microbiota activity. (A) Summed relative abundance of *Bacteroidales* (blue) and *Clostridiales* (green) taxa in the two sets of nasal samples (nasal swabs from 24 farm animals, left; surgical samples from 5 animals, right; see methods). Other taxa (order level) with > 1% mean relative abundance are shown as “Other” (pink). Taxa with <1% relative abundance are summed in the category “Low abundant”). (B) Relative

abundance in RNA/DNA samples of Bacteroidales (blue), Clostridiales (green), and nasal colonizers (*Lactobacillaceae*, *Moraxellaceae*, *Pasteurellaceae*, *Streptococcaceae* and *[Weeksellaceae]*, pink). Each dot corresponds to the abundances of the mentioned taxa in one individual animal. Open circles: invasive nasal samples from 5 animals (deep and middle nasal cavity, see Figure 3). Filled circles: standard nasal swabs from 24 animals across 3 farms. Shown are mean values across all samples. (C) Relative abundance in DNA and RNA samples (as determined by 16S rRNA sequencing of DNA and cDNA, see main text) of the most relatively abundant genera, selected as those > 0.1% mean abundance in DNA or RNA farm samples. Gut-microbiota associated taxa (*Bacteroidales* and *Clostridiales*) are labelled in blue and green, respectively. Genera from other orders are shown in pink. Open circles: deep nasal samples obtained surgically from 5 animals (deep and middle nasal cavity, see Figure 3). Filled circles: standard nasal swabs from 24 animals across 3 farms. Shown are mean values across all samples.

Discussion

In this study, we aimed to elucidate the role of typically gut-microbiota associated taxa in the nasal cavity of domestic pigs. Specifically, we asked three questions: 1) how prevalent are these taxa *in vivo* and across different anatomical sites in the nasal cavity?; 2) do these taxa indeed stem from the gut microbiota?; and 3) are these taxa active in the pig nasal cavity? To answer these questions, we used a combination of regular swab and surgical deep sampling of the pig nasal microbiota and inferred its composition/activity by 16S rRNA gene DNA and cDNA sequencing, respectively. These efforts yielded three key findings.

First, we found that gut-microbiota associated taxa constitute a substantial fraction of the nasal microbiota both across individual animals, as well as across different sites within the nasal cavity. This finding is in agreement with previous studies, which sampled the pig nasal microbiota, and is also consistent with other studies relying on sampling of the lower respiratory tract of pigs, where *Clostridium* and *Prevotella* genera were prevalent (26,28,47). Thus, our results confirm that gut-microbiota associated taxa are part of the

swine nasal respiratory microbiota and are not a sampling artifact. Nevertheless, the sampling of environments that do not have high microbial biomass, including many anatomical sites in the respiratory tract (182,183), can pose technical challenges such as the false detection of transient environmental microbes (183). Although we controlled for some potential sources of contamination (sequencing blank controls and getting undisturbed nasal samples) more studies are needed to validate the inhabitants of the pig nasal microbiota.

Second, our analysis of matched nasal and rectal microbiota samples from different farms showed a large overlap between gut-microbiota associated taxa from these two sites at ASV level. This finding suggests a common source of these taxa in both anatomic sites, for example in the form of solid fecal matter that enters the nasal cavity (13,57), or through the air within a farm. In support of this hypothesis, a recent study found that gut-microbiota associated taxa are highly abundant in the air of pig farms, as well as in the nasal microbiota of pig farmers (compared to non-exposed individuals) (184,185), suggesting substantial flow of bacterial material within animal farms. In contrast to gut-microbiota associated taxa, many other dominant taxa in the nasal microbiota, such as *Moraxellaceae*, *Pasteurellaceae,* and *Weeksellaceae*, showed a low degree of overlap between the two body sites, indicating that they may be professional nasal colonizers with reduced ability to colonize other (i.e. anaerobic) niches. Interesting exceptions were the *Lactobacillus* and *Streptococcus* genera, which also overlapped substantially at ASV level between these two body sites. Given that both *Lactobacillus* and *Streptococcus* are facultative anaerobes, it is tempting to speculate whether these genera may in fact predominantly reside in the gut, but are also able to colonize the nasal cavity (and possibly other sites) if presented with the opportunity.

Third, our quantification of 16S rRNA expression showed that gut-microbiota associated taxa are indeed active in the pig nasal cavity. To our knowledge,

this is the first study which quantified pig nasal microbiota activity *in vivo*, and it suggests that these gut-microbiota associated taxa are not merely inactive transient members in the nasal cavity. Notably, recent metatranscriptomics analyses did report expression of some gut-microbiota associated taxa (e.g. *Prevotella*) in nasal samples taken from children (186) and adult patients with asthma (187), suggesting that these taxa may also be active in the human nasal cavity. One caveat is that we quantified 16S rRNA gene transcripts to determine the activity of gut-microbiota associated taxa in the pig nasal microbiota, which is an imperfect measure of cellular activity (108). Future efforts could examine the activity of these taxa with methods that quantify metabolic activity more directly, for example by using fluorescently-tagged metabolic probes (188–190) or metaproteomics (191).

This study has several limitations. First, we sampled animals at an early age (i.e. pre-weaning stage at 2-3 weeks of age). Our choice was motivated by the fact that pigs at this age are most susceptible to respiratory pathogens, while the nasal microbiota is still rather unstable (10,16). Nevertheless, our analysis of other published studies did reveal that gut-microbiota associated taxa are also prevalent at later stages in the animals' life (38,57,178), and future studies may examine to which extent the findings presented herein also hold in adult pigs and other species.

Second, although our comparison of abundant 16S DNA ASVs does indicate a shared origin of gut-microbiota associated taxa in the nasal and gut microbiota, our analysis was restricted to one section of the 16S DNA (variable regions V3-V4). Therefore, strains that share this DNA region, but differ in other parts of their 16S DNA (or other parts of their genome), cannot be distinguished with the approach chosen here. Given that we only identified few ubiquitous ASVs (that is, ASVs conserved across all three tested farms) within these taxa, we argue that our approach has the necessary resolution to detect strain differences. Nevertheless, to further validate that gut-microbiota associated strains are indeed shared by the nasal and gut

microbiota, and to identify the species these strains belong to, future studies may use either full-length 16S DNA sequencing, or metagenomics, of matched nasal and gut microbiota samples.

Third, although our surgical samples suggest that gut-microbiota associated taxa are present throughout the pig nasal cavity, the methods used in this study are not able to resolve whether these taxa are located in any specific anatomical niches. Scanning electron microscopy imaging of the upper respiratory tract of piglets has recently suggested the absence of anaerobic crypts (13), but it is conceivable that there may still be anaerobic substructures e.g. within or under the mucosal layer. Such potential substructures or locations may explain why strictly anaerobic bacterial taxa survive and are active in a basically aerobic environment such as nasal cavity. To resolve these questions, future efforts may use *in vivo* imaging methods like these (192,193) to resolve the microstructure of the nasal microbiota in more detail.

Fourth, in this study we analyzed various nasal microbiota samples, but did not specifically address the impact of environmental factors. For example, while we found that gut-microbiota associated taxa are prevalent in the pig nasal microbiota across many studies (see **Supplementary Figure 3**), their relative abundance varied widely across these samples. Moreover, the degree of overlap of gut-microbiota associated ASVs between nasal and gut microbiota samples varied between the three different farms (i.e. with farm 2 showing a much lower overlap, an observation for which we currently do not have an explanation). Further epidemiological studies may use this work as a starting point to assess the impact of environmental factors such as feeding strategies, housing conditions, or sanitation protocols, on the presence of gut-microbiota associated taxa in the nasal microbiota.

Finally, while our results do suggest that gut-microbiota associated taxa are indeed active in the pig nasal cavity, we did not examine the *in vivo* function

of these taxa within the pig nasal microbiota. Recent *in vivo* studies (16,17,40,46,47,50,113) have shown that the presence of these taxa is associated with disease outcomes and dysbiosis in various farm animals, but also in humans (187). For example, in humans with chronic rhinosinusitis, *Prevotella* (genus belonging to *Bacteroidales*) was associated with increased inflammatory severity (194). Future studies may use the results presented here as a starting point to elucidate the mechanistic underpinnings of these associations.

In conclusion, this study suggests that gut-microbiota associated taxa are indeed present and active in the nasal cavity of domestic pigs. These findings may serve as a starting point for future research aiming at elucidating the *in vivo* function of these taxa within the nasal microbiota.

Methods

Animal experimentation and ethics approval

Animal experimentation was performed following proper veterinary practices, in accordance with European (Directive 2010/63/EU) and Spanish (Real Decreto 53/2013) regulation, and in compliance with the ARRIVE guidelines (https://arriveguidelines.org/about). Sampling in farms was done with the approval of the Ethics Commission in Animal Experimentation of the Generalitat de Catalunya (Protocol number 11213). For the surgical sampling of five animals' different nasal depths, euthanasia was performed following good veterinary practices. According to European (Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes) and Spanish (*Real Decreto* 53/2013) normative, this latter procedure did not require specific approval by an Ethical Committee (Chapter I, Article 3. 1 of 2010/63/EU).

Sample collection

Farm Samples: Matched nasal and rectal swabs were taken from 24 healthy 2-3-week-old piglets from three commercial farms from Spain without reported respiratory diseases (termed Farm 1, Farm 2, Farm 3). After sampling, swabs were placed into sterile tubes filled with 1000µL DNA/RNA shield (Zymo Research) and transported to the laboratory under refrigeration.

Surgical Samples: Deep surgical samples of the nasal cavity were obtained as follows: five healthy piglets of 2-3 weeks of life were moved to IRTA-CReSA facilities and euthanized by means of an overdose of sodium pentobarbital (Dolethal). Four types of nasal swabs per piglet were taken (labeled standard, external, middle, and deep) as described below. First, a nasal swab from one nostril was taken ("standard" swab, Figure 3A). Second, an external swab was taken by introducing the swab superficially in the second nostril. Afterwards, deep and middle swabs were taken (at positions outlined in Figure 3A) from the second nostril after longitudinally cutting and

separating the skin of each piglet's head (to prevent contaminations from the skin surface) and subsequent cutting the skull, directly from the nasal turbinate without touching any other part of the piglet nose to avoid contamination and after removing the cartilaginous nasal wall. Additionally, rectal swabs were obtained from the same animals. Sample swabs were collected in sterile plastic tubes filled with 800µL of DNA/RNA shield (Zymo Research) to ensure the stability and preservation of the genetic material. Negative controls (sterile DNA/RNA shield without a swap) were included in each extraction. Samples were stored at -20ºC until extraction.

DNA/RNA extraction

Metagenomic DNA and RNA were extracted starting from 350 µL of DNA/RNA shield swab sample (previously vortexed) and following a modified protocol of ZymoBIOMICS DNA/RNA Miniprep kit (Zymo Research) in which the lysis was performed only chemically using 700 µL of ZymoBIOMICS DNA/RNA lysis buffer (2 volumes of lysis buffer per 1 volume of sample). The RNA fraction of the samples was treated with 80 µL of DNAse I (included in the kit) at room temperature for 20 minutes. Elution of both DNA and RNA was done in 50 µL of elution buffer. DNA and RNA concentration was measured using BioDrop DUO (BioDrop Ltd). DNA and RNA were stored at -80ºC until sequencing.

16S rRNA gene sequencing

16S rRNA gene libraries were prepared from the total extracted DNA and cDNA from RNA samples and sequenced at *Servei de Genòmica, Universitat Autònoma de Barcelona* (Illumina pair-end 2X300 bp, MS-102-2003 MiSeq Re-agent Kit v2, 500 cycle), using Illumina recommended primers for variable regions V3V4 of 16S rRNA gene:

fwd
5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGC
WGCAG

rev
5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGT
ATCTAATCC

The size of the amplicons was verified on a Bioanalyzer DNA 1000 chip (Agilent), as expected amplicon length was approximately 460bp. Finally, the sequences were sorted into samples and used as input for bioinformatic analysis.

The raw data used in this study are publicly available at NCBI's SRA database under BioProject ID PRJNA981084. Processed data sets are available as supplementary tables 1-7 in individual files as described below.

Bioinformatic analysis of 16S rRNA sequencing data

The microbiota composition of the samples was analyzed with quantitative insights into microbial ecology (QIIME) 2 software version 2022.2 (115). The detailed pipeline followed from raw reads to obtain tables of abundances at ASV and other taxonomic levels, the diversity analyses and the statistical tests to determine the differential abundances, can be found at https://zenodo.org/record/8013997 (Zenodo ID 8013997). Briefly, the pre-processing of the reads was done separately for each sequencing run. At first, raw demultiplexed reads with quality (fastq) were imported and their quality was evaluated with *qiime2 demux* plugin. Primers from the variable region V3-V4 were extracted using *qiime2 cutadapt* plugin (195) Sequences that did not contain primers and thus, were not a sequencing product, were removed from the analysis. DADA2 software package (116) was used under the parameters detailed in the pipeline to quality filter, paired-end merge, remove chimeras and sort reads into ASVs. Contaminant artifactual amplicons from non-prokaryotic origin were identified with *qiime2 quality control* plugin (117) by aligning with VSEARCH (119) all ASVs against Greengenes database Vs.

13.8 (118) clustered with 88% identity (available at https://docs.qiime2.org/2022.2/data-resources/). Unmatched sequences were filtered out from the analysis. Additionally, ASVs identified as *Archea*, *Mitochondria* or *Chloroplast* (which also contain 16S rRNA), were also discarded. The alignment of the remaining sequences was performed with MAFFT (121), and the hypervariable positions were masked (122) with *qiime2 alignment* plugin. Finally, ASVs found in the negative control of the deep nose dataset were removed from the analysis (21 and 41 for DNA and RNA controls, respectively, **Supplementary table 7**). The taxonomic classification of the ASVs was performed using a naïve Bayes classifier with scikit-learn Python module for machine learning (120). In order to increase the classifier accuracy (165), it was previously trained against prokaryotic 16S rRNA gene V3-V4 region, extracted from Greengenes database (13.8 version) clustered at 99% sequence identity. This qiime2 feature classifier artifact can be found at https://zenodo.org/record/8013997 (Zenodo ID 8013997). Square brackets in taxonomical assignations indicate contested names in the reference Greengenes database used (13.8).

In order to normalize uneven sampling depths (153), the diversity analyses of the samples from the surgical sampling (supplementary text 2) were performed at a normalized depth of 2291, corresponding to the lowest depth sample. The alpha diversity of the samples was estimated with Chao1 (130) and Shannon (131) indexes. Significant differences were found with pairwise non-parametric t-tests (999 random permutations) using *qiime2 diversity alpha-group-significance* plugin (126). Beta diversity was calculated with Jaccard (134) and Bray-Curtis (133) dissimilarity indexes for the qualitative and quantitative analyses, respectively. *Qiime2 core-metrics* plugin (164,196) was used to compute principal coordinate (PCoA) analysis. The percentage of explanation of the variables under study was estimated with the Adonis function from the Vegan package, in R software (128). The significance of beta diversity analyses was calculated by PERMANOVA pairwise test (999

random permutations) using *qiime2 diversity beta-group-significance* plugin (127). Differently abundant taxa between groups were found with ANCOM-BC algorithm (197). For all the stated analyses, the significance threshold *p* value was set to 0.05. To perform the phylogenetic analyses, the ASVs of interest were multiple aligned with MAFFT (121), from where the phylogeny was built with the IQtree (198) online tool (available in: http://iqtree.cibiv.univie.ac.at/) using automatic selection of the substitution model, with 1,000 bootstrap alignments and iterations. Output microbiota data processing and plot generation was performed using R script language version 4.2.2 in RStudio environment version 2022.07.0 (199), using the packages qiime2r (200), reshape2 (201), ggplot2 (202), tidyverse (203), and ggtree (204), as well as MATLAB (version 2021A).

Quantification of total 16S rRNA gene concentration

Total 16S rRNA gene concentrations (as a proxy for microbial load in swabs) were quantified as follows. Briefly, the reaction was prepared in a volume of 20 µL consisting in 2 µL of the template DNA and 18 µL of Femto Bacterial qPCR Premix, which includes a primer mix targeting the 16S rRNA (Femto Bacterial DNA Quantification Kit, Zymo Research). The PCR reaction was performed using a 7500 Fast Real-Time PCR System (ThermoFisher Scientific) at 95ºC for 10 min, 40 cycles of 95ºC for 30s, 50ºC for 30s, 72ºC for 1 min, followed by a melting curve and a final extension of 72ºC for 7 min. Each run contained two replicates per sample, a standard curve of 6 points (2 to 0.00002 ng), a negative and a positive extraction control as well as PCR negative controls. Graphpad 8.3 (538) Prism software (Dotmatics, San Diego CA) was used to analyze the data obtained from the 16s rRNA qPCR of DNA extracted from the swabs.

Supplementary materials

Supplementary text 1 – Identifying gut-associated taxa in the pig nasal microbiota

To identify taxa in the pig nasal microbiota that may potentially stem from fecal contamination, we initially focused on two orders, namely *Bacteroidales* and *Clostridiales*. These two orders represent the largest fractions of the gut microbiota in humans and pigs. Moreover, at least *Clostridiales* are considered to be obligate anaerobes (and *Bacteroidales* include obligate and facultative anaerobes), and therefore are not expected to thrive in the largely aerobic environment of the nasal cavity (see e.g., (205) for O2 levels in human respiratory tract).

To test whether the pig nasal microbiota contains other potentially gut-microbiota associated taxa not included in these two orders, we used a reference gut microbiota study from Xiao et al (176) and identified taxa that are frequently found in the microbiota of 288 healthy piglets sampled in different countries. We identified 36 taxonomical families that are present (relative abundance >= 0.01%) in at least 10% of the animals (using processed data obtained from http://gigadb.org/dataset/view/id/100187). We used these low thresholds to include also families with low prevalence. Of these taxonomical families, 33 were also detected in the pig nasal microbiota estimated from 24 farm samples. Four of these families (*Streptococcaceae*, *Staphylococcaceae*, *Lactobacillaceae*, *Pasteurellaceae*) represent well-established pig nasal commensals and were excluded from further analyses. We found that the summed relative abundance of these more refined gut-associated taxa in the nasal microbiota samples was very similar to the metric used above (i.e., labeling all *Clostridiales*/*Bacteroidales* as gut-microbiota associated), and was similarly dominated by *Clostridiales* and *Bacteroidales* (**Supplementary Figure 1**). Moreover, some taxa found in our samples belonging to these two orders could not be compared to these reference gut-

microbiota taxa due to unresolved classification at the family level. To avoid potential biases by such unresolved classification at the family level, we decided to use *Clostridiales* and *Bacteroidales* orders as our indicators of potential gut-microbiota associated taxa throughout this study.

Supplementary text 2 – Compositional analysis of invasive pig nasal microbiota samples

During the characterization of the different nose depth samples of the five animals necropsied, the bacterial load, measured by 16S rRNA gene concentration (ng/µL), was detected in a decreasing gradient from outside to inside the nasal cavity (Mann-Whitney test $p < 0.05$, **Supplementary Figure 5A**). The external nasal swabs contained the highest concentration of all nasal samples (mean of 0.081 ± 0.205 ng/µL), followed by the standard nasal swabs (0.015 ± 0.014 ng/µL), the middle nasal swabs (0.003 ± 0.002 ng/µL) and finally, the deep nasal swabs (0.001 ± 0.001 ng/µL). The concentration of bacteria was even higher in rectal samples (2.211 ± 1.986 ng/µL). Similarly, the number of reads obtained from 16S rRNA gene sequencing from each type of samples also decreased through the depth of the nose, while the representation of the ASVs found in the negative control correlated inversely as these represented a mean relative abundance of 18.3%, 3.2%, 1.5%, 0.6% and 0% in deep, middle, standard, external and rectal samples. (**Supplementary Figure 5B**).

When the diversity between the five types of samples was compared (beta diversity) a strong effect size was found in both qualitative (23%) and quantitative (29%) analyses (Adonis function R2, $p < 0.05$), as rectal samples formed a differential cluster. In order to exclusively compare the different types of nasal samples, we excluded the rectal samples from the diversity analysis. Interestingly, despite some individual pairwise differences, the quantitative beta diversity analysis became non-significant statistically

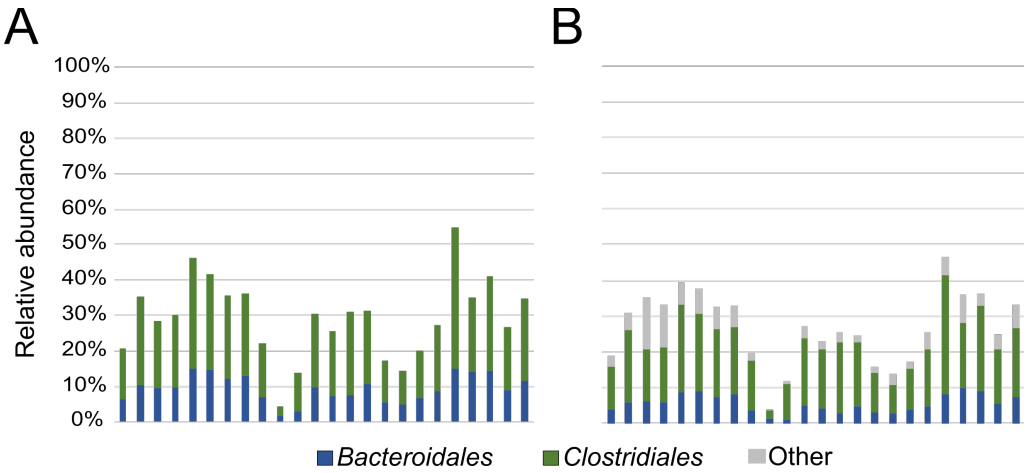
(PERMANOVA $p > 0.05$). On the contrary, the qualitative analysis still explained the 19% of group differences (Adonis test R2, $p = 0.003$). In the pairwise analysis, we detected that most differences occurred between the deep nose and the external samples in both qualitative and quantitative analyses (**Supplementary Figure 6A**). Regarding the alpha diversity, there were no differences between nose locations (Shannon index $p > 0.05$), but deep nasal samples reported a lower species richness than the external and standard nasal samples (Chao1 index, **Supplementary Figure 6A**, $p < 0.05$).

In order to characterize the most prevalent bacteria in the different parts of the swine nasal tract, we focused in the most dominant taxa (**Supplementary Figure 6B**). The most abundant taxa in the deep nose belonged to the orders *Lactobacillales*, such as *Streptococcus* and *Lactobacillus*; *Clostridiales* (composed of lower abundant genera within the families *Ruminococcaceae* and *Lachnospiraceae*); and *Pseudomonadales*, mainly represented by *Moraxella* and *Pseudomonas*. Other relatively abundant genera were *Haemophilus*, *Chitinophagaceae (uncl.)* and *Staphylococcus*. Samples from the middle nose were dominated by *Clostridiales* (with a similar composition of many low abundant genera as the deep nose), *Lactobacillales* (mainly *Streptococcus* and *Lactobacillus*) and *Bacteroidales*, with *Prevotella* and *Bacteroides* as the most abundant. *Neisseria* was the dominating genus in only one sample. One sample was fully composed of the *Chitinophagaceae (uncl.)* genus (94%). The most relatively abundant taxa in the samples from the external nose were *Neisseria*, *Streptococcus*, *Rothia*, *Enhydrobacter* (*Pseudomonadales*), *Moraxella* and *Lactobacillus*, with their respective orders as the most prevalent. Again, *Clostridiales* and *Bacteroidales*, which were among the most abundant orders were composed of low abundant genera. Regarding the standard samples, Neisseria was the predominant genus in three of the samples, while *Chitinophagaceae (uncl.)* was in another. *Streptococcus*, *Lactobacillus*, *Rothia*, *Moraxella* and *Bergeyella* were among the most abundant genera, with less variability between animals. Although

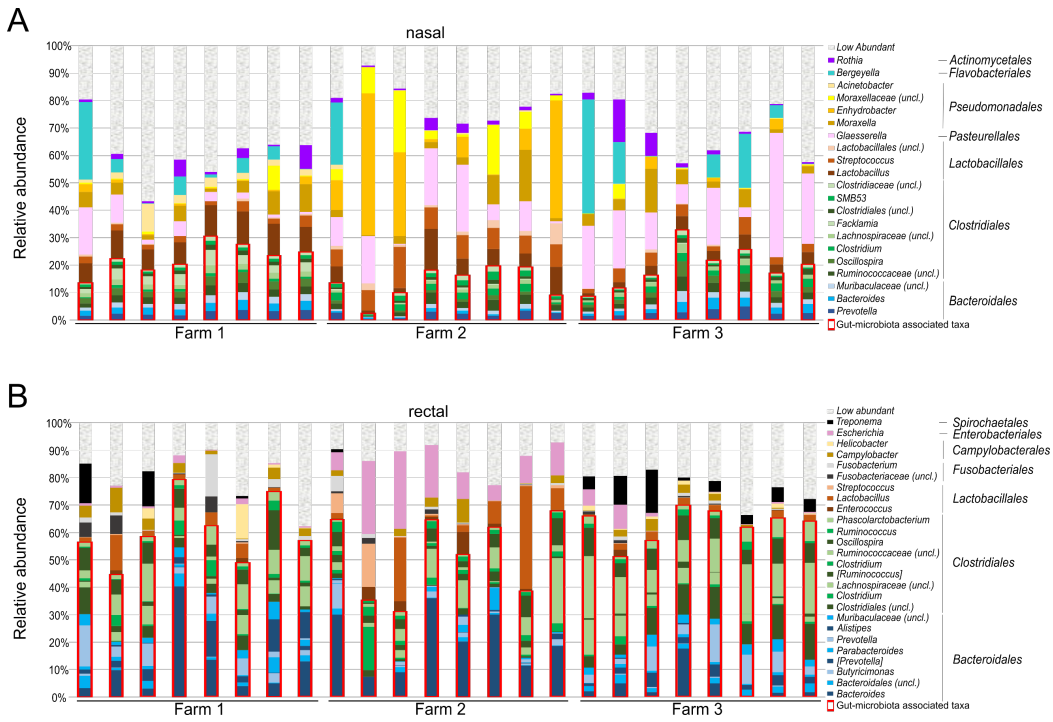
Clostridiales and *Bacteroidales* were the 4th and 5th most abundant orders, respectively; they were composed of several low abundant genera.

The differential composition of the microbiota from the deep nose samples was compared with the rest of nasal samples using ANCOM-BC at order, family, genus, species and ASV level. Although the reduced number of samples and the variability between animals from this dataset complicated the search of differently abundant taxa, some differences were identified. Among the differently abundant taxa identified in the external part of the nose, the most relatively abundant were *Flavobacteriales*, *Dyella*, and a highly abundant *Neisseria shayeganii* ASV that was absent in the deep samples. On the contrary, Pseudomonas seemed to constitute a bigger portion in the microbiota of the deep nose compared to the other nasal cavities. All the differentially abundant taxa identified comparing the deep and other nasal samples are shown in **Supplementary Table 4**.

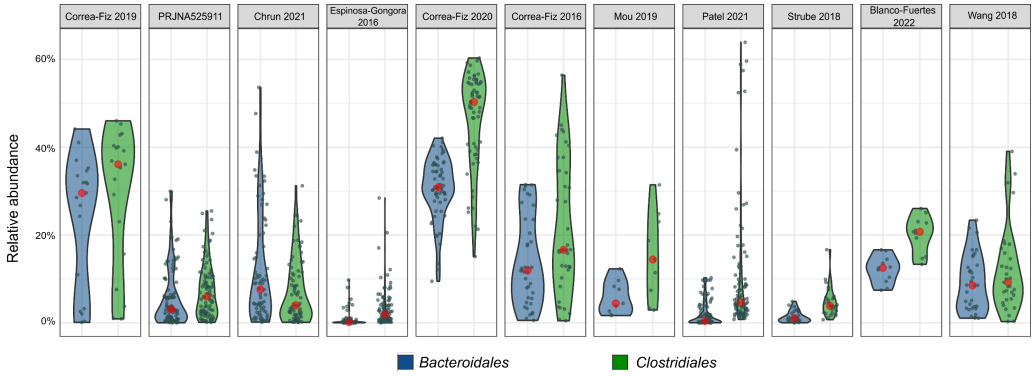
Supplementary Figures



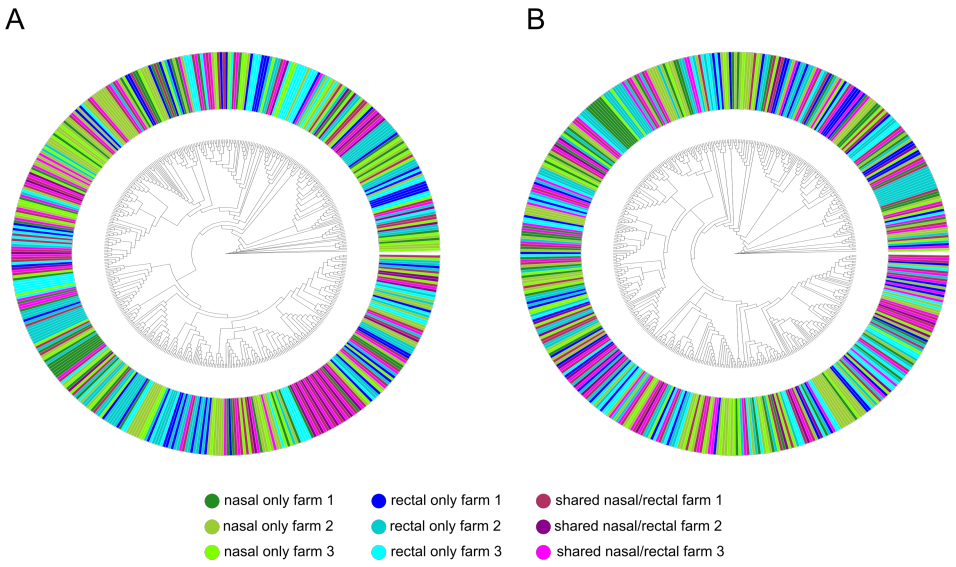
Supplementary Figure 1. Detection of gut-microbiota associated taxa in nasal swabs. A) Relative abundance of taxa belonging to *Bacteroidales* (blue) and *Clostridiales* (green) orders in nasal swabs of 24 animals. B) Order-level relative abundance of gut-associated taxa using an alternative metric (prevalence of respective family in a reference gut microbiota data set, see supplementary text 1). Data are summed by order, namely *Bacteroidales* (blue), *Clostridiales* (green), and all other orders (grey).



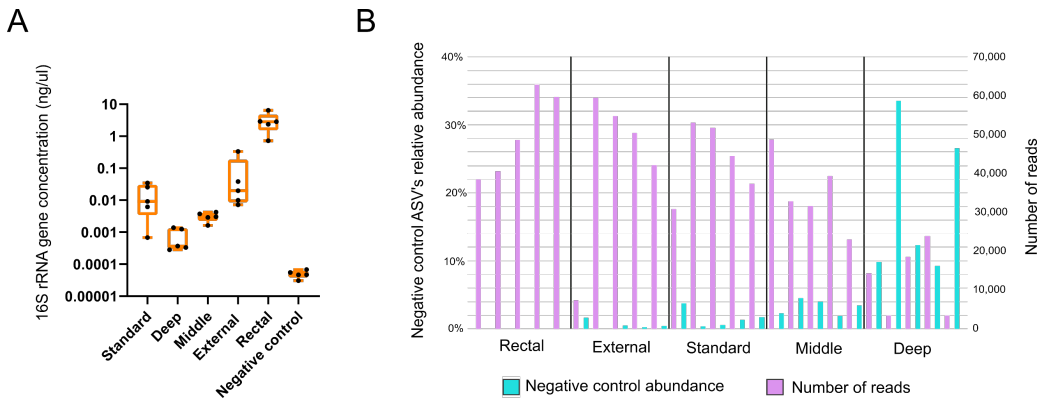
Supplementary Figure 2. Composition of nasal (A) and rectal (B) microbiota in 24 individual animals across 3 different farms at genus level. Highlighted in red: gut-microbiota associated taxa. Note that only taxa with > 1% relative abundance are labeled (taxa with <1% relative abundance are summed in the category “Low abundant”).



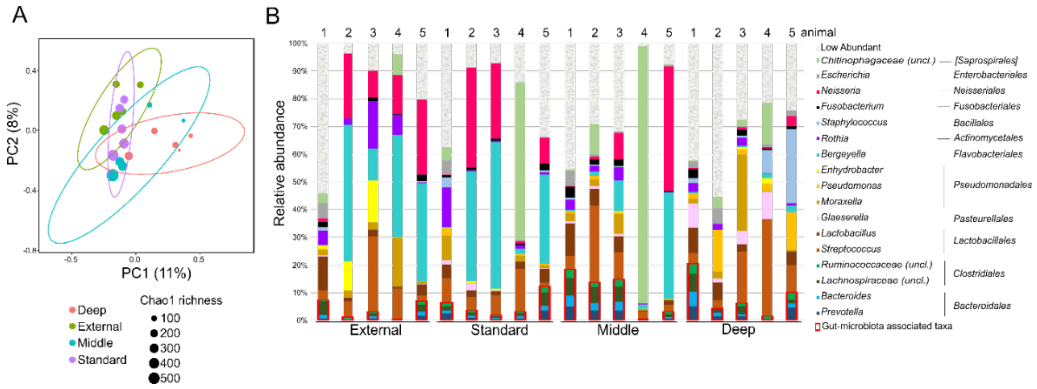
Supplementary Figure 3. Fraction of gut-microbiota associated taxa (i.e. taxa from *Clostridiales* and *Bacteroidales* orders) in nasal microbiota samples from 11 publicly available data sets. Data sets are listed in Supplementary Table 1. Distribution of data across individual animals is represented in separate violin plots for each data set. Red circle: median. Grey dots: abundance in each individual sample.



