

Sequencing of five poultry strains elucidates phylogenetic relationships and divergence in virulence genes in *Morganella morganii*

Nicola Palmieri

Veterinarmedizinische Universität Wien

Claudia Hess

Veterinarmedizinische Universität Wien

Michael Hess

Veterinarmedizinische Universität Wien

Merima Alispahic (✉ merima.alispahic@vetmeduni.ac.at)

Veterinarmedizinische Universität Wien <https://orcid.org/0000-0002-2347-2030>

Research article

Keywords: *Morganella morganii*, poultry, NGS data, MALDI-TOF MS, antimicrobial resistance, virulence genes, phylogeny

Posted Date: April 9th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-21281/v1>

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

[Read Full License](#)

Version of Record: A version of this preprint was published on August 24th, 2020. See the published version at <https://doi.org/10.1186/s12864-020-07001-2>.

Abstract

Background *M. morganii* is a bacterium frequently associated with urinary infections in humans. While many human strains are sequenced, only the genomes of few poultry strains are available. Here, we performed a detailed characterization of five highly resistant *Morganella morganii* strains isolated in association with *Escherichia coli* from diseased domestic Austrian poultry flocks, namely geese, turkeys and chicken layers. Additionally, we sequenced the genomes of these strains by NGS and analyzed phylogenetic clustering, resistance and virulence genes in the context of host-specificity. Results Two strains were identified to be Extended Spectrum Beta Lactamase (ESBL) and one as AmpC beta-lactamases (AMP-C) phenotype, while two were ESBL negative. By integrating the genome sequences of these five poultry strains with all the available *M. morganii* genomes, we constructed a phylogenetic tree that clearly separates the *Morganella* genus into two clusters (M1 and M2), which approximately reflect the proposed subspecies classification (*morganii* and *sibonii*). Additionally, we found no association between phylogenetic structure and host, suggesting interspecies transmission. All five poultry strains contained genes for resistance to aminocoumarins, beta-lactams, colistin, elfamycins, fluoroquinolones, phenicol, rifampin and tetracycline. A comparative genomics analysis of virulence genes showed acquisition of novel virulence genes involved in secretion system and adherence in cluster M2. We showed that some of these genes were acquired by horizontal gene transfer from closely related *Morganellaceae* species and propose that novel virulence genes could be responsible for expansion of tissue tropism in *M. morganii*. Finally, we detected variability in copy number and high sequence divergence in toxin genes and provided evidence for positive selection in insecticidal toxins genes, likely reflecting host-related adaptations. Conclusions In summary, this study describes i) the first isolation and characterization of *M. morganii* from goose and turkey, ii) a large-scale genetic analysis of *M. morganii* and an attempt to generate a global picture of the *M. morganii* intraspecific phylogenetic structure.

Background

Morganella morganii is a facultative anaerobic Gram-negative enteric rod-shaped bacterium that used to belong to the *Enterobacteriaceae* family [1]. Recently, *Morganella* was suggested to be a type genus of a novel *Morganellaceae* family that consists of the following eight genera: *Arsenophonus*, *Cosenzaea*, *Moellerella*, *Morganella*, *Photorhabdus*, *Proteus*, *Providencia*, and *Xenorhabdus* [2]. *M. morganii* is reported as an opportunistic pathogen in humans and various other animal species [3]. In reported human cases the disease spectrum associated with *M. morganii* infections is broad and the mortality following an infection remains high [3]. In humans it has also been linked with animal bite wound infections, which underlines that *M. morganii* causes zoonotic infectious diseases [3]. However, the bacterium is scarcely reported in poultry: in Nigeria and China, *M. morganii* was isolated from dead broilers and its pathogenicity confirmed by infecting day-old chicks [4, 5]. Furthermore, *M. morganii* was isolated from chicken carcasses in US retail raw meat [6] and from a 13-day old broiler from a farm in Portugal [7]. Several factors can affect the progression and severity of an infection with antibiotic resistance as a main factor in both human and veterinary medicine. The World Health Organization has

recently published a global priority list of antibiotic-resistant bacteria; “Priority 1: CRITICAL” ranking include extended spectrum β -lactamase (ESBL) producing and carbapenemase producing *M. morganii* (WHO 2019). *M. morganii* has a natural resistance to beta-lactam antibiotics (e.g. ampicillins, amoxicillin, most of the 1st and 2nd generation cephalosporins) and colistin [3]. In addition, the presence of pathogenicity determinants is essential to the success of *M. morganii* in any environment. Virulence genes cluster such as urease, flagellar and haemolysin were all found in *M. morganii* characterizing its potential as a pathogen and the ability to colonise different hosts [7–10]. In other *Enterobacteriaceae*, association between host and phylogenetic structure was observed in *E. coli* [11] but not in *Salmonella enteritidis* [12], neither in *Providencia* species [13], while the situation for *M. morganii* is presently unknown. A total of 47 complete genome sequences from *M. morganii* strains are currently available from NCBI database mainly isolated from humans, but also one from a broiler, salamander, dolphin, plants and few where the host is unknown. To date, a detailed genome analysis has been carried out for only few of these strains (KT, F675, INSRALV892a, MM 1, MM 4, and MM 190) (Chen et al., 2012; Olaitan et al., 2014; Jones-Dias et al., 2016, Minnullina et al 2019) and detailed phylogeny is lacking. These studies reported peculiar strain-specific innovations such as: i) acquisition of resistance genes from *Acinetobacter* spp. in the F675 strain [9], ii) introduction of toxins (hlyA) through lateral transfer from phages in the MM1, MM4 and MM190 strains [10], and iii) plasmid-mediated acquisition of quinolone resistance in the INSRALV892a strain [7].

