

RESEARCH

Open Access



# The genus *Akkermansia* is populated by a multitude of biological species with a wide distribution in the animal kingdom

Cristian Molteni<sup>1\*</sup>, Diego Forni<sup>1</sup>, Rachele Cagliani<sup>1</sup> and Manuela Sironi<sup>2</sup>

## Abstract

**Background** The mucin-degrading bacterium *Akkermansia muciniphila* has attracted enormous interest for its beneficial effects on human health. However, growing evidence suggests that the *Akkermansia* genus is populated by several species that differ in phenotypic characteristics and association with human traits.

**Results** We present the most comprehensive phylotaxonomic analysis of *Akkermansia* genomes in terms of sample size and host representation. By applying approaches based on average nucleotide identities and on the biological species concept, we show that the *Akkermansia* genus comprises at least 31 species, 13 of which can be detected in humans. The largest species diversity is contributed by non-human and non-mouse animals, and limited evidence of species-specificity is evident, with several *Akkermansia* species detected in phylogenetically distant animals. Analysis of accessory gene content among species also failed to reveal species-specific or diet-specific associations, but rather reflected genome size. Thus, *A. muciniphila* and *A. ignis* have, on average, small genomes and retain a part of genes that characterize either *A. massiliensis* or *A. sp004167605/A. biwaensis*. Finally, investigation of the population structure of *A. muciniphila*, the species that has been more intensely investigated due to its effects on human health, clearly distinguished two phylogroups corresponding to Amla and Amlb. However, analysis of laboratory mouse-derived genomes revealed that additional populations, specific to these animals, exist. Such populations show limited evidence of admixture, suggesting bottleneck or competition effects.

**Conclusions** Our data support the concept that the genetic diversity of *Akkermansia* should be taken into account in experimental settings. They also call for sequencing efforts to characterize the wider genetic diversity of *Akkermansia* bacteria.

**Keywords** *Akkermansia*, Population structure, Genetic diversity, Species classification, Gene flow

## Background

*Akkermansia muciniphila* (phylum *Verrucomicrobiota*), a Gram-negative, mucin-degrading bacterium common in the mammalian gastrointestinal tract, was first characterized two decades ago [1]. The second species described in the genus *Akkermansia* was *A. glycaniphila*, isolated from the feces of a captive reticulated python in 2016 and, later, from human stool [2, 3]. As more genomes have become available over the years, several authors have proposed that the *Akkermansia* genus includes multiple species or subspecies, often referred to

\*Correspondence:

Cristian Molteni  
c.r.molteni@gmail.com

<sup>1</sup> Scientific Institute IRCCS E. MEDEA, Bioinformatics, Bosisio Parini 23842, Italy

<sup>2</sup> School of Medicine and Surgery, University of Milano-Bicocca, Monza 20900, Italy



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

as AmI (sometimes split into AmIa and AmIb) to AmVI [4, 5]. Consistently, in a large-scale metagenomic analysis, Karker and coworkers identified four putative species in the human gut in addition to *A. muciniphila* [6]. Recently, new species were formally proposed, including *A. biwaensis*, *A. massiliensis*, *A. ignis*, *A. durhamii*, *A. timonensis*, *A. sp004167605*, and *A. sp001580195* [7–10]. Moreover, bacteria in the *Akkermansia* genus have been isolated from a number of mammalian and non-mammalian hosts, both in captivity and in the wild [11]. On one hand, *Akkermansia* bacteria derived from non-human animals appear to be divergent so that co-evolution with their hosts was suggested [12, 13]. On the other hand, the failure to detect *Akkermansia* in the gut microbiomes of wild rodents led to the suggestion that the bacterium is present in the intestines of laboratory mice due to acquisition from humans and that it has no evolutionary relationship with mice [14–18]. Nonetheless, a systematic analysis of *Akkermansia* genomes sampled in non-human hosts is presently missing.

In recent years, *Akkermansia*, which reaches an average estimated abundance of approximately 3–5% of the gut microbial community, has attracted great interest due to its association with beneficial effects for human health [19, 20]. Metagenomic studies have shown an inverse correlation between the abundance of *Akkermansia* and the occurrence of metabolic diseases, including diabetes, obesity, and fatty liver disease [10]. Indeed, in experimental mouse models and in a proof-of-concept human study, *A. muciniphila* administration was shown to decrease insulin resistance and cholesterol levels [10]. Also, the bacterium was reported to increase gut barrier integrity and to exert anti-inflammatory effects [10], as well as to potentiate antitumor efficacy with chemotherapy or immunotherapy in different cancer types [20, 21]. Negative associations of *Akkermansia* with human health were also reported, though, including the increased intestinal abundance of the bacterium in subjects suffering from Parkinson's disease or IgA nephropathy [20, 22]. Whether the bacterium exerts beneficial or detrimental effects in the pathogenesis of multiple sclerosis is still controversial [23–25]. Because of these observations, and due to its intrinsic ability to degrade the mucin layer, several studies have thus warned that the use of *Akkermansia* as a probiotic needs further investigation [10, 19, 20, 22]. A critical aspect in this respect is that multiple *Akkermansia* species and strains can colonize the same individual and that different species or phylogroups were shown to display distinct biological properties, especially in relation to metabolic traits [10, 20]. As a consequence, strain- or species-specific phenotypic differences and effects on the host were reported [4, 6, 9, 26–28]. For instance, a study showed that *A. muciniphila*, but not

other *Akkermansia* species, associates with lower body mass index (BMI) [6], whereas a re-analysis of metagenomic studies found *A. biwaensis* to be less common in the fecal samples of obese children compared to controls [9]. Also, the culturing *Akkermansia* species remains technically challenging [29]. Thus, the isolates that have been phenotypically characterized in vitro or in mouse models may represent only the ones that grow more readily in the current culture methods, rather than the full breadth of strains harbored in the animal hosts.

It is thus of paramount importance to gain insight into the genetic diversity and relationships within the *Akkermansia* genus, so that molecular profiling approaches can be developed to investigate the epidemiology of *Akkermansia*-associated host effects in humans and other animals.

## Methods

### Bacterial genomes and core gene sequences

The collection of *Akkermansia* genome accession IDs was sourced from the BV-BRC database (<https://www.bv-brc.org/>), selecting only entries marked as having "good" genome quality. Specifically, the BV-BRC database evaluates genome quality using metrics that evaluate the completeness, contamination, and consistency of genome annotations [30]. However, genomes labeled as "good" do not necessarily have 100% coverage. Genome sequences, including both complete and draft assemblies, were downloaded from the NCBI database using the getGenome function provided by the R package biomaRt [31], resulting in a list of 1,887 bacterial samples. A subsequent selection was made using the Mash tool [32] implemented in the R package PATO [33]. After discarding strains with an average similarity to all other strains lower than 50%, we resulted in a final dataset of 1,626 strains (Supplementary Table 1). Taxonomic classification was performed using the Genome Taxonomy Database Toolkit (GTDB-Tk) [34]. This tool classifies bacterial sequences based on a standardized set of 120 single-copy marker proteins and extracts both nucleotide and protein sequences for each marker from the input genomes. The nucleotide sequences were then used for subsequent analyses. Due to incomplete genome coverage in some samples, not all markers could be retrieved from every genome. Missing genes in individual genomes were treated as gaps.