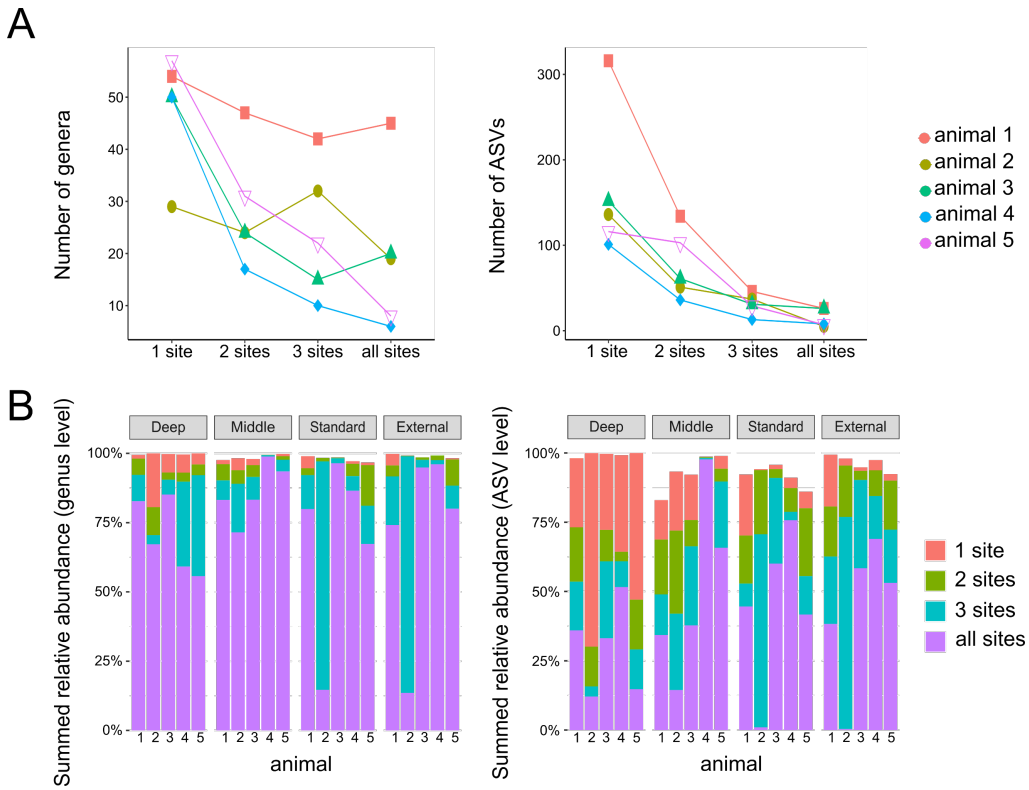
Supplementary Figure 4. Phylogenetic analysis of the 100 most abundant Bacteroidales (A) and Clostridiales (B) ASVs in nasal and rectal samples. Phylogenetic trees were inferred for the ASVs shown in Figure 2A (see methods). Color palette code follows the body site these ASVs are most abundant in: green for nasal samples, blue for rectal samples and purple for both sites.



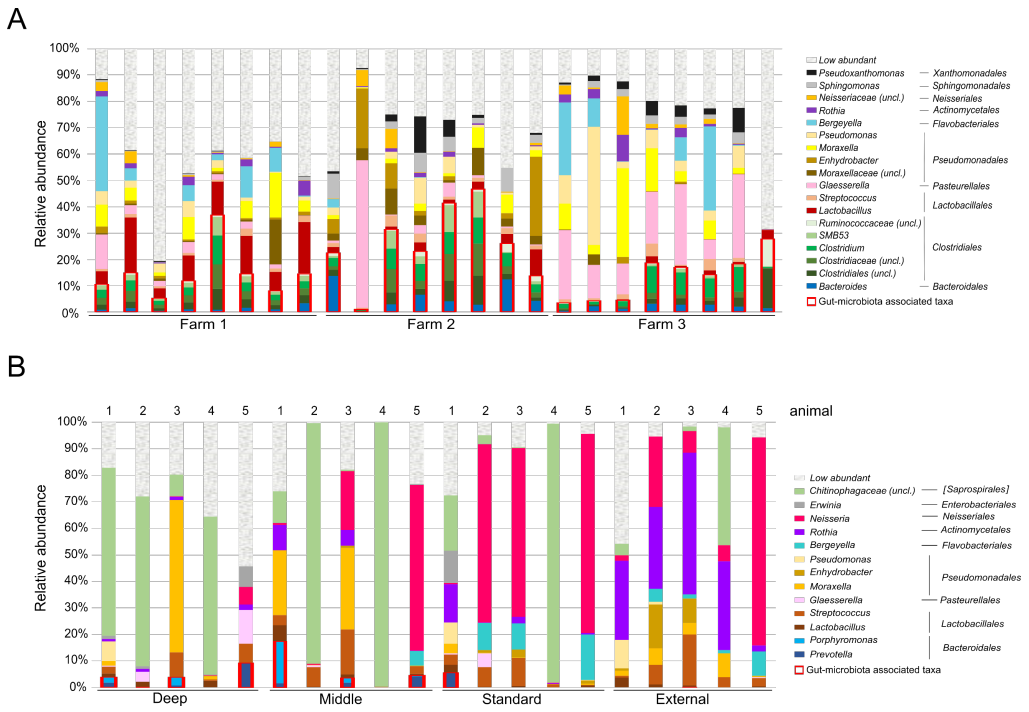
Supplementary Figure 5. Biomass abundance assessment in surgical samples. A) Total 16S rRNA quantification of the samples from DNA extracted from the nasal invasive dataset (see methods). In orange, samples obtained from nasal standard swabs. qPCR was performed twice for each sample (data shown is mean of technical replicates). All pairwise differences are statistically significant with Mann-Whitney test ($p < 0.05$) B) Summed relative abundance of the negative control ASVs (turquoise) and number of reads (purple) in each of the same samples.



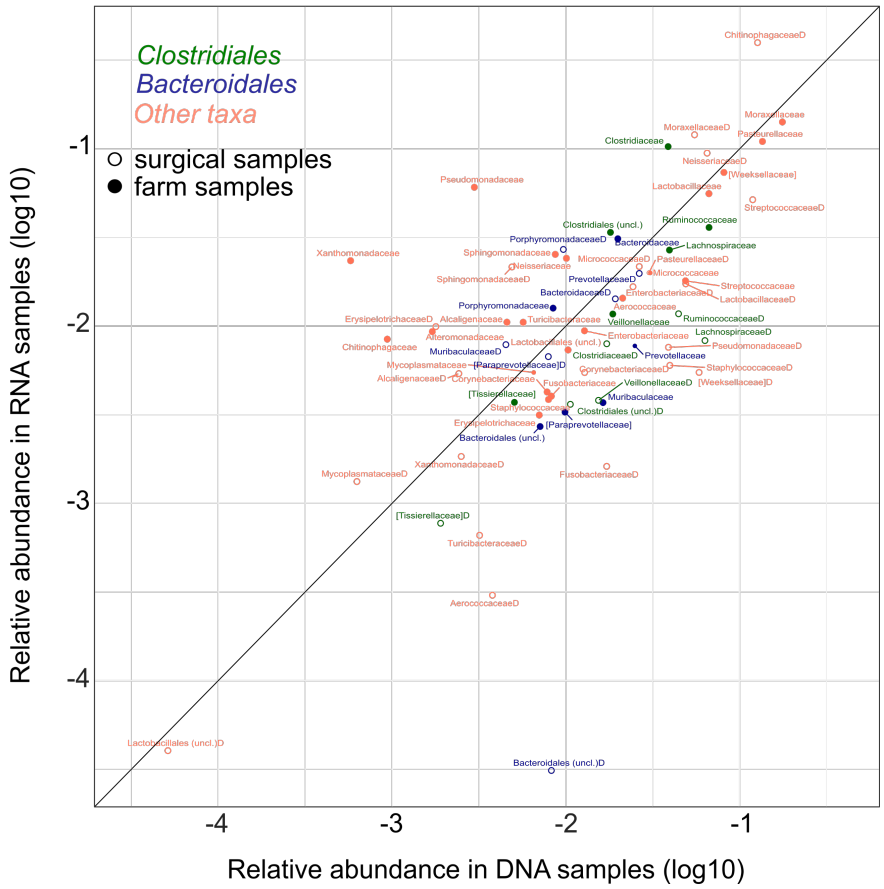
Supplementary Figure 6. Composition of surgical nasal cavity samples. A) Principal Component Analysis of microbiota samples performed with Jaccard dissimilarity index (shown here: first two principal components with respective explained variance). Each circle denotes an individual animal, and circle size denotes the alpha diversity estimated by Chao1 richness index. Group ellipses are calculated with the euclidean distances of the samples within each group. B) Detailed composition of individual samples at genus level. The order level for each genus is also included in the legend for clarification. Highlighted with red squares: gut-microbiota associated taxa. Note that only taxa with > 1% relative abundance are labeled (taxa with <1% relative abundance are summed in the category “Low Abundant”).



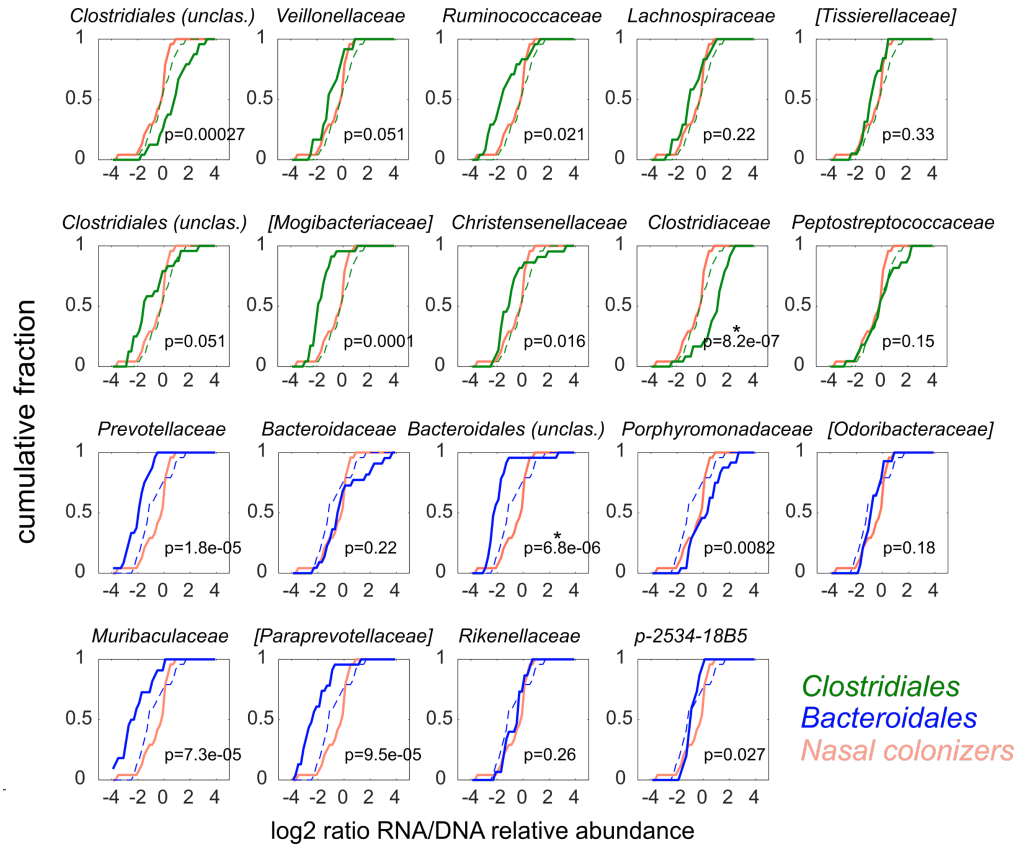
Supplementary Figure 7. Prevalence of genera and ASVs at different nasal sampling sites. A) Number of genera (left) or ASVs (right) detected at different sites (from “detected at 1 site only” to “detected at all 4 sites”) in each individual animal by accounting the four types of nasal samples from the invasive dataset (see methods). Only genera and ASVs detected >0.1% relative abundance in at least one sample were considered as “detected”. Genera/ASVs that never exceeded these thresholds are not shown. Note that only a small fraction of genera/ASVs are typically detected at all sites. B) Respective summed relative abundance for genera (left) and ASVs (right) detected at different sites (determined as in A). Note that summed relative abundances are dominated by the small fraction of genera/ASVs that are detected at all sites (purple bars).



Supplementary Figure 8. RNA-based Nasal microbiota composition as determined by 16S rRNA cDNA sequencing at genus level. A) Nasal swabs of 24 animals sampled from 3 farms (see methods). B) Surgical samples taken at different sites in the nasal cavity from 5 animals. Highlighted in red: gut-microbiota associated taxa. Note that only taxa with > 1% relative abundance are labeled (taxa with <1% relative abundance are summed in the category “Low abundant”).



Supplementary Figure 9. Comparison of 16S rDNA and rRNA mean abundances for the most abundant families in farm and surgical samples. Relative abundance in DNA and RNA samples was determined by 16S rRNA sequencing of DNA and cDNA, see main text for families with > 0.5% mean abundance in DNA or RNA farm samples. The abundances of the selected families in deep surgical nasal samples are shown as well (with an extra "D" in the label). Gut-microbiota associated taxa (*Bacteroidales* and *Clostridiales*) are labelled in blue and green, respectively. Families from other orders are shown in pink. Open circles: deep nasal samples obtained surgically from 5 animals (deep and middle nasal cavity, see Figure 3). Filled circles: standard nasal swabs from 24 animals across 3 farms. Shown are mean values across all samples.



Supplementary Figure 10. Distribution of RNA/DNA ratios in gut-microbiota associated taxa at family level in nasal swabs from 24 farm animals. Top two rows: *Clostridiales*, bottom two rows: *Bacteroidales*. Black: ratio for well-established nasal colonizers (combining *Moraxellacae*, *Pasteurellaceae*, *Streptococcaceae*, *Lactobacillaceae*, [*Weeksellaceae*]). Green/Blue: ratio for gut-microbiota associated family of interest (dashed line: corresponding order level data). P = p-value of two-sample Kolmogorov-Smirnov test (testing whether the two samples stem from the same continuous distribution). Plots denoted with *: p < 10-5. Square brackets in taxonomical assignations indicate contested names in the reference Greengenes database used (Version 13.8). Only taxa which had a relative abundance (at DNA level) > 0.1% in at least 50% of the samples were considered.

Supplementary tables

Processed data sets are available as supplementary tables 1-7 in individual files, available within the publication in:
https://www.nature.com/articles/s41598-024-58681-9

CHAPTER 4

Pig nasal and rectal microbiotas are involved in the antibody response to *Glaesserella parasuis*

Pau Obregon-Gutierrez, Yasser Mahmmod, Emili Barba-Vidal, Marina Sibila, Florencia Correa-Fiz, Virginia Aragón. Sci Rep. 2025;15(1):2347. doi:10.1038/s41598-025-85867-6

Abstract

Vaccination stands as one of the most sustainable and promising strategies to control infectious diseases in animal production. Nevertheless, the causes for antibody response variation among individuals are poorly understood. The animal microbiota has been shown to be involved in the correct development and function of the host immunity, including the antibody response. Here, we studied the nasal and rectal microbiota composition in association with the antibody response against the pathobiont *Glaesserella parasuis*. The nasal and rectal microbiotas of 24 piglets were sampled in two farms before vaccination and in one unvaccinated farm (naturally exposed to the pathobiont) at similar time. Microbiota composition was inferred by V3V4 16S rRNA gene sequencing and bioinformatics analysis and the antibody response was quantified using the variation between the levels before and after vaccination (normalized per farm). Piglets with higher antibody responses showed more diverse nasal and rectal microbial communities compared to piglets with lower responses. Moreover, swine nasal core microbiota colonizers were associated with higher antibody levels, such as several members from *Bacteroidales* and *Clostridiales* orders and genera including *Moraxella*, *Staphylococcus*, *Fusobacterium* and *Neisseria*. Regarding taxa found in the rectal microbiota, associations with antibody responses were detected only at order level, pointing towards a positive role for *Clostridiales* while negative for *Enterobacteriales*. Altogether, these results suggest that the microbiota is associated with the antibody response to *G. parasuis* (and probably to other pathogens) and serves as starting point to understand the factors that contribute to immunization in pigs.

Introduction

Animals live in constant contact with a variety of pathogens, which are controlled by the immune system. Antibodies, which target specific pathogens, constitute one of the main effector responses of adaptative immunity (206). The adaptive immune response can hold memory through antigen-specific memory cells, permitting a more effective response in subsequent encounters with the pathogen (206,207). One of the best ways to stimulate immunization and memory against pathogens is vaccination (207), which represents an efficacious strategy to control infectious diseases nowadays (208). Vaccines have the potential to reduce disease severity, eliminate pathogens locally and even eradicate them globally (208) . In consequence, they also contribute to the reduction in the use of antibiotics, which is particularly needed for minimizing the emergence of multidrug resistant bacteria (209). However, not all individuals exhibit the same level of response to vaccination. Several factors such as maternal immunity, host genetics and environmental factors, among others, can be responsible for vaccination failure (210). Therefore, it is crucial to further explore why the immune response (either after natural exposure to the pathogen or after vaccination) is variable among individuals (209–213), and to identify the factors that contribute to a robust response (212,213) .

The animal microbiota, the community of microorganisms inhabiting different niches of the animal host, is known to have an important role in the immune system maturation and modulation (212). The microbiota is involved in the local immune response, such as the stimulation of immune cells with bacterial compounds in the intestine or airway mucosae, but also in systemic immune responses via dissemination of microbial products and/or immune signals and cells through the whole organism (212). Moreover, the stated systemic effect can be enhanced by the constant crosstalk of the different microbiotas in an organism (214). In agreement, poor or deficient immune responses in the context of microbiota dysbiosis have been reported (212). The relationship

between the microbiota and the immune response is still to be unveiled, making the study of the microbiota a key point in vaccination efficacy studies (215).

Mounting evidence shows that microbiota composition influences responses to vaccination in humans (212,215–224), where several studies showed that patients with disrupted microbiota tend to show a reduced response to vaccination (224,225).

In the case of swine, vaccination programs are essential to face the strong institutional call towards the reduction in the use of antibiotics in animal farming (91) and are critical to control infectious diseases, such as Glässer's disease (8), an endemic disease caused by *Glaesserella parasuis*, a pathobiont member of the porcine nasal microbiota that colonizes young piglets early after birth (76). Few studies assess the microbiota in relation with vaccine response in pigs. In 2019 and 2020, Munyaka *et al*. studied the fecal microbiota as a predictor of high and low vaccine response, measured by the levels of specific antibodies against *Mycoplasma hyopneumoniae* in serum (226,227). They were able to identify several taxa that discriminated the high vaccine-responders and whose presence was positively correlated with antibody titters. In a different study, Sanglard, *et al*. showed that the composition of the vaginal microbiota of sows discriminated between high and low antibody-responders against a porcine reproductive and respiratory syndrome virus (PRRSV) vaccine (228).

Here, using Glässer's disease as a model where antibodies are important for protection, we studied the composition of the nasal and rectal microbiota in pigs and identified taxa associated with different level of antibodies after bacterin vaccination and/or natural exposure to the pathogen.

Methods

Samples included in the study

Twenty-four piglets were randomly selected from different litters from three Spanish farms (eight per farm). Two of the farms (farms 1 and 2) performed vaccination against Glässer's disease with a commercial bacterin (HIPRASUIS® GLÄSSER) by injecting 2ml/piglet at 3 and 6 weeks or at 2 and 5 weeks of age, respectively. In addition, in farm 2, sows were also vaccinated with the same vaccine before farrowing and penicillin was administered to the piglets at first day of life. To measure antibody levels, serum samples were taken within the 24h before the first vaccination and 3 weeks after the second vaccination (9 and 8 weeks of age for farm 1 and 2, respectively). Samples were transported under refrigeration to the laboratory and were processed within 48h after collection. Nasal and rectal microbiota samples were obtained from both nostrils and rectum using thin aluminum cotton swabs (Deltalab) taken before the first vaccination (together with the pre-vaccination serum samples). Similar times were used for the sampling of non-vaccinated pigs in a third farm (farm 3); i.e. microbiota samples at 1 week and serum samples at 1 and 9 weeks of age. Swabs were stored in 1000 µL DNA/RNA shield (Zymo Research) at 4ºC before further processing. These samples were also used in a previous study assessing gut-associated components of the nasal microbiota (229).

Animal experimentation was performed following proper veterinary practices, in accordance with European (Directive 2010/63/EU) and Spanish (Real Decreto 53/2013) regulation, in compliance with the ARRIVE guidelines (https://arriveguidelines.org/about). Animal sampling in farms was approved by the Ethics Commission in Animal Experimentation of the Generalitat de Catalunya (Protocol number 11213) and the owners of the farms.

Antibody levels

Antibodies against *G. parasuis* were measured in serum using an in-house ELISA previously described (76). Plates were coated overnight at 4ºC with 250 ng of F4 (a protein fragment from the outer membrane proteins VtaA of *G. parasuis*) in 50 µl of carbonate-bicarbonate buffer per well. After washing, wells were blocked with 1% casein in phosphate-buffered saline (PBS) with 0.05% Tween 20 (PBS-Tw20). Sera were diluted 1:100 in blocking solution and added to the wells. After 1 h of incubation at 37ºC, wells were washed and incubated with a goat anti-porcine IgG HRP-conjugated antibody (Sigma-Aldrich, Madrid, Spain) diluted 1:10,000. Positive reactions in the ELISA were developed using the 3,3,3,5-tetramethylbenzidine (TMB) substrate (Sigma-Aldrich, Madrid, Spain) and the reactions were stopped with 1 N sulfuric acid. Plates were then read in a Power Wave XS spectrophotometer (Biotech, Winooski, VT, USA) at 450 nm.

A preliminary classification into good (higher response) or bad (lower response) responders was performed in each group according to the variation between the initial level of antibodies and the level of antibodies at 8-9 weeks of age (delta antibody value, ΔAb). The piglets showing variations above the median of the group were considered as good responders and those below the median as bad responders. To deal with farm variability, this classification was done independently within each farm. However, among the 24 piglets included in the study, two piglets were not classified following these criteria. One piglet from farm 2 was considered a bad responder despite showing a ΔAb above the median, since the clear reduction showed in the level of antibodies (from 1.197 to 0.605; delta = -0.592). For one piglet from farm 3, we did not have the initial levels of antibodies and therefore it was not included in the analyses using the ΔAb. Nevertheless, it was considered a bad responder according to its final level of antibodies in comparison to the rest of the animals in the same farm. Hence, out of the twenty-four piglets, ten

were considered to have a good response while fourteen were considered bad responders.

DNA extraction

DNA was extracted from the nasal and rectal swabs following ZymoBIOMICS protocol, with the following modifications. Microbial cell lysis was performed by adding 700 µL of lysis buffer to 350 µL of sample, and DNA was eluted in 50 µL of elution buffer. DNA concentration was measured with a BioDrop DUO (BioDrop Ltd) and stored at -80ºC. A negative sample consisting of DNA/RNA shield alone was included as control.

Detection of G. parasuis by PCR

The presence of virulent and non-virulent strains of *G. parasuis* in the nasal samples was confirmed by PCR of the specific *vtaA* leader sequence, as described in Galofré-Milà, *et al*. (72), that allows the differentiation of virulent and non-virulent *G. parasuis*.

Microbiota sequencing

The library preparation and Illumina sequencing were performed at Servei de Genòmica, Universitat Autònoma de Barcelona. Variable regions 3 and 4 (V3-V4) from 16S rRNA gene were sequenced from genomic libraries prepared using Illumina recommended primers for these variable regions of the gene and following Illumina protocol (Illumina pair-end 2X250 bp, MS-102-2003 MiSeq Re285 agent Kit v2, 500 cycle).

fwd

5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG

rev

5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC

Sequenced amplicons lengths were checked on a Bioanalyzer DNA 1000 chip (Agilent).

The raw sequencing data used in this study can be found at NCBI's SRA database under BioProject ID PRJNA981084.

Microbiota bioinformatic analysis

The analysis of the 16S rRNA gene amplicons was performed using Quantitative Insights into Microbial Ecology (QIIME) 2 software in its 2023.9 version (115). After importing the paired-end raw reads, primers were removed with *q2 cutadapt* (195), with the option to discard any sequence not containing the primers. Reads were quality filtered, denoised, paired-end merged, sorted into Amplicon Sequence Variants (ASVs) and chimera cleaned with DADA2 (116) used as a qiime2 plugin. Additionally, low quality positions at the 3' end of the reads were removed. Three extra filtering steps were applied after DADA2. The first one, to remove non-prokaryotic sequences with *q2 quality control* (117) by discarding all ASVs that did not match the Greengenes 13_8 database (118) clustered at 88% identity, after aligning with VSEARCH (119) under permissive parameters (65% identity and 50% query coverage). Taxonomy was assigned to the remaining ASVs using a scikit-learn naïve Bayes classifier (120), previously trained against

V3-V4 16S rRNA gene region to increase its accuracy (165) using the same Greengenes 13_8 database (clustered at 99% identity). A second filter was applied to discard all sequences classified as *Archaea*, *Chloroplast* or *Mitochondria*. Thirdly, all ASVs present in the negative control sample were removed from the analysis (21 ASVs; as previously done (229)) using ID-based filtering. The phylogenetic tree was built using FastTree (123) after aligning the remaining ASVs with MAFFT (121). Functional genes present in the predicted metagenome were inferred with PICRUST2 (106) and mapped to KEGG database (230) modules.

The diversity analysis was done at a common depth of 76,253, corresponding to the lowest sampling depth using *q2 diversity*. The alpha diversity of the samples was measured with Shannon (131) and Chao1 (130) indexes and the beta diversity with Jaccard (134) and Bray-Curtis (133) dissimilarity indexes (qualitative and quantitative, respectively). We tested linear correlations between the microbiota diversity and the antibody levels using Spearman correlation coefficient (231). The correlation of the vaccine response (antibody delta) with alpha diversity indexes was calculated with *cor.test* function in Stats R package (232) version 4.3.1. Correlation between beta diversity and vaccine response was computed with Mantel test (999 permutations) included in *q2 diversity beta-correlation* (233). Correlations between taxa/functional modules and antibody response were inferred using Maaslin2 R package (234) filtering taxa with less than 0.01 abundance and the rest of the parameters set to default, using abundances normalized relatively per sample as input (Total Sum Scaling). After Benjamini-Hochberg correction, significance was considered when $q < 0.05$.

In the discrete comparative analysis (good against bad responders), alpha diversity differences between groups were estimated by Kruskal-Wallis tests (999 random permutations) using *q2 diversity alpha-group significance* (126). In the beta diversity discrete analysis, the Principal Coordinate Analysis (PCoA) was computed with *q2 diversity core-metrics* (164,196), which was

visualized in R with qiime2R package (200). The significance of beta diversity comparisons was tested by PERMANOVA pairwise tests (999 permutations), and the percentage of explanation of the studied variables using Adonis function from Vegan R software package (128), using q2 diversity beta-group-significance (127). For all mentioned tests, *P* values lower than 0.05 were considered significative.

The healthy swine nasal core microbiota was obtained by filtering all ASVs from the most prevalent genera in healthy pigs as described previously (33). Briefly, all ASVs not classified within the stated genera were filtered out using *q2 feature-table* filtering options. The prevalence threshold to consider a specific genus as core-microbiota was set to 80%. To find taxa differentially abundant on the remaining nasal core-microbiota between good and bad responders we used Lefse (235) under its default parameters to perform a linear siscriminant analysis (LDA). Taxa showing a LDA score > 2 between the two groups were considered as significantly different.

R script language (version 4.2.2) was used in RStudio environment (version 2022.07.0) (199) to process Qiime2 microbiome generated data as well as to generate plots using qiime2r, ggplot2 (202), tidyverse (203) and reshape2 (201) packages.

Statistical modelling

Prior to undertaking statistical analysis, microbiota composition at order level was screened for unlikely or missing values. No data were excluded on this basis. Subsequently, a descriptive statistical analysis was carried out to the nasal and rectal microbiota composition based on the relative abundance of ASVs for both DNA and RNA. We ran two different statistical models of nasal microbiota including a multivariable logistic regression model with the variation (ΔAb) in vaccine response into good (better response) or bad (worse response) responders as the outcome variable and a multivariable linear

regression with the actual value of the ΔAb in vaccine response as the continuous outcome variable. This has been applied for both microbiota composition based on nasal and rectal samples. Initially, a univariable model was carried out to test the unconditional associations between dependent and various independent variables (relative abundances of members of the nasal microbiota). Only independent variables with $P \leq 0.25$ in this initial screening were included in multivariable logistic and linear regression models in accordance with Dohoo *et al.* (236). To account for the farm variations, an additional model was developed accounting for farm as a random effect for each of nasal and rectal DNA using generalized mixed models. For each model, the significant independent variables from the univariable analysis were then offered to a multivariable model and a manual backward elimination was implemented, to obtain a final model that exclusively included variables with a P value < 0.05, considered as significant. The P value and the regression coefficient (b) with a 95% confidence interval (95% CI) were reported for each variable. In a similar approach, we developed a model following the same analysis for relative abundances of members of the rectal microbiota. In all statistical analyses, the results were regarded as significant at $P \leq 0.05$. All statistical analyses were conducted using the R version 3.3.3 software.

Results

The increase in antibody levels correlated with increase in nasal microbiota diversity

To examine the response to vaccine administration or to natural exposure to *G. parasuis*, we determined the antibody levels from the piglets before and after vaccination or in equivalent times for those non-vaccinated but naturally exposed to the pathobiont, which was detected by PCR in the nasal samples (**Table 1**). All piglets showed maternal derived antibodies, although at different levels (farm 2 showed higher maternal antibodies due to sow

vaccination). Since the levels of antibodies were diverse at both tested timepoints, the dynamics of the antibody levels (response) were measured using the difference between both values in each animal (delta antibody value, ΔAb). This value was used to preliminary classify the piglets into good or bad responders, i.e., piglets showing variations above the median of the group or below the median, respectively (see **Methods**, **Supplementary Figure 1 and Table 1**).

Table 1. Level of antibodies against *G. parasuis* (shown as A_{450nm} from ELISA) and presence of virulent (Vir) and non-virulent (Nvir) *G. parasuis* by PCR.

Farm	Vaccina-tion	Antibodies at 1-3 weeks of age	Antibodies at 8-9 weeks of age	Vir PCR	Nvir PCR	Variation (ΔAb)*	Resp-onse**
Farm 1	Weeks 3 and 6	0.265	0.65	-	+	0.385	Bad
		0.699	0.682	-	+	-0.017	Bad
		0.62	1.55	-	+	0.93	Good
		0.837	1.619	-	+	0.782	Good
		0.859	0.933	+	+	0.074	Bad
		0.54	1.496	-	+	0.956	Good
		1.511	0.959	-	+	-0.552	Bad
		0.757	1.547	-	+	0.79	Good
Farm 2	Weeks 2 and 5	2.083	0.906	-	+	-1.177	Bad
		2.474	0.719	-	+	-1.755	Bad
		2.243	0.835	-	-	-1.408	Bad
		1.927	1.913	+	+	-0.014	Good
		1.197	0.605	-	+	-0.592	Bad
		0.656	1.715	-	+	1.059	Good
		1.061	1.138	-	+	0.077	Good
		2.196	1.069	-	+	-1.127	Bad
		1.56	0.275	-	+	-1.285	Bad
		0.448	0.623	+	+	0.175	Good
		0.429	0.322	+	+	-0.107	Bad

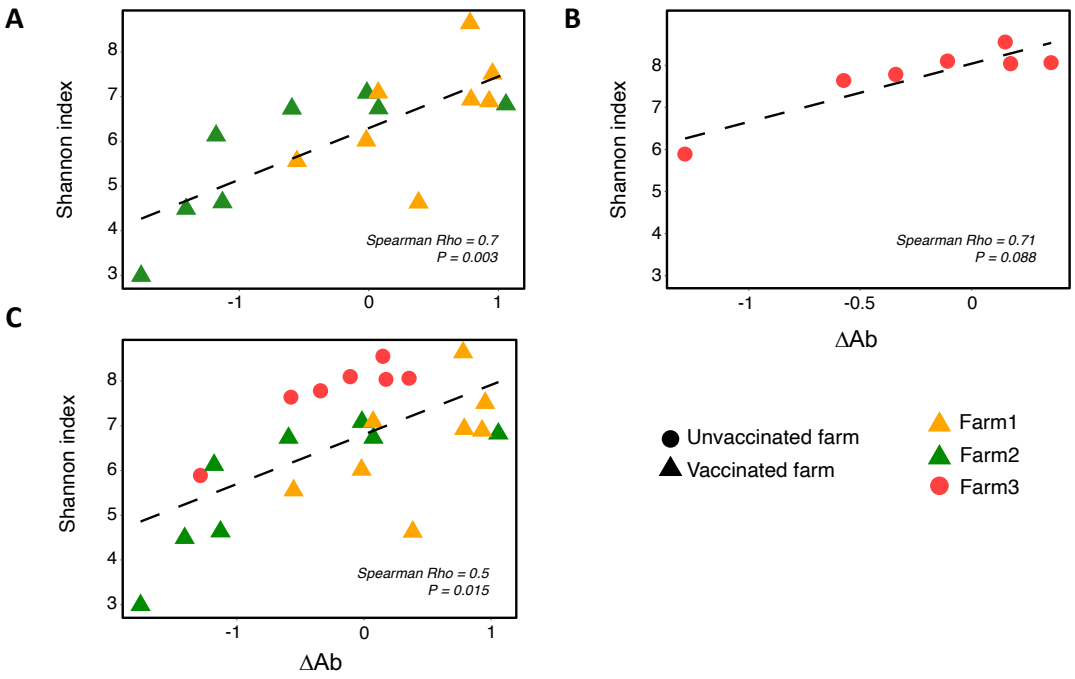
Farm 3	Unvaccinated	1.006	0.433	+	+	-0.573	Bad
		0.363	0.514	+	+	0.151	Good
		0.325	0.682	+	+	0.357	Good
		N.A.	0.42	+	+	N.A.	Bad
		0.461	0.121	-	+	-0.34	Bad

* The variation between the level of antibodies at the two timepoints (ΔAb).

** A preliminary classification into good or bad antibody responders (see **Methods**).

N.A.: Not available

The association between the nasal microbiota composition and the response to vaccination was initially evaluated in samples from the two vaccinated farms. A strong positive correlation of the alpha diversity estimated through Shannon index was observed with the antibody variation levels, ΔAb (Spearman Rho = 0.7, *P* = 0.003, **Figure 1A**). A similar but more moderate tendency was observed for Chao1 index (Spearman Rho = 0.48, *P* = 0.06), indicating that piglets with a more diverse nasal microbiota at vaccination time,



responded better to vaccination. Regarding the beta diversity analysis, weak and moderate correlations were detected between the qualitative and quantitative distance matrices and the ΔAb (Jaccard and Bray-Curtis Spearman Rho = 0.29 and 0.42, respectively; Mantel test $P < 0.007$), showing that larger community differences between samples (primarily quantitative) were associated with greater differences in the ΔAb.

Figure 1. Correlation between nasal microbiota alpha diversity and antibody response. Spearman correlation between alpha diversity of the nasal microbiota (measured by Shannon index) and the antibody response (ΔAb, measured as the difference between the level of antibodies after and before vaccination or equivalent times in non-vaccinated piglets) are shown A) in both farms vaccinated against G. parasuis (yellow and green triangles), B) in the unvaccinated farm (red spheres) and C) in the three farms together. Each tendency line depicted in the graphs (dashed lines) was generated using geom_smooth function (ggplot2) using linear model (lm) as the method.

To examine if the influence of the nasal microbiota was specific for vaccinated animals, we included in the analysis unvaccinated animals with natural exposure to *G. parasuis* (farm 3), since strains of this bacterium, including virulent ones, can colonize the nasal cavity of healthy piglets without causing clinical disease. In fact, virulent and non- virulent strains of *G. parasuis* were detected in piglets from farm 3 by specific PCR. Although we could not detect a significant correlation between the ΔAb and the alpha diversity (Shannon index), probably due to the low number of samples, the tendency observed was similar to that detected in vaccinated piglets, *i.e.* animals with more diverse nasal microbiota showed increased levels of specific antibodies against *G. parasuis* (Spearman Rho = 0.71, $P = 0.088$, **Figure 1B**). When the three farms were analyzed together, the Shannon index correlated positively with the ΔAb (Spearman Rho = 0.5, $P = 0.015$, **Figure 1C**), while the Chao1 index showed a weaker tendency (Spearman Rho = 0.38, $P = 0.07$). These findings support that animals with more diverse nasal microbiotas tend to exhibit higher level of antibodies against *G. parasuis* even when they are not

vaccinated and naturally encounter the pathobiont. There was a weak correlation between the beta diversity distance matrix and the antibody variation under both assessed metrics, possibly because of the introduction of another farm and therefore, more variation (Jaccard and Bray-Curtis Spearman Rho = 0.27 and 0.28, respectively; Mantel test $P < 0.004$).