Here, we describe five poultry strains of *M. morganii* isolated for the first time from geese, a fattening turkey and a layer chicken parent flock. We characterized their resistance profile phenotypically and analyzed resistance genes, mobile elements and virulence genes through genome sequencing by Illumina NGS. By combining the genomic data of these poultry strains with all the available genomes we constructed a large-scale phylogeny of *M. morganii*, which we used to test the hypothesis of association between host and phylogeny and to highlight phenomena of lineage-specific divergence in virulence genes.

Methods

Necropsy, bacteria isolation and identification

Carcasses from three geese flocks (A-C), one-day old turkey (D) and one-layer chicken parent (E) flocks (Table 1) were sent to the Clinic of Poultry and Fish Medicine (Vetmeduni Vienna, Austria) for necropsy and bacteriological investigation. Necropsy was performed according to a standard protocol. For bacteriological investigation, direct smears were taken aseptically from heart, liver, intestine, brain, spleen and lungs. Samples from each organ were cultivated on Columbia (COS) agar containing 5% sheep blood (BioMerieux, Vienna, Austria), McConkey agar (Bertoni, Vienna, Austria), Chromocult Coliform agar (Merck, Vienna, Austria) and subsequently incubated aerobically at 37 °C for 24hrs. Additionally, intestinal samples were cultivated on Schaedler (SCS) agar containing 5% sheep blood (BioMerieux, Vienna, Austria) for anaerobic growth (Genbox anaer BioMerieux) incubated at 37 °C for 24hrs as well as for

fungal growth on Sabouraud Gentamicin Chloramphenicol (SGC2) agar (BioMerieux, Vienna, Austria) incubated at 42 °C for 48hrs. All bacterial isolates were identified by Matrix Assisted Laser Desorption/Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS) (Microflex LT instrument, Bruker Daltonic) against the reference library from Bruker Daltonic version 4.1.80 according to the manufacturer's instructions, with sample preparation performed as previously described [14].

Antibiotic susceptibility testing

Antibiotic susceptibility testing was performed by disc diffusion method according to the EUCAST guidelines (Bauer et al 1966, EUCAST 2019). Phenotypic detection of ESBL type A β -lactamases, AmpC type C cephalosporinases, KPC type A carbapenamases and MBL metallo- β -lactamases was performed by broth microdilution method using the commercially available MICRONAUT-S β -Lactamases plateE1-111-040 (MERLIN Diagnostika GmbH, Bornheim-Hersel, Germany). The antimicrobial substances and their concentration are given in Additional File1: Supplementary Table 1. The analysis was done according the manufacturer's instructions. Results were evaluated with the MCN6 Software Version 6.00 Release 71 (MERLIN Diagnostika GmbH, Bornheim-Hersel, Germany).

Genome sequencing and annotation

DNA was extracted using QIAamp DNA Mini Kit following the manufacturer's guidelines (Qiagen, Hilden, Germany). The five samples were sent to the Vienna BioCenter Core Facility, where they were sequenced on an Illumina NextSeq machine using a paired-end PCR-free library (150 bp read length). Resulting reads were assembled into contigs using CLC Genomics Workbench 12.0 (<https://www.qiagenbioinformatics.com/>) by the De Novo Assembly workflow (default parameters) and annotated using PROKKA 1.13 [15]. Plasmids were detected using PlasmidFinder 2.0 [16] and by BLAST searches against a local version of the *Enterobacteriaceae* plasmid database [17]. Prophages were annotated using the recent tool Prophage Hunter [18], which is also able to differentiate active from inactive prophages. Integrons were annotated by IntegronFinder 1.5.1 [19]. Pathogenicity islands were identified by IslandViewer 4 [20] using the *M. morganii* KT strain [8] as a reference.

Phylogenetic trees construction

For tree construction, the five sequenced assemblies were employed together with all complete *M. morganii* genomes downloaded from NCBI, encompassing a total of 52 strains. For each strain, the genome sequence was concatenated to form a single artificial sequence, including plasmids and unscaffolded sequences. A k-mer phylogenetic tree was constructed using CLC Microbial Genomics 12.0 by the Create K-mer Tree workflow (default parameters). An additional phylogenetic tree was constructed based on multiple genome alignment of all the strains with the program parsnp [21], using the KT strain [8] as a reference.