### Average nucleotide identity and core gene network

The average nucleotide identity (ANI) of all 1,626 strains was computed using Pyani (v.0.2.12), a Python module designed for microbial whole-genome classification [35]. The analysis specifically utilized the ANIm method, which is based on the MUMmer aligner to calculate

pairwise genome similarities [36]. To visualize the results, a heatmap plot was generated using the pheatmap R package.

A nucleotide core gene alignment based on the 120 single-copy markers was constructed using MAFFT with default parameters [37]. A neighbor-net split network was generated throughout SplitsTree4 [38]. Distances between taxa were calculated using the uncorrected P method, that is the proportion of positions at which the two sequences differ.

### Gene flow estimation

Ecologically meaningful populations can be defined by the amount of gene flow shared by pairs of strains. In particular, we applied the PopCOGenT tool [39], that defines species boundaries based on recent horizontal gene transfer by searching for stretches of high sequence identity. PopCOGenT estimates a metric defined as “length bias”, which represents the observed length distribution of identical regions between pairs of genomes compared to a null expectation of non-recombinogenic evolution [39]. Based on this metric, PopCOGenT infers microbial populations and generates networks of gene flow, with strains as nodes and the length bias as a measure of their relationship. To make the analysis computationally feasible, we selected 409 *Akkermansia* strains based on their distribution in the analyzed species (Supplementary Table 1). In particular, we downsampled *A. muciniphila* ( $n=1,066$ ) and *A. ignis* ( $n=312$ ) to the same number of the third most represented species (*A. sp004167605*,  $n=80$ ). To ensure that the results were robust to genome number and choice, we repeated the analysis by including 40 or 160 genomes for *A. muciniphila* and *A. ignis*.

The gene flow networks generated by PopCOGenT were visualized using Cytoscape v3.9.1 [40].

### Accessory gene identification

Complete and draft genomes were used to identify accessory genes using PPanGGOLiN [41]. PPanGGOLiN constructs pangenomes using a graphical model combined with a statistical approach to categorize gene families into three components: persistent, shell, and cloud genomes. Genes belonging to the cloud category (genes shared by at most 15% of the total strains) were considered for subsequent analysis as accessory genes ( $n=63,284$ ).

### PCA analyses on core and accessory genes

The core gene alignment described previously was used to extract biallelic (97% of the total) parsimony-informative (PI) sites; in particular, we focused on sites that had a minimum frequency of two, and we considered only those genomic positions where at least 50% of the sequences had non-missing data. Positions with gaps or

any nonstandard nucleotide bases were excluded and treated as missing values. This approach ensured the inclusion of reliable data while accounting for potential sequencing errors or inconsistencies. This generated a list of 29,622 variable positions. Principal component analysis (PCA) was carried out using the mixOmics R package [42], with the PI matrix serving as the input. This approach enabled the reduction of dimensionality and the visualization of the main sources of variation within the dataset, helping to identify patterns and relationships among the data points. Additionally, a matrix based on the presence and absence of accessory genes in our bacterial dataset was constructed using PPanGGOLiN and used for a second PCA analysis.

### Population structure

A total of 519 *A. muciniphila* strains were selected to create a new concatenated core gene alignment. This dataset included bacterial strains sampled from all non-human hosts (excluding mouse), all human-derived samples, and a random selection of mouse-derived samples equal to human-derived ones (Supplementary Table 1). From this new alignment, biallelic parsimony-informative sites were extracted as above. This resulted in a list of 8,424 variable positions. This data was then used for STRU CTURE analysis [43, 44]. Initially, the software was run with  $K=1$  to estimate the frequency spectrum parameter ( $\lambda$ ), as recommended. The estimated value for  $\lambda$  was 0.3087. Using this parameter, the linkage model with correlated allele frequencies was run for different values of  $K$  (ranging from 1 to 12). For each  $K$  value, ten runs were performed, each with a total Markov chain Monte Carlo (MCMC) chain length of 500,000 iterations and 50,000 iterations of burn-in. The optimal  $K$  was determined using Evanno’s method [45], implemented in the HARVEST tool [46]. Replicate runs for each  $K$  were combined using the CLUMPAK software [47] to generate the membership coefficient matrix.

## Results

### Bacteria in the *Akkermansia* genus can be classified in several distinct species

To perform a comprehensive analysis of bacteria that tentatively belong to the *Akkermansia* genus, we retrieved sequence information for 1,887 genomes from the BV-BRC database. Specifically, we selected all complete/almost complete genomes that were labeled as “*Akkermansia*”, irrespective of host or isolation source (Supplementary Table 1). To obtain a general overview of the diversity among these sequences, we applied Mash, which uses a dimensionality-reduction approach to compute an all-pairs distance matrix [32]. From this matrix, it was evident that some strains were largely

divergent (score > 0.5) and were thus not included in the following analyses, which were performed on a set of 1,626 genomes (Supplementary Fig. 1). Most of these genomes were sampled from humans (n = 352) and mice (n = 1118), but several also derived from a wide number of other animal species (Supplementary Table 1). For these ~1,600 genomes, an average nucleotide identity (ANI) matrix was calculated (Fig. 1A). A 96% cutoff is often used to define bacterial species in ANI analyses [10, 48]. Using this criterion, we identified a number of candidate species, five of which comprised most genomes. Based on the taxonomy proposed in previous works [7–10], we found that the most populous species correspond to *A. muciniphila* (AmI), *A. ignis* (AmV), *A. massiliensis* (AmII), *A. biwaensis* (AmIV), and *A. sp004167605* (hereafter referred to as “common species”). These results were confirmed by a taxonomic classification performed using the Genome Taxonomy Database toolkit [34] on a set of 120 conserved genes.

To explore in more detail the representation of the less abundant species, we repeated ANI analysis after removing genomes of the five common species. The results revealed a quite intricate scenario. Other described species were detected, namely *AmIII*, *A. intestinavium*, *A. glycaniphyla*, and *A. intestinigallinarum* (Fig. 1B). The latter had high similarity to 9 genomes derived from Chinese giant salamanders (*Andrias davidianus*), which, in turn, showed a complex pattern of similarity with strains sampled in different mammals. Moreover, sequences clustering with *A. intestinavium* were sampled from either chickens or humans. As mentioned above, *A. glycaniphyla* was also isolated from humans and other animals (pythons) [2, 3]. In addition, four other putative *Akkermansia* species (each represented by few genomes) were sampled from humans and other animals. Finally, some putative new species were accounted for by a single genomic sequence. The human-derived genomes that contribute to these undescribed species were sampled in different geographic locations (Supplementary Table 1). Overall, these results indicate that the *Akkermansia* genus is divided in many species, several of which are also found in humans. They also suggest that *Akkermansia* bacteria display limited or no species-specificity.