Nasal microbiota taxa associated with response to G. parasuis

To identify nasal microbiota members associated with the antibody response, we investigated the correlations between the taxa and the ΔAb in the three farms. First, we examined if the response to *G. parasuis* correlated with the abundance of this bacterium in the nasal microbiota of the piglets and found no association (**Supplementary Figure 2**). When we analyzed the global microbiota composition, 6 orders and 22 genera showed to correlate with the ΔAb, most of them positively. *Clostridiales* and *Bacteroidales* were the most abundant orders identified with positive correlation with the ΔAb ($q < 0.05$, **Figure 2**). Other orders were also positively correlated, namely *Enterobacteriales*, *Bacillales* and *Fusobacteriales* ($q < 0.05$, **Figure 2 and Supplementary Table 1**), while the only order that correlated negatively with the ΔAb was *Pseudomonadales* ($q = 0.043$). Correlations at genus level also followed the same dynamics observed for the corresponding orders (**Supplementary Figure 3** and **Supplementary Table 1**), such as *Prevotella* (*Bacteroidales*) and several members of *Clostridiaceae*, *Lachnospiraaceae*, *Ruminococcaceae* and *Veillonellaceae* families (belonging to order *Clostridiales*), which correlated positively with the delta antibody. Other genera with a positive correlation with the antibody variation were *Klebsiella*, *Staphylococcus, Fusobacterium*, *Pasteurellaceae* (uncl.), *Escherichia*, *Moraxella,* and *Corynebacterium*. Within *Pseudomonadales*, only a divergent *Moraxella* (originally classified as *Enhydrobacter* in the used database but confirmed as *Moraxella* by BLASTn) negatively correlated with ΔAb (see

Supplementary Table 1 for all correlations). To further investigate whether these taxa associations with the antibody response might also be reflected in the microbiota functionality, the association between the inferred metagenome functional modules and the ΔAb was also evaluated. Three modules were found to be positively associated with ΔAb ($q < 0.05$), i.e., iron complex transport system, glycolysis, and simple sugar transport system (**Supplementary Table 1**).

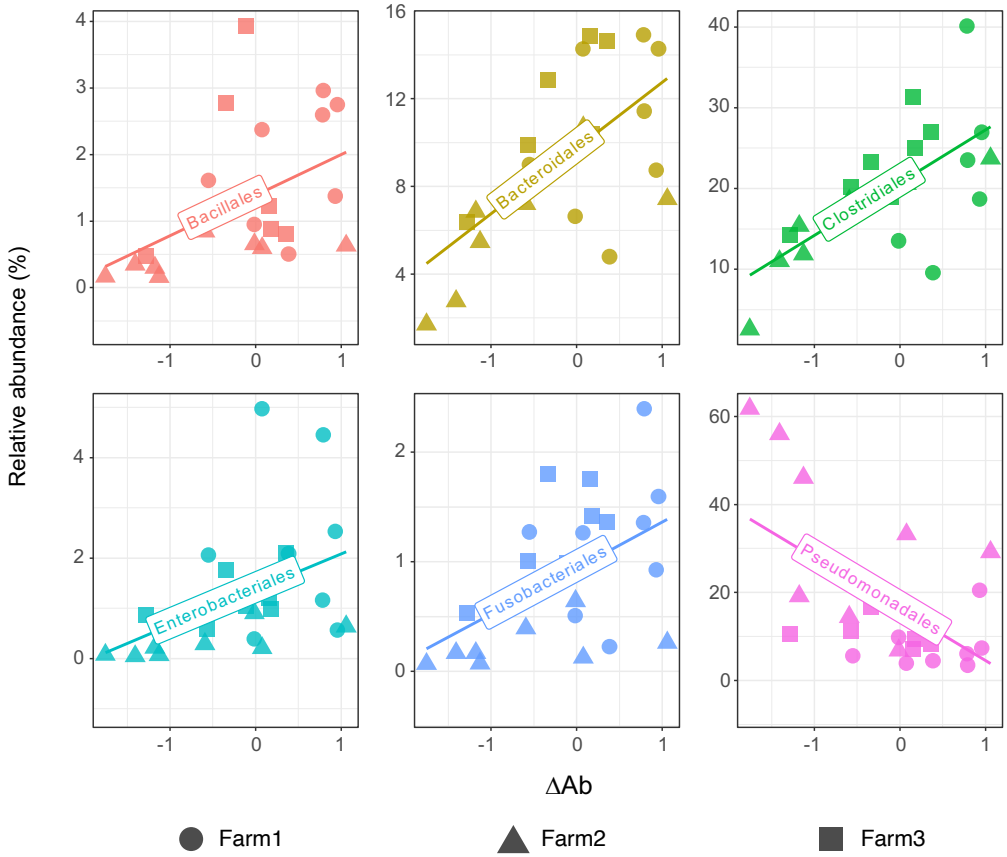


Figure 2. Correlations with the antibody response and nasal microbiota at order level. Scatter plots show the relative abundance versus ΔAb, measured as the difference between the level of antibodies after and before vaccination or equivalent times in non-vaccinated piglets. Only orders found significant with Maaslin2 are shown. Each tendency line depicted in the graphs was generated using geom_smooth function (ggplot2) using linear model (lm) as the method.

Some of the previous associations at order level were also identified using multivariable linear regression with the actual value of the antibody variation (ΔAb) as a continuous outcome variable, including two models with and without farm as a random effect which results are shown on **Table 2**. Again, the presence of *Bacteroidales*, *Clostridiales,* together with *Spirochaetales* in the nasal mucosa of piglets was positively associated with the antibody response to *G. parasuis*. Whereas the presence of *Pseudomonadales* in the nasal mucosa of piglets was negatively associated with the antibody response to *G. parasuis*.

Table 2: Nasal microbiota members associated with the variation in antibody response (measured as a continuous outcome variable) to *G. parasuis* (*P* ≤ 0.05) in 23 animals in 3 farms using a mixed effects generalized linear regression model.

Variable	**Coefficient Estimate**	**95% CI**	***P* value***
Pseudomonadales	-1.77	(-4.02) – 0.48	0.016
Bacteroidales	11.71	(-12.99) – 36.41	0.005
Enterobacteriales	18.27	(-26.36) – 62.91	0.845
Clostridiales	5.03	(-6.52) – 16.57	0.028
Erysipelotrichales	-216.01	(-468.99) – 36.98	0.518
Fusobacteriales	89.13	(-31.95) – 210.21	0.272
Spirochaetales	161.34	(-79.71) – 402.39	0.037
Desulfovibrionales	-229.34	(-591.80) – 133.13	0.172
Coriobacteriales	-116.41	(-427.39) – 194.59	0.099
Rhizobiales	-195.29	(-491.74) – 101.16	0.087

*Taxa with *P* > 0.05 were kept in the final model due to significant confounding effect with other significantly associated taxa

Rectal microbiota diversity also correlated with antibody response

Different microbiotas can crosstalk with the immune system and have a systemic effect. Since previous studies have highlighted the role of the gut microbiota with the antibody response, we aimed to analyze whether similar associations to those detected between the nasal microbiota and the antibody response were also occurring in the rectal microbiota. The alpha diversity of the rectal microbiota of animals from the three farms moderately correlated with the antibody response (Spearman Rho = 0.44 and 0.6; *P* = 0.036 and 0.003 for Chao1 and Shannon indexes, respectively**, supplementary Figure 4**). The beta diversity of the rectal microbiota was weakly correlated with the ΔAb (Jaccard and Bray-Curtis Spearman Rho = 0.21; Mantel test *P* < 0.006). No significant correlations were found between taxa in the rectal samples and the ΔAb using Maaslin2, except for a *Mogibacteriaceae* unclassified genus (*Clostridiales*), positively associated with the ΔAb (q = 0.021). Nevertheless, four orders were identified as significantly correlating with the delta antibody in both multivariable linear regression models using the ΔAb as a continuous outcome variable (accounting for farm variations as a random effect or not). The presence of *GMD14H09* and *Clostridiales* in the rectal microbiota of piglets was associated with better antibody response to *G. parasuis,* whereas the presence of *Pasteurellales* and *Enterobacteriales* was negatively associated with the antibody response to *G. parasuis* (**Supplementary Figure 5 and Table 3**). The farm effect was negligible.

Table 3: Rectal microbiota members associated with the variation in antibody response (as a continuous outcome variable) to *G. parasuis* (*P* ≤ 0.05) in 23 animals in 3 farms using a mixed effects generalized linear regression model.

Variable	Coefficient Estimate	95% CI	*P* value
Clostridiales	0.301	(-2.327) – 2.929	0.0075
GMD14H09	76.900	2.731 – 151.069	0.036
Pasteurellales	-42.895	(-84.076) – (-1.715)	0.036
Enterobacteriales	-4.045	(-7.247) – (-0.842)	0.016

Microbiota associations with piglets classified as good antibody responders

To further investigate and validate the associations with a good antibody response, we used the preliminary classification according to the ΔAb (see **Methods**, **Supplementary Figure 1 and Table 1**).

Using this classification, a more diverse and richer microbiota was observed in the good responders compared to the bad ones (Shannon and Chao1 indexes, $P < 0.05$; **Figure 3A**). In the beta diversity analysis, no significant differences were detected between good and bad responders. A strong environmental effect caused by the analysis of three different farms together was observed in the clustering of the samples, quantified to be 29% qualitatively and 48% quantitatively (Adonis test R^2 using Jaccard and Bray-Curtis indexes, respectively, $P = 0.001$**, Figure 3B**). However, the differences between responders were still non-significant when this effect was evaluated as a nested variable after considering farm as the main effect.

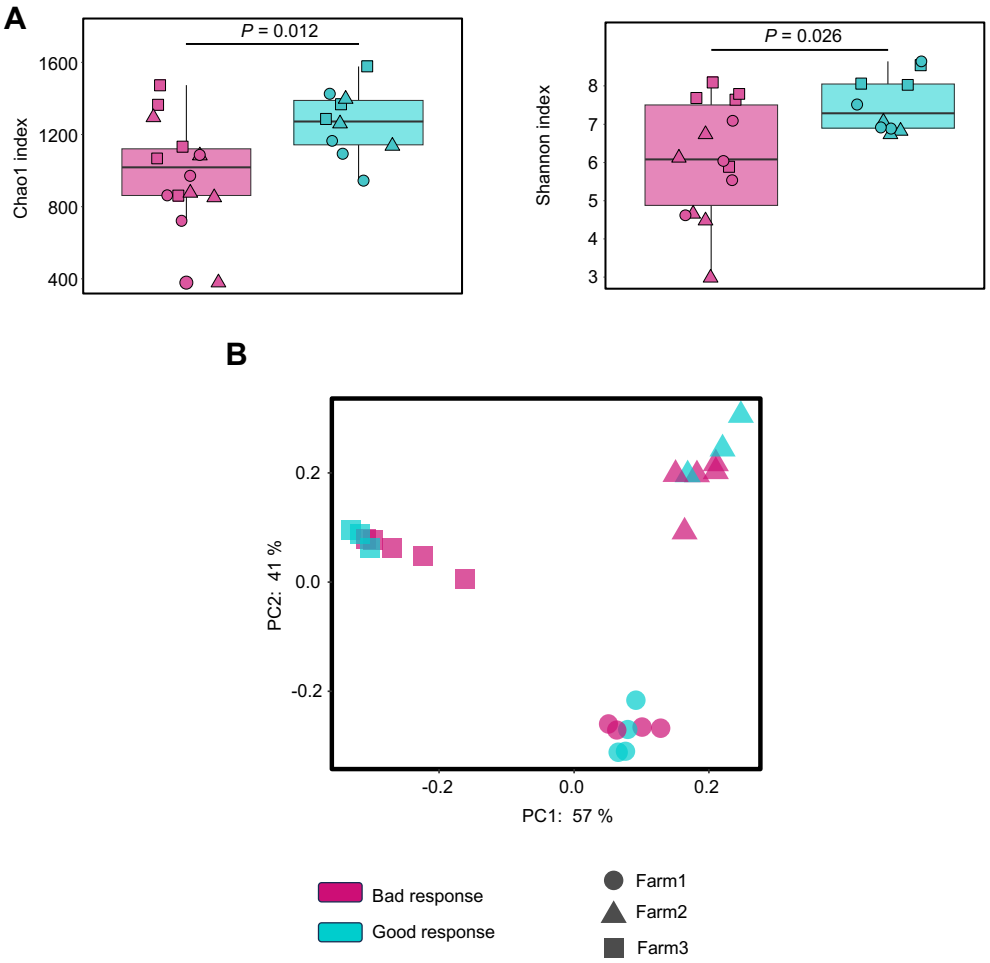


Figure 3. Alpha and beta diversity of nasal microbiota from good (turquoise) and bad (mauve) responders. A) Alpha diversity measured by Chao1 and Shannon indexes. B) Beta diversity estimated through Bray-Curtis dissimilarity index. The shapes indicate different farms.

With the aim to reduce the effect of the variability among farms and knowing that is frequent to detect transient environmental taxa in nasal microbiota samples, the ASVs from genera not included in the core of the swine nasal microbiota were filtered out (see **Methods**). Interestingly, the significance of the alpha diversity comparison between good and bad responders increased (Shannon and Chao1 $P < 0.009$) indicating that differences observed were not caused by transient taxa and truly relied on common swine nasal colonizers. To reveal nasal taxa that were associated with good or bad responders to *G. parasuis* beyond the farm environmental effect, the filtered microbiota composition from all samples was submitted to linear discriminant analysis (Lefse) at different taxonomic levels (**Figure 4 and Supplementary Figure 6**). In nasal samples, 4 orders and 20 genera from the core of the nasal microbiota were found to significantly discriminate good responders (LDA score > 2), independently of the farm of origin. Among the detected taxa, *Clostridiales* and *Bacteroidales* were also found to be associated with good response in the previous analysis using the numeric antibody delta, while *Erysipelotrichales* was only detected in this discrete analysis (Lefse). Most genera associated with good vaccination response belonged to the above orders, with several genera of the *Ruminococcaceae* and *Lachnospiraceae* families (among other *Clostridiales*), as well as *Prevotella* (*Bacteroidales*). *Neisseria* was also associated with the group with better response. No taxa with a LDA score > 2 were associated with a bad response.

Regarding the rectal microbiota, no differences were detected in alpha, beta diversity or taxonomical composition when good and bad responders were compared.

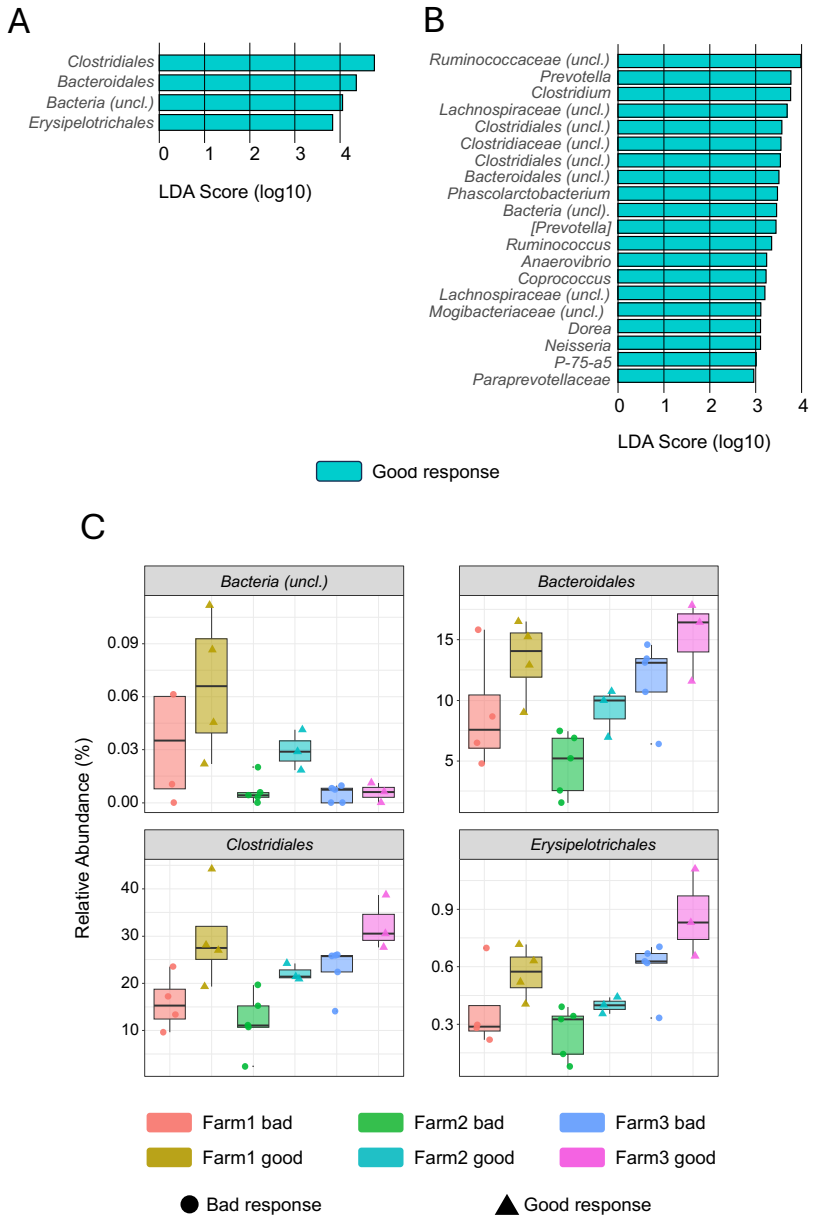


Figure 4. Nasal microbiota taxa discriminating good responders. A) At order level. B) At genus level. C) Relative abundance of the orders identified as associated with good responders (A) in the nasal microbiota. The analysis was done using Lefse. No taxa discriminating bad responders were found.

Discussion

In this study, we evaluated the swine nasal and rectal microbiota composition according to the response to vaccination against *G. parasuis* or natural exposure to this bacterium. Despite all animals in this study being healthy and, therefore, no major changes were expected in their microbiota composition, several differences were identified according to the antibody response. To our knowledge, this is the first study that correlates the porcine nasal microbiota composition as a predictor for the response to vaccination providing some notable findings that highlight the relationship between the diversity of the microbiota and the immune response. Moreover, the introduction of an unvaccinated control farm emphasized that these claims could also stand for animals that naturally encounter pathogens.

The initial levels of antibodies were variable between farms, but also within the same farm indicating variable level of maternal immunity at the initial sampling time of the piglets (1-3 weeks of age). Nevertheless, independently of the starting level of antibodies in each farm, piglets exhibited variable dynamics in the levels of antibodies. To analyze the antibody response, and also deal with variability between animals, we decided to measure the response to *G. parasuis* using the difference between the two time points i.e., after and before vaccination or in equivalent times for the unvaccinated farm (delta antibody value, or ΔAb).

In agreement with previous reports (213,218), piglets in this study with more diverse microbiota showed higher antibody response. In the case of nasal samples, these differences were clearer when transient taxa were removed from the analysis (by keeping the healthy core microbiota), highlighting that the major drivers were swine nasal commensals. Moreover, the fact that stronger correlations were observed when alpha diversity was measured with Shannon index rather than Chao1 suggests that not only species richness but also community evenness are important for better immune responses. We

found three metabolic modules inferred from the microbiota composition that were positively correlated with the antibody response, that can possibly be related with an increased alpha diversity (these three modules may be associated with more bacterial activity). Several investigations targeting the swine microbiota in different scenarios also reported higher alpha diversity in healthy groups compared to different diseased animals or environments (9,16,17) and associated it with a better immune status of the animals. The well-known notion that the microbiota participates in the maturation of the immune system (212) also agrees with these results. Contrarily, other studies on gut microbiota did not report any relationship between alpha diversity and response to vaccination (216,221,223,224,226,227). This apparent contradiction may be explained by different factors, such as the time when the microbiota was evaluated, the type of vaccine or immune stimulus or the host studied, which all can affect the output of the analyses.

Since samples came from three different environments (farms), we did not focus the analysis on ASVs but searched for associations at higher taxonomic levels (order and genus). We detected several taxa in the nasal cavities of the piglets that positively correlated with a higher increase in the level of antibodies. Among these, taxa frequently found in the gut microbiota (*Bacteroidales* and *Clostridiales*) appeared as the most associated. It is not unusual to find these taxa in the swine respiratory tract (26,229), most are included in the nasal core, and their presence has been recently discussed in a previous study from our group that confirms these bacteria are indeed present in this body site (229). Nevertheless, the role of these gut microbiota-related taxa in the upper respiratory tract is still uncertain (13,229), and future studies focusing on these microbes may help understanding their connection with the immune response. Interestingly, genera within these two orders appeared in higher abundances in farms without respiratory disease condition compared to farms with Glässer's disease (16), as well as when compared with farms with polyserositis caused by *Mycoplasma hyorhinis* (17). Besides

these, other taxa frequently found in the swine nasal microbiota, such as *Fusobacterium*, *Staphylococcus*, *Pasteurellaceae* (uncl.), *Moraxella* and *Neisseria* (26) were positively associated with the increase of antibodies against *G. parasuis*. The presence of these taxa may be important for the immune system stimulation and therefore, for the immune response. *Moraxella* has already been proved as a potential nasal probiotic within a five-bacteria cocktail (20). However, very few nasal colonizers are being considered for the moment as probiotics to enhance the immune response (237), which deserve further investigation. Regarding taxa negatively correlated with the antibody response, only a different group from *Moraxella* genus negatively correlated with *Glaesserella* antibody response. In a previous study (16), higher abundances of *Moraxellaceae* (*Moraxella* and *Enhydrobacter*) were found in farms with Glässer's disease compared to control farms. The variability within the genus *Moraxella* may explain the different role in health status depending on the species or even the specific strain, as it has been previously observed regarding the virulence of these bacteria (238). Similarly, in gut microbiota, immune stimulation by probiotics is highly influenced by variables like the specific strain, host or environment (also explaining differences across farms in this study) (239).

In the case of rectal samples, fewer taxa were identified as related with the response to vaccination, possibly because this microbiota can be more resilient to exposure to environmental taxa compared to the nasal microbiota. Two gut-microbiota core orders (*Clostridiales* and *Enterobacteriales*) (240) were identified as positively and negatively correlated with the antibody response, respectively. Munyaka, *et al*. showed that in vaccination against *Mycoplasma hyopneumoniae*, Operational Taxonomic Units (OTUs) from *Bacteroidales* (mainly *Prevotella*) found in the fecal microbiota at the moment of vaccination were positively correlated with higher antibody titters (226,227). OTUs classified within *Clostridiales* were associated with both high and low responses, while no associations within *Enterobacteriales* were found. The

fact that in these studies samples came from the same environment (farm) and the analyses were performed at OTU level could explain the differences with this study (three farms and associations evaluated at higher taxonomical levels). On the other hand and in agreement with our results, studies on the role of the human microbiota also associated several taxa within *Clostridiales* with better immune responses (213,217,218,223), possibly through the production of short-chain fatty acids that stimulate the immune system (213,241). However, other studies found taxa within *Clostridiales* associated with immune responses below average (224), suggesting differences in the role that taxa within this order may have. Aligned with our findings, higher abundances of *Enterobacteriaceae* (221) and *Enterobacteriales* (219) were associated with lower responses to vaccination. In any case, although bacteria within *Clostridiales* are frequently associated with health, while some *Enterobacteriales* can be associated with disease (242), both orders contain taxa that contribute metabolically to immune system homeostasis and stimulation (216,243). Two lower abundance orders (GMD14H09 and *Pasteurellales*) were also identified in this study as positively and negatively associated with the antibody response, respectively; but their presence and role are unclear in this microbiota. Finally, despite not finding any associations within *Bacteroidetes* and *Actinobacteria*, such as *Prevotella* and *Bifidobacterium*, these orders may play a role in immune response in humans and pigs (216,218–220,226,227), and also deserve further study.

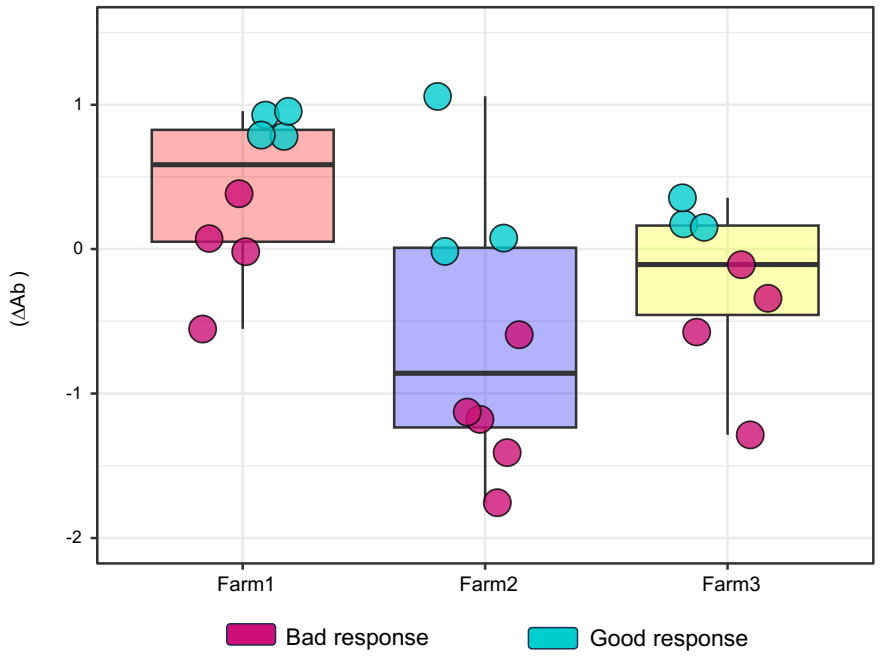
There are some limitations to be considered in this study. The main limitation is the variation given by the three farms, which introduced substantial environmental variability complicating the analysis. Similarly, there may be a certain caveat in comparing the antibody responses between animals from different farms (variable antibody levels and dynamics, maternal immunity, etc.). To address these issues, we classified the piglets as good or bad responders within each farm using a fixed criterion. The discrete analysis using this classification into good or bad responders agreed with the results

obtained in the numeric correlation analysis between the microbiota composition and the level of antibodies, which otherwise proved to be a fruitful strategy with higher statistical power, as expected. Also, in order to diminish the limitation of not accounting for collinearity in the regression models, the variables were examined for the possibility of interaction and confounding effects. Nevertheless, other statistical models could be explored to improve the analysis of these high-dimensional and compositional data. Finally, we analyzed the microbiota composition before weaning, when it may not be yet stable, giving rise to further variability. However, the fact that we obtained significant results using these samples can also be considered a strength, as the associations found can be considered more representative since they were observed in three different farms with independent management.

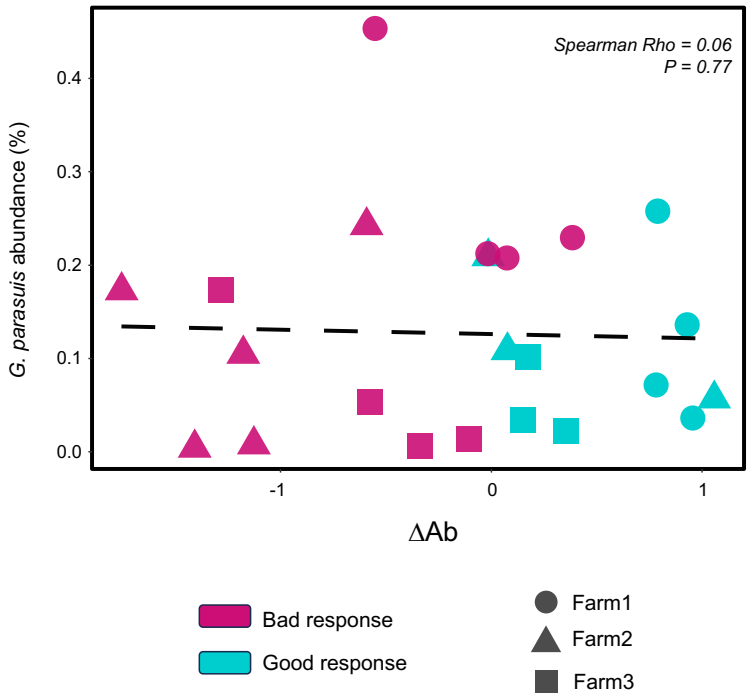
In conclusion, this study highlights the importance of the diversity of the nasal and rectal microbiotas in the antibody response when facing an antigen (naturally or in vaccination). The associations of several bacterial species with a better response may serve as a starting point to investigate how these taxa stimulate the immune system from the nasal cavity. Additionally, it may help to design targeted interventions to enhance immunization and protection in animals.

Supplementary materials

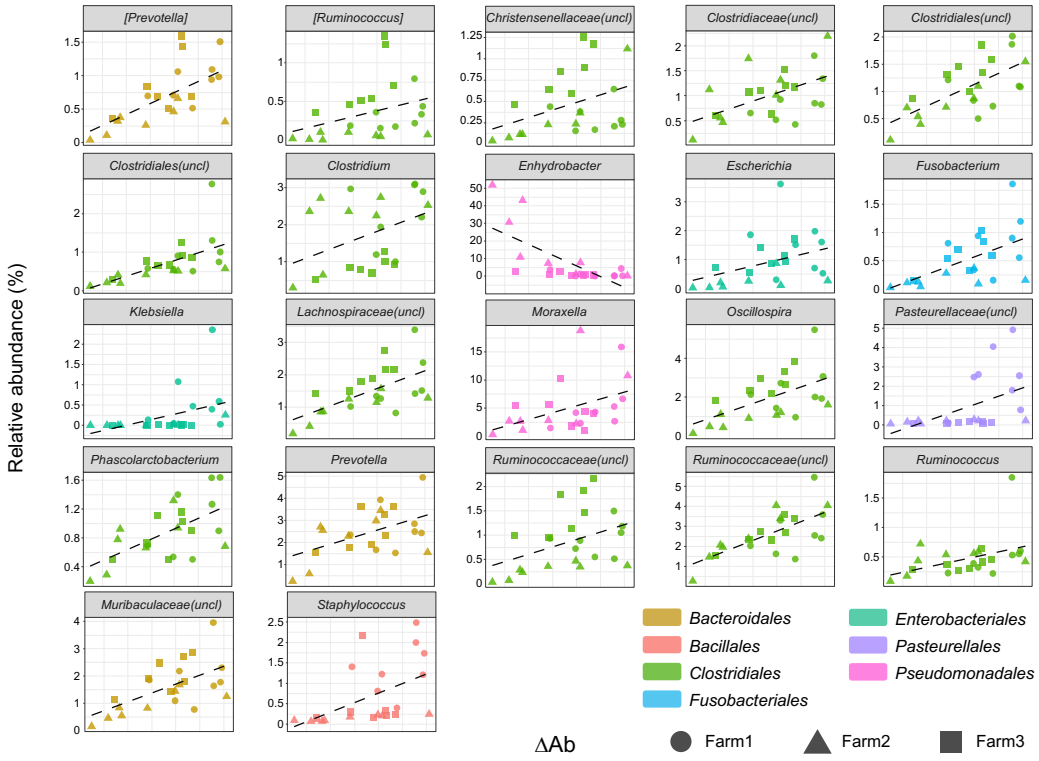
Supplementary figures



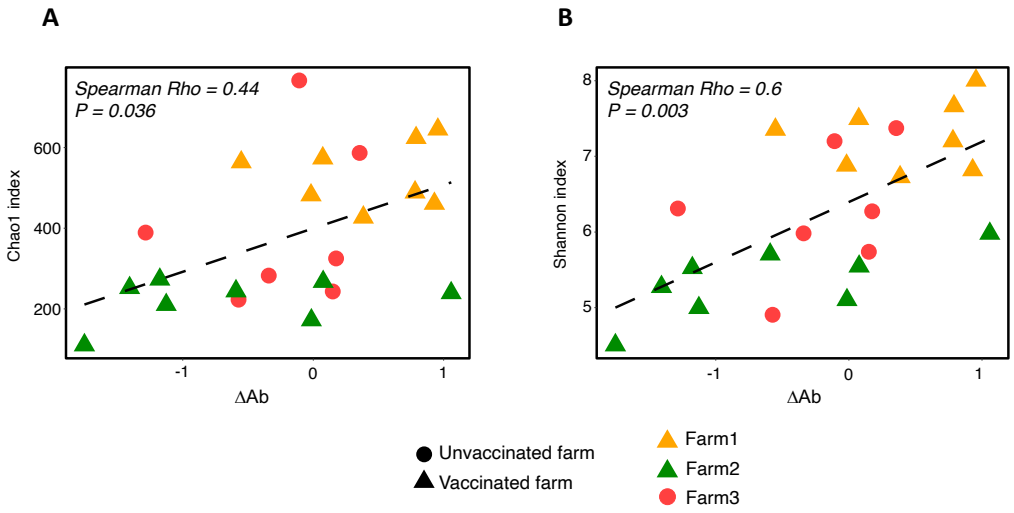
Supplementary Figure 1. Antibody variation of the samples under study (DAb). The antibody variation is measured as the difference between the level of antibodies after and before vaccination or equivalent times in non-vaccinated piglets) shown per farm. Classification into good (turquoise) or bad (mauve) responders is also provided.



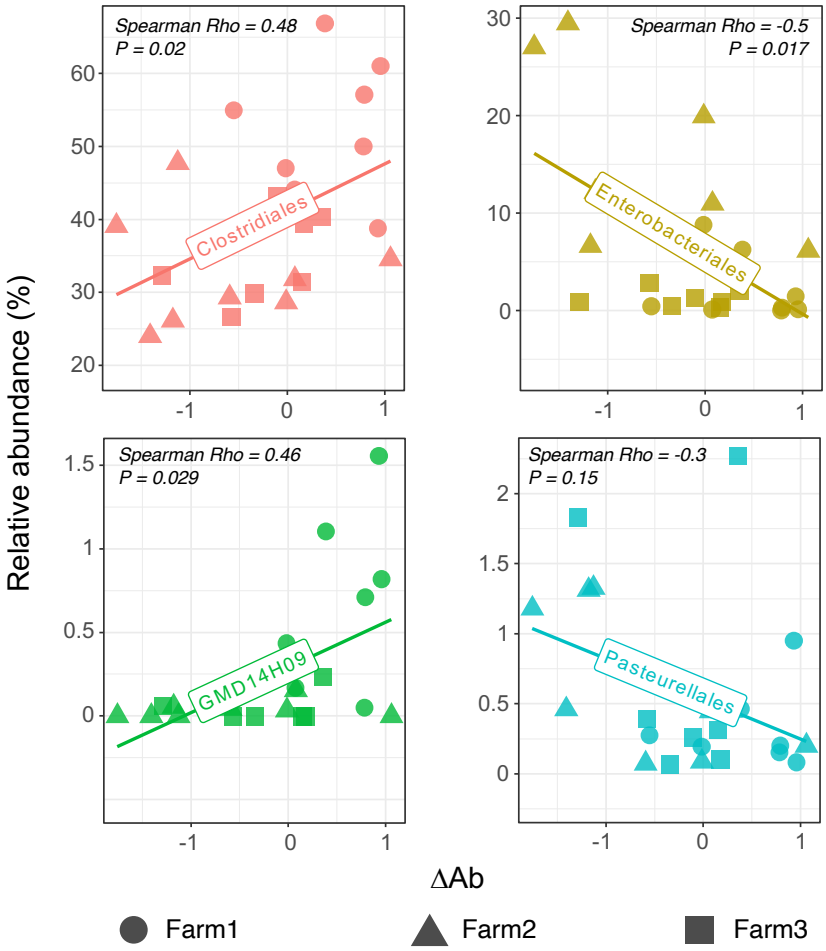
Supplementary Figure 2. Correlation between *G. parasuis* and antibody response. Spearman correlation between *G. parasuis* relative abundance (all ASVs collapsed) and antibody response (DAb measured as the difference between the level of antibodies after and before vaccination or equivalent times in non-vaccinated piglets) are shown for nasal microbiota samples from the three farms. The tendency line depicted in the graph (dashed line) was generated using geom_smooth function (ggplot2) using linear model (lm) as the method.