Identification of resistance genes

The protein sequences of each strain were BLASTed (blastp -evalue 1e-05) to a local version of the MEGARes database [22], which contains more than 4,000 hand-curated antimicrobial resistance genes.

Hits with a query coverage > 40% were kept for further analyses. Additional screening for resistance genes that were absent in the MEGARes database (*lpx* genes, *tetB*) was performed using the same BLAST-based approach.

Identification of virulence genes

The protein sequences of virulence genes for the *M. morganii* KT strain [8] were retrieved using the 4-letter gene-id from the NCBI Protein database and BLASTed (blastp -evalue 1e-05) to the proteomes of the five sequenced poultry strains. Hits with a query coverage > 40% were kept and processed. Novel virulence genes were identified by BLASTing the protein sequences of the five strains to a local version of the VFDB (Virulence Factor DataBase) using stringent criteria (blastp -evalue 1e-10, query coverage > 40%), in order to get a conservative set of novel virulence genes. Annotation of functional classes for virulence genes was based on the intra-genera VFs comparison tables from VFDB (http://www.mgc.ac.cn/VFs/Down/Comparative_tables_from_VFDB.tar.gz). Gene families were computed by clustering the protein sequences of each strain using the sequence clustering program MMseqs2 [23] with a minimum 80% similarity threshold.

Positive-selection scans

Coding sequences for orthologous groups of genes of interest were aligned using PRANK [24]. Positive selection scans were performed with the codeml module from PAML [25] by comparing a model with a fixed ω for the whole tree (model = 2, NSsites = 2, fix_omega = 1) with a model allowing positive selection in the PA17/10312 branch (model = 2, NSsites = 2, fix_omega = 0). P-values were calculated by standard log-likelihood ratio tests.

Construction of metabolic networks

Protein sequences were submitted to BlastKOALA [26] to reconstruct the metabolic network of each strain. The metabolic networks of the different strains were compared through the KEGG Pathway interface from the KEGG database [27] in order to discover strain-specific pathways.

Results And Discussion

Necropsy, bacteria isolation and identification

General fibrinous serositis was observed in dead birds from flocks A and B. In case of flock C, fibronectic typhilitis was the main gross pathological lesion found, a similar observation as in the 13-day old broiler from Portugal [7]. Omphalitis was diagnosed in turkey poults from flock D, while the chicken from a layer parent flock E did lack signs of bacterial infections but had a fracture of the leg. *Escherichia coli* was isolated from all cases investigated as reported by [7] who detected the same bacterium with *M. morganii* in liver and spleen of a broiler. *Riemerella anatipestifer* was detected in some organ samples from birds of flock B and *Clostridium perfringens* was found in association with a fibronectic typhilitis. When investigating the *E. coli* isolates for antibiotic resistance by disc diffusion

method, smaller colonies were observed within the inhibition zone of ampicillin, amoxicillin and colistin, pointing towards a co-infection of *E. coli* and *M. morganii*. MALDI-TOF MS identified all strains to be *M. morganii* species, of which the strain PA17/10312 was classified as subspecies *sibonii* and the other four strains as subspecies *morganii* with a log score value above 2.5 (Table 1). On COS agar *M. morganii* could not be differentiated from *E. coli* or *Salmonella* spp., as both grow in round shape and appear as greyish smooth colonies. But on McConkey and Coliform agar, colonies were colourless and white, respectively, also very similar to *Salmonella* spp. Thus, morphological similarities to *E. coli* and *Salmonella* might have an impact on identification of *M. morganii* in routine diagnostics.

Antibiotic susceptibility testing

All isolates were found to be multidrug resistant (resistance to more than three antimicrobial substances) by disc diffusion method with resistance patterns ranging from five to ten antibiotics (Table 2). All were resistant to amoxicillin, ampicillin, and colistin, but also to tilmicosin and tylosin. The combination of trimethoprim/sulfamethoxazole was the only antibiotic to which all five strains were susceptible. Resistance to β -lactam antibiotics in *Morganella* species is usually mediated by the presence of chromosomally encoded β -lactamases belonging to the AmpC β -lactamase family [28]. These β -lactamases are typically induced in the presence of β -lactam antibiotics. The strain PA17/10312 (flock A) was confirmed to be an AMP-C phenotype. In contrast to ESBLs, AmpC hydrolyses broad and extended-spectrum cephalosporins (cephamycins as well as to oxyimino- β -lactams) but are not inhibited by β -lactamase inhibitors such as clavulanic acid [29]. The strains PA18/15564 (flock B) and PA18/25921 (flock D) were confirmed to be ESBL phenotype and possible producers of type D carbapenemases or impermeability/porin loss, whereas the strains PA18/16407 (flock C) and strain PA19/9695 (flock E) were ESBL negative. Overall, the results are in agreement with earlier data describing ESBL producers as multiple drug resistant [30].