### The *Akkermansia* genus includes at least 31 biological species

ANI analysis is widely used to define bacterial species. However, this approach is based on an arbitrary threshold of identity. As a consequence, strategies based on the Biological Species Concept (BSC) have been proposed as an alternative to ANI in the delineation of species boundaries. The BSC defines a species as a group of interbreeding individuals that remain reproductively

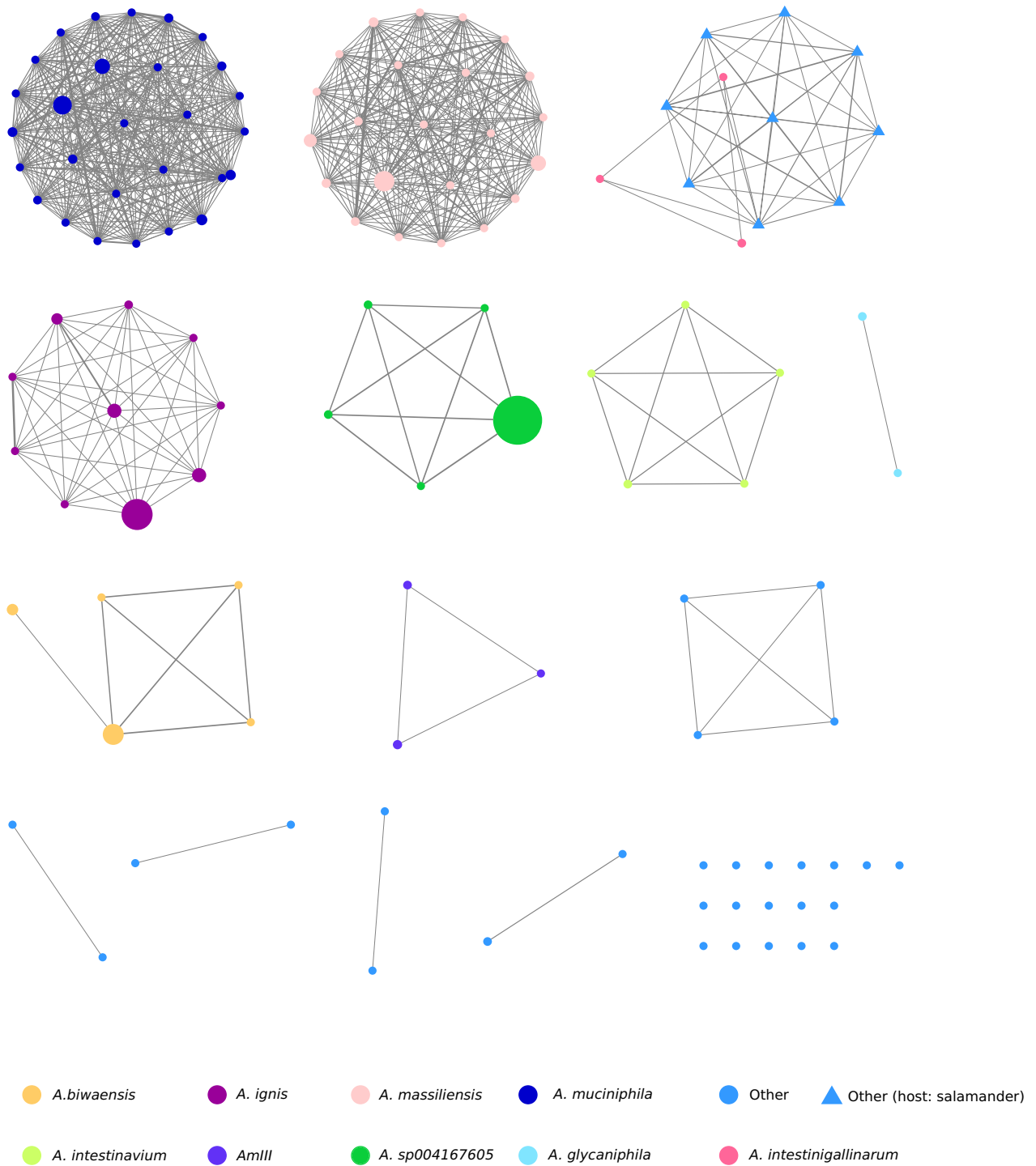
isolated from other groups. In the case of bacteria, the absence of gene-flow can be used to define biological species. We thus applied PopCOGenT (populations as clusters of gene transfer), an approach based on the BSC, to investigate the composition of biological species within the *Akkermansia* genus. PopCOGenT is based on the detection of recent gene-flow discontinuities that delineate species. Specifically, it compares the length distribution of identical regions between pairs of genomes to that expected under a model of clonal evolution, and it provides a measure referred to as “length bias”. Because this approach performs pairwise comparisons of the entire genomes, we pruned our dataset to reduce the computational burden. In particular, we included all genomes of uncommon species and 80 genomes for each of the common ones (see Methods). Using this dataset of 409 genomes, PopCOGenT identified 31 genetically isolated ecological units and revealed clusters with excellent congruence to ANI-defined species (Fig. 2). To assess whether these results were influenced by sequence selection and number, we repeated the analysis using 329 genomes (40 genomes for each of the common species) and 569 genomes (160 for each of the common species). The same number of ecological units was detected in all analyses (Supplementary Fig. 2). Thus, using PopCOGenT, the five common species were confirmed to represent distinct biological units with no detectable gene flow, although *A. biwaensis* genomes were found to belong to two different populations. PopCOGenT also confirmed that *A. glycaniphyla*, *AmIII*, and *A. intestinavium* denote independent biological species. In agreement with ANI analysis, PopCOGenT detected some gene flow between bacteria hosted by giant salamanders and those classified as *A. intestinigallinarum* (Fig. 2). Many other *Akkermansia* species instead consisted of a single genome or very few genomes.

### Genetic relationships of core *Akkermansia* genomes and host associations

To gain further insight into the relationships among *Akkermansia* species and their host associations, we used the Genome Taxonomy Database Toolkit to extract the sequences of 120 core genes from genomes in our dataset. The neighbor-net split network of the core genome showed a complex reticulation pattern, suggestive of ancestral recombination (Fig. 3A). In line with previous reports, *A. muciniphila* was more closely related to *A. ignis* than to other known species [9]. *A. massiliensis* and *AmIII* showed close relationships, as did *A. biwaensis* and *A. sp004167605*. Most of the unclassified core genomes, as well as *A. glycaniphyla*, *A. intestinavium*, and *A. intestinigallinarum*, formed a long tail of highly divergent sequences (Fig. 3A).