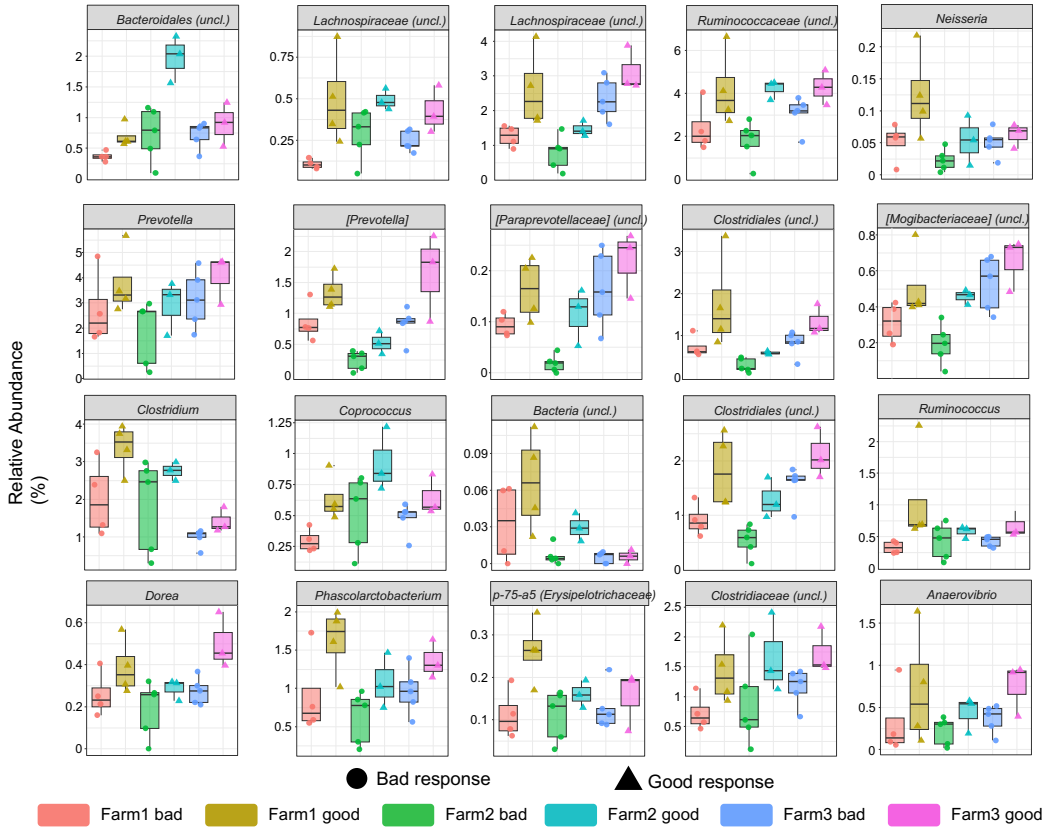
Supplementary Figure 3. Correlations with the antibody response and nasal microbiota at genus level. Scatter plots show the relative abundance versus DAb, measured as the difference between the level of antibodies after and before vaccination or equivalent times in non-vaccinated piglets. Only genera found significant with Maaslin2 are shown. Each tendency line depicted in the graphs (dashed lines) was generated using geom_smooth function (ggplot2) using linear model (lm) as the method. Color code corresponds to orders englobing these genera.



Supplementary Figure 4. Correlation between rectal microbiota alpha diversity and antibody response. Spearman correlation and the tendency as a linear model (dashed lines) between alpha diversity in the rectal microbiota samples are shown, measured by Chao1 (A) or Shannon (B) indexes, and the antibody response (measured as the difference between the level of antibodies after and before vaccination or equivalent times in non-vaccinated piglets, or DAb). Each tendency line depicted in the graphs (dashed lines) was generated using geom_smooth function (ggplot2) using linear model (lm) as the method.



Supplementary Figure 5. Correlations with the antibody response and rectal microbiota at order level. Scatter plots show the relative abundance versus DAb, measured as the difference between the level of antibodies after and before vaccination or equivalent times in non-vaccinated piglets. Only orders found significant with Maaslin2 are shown. Each tendency line depicted in the graphs was generated using geom_smooth function (ggplot2) using linear model (lm) as the method. Spearman correlations values are also depicted in each plot.



Supplementary Figure 6. Nasal microbiota taxa discriminating good responders at genus level. Relative abundance of the genera identified as associated with good responders in the nasal microbiota in the discrete analysis using Lefse (Figure 4B) are shown.

Supplementary tables

Supplementary tables are available within the publication in:

https://www.nature.com/articles/s41598-025-85867-6

CHAPTER 5

Insights into the *Mycoplasma hyorhinis* pangenome

P Obregon-Gutierrez, J Nogales, C González, E Huerta, A Rubio, M. Domingo, J Segalés, K Kochanowski, AJ Pérez-Pulido, V Aragón, F Correa-Fiz[#], M Sibila[#]

shared last authors

Abstract

*Mycoplasma hyorhinis (*homotypic synonym of *Mesomycoplasma hyorhinis)* is a pathobiont commonly found in the upper respiratory tract of pigs. Under unclear circumstances, this bacterium is able to disseminate systemically and cause disease. Despite some studies have described different infectious capabilities, no factors have been linked specifically to virulent strains. The goal of this study was to analyse the pangenome of all available *M. hyorhinis* strains to identify potential virulence factors.

We isolated strains from the nasal cavity of healthy animals (n=10) as well as from systemic lesions of diseased animals (n=8). The pangenome of 110 *M. hyorhinis* strains was characterized using strains isolated from healthy animals (nasal cavity), diseased animals (systemic organs, nasal cavity and lung), and animals with unknown health status (nasal cavity and lung). Comparative studies were performed according to their different clinical backgrounds and the metabolic capabilities of the strains were also inferred using genome-scale metabolic models.

Although most putative virulence genes were shared across strains, we identified some genes encoding proteins involved in catalytic activity of DNA absent in a cluster of six health-associated strains, such as two genes of the *hsd* restriction/modification system (*hsdM* and *hsdR*) and different helicases. In addition, the variable lipoprotein (*vlp*) genes were studied due to their reported involvement in the infectious process. Although the prevalence of *vlp* genes did not vary between different clinical backgrounds, we found more repetitions in region III of *vlpF* and *vlpC* (tendency) in strains isolated from systemic organs compared with nasal strains. Besides, all strains exhibited comparable reactomes, and consequently, the predicted growth capabilities and auxotrophies were highly conserved.

In conclusion, although all *M. hyorhinis* strains may possess the necessary gene set to cause disease, nasal strains associated with health lacked genes

involved in DNA processing and exhibited different VLP patterns. Future studies including more strains with complete metadata, especially from healthy animals could help clarifying whether these variations are associated with virulence.

Background

Mycoplasma hyorhinis is a colonizer of the upper respiratory tract of healthy pigs. However, under some circumstances, this bacterium is able to escape the host defence barriers and disseminate systemically causing polyserositis (including polyarthritis, pericarditis, pleuritis, peritonitis and meningitis), mainly in nursery pigs (77,78) . This disease is an increasing cause of concern for pig producers worldwide (60,244,245) as compromises the animals' welfare and lead to economically significant expenses derived from the antibiotic treatment costs and the associated productive losses (82). In addition, *M. hyorhinis* has also been associated, although to a lesser extent, with other pathologies such as otitis (83,246), conjunctivitis (84,85) and abortions (86). This *Mycoplasma* is also frequently detected in lower respiratory samples from apparently healthy but also from pneumonic pigs (247,248).

Previous studies detected variable clinical outcomes among *M. hyorhinis* strains, suggesting the existence of different degree of virulence. For instance, Lin *et al*. reported the development of pneumonia after inoculation with clinical isolates but not with a non-clinical one (80), while Gois *et al*. detected different extent of lesions when inoculating intranasally three strains isolated from pneumonic lungs of pigs (81). In agreement, Wang *et al*. (82) observed that the clinical outcome of an experimental inoculation using the strain isolated from the tonsil of an asymptomatic animal from a farm without polyserositis, was significantly milder than the one observed using a strain isolated from lesions of a clinically affected animal. In a previous study performed by our group in a farm with polyserositis, only two of the different *M. hyorhinis* amplicon sequence variants (n=45) detected in the nose of weaning piglets showed a 99-100% of homology with the strain isolated from the lesions (17). This result suggested a role for specific strains in the development of disease and highlighted the need to study the pathogenesis of this bacterium at strain level. Nevertheless, different molecular typing techniques used to

characterize *M. hyorhinis* strains isolated from clinical cases showed no relationship between the site of isolation and/or virulence with specific genotypes (79,249–253).

Although there are several environmental factors that can facilitate *M. hyorhinis* systemic spread, such as stress (254), the nasal microbiota composition at weaning (17) or the presence of other pathogens (255), the underlying intrinsic mechanisms involved in this process remain elusive. This process is probably linked to various factors related to adhesion, nutrient uptake/transportation, homeostasis maintenance, antigenic variability and exoenzymatic activity, as described in other *Mycoplasma* species (256–263). In particular, in *M. hyorhinis*, the surface variable lipoproteins (VLPs) are postulated as an important virulence factor potentially involved in adhesion and antigenic variability, (256,264–267). However, to date no markers have been identified that reliably distinguish virulent from non-virulent *M. hyorhinis* strains.

In this study, we aimed to perform a genomic comparison of *M. hyorhinis* strains according to their clinical background. To this end, we included the genomes of nasal isolates obtained from healthy animals, a source not previously represented in public databases, alongside strains isolated from systemic lesions in clinically affected animals. We carried out a pangenome analysis with these genomes together with those currently available in public databases to identify strain-level differences and candidate virulence markers.

Methods

Isolation of strains from systemic lesions from animals with fibrinous polyserositis or from the nasal cavity of healthy animals

Swabs from systemic lesions from necropsied animals (at Diagnostic Service from Veterinary Faculty of Autonomous University of Barcelona) showing

fibrinous polyserositis lesions (N=8) were taken. These animals came from 7 different farms and were 5-9 weeks of age. Moreover, strains isolated from swabs from the nasal cavity of healthy animals (N=10) coming from 4 different farms with no history of polyserositis in nurseries were also included in the study. All samples were confirmed to be positive to *M. hyorhinis* by isolation in *Modified Friis medium* liquid culture and a specific qPCR (Clavijo et al., 2014). In parallel, the presence of *Glaesserella parasuis* (72) and *Streptococcus suis* (268), two other common pathogens causing polyserositis in nursery pigs was discarded by bacterial isolation and PCR.

For nasal *M. hyorhinis* strains, a further isolation step was performed on agar plates. To do so, 50-100µl of each *M hyorhinis* liquid culture were plated on *Modified Friis medium* agar plates and grown for 1-7 days at 37ºC. Afterwards, one or two *M. hyorhinis* colonies per culture were individually scaled up to 40 ml in the same liquid media. Cultures were centrifuged at 20,817 x *g* for 5 min at room temperature and the pellet was used for DNA extraction with DNeasy UltraClean Microbial Kit (QIAGEN) or Nucleospin Blood (Macherey Nagel), following the manufacturer's instructions. Quality and quantity assessment of the DNA was done using Qubit dsDNA BR Assay Kit (Thermo Fisher) and the Qubit 3.0 Fluorometer (Invitrogen).

Genome sequencing and de novo assembly

DNA samples were sent to Novogene (Cambridge Sequencing Center, UK) for sequencing using Illumina NovaSeq 6000 (WGS) (3 from nasal cavities from healthy pigs and 5 from systemic lesions) and to NanoHealth for Oxford Nanopore Long-read sequencing (all 18 strains). Long read adapters were removed using Porechop (269) and quality and length filtered with chopper (270). Read-quality was assessed with Fastqc (271). Raw reads were assembled *de novo* using Unicycler v0.4.8 (272) for strains sequenced with short and long-reads (hybrid assembly) or Canu v2.2 (273) for long-read only, in both cases under default parameters (just adding the expected genome

size). Long-read assemblies were polished using Medaka v1.11.3 (https://github.com/nanoporetech/medaka), while this tool was used in combination with Polypolish v0.5.0 for hybrid assemblies. The genomes of the strains isolated in this study are available in NCBI BioProject PRJNA1281057.

Genomes included

All publicly available *M. hyorhinis* strains (n=99) and their related information were downloaded from National Centre for Biotechnology Information (NCBI), accessed on November of 2023, using their command line tools (datasets and dataformat) (274) by retrieving any genome classified as "*Mesomycoplasma hyorhinis or Mycoplasma hyorhinis*". Additionally, the genomes of the reference strains of other *Mycoplasma* species infecting pigs (*M. hyopneumoniae*, *M. flocculare and M. hyosynoviae)* or other animal species (*M. conjuctivae, M. ovipneumoniae* and *M. dispar)* were added to the initial phylogeny.

Metadata available for each public genome regarding the isolation site (nasal cavity, systemic organ or lung), animal health status (healthy, diseased or unknown) and country of origin (Spain, Austria, France, Germany or others) was included. All strains isolated from systemic sites were considered to come from diseased animals. Nasal strains came from animals with different clinical backgrounds: i) healthy animals (this study); ii) diseased animals (when systemic isolates were available for the same animal); iii) animals with unknown health status (no available information). Similarly, lung strains also came from diseased or unknown health-status animals. The information of all strains included in this study is provided in **supplementary Table 1**.

Genome annotation and pangenome characterization

Genomes were annotated with Prokka v1.14.6 (275) under the default parameters and using translation with *Mycoplasma* genetic code (translation table 4). After, functional annotation was done with Sma3s v2 (276) using Uniprot 2023.8 database (277).

Roary v3.12.0 (278) was used to infer the *M. hyorhinis* pangenome from Prokka predicted encoding genes using an identity of 90% (using flag -i), meaning that genes with more or equal of this sequence identity were considered in the same group (orthologs) and with the same reference sequence. We used the additional flags to align the core genes (-e -n), do not split paralogs (-s) and use the translation table 4 (-t).

Differences between groups of study based on gene presence/absence were evaluated with PERMANOVA pairwise tests (1000 permutations) on Jaccard distances (134) between groups using Vegan (128) and pairwise Adonis (279) packages in Rstudio (199). Samples from groups with very low number of strains (i.e. countries with two or less strains isolated) were removed for the PERMANOVA analysis. The robustness of the significant results in comparisons with low number of samples (i.e. when only strains from Spain were considered) was verified using also a leave-one-out strategy and a bootstrap resampling by randomly relabelling the sample metadata and compute the same PERMANOVA test for the same number of permutations (1,000) to see if differences were still significant.

Identification of point mutations, variable lipoproteins, resistance genes and other virulence factors

The identification of genes possibly implicated in the virulence of *M. hyorhinis* was done by selecting protein-encoding genes that were annotated into any of the categories compiled in the Virulence Factor Database (VFDB) (280) or in the literature specifying processes and proteins involved in *Mycoplasma* pathogenesis (257–259,281). Putative exoenzymes search was done by

combining the function (i.e. nuclease, protease or kinase) with membrane or extracellular presence. Specific genes mentioned in *Mycoplasma* literature were grouped under the category "*mycoplasma*-associated". The gene prevalence was estimated as the percentage of strains within a group containing the gene. Genes were considered as differential when differences of prevalence (relative frequency) between groups were greater than 0.5 or lower than -0.5. Additionally, SignalP 6.0 (282) was also used to identify lipoprotein or secretory signal peptides in the predicted proteomes of each group (selecting genes present >80% in the strains of each group).

Due to their highly variable and repetitive sequences, variable lipoproteins (VLPs) were studied in newly sequenced strains, taking advantage of the long-read data generated in this study. Putative VLPs were not automatically annotated in the pangenome. Therefore, we extracted all predicted proteins containing the VLP signal (shared consensus starting motif: KKSIFSKKLLVSFGS) obtained by aligning all VLPs from *M. hyorhinis* HUB-1 reference strain from Uniprot (277) using MAFFT (121). The repetitions within the VLP region III (264) were identified in the annotated genomes, quantified and used to distinguish the VLPs. The number of the VLP repetitions was compared using non-parametric Mann-Whitney U test (283).

Resistance genes were searched in the Comprehensive Antibiotic Resistance Database (CARD) (284) and NCBI (285) databases (accessed on December 2024) using blast (117). Only hits with identity > 80%, alignment length > 80% and an E value < 10^{-5} were considered. ResFinder (286) online tool was also used to find resistance genes under default parameters. Ribosomal *RNA 23S* and *16S* subunit genes were extracted from annotated genomes and aligned with MAFFT (121). Single Nucleotide Polymorphisms (SNPs) were identified using snp-sites (287) and processed with vcftools (288), vcfR (289) and VariantAnnotation (290) R packages. Significant SNPs were selected using the allele frequency (>5%). Genes of DNA gyrase (*gyrA* and *gyrB*) and topoisomerase (*parC* and *parE*) were also extracted from annotated genomes

and multiple aligned. To account for the position of the described mutations in *Mycoplasma* (291,292) the same proteins of *Escherichia coli str. K-12 substr. MG1655* were included (available in NCBI, Taxonomy ID: 511145). All alignments were processed and visualized using Aliview (293).

Phylogenetic analysis

An initial phylogeny of all strains (as well as other *Mycoplasma* reference strains) was built from the whole genome sequences using Realphy (294) online tool (https://realphy.unibas.ch/realphy/) under the default parameters using type strain ATCC 17981 (BTS-7) as reference (isolated from the nasal cavity of a pig with atrophic rhinitis (295). The alignment was bootstrapped using Iqtree2 v2.4.0 (296) and visualized using iTOL v7.1 (297). Average Nucleotide Identity (ANI) was calculated using pyani version 0.2.12 (298) under its default parameters.

The phylogeny of the strains finally included in the study was inferred from the core-gene alignment outputted from Roary. Recombination and highly conserved sites were eliminated using Gubbins v3.1.0 (299) and the final alignment was trimmed with Trimal v1.5.0 (300). Phylogenetic tree was generated using Iqtree2 v2.4.0 (296) with automatic model selection (-MFP) and 1000 rounds of standard nonparametric bootstrap. The tree was processed and visualized using iTOL (297) , where it was rooted at midpoint, and branches with bootstrap < 70 were collapsed. The phylogenetic tree inferred from 23S SNPs included ascertainment bias correction in Iqtree2 (-MFP+ASC), and was visualized using Ggtree (204) and Phytools (301).

Differently abundant genes between clusters were identified by comparing the prevalence as explained previously (prevalence differences > 0.5 or < -0.5). Additionally, genes absent in a group and found over 25% strains in the other were included. Some of the associated proteins had uncomplete or absent

functional annotation, which were further searched using blastp (117) against Uniprot (277) and NCBI databases.

Enrichment analysis was performed in String (302) using the sequences of the selected genes as input and selecting *M. hyorhinis* as organism. Significantly enriched biological processes were selected when false discovery rate was $P < 0.05$.

Genome scale metabolic modelling

Genome-scale metabolic models were generated from annotated genomes using Carveme v1.6 (79), by adapting the biomass equation to *M. hyorhinis* using the information in Ferrarini, et al. (303). Models were gapfilled to ensure metabolic viability of the strains in a simulated rich cultivation medium resembling typical Mycoplasma cultivation media, with a similar number of gapfilled reactions (mean of 20 +- X per strain). Reaction prevalence analyses were performed using custom MATLAB scripts (Version R2021a), only taking reactions with a gene associated into consideration. Flux-Balance-Analysis (FBA) simulations using these models (i.e. reaction essentiality analysis) were performed using custom MATLAB (Version 2021a) scripts and commands from the COBRA toolbox (304) (Version 3.4, downloaded from https://github.com/opencobra/cobratoolbox in April of 2024), and using Gurobi (Version 9.1.2) as a linear solver.

Data processing and figures

Output data was processed using in-house bash scripts and using R script language version 4.2.2 in RStudio environment version 2022.07.0 (199) and ggplot (202), and reshape2 (201) packages. In all tests, significance thresholds were set at 0.05.

Results

Definition of M. hyorhinis pangenome

To build the pangenome of *M. hyorhinis,* the genome assemblies of all publicly available *M. hyorhinis* strains (N=99) were downloaded. Additional strains isolated from the nasal cavity of healthy animals (N=10) and from a variety of systemic organs of diseased pigs (N=8) were sequenced, assembled *de novo* and added to the dataset **(Table 1).**

Table 1. *Mycoplasma hyorhinis* strains isolated in this study. The isolation sites, animal health status, sequencing platform, number of contigs and length of the assembly are shown.

STRAIN*	ISOLATION SITE	ANIMAL HEALTH STATUS	SEQUENCING PLATFORM	CONTIGS	TOTAL LENGTH (bp)
103-2.C1	Nasal cavity	Healthy	NANOPORE	1	824556
103-2.C2	Nasal cavity	Healthy	NANOPORE	1	827863
107-1	Nasal cavity	Healthy	NANOPORE	1	868567
107-5.C1	Nasal cavity	Healthy	NANOPORE	30	816783
107-5.C2	Nasal cavity	Healthy	NANOPORE	2	867457
109-2	Nasal cavity	Healthy	NANOPORE	1	855111
82-1	Nasal cavity	Healthy	ILLUMINA+NANOPORE	2	868017
82-2	Nasal cavity	Healthy	ILLUMINA+NANOPORE	2	861055
82-3	Nasal cavity	Healthy	NANOPORE	1	868628
82-6	Nasal cavity	Healthy	ILLUMINA+NANOPORE	1	900431
101-8	Pericardium	Diseased	NANOPORE	2	893176
104-2	Pleura	Diseased	NANOPORE	1	863659
61-2	Joint	Diseased	ILLUMINA+NANOPORE	2	883462
62-1	Peritoneum	Diseased	ILLUMINA+NANOPORE	13	862897
74-2	Joint	Diseased	ILLUMINA+NANOPORE	5	862971
83-1	Pleura	Diseased	ILLUMINA+NANOPORE	7	901416
83-4	Pericardium	Diseased	NANOPORE	1	909194
RM5-5	Joint	Diseased	ILLUMINA+NANOPORE	1	898669

* The strain name corresponds to “farm ID - animal ID”. C1 and C2 correspond to two different colonies isolated from the same sample.

Initially, the phylogenetic relationship of all 117 available genomes (18 sequenced in this study plus the 99 publicly available strains, including the *M. hyorhinis* reference strain BTS-7), was estimated together with the reference strains of other *Mycoplasmas* (i.e., *M. ovipneumoniae*, *M. dispar*, *M. flocculare, M. hyopneumoniae, M. conjuctivae* and *M. hyosynoviae*). All *M. hyorhinis* strains showed an ANI between them of 99.53% (range 99.08% to 99.65%), while with other *Mycoplasma* species the mean identity was between 86.45% and 92.32%. We confirmed that all strains included in the pangenome analysis belonged to the *M. hyorhinis* cluster in the phylogenetic tree (**supplementary Figure 1A**). The analysis of the preliminary pangenome showed a mean number of genes per strain of 663 ± 12.7 and a mean number of shared genes of 636 ± 7.2. One strain isolated from lung in France, shared an unusual low number of genes with the other strains (591.67, **supplementary Figure 1B**) and was discarded from the analysis after confirming that its genome was uncomplete (286 contigs). Although one strain from the nasal cavity isolated in this study (strain 107-5.C1) showed a remarkable high number of predicted genes (719), it was kept in the downstream analysis because the number of shared genes with other strains was within the normal range. Finally, 4 strains isolated from cell culture, 1 strain from *Bos taurus* and 1 from unknown host (**supplementary Figure 1C**) were also discarded. After removing the stated seven strains from the analysis, a final number of 110 strains were kept in the study. Among the strains finally included in the study, 36 were isolated from a variety of systemic organs, 44 from lungs and 30 from the nasal cavity. The specific isolation sites and the associated animal health status are shown in **supplementary Figure 1D**. The final *M. hyorhinis* pangenome was computed from the refined pool of genomes. All strains included in the pangenome exhibited a similar number of genes mostly shared across strains (mean number of genes found per strain was 662.6, mean number of shared genes between strains was 635.4, **Figure 1A**). The core-genome was represented by 588 genes, while 452 were found to be accessory genes (**Figure 1B**), summing up a total amount of

1,040 genes in the pangenome (**Figure 1C**). From these, 1,014 were associated to proteins where 669 were functionally annotated (66%).

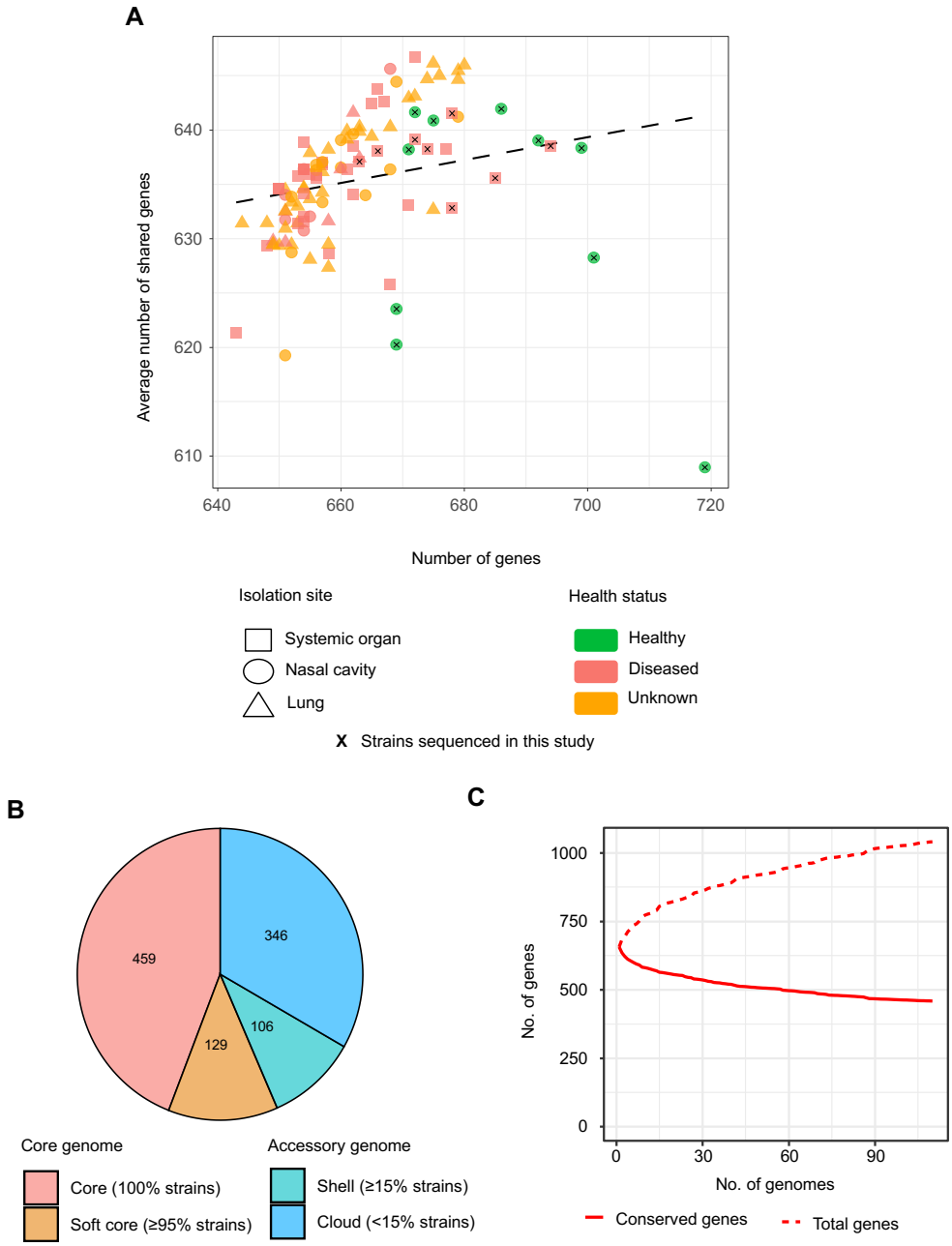


Figure 1. Overview of the M. hyorhinis pangenome. A) Number of genes and mean number of shared genes of each strain under study. Each dot represents one strain. Shape stands for isolation site and colour code follows the host health status. Strains sequenced in this study are indicated with crosses. B) Number of genes in the pangenome. Core-genes are shown in red and orange while accessory-genes are shown in turquoise and blue. C) Number of conserved (solid line) versus total genes (dashed line) appearing in the pangenome through all the strains.

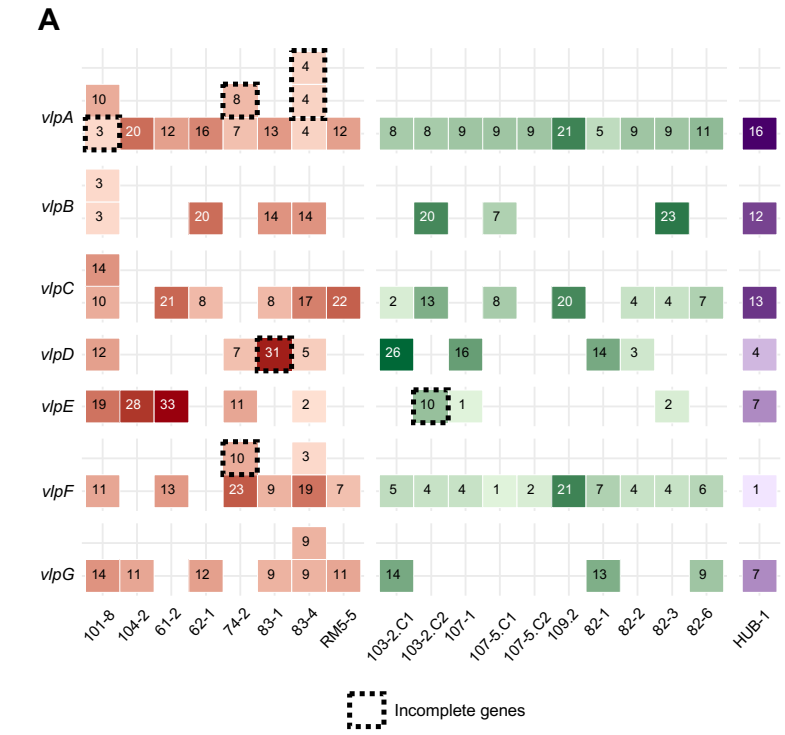
M. hyorhinis pangenome analysis did not reveal differences related to pathogenicity

We analysed the pangenome to identify genomic differences (in terms of presence /absence of genes) when comparing strains from different isolation sites or animal health status. Nevertheless, significant clustering was found only when considering the country of origin of the strains (PERMANOVA *P*=0.003 **supplementary figure 2**).

To investigate factors that could be responsible for pathogenicity, the presence of genes possibly involved in virulence was compared between healthy nasal strains and systemic strains. Proteins classified as virulence genes and/or possibly related with adhesion, cell migration, competence, effectors, exoenzymes, helicases, immune modulation, lipoproteins, metal uptake, mobile elements, molecule modification, peroxidase, phages, regulation, secretion, signalling, stress and toxins were identified in the pangenome (**supplementary table 2**). While a *recD* helicase, some variants of the *dcm* DNA modification system, and *vlpA* were more prevalent in nasal strains, the *hsdM/R* modification system occurred at a lower frequency in these strains. The virulence factors present in strains with unknown health status either from lung or nasal cavities were also compared with systemic strains, showing no differentially prevalent genes (see **supplementary table 2** for gene prevalence comparison). A more specific search of lipoproteins and secreted proteins done using SignalP resulted in similar observations. A few genes were predicted as lipoproteins or contained a signal peptide and

were found equally in all groups (**supplementary table 2**). Only three predicted proteins were exclusively detected in the strains from the nasal cavity of healthy animals (>80% strains), but not in strains from systemic organs, nasal cavity of unknown health-status animals and lungs. *MOS_610* and *vlpA* were already detected in the previous analysis, whereas *unknown99* was not; it was later classified as a hypothetical protein with similarity to other *Mycoplasma* lipoproteins.

A detailed analysis of the *vlp* genes revealed that they were misannotated or appeared as hypothetical proteins in the pangenome. No significant differences in the prevalence of any of the seven *vlp* genes were observed between the two groups of strains sequenced in this study (**Figure 2A**). When the number of repetitions was compared between groups (**Figure 2B**), the genes coding for *vlpF* and *vlpC* tended to show more repetitions only in systemic strains (only significant for *vlpF*). Notably, we observed that *vlp*s were present in multiple copies in several systemic strains (8 out of 39), but in none of the strains isolated from the nose of healthy animals. In some cases, difficulties in sequencing and assembly of these regions did not allow a fully characterization of the genes (indicated with dashed lines in **Figure 2A**). A *vlpA* in strain 1001-8 and a *vlpF* in strain 74-2 were incomplete for being in the beginning of a contig. Also, *vlpA* from strain 74-2, *vlpD* from 83-1 and *vlpE* from 103-2.C2 showed unannotated repetitions possibly due to a frameshift. Strain 83-4 had two *vlpA* fragments (identified by repetitions) separated by an insertion, which may represent a unique gene, as only one of the fragments had the signal sequence. Nevertheless, including the unannotated repeats in the comparison of repetitions did not alter the results in any Vlp.