Sequencing and annotation of five poultry strains

The genome sequences of the five poultry strains displayed substantial variation in length and number of genes (Table 3), with the PA17/10312 being longer and having more genes compared to the other strains. To evaluate whether these differences might be due to missing annotations, we computed genome completeness using the program CheckM [31] for each assembly and obtained values of 100% for all five strains, implying that the observed variation in gene content cannot be attributed to uneven coverage of the assembled strains. Genome size and gene content were also comparable to the assemblies of other published strains [7–10]. Additionally, we observed variations in genome structure among the five strains (Fig. 1A): PA18/15564 and PA18/16407 showed high colinearity compared to the closest reference strain (Additional File 2: Supplementary Fig. 1A), while PA17/10312 displayed more genomic rearrangements (Additional File 2: Supplementary Fig. 1B). Finally, the genome of the five strains was aligned to the reference strain KT, for which the most comprehensive annotation is available [8]. A multiple alignment in circular form is displayed in Fig. 1B and shows a high level of conservation among the five poultry strains, with few shared gaps, corresponding to the rRNA ribosomal gene clusters, which are notoriously difficult to assemble due to their repetitive nature.

Phylogenetic analysis

To test the hypothesis of a phylogenetic clustering based on the host in *M. morganii*, a phylogenetic tree was constructed including the five sequenced poultry isolates together with all available complete genomes from NCBI, encompassing a total of 52 strains. The genome-based k-mer tree displayed two distinct subtrees (M1 and M2), with no apparent clustering according to the host they were isolated from (Fig. 2). Lack of association with host was also previously reported for *Salmonella enteritidis* [12] and *Providencia* species [13]. The resulting tree places the strains PA18/16407, PA18/15564, PA19/9695, PA18/25921 in cluster M1 and the PA17/10312 strain in cluster M2, consistent with the differences in genome size and number of genes previously described. Historically, *M. morganii* has been divided into two subspecies: *morganii* and *sibonii*, based on the ability to ferment trehalose and other biochemical features [32]. Thus, we hypothesized that clusters M1 and M2 simply reflect the known subspecies classification. To test this idea, the presence of the trehalose operon (*treR*, *treB* and *treP*) in all the strains was investigated. By this approach, the subspecies *sibonii* was assigned to the strains in which the trehalose operon was present. Comparing this classification to the observed phylogenetic clustering, we found that the *sibonii* strains were entirely allocated in cluster M2, except the H1R strain (Fig. 2), confirming that differentiation of subspecies can be mostly performed based on phylogenetic grouping. These results are consistent with the subspecies assignment achieved by MALDI-TOF MS (Table 1). To validate the accuracy of the k-mer tree reconstruction, an additional tree was built using a whole-genome based approach (Additional File 3: Supplementary Fig. 2). Comparison of both trees showed high similarity in topology and few differences: i) the NCTC12358 and TW17014 strains of cluster M2 are swapped in the whole-genome based tree compared to the k-mer tree; ii) the 640_MMOR strain is grouped together with NCTC12286 and 8066 in the whole-genome based tree, while in the k-mer tree constitutes a separate cluster. As these differences do not affect the conclusions derived from following analyses, the k-mer tree was used as a guide for the remainder of the analyses.

Detection of resistance genes

The sequences of all five poultry strains contained genes for resistance to aminocoumarins, beta-lactams, elfamycins, fluoroquinolones, rifampin and tetracycline. Resistance to phenicol was present in all the sequenced poultry strains, except in the chicken strain PA19/9695. Based on BLAST searches against the MEGARes database, the investigation of presence/absence of resistance genes in all 52 strains from Fig. 2 revealed that resistance to certain antibiotics is widespread in *Morganella*. This includes aminocoumarins, beta-lactams, elfamycins, fluoroquinolones, phenicol, rifampin and tetracyclines (Fig. 3). Other types of resistance show a more uneven distribution along the tree, with multiple cases of strain-specific gains, like resistance to glycopeptides, which is present in 5/52 strains with no apparent phylogenetic clustering. While this might suggest a putative origin by horizontal gene transfer, a comprehensive analysis of mechanisms of originations of resistance genes is outside the scope of this study. Additionally, we looked at resistance genes that were absent in the MEGARes database, but that were previously reported for other strains of *M. morganii*, such as genes involved in lipid A biosynthesis (*lpxA*, *lpxB*, *lpxC*, *lpxD*, *lpxH*, *lpxK*, *lpxL*, *lpxM*, *lpxP* and *lpxT*), which are responsible for

colistin resistance [9]. Such analyses detected *lpxABCDHKLMP*T in both geese strains, *lpxABCDHKLMT* in the chicken strain PA19/9695, *lpxABCDHKLMT* in the turkey strain PA18/25921 and *lpxABCDHLMT* in the goose strain PA17/10312. An additional copy of the *lpxP* gene was found in PA18/25921 and PA17/10312. Also, the *tetB* gene responsible for tetracycline resistance was found in all five poultry strains. No association was detected between resistance genes of a certain class and a specific phylogenetic cluster. These results are concordant with the phenotypic tests of antibiotic resistance (Table 2). Overall, the classes of resistance determined are consistent with a previously compiled review of resistance genes in *M. morganii* [3].