**Fig. 1** Average Nucleotide Identity among *Akkermansia* genomes. **A** The ANI heatmap is shown with a cut-off equal to 0.96 of identity. *Akkermansia* species and their hosts are displayed with different colors, as per legend. **B** An enlargement of ANI heatmap for the unknown species is shown with colors as in panel A



**Fig. 2** Gene flow among *Akkermansia* species. Gene flow network of *Akkermansia* genomes. Each node represents a bacterial strain or a clonal strain (all the strains that are closely related, <0.035% divergence) and the size of the node is proportional to the size of the clonal cluster. Edges indicate the inferred magnitude of gene flow between nodes (i.e., the length bias). The width of the edges corresponds to the extent of gene flow between genome pairs

To further explore the diversity of core genomes, we extracted parsimony-informative (PI) sites from the core gene alignment. PI information was used as the input for principal component analysis (PCA). In agreement with the neighbor-net split tree, the first PC, which explained 36% of the variance, separated *A. muciniphila* and *A. ignis* from all other genomes (Fig. 3B). The second component (17% of variance explained) separated most of the common species from the uncommon ones. Overall, the unclassified genomes showed closer relationships to each other than to the common species. In terms of host representation, human- and rodent-derived sequences populated all the common species clusters (Fig. 3C). Conversely, bacteria sampled from other animals strongly contributed to the large diversity of uncommon *Akkermansia* genomes. Some exceptions were also evident, with genomes derived from koalas, wombats, rabbits, and water buffaloes belonging to the common *Akkermansia* species (Fig. 3B). Because all these animals are domestic or were held in captivity, they may have acquired the bacteria from human hosts [49, 50]. When we assigned species (as defined by PopCOGenT) to hosts, we found that the largest number of different species was detected in humans (n=13). After humans, rabbits and koalas were found to host the largest number of species, followed by mice, goats, and chickens (Fig. 3D).

#### Accessory gene content depends on species membership and genome size, not on host association

We next sought to investigate how the genetic relationships established using the core genomes related to accessory gene content and genome size. We thus used the PPanGGOLiN toolkit [41] to extract accessory genes from the *Akkermansia* genomes. A total of 63,284 accessory genes were obtained, which were used for a PCA analysis. Results indicated a very different picture than the one obtained using core genome sequences (Fig. 4A). Genomes mostly grouped by species, but *A. muciniphila* and *A. ignis* clustered together with the majority of uncommon species genomes, including *A. glycaniphila*, *A. intestinavium*, and *A. intestinigallinarum*. The most divergent species in terms of accessory gene content were *A. sp004167605* plus *A. biwaensis* and *A. massiliensis*, that were separated along both components (Fig. 4A).

Virtually no effect of host association on accessory gene content was evident.

Analysis of genome sizes indicated that most of the species that cluster together in the accessory protein PCA (*A. muciniphila*, *A. ignis*, *A. intestinavium*, *A. intestinigallinarum*) have, on average, small genomes (Fig. 4B). The uncommon, unclassified species, despite showing a wide heterogeneity, had also relatively short genomes if compared to those of *A. biwaensis*, *A. massiliensis*, and *A. sp004167605*. Thus, the results of the PCA seem to be explained by the fact that some *Akkermansia* species have more essential genomes compared to others that display larger genomes and different accessory proteins. Thus, *A. muciniphila* and *A. ignis* (plus the uncommon species) probably retain a part of genes that characterize either *A. massiliensis* or *A. sp004167605/A. biwaensis*.

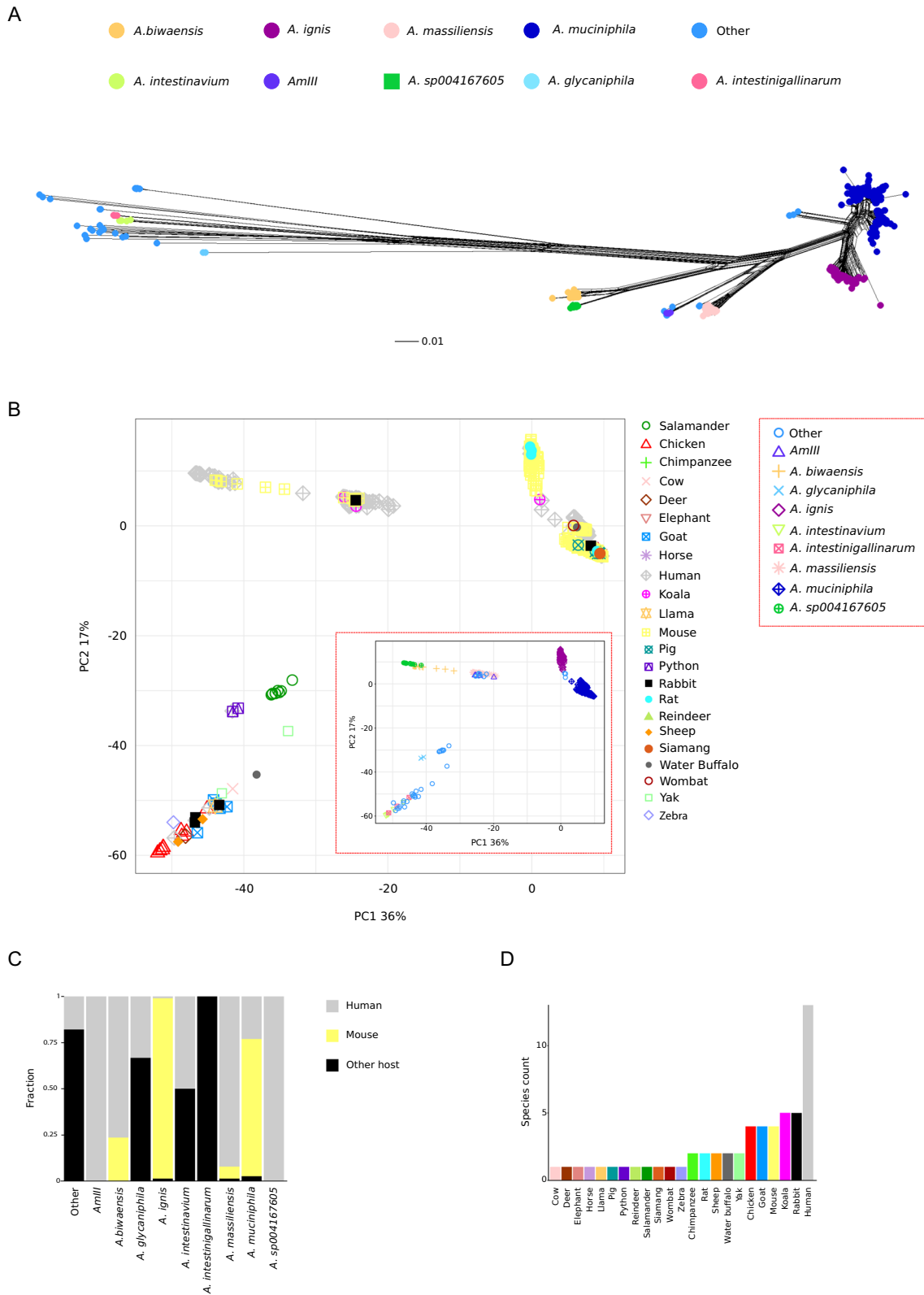
#### Population structure of *A. muciniphila*

We next sought to investigate the population structure of *A. muciniphila*, the *Akkermansia* species that has been more intensely investigated due to its potential effects on human health. To this aim, we used STRUCTURE [43], a program that relies on a Bayesian statistical model for clustering genotypes into populations. STRUCTURE can identify distinct clusters or ancestral subpopulations that account for the ancestry of individuals in the extant population. In our dataset, *A. muciniphila* accounted for the majority of genomes, which were sampled from humans (n=246), mice (n=800), and other animals (n=27). To reduce the computation load, we sub-sampled the *A. muciniphila* genomes by including all those from humans and animals other than mice, plus 246 randomly selected genomes derived from mice. Using this dataset, we applied the linkage model with correlated allele frequencies, which assumes that discrete genome “chunks” are inherited from K ancestral populations [44]. The optimal number of populations (K) was estimated using the  $\Delta K$  method [45] and resulted equal to 7 (Supplementary Fig. 3).