A
vlpA
vlpB
vlpC
vlpD
vlpE
vlpF
vlpG
101-8
104-2
61-2
62-1
74-2
83-1
83-4
RM5-5
103-2.C1
103-2.C2
107-1
107-5.C1
107-5.C2
109-2
82-1
82-2
82-3
82-6
HUB-1
Incomplete genes

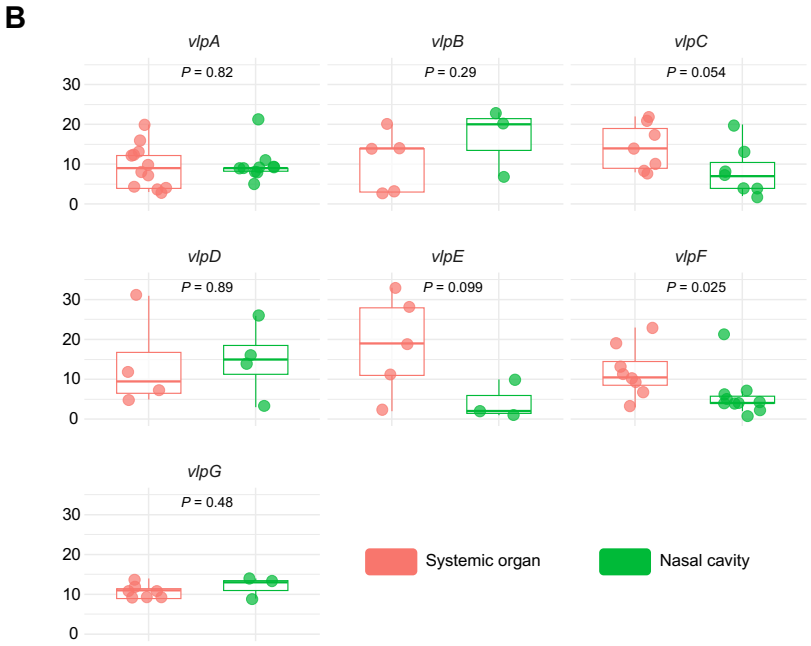

B
vlpA
P = 0.82
vlpB
P = 0.29
vlpC
P = 0.054
vlpD
P = 0.89
vlpE
P = 0.099
vlpF
P = 0.025
vlpG
P = 0.48
30
20
10
0
Systemic organ
Nasal cavity

Figure 2. Identification of the variable lipoproteins (VLPs) described in M. hyorhinis HUB-1 in the strains characterized in this study. A) Presence and number of repetitions of the VLPs identified. Colour code follows the host health status / isolation site, while the intensity corresponds to the number of repetitions found in region III. Incomplete VLPs are shown in dashed squares (see main text). For improved visualization, numbers on darker backgrounds are shown in white. B) Number of repetitions by VLP type compared between nasal and systemic strains with Mann-Whitney U tests.

We next evaluated possible antibiotic resistances in the *M. hyorhinis* pangenome using different antimicrobial resistance databases yielding no evidence of resistance genes. When we studied point mutations in particular genes previously found to confer resistance in other *Mycoplasma* species, twelve significant SNPs in the *23S rRNA* gene were found (**Figure 3)**, with some SNPs specifically associated to healthy animals (positions 727, 879, 1246 and 2134, indicated in green). A SNP-tree inferred from these polymorphisms showed also a strong clustering of strains isolated from the nasal cavity of healthy animals. Additionally, a G2057A substitution, previously described to confer resistance to erythromycin in other swine *Mycoplasmas* (291), was found only in strains isolated from animals with diseased or unknown health status, corresponding to 72.7% of all the strains. On the contrary, the analysis for 16S rRNA gene revealed only one significant SNP with no association to a particular cluster. Finally, no associations were established between study groups and mutations in *gyrA*/*gyrB* (DNA gyrase) and *parC*/*parE* (DNA topoisomerase), given their limited or widespread occurrence across M. hyorhinis strains. (**Supplementary Table 3**).

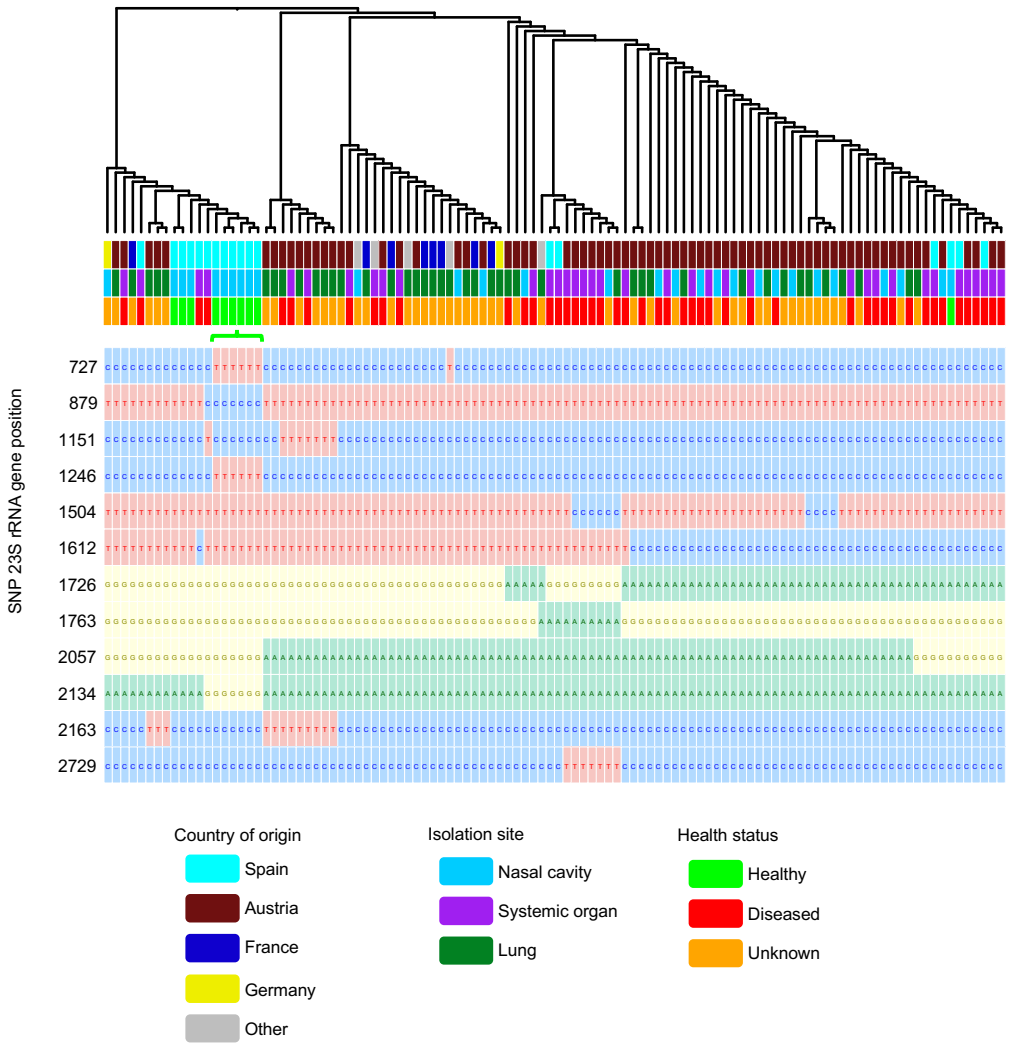


Figure 3. SNP-tree of all strains computed using significant SNPs found in the 23S rRNA gene. Colour code follows the country of origin, isolation site and host health status (from top to bottom). Significant SNPs in each strain are attached below. Each allele is shown in one different colour: green for Adenine, blue for Cytosine, yellow for Guanine and red for Timine. SNPs specific to health-associated strains are highlighted with a green bracket.

Core-genome phylogeny identified a cluster of six strains isolated from healthy animals

To further investigate the genetic relationship of the strains, the genes present in the core were used to infer the phylogeny (**Figure 4**). Although no global clustering according to any of the categories under study was detected, we found a clade of six genetically close strains isolated from the nasal cavity of healthy animals 82-1, 82-2, 82-3, 107-1, 107-5.C1 and 107-5.C2 (NH cluster indicated in green in **Figure 4**).

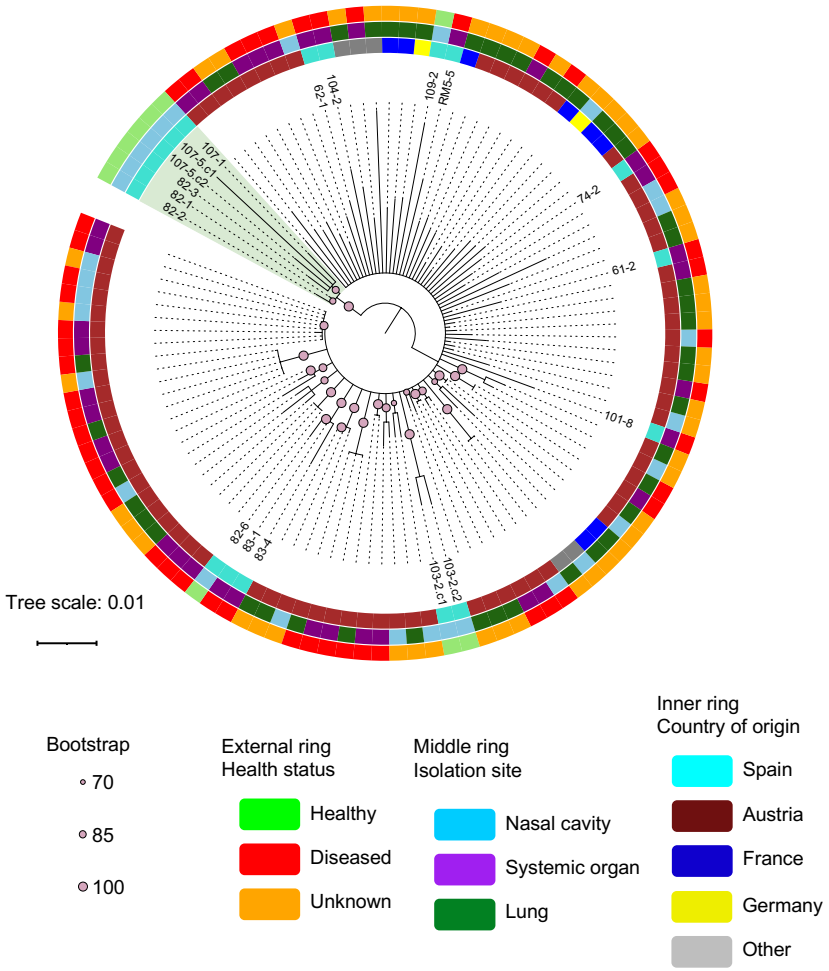
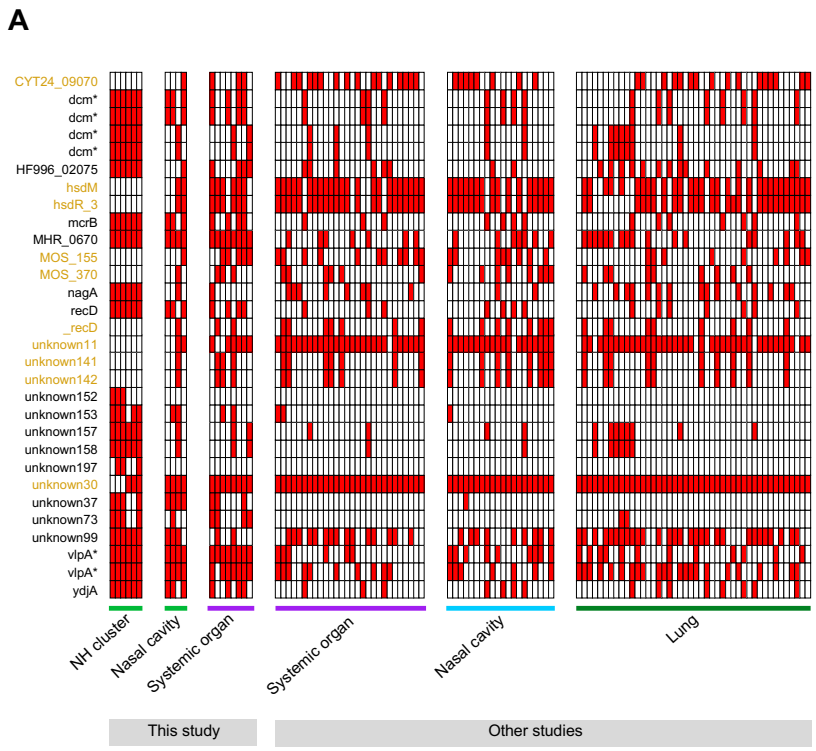


Figure 4. Phylogenetic tree computed using the alignment from core-genes. Colour code follows the host health status, isolation site and country of origin (from outside to inside). The clade of six strains isolated from the nasal cavity of healthy animals is highlighted in green. Branches with bootstrap support < 70 have been collapsed. The names of the strains characterized in this study are shown.

To evaluate the relevance of the NH cluster, their whole genomes were compared with the ones from systemic strains, yielding 30 potential genes with different prevalence. Then, we examined the presence of these 30 genes and the Jaccard distances throughout all the strains (**Figure 5A**). Apart from reporting differences by country of origin (PERMANOVA $P<0.05$), strains isolated from the nasal cavity of healthy animals in Spain showed significant differences with strains isolated from systemic organs, including those isolated in the same country (PERMANOVA $P=0.04$, **Figure 5B).** Interestingly, the significance of this comparison increased when only the strains from Spain were analysed ($P=0.003$). Despite the low number of samples, the robustness of the results was supported by additional analyses: all leave-one-out PERMANOVA tests yielded *P* values below 0.05, and only 4.5% of bootstrap iterations with randomly shuffled metadata showed significant differences (global empirical P = 0.004).

Remarkably, some genes absent in NH cluster emerged as possible virulence markers (highlighted in yellow in **Figure 5A**). Among these, we found two *hsd* type 1 restriction-modification system genes (*hsdM* and *hsdR*) that were almost absent in all samples from the nasal cavity of healthy animals. Besides, helicase genes, such as *CYT24_09070*, *_recD*, *MOS_370*, *unknown141* and *unknown142*. Notably, _recD was confirmed to be distinct from recD (i.e., not a variant) but was functionally annotated as _recD due to its similarity to recD-like helicases. Moreover, it had similarity with DEAD/DEAH box helicase family proteins. Therefore, it was included among the characterized genes and considered a separate protein. Additionally, *MOS_155* mobile element

protein, a family envelope stress response protein annotated as *unknown30* and hypothetical protein *unknown11* were also totally or partially missing in NH cluster. Interestingly, some of these genes were consistently located in contiguous regions within the contigs, suggesting the absence of a specific genomic region in the nasal strains. Although only 8 of these genes were identified in the String database, enrichment analysis confirmed that most associations between them were due to gene neighbourhood, except *hsdM* and *hsdR* that also showed other types of associations (cooccurrence, coexpression and/or interaction of homologs in other organisms, etc.). The whole network had a functional enrichment of catalytic activity on DNA (*P*=0.004) dependant on ATP (*P*=0.04).



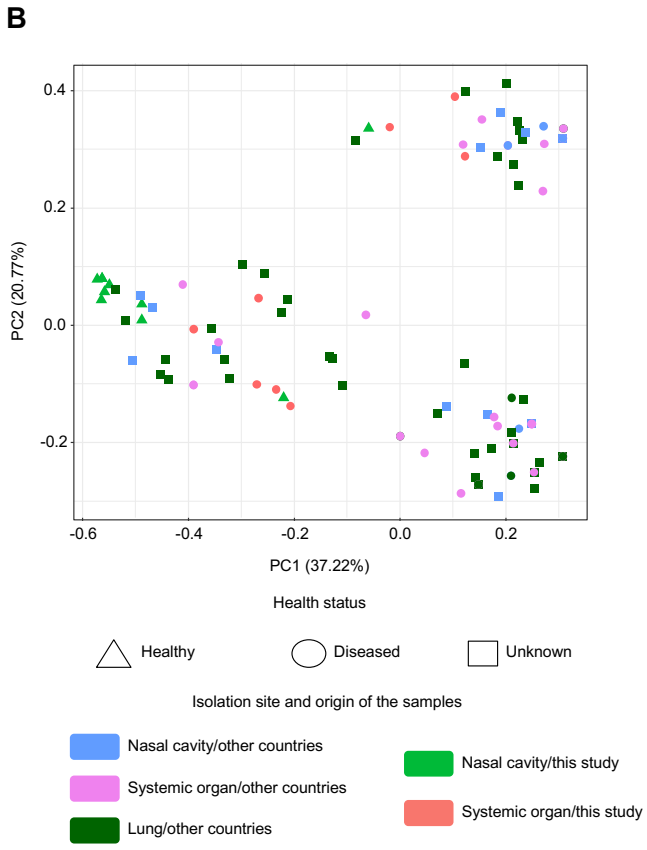
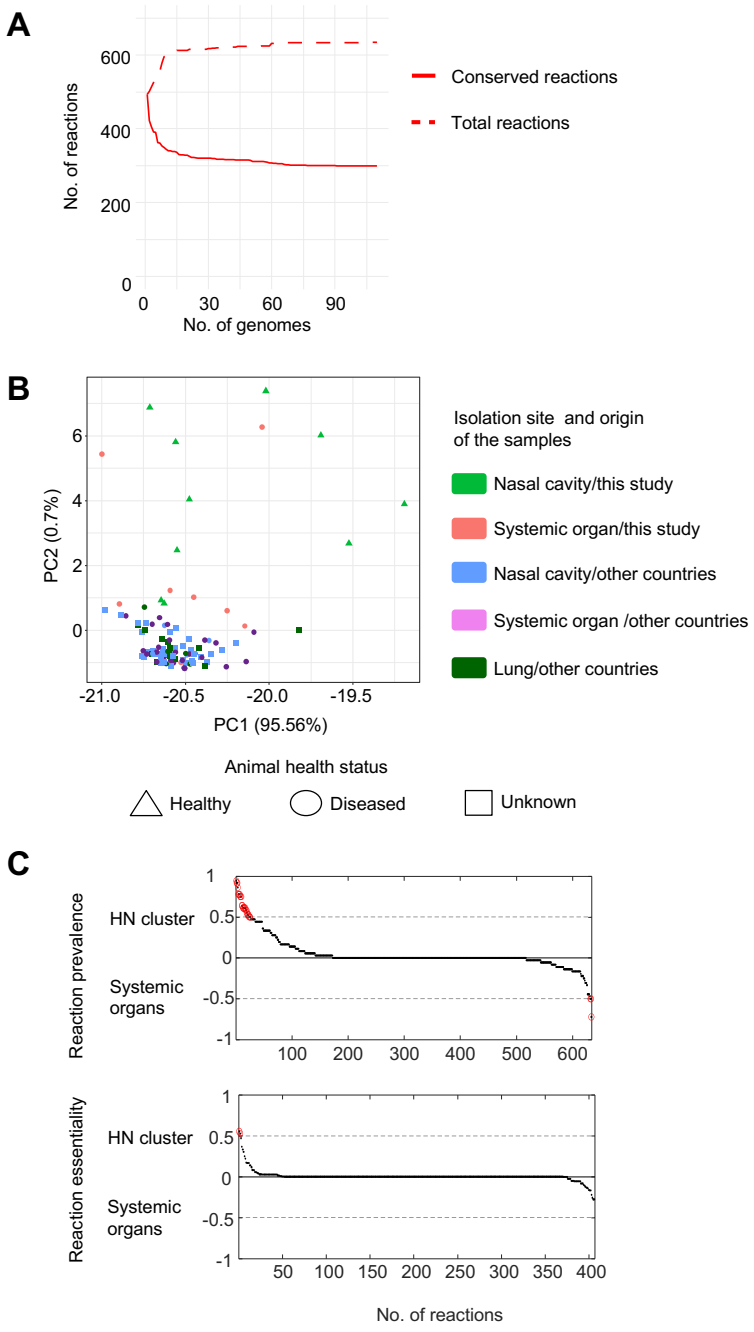


Figure 5. Genes differently prevalent between NH cluster and systemic strains. A) Presence and absence across all strains (ordered by isolation site). Selected markers for deeper study are highlighted in yellow. B) Principal Coordinate Analysis considering all differential genes using Jaccard distances. Shape stands for the host health status and colour code follows isolation site/origin of the samples.

Genome-scale metabolic modelling revealed conserved reactome across M. hyorhinis strains

The metabolic capabilities of the *M. hyorhinis* strains were computationally predicted using genome-scale metabolic models to search differences between groups of strains at reaction level (panreactome). The adapted biomass equation used to generate the *M. hyorhinis* models is shown in

supplementary table 4. All strains reported a similar number of reactions (mean of 441 ± 16.1), giving a final number of 634 total reactions of which 299 were shared between all strains (**Figure 6A**). In agreement with the genomic information, distances between strains based on reaction presence/absence showed differences according to the country of origin (Jaccard distance, PERMANOVA $P < 0.05$), but not according to the isolation site or associated host health status. Most strains isolated from the nasal cavities of healthy animals showed greater variability in their metabolic reactions, suggesting the presence of distinct reactions compared to the main pool of *M. hyorhinis*. However, due to the high variability among them, they did not appear to share a consistent set of differential reactions (**Figure 6B**). The comparison of reaction presence between the NH cluster and strains isolated from systemic organs revealed that most reactions were shared across all strains and only showed some transport reactions (mostly of peptides) more prevalent in strains coming from healthy animals (**Figure 6C top and supplementary table 5**). Since these groups were uneven, other comparisons were made (i.e. using all strains from healthy animals or only systemic strains isolated in this study), yielding similar results (**supplementary table 5**). The reaction essentiality was also compared between groups, by calculating the impact of restricting the flux of each individual reaction to zero (equivalent to simulating a reaction "knock-out") in each group. The few differences in the predicted relative growth rate indicated no differences in reaction essentiality between the different groups (**figure 6C bottom**), meaning that each reaction was equally essential or dispensable for all the models, also when evaluating the other comparisons. Finally, although our models support the possibility that *M. hyorhinis* may exhibit multiple auxotrophies, we did not detect any significant differences in the predicted auxotrophies among the groups analysed (**figure 6D**).


A
600
400
200
0
No. of reactions
0
30
60
90
No. of genomes
Conserved reactions
Total reactions
B
6
4
2
0
PC2 (0.7%)
-21.0
-20.5
-20.0
-19.5
PC1 (95.56%)
Isolation site and origin of the samples
Nasal cavity/this study
Systemic organ/this study
Nasal cavity/other countries
Systemic organ /other countries
Lung/other countries
Animal health status
Healthy
Diseased
Unknown
C
Reaction prevalence
HN cluster
Systemic organs
1
0.5
0
-0.5
-1
100
200
300
400
500
600
Reaction essentiality
HN cluster
Systemic organs
1
0.5
0
-0.5
-1
50
100
150
200
250
300
350
400
No. of reactions

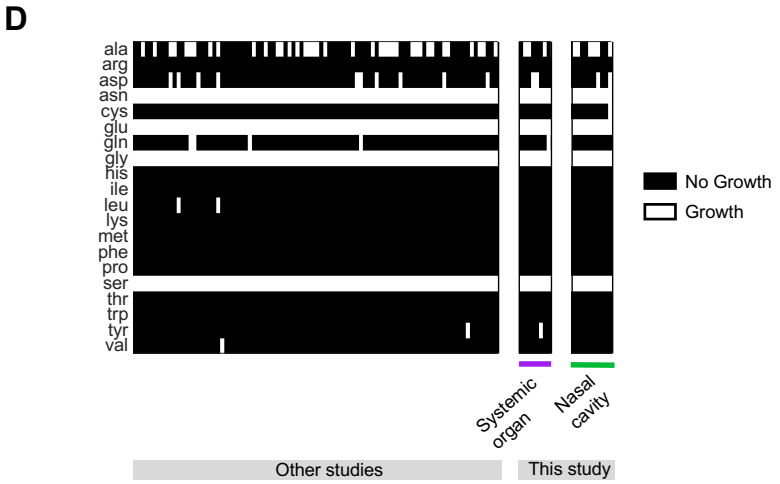


Figure 6. Genome-scale metabolic modelling of M. hyorhinis strains. A) Number of conserved (solid line) versus total reactions (dashed line) appearing in the predicted panreactome. B) PCA analysis of the reaction presence/absence in the strains (dots). Shape stands for the host health status and colour code follows isolation site/origin of the samples. C) Top: Reaction prevalence difference in the groups compared; clade of six strains isolated from the nasal cavity versus strains isolated from systemic organs. Reactions differing between the groups compared are highlighted in red (explained in supplementary table 5). Bottom: Difference in reaction essentiality between the same groups compared. Reactions differing between the groups compared are highlighted in red. Only reactions present at least in the 90% of the strains are accounted. D) Predicted growth of the strains in the absence of each amino acid.

Discussion

In this study, we performed a *M. hyorhinis* pangenome analysis to identify characteristics that could help discriminating between commensal and pathogenic strains. Previous reports have shown that strains of *M. hyorhinis* can exhibit different capacity to develop disease, but no clear marker allowing the distinction of virulent strains has been defined so far (80–82). Most of the studies performing molecular characterization of *M. hyorhinis*, analysed strains isolated from systemic organs and respiratory tract sites from

diseased pigs (79,249–253). Only Clavijo *et al*. (2019) included few strains isolated from the nasal cavity and bronchus of healthy animals in the study (79). Using these techniques based on the detection of targeted house-keeping or specific genes, no distinctive clusters of *M. hyorhinis* strains associated with different clinical manifestations were identified. In this study, we compared the whole genome composition incorporating newly isolated strains from the nares of healthy animals as well as strains from systemic lesions.

Analyses of gene presence/absence detected few differences among strains with different virulence. Nasal strains isolated from healthy pigs that constituted the NH cluster lacked two of the *hsd* system genes (*hsdM* methylase and *hsdR* nuclease). In *Mycoplasma,* these genes are related with phase-variable DNA methylation allowing to respond to the environment and participate in the expression of virulence factors (305) and are widely associated with bacterial pathogenesis through virulence regulation (306). Moreover, Pobeguts *et al*. identified HsdM in clinical isolates of *M. hominis* but not in laboratory strains (305). Additionally, although systemic strains showed higher prevalence of genes that could also be related with virulence such as helicases (258,307,308), mobile elements or stress response activities (280), their incomplete annotation prevents drawing further conclusions.

Analysis of the *vlp* genes revealed that strains isolated from systemic organs showed more repetitions in the region III of VlpF (and a tendency for VlpC). Indeed, an increase in length of this region of different variable lipoproteins has been associated to pathogenicity. Longer region III of different *M. hyorhinis* Vlps (VlpsA, VlpB or VlpC) conferred resistance to grow-inhibition host antibodies (309) but decreased cytoadhesion capabilities and biofilm formation (264). Similar observations have been reported in *M. pulmonis*, where more tandem repeats in Vsa lipoprotein were associated to resistance to complement (310) or phagocytosis by macrophages (311) but with less

adhesion (312). All these findings may suggest that systemic strains may have higher dissemination capabilities. Moreover, some strains isolated from systemic organs showed more *vlp* gene copies compared to nasal strains coming from healthy animals. It is still unknown if gene copy number is actually related with pathogenicity, but we hypothesize that this could increase the antigenic diversity of the strains and increase their immune evasion capacity, as suggested in previous studies (256,267,313). However, genomic studies targeting surface variable lipoproteins in mycoplasmas should be carefully interpreted because of the recombination, slipped-strand mispairing, phase variation or variable expression events occurring in these genes (256,267,311,313,314). Here, we identified *vlps* by searching both for a conserved sequence in the signal region (266,315) and variable repetitive (264) regions, what proved a fruitful strategy because of the difficulties to assemble and annotate such proteins. Additional studies, as proteomics, comparing strains with different clinical background would help to understand the role of Vlps variability in *M. hyorhinis* pathogenicity.

Other determinants under investigation included the genes associated with resistance to antimicrobial agents. We did not find antimicrobial resistance genes using different databases, similarly to Gaeta, *et al*. in *M. ovipneumoniae* (259). However, many point mutations in DNA gyrases and topoisomerases previously associated with quinolone resistance (259,291,292,316) were detected in all *M. hyorhinis* strains, suggesting intrinsic and conserved resistances in this species. In agreement, *M. hyorhinis* species tends to present higher MIC values against fluoroquinolones compared to other swine *Mycoplasmas* (291). Moreover, analysis of the 16S and 23S rRNA genes was included, as mutations in these regions may also contribute to antibiotic resistance (291). Four SNPs were identified in the 23S rRNA gene that discriminated most strains isolated from the nasal cavity of healthy animals (NH cluster). Also, a G2057A substitution, previously described to confer intrinsic resistance to erythromycin in other

swine *Mycoplasmas* (291) , was only identified in the majority of the strains isolated from diseased animals.

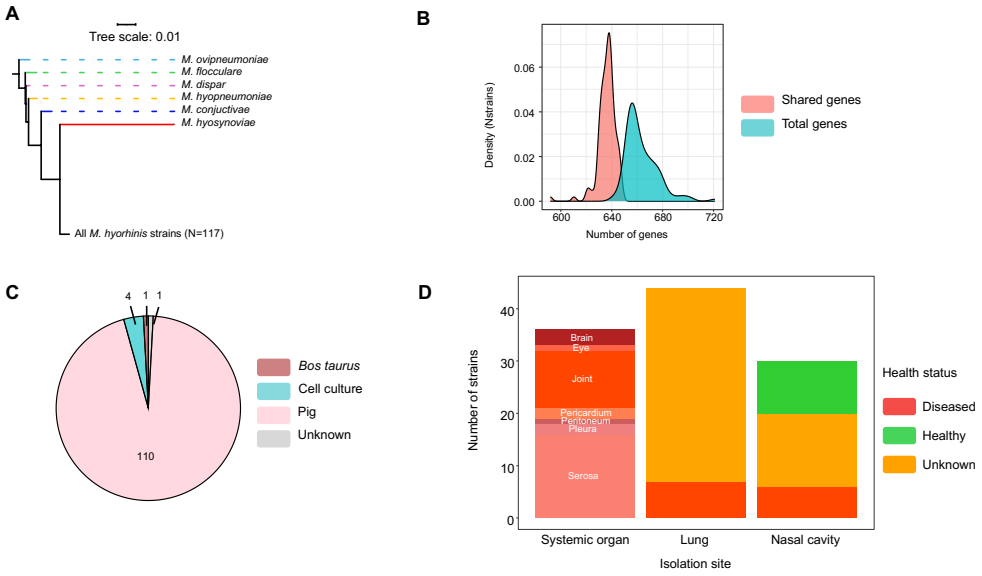
The panreactome analysis using genome-scale metabolic models also identified few differential reactions, e.g. some transport reactions (mainly of peptides) were more prevalent in nasal strains isolated from healthy animals. However, the observed differences in transport reactions did not result in significant variations in metabolic capabilities or auxotrophic requirements based on in silico simulations. These types of models were also studied for other swine mycoplasmas (*M. flocculare* and *M. hyopneumoniae*, together with *M. hyorhinis*) by Ferrarini, *et al*. where, although some differences were established between species, no intra-species variations were investigated (303). Genome-scale metabolic models have been useful to predict different metabolic and growth capabilities associated with variable traits between strains in other bacterial species (317), but our results showed no variations among *M. hyorhinis* strains.

The main limitation of this study is the number of strains analysed, along with the imbalance between strains isolated from healthy and diseased animals. The number of strains from nasal cavities of healthy pigs available in public data base is low probably due to the intrinsic difficulties in isolating unique *M. hyorhinis* strains in nasal cavities where more than one strain can be found (17) and because the lack of interest of this sample from a diagnostic point of view. Moreover, due to the scarce representation of *M. hyorhinis* in public databases, many genes remained unannotated, and potential antimicrobial resistance genes could not be identified. A larger number of strains, particularly from healthy animals, with complete metadata would help to confirm our findings. However, this study serves as a starting point to explore whether genomic information of the available *M. hyorhinis* strains can be linked to strain virulence.

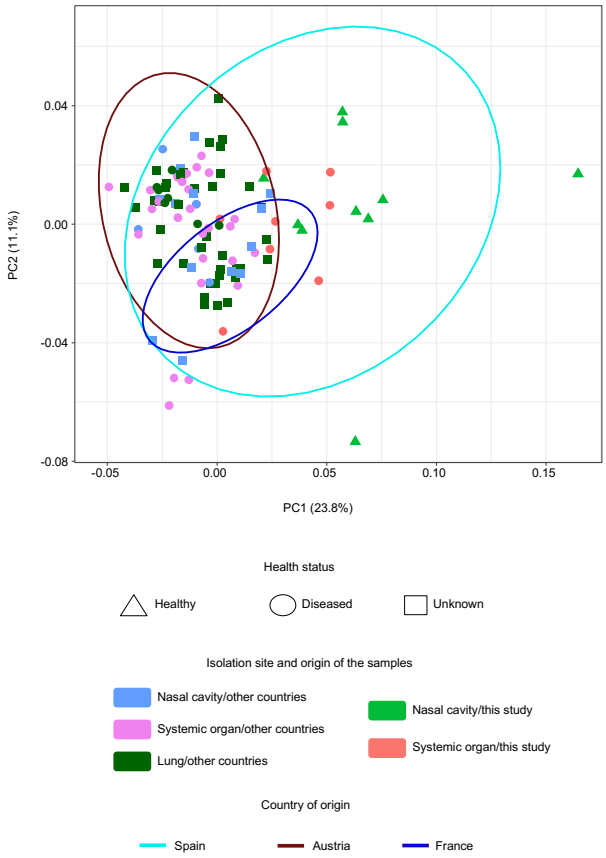
In conclusion, comparison of multiple *M. hyorhinis* strains yielded few differences that can be associated with virulence and suggests that other mechanisms, such as gene expression regulation, may be also determinant for virulence. Phylogenetic relationships allowed to discriminate a group of a few strains isolated from the nares of healthy pigs with some differences in genomic composition. Future efforts with higher number of strains, especially from healthy animals, will be determinant to confirm whether such differences are related with strain virulence and unveil the mechanisms that allow *M. hyorhinis* to disseminate systemically.

Supplementary materials

Supplementary figures



Supplementary figure 1. Preliminary analysis of the pool of M. hyorhinis strains. A) Phylogeny of all strains initially added in the study. The reference strains of other Mycoplasmas were added to the phylogeny (i.e., M. ovipneumoniae, M. dispar, M. floculare, M. hyopneumoniae, M. conjuctivae and M. hynosynoviae). All shown branch divisions are supported by bootstrap analysis. B) Distribution of the population of strains using their total number of genes (turquoise) and their mean number of shared genes with the other strains (red). C) Isolation host of the remaining 116 strains after eliminating one because of the low number of shared genes. D) Isolation site and host health status (colour coded) of the remaining pool of 110 strains used in this study.



Supplementary figure 2. Principal Coordinate Analysis of all strains using Jaccard distances computed from the presence and absence of all genes in the pangenome, coloured by isolation site/origin of the samples. Shape stands for the host health status. Ellipses of confidence using Euclidean distances are shown by country of origin of the strains (only for countries > 2 strains).

Supplementary tables

Supplementary tables are available in the following Zenodo repository:

https://zenodo.org/records/15769032

General discussion

Understanding the microbiota composition, as well as its interactions with the host immunity and opportunistic pathogens, is essential for improving pig health and reducing disease outbreaks in the swine industry. The early-life microbiota is strongly shaped by external factors and tends to develop a stable core composition with important functional effects. While the global microbial composition appears to modulate the risk of disease, the contribution of different strains within each bacterial species needs to be better defined, especially those with different degree of virulence.