Detection of known virulence genes

In an initial approach, the sequences from the poultry strains were screened for known virulence genes based on the comparison with the KT strain, for which an exhaustive list of virulence genes is available. However, the gene-id based detection of genes from the KT strain was unfeasible due to incongruences between gene-ids from the KT strain and the ones deposited in the NCBI gene database, thus the 4-letter gene name was used to retrieve the genes of interest, resulting in a partial set of 125 genes for this analysis. Most of the known virulence genes present in this set were found in the sequences of the five poultry strains (Fig. 4). This not only indicates that virulence genes are very conserved, but also suggests that the annotations of the five poultry strains is highly complete. As this gene set did not include all known virulence genes, we additionally looked for the presence of specific classes of virulence genes of importance. Hemolysins are essential for pore-formation during invasion of host cells [33] and encoded by the *hlyCABD* operon. This operon was present and intact only in the chicken strain PA18/25921, while absent in PA18/16407, PA18/15564 and PA19/9695. In the sibonii strain PA17/10312, only *hlyC* and *hlyD* were present and distributed on different contigs. Variation in hemolysin composition was also observed in three recently sequenced human strains [10]. Ureases are also typically organized into the *ureABCFGD* operon and contribute to formation of urinary stones [34]. This operon was present and intact in all five strains. Finally, the capsule synthesis regulation genes *rscB*, *rscC*, *rscD* and *rscF* involved in host immune response [35] were also detected in all five strains, with an additional copy of *rscB* in PA18/25921. Interestingly, virulence genes in the category “toxins” displayed high variation in copy numbers among the five strains (Fig. 4). Manual scanning of the alignments also showed substantial variation in coding sequences exemplarily shown for the RtxA toxin in Additional File 4: Supplementary Fig. 3. For these reasons, a positive selection scan using PAML (see Materials and methods) was performed based on the hypothesis that toxins might evolve under positive selection in *M. morganii*. For this analysis only toxins present in all five strains were selected (Fig. 4) and it was found that four out of nine toxins (*tcaC*/*tcdB2*, *xptA1*, *xptB1*, *xptC1*) showed signs of positive selection in the PA17/10312 branch (Table 4). Notably, the insecticidal proteins *XptA1B1C1* had the strongest signal of positive selection. While the precise function of these toxins in *M. morganii* is unknown, it was shown in *Xenorhabdus spp.* and *Photorhabdus luminescens* that they facilitate the killing of the insect host via nematodes living in symbiosis with these bacteria [36]. Thus, positive selection in insecticidal toxins might reflect an evolutionary arms race among species.

Detection of novel virulence genes

To identify novel virulence genes, the proteomes of the five poultry strains were BLASTed to a curated set of *Enterobacteriaceae* virulence factors generated from the Virulence Factor DataBase (VFDB). For each strain, presence/absence of each virulence gene was computed, and virulence genes were classified according to different functional categories based on the VFDB functional annotation. To get an overview of the variation in individual virulence genes across the strains, the number of virulence genes of different functional categories in each strain was plotted sorted according to the phylogenetic position on the tree (Fig. 5). An increase in number of genes in strains of cluster M2 compared to the cluster M1 in the virulence categories adhesion, secretion system and hypothetical was observed (Fig. 5). As the strain PA17/10312 is also located in cluster M2, these results might explain the higher number of genes in the PA17/10312 strain compared to the other strains. Novel virulence genes were defined as those specific to cluster M2: a total of 39 novel virulence genes were found in this way (Table 4). To validate specificity to cluster M2, these genes were BLASTed to two outgroup species: *Providencia stuartii* and *Proteus mirabilis*, and only genes with no hits to both species were retained. This resulted in a final set of 26 novel virulence genes. Among these, *afaAB* [37], *pixC* [38], *sfaG* and *fyuA* [39] are known to be associated with adhesion and colonization of urinary tract in *E. coli*, while the intimin *fdeC* [40], the Shiga toxins *stx2A/stxA* [41], *escN* [42], *espD* [43] are connected with enteropathogenic strains of *E. coli*. This might suggest that some *M. morganii* strains of the *sibonii* subspecies might have acquired the ability to invade the enteric tract. With regard to phylogenetic distribution, 10/26 genes were shared by all strains in cluster M2, consistent with ancestral gain in cluster M2 and 12/26 genes were specific to a single strain, likely due to strain-specific horizontal gene transfer events. To find evidence for horizontal gene transfer the novel virulence genes were BLASTed against the whole bacterial protein database. Then, using the program Alienness [44], evidence for horizontal gene transfer (HGT) was detected in 6/26 (23%) of the novel virulence genes (*afaB*, *escU*, *escV*, *SG1030*, *stx2A*, *ycbV*), having an Alien Index (AI) > 30 (Additional File 5: Supplementary Table 2). These genes were assigned to the different donors of the class *Morganellaceae* (Additional File 5: Supplementary Table 2). While horizontal gene transfer of resistance genes was previously described in *M. morganii* [9], to our knowledge this is the first report of HGT of virulence genes in *M. morganii*.