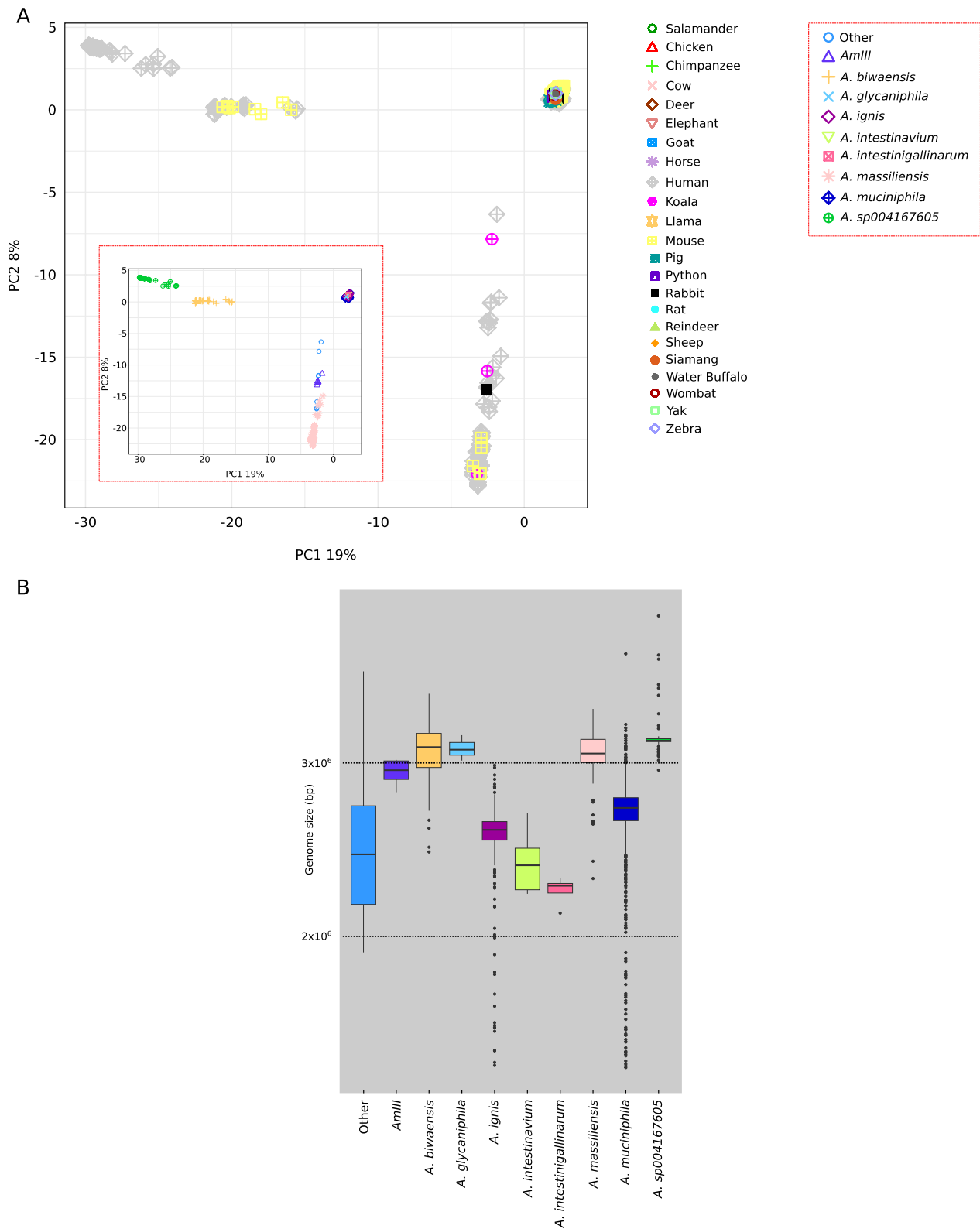
Previous studies indicated that *A. muciniphila* genomes belong to two different phylogroups, designated as AmIa and AmIb [4, 9]. Analysis of ancestry components revealed that this distinction is clearly visible: human-derived genomes previously classified as AmIb have most of their ancestry contributed by two components

(See figure on next page.)

**Fig. 3** Relationships among *Akkermansia* species and host associations. **(A)** Neighbor-net split network based on all variant sites of the core genome alignment. Each sequence is shown as a dot, color-coded by species. **(B)** PCA plot based on parsimony-informative sites of the core genome alignment. The first two principal components are shown. Each strain is represented as a dot and colored based on sampling host (large image) or *Akkermansia* species (insert). **(C)** Barplot of host representation for different *Akkermansia* species. Hosts other than humans and mice are grouped together. **(D)** Barplot showing the number of different *Akkermansia* species in each host