Standardization and optimization of protocols are essential for microbiota analysis. The pipeline developed in this thesis addresses and standardizes key steps such as read filtering and taxonomic assignment through an ASV-based approach, which can be among the steps introducing biases between studies (99,100). This methodological contribution was essential for ensuring the reliability and comparability of microbial profiles across different studies. Differences in microbiota composition between farms or batches were consistently observed, indicating that local environmental conditions and farm-specific practices, are structuring forces. These environmental effects were not only biologically meaningful but also analytically relevant, as they had to be accounted in the design of statistical models and bioinformatics pipelines to ensure robust interpretations of the results. Studies focused on pig URT microbiotas have highlighted differences according to farms, batches or co-housing environments (17,30). This phenomenon is also evident in human studies, where cohabitation and shared environments facilitate the transmission of microbial strains across individuals (318).

Studies on the swine nasal microbiota composition in health and in altered states (diseases, treatments, environments, etc.) converge on several key points. The swine nasal microbiota exhibits a relatively consistent composition, with recurring genera and species across individuals (26,319).

This composition is affected by environmental conditions (9,26) and disturbances or changes in the microbial structure commonly correlate with compromised health and greater pathogen incidence (16,17,32,46–48,57,58). Collectively, the studies on the swine nasal microbiota have identified a group of taxa that constitute the core members in healthy animals (26,319). Interestingly, several of the bacterial genera identified in the swine respiratory tract are commonly found in the human upper respiratory tract, highlighting a notable overlap in microbial colonizers across species (170). This cross-species similarity suggests that certain commensal taxa may play conserved roles in maintaining respiratory health and immune homeostasis in mammals. Taxa commonly found in the core nasal microbiota of healthy piglets, such as *Moraxella*, *Streptococcus*, *Lactobacillus*, *Rothia*, *Corynebacterium*, *Neisseria* and *Bergeyella* (among others), were identified to be transmitted by the sows, highlighting the importance of the sow-piglet contact for the proper establishment of this microbiota. In agreement with the beneficial role of the core taxa, *Moraxella*, *Staphylococcus*, and *Neisseria* were positively associated with higher antibody responses against *G. parasuis* (Chapter 4), suggesting an immunomodulatory role within the nasal microbiota. In addition, increased relative abundance of *Bergeyella*, *Neisseria* (*Eikenella*), *Glaesserella* and *Moraxella* (*Moraxella*_A) was found in the nose of piglets that survived an infection with a highly virulent PRSS virus. The beneficial role of this core taxa is further supported by the fact that several species from these genera, including *Moraxella pluranimalium*, *Rothia nassimurium*, non-virulent *G. parasuis* and *Streptococcus pluranimalium*, were used as nasal colonizers for early-life inoculation in piglets, demonstrating their ability to establish in the upper respiratory tract and to partially restore microbiota composition after an antibiotic-induced disruption (20). Notably. this intervention was associated with a reduction in clinical signs and an overall improvement of microbiota structure, reinforcing their value as candidates for microbiota-based strategies to support respiratory health. Most of these bacteria were included in a defined porcine nasal consortium (PNC) that

represents the core nasal microbiota of piglets and encompasses the majority of its metabolic capacities (319). The ecological specificity of these taxa as nasal inhabitants is supported by the observation that ASVs classified as *Moraxellaceae*, *Pasteurellaceae*, and *Weeksellaceae* (families of *Moraxella*, *Glaesserella*, and *Bergeyella*, respectively) were virtually absent in matched-nasal rectal samples from (Chapter 3). On the other hand, *Lactobacillus* and *Streptococcus*, among others, were found in both types of samples, and also in other microbiotas (27,320,321), suggesting that these are more generalist bacteria able to thrive in a wide range of different niches.

Besides these specialized nasal colonizers, we reported anaerobic taxa (strict or facultative) typically associated with the gastrointestinal tract in the swine nasal microbiota, in agreement with previous studies (13,15–17,27,40,46,47,50). In the present work, we confirmed not only the presence of these microbes but also their likely fecal origin. This gut-associated fraction is mostly represented members of the orders *Clostridiales* and *Bacteroidales*, including taxa such as *Ruminococcaceae*, *Lachnospiraceae* and *Veillonellaceae* families, or *Clostridium, Prevotella* and *Bacteroides*, among others. The detection of shared and abundant ASV between these matched sites supports a strain-level connection and suggests a fecal origin, that can be explained by the rooting behavior of piglets (13,174,175). Interestingly, we found that these gut-associated anaerobes showed higher relative abundance in piglets with altered sow-contact (Chapter 2), possibly due to the lack of the nasal colonizers, transmitted by the sows. In agreement, we also detected higher abundances of gut-associated taxa in two different scenarios of microbiota dysbiosis, i.e. in piglets born to antibiotic-treated sows (33) or piglets that did not survive an infection with a highly virulent PRRS virus strain (58). Similarly, Zeineldin *et al.* reported increased abundances of *Clostridium,* together with a decrease of typical nasal colonizers, after parenteral ceftiofur administration (36). Moreover, we observed that these microbes appeared also in the deep nasal cavity and with similar activity

levels as other commensals (RNA/DNA ratio), indicating that these anaerobes could be more than transient passengers in the nose. Remarkably, a positive correlation was observed between the abundance of gut-associated microbes and antibody responses to *G. parasuis*, probably through immune system stimulation. Future studies are needed to unravel the role and functionality that these gut-associated microbes provide to the nasal microbiota, and to determine their precise localization within the respiratory mucosa, as the survival of anaerobic bacteria, and especially strict anaerobes, depends on microenvironments protected from oxygen exposure. The presence and location of anaerobes was only discussed by Peña-cortes, et al., who observed a progressive increase in *Clostridiales* when analyzing the tonsillar microbiota throughout pig development. They suggested that this might be due to the formation of deeper tonsillar crypts, either as a natural consequence of aging or induced by the activity of the bacteria themselves, and wondered whether these taxa might play roles similar to those observed in the gut microbiota, where they have been associated with resistance to pathogen colonization (13). Interestingly, some of these genera are also detected in the human nasal microbiota, often represented by different strains (322), although in lower prevalence and abundance than those observed in piglets.

In early life, beyond the gut-nasal connection, other maternal microbiotas, such as the vaginal and skin microbiota, may also contribute to nasal community assembly. Several genera identified in noses (most of them generalists), such as *Streptococcus, Staphylococcus, Corynebacterium* or *Lactobacillus,* are well-documented members of both vaginal and skin microbiotas in sows (320,321,323). These communities were shown to share taxa with the early tonsillar microbiota of piglets (27), supporting maternal contact at birth and during lactation as a relevant source of initial colonizers. Likewise, Bugenyi et al. (324) and Zang et al. (325) reported overlaps between vaginal, rectal and oropharyngeal microbiotas, indicating that these

environments are not independent and may contribute jointly to the microbial landscape of piglets. Notably, several of the anaerobic taxa identified in the nasal cavity and suspected to originate from fecal microbiota, such as *Clostridium, Bacteroides, Prevotella, Ruminococcaceae, Lachnospiraceae* and *Veillonellaceae*, have also been detected in the vaginal microbiota (320,324), and to a lesser extent in skin (321). This suggests that not only the gut, but also the vaginal and skin microbiota may serve as sources of colonizers that reach the nasal cavity, especially at early life stages when piglets are with the sows and still have an unmature microbiota. In contrast, other taxa such as *Moraxella, Glaesserella, Bergeyella* and *Rothia* (to a lesser extent) which were consistently associated with the nasal cavity in our studies and previous literature, were found in very low abundance or absent entirely in the vaginal and skin microbiotas (320,321). This supports the idea that these genera constitute specialized members of the upper respiratory microbiota, adapted to its specific ecological conditions and selective environmental pressures. Similar patterns have been observed in humans, where strain-resolved metagenomic studies have pointed out to the mothers as the main source of early-life microbial strains, with infant initial oral colonization reflected the microbial composition of maternal vaginal, skin, and oral sites (326). A recent study specifically highlighted that the main source of the nasal microbiota in infants during early life is the mother's nasal microbiota (29).

Beyond taxonomic composition, microbial diversity emerged as another key feature for respiratory health, also via immune modulation. We described that higher alpha diversity was associated with higher antibody response against *G. parasuis*). The association with health is reinforced across several independent studies, which reported that increased microbial richness and/or evenness were consistently associated with more resilient communities and/or better health status (15–17,20,26,33–35,40,49,58,327,328). However, not all forms of diversity are equally beneficial. For instance, piglets reared in

high biosecurity facilities, particularly those with absent or limited sow contact, exhibited a high alpha diversity driven by an unusual microbial community (Chapter 2). We hypothesize that this pattern is related to both the overrepresentation of transient environmental taxa and lack of common professional colonizers and thus, not linkable with a status of health. A similar phenomenon was also observed in ceftiofur-treated piglets and in piglets born to ceftiofur-treated sows, which showed a transient high diversity associated to the presence of environmental taxa (20).

Our findings align with a broader body of research on the human digestive and nasal microbiome, which consistently demonstrates that each body site harbors a distinct and characteristic microbial community (i.e. a site-specific core microbiota) composed of a conserved set of bacterial genera shared across healthy individuals (322,329–331). This core is maintained regardless of minor environmental influences, and while inter-individual differences may occur at finer taxonomic levels, such as species or strain level, the functional profiles of these communities tend to remain consistent across individuals, suggesting that both taxonomic stability and functional conservation are key features of a healthy microbiota (322,330,331). In humans, high microbial diversity is also a signature of health. Reduced diversity has been repeatedly associated with different diseases, such as inflammatory bowel disease, obesity and chronic respiratory conditions, among many others, often reflecting a microbial imbalance or loss of ecological resilience (5,7,330,332–334). This indicates that diversity, together with a stable core composition and preserved functional potential, plays a central role in maintaining host–microbiota homeostasis across animal species.

We also aimed to explore microbes at a strain level, since studies performed using 16S profiling indicated ambiguous roles within a genus. For instance, we reported different *Moraxella* clades associated in opposite ways with the immune response to *G. parasuis* (Chapter 4), and different ASVs classified within *Lactobacillus, Streptococcus, Prevotella, Moraxella, Clostridium* or

Fusobacterium genera (among others) associated to either surviving or succumbing outcomes after a highly virulent PRRS virus infection (58) (. Studies at strain level are especially necessary in the case of pathobiont species, since these can be responsible for clinical outbreaks and great welfare and economic loses (8,21,22). Among the three pathobionts that can act as causative agents for polyserositis of weaned piglets, strains of different virulence have been already described for *S. suis* and *G. parasuis*, together with virulence markers for their differentiation (8,21,22,66,335,336). Here, we focused on the other pathogen capable of causing this relevant disease, *M. hyorhinis* whose mechanisms of pathogenicity are less understood. Although previous studies have suggested varying degrees of virulence among different *Mycoplasma hyorhinis* strains (17,80–82), no specific genetic markers or genomic patterns have been identified that correlate with virulence or the site of isolation (79,249–253). Our genomic analysis, despite employing extensive comparative analyses, revealed no strong genomic signatures that allowed to differentiate strains of different virulence or organ of isolation. Interestingly, we identified a phylogenetic cluster of nasal strains from healthy animals, with distinct characteristics that seems to represent a genetically conserved lineage associated with a commensal behavior. This nasal cluster lacked a few genes possibly linked with virulence and showed different point mutations. However, we cannot rule out that the ability of *M. hyorhinis* to cause systemic disease may depend less on its genome and more on other factors. In particular, phenotypic heterogeneity derived from mechanisms like phase variation, antigenic switching, or variable gene expression may generate subpopulations within genetically similar strains that differ in traits such as adhesion, immune evasion, or tissue tropism (256,264,266,267). Such flexibility could allow the bacterium to explore different host niches or adapt rapidly to changing conditions without requiring permanent genetic changes. Moreover, the surrounding microbiota may act as a key ecological determinant in this process. As demonstrated here or in many other studies, the microbiota influences both pathogen resistance and host immunity, and

may directly or indirectly modulate the behavior of *M. hyorhinis*. A balanced microbiota could have a role in inhibiting activation. Hyorhinis either through direct and indirect competition, or immune stimulation, while disruptions to this community might create permissive conditions for invasion.

This thesis shows that the pig nasal microbiota comprises a stable microbial community consisting of both professional specialist colonizers and more generalist microbes connected with other body sites and environmental sources. Taken together, the findings presented in this thesis contribute to a more integrative and functional understanding of the porcine nasal microbiota. By combining taxonomic, immunological, and genomic analyses across diverse experimental contexts, this work helps underlines the nasal microbiota as an active player in host physiology and disease susceptibility. These results set the stage for future applications of microbiota-directed strategies, ranging from microbial (strain) monitoring to targeted interventions, that could support respiratory health and resilience in swine.

Future perspectives

While this thesis has laid a strong foundation in the field, it simultaneously highlights several promising lines of future investigation. These lines of investigation can be broadly categorized into two main areas. The first involves a deeper study of the swine respiratory microbiome, aiming to identify its key microbial inhabitants in healthy conditions, understand how they are influenced by intrinsic and extrinsic factors, and uncover the mechanisms by which they support host health. The second focuses on developing and applying therapeutic strategies that manipulate the microbiota to enhance host well-being, offering a potentially effective alternative to conventional treatments. Building on the findings of this thesis and taking into account its limitations, several important research questions have emerged that deserve further investigation.

What are the consequences of losing part of the commensal microbiota in piglets without contact with sows?

In chapter 2 piglets reared without normal contact with sows exhibited a dysbiotic nasal microbiota marked by a lack of common commensals and the over-detection of environmental taxa, indicating that sows are a fundamental source of microbiota. In further studies, most of the diminished populations have been associated with better immune responses and / or better disease outcomes. Thus, it would be of interest to evaluate the performance or the disease susceptibility of such piglets later in life, reinforcing the role of the sows in the rearing period, but also the beneficial role of several nasal commensals.

What are exactly gut-associated microbes doing in the nasal cavity?

Chapter 3 demonstrated gut-associated microbes were present through all the nasal passage of piglets, and that these taxa were equally active than other commensals. Approaches such as WGS would allow to unravel the full genomic composition of these strains (99) and thus, what capacities they can bring to the nasal microbiota. Besides, transcriptomic and proteomic studies could reveal what are actually these anaerobes doing. Moreover, shifting the focus to these microbes in the nasal cavity raises additional questions from an ecological perspective, particularly, how these organisms survive and where they encounter low-oxygen niches within an environment that is predominantly aerobic (205). To localize these anaerobes, it would be useful to establish specific techniques, such as in situ hybridization, for their detection in the tissue.

Is there a true strain-connection between microbial sources?

In this thesis, connection between microbiotas was made at ASV level (i.e. nasal and rectal ASVs overlap in chapter 3). Studies carried in humans at strain level using WGS were able to quantify the bacterial strains shared between mother and infant (326,337), co-habitants (318) and populations (338). The application of this technique to pigs will allow a precise resolution to conclusively define the strains shared between different microbiotas and sources, shedding light into the development and shaping of the microbiota in pigs.

How do beneficial taxa stimulate the immune system?

Chapter 4 connected increased abundances of several nasal colonizers with higher antibody responses. Mechanisms through which gut microbiota stimulate the immune system have been described, such as activating the inflammasome (339), activating TLRs through recognition of MAMPs (340), promoting NK cells (341) and memory B cells (342) or maintaining regulatory T cells (343). There are less studies focusing on pigs, and some of them using gnotobiotic pigs as a human model. Still, mechanisms such as modulation of TLR expression, cytokine profiles, adaptive immune cells (like T cells, NK cells), dendritic–T cell crosstalk, and maintenance of epithelial barrier integrity have been described (344–348). Notably, these studies do not target the nasal epithelium and are rather focused on the gut. In addition, most studies focus on correlations while those addressing the mechanistic pathways remain scarce. Future efforts examining specific compounds from commensal microbes of interest will help clarifying how the nasal microbiota is able to modulate immunity, especially in the case of the pigs' respiratory mucosa.

How do beneficial commensals interact with pathobionts?

Many of the commensal microbes described here have also been associated with a reduced incidence of diseases, often through direct effects on pathogens via mechanisms such as direct or indirect killing, or competitive exclusion (4–6). Future research focusing on the nasal microbiota of pigs should aim to elucidate the complex interactions between commensal microbes and pathobionts. This includes investigating which taxa outcompete others for key nutrients or niches, identifying commensals capable of inhibiting or killing pathobionts, and determining how these dynamics vary across host conditions or disease states. This new knowledge could pave the way for microbiota-based interventions to promote resilience against pathogens. Ongoing efforts are aiming to uncover the interactions among eight of the most abundant members of the nasal microbiota, which are expected to shed light on some of these questions (319).

Are *M. hyorhinis* strains variable in virulence?

Finally, chapter 5 yielded few genomic differences between *M. hyorhinis* strains according to their clinical origin, with a few distinct traits in some strains isolated from the nares of healthy animals. A clear limitation of the analysis is the reduced number of “non-invasive” strains associated to colonization of healthy piglets, since most of the *M. hyorhinis* available strains had been isolated from diseased animals. Moreover, the global number of strains compared to other pangenome / panreactome analyses is still very low (66,317,349). Therefore, it is necessary to increase the number of *M. hyorhinis* strains, especially those isolated from the nasal cavity of healthy animals. In addition, there might be other mechanisms implicated in *M. hyorhinis* virulence regulation, such as gene expression or interactions with other microbes. Hence, to gain a comprehensive understanding of these regulatory processes, it would be valuable to perform transcriptomics,

proteomics, and metabolomics studies. These multi-omics approaches may reveal key pathways involved in virulence, identify potential biomarkers, and uncover how *M. hyorhinis* interacts with the host and adapts to different organs.

Conclusions

1. The robust pipeline that was established in this thesis facilitates a comprehensive characterization of the nasal microbiota in pigs under different scenarios.

2. Sows play a crucial role as primary sources of early-life colonizers of the piglets' nasal cavity.

3. Artificially rearing piglets in highly controlled environments, without the presence of sows, can significantly influence the nasal microbiota of weaning piglets and may introduce bias into microbiome research.

4. Anaerobic taxa are present and metabolically active in the nasal microbiota of pigs.

5. The composition and diversity of the piglets' nasal and rectal microbiotas are associated to the immune response to vaccination and / or natural exposure to *Glaesserella parasuis*.

6. Strains of *Mycoplasma hyorhinis* share most of their genomic content with few variations potentially associated with virulence.

Bibliography

1. Whipps JM, Lewis K, Cooke RC. Mycoparasitism and plant disease control 161– 187. In: Burge, NM (editor), Fungi in Biological Control Systems. Manchester University Press; 1988. P. 176.

2. Berg, G., Rybakova, D., Fischer, D. et al. Microbiome definition re-visited: old concepts and new challenges. Microbiome 8, 103 (2020). https://doi.org/10.1186/s40168-020-00875-0.

3. Lloyd-Price, J., Abu-Ali, G., & Huttenhower, C. (2016). The healthy human microbiome. Genome medicine, 8(1), 51. https://doi.org/10.1186/s13073-016-0307-y.

4. Pickard, J. M., Zeng, M. Y., Caruso, R., & Núñez, G. (2017). Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. Immunological reviews, 279(1), 70–89. https://doi.org/10.1111/imr.12567.

5. Hou, K., Wu, ZX., Chen, XY. et al. Microbiota in health and diseases. Sig Transduct Target Ther 7, 135 (2022). https://doi.org/10.1038/s41392-022-00974-4.

6. Adak, A., & Khan, M. R. (2019). An insight into gut microbiota and its functionalities. Cellular and molecular life sciences : CMLS, 76(3), 473–493. https://doi.org/10.1007/s00018-018-2943-4.

7. Martel, J., Chang, S. H., Ko, Y. F., Hwang, T. L., Young, J. D., & Ojcius, D. M. (2022). Gut barrier disruption and chronic disease. Trends in endocrinology and metabolism: TEM, 33(4), 247–265. https://doi.org/10.1016/j.tem.2022.01.002.

8. Costa-Hurtado M, Barba-Vidal E, Maldonado J, Aragon V. Update on Glässer's disease: How to control the disease under restrictive use of antimicrobials. Vet Microbiol. 2020 Mar;242:108595. doi: 10.1016/j.vetmic.2020.108595. Epub 2020 Jan 25. PMID: 32122599.

9. Niederwerder M.C. Role of the Microbiome in Swine Respiratory Disease. Vet. Microbiol. 2017;209:97–106. doi: 10.1016/j.vetmic.2017.02.017.

10. Nowland, T. L., Plush, K. J., Barton, M., & Kirkwood, R. N. (2019). Development and Function of the Intestinal Microbiome and Potential Implications for Pig Production. Animals : an open access journal from MDPI, 9(3), 76. https://doi.org/10.3390/ani9030076.

11. Wang, M., Radlowski, E. C., Monaco, M. H., Fahey, G. C., Jr, Gaskins, H. R., & Donovan, S. M. (2013). Mode of delivery and early nutrition modulate microbial colonization and fermentation products in neonatal piglets. The Journal of nutrition, 143(6), 795–803. https://doi.org/10.3945/jn.112.173096.

12. Campbell J.M., Crenshaw J.D., Polo J. The Biological Stress of Early Weaned Piglets. J. Anim. Sci. Biotechnol. 2013;4:19. doi: 10.1186/2049-1891-4-19.

13. Pena Cortes, L. C., LeVeque, R. M., Funk, J. A., Marsh, T. L., & Mulks, M. H. (2018). Development of the Tonsil Microbiome in Pigs and Effects of Stress on the Microbiome. Frontiers in veterinary science, 5, 220. o.

14. Guevarra RB, Lee JH, Lee SH, et al. Piglet gut microbial shifts early in life: causes and effects. J Anim Sci Biotechnol. 2019;10:1. Published 2019 Jan 14. doi:10.1186/s40104-018-0308-3.

15. Slifierz, M. J., Friendship, R. M., & Weese, J. S. (2015). Longitudinal study of the early-life fecal and nasal microbiotas of the domestic pig. BMC microbiology, 15(1), 184. https://doi.org/10.1186/s12866-015-0512-7.

16. Correa-Fiz F, Fraile L, Aragon V. Piglet nasal microbiota at weaning may influence the development of Glässer's disease during the rearing period. BMC Genomics. 2016 May 26;17:404. doi: 10.1186/s12864-016-2700-8. PMID: 27230662; PMCID: PMC4881051.

17. Blanco-Fuertes M, Correa-Fiz F, Fraile L, Sibila M, Aragon V. Altered Nasal Microbiota Composition Associated with Development of Polyserositis by Mycoplasma hyorhinis. Pathogens. 2021 May 14;10(5):603. doi: 10.3390/pathogens10050603. PMID: 34069250; PMCID: PMC8156107.

18. Tang X, Xiong K, Fang R, Li M. Weaning stress and intestinal health of piglets: A review. Front Immunol. 2022;13:1042778. Published 2022 Nov 24. doi:10.3389/fimmu.2022.1042778.

19. Saladrigas-García, M., Durán, M., D'Angelo, M. et al. An insight into the commercial piglet's microbial gut colonization: from birth towards weaning. anim microbiome 4, 68 (2022). https://doi.org/10.1186/s42523-022-00221-9.

20. Blanco-Fuertes, M., Sibila, M., Franzo, G., Obregon-Gutierrez, P., Illas, F., Correa-Fiz, F., & Aragón, V. (2023). Ceftiofur treatment of sows

results in long-term alterations in the nasal microbiota of the offspring that can be ameliorated by inoculation of nasal colonizers. Animal microbiome, 5(1), 53. https://doi.org/10.1186/s42523-023-00275-3.

21. Martelli, P., Segalés, J., Torremorell, M., Canelli, E., Maes, D., Nathues, H., Brockmeier, S., Gottschalk, M., & Aragón, V. (2019). Swine respiratory disease. Servet.

22. Zimmerman, J. J., Karriker, L. A., Ramirez, A., Schwartz, K. J., Stevenson, G. W., & Zhang, J. (Eds.). (2019). Diseases of swine (11ª ed.). Wiley.

23. Salzano FA, Marino L, Salzano G, et al. Microbiota Composition and the Integration of Exogenous and Endogenous Signals in Reactive Nasal Inflammation. J Immunol Res. 2018;2018:2724951. Published 2018 Jun 3. doi:10.1155/2018/2724951.

24. Li Y, Yang C, Jiang Y, et al. Characteristics of the nasal mucosa of commercial pigs during normal development. Vet Res. 2023;54(1):37. Published 2023 Apr 24. doi:10.1186/s13567-023-01164-y.

25. Scherzad A, Hagen R, Hackenberg S. Current Understanding of Nasal Epithelial Cell Mis-Differentiation. J Inflamm Res. 2019;12:309-317. Published 2019 Dec 13. doi:10.2147/JIR.S180853.

26. Pirolo, M., Espinosa-Gongora, C., Bogaert, D., & Guardabassi, L. (2021). The porcine respiratory microbiome: recent insights and future challenges. Animal microbiome, 3(1), 9. https://doi.org/10.1186/s42523-020-00070-4.

27. Pena Cortes, L. C., LeVeque, R. M., Funk, J., Marsh, T. L., & Mulks, M. H. (2018). Development of the tonsillar microbiome in pigs from newborn through weaning. BMC microbiology, 18(1), 35. https://doi.org/10.1186/s12866-018-1176-x.

28. Pirolo, M., Espinosa-Gongora, C., Alberdi, A., Eisenhofer, R., Soverini, M., Eriksen, E. Ø., Pedersen, K. S., & Guardabassi, L. (2023). Bacterial topography of the upper and lower respiratory tract in pigs. Animal microbiome, 5(1), 5. https://doi.org/10.1186/s42523-023-00226-y.

29. Quinn BE, Rodríguez JAR, Kweku Sam E, et al. The nasal microbiome in early infancy is primarily shaped by the maternal nasal microbiome. J Allergy Clin Immunol. Published online May 16, 2025. doi:10.1016/j.jaci.2025.05.004.

30. Fredriksen S, Guan X, Boekhorst J, Molist F, van Baarlen P, Wells JM. Environmental and maternal factors shaping tonsillar microbiota development in piglets. BMC Microbiol. 2022;22(1):224. Published 2022 Sep 26. doi:10.1186/s12866-022-02625-8.

31. Dickson RP, Erb-Downward JR, Freeman CM, et al. Bacterial Topography of the Healthy Human Lower Respiratory Tract. mBio. 2017;8(1):e02287-16. Published 2017 Feb 14. doi:10.1128/mBio.02287-16.

32. Siqueira FM, Pérez-Wohlfeil E, Carvalho FM, Trelles O, Schrank IS, Vasconcelos ATR, et al. Microbiome overview in swine lungs. PLoS One. 2017;12:e0181503. doi: 10.1371/journal.pone.0181503.

33. Bonillo-Lopez L, Obregon-Gutierrez P, Huerta E, Correa-Fiz F, Sibila M, Aragon V. Intensive antibiotic treatment of sows with parenteral crystalline ceftiofur and tulathromycin alters the composition of the nasal microbiota of their offspring. Vet Res. 2023 Nov 24;54(1):112. doi: 10.1186/s13567-023-01237-y.

34. Correa-Fiz, F., Gonçalves Dos Santos, J. M., Illas, F., & Aragon, V. (2019). Antimicrobial removal on piglets promotes health and higher bacterial diversity in the nasal microbiota. Scientific reports, 9(1), 6545. https://doi.org/10.1038/s41598-019-43022-y.

35. Mou KT, Allen HK, Alt DP, Trachsel J, Hau SJ, Coetzee JF, et al. Shifts in the nasal microbiota of swine in response to different dosing regimens of oxytetracycline administration. Vet Microbiol. 2019;108386:108386.

36. Zeineldin M, Aldridge B, Blair B, Kancer K, Lowe J. Microbial shifts in the swine nasal microbiota in response to parenteral antimicrobial administration. Microb Pathog. 2018;121:210–217. doi: 10.1016/j.micpath.2018.05.028.

37. Blanco-Fuertes M, Correa-Fiz F, López-Serrano S, Sibila M, Aragon V. Sow vaccination against virulent Glaesserella parasuis shapes the nasal microbiota of their offspring. Sci Rep. 2022 Mar 1;12(1):3357. doi: 10.1038/s41598-022-07382-2.

38. Chrun T, Leng J, La Ragione RM, Graham SP, Tchilian E. Changes in the Nasal Microbiota of Pigs Following Single or Co-Infection with Porcine Reproductive and Respiratory Syndrome and Swine Influenza A Viruses. Pathogens. 2021 Sep 22;10(10):1225. doi: 10.3390/pathogens10101225.

39. Nielsen, D. W., Hau, S. J., Mou, K. T., Alt, D. P., & Brockmeier, S. L. (2023). Shifts in the swine nasal microbiota following Bordetella bronchiseptica challenge in a longitudinal study. Frontiers in microbiology, 14, 1260465. https://doi.org/10.3389/fmicb.2023.1260465.

40. Jiang N, Liu H, Wang P, Huang J, Han H, Wang Q. Illumina MiSeq sequencing investigation of microbiota in bronchoalveolar lavage fluid and cecum of the swine infected with PRRSV. Curr Microbiol. 2019;76:222–230. doi: 10.1007/s00284-018-1613-y.

41. Hau SJ, Nielsen DW, Mou KT, Alt DP, Kellner S, Brockmeier SL. Resilience of swine nasal microbiota to influenza A virus challenge in a longitudinal study. Vet Res. 2023 May 2;54(1):38. doi: 10.1186/s13567-023-01167-9. PMID: 37131235; PMCID: PMC10152739.

42. Sonalio, K., Almeida, H. M. S., Mechler-Dreibi, M. L., Storino, G. Y., Haesebrouck, F., Maes, D., & de Oliveira, L. G. (2022). Influence of Mycoplasma hyopneumoniae natural infection on the respiratory microbiome diversity of finishing pigs. Veterinary research, 53(1), 20. https://doi.org/10.1186/s13567-022-01038-9.

43. Almeida HMdS, Sonalio K, Mechler-Dreibi ML, Petri FAM, Storino GY, Maes D, de Oliveira LG. Experimental Infection with Mycoplasma hyopneumoniae Strain 232 in Swine Influences the Lower Respiratory Microbiota. Veterinary Sciences. 2022; 9(12):674. https://doi.org/10.3390/vetsci9120674.

44. Rampelotto PH, dos Santos ACR, Muterle Varela AP, Takeuti KL, Loiko MR, Mayer FQ, Roehe PM. Comparative Analysis of the Upper Respiratory Bacterial Communities of Pigs with or without Respiratory Clinical Signs: From Weaning to Finishing Phase. Biology. 2022; 11(8):1111. https://doi.org/10.3390/biology11081111.

45. Niazy, M., Hill, S., Nadeem, K. et al. Compositional analysis of the tonsil microbiota in relationship to Streptococcus suis disease in nursery pigs in Ontario. anim microbiome 4, 10 (2022). https://doi.org/10.1186/s42523-022-00162-3.

46. Wang, Q., Cai, R., Huang, A., Wang, X., Qu, W., Shi, L., Li, C., & Yan, H. (2018). Comparison of Oropharyngeal Microbiota in Healthy Piglets and Piglets With Respiratory Disease. Frontiers in microbiology, 9, 3218. https://doi.org/10.3389/fmicb.2018.03218.

47. Huang, T., Zhang, M., Tong, X., Chen, J., Yan, G., Fang, S., Guo, Y., Yang, B., Xiao, S., Chen, C., Huang, L., & Ai, H. (2019). Microbial

communities in swine lungs and their association with lung lesions. Microbial biotechnology, 12(2), 289–304. https://doi.org/10.1111/1751-7915.13353.

48. Li Z, Wang X, Di D, Pan R, Gao Y, Xiao C, et al. Comparative analysis of the pulmonary microbiome in healthy and diseased pigs. Mol Genet Genomics. 2020. 10.1007/s00438-020-01722-5.

49. Wang T, He Q, Yao W, Shao Y, Li J, Huang F. The variation of nasal microbiota caused by low levels of gaseous ammonia exposure in growing pigs. Front Microbiol. 2019;10:1083. doi: 10.3389/fmicb.2019.01083.

50. Megahed A, Zeineldin M, Evans K, Maradiaga N, Blair B, Aldridge B, et al. Impacts of environmental complexity on respiratory and gut microbiome community structure and diversity in growing pigs. Sci Rep. 2019;9:13773. doi: 10.1038/s41598-019-50187-z.

51. Weese JS, Slifierz M, Jalali M, Friendship R. Evaluation of the nasal microbiota in slaughter-age pigs and the impact on nasal methicillin-resistant Staphylococcus aureus (MRSA) carriage. BMC Vet Res. 2014;10:69. doi: 10.1186/1746-6148-10-69.

52. Correa-Fiz, F., Neila-Ibáñez, C., López-Soria, S., Napp, S., Martinez, B., Sobrevia, L., Tibble, S., Aragon, V., & Migura-Garcia, L. (2020). Feed additives for the control of post-weaning Streptococcus suis disease and the effect on the faecal and nasal microbiota. Scientific reports, 10(1), 20354. https://doi.org/10.1038/s41598-020-77313-6.

53. Yang Y, Jiang X, Cai X, et al. Deprivation of Dietary Fiber Enhances Susceptibility of Piglets to Lung Immune Stress. Front Nutr. 2022;9:827509. Published 2022 Feb 10. doi:10.3389/fnut.2022.827509.

54. Yang, J., Wang, J., Shang, P., Liu, Z., Zhang, B., Yang, D., & Zhang, H. (2024). Translocation of probiotics via gut-lung axis enhanced pulmonary immunity of weaned piglets exposed to low concentrations of ammonia. Ecotoxicology and environmental safety, 284, 116821. https://doi.org/10.1016/j.ecoenv.2024.116821.

55. Schuijt TJ, Lankelma JM, Scicluna BP, et al. The gut microbiota plays a protective role in the host defence against pneumococcal pneumonia. Gut. 2016;65(4):575-583. doi:10.1136/gutjnl-2015-309728.

56. Thorburn AN, McKenzie CI, Shen S, Stanley D, Macia L, Mason LJ, et al. Evidence that asthma is a developmental origin disease influenced

by maternal diet and bacterial metabolites. Nat Commun. (2015) 6:1–13. 10.1038/ncomms8320.

57. Espinosa-Gongora C, Larsen N, Schønning K, Fredholm M, Guardabassi L. Differential analysis of the nasal microbiome of pig carriers or non-carriers of Staphylococcus aureus. PLoS One. 2016;11:e0160331.

58. Obregon-Gutierrez, P., Cortey, M., Martín-Valls, G. E., Clilverd, H., Correa-Fiz, F., Aragón, V., & Mateu, E. (2025). Nasal microbial diversity is associated with survival in piglets infected by a highly virulent PRRSV-1 strain. Animal microbiome, 7(1), 9. https://doi.org/10.1186/s42523-024-00371-y.

59. Bugenyi AW, Cho H-S, Heo J. Association between oropharyngeal microbiome and weight gain in piglets during pre and post weaning life. J Anim Sci Technol. 2020;62:247–62.

60. Salogni C, Capucchio MT, Colombino E, Pozzi P, Pasquali P, Alborali GL. Bacterial polyarthritis in post-weaning pigs in a high-density swine breeding area in Italy. J Vet Diagn Invest. 2022;34(4):709-711. doi:10.1177/10406387221090903.

61. Haas, B., & Grenier, D. (2018). Understanding the virulence of Streptococcus suis: A veterinary, medical, and economic challenge. Medecine et maladies infectieuses, 48(3), 159–166. https://doi.org/10.1016/j.medmal.2017.10.001.