Prediction of pathogenic potential for humans

Using PathogenFinder 1.1 [45] the pathogenic potential of the five poultry strains for humans was estimated. The approach used by this tool is based on the presence of group of genes that are frequently associated with human pathogenic bacteria. Results showed that the PA17/10312 was predicted to be pathogenic in humans (Probability = 0.64), while the four other strains were classified as non-pathogenic in humans (Probability = 0.57 for PA18/15564, 0.52 for PA18/16407, 0.57 for PA18/25921 and 0.54 for PA19/9695). These results are consistent with the higher number of virulence factors found in PA17/10312 compared to the other strains. It remains to be determined in an animal trial whether these findings also reflects the host-specific pathogenicity.

Annotation of mobile elements

Mobile elements such as plasmids, prophages and integrons are often responsible for transfer of resistance genes in bacteria [46, 47]. No evidence for plasmids was found in any of the five poultry strains using two independent approaches (see Materials and methods). In the published strains, no plasmids were found in the KT strain [8] but were detected in the INSRALV892 strain [7], while no information is available for the other sequenced strains [9, 10]. A total of the four active prophages were detected in PA17/10312, three in PA18/15564, two in PA18/16407, five in PA18/25921 and one in PA19/9695 (Additional File 6: Supplementary Table 3), none of them showing association with resistance genes. Two integrons were found in the PA19/9695 strain and one integron was found in the PA17/10312 strain, also not associated with resistance genes. Pathogenicity islands are another class of mobile elements, which are linked with virulence genes in a large number of bacteria [48]. A total of 18 pathogenicity islands were detected in PA17/10312, 16 in PA18/15564, 21 in PA18/16407, 22 in PA18/25921 and 23 in PA19/9695 (Additional File 7: Supplementary Table 4), associated, respectively, with 25, 9, 23, 16 and 24 virulence genes. The most abundant virulence categories associated with these genes were: “adherence” for 20/25 genes in PA17/10312, “adherence” for 5/9 genes in PA18/15564 and “iron uptake” for 11/23 genes in PA18/16407, “toxins” for 5/16 genes in PA18/25921 and “adherence” for 11/24 in PA19/9695. In summary, these results suggest high variability in gene content of pathogenicity islands, probably due to phenomena of genome rearrangements and lateral transfer.

Identification of strain-specific metabolic pathways

To study metabolic pathways, the annotations of the five poultry strains were submitted to BlastKOALA providing a global overview of metabolic pathways displayed as a metabolic network. This program also allows an easy comparison of metabolic networks among different organisms or strains. All metabolic pathways were entirely conserved among the five poultry strains. In addition, two novel pathways specific to the PA17/10312 strain were discovered (Fig. 6): sucrose catabolism and biosynthesis of betaine from choline. Betaine is an osmoprotectant that allows organisms to grow in high-osmolarity environments [49]. Some bacteria can synthesize betaine de novo from intermediates of central metabolism whereas others can synthesize it only from exogenously supplied choline [50]. Screening all the 52 *M. morganii* strains included in Fig. 2 revealed that only the FAM24091 strain (isolated from Switzerland cheese), which is the closest strain to PA17/10312, possess genes for betaine biosynthesis (BetA and BetB). No evidence for horizontal gene transfer was found, based on the absence of proteins from closely related species which would cluster together with BetA and BetB from PA17/10312 based upon BLAST analyses. Whether the determined pathways are beneficial for the bacteria *in vitro* and *in vivo* remains speculative for the moment and needs to be determined in additional studies.

Conclusion

In the actual study, five highly resistant *Morganella morganii* strains from diseased poultry flocks—geese, turkeys and chickens—were isolated in association with *E. coli*. Strains were identified by MALDI-TOF MS down to subspecies level. Micronaut revealed two strains to be ESBL and one strain to be AMP-C