**Fig. 3** (See legend on previous page.)



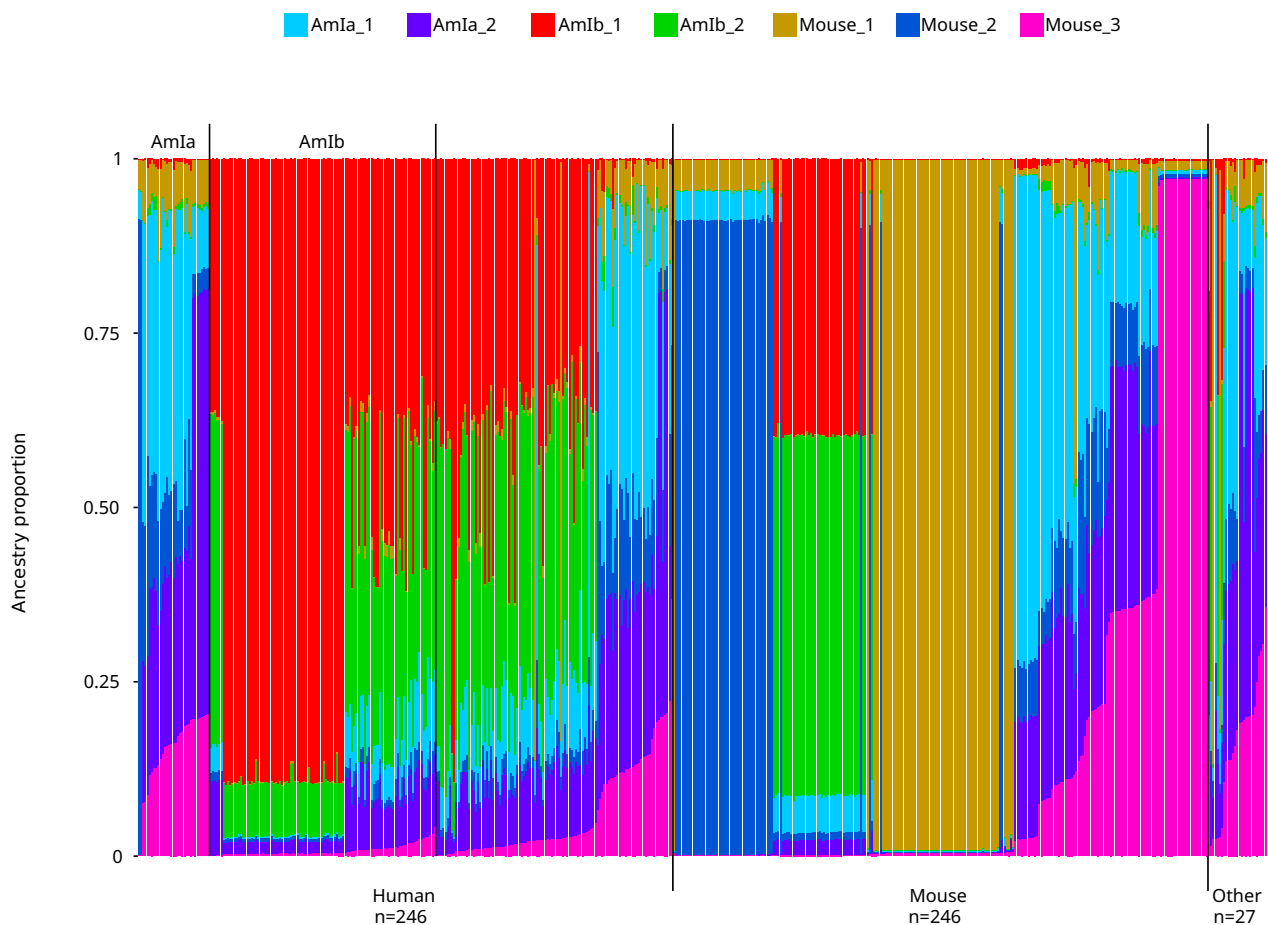
**Fig. 4** Accessory genes content and genomes size variability. **(A)** PCA plot based on accessory genes presence/absence matrix. The first two principal components are shown, each strain is represented as a dot and colored based on sampling host (larger image) or *Akkermansia* species (insert). **(B)** Box and whiskers plot of genome size in *Akkermansia* species

(AmIb\_1 and AmIb\_2), whereas AmIa genomes are more admixed with two components (AmIa\_1 and AmIa\_2) being the most common (Fig. 5). Analysis of mouse-derived genomes showed that some of these belong to the AmIa and AmIb phylogroups. However, additional populations with little evidence of admixture were evident. In particular, three components (mouse\_1, mouse\_2, and mouse\_3) contributed most of the ancestry of several genomes, whereas they were relatively uncommon in genomes derived from humans (Fig. 5). In non-human and non-mouse animals, most of the genomes belong to the AmIa phylogroup. Overall, these results indicate that humans and mice host distinct *A. muciniphila* populations.

**Discussion**

In this study, we analyzed a large dataset of *Akkermansia* genomes with the aim of determining the phylotaxonomic relationships within the genus. Previous studies of *Akkermansia* genetic diversity mainly focused either on

human- (and mouse-) derived sequences or on genomes obtained from captive non-human mammals [3, 6, 12, 13]. Herein, in order to investigate host associations, we assembled a genome collection derived from a wide diversity of animals, including mammals, birds, reptiles, and amphibians. We show that the genus consists of a large number of species. In particular, ANI analysis and the application of a species definition criterion based on the BSC were highly concordant in revealing the presence of 31 biological species, at least 13 of which can colonize the human gut. In line with previous reports [12, 13], the largest species diversity was contributed by non-human and non-mouse animals, and the fact that several species were represented by a single or very few genomes clearly indicates the need of implementing sequencing efforts to characterize the wider genetic diversity of *Akkermansia* bacteria. In this respect, it is worth mentioning that, in order to generate interpretable results, we excluded from our analyses the most divergent sequences. Further analyses will be required to assess whether these genomes



**Fig. 5** *Akkermansia muciniphila* population structure. Bar plot representing the proportion of ancestral population components in a subset of *A. muciniphila* core genomes. Each vertical line represents a strain and it is colored by the proportion of sites that have been assigned to one of the seven populations estimated by STRUCTURE. Ancestry components are named based on the genomes where they are more prevalent

represent additional species in the genus *Akkermansia* and if they have any effect on human health.

In terms of host associations, we found very little evidence of species-specificity or adaptation, as several species were detected in phylogenetically distant animals. This contrasts with previous data showing that microbial communities in the vertebrate gut differ considerably as an adaptation to host diet and other physiological parameters (e.g., body temperature) [51]. We thus investigated whether the repertoire of accessory proteins showed host-specific variations, but again we detected no such effect. As previously suggested, the presence of *Akkermansia* across multiple vertebrate hosts may relate to the specialization of the bacterium for the mucosal niche, where it can thrive using mucin as the sole nitrogen and carbon source [10, 12]. This may allow *Akkermansia* bacteria to colonize the mucus layer of different hosts, irrespective of diet or other physiological characteristics. It should also be noted that most of the bacterial strains we analyzed were sequenced from captive animals. It is thus impossible to ascertain whether the *Akkermansia* species we identified in these hosts represent long-standing evolutionary relationships or recent acquisitions from humans. However, because *A. muciniphila* is the most abundant species in the human gut [10], we would expect it to be preferentially transferred to captive animals, as is the case of laboratory mice. Instead, a large diversity of *Akkermansia* species was evident in these hosts, with some animals (e.g., rabbits and koalas) carrying multiple species. In this respect, it is somehow surprising that *Akkermansia* has never been found in the microbiome of wild mice [16, 17]. The gut of laboratory mice can be colonized by *Akkermansia* (indeed mouse-derived strains represented the majority of genomes in our database) and these rodents have a typical commensal behavior that may facilitate the exposure to human-derived bacteria. Whether *Akkermansia* is out-competed in wild mice by other bacteria or the lack of detection simply reflects limited sampling efforts remains to be determined.

The analysis of accessory proteins clearly showed that differences among species largely reflect genome size. Three of the common *Akkermansia* species (*A. biwaensis*, *A. massiliensis*, and *A. sp004167605*) have, on average, much larger genomes than *A. muciniphila* and *A. ignis*, and the latter retain a portion of the accessory gene content of the former. Differences in the pangenomes of *A. biwaensis*, *A. massiliensis*, and *A. muciniphila* were previously reported and were suggested to result in phenotypic differences [4]. Indeed, the three species display distinctive traits, in terms of doubling time, oxygen tolerance, adherence to epithelial cells, oligosaccharide usage, and optimal growth temperature [4, 10]. The ecological reasons why these species have diverged genetically and

phenotypically are presently unclear, but may relate to competition among themselves or with other bacteria. Indeed, *A. biwaensis* outcompetes *A. muciniphila* and *A. massiliensis* in the colonization of the murine gut [4]. Whatever the underlying reasons, even subtler differences are likely to have an impact on human health. For instance, recent evidence indicated that *A. muciniphila* AmIa protects from Crohn's disease and ulcerative colitis, whereas AmIb is only protective from ulcerative colitis. Likewise, AmIa and AmIb differ in their modulation of the efficacy of immune checkpoint inhibitor treatment [9]. In this respect, we should also add that the sequencing of *A. muciniphila* genomes revealed that some isolates carry large inversions and rearrangements [52]. In different bacterial species, inversions were shown to change gene expression profiles and to modulate phenotypic traits, possibly enabling adaptation to stressful conditions [53–57]. Thus, it is also possible that the genotypic/phenotypic characteristics of different isolates belonging to the same *Akkermansia* species have distinct effects on human health.