62. Gottschalk, M. and Segura, M. (2019). Streptococcosis. In Diseases of Swine (eds J.J. Zimmerman, L.A. Karriker, A. Ramirez, K.J. Schwartz, G.W. Stevenson and J. Zhang). https://doi.org/10.1002/9781119350927.ch61.

63. Hill J.E., Gottschalk M., Brousseau R., Harel J., Hemmingsen S.M., Goh S.M. Biochemical analysis, cpn60 and 16C rDNA sequence data indicate that Streptococcus suis serotypes 32 and 34, isolated from pigs, are Streptococcus orisratti. Vet. Microbiol. 2005;107:63–69. doi: 10.1016/j.vetmic.2005.01.003.

64. Le Tien H.T., Nishibori T., Nishitani Y., Nomoto R., Osawa R. Reappraisal of the taxonomy of Streptococcus suis serotypes 20, 22, 26, and 33 based on DNA-DNA homology and soda and recN phylogenies. Vet. Microbiol. 2013;162:842–849. doi: 10.1016/j.vetmic.2012.11.001.

65. Lun, Z. R., Wang, Q. P., Chen, X. G., Li, A. X., & Zhu, X. Q. (2007). Streptococcus suis: an emerging zoonotic pathogen. The Lancet. Infectious diseases, 7(3), 201–209. https://doi.org/10.1016/S1473-3099(07)70001-4.

66. Murray GGR, Hossain ASMM, Miller EL, et al. The emergence and diversification of a zoonotic pathogen from within the microbiota of intensively farmed pigs. Proc Natl Acad Sci U S A. 2023;120(47):e2307773120. doi:10.1073/pnas.2307773120.

67. Neila-Ibáñez, C., Napp, S., Pailler-García, L., Franco-Martínez, L., Cerón, J. J., Aragon, V., & Casal, J. (2023). Risk factors associated with Streptococcus suis cases on pig farms in Spain. The Veterinary record, 193(5), e3056. https://doi.org/10.1002/vetr.3056.

68. M. Cerdà-Cuéllar, J.F. Naranjo, A. Verge, M. Nofrarias, M. Cortey, A. Olvera, J. Segales, V. Aragon Sow vaccination modulates the colonization of piglets by Haemophilus parasuis Vet. Microbiol., 145 (2010), pp. 315-320.

69. V. Aragon, J. Segales, A.W. Tucker Glässer's disease J.J. Zimmerman, L.A. Karriker, A. Ramirez, K.J. Schwartz, G.W. Stevenson, J. Zhang (Eds.), Diseases of Swine, Wiley-Blackwell, New Jersey (2019), pp. 844-853.

70. Kielstein, P., & Rapp-Gabrielson, V. J. (1992). Designation of 15 serovars of Haemophilus parasuis on the basis of immunodiffusion using heat-stable antigen extracts. Journal of clinical microbiology, 30(4), 862–865. https://doi.org/10.1128/jcm.30.4.862-865.1992.

71. Rafiee, M., Bara, M., Stephens, C. P., & Blackall, P. J. (2000). Application of ERIC-PCR for the comparison of isolates of Haemophilus parasuis. Australian veterinary journal, 78(12), 846–849. https://doi.org/10.1111/j.1751-0813.2000.tb10507.x.

72. Galofré-Milà N, Correa-Fiz F, Lacouture S, Gottschalk M, Strutzberg-Minder K, Bensaid A, Pina-Pedrero S, Aragon V. 2017. A robust PCR for the differentiation of potential virulent strains of Haemophilus parasuis. BMC Veterinary Research 13:124.

73. Bello-Orti B, Costa-Hurtado M, Martinez-Moliner V, et al. 2014a. Vet Microbiol 170:430–437.

74. Olvera A, Ballester M, Nofrarias M, et al. 2009. Vet Res 40:24.

75. Cerdà-Cuéllar, M., & Aragon, V. (2008). Serum-resistance in Haemophilus parasuis is associated with systemic disease in swine. Veterinary journal (London, England : 1997), 175(3), 384–389. https://doi.org/10.1016/j.tvjl.2007.01.016.

76. López-Serrano S, et al. Sow vaccination with a protein fragment against virulent Glaesserella (Haemophilus) parasuis modulates immunity traits in their offspring. Vaccines. 2021;9:534. doi: 10.3390/vaccines9050534.

77. Palzer et a., 2020 Mycoplasma hyorhinis and Mycoplasma hyosynoviae Chapter 13 pp 247-265 Mycoplasmas in Swine, 1st Edition, 2020 Edited By: Dominiek Maes, Marina Sibila and Maria Pieters Acco.

78. Pieters and Maes, 2019 Mycoplasmosis Chapter 56 pp 863-883 Disease of Swine, 11th Edition, 2019 Edited By: Jeffrey J. Zimmerman, Locke A. Karriker, Alejandro Ramirez, Kent J. Schwartz, Gregory W. Stevenson, Jianqiang Zhang Wiley Blacwell.

79. Clavijo, M.J., Sreevatsan, S., Johnson, T.J., Rovira, A., 2019. Molecular epidemiology of Mycoplasma hyorhinis porcine field isolates in the United States. PLoS One 14, e0223653. https://doi.org/10.1371/journal.pone.0223653.

80. Lin JH, Chen SP, Yeh KS, Weng CN. Mycoplasma hyorhinis in Taiwan: diagnosis and isolation of swine pneumonia pathogen. Vet Microbiol. 2006;115(1-3):111-116. doi:10.1016/j.vetmic.2006.02.004.

81. Gois M, Pospisil Z, Cerny M, Mrva V. 1971. Production of pneumonia after intransal inoculation of gnotobiotic piglets with three strains of Mycoplasma hyorhinis. J Comp Pathol 81:401–410. https://doi.org/10.1016/ 0021-9975(71)90028-4.

82. Wang, J., Hua, L., Gan, Y., Yuan, T., Li, L., Yu, Y., Xie, Q., Olaniran, A. O., Chiliza, T. E., Shao, G., Feng, Z., Pillay, B., & Xiong, Q. (2022). Virulence and Inoculation Route Influence the Consequences of Mycoplasma hyorhinis Infection in Bama Miniature Pigs. Microbiology spectrum, 10(3), e0249321. https://doi.org/10.1128/spectrum.02493-21.

83. Friis NF, Kokotovic B, Svensmark B. Mycoplasma hyorhinis isolation from cases of otitis media in piglets. Acta Vet Scand. 2002;43(3):191-193.

84. Hennig-Pauka I, Sudendey C, Kleinschmidt S, Ruppitsch W, Loncaric I, Spergser J. Swine Conjunctivitis Associated with a Novel Mycoplasma

Species Closely Related to Mycoplasma hyorhinis. Pathogens. 2020;10(1):13. Published 2020 Dec 25. doi:10.3390/pathogens10010013.

85. Resende TP, Pieters M, Vannucci FA. Swine conjunctivitis outbreaks associated with Mycoplasma hyorhinis. J Vet Diagn Invest. 2019;31(5):766-769. doi:10.1177/1040638719865767.

86. Shin, J.H., Joo, H.S., Lee, W.H., Seok, H.B., Calsamig, M., Pijoan, C., Molitor, T.W., 2003. Identification and characterization of cytopathogenic Mycoplasma hyorhinis from swine farms with a history of abortions. J. Vet. Med. Sci. 65, 501–509. https://doi.org/10.1292/jvms.65.501.

87. Martin MJ, Thottathil SE, Newman TB. Antibiotics Overuse in Animal Agriculture: A Call to Action for Health Care Providers. Am J Public Health. 2015;105(12):2409-2410. doi:10.2105/AJPH.2015.302870.

88. Tang, K. L., Caffrey, N. P., Nóbrega, D. B., Cork, S. C., Ronksley, P. E., Barkema, H. W., Polachek, A. J., Ganshorn, H., Sharma, N., Kellner, J. D., & Ghali, W. A. (2017). Restricting the use of antibiotics in food-producing animals and its associations with antibiotic resistance in food-producing animals and human beings: a systematic review and meta-analysis. The Lancet. Planetary health, 1(8), e316–e327. https://doi.org/10.1016/S2542-5196(17)30141-9.

89. D.B. Holman, M.R. Chénier, Antimicrobial use in swine production and its effect on the swine gut microbiota and antimicrobial resistance, Can. J. Microbiol. 61 (2015) 785–798, http://dx.doi.org/10.1139/cjm-2015-0239.

90. Blaser, M. (2011). Stop the killing of beneficial bacteria. Nature, 476(7361), Article 7361. https://doi.org/10.1038/476393a.

91. Guidelines for the Prudent Use of Antimicrobials in Veterinary Medicine. Official Journal of the European Union (2015/C 299/04) [(accessed on 26 March 2021)]; Available online: https://ec.europa.eu/health/sites/health/files/antimicrobial_resistance/docs/2015_prudent_use_guidelines_en.pdf.

92. Ji J, Jin W, Liu SJ, Jiao Z, Li X. Probiotics, prebiotics, and postbiotics in health and disease. MedComm (2020). 2023;4(6):e420. Published 2023 Nov 4. doi:10.1002/mco2.420.

93. Paule, A., Frezza, D., & Edeas, M. (2018). Microbiota and Phage Therapy: Future Challenges in Medicine. Medical sciences (Basel, Switzerland), 6(4), 86. https://doi.org/10.3390/medsci6040086.

94. Martens, K., Pugin, B., De Boeck, I., Spacova, I., Steelant, B., Seys, S. F., Lebeer, S., & Hellings, P. W. (2018). Probiotics for the airways: Potential to improve epithelial and immune homeostasis. Allergy, 73(10), 1954–1963. https://doi.org/10.1111/all.13495.

95. Liao, S. F., & Nyachoti, M. (2017). Using probiotics to improve swine gut health and nutrient utilization. Animal nutrition (Zhongguo xu mu shou yi xue hui), 3(4), 331–343. https://doi.org/10.1016/j.aninu.2017.06.007.

96. Rattigan R, Wajda L, Vlasblom AA, Wolfe A, Zomer AL, Duim B, et al. Safety Evaluation of an intranasally applied cocktail of Lactococcus lactis strains in pigs. Animals. 2023 Nov 8;13(22):3442.

97. Yang Y, Jing Y, Yang J, Yang Q. Effects of intranasal administration with Bacillus subtilis on immune cells in the nasal mucosa and tonsils of piglets. Exp Ther Med. 2018 Jun;15(6):5189-5198.

98. Scholz, M. B., Lo, C. C., & Chain, P. S. (2012). Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. Current opinion in biotechnology, 23(1), 9–15. https://doi.org/10.1016/j.copbio.2011.11.013.

99. Weinstock G. M. (2012). Genomic approaches to studying the human microbiota. Nature, 489(7415), 250–256. https://doi.org/10.1038/nature11553.

100. Pollock, J., Glendinning, L., Wisedchanwet, T., & Watson, M. (2018). The Madness of Microbiome: Attempting To Find Consensus 'Best Practice' for 16S Microbiome Studies. Applied and environmental microbiology, 84(7), e02627-17. https://doi.org/10.1128/AEM.02627-17.

101. Durazzi, F., Sala, C., Castellani, G., Manfreda, G., Remondini, D., & De Cesare, A. (2021). Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. Scientific reports, 11(1), 3030. https://doi.org/10.1038/s41598-021-82726-y.

102. Marotz CA, Sanders JG, Zuniga C, Zaramela LS, Knight R, Zengler K. Improving saliva shotgun metagenomics by chemical host DNA depletion. Microbiome. 2018;6(1):42. Published 2018 Feb 27. doi:10.1186/s40168-018-0426-3.

103. Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., Brown, C. T., Porras-Alfaro, A., Kuske, C. R., & Tiedje, J. M. (2014). Ribosomal Database Project: Data and tools for high throughput rRNA analysis. Nucleic Acids Research, 42(D1), D633–D642. https://doi.org/10.1093/nar/gkt1244.

104. McDonald D, Jiang Y, Balaban M, Cantrell K, Zhu Q, Gonzalez A, et al. Greengenes2 unifies microbial data in a single reference tree. Nature Biotechnology [Internet]. 2023 Jul 27; Available from: https://doi.org/10.1038/s41587-023-01845-1

105. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Research, 41(D1), D590–D596. https://doi.org/10.1093/nar/gks1219.

106. Douglas GM, et al. PICRUSt2 for prediction of metagenome functions. Nat. Biotechnol. 2020;38:685–688.

107. Wemheuer, F., Taylor, J. A., Daniel, R., Johnston, E., Meinicke, P., Thomas, T., & Wemheuer, B. (2020). Tax4Fun2: prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene sequences. Environmental microbiome, 15(1), 11. https://doi.org/10.1186/s40793-020-00358-7.

108. Blazewicz SJ, Barnard RL, Daly RA, Firestone MK. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. ISME J. 2013;7(11):2061-2068. doi:10.1038/ismej.2013.102.

109. Sarangi AN, Goel A, Aggarwal R (2019) Methods for studying gut microbiota: a primer for physicians. J Clin Exp Hepatol 9(1):62–73.

110. Bokulich NA, Ziemski M, Robeson MS, Kaehler BD (2020) Measuring the microbiome: best practices for developing and benchmarking microbiomics methods. Comput Struct Biotechnol J 18:4048–4062.

111. Gotschlich EC, Colbert RA, Gill T (2019) Methods in microbiome research: past, present and future. Best Pract Res Clin Rheumatol 33(6):101498.

112. Xue Z, Kable ME, Marco ML (2018) Impact of DNA sequencing and analysis methods on 16S rRNA gene bacterial community analysis of dairy products. mSphere 3(5):e00410–e00418.

113. Obregon-Gutierrez, P., Aragon, V., & Correa-Fiz, F. (2021). Sow Contact Is a Major Driver in the Development of the Nasal Microbiota of Piglets. Pathogens (Basel, Switzerland), 10(6), 697. https://doi.org/10.3390/pathogens10060697.

114. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M et al (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 41(1):e1.

115. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA et al (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37(8):852–857.

116. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–583. doi: 10.1038/nmeth.3869.

117. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K et al (2009) BLAST+: architecture and applications. BMC Bioinform 10(1):421.

118. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A et al (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 6(3):610–618.

119. Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: a versatile open source tool for metagenomics. PeerJ 18:e2584.

120. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V et al (2011) Scikit-Learn: Machine Learning in Python. J Mach Learn Res 12:2825–2830.

121. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30(4):772–780.

122. Lane DJ (1991) 16s/23s rRna Sequencing. Nucleic Acid Techniques in Bacterial Systematics. John Wiley and Sons; New York, NY, USA pp. 115–175.

123. Price MN, Dehal PS, Arkin AP (2010) FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS One 5(3):e9490.

124. Vázquez-Baeza Y, Pirrung M, Gonzalez A, Knight R (2013) EMPeror: a tool for visualizing high-throughput microbial community data. Gigascience 2(1):16.

125. Halko N, Martinsson PG, Shkolnisky Y, Tygert M (2010) An algorithm for the principal component analysis of large data sets. SIAM J Sci Comput 33:2580–2594.

126. Kruskal WH, Wallis WA (1952) Use of ranks in one-criterion variance analysis. J Am Stat Assoc 47(260):583–621.

127. Anderson M (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26:32–46.

128. Oksanen F, Blanchet G, Friendly M, Kindt R, Legendre P, McGlinn D et al (2020) Vegan: community ecology package. R Package Version 2:5–7. Available online: https://CRAN.R-project.org/package=vegan.

129. Rideout JR, Chase JH, Bolyen E, Ackermann G, González A, Knight R et al (2016) Keemei: cloud-based validation of tabular bioinformatics file formats in Google sheets. Gigascience 5:27.

130. Chao A (1984) Nonparametric estimation of the number of classes in a population. Scand J Stat 4:265–270.

131. Shannon C, Weaver W (1948) The mathematical theory of communication. Bell Syst Tech J 27:379–423.

132. Simpson EH (1949) Measurement of diversity. Nature 163(4148):688–688.

133. Bray JR, Curtis JT (1957) An ordination of upland forest communities of southern Wisconsin. Ecol Monogr 27:325–349.

134. Jaccard P (1908) Nouvelles Recherches sur la Distribution Florale. Bull Société Vaud Des Sci Nat 44:223–270.

135. Lozupone CA, Hamady M, Kelley ST, Knight R (2007) Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. Appl Environ Microbiol 73(5):1576–1585.

136. Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol 71(12):8228–8235.

137. Barko P.C., McMichael M.A., Swanson K.S., Williams D.A. The Gastrointestinal Microbiome: A Review. J. Vet. Intern. Med. 2018;32:9–25. doi: 10.1111/jvim.14875.

138. Lazar V., Ditu L.-M., Pircalabioru G.G., Gheorghe I., Curutiu C., Holban A.M., Picu A., Petcu L., Chifiriuc M.C. Aspects of Gut Microbiota and Immune System Interactions in Infectious Diseases, Immunopathology, and Cancer. Front. Immunol. 2018;9:1830. doi: 10.3389/fimmu.2018.01830.

139. Rawls M., Ellis A.K. The microbiome of the nose. Ann. Allergy Asthma Immunol. 2019;122:17–24. doi: 10.1016/j.anai.2018.05.009.

140. Fouhse J.M., Zijlstra R.T., Willing B.P. The Role of Gut Microbiota in the Health and Disease of Pigs. Anim. Front. 2016;6:30–36. doi: 10.2527/af.2016-0031.

141. Wang S., Ryan C.A., Boyaval P., Dempsey E.M., Ross R.P., Stanton C. Maternal Vertical Transmission Affecting Early-Life Microbiota Development. Trends Microbiol. 2020;28:28–45. doi: 10.1016/j.tim.2019.07.010.

142. Meat—Food and Agriculture Organization of the United Nations. [(accessed on 26 March 2021)];2020 Available online: http://www.fao.org/3/ca8861en/Meat.pdf.

143. Cheng G., Hao H., Xie S., Wang X., Dai M., Huang L., Yuan Z. Antibiotic Alternatives: The Substitution of Antibiotics in Animal Husbandry? Front. Microbiol. 2014;5:217. doi: 10.3389/fmicb.2014.00217.

144. Jiang L., Feng C., Tao S., Li N., Zuo B., Han D., Wang J. Maternal Imprinting of the Neonatal Microbiota Colonization in Intrauterine Growth Restricted Piglets: A Review. J. Anim. Sci. Biotechnol. 2019;10:1–8. doi: 10.1186/s40104-019-0397-7.

145. Sakwinska O., Foata F., Berger B., Brüssow H., Combremont S., Mercenier A., Dogra S., Soh S.-E., Yen J.C.K., Heong G.Y.S., et al. Does the Maternal Vaginal Microbiota Play a Role in Seeding the Microbiota of Neonatal Gut and Nose? Benef. Microbes. 2017;8:763–778. doi: 10.3920/BM2017.0064.

146. Baekbo P., Kristensen C.S., Larsen L.E. Porcine Circovirus Diseases: A Review of PMWS: Porcine Circovirus Diseases. Transbound. Emerg. Dis. 2012;59:60–67. doi: 10.1111/j.1865-1682.2011.01288.x.

147. Jayaraman B., Nyachoti C.M. Husbandry Practices and Gut Health Outcomes in Weaned Piglets: A Review. Anim. Nutr. 2017;3:205–211. doi: 10.1016/j.aninu.2017.06.002.

148. Murase K., Watanabe T., Arai S., Kim H., Tohya M., Ishida-Kuroki K., Võ T.H., Nguyễn T.P.B., Nakagawa I., Osawa R., et al. Characterization of Pig Saliva as the Major Natural Habitat of Streptococcus suis by Analyzing Oral, Fecal, Vaginal, and Environmental Microbiota. PLoS ONE. 2019;14:e0215983. doi: 10.1371/journal.pone.0215983.

149. Larivie G. Salmonella Shedding Status of the Sow Affects the Microbiota of Their Piglets at Weaning. J. Appl. Microbiol. 2019;126:411–423. doi: 10.1111/jam.14139.

150. Blanco I., Galina-Pantoja L., Oliveira S., Pijoan C., Sánchez C., Canals A. Comparison between Haemophilus parasuis Infection in Colostrums-Deprived and Sow-Reared Piglets. Vet. Microbiol. 2004;103:21–27. doi: 10.1016/j.vetmic.2004.06.011.

151. Opriessnig T., Gerber P.F., Halbur P.G. Refinement of a Colostrum-Deprived Pig Model for Infectious Disease Research. MethodsX. 2018;5:403–413. doi: 10.1016/j.mex.2018.03.010.

152. Vahle J.L., Haynes J.S., Andrews J.J. Interaction of Haemophilus parasuis with Nasal and Tracheal Mucosa Following Intranasal Inoculation of Cesarean Derived Colostrum Deprived (CDCD) swine. Can. J. Vet. Res. 1997;61:200–206.

153. Weiss S., Xu Z.Z., Peddada S., Amir A., Bittinger K., Gonzalez A., Lozupone C., Zaneveld J.R., Vázquez-Baeza Y., Birmingham A., et al. Normalization and Microbial Differential Abundance Strategies Depend upon Data Characteristics. Microbiome. 2017;5:27. doi: 10.1186/s40168-017-0237-y.

154. Correa-Fiz F., Blanco-Fuertes M., Navas M.J., Lacasta A., Bishop R.P., Githaka N., Onzere C., Le Potier M.-F., Almagro-Delgado V., Martinez J., et al. Comparative Analysis of the Fecal Microbiota from Different Species of Domesticated and Wild Suids. Sci. Rep. 2019;9:13616. doi: 10.1038/s41598-019-49897-1.

155. Lowe B.A., Marsh T.L., Isaacs-Cosgrove N., Kirkwood R.N., Kiupel M., Mulks M.H. Microbial Communities in the Tonsils of Healthy Pigs. Vet. Microbiol. 2011;147:346–357. doi: 10.1016/j.vetmic.2010.06.025.

156. De Arriba L.M., Lopez-Serrano S., Galofre-Mila N., Aragon V. Characterisation of Bergeyella Spp. Isolated from the Nasal Cavities of Piglets. Vet. J. 2018;234:1–6. doi: 10.1016/j.tvjl.2018.01.004.

157. Maes D., Sibila M., Kuhnert P., Segalés J., Haesebrouck F., Pieters M. Update on Mycoplasma hyopneumoniae Infections in Pigs: Knowledge Gaps for Improved Disease Control. Transbound. Emerg. Dis. 2018;65:110–124. doi: 10.1111/tbed.12677.

158. Vacca M., Celano G., Calabrese F.M., Portincasa P., Gobbetti M., De Angelis M. The Controversial Role of Human Gut Lachnospiraceae. Microorganisms. 2020;8:573. doi: 10.3390/microorganisms8040573.

159. Gardiner G.E., Metzler-Zebeli B.U., Lawlor P.G. Impact of Intestinal Microbiota on Growth and Feed Efficiency in Pigs: A Review. Microorganisms. 2020;8:1886. doi: 10.3390/microorganisms8121886.

160. Amat S., Lantz H., Munyaka P.M., Willing B.P. Prevotella in Pigs: The Positive and Negative Associations with Production and Health. Microorganisms. 2020;8:1584. doi: 10.3390/microorganisms8101584.

161. Ormerod K.L., Wood D.L.A., Lachner N., Gellatly S.L., Daly J.N., Parsons J.D., Dal'Molin C.G.O., Palfreyman R.W., Nielsen L.K., Cooper M.A., et al. Genomic Characterization of the Uncultured Bacteroidales Family S24-7 Inhabiting the Guts of Homeothermic Animals. Microbiome. 2016;4:36. doi: 10.1186/s40168-016-0181-2.

162. Wexler H.M. Bacteroides: The Good, the Bad, and the Nitty-Gritty. CMR. 2007;20:593–621. doi: 10.1128/CMR.00008-07.

163. Mortensen M.S., Rasmussen M.A., Stokholm J., Brejnrod A.D., Balle C., Thorsen J., Krogfelt K.A., Bisgaard H., Sørensen S.J. Modeling Transfer of Vaginal Microbiota from Mother to Infant in Early Life. eLife. 2021;10:e57051. doi: 10.7554/eLife.57051.

164. Legendre P., Legendre L. Numerical Ecology. 3rd ed. Elsevier; Amsterdam, The Netherlands: 2012. p. 499.

165. Werner J.J., Koren O., Hugenholtz P., DeSantis T.Z., Walters A.W., Caporaso J.G., Angenent L.T., Knight R., Ley E.R. Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. ISME J. 2012;6:94–103. doi: 10.1038/ismej.2011.82.

166. Mandal S., Van Treuren W., White R.A., Eggesbø M., Knight R., Peddada S.D. Analysis of Composition of Microbiomes: A Novel

Method for Studying Microbial Composition. Microb. Ecol. Health Dis. 2015;26:27663. doi: 10.3402/mehd.v26.27663.

167. Jiang L., Amir A., Morton J.T., Heller R., Arias-Castro E., Knight R. Discrete False-Discovery Rate Improves Identification of Differentially Abundant Microbes. mSystems. 2017;2:e00092-17. doi: 10.1128/mSystems.00092-17.

168. Cho DY, Hunter RC, Ramakrishnan VR. The microbiome and chronic rhinosinusitis. Immunol. Allergy Clin. N. Am. 2020;40(2):251–263. doi: 10.1016/j.iac.2019.12.009.

169. Kumpitsch C, Koskinen K, Schöpf V, Moissl-Eichinger C. The microbiome of the upper respiratory tract in health and disease. BMC Biol. 2019;17(1):87. doi: 10.1186/s12915-019-0703-z.

170. Bomar L, Brugger SD, Lemon KP. Bacterial microbiota of the nasal passages across the span of human life. Curr Opin Microbiol. 2018;41:8-14. doi:10.1016/j.mib.2017.10.023.

171. Bond S, McMullen C, Timsit E, Léguillette R. Topography of the respiratory, oral, and guttural pouch bacterial and fungal microbiotas in horses. J. Vet. Intern. Med. 2023;37(1):349–360. doi: 10.1111/jvim.16612.

172. McDaneld TG, Kuehn LA, Keele JW. Microbiome of the upper nasal cavity of beef calves prior to weaning12. J. Anim. Sci. 2019;97(6):2368–2375. doi: 10.1093/jas/skz119.

173. Nicola I, Cerutti F, Grego E, et al. Characterization of the upper and lower respiratory tract microbiota in Piedmontese calves. Microbiome. 2017;5(1):152. doi: 10.1186/s40168-017-0372-5.

174. Soave O, Brand CD. Coprophagy in animals: A review. Cornell Vet. 1991;81(4):357–364.

175. Sansom BF, Gleed PT. The ingestion of sow's faeces by suckling piglets. Br. J. Nutr. 1981;46(3):451–456. doi: 10.1079/bjn19810053.

176. Xiao L, Estellé J, Kiilerich P, et al. A reference gene catalogue of the pig gut microbiome. Nat. Microbiol. 2016;1:16161. doi: 10.1038/nmicrobiol.2016.161.

177. Patel S, Vlasblom AA, Verstappen KM, et al. Differential analysis of longitudinal methicillin-resistant Staphylococcus aureus colonization in

relation to microbial shifts in the nasal microbiome of neonatal piglets. mSystems. 2021;6(4):e0015221. doi: 10.1128/mSystems.00152-21.

178. Strube ML, Hansen JE, Rasmussen S, Pedersen K. A detailed investigation of the porcine skin and nose microbiome using universal and Staphylococcus specific primers. Sci. Rep. 2018;8(1):12751. doi: 10.1038/s41598-018-30689-y.

179. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. 2017;11(12):2639–2643. doi: 10.1038/ismej.2017.119.

180. Shaffer JP, Marotz C, Belda-Ferre P, et al. A comparison of DNA/RNA extraction protocols for high-throughput sequencing of microbial communities. Biotechniques. 2021;70(3):149–159. doi: 10.2144/btn-2020-0153.

181. Ya, W., Kelsey, N. T., Yan, Y. et al. Characterizing microbial community viability with RNA-based high-throughput sequencing, PREPRINT (Version 1) available at Research Square. 10.21203/rs.3.rs-1870950/v1 (2022).

182. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The microbiome and the respiratory tract. Annu. Rev. Physiol. 2016;78:481–504. doi: 10.1146/annurev-physiol-021115-105238.

183. Kennedy KM, de Goffau MC, Perez-Muñoz ME, et al. Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies. Nature. 2023;613(7945):639–649. doi: 10.1038/s41586-022-05546-8.

184. Kraemer JG, Aebi S, Oppliger A, Hilty M. The indoor-air microbiota of pig farms drives the composition of the pig farmers' nasal microbiota in a season-dependent and farm-specific manner. Appl. Environ. Microbiol. 2019;85(9):e0303818. doi: 10.1128/AEM.03038-18.

185. Kraemer JG, Ramette A, Aebi S, Oppliger A, Hilty M. Influence of pig farming on the human nasal microbiota: Key role of airborne microbial communities. Appl. Environ. Microbiol. 2018;84(6):e0247017. doi: 10.1128/AEM.02470-17.

186. Rajagopala SV, Bakhoum NG, Pakala SB, et al. Metatranscriptomics to characterize respiratory virome, microbiome, and host response directly from clinical samples. Cell Rep. Methods. 2021;1(6):100091. doi: 10.1016/j.crmeth.2021.100091.

187. Castro-Nallar E, Shen Y, Freishtat RJ, et al. Integrating microbial and host transcriptomics to characterize asthma-associated microbial communities. BMC Med. Genom. 2015;8:50. doi: 10.1186/s12920-015-0121-1.

188. Campisciano G, Biffi S. Microbiota in vivo imaging approaches to study host-microbe interactions in preclinical and clinical setting. Heliyon. 2022;8(12):e12511. doi: 10.1016/j.heliyon.2022.e12511.

189. Geva-Zatorsky N, Alvarez D, Hudak JE, et al. In vivo imaging and tracking of host-microbiota interactions via metabolic labeling of gut anaerobic bacteria. Nat. Med. 2015;21(9):1091–1100. doi: 10.1038/nm.3929.

190. Lin L, Song J, Li J, et al. Imaging the in vivo growth patterns of bacteria in human gut Microbiota. Gut Microbes. 2021;13(1):1960134. doi: 10.1080/19490976.2021.1960134.

191. Kim YS, Han D, Mo JH, Kim YM, Kim DW, Choi HG, Park JW, Shin HW. Antibiotic-dependent relationships between the nasal microbiome and secreted proteome in nasal polyps. Allergy Asthma Immunol. Res. 2021;13(4):589–608. doi: 10.4168/aair.2021.13.4.589.

192. Mark Welch JL, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG. Biogeography of a human oral microbiome at the micron scale. Proc. Natl. Acad. Sci. U. S. A. 2016;113(6):E791–E800. doi: 10.1073/pnas.1522149113.

193. Mark Welch JL, Ramírez-Puebla ST, Borisy GG. Oral microbiome geography: Micron-scale habitat and niche. Cell Host Microbe. 2020;28(2):160–168. doi: 10.1016/j.chom.2020.07.009.

194. Kuhar HN, Tajudeen BA, Mahdavinia M, et al. Relative abundance of nasal microbiota in chronic rhinosinusitis by structured histopathology. Int. Forum Allergy Rhinol. 2018;8(12):1430–1437. doi: 10.1002/alr.22192.

195. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17(1):10–12. doi: 10.14806/ej.17.1.200.

196. Nathan, H., Per-Gunnar, M., Yoel, S. & Mark, T. An algorithm for the principal component analysis of large data sets. arXiv:1007.5510.

197. Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. Nat. Commun. 2020;11(1):3514. doi: 10.1038/s41467-020-17041-7.

198. Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ. W-IQ-TREE: A fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 2016;44(W1):W232–W235. doi: 10.1093/nar/gkw256.

199. RStudio|Open Source & Professional Software for Data Science Teams. https://rstudio.com/ (2023).

200. Bisanz J. Tutorial: integrating QIIME2 and R for data visualization and analysis using qiime2R (v0.99.6). 2021.

201. Wickham H. Reshaping data with the reshape package. J. Stat. Softw. 2007;21(12):1–20. doi: 10.18637/jss.v021.i12.

202. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2009.

203. Wickham H, et al. Welcome to the tidyverse. J. Open Source Softw. 2019;4:1686. doi: 10.21105/joss.01686.

204. Yu G, Smith D, Zhu H, Guan Y, Lam TT. ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. 2017;8:28–36. doi: 10.1111/2041-210X.12628.

205. Man WH, de Steenhuijsen Piters WA, Bogaert D. The Microbiota of the Respiratory Tract: Gatekeeper to Respiratory Health. Nat Rev Microbiol. 2017;15(5):259-270. Doi:10.1038/Nrmicro.2017.14.

206. Chaplin DD. Overview of the immune response. Journal of Allergy and Clinical Immunology. 2010 Feb;125(2):S3–23.

207. Bartlett BL, Pellicane AJ, Tyring SK. Vaccine immunology. Dermatologic Therapy. 2009 Mar;22(2):104–9.

208. Andre F, Booy R, Bock H, Clemens J, Datta S, John T, et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. Bull World Health Org. 2008 Feb 1;86(2):140–6.

209. Lynn DJ, Benson SC, Lynn MA, Pulendran B. Modulation of immune responses to vaccination by the microbiota: implications and potential mechanisms. Nat Rev Immunol. 2022 Jan;22(1):33–46.

210. Roth JA. Mechanistic Bases for Adverse Vaccine Reactions and Vaccine Failures. In: Advances in Veterinary Medicine [Internet]. Elsevier; 1999 [cited 2024 May 21]. p. 681–700. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0065351999800536

211. El T, Bj T, Tb B. Comparison of antibody production, lymphocyte stimulation, and protection induced by four commercial Mycoplasma hyopneumoniae bacterins. Swine Health and Production. 1998;6(3).

212. De Jong SE, Olin A, Pulendran B. The Impact of the Microbiome on Immunity to Vaccination in Humans. Cell Host & Microbe. 2020 Aug;28(2):169–79.

213. Seong H, Choi BK, Han YH, Kim JH, Gim JA, Lim S, et al. Gut microbiota as a potential key to modulating humoral immunogenicity of new platform COVID-19 vaccines. Sig Transduct Target Ther. 2023 May 3;8(1):178.

214. Rojas C, Gálvez-Jirón F, De Solminihac J, Padilla C, Cárcamo I, Villalón N, et al. Crosstalk between Body Microbiota and the Regulation of Immunity. Xu Z, editor. Journal of Immunology Research. 2022 May 19;2022:1–13.

215. Ruck CE, Odumade OA, Smolen KK. Vaccine Interactions With the Infant Microbiome: Do They Define Health and Disease? Front Pediatr. 2020 Nov 26;8:565368.

216. Harris VC, Armah G, Fuentes S, Korpela KE, Parashar U, Victor JC, et al. Significant Correlation Between the Infant Gut Microbiome and Rotavirus Vaccine Response in Rural Ghana. The Journal of Infectious Diseases. 2017 Jan 1;215(1):34–41.

217. Harris V, Ali A, Fuentes S, Korpela K, Kazi M, Tate J, et al. Rotavirus vaccine response correlates with the infant gut microbiota composition in Pakistan. Gut Microbes. 2018 Mar 4;9(2):93–101.

218. Eloe-Fadrosh EA, McArthur MA, Seekatz AM, Drabek EF, Rasko DA, Sztein MB, et al. Impact of Oral Typhoid Vaccination on the Human Gut Microbiota and Correlations with S. Typhi-Specific Immunological Responses. Gilbert JA, editor. PLoS ONE. 2013 Apr 24;8(4):e62026.