phenotypes. To gather more detailed information about their association with the host, resistance profile, virulence and zoonotic potential, the genomes of the five strains were sequenced and compared with all available complete *M. morganii* genomes. The phylogenetic structure of *M. morganii* separates all strains into two main phylogenetic clusters (M1 and M2) that approximately reflect the subdivision of the *Morganella* clade into the known subspecies *morganii* and *sibonii*. However, the presence of one *sibonii* strain (H1r) in the M1 cluster implies that the characterization of subspecies in *M. morganii* purely based on the presence of the trehalose fermentation genes can be sometimes imprecise. We found that four of the avian strains (PA18/15564, PA18/16407, PA18/25921 and PA19/9695) were assigned as subsp. *morganii*, while the strain PA17/10312 was assigned to subsp. *sibonii*, implying no association between phylogenetic structure and host. Additionally, we showed that cluster M2 diverged through acquisition of novel virulence genes involved in adherence and secretion and that some of these genes might have originated by horizontal gene transfer from closely related *Morganellaceae* species. Furthermore, high sequence divergence in toxins was detected and evidence for positive selection for some of these genes was provided. Finally, we reported detailed resistance profiles for all so far sequenced *M. morganii* strains complementing the phenotypic resistance profile of the five poultry strains. Altogether, this study represents the first large-scale comparative study in *M. morganii* and will be a useful resource to aid future clinical management of infections and for advancing the field of comparative genomics of *Morganella*. Performing experimental studies would be crucial to further clarify the nature of *M. morganii* as an avian pathogen and its zoonotic potential.

Declarations

Ethics approval and consent to participate

Sampling was performed during post-mortem investigations and complies with national legislation (Tierversuchsgesetz – TVG 2012, §1). Furthermore, the study was performed in co-operation with veterinarians in charge of the respective farms who have agreements with farm owners for applying veterinary procedures.

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the NCBI repository (Accession IDs of the five sequenced poultry strains: WCIS00000000, WCIT00000000, WCIU00000000, WCIV00000000, WCIW00000000).

Competing interests

The authors declare that they have no competing interests

Funding

Not applicable

Authors' contributions

MA conceived and designed the study. CH collected the samples. MA performed the laboratory analyses. NP performed the bioinformatics analysis. MA and NP drafted the first version of the manuscript. All authors critically evaluated and interpreted the results. All authors provided feedback on the manuscript and read and approved the final version.

Acknowledgements

We thank Delfina Jandreski-Cvetkovics for her technical assistance.

References

1. O'Hara CM, Brenner FW, Miller JM. Classification, identification, and clinical significance of *Proteus*, *Providencia*, and *Morganella*. *Clinical Microbiology Reviews*. 2000.
2. Adeolu M, Alnajar S, Naushad S, Gupta RS. Genome-based phylogeny and taxonomy of the 'Enterobacteriales': Proposal for enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morgane. *Int J Syst Evol Microbiol*. 2016.
3. Liu H, Zhu J, Hu Q, Rao X. *Morganella morganii*, a non-negligent opportunistic pathogen. *International Journal of Infectious Diseases*. 2016.
4. Habila Mamman P, Kazeem HM, Raji MA, Nok AJ, Kwada J, Kwaga P. Preliminary Report on Isolation of *Morganella morganii* from Fatal Infections of Chickens in Kaduna State, Nigeria. *Int Invent J Med Med Sci*. 2014.
5. Zhao C, Tang N, Wu Y, Zhang Y, Wu Z, Li W, et al. First reported fatal *Morganella morganii* infections in chickens. *Vet Microbiol*. 2012.
6. Kilonzo-Nthenge A, Rotich E, Nahashon SN. Evaluation of drug-resistant Enterobacteriaceae in retail poultry and beef. *Poult Sci*. 2013.
7. Jones-Dias D, Clemente L, Moura IB, Sampaio DA, Albuquerque T, Vieira L, et al. Draft genomic analysis of an avian multidrug resistant *Morganella morganii* isolate carrying qnrD1. *Front Microbiol*. 2016.
8. Chen Y-T, Peng H-L, Shia W-C, Hsu F-R, Ken C-F, Tsao Y-M, et al. Whole-genome sequencing and identification of *Morganella morganii* KT pathogenicity-related genes. *BMC Genomics*. 2012.
9. Olaitan AO, Diene SM, Gupta SK, Adler A, Assous MV, Rolain JM. Genome analysis of NDM-1 producing *Morganella morganii* clinical isolate. *Expert Rev Anti Infect Ther*. 2014.
10. Minnullina L, Pudova D, Shagimardanova E, Shigapova L, Sharipova M, Mardanov A. Comparative Genome Analysis of Uropathogenic *Morganella morganii* Strains. *Front Cell Infect Microbiol*. 2019;9:167. doi:.

11. Ishii S, Meyer KP, Sadowsky MJ. Relationship between phylogenetic groups, genotypic clusters, and virulence gene profiles of *Escherichia coli* strains from diverse human and animal sources. *Appl Environ Microbiol.* 2007.
12. Lawson B, Franklinos LHV, Rodriguez-Ramos Fernandez J, Wend-Hansen C, Nair S, Macgregor SK, et al. *Salmonella enteritidis* ST183: Emerging and endemic biotypes affecting western European hedgehogs (*Erinaceus europaeus*) and people in Great Britain. *Sci Rep.* 2018.
13. Galac MR, Lazzaro BP. Comparative genomics of bacteria in the genus *Providencia* isolated from wild *Drosophila melanogaster*. *BMC Genomics.* 2012.
14. Alispahic M, Christensen H, Bisgaard M, Hess M, Hess C. MALDI-TOF mass spectrometry confirms difficulties in separating species of the *Avibacterium* genus. *Avian Pathol.* 2014.
15. Seemann T. Prokka. Rapid prokaryotic genome annotation. *Bioinformatics.* 2014.
16. Carattoli A, Zankari E, Garcíá-Fernández A, Larsen MV, Lund O, Villa L, et al. In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother.* 2014.
17. Orlek A, Phan H, Sheppard AE, Doumith M, Ellington M, Peto T, et al. A curated dataset of complete Enterobacteriaceae plasmids compiled from the NCBI nucleotide database. *Data Br.* 2017.
18. Song W, Sun H-X, Zhang C, Cheng L, Peng Y, Deng Z, et al. Prophage Hunter: an integrative hunting tool for active prophages. *Nucleic Acids Res.* 2019.
19. Cury J, Jové T, Touchon M, Néron B, Rocha EP. Identification and analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Res.* 2016.
20. Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, et al. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* 2017.
21. Treangen TJ, Ondov BD, Koren S, Phillippy AM. Rapid Core-Genome Alignment and Visualization for Thousands of Microbial Genomes. *bioRxiv.* 2014.
22. Lakin SM, Dean C, Noyes NR, Dettenwanger A, Ross AS, Doster E, et al. MEGARes: An antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Res.* 2017.
23. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol.* 2017.
24. Löytynoja A. Phylogeny-aware alignment with PRANK. *Methods Mol Biol.* 2014.
25. Yang Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 2007.
26. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J Mol Biol.* 2016.
27. Kanehisa M. The KEGG Database. 2003.
28. Jacoby GA. AmpC-Lactamases. *Clin Microbiol Rev.* 2009.
29. Bush LM, Calmon J, Johnson CC. Newer penicillins and beta-lactamase inhibitors. *Infect Dis Clin North Am.* 1995.

30. Bradford PA. Extended-spectrum β -lactamases in the 21st century: Characterization, epidemiology, and detection of this important resistance threat. *Clinical Microbiology Reviews*. 2001.
31. 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.
32. Truberg Jensen K, Frederiksen W, Hickman-Brenner FW, Steigerwalt AG, Riddle CF, Brenner DJ. Recognition of *Morganella* Subspecies, with Proposal of *Morganella morganii subsp. morganii subsp. nov.* and *Morganella morganii subsp. sibonii subsp. nov.* *Int J Syst Bacteriol*. 1992.
33. Minnullina LF, Al Allak MH, Khaitlina SY, Sharipova MR, Mardanov AM. The hemolytic properties of clinical isolates of *Morganella morganii*. *Res J Pharm Biol Chem Sci*. 2016.
34. Jones BD, Mobley HLT. Genetic and biochemical diversity of ureases of *Proteus*, *Providencia*, and *Morganella* species isolated from urinary tract infection. *Infect Immun*. 1987.
35. Willis LM, Whitfield C. Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by ABC transporter-dependent pathways. *Carbohydr Res*. 2013.
36. Morgan JAW, Sergeant M, Ellis D, Ousley M, Jarrett P. Sequence Analysis of Insecticidal Genes from *Xenorhabdus nematophilus* PMFI296. *Appl Environ Microbiol*. 2001.
37. Servin AL. Pathogenesis of Afa/Dr diffusely adhering *Escherichia coli*. *Clinical Microbiology Reviews*. 2005.
38. Worpel DJ, Beatson SA, Totsika M, Petty NK, Schembri MA. Chaperone-usher fimbriae of *Escherichia coli*. *PLoS One*. 2013;8:e52835–5. doi:.
39. 10.1128/IAI.00752-12
Spurbeck RR, Dinh PC, Walk ST, Stapleton AE, Hooton TM, Nolan LK, et al. *Escherichia coli* Isolates That Carry vat, fyuA, chuA, and yfcV Efficiently Colonize the Urinary Tract. *Infect Immun*. 2012;80:4115 LP – 4122. doi:.
40. Easton DM, Allsopp LP, Phan MD, Moriel DG, Goh GK, Beatson SA, et al. The intimin-like protein FdeC is regulated by H-NS and temperature in enterohemorrhagic *Escherichia coli*. *Appl Environ Microbiol*. 2014.
41. Bai X, Fu S, Zhang J, Fan R, Xu Y, Sun H, et al. Identification and pathogenomic analysis of an *Escherichia coli* strain producing a novel Shiga toxin 2 subtype. *Sci Rep*. 2018;8:6756. doi:.
42. Andrade A, Pardo JP, Espinosa N, Perez-Hernandez G, Gonzalez-Pedrajo B. Enzymatic characterization