The observations above prompted us to analyze the population structure of *A. muciniphila* (although with this approach rearrangements and inversions cannot be analyzed). Inspection of ancestry components in human-derived genomes clearly distinguished the AmIa and AmIb phylogroups, characterized by variable levels of admixture. However, a different picture emerged in genomes sampled from mice. Whereas some of these belong to the AmIa and AmIb phylogroups, others had ancestry components specific for these animals and very little evidence of admixture. The major components mouse\_1, \_2 and \_3 contribute a very minor ancestry fraction of AmIa strains. The most likely explanation for these findings is that AmIa and AmIb sampled from mice represent recent, ongoing transfers from humans. Conversely, the mouse specific components are the result of bottlenecks (probably due to transfer or competition) experienced by AmIa bacteria that generated genetically drifted, mouse-specific populations over years of husbandry. Strains of *Akkermansia* show competitive exclusion in the colonization of the mouse intestine, which can only occur in animals with a microbiome devoid of *A. muciniphila* [58]. This creates the conditions for the emergence of mouse-specific populations. Whether such populations are also adapted to their hosts or simply derive from drift remains to be evaluated. In any case, their existence contributes to expand the concept that the genetic diversity of *Akkermansia*, not only in terms of species, but also of populations or phylogroups, should be taken into account in experimental settings [20]. Variability in the microbiome of laboratory mice among vivaria is considered a major confounding effect

and a cause of replication failure [59]. Thus, our findings are relevant to the translatability of the mouse model to human *Akkermansia* research. Also, the identification of several novel species, including some hosted by humans, suggests that some effects initially attributed to the *A. muciniphila* species may need to be recharacterized.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13062-025-00680-5>.

Additional file 1  
Additional file 2  
Additional file 3  
Additional file 4

## Author contributions

Conceptualization, M.S. and D.F.; Methodology, M.S., C.M., and D.F.; Investigation, C.M., D.F., R.C.; Writing Original Draft, M.S. C.M.; Writing Review & Editing, M.S., D.F., and R.C.; Funding Acquisition, R.C.

## Funding

This work was supported by the Italian Ministry of Health ("Ricerca Corrente").

## Availability of data and materials

All *Akkermansia* strain GenBank accession IDs are listed in Supplementary Table S1.

## Declarations

## Competing interests

The authors declare no competing interests.

Received: 20 May 2025 Accepted: 17 July 2025

Published online: 24 July 2025

## References

- Derrien M, Vaughan EE, Plugge CM, De Vos WM. *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int J Syst Evol Microbiol*. 2004;54:1469–76.
- Ouwerkerk JP, Aalvink S, Belzer C, De Vos WM. *Akkermansia glycaniphila* sp. nov., an anaerobic mucin-degrading bacterium isolated from reticulated python faeces. *Int J Syst Evol Microbiol*. 2016;66:4614–20.
- Lv Q-B, Li S, Zhang Y, Guo R, Wang Y-C, Peng Y, et al. A thousand metagenome-assembled genomes of *Akkermansia* reveal phylogroups and geographical and functional variations in the human gut. *Front Cell Infect Microbiol*. 2022;12: 957439.
- Becken B, Davey L, Middleton DR, Mueller KD, Sharma A, Holmes ZC, et al. Genotypic and phenotypic diversity among human isolates of *Akkermansia muciniphila*. McFall-Ngai MJ, editor. *MBio*. 2021;12:e00478–e521.
- Kelly C, Jawahar J, Davey L, Everitt JJ, Galanko JA, Anderson C, et al. Spontaneous episodic inflammation in the intestines of mice lacking HNF4A is driven by microbiota and associated with early life microbiota alterations. Dominguez Bello MG, editor. *MBio*. 2023;14:e01504–e1523.
- Karcher N, Nigro E, Punčochář M, Blanco-Míguez A, Ciciani M, Manghi P, et al. Genomic diversity and ecology of human-associated *Akkermansia* species in the gut microbiome revealed by extensive metagenomic assembly. *Genome Biol*. 2021;22:209.
- Ndongo S, Armstrong N, Raoult D, Fournier P-E. Reclassification of eight *Akkermansia muciniphila* strains and description of *Akkermansia massiliensis* sp. nov. and *Candidatus Akkermansia timonensis*, isolated from human feces. *Sci Rep*. 2022;12(1): 21747.
- Kobayashi Y, Kawahara T, Inoue S, Kohda N. *Akkermansia biwaensis* sp. Nov., an anaerobic mucin-degrading bacterium isolated from human faeces. *Int J Syst Evol Microbiol*. 2023;73:005697.
- Mueller KD, Panzetta ME, Davey L, McCann JR, Rawls JF, Flores GE, et al. Pangenomic analysis identifies correlations between *Akkermansia* species and subspecies and human health outcomes. *Microbiome Res Rep*. 2024;3:33.
- Ioannou A, Berkhout MD, Geerlings SY, Belzer C. *Akkermansia muciniphila*: biology, microbial ecology, host interactions and therapeutic potential. *Nat Rev Microbiol*. 2025;23:162–77.
- Belzer C, De Vos WM. Microbes inside—from diversity to function: the case of *Akkermansia*. *ISME J*. 2012;6:1449–58.
- Geerlings SY, Ouwerkerk JP, Koehorst JJ, Ritari J, Aalvink S, Stecher B, et al. Genomic convergence between *Akkermansia muciniphila* in different mammalian hosts. *BMC Microbiol*. 2021;21:298.
- González D, Morales-Olavarria M, Vidal-Veuthey B, Cárdenas JP. Insights into early evolutionary adaptations of the *Akkermansia* genus to the vertebrate gut. *Front Microbiol*. 2023;14:1238580.
- Cardilli A, Hamad I, Dyczko A, Thijs S, Vangronsveld J, Müller DN, et al. Impact of high salt-intake on a natural gut ecosystem in Wildling mice. *Nutrients*. 2023;15:1565.
- Rosshart SP, Vassallo BG, Angeletti D, Hutchinson DS, Morgan AP, Takeda K, et al. Wild mouse gut microbiota promotes host fitness and improves disease resistance. *Cell*. 2017;171:1015–1028.e13.
- Rosshart SP, Herz J, Vassallo BG, Hunter A, Wall MK, Badger JH, et al. Laboratory mice born to wild mice have natural microbiota and model human immune responses. *Science*. 2019;365: eaaw4361.
- Segura Munoz RR, Mantz S, Martínez I, Li F, Schmaltz RJ, Pudlo NA, et al. Experimental evaluation of ecological principles to understand and modulate the outcome of bacterial strain competition in gut microbiomes. *ISME J*. 2022;16:1594–604.
- Hild B, Dreier MS, Oh JH, McCulloch JA, Badger JH, Guo J, et al. Neonatal exposure to a wild-derived microbiome protects mice against diet-induced obesity. *Nat Metab*. 2021;3:1042–57.
- Menjivar C, Pagella E, Biswas I. *Akkermansia muciniphila*. *Trends Microbiol*. 2024;32:1143–4.
- Panzetta ME, Valdivia RH. *Akkermansia* in the gastrointestinal tract as a modifier of human health. *Gut Microbes*. 2024;16:2406379.
- Li L, McAllister F. Too much water drowned the miller: *Akkermansia* determines immunotherapy responses. *Cell Rep Med*. 2022;3: 100642.
- Chiantera V, Laganà AS, Basciani S, Nordio M, Bizzarri M. A critical perspective on the supplementation of *Akkermansia muciniphila*: benefits and harms. *Life*. 2023;13:1247.
- Cox LM, Maghzi AH, Liu S, Tankou SK, Dhang FH, Willocq V, et al. Gut microbiome in progressive multiple sclerosis. *Ann Neurol*. 2021;89:1195–211.
- Hindson J. A possible link between multiple sclerosis and gut microbiota. *Nat Rev Neurol*. 2017;13:705–705.
- Vallino A, Dos Santos A, Mathé CV, Garcia A, Morille J, Dugast E, et al. Gut bacteria *Akkermansia* elicit a specific IgG response in CSF of patients with MS. *Neurol Neuroimmunol Neuroinflamm*. 2020;7: e688.
- Kumar R, Kane H, Wang Q, Hibberd A, Jensen HM, Kim H-S, et al. Identification and characterization of a novel species of genus *Akkermansia* with metabolic health effects in a diet-induced obesity mouse model. *Cells*. 2022;11:2084.
- Zhai R, Xue X, Zhang L, Yang X, Zhao L, Zhang C. Strain-specific anti-inflammatory properties of two *Akkermansia muciniphila* strains on chronic colitis in mice. *Front Cell Infect Microbiol*. 2019;9:239.
- Liu Q, Lu W, Tian F, Zhao J, Zhang H, Hong K, et al. *Akkermansia muciniphila* exerts strain-specific effects on DSS-induced ulcerative colitis in mice. *Front Cell Infect Microbiol*. 2021;11: 698914.
- Li Z, Hu G, Zhu L, Sun Z, Jiang Y, Gao M, et al. Study of growth, metabolism, and morphology of *Akkermansia muciniphila* with an in vitro advanced bionic intestinal reactor. *BMC Microbiol*. 2021;21:61.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25:1043–55.

31. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc.* 2009;4:1184–91.
32. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* 2016;17:132.
33. Fernández-de-Bobadilla MD, Talavera-Rodríguez A, Chacón L, Baquero F, Coque TM, Lanza VF. PATO: pangenome analysis toolkit. Luigi Martelli P, editor. *Bioinformatics.* 2021;37:4564–6.
34. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics.* 2022;38:5315–6.
35. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods.* 2016;8:12–24.
36. Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. MUMmer4: A fast and versatile genome alignment system. Darling AE, editor. *PLoS Comput Biol.* 2018;14:e1005944.
37. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30:772–80.
38. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol.* 2006;23:254–67.
39. Arevalo P, VanInsberghe D, Elsherbini J, Gore J, Polz MF. A reverse ecology approach based on a biological definition of microbial populations. *Cell.* 2019;178:820–834.e14.
40. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13:2498–504.
41. Gautreau G, Bazin A, Gachet M, Planel R, Burlot L, Dubois M, et al. PPanG-GOLIN: depicting microbial diversity via a partitioned pangenome graph. Ouzounis CA, editor. *PLoS Comput Biol.* 2020;16:e1007732.
42. Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol.* 2017;13: e1005752.
43. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000;155:945–59.
44. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics.* 2003;164:1567–87.
45. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol.* 2005;14:2611–20.
46. Earl DA, vonHoldt BM. Structure harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour.* 2012;4:359–61.
47. Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour.* 2015;15:1179–91.
48. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA.* 2009;106:19126–31.
49. Blyton MDJ, Soo RM, Hugenholtz P, Moore BD. Maternal inheritance of the koala gut microbiome and its compositional and functional maturation during juvenile development. *Environ Microbiol.* 2022;24:475–93.
50. Shiffman ME, Soo RM, Dennis PG, Morrison M, Tyson GW, Hugenholtz P. Gene and genome-centric analyses of koala and wombat fecal microbiomes point to metabolic specialization for Eucalyptus digestion. *PeerJ.* 2017;5: e4075.
51. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE, et al. Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci USA.* 2013;110:3229–36.
52. Ouwerkerk JP, Tytgat HLP, Elzinga J, Koehorst J, Van Den Abbeele P, Henrissat B, et al. Comparative Genomics and Physiology of Akkermansia muciniphila Isolates from Human Intestine Reveal Specialized Mucosal Adaptation. *Microorganisms.* 2022;10:1605.
53. Cui L, Neoh H, Iwamoto A, Hiramatsu K. Coordinated phenotype switching with large-scale chromosome flip-flop inversion observed in bacteria. *Proc Natl Acad Sci USA.* 2012. <https://doi.org/10.1073/pnas.1204307109>.
54. Irvine S, Bunk B, Bayes HK, Spröer C, Connolly JPR, Six A, et al. Genomic and transcriptomic characterization of *Pseudomonas aeruginosa* small colony variants derived from a chronic infection model. *Microbial Genomics.* 2019;5:e000262.
55. Guérrillot R, Kostoulas X, Donovan L, Li L, Carter GP, Hachani A, et al. Unstable chromosome rearrangements in *Staphylococcus aureus* cause phenotype switching associated with persistent infections. *Proc Natl Acad Sci USA.* 2019;116:20135–40.
56. Kojic M, Jovcic B, Miljkovic M, Novovic K, Begovic J, Studholme DJ. Large-scale chromosome flip-flop reversible inversion mediates phenotypic switching of expression of antibiotic resistance in lactococci. *Microbiol Res.* 2020;241: 126583.
57. Fitzgerald SF, Lupolova N, Shaaban S, Dallman TJ, Greig D, Allison L, et al. Genome structural variation in *Escherichia coli* O157:H7. *Microb Genom.* 2021. <https://doi.org/10.1099/mgen.0.000682>.
58. Segura Munoz RR, Mantz S, Martínez I, Li F, Schmaltz RJ, Pudlo NA, et al. Experimental evaluation of ecological principles to understand and modulate the outcome of bacterial strain competition in gut microbiomes. *ISME J.* 2022;16:1681–2.
59. Servick K. Of mice and microbes. *Science.* 2016;353:741–3.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.