219. Huda MN, Lewis Z, Kalanetra KM, Rashid M, Ahmad SM, Raqib R, Qadri F, Underwood MA, Mills DA, Stephensen CB. Stool microbiota and vaccine responses of infants. Pediatrics. 2014 Aug;134(2):e362-72. doi: 10.1542/peds.2013-3937. Epub 2014 Jul 7. PMID: 25002669; PMCID: PMC4187229.

220. Huda MN, Ahmad SM, Alam MJ, Khanam A, Kalanetra KM, Taft DH, Raqib R, Underwood MA, Mills DA, Stephensen CB. Bifidobacterium Abundance in Early Infancy and Vaccine Response at 2 Years of Age. Pediatrics. 2019 Feb;143(2):e20181489. doi: 10.1542/peds.2018-1489. PMID: 30674610; PMCID: PMC6361348.

221. Fix J, Chandrashekhar K, Perez J, Bucardo F, Hudgens MG, Yuan L, et al. Association between Gut Microbiome Composition and Rotavirus Vaccine Response among Nicaraguan Infants. The American Journal of Tropical Medicine and Hygiene. 2020 Jan 8;102(1):213–9.

222. Ng SC, Peng Y, Zhang L, Mok CK, Zhao S, Li A, et al. Gut microbiota composition is associated with SARS-CoV-2 vaccine immunogenicity and adverse events. Gut. 2022 Jun;71(6):1106–16.

223. Tang B, Tang L, He W, Jiang X, Hu C, Li Y, et al. Correlation of gut microbiota and metabolic functions with the antibody response to the BBIBP-CorV vaccine. Cell Reports Medicine. 2022 Oct;3(10):100752.

224. Alexander JL, Mullish BH, Danckert NP, Liu Z, Olbei ML, Saifuddin A, et al. The gut microbiota and metabolome are associated with diminished COVID-19 vaccine-induced antibody responses in immunosuppressed inflammatory bowel disease patients. eBioMedicine. 2023 Feb;88:104430.

225. Hagan T, Cortese M, Rouphael N, Boudreau C, Linde C, Maddur MS, et al. Antibiotics-Driven Gut Microbiome Perturbation Alters Immunity to Vaccines in Humans. Cell. 2019 Sep;178(6):1313-1328.e13.

226. Munyaka PM, Kommadath A, Fouhse J, Wilkinson J, Diether N, Stothard P, et al. Characterization of whole blood transcriptome and early-life fecal microbiota in high and low responder pigs before, and after vaccination for Mycoplasma hyopneumoniae. Vaccine. 2019 Mar;37(13):1743–55.

227. Munyaka PM, Blanc F, Estellé J, Lemonnier G, Leplat JJ, Rossignol MN, et al. Discovery of Predictors of Mycoplasma hyopneumoniae Vaccine Response Efficiency in Pigs: 16S rRNA Gene Fecal Microbiota Analysis. Microorganisms. 2020 Jul 29;8(8):1151.

228. Sanglard LP, Schmitz-Esser S, Gray KA, Linhares DCL, Yeoman CJ, Dekkers JCM, et al. Investigating the relationship between vaginal microbiota and host genetics and their impact on immune response and farrowing traits in commercial gilts. J Animal Breeding Genetics. 2020 Jan;137(1):84–102.

229. Obregon-Gutierrez P, Bonillo-Lopez L, Correa-Fiz F, Sibila M, Segalés J, Kochanowski K, et al. Gut-associated microbes are present and active in the pig nasal cavity. Sci Rep. 2024 Apr 11;14(1):8470.

230. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30.

231. Spearman, C. The proof and measurement of association between two things. Am J. Psychol.15, 72–101. 10.2307/1412159 (1904).

232. R Core Team (2023). _R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

233. Mantel N. The detection of disease clustering and a generalized regression approach. Cancer Res. 1967;27:209–220.

234. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, Tickle TL, Weingart G, Ren B, Schwager EH, Chatterjee S, Thompson KN, Wilkinson JE, Subramanian A, Lu Y, Waldron L, Paulson JN, Franzosa EA, Bravo HC, Huttenhower C (2021). Multivariable Association Discovery in Population-scale Meta-omics Studies. PLoS Computational Biology, 17(11):e1009442.

235. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. Metagenomic biomarker discovery and explanation. Genome Biol. 2011 Jun 24;12(6):R60. doi: 10.1186/gb-2011-12-6-r60. PMID: 21702898; PMCID: PMC3218848.

236. Dohoo, I.; Martin, W.; Stryhn, H. 2009. Veterinary Epidemiologic Research; MacPike, S.M., Ed.; VER Inc.: Charlottetown, PE, Canada.

237. Zhang, Y., Zhang, Y., Liu, F. et al. Mechanisms and applications of probiotics in prevention and treatment of swine diseases. Porc Health Manag 9, 5 (2023). https://doi.org/10.1186/s40813-022-00295-6.

238. López-Serrano S, Galofré-Milà N, Costa-Hurtado M, Pérez-de-Rozas AM, Aragon V. Heterogeneity of Moraxella isolates found in the nasal cavities of piglets. BMC Vet Res. 2020 Dec;16(1):28.

239. Knecht D, Cholewińska P, Jankowska-Mąkosa A, Czyż K. Development of Swine's Digestive Tract Microbiota and Its Relation to Production Indices—A Review. Animals. 2020 Mar 21;10(3):527.

240. Chen, C., Zhou, Y., Fu, H., Xiong, X., Fang, S., Jiang, H., Wu, J., Yang, H., Gao, J., & Huang, L. (2021). Expanded catalog of microbial genes

and metagenome-assembled genomes from the pig gut microbiome. Nature communications, 12(1), 1106. https://doi.org/10.1038/s41467-021-21295-0.

241. Camille Martin-Gallausiaux,LudovicaMarinelli,HerveH.Blottiere,PierreLarraufie,Nicolas La- paque. ShortChainFattyAcids–mechanismsandfunctionalimportanceinthegut.Proceedingsof the NutritionSociety,2021,80(1),pp.37-49. 10.1017/S0029665120006916.

242. Sun, J., Du, L., Li, X. et al. Identification of the core bacteria in rectums of diarrheic and non-diarrheic piglets. Sci Rep 9, 18675 (2019). https://doi.org/10.1038/s41598-019-55328-y.

243. Wang, J., Zhu, N., Su, X., Gao, Y., & Yang, R. (2023). Gut-Microbiota-Derived Metabolites Maintain Gut and Systemic Immune Homeostasis. Cells, 12(5), 793. https://doi.org/10.3390/cells12050793.

244. Klein U, Földi D, Belecz N, et al. Antimicrobial susceptibility profiles of Mycoplasma hyorhinis strains isolated from five European countries between 2019 and 2021. PLoS One. 2022;17(8):e0272903. Published 2022 Aug 11. doi:10.1371/journal.pone.0272903.

245. Silva, A. P. S. P., Almeida, M., Michael, A., Rahe, M. C., Siepker, C., Magstadt, D. R., Piñeyro, P., Arruda, B. L., Macedo, N. R., Sahin, O., Gauger, P. C., Krueger, K. M., Mugabi, R., Streauslin, J. S., Trevisan, G., Linhares, D. C. L., Silva, G. S., Fano, E., Main, R. G., Schwartz, K. J., … Clavijo, M. J. (2023). Detection and disease diagnosis trends (2017-2022) for Streptococcus suis, Glaesserella parasuis, Mycoplasma hyorhinis, Actinobacillus suis and Mycoplasma hyosynoviae at Iowa State University Veterinary Diagnostic Laboratory. BMC veterinary research, 19(1), 268. https://doi.org/10.1186/s12917-023-03807-w.

246. Morita T, Ohiwa S, Shimada A, Kazama S, Yagihashi T, Umemura T. Intranasally inoculated Mycoplasma hyorhinis causes eustachitis in pigs. Vet Pathol. 1999;36(2):174-178. doi:10.1354/vp.36-2-174.

247. Renzhammer, R.; Auer, A.; Loncaric, I.; Entenfellner, A.; Dimmel, K.; Walk, K.; Rümenapf, T.; Spergser, J.; Ladinig, A. Retrospective Analysis of the Detection of Pathogens Associated with the Porcine Respiratory Disease Complex in Routine Diagnostic Samples from Austrian Swine Stocks. Vet. Sci. 2023, 10, 601. https://doi.org/10.3390/vetsci10100601.

248. Thakor JC, Sahoo M, Karam Pal Singh, et al. Porcine respiratory disease complex (PRDC) in Indian pigs: a slaughterhouse survey. Vet Ital. 2023;59(1):23-38. Published 2023 Mar 31. doi:10.12834/VetIt.2935.20591.2.

249. Tocqueville, V., Ferr´e, S., Nguyen, N.H.P., Kempf, I., Marois-Cr´ehana, C., 2014. Multilocus sequence typing of Mycoplasma hyorhinis strains identified by a real-time TaqMan PCR assay. J. Clin. Microbiol. 52, 1664–1671. https://doi.org/10.1128/ JCM.03437-13.

250. Trüeb, B., Catelli, E., Luehrs, A., Nathues, H., Kuhnert, P., 2016. Genetic variability and limited clonality of Mycoplasma hyorhinis in pig herds. Vet. Microbiol. 191, 9–14. https://doi.org/10.1016/j.vetmic.2016.05.015.

251. Foldi, ¨ D., Beko, ˝ K., Felde, O., Kreizinger, Z., Kovacs, ´ A.B., ´ Toth, ´ F., B´ anyai, K., Kiss, K., Biksi, I., Gyuranecz, M., 2020. Genotyping Mycoplasma hyorhinis by multi-locus sequence typing and multiple-locus variable-number tandem-repeat analysis. Vet. Microbiol. 249 https://doi.org/10.1016/j.vetmic.2020.108836.

252. Bünger M, Posch M, Wiesauer J, et al. A core genome multilocus sequence typing scheme for Mycoplasma hyorhinis. Vet Microbiol. 2021;262:109249. doi:10.1016/j.vetmic.2021.109249.

253. Dos Santos, L.F., Clavijo, M.J., Sreevatsan, S., Rovira, A., Moreira, M.A., Pieters, M., 2015. Genotyping of Mycoplasma hyorhinis using multiple-locus variable number tandem repeat analysis. J. Microbiol. Methods 111, 87-92.

254. Kinne J, Johannsen U, Neumann R, Mehlhorn G, Pfutzner H. 1991. The pathology and pathogenesis of experimental Mycoplasma hyorhinis infection of piglets with and without thermomotoric stress. 1. Pathologico-anatomic, histologic and immunomorphologic study results. Zentralbl Veterinarmed A 38:306–320. https://doi.org/10.1111/j.1439-0442.1991.tb01017.x.

255. Palzer, A.; Ritzmann, M.; Wolf, G.; Heinritzi, K. Associations between Pathogens in Healthy Pigs and Pigs with Pneumonia. Vet. Rec. 2008, 162, 267–271.

256. Christine Citti; Kim S. Wise (1995). Mycoplasma hyorhinis vlp gene transcription: critical role in phase variation and expression of surface lipoproteins. , 18(4), 649–660. doi:10.1111/j.1365-2958.1995.mmi_18040649.x.

257. Henrique Bunselmeyer Ferreira; Luiza Amaral de Castro. A preliminary survey of M. hyopneumoniae virulence factors based on comparative genomic analysis. Genet. Mol. Biol. 30 (1 suppl). 2007. https://doi.org/10.1590/S1415-47572007000200012.

258. Tavares BADR, Paes JA, Zaha A, Ferreira HB. Reannotation of Mycoplasma hyopneumoniae hypothetical proteins revealed novel potential virulence factors. Microb Pathog. 2022 Jan;162:105344. doi: 10.1016/j.micpath.2021.105344. Epub 2021 Dec 2. PMID: 34864146.

259. Gaeta, N. C., de Sá Guimarães, A. M., Timenetsky, J., Clouser, S., Gregory, L., & Ganda, E. (2024). Comparative genomic analysis of Brazilian Mesomycoplasma ovipneumoniae strains revealed genomic differences associated with the geographic origin and health status and mutations in the gyrA. Veterinary microbiology, 295, 110158. https://doi.org/10.1016/j.vetmic.2024.110158.

260. Maksimović, Z., Rifatbegović, M., Loria, G. R., & Nicholas, R. A. J. (2022). Mycoplasma ovipneumoniae: A Most Variable Pathogen. Pathogens (Basel, Switzerland), 11(12), 1477. https://doi.org/10.3390/pathogens11121477.

261. Ionas, G.; AJ, M.; Alley, M.; Clarke, J.; Robinson, A.; Marshall, R. Colonisation of the Respiratory Tract of Lambs by Strains of Mycoplasma ovipneumoniae. Vet. Microbiol. 1985, 10, 533–539.

262. Liu W, Xiao S, Li M, et al. Comparative genomic analyses of Mycoplasma hyopneumoniae pathogenic 168 strain and its high-passaged attenuated strain. BMC Genomics. 2013;14:80. Published 2013 Feb 5. doi:10.1186/1471-2164-14-80.

263. Vilei, E. M., & Frey, J. (2001). Genetic and biochemical characterization of glycerol uptake in mycoplasma mycoides subsp. mycoides SC: its impact on H(2)O(2) production and virulence. Clinical and diagnostic laboratory immunology, 8(1), 85–92. https://doi.org/10.1128/CDLI.8.1.85-92.2001.

264. Xiong Q, Zhang B, Wang J, Ni B, Ji Y, Wei Y, et al. Characterization of the role in adherence of Mycoplasma hyorhinis variable lipoproteins containing different repeat unit copy numbers. Veterinary Microbiology. 2016 Dec;197:39–46.

265. Li, J., Wang, J., Shao, J., Li, Y., Yu, Y., Shao, G., Feng, Z., & Xiong, Q. (2022). The variable lipoprotein family participates in the interaction of Mycoplasma hyorhinis with host extracellular matrix and plasminogen.

Veterinary microbiology, 265, 109310. https://doi.org/10.1016/j.vetmic.2021.109310.

266. Yogev D, Rosengarten R, Watson-McKown R, Wise KS. Molecular basis of Mycoplasma surface antigenic variation: a novel set of divergent genes undergo spontaneous mutation of periodic coding regions and 5' regulatory sequences. EMBO J. 1991;10(13):4069-4079. doi:10.1002/j.1460-2075.1991.tb04983.x.

267. Rosengarten, R; Wise, K. (1990). Phenotypic switching in mycoplasmas: phase variation of diverse surface lipoproteins. Science, 247(4940), 315–318. doi:10.1126/science.1688663.

268. Ishida S, Tien le HT, Osawa R, Tohya M, Nomoto R, Kawamura Y, Takahashi T, Kikuchi N, Kikuchi K, Sekizaki T. 2014.Development of an appropriate PCR system for the reclassification of Streptococcus suis. J Microbiol Methods. 107:66-70.

269. Wick RR, Judd LM, Gorrie CL, Holt KE. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom. 2017;3(10):e000132. Published 2017 Sep 14. doi:10.1099/mgen.0.000132.

270. Wouter De Coster, Rosa Rademakers, NanoPack2: population-scale evaluation of long-read sequencing data, Bioinformatics, Volume 39, Issue 5, May 2023.

271. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data.

272. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 2017.

273. Koren S, Walenz BP, Berlin K, Miller JR, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Research. (2017). doi:10.1101/gr.215087.116.

274. O'Leary NA, Cox E, Holmes JB, Anderson WR, Falk R, Hem V, Tsuchiya MTN, Schuler GD, Zhang X, Torcivia J, Ketter A, Breen L, Cothran J, Bajwa H, Tinne J, Meric PA, Hlavina W, Schneider VA. Exploring and retrieving sequence and metadata for species across the tree of life with NCBI Datasets. Sci Data. 2024 Jul 5;11(1):732. doi: 10.1038/s41597-024-03571-y. PMID: 38969627; PMCID: PMC11226681.

275. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014 Jul 15;30(14):2068-9. PMID:24642063.

276. Carlos S. Casimiro-Soriguer, Antonio Muñoz-Mérida, Antonio J. Pérez-Pulido (2017) Sma3s: a universal tool for easy functional annotation of proteomes and transcriptomes. Proteomics. 17(12).

277. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 51:D523–D531.

278. Andrew J. Page, Carla A. Cummins, Martin Hunt, Vanessa K. Wong, Sandra Reuter, Matthew T. G. Holden, Maria Fookes, Daniel Falush, Jacqueline A. Keane, Julian Parkhill, 'Roary: Rapid large-scale prokaryote pan genome analysis', Bioinformatics, 2015;31(22):3691-3693 doi:10.1093/bioinformatics/btv421.

279. Martinez Arbizu, P. (2020). pairwiseAdonis: Pairwise multilevel comparison using adonis. R package version 0.4.

280. Liu, B., Zheng, D., Zhou, S., Chen, L., & Yang, J. (2022). VFDB 2022: a general classification scheme for bacterial virulence factors. Nucleic acids research, 50(D1), D912–D917. https://doi.org/10.1093/nar/gkab1107.

281. Leal Zimmer, F. M. A., Paes, J. A., Zaha, A., & Ferreira, H. B. (2020). Pathogenicity & virulence of Mycoplasma hyopneumoniae. Virulence, 11(1), 1600–1622. https://doi.org/10.1080/21505594.2020.1842659.

282. Teufel, F., Almagro Armenteros, J.J., Johansen, A.R. et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 40, 1023–1025 (2022). https://doi.org/10.1038/s41587-021-01156-3.

283. Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics, 18(1), 50–60.

284. Alcock et al. 2023. CARD 2023: Expanded Curation, Support for Machine Learning, and Resistome Prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Research, 51, D690-D699.

285. National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; [1988] – [cited 2017 Apr 06]. Available from: https://www.ncbi.nlm.nih.gov/.

286. Florensa AF, Kaas RS, Clausen PTLC, Aytan-Aktug D, Aarestrup FM. ResFinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. Microb Genom. 2022 Jan;8(1):000748. doi: 10.1099/mgen.0.000748.

287. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments", Andrew J. Page, Ben Taylor, Aidan J. Delaney, Jorge Soares, Torsten Seemann, Jacqueline A. Keane, Simon R. Harris, Microbial Genomics 2(4), (2016).

288. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. Bioinformatics. 2011;27(15):2156-2158. doi:10.1093/bioinformatics/btr330.

289. Knaus BJ, Grünwald NJ (2017). "VCFR: a package to manipulate and visualize variant call format data in R." Molecular Ecology Resources, 17(1), 44–53. ISSN 757, https://dx.doi.org/10.1111/1755-0998.12549.

290. Obenchain V, Lawrence M, Carey V, Gogarten S, Shannon P, Morgan M (2014). "VariantAnnotation: a Bioconductor package for exploration and annotation of genetic variants." Bioinformatics, 30(14), 2076-2078. doi:10.1093/bioinformatics/btu168.

291. Gautier-Bouchardon A. V. (2018). Antimicrobial Resistance in Mycoplasma spp. Microbiology spectrum, 6(4), 10.1128/microbiolspec.arba-0030-2018. https://doi.org/10.1128/microbiolspec.ARBA-0030-2018.

292. Li J, Wei Y, Wang J, Li Y, Shao G, Feng Z, Xiong Q. Characterization of Mutations in DNA Gyrase and Topoisomerase IV in Field Strains and In Vitro Selected Quinolone-Resistant Mycoplasma hyorhinis Mutants. Antibiotics (Basel). 2022 Apr 7;11(4):494. doi: 10.3390/antibiotics11040494. PMID: 35453245; PMCID: PMC9024574.

293. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30(22):3276-3278. doi:10.1093/bioinformatics/btu531.

294. Frederic Bertels, Olin K. Silander, Mikhail Pachkov, Paul B. Rainey, Erik van Nimwegen, Automated Reconstruction of Whole-Genome Phylogenies from Short-Sequence Reads , Molecular Biology and Evolution, Volume 31, Issue 5, May 2014, Pages 1077–1088, https://doi.org/10.1093/molbev/msu088.

295. SWITZER WP. Studies on infectious atrophic rhinitis. IV. Characterization of a pleuropneumonia-like organism isolated from the nasal cavities of swine. Am J Vet Res. 1955;16(61 Part 1):540-544.

296. Bui Quang Minh, Heiko A Schmidt, Olga Chernomor, Dominik Schrempf, Michael D Woodhams, Arndt von Haeseler, Robert Lanfear, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era, Molecular Biology and Evolution, Volume 37, Issue 5, May 2020, Pages 1530–1534, https://doi.org/10.1093/molbev/msaa015.

297. Letunic I and Bork P (2024) Nucleic Acids Res doi: 10.1093/nar/gkae268 Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool.

298. Pritchard et al. (2016) 'Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens' Anal. Methods, 2016,8, 12-24 DOI: 10.1039/C5AY02550H.

299. Croucher N. J., Page A. J., Connor T. R., Delaney A. J., Keane J. A., Bentley S. D., Parkhill J., Harris S.R. 'Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins'. doi:10.1093/nar/gku1196, Nucleic Acids Research, 2014.

300. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009 Aug 1;25(15):1972-3. doi: 10.1093/bioinformatics/btp348. Epub 2009 Jun 8. PMID: 19505945; PMCID: PMC2712344.

301. Revell L (2024). "phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things)." PeerJ, 12, e16505. doi:10.7717/peerj.16505.

302. Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(D1):D638-D646. doi:10.1093/nar/gkac1000.

303. Ferrarini MG, Siqueira FM, Mucha SG, Palama TL, Jobard É, Elena-Herrmann B, R Vasconcelos AT, Tardy F, Schrank IS, Zaha A, Sagot MF. Insights on the virulence of swine respiratory tract mycoplasmas through genome-scale metabolic modeling. BMC Genomics. 2016 May 13;17:353. doi: 10.1186/s12864-016-2644-z. PMID: 27178561; PMCID: PMC4866288.

304. Heirendt L, Arreckx S, Pfau T, et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. Nat Protoc. 2019;14(3):639-702. doi:10.1038/s41596-018-0098-2.

305. Pobeguts OV, Galaymina MA, Sikamov KV, et al. Unraveling the adaptive strategies of Mycoplasma hominis through proteogenomic profiling of clinical isolates. Front Cell Infect Microbiol. 2024;14:1398706. Published 2024 May 2. doi:10.3389/fcimb.2024.1398706.

306. Heusipp G, Fälker S, Schmidt MA. DNA adenine methylation and bacterial pathogenesis. Int J Med Microbiol. 2007;297(1):1-7. doi:10.1016/j.ijmm.2006.10.002.

307. Redder P, Linder P. DEAD-box RNA helicases in gram-positive RNA decay. Methods Enzymol. 2012;511:369-383. doi:10.1016/B978-0-12-396546-2.00017-6.

308. Hausmann S, Gonzalez D, Geiser J, Valentini M. The DEAD-box RNA helicase RhlE2 is a global regulator of Pseudomonas aeruginosa lifestyle and pathogenesis. Nucleic Acids Res. 2021;49(12):6925-6940. doi:10.1093/nar/gkab503.

309. Citti, C., Kim, M.F., Wise, K.S., 1997. Elongated versions of Vlp surface lipoproteins protect Mycoplasma hyorhinis escape variants from growth-inhibiting host antibodies. Infect. Immun. 65, 1773–1785.

310. Simmons WL, Dybvig K. The Vsa proteins modulate susceptibility of Mycoplasma pulmonis to complement killing, hemadsorption, and adherence to polystyrene. Infect Immun. 2003;71(10):5733-5738. doi:10.1128/IAI.71.10.5733-5738.2003.

311. Shaw BM, Simmons WL, Dybvig K. The Vsa shield of Mycoplasma pulmonis is antiphagocytic. Infect Immun. 2012;80(2):704-709. doi:10.1128/IAI.06009-11.

312. Bolland JR, Dybvig K. Mycoplasma pulmonis Vsa proteins and polysaccharide modulate adherence to pulmonary epithelial cells. FEMS Microbiol Lett. 2012;331(1):25-30. doi:10.1111/j.1574-6968.2012.02551.x.

313. Citti, C., Nouvel, L. X., & Baranowski, E. (2010). Phase and antigenic variation in mycoplasmas. Future microbiology, 5(7), 1073–1085. https://doi.org/10.2217/fmb.10.71.

314. Chopra-Dewasthaly, R., Baumgartner, M., Gamper, E., Innerebner, C., Zimmermann, M., Schilcher, F., Tichy, A., Winter, P., Jechlinger, W., Rosengarten, R., & Spergser, J. (2012). Role of Vpma phase variation in Mycoplasma agalactiae pathogenesis. FEMS immunology and medical microbiology, 66(3), 307–322. https://doi.org/10.1111/j.1574-695X.2012.01010.x.

315. Yogev D, Watson-McKown R, Rosengarten R, Im J, Wise KS. Increased structural and combinatorial diversity in an extended family of genes encoding Vlp surface proteins of Mycoplasma hyorhinis. J Bacteriol. 1995;177(19):5636-5643. doi:10.1128/jb.177.19.5636-5643.1995.

316. Murray GL, Bradshaw CS, Bissessor M, Danielewski J, Garland SM, Jensen JS, Fairley CK, Tabrizi SN. Increasing Macrolide and Fluoroquinolone Resistance in Mycoplasma genitalium. Emerg Infect Dis. 2017 May;23(5):809-812. doi: 10.3201/eid2305.161745. PMID: 28418319; PMCID: PMC5403035.

317. Norsigian, C. J., Fang, X., Palsson, B. O., & Monk, J. M. (2020). Pangenome Flux Balance Analysis Toward Panphenomes. In H. Tettelin (Eds.) et. al., The Pangenome: Diversity, Dynamics and Evolution of Genomes. (pp. 219–232). Springer.

318. Valles-Colomer, M., Blanco-Míguez, A., Manghi, P., Asnicar, F., Dubois, L., Golzato, D., Armanini, F., Cumbo, F., Huang, K. D., Manara, S., Masetti, G., Pinto, F., Piperni, E., Punčochář, M., Ricci, L., Zolfo, M., Farrant, O., Goncalves, A., Selma-Royo, M., Binetti, A. G., … Segata, N. (2023). The person-to-person transmission landscape of the gut and oral microbiomes. Nature, 614(7946), 125–135. https://doi.org/10.1038/s41586-022-05620-1.

319. Bonillo-Lopez, L., Rouam el Khatab, O., Obregon-Gutierrez, P., Almansa Fernandez-Villacañas, J., Florez-Sarasa, I., Correa-Fiz, F., Sibila, M., Aragon#, V., Kochanowski#, K. In vitro metabolic interaction network of a rationally designed nasal microbiota community. iScience (accepted).

320. Poor AP, Moreno LZ, Monteiro MS, et al. Vaginal microbiota signatures in healthy and purulent vulvar discharge sows. Sci Rep. 2022;12(1):9106. Published 2022 Jun 1. doi:10.1038/s41598-022-13090-8.

321. Wei M, Flowers L, Knight SAB, et al. Harnessing diversity and antagonism within the pig skin microbiota to identify novel mediators of

colonization resistance to methicillin-resistant Staphylococcus aureus. mSphere. 2023;8(4):e0017723. doi:10.1128/msphere.00177-23.

322. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486(7402):207-214. Published 2012 Jun 13. doi:10.1038/nature11234.

323. Wang, J., Li, C., Nesengani, L. T., Gong, Y., Zhang, S., & Lu, W. (2017). Characterization of vaginal microbiota of endometritis and healthy sows using high-throughput pyrosequencing of 16S rRNA gene. Microbial pathogenesis, 111, 325–330. https://doi.org/10.1016/j.micpath.2017.08.030.

324. Bugenyi AW, Lee MR, Choi YJ, et al. Oropharyngeal, proximal colonic, and vaginal microbiomes of healthy Korean native black pig gilts. BMC Microbiol. 2023;23(1):3. Published 2023 Jan 5. doi:10.1186/s12866-022-02743-3.

325. Zhang L, Wang L, Dai Y, et al. Effect of Sow Intestinal Flora on the Formation of Endometritis. Front Vet Sci. 2021;8:663956. Published 2021 Jun 18. doi:10.3389/fvets.2021.663956.

326. Ferretti P, Pasolli E, Tett A, et al. Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome. Cell Host Microbe. 2018;24(1):133-145.e5. doi:10.1016/j.chom.2018.06.005.

327. Ober RA, Thissen JB, Jaing CJ, Cino-Ozuna AG, Rowland RRR, Niederwerder MC. Increased microbiome diversity at the time of infection is associated with improved growth rates of pigs after co-infection with porcine reproductive and respiratory syndrome virus (PRRSV) and porcine circovirus type 2 (PCV2). Vet Microbiol. 2017 Sep;208:203-211. doi: 10.1016/j.vetmic.2017.06.023. Epub 2017 Aug 18. PMID: 28888639.

328. Niederwerder MC, Jaing CJ, Thissen JB, Cino-Ozuna AG, McLoughlin KS, Rowland RR. Microbiome associations in pigs with the best and worst clinical outcomes following co-infection with porcine reproductive and respiratory syndrome virus (PRRSV) and porcine circovirus type 2 (PCV2). Vet Microbiol. 2016;30(188):1–11. 10.1016/j.vetmic.2016.03.008.

329. Segata N, Haake SK, Mannon P, et al. Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces,

tonsils, throat and stool samples. Genome Biol. 2012;13(6):R42. Published 2012 Jun 14. doi:10.1186/gb-2012-13-6-r42.

330. Biswas K, Hoggard M, Jain R, Taylor MW, Douglas RG. The nasal microbiota in health and disease: variation within and between subjects. Front Microbiol. 2015;9:134. Published 2015 Mar 2. doi:10.3389/fmicb.2015.00134.

331. Chen, CH., Liou, ML., Lee, CY. et al. Diversity of nasal microbiota and its interaction with surface microbiota among residents in healthcare institutes. Sci Rep 9, 6175 (2019). https://doi.org/10.1038/s41598-019-42548-5.

332. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. Nature. 2012;489(7415):220-230. doi:10.1038/nature11550.

333. Van Hul M, Cani PD, Petitfils C, De Vos WM, Tilg H, El-Omar EM. What defines a healthy gut microbiome?. Gut. 2024;73(11):1893-1908. Published 2024 Oct 7. doi:10.1136/gutjnl-2024-333378.

334. Wypych TP, Wickramasinghe LC, Marsland BJ. The influence of the microbiome on respiratory health. Nat Immunol. 2019;20(10):1279-1290. doi:10.1038/s41590-019-0451-9.

335. King SJ, Leigh JA, Heath PJ, Luque I, Tarradas C, Dowson CG, et al. Development of a multilocus sequence typing scheme for the pig pathogen Streptococcus suis: identification of virulent clones and potential capsular serotype exchange. J Clin Microbiol. 2002. October 1;40(10): 3671–80. 10.1128/JCM.40.10.3671-3680.2002.

336. Olvera A, Cerda-Cuellar M, Aragon V. Study of the population structure of Haemophilus parasuis by multilocus sequence typing. Microbiology. 2006. December 1;152(Pt 12): 3683–90. 10.1099/mic.0.29254-0.

337. Asnicar F, Manara S, Zolfo M, Truong DT, Scholz M, Armanini F, Ferretti P, Gorfer V, Pedrotti A, Tett A, Segata N.2017.Studying Vertical Microbiome Transmission from Mothers to Infants by Strain-Level Metagenomic Profiling. mSystems2:e00164-16.https://doi.org/10.1128/mSystems.00164-16.

338. Truong DT, Tett A, Pasolli E, Huttenhower C, Segata N. Microbial strain-level population structure and genetic diversity from metagenomes. Genome Res. 2017;27(4):626-638. doi:10.1101/gr.216242.116.

339. Ichinohe T, Pang IK, Kumamoto Y, Peaper DR, Ho JH, Murray TS, et al (2011) Microbiota regulates immune defense against respiratory tract influenza A virus infection. Proceedings of the National Academy of Sciences of the United States of America 108: 5354–5359. doi: 10.1073/pnas.1019378108 PMID: 21402903.

340. Oh JZ, Ravindran R, Chassaing B, et al. TLR5-mediated sensing of gut microbiota is necessary for antibody responses to seasonal influenza vaccination. Immunity. 2014;41(3):478-492. doi:10.1016/j.immuni.2014.08.009.

341. Yaqoob P (2014) Ageing, immunity and influenza: a role for probiotics? The Proceedings of the Nutrition Society 73: 309–317. doi: 10.1017/S0029665113003777 PMID: 24300282.

342. Chac D, Bhuiyan TR, Saha A, Alam MM, Salma U, Jahan N, Chowdhury F, Khan AI, Ryan ET,,LaRocque R,Harris JB,Qadri F, Weil AA. 2021. Gut Microbiota and Development of Vibrio cholerae-Specific Long-Term Memory B Cells in Adults after Whole-Cell Killed Oral Cholera Vaccine. Infect Immun 89:. https://doi.org/10.1128/iai.00217-21.

343. Hall JA, Bouladoux N, Sun CM,Wohlfert EA, Blank RB, Zhu Q, et al. (2008)Commensal DNA limits regulatoryT cell conversion and is a natural adjuvant of intestinal immune responses. Immunity 29: 637–649. doi: 10.1016/j.immuni.2008.08.009 PMID: 18835196.

344. Schachtschneider KM, Yeoman CJ, Isaacson RE, White BA, Schook LB, Pieters M (2013) Modulation of Systemic Immune Responses through Commensal Gastrointestinal Microbiota. PLoS ONE 8(1): e53969. https://doi.org/10.1371/journal.pone.0053969.

345. Wen K, Li G, Bui T, et al. High dose and low dose Lactobacillus acidophilus exerted differential immune modulating effects on T cell immune responses induced by an oral human rotavirus vaccine in gnotobiotic pigs. Vaccine. 2012;30(6):1198-1207. doi:10.1016/j.vaccine.2011.11.107.

346. Schrock, J., Yan, M., Dolatyabi, S. et al. Human Infant Fecal Microbiota Differentially Influences the Mucosal Immune Pathways Upon Influenza Infection in a Humanized Gnotobiotic Pig Model. Curr Microbiol 81, 267 (2024). https://doi.org/10.1007/s00284-024-03785-8.

347. Geervliet M, de Vries H, Jansen CA, Rutten VPMG, van Hees H, Wen C, Skovgaard K, Antonello G, Savelkoul HFJ, Smidt H, Tijhaar E and Wells JM (2022) Effects of Escherichia coli Nissle 1917 on the Porcine

Gut Microbiota, Intestinal Epithelium and Immune System in Early Life. Front. Microbiol. 13:842437. doi: 10.3389/fmicb.2022.842437.

348. Gebert S, Davis E, Rehberger T, Maxwell CV. Lactobacillus brevis strain 1E1 administered to piglets through milk supplementation prior to weaning maintains intestinal integrity after the weaning event. Benef Microbes. 2011;2(1):35-45. doi:10.3920/BM2010.0043.

349. Mangas EL, Rubio A, Álvarez-Marín R, et al. Pangenome of Acinetobacter baumannii uncovers two groups of genomes, one of them with genes involved in CRISPR/Cas defence systems associated with the absence of plasmids and exclusive genes for biofilm formation. Microb Genom. 2019;5(11):e000309. doi:10.1099/mgen.0.000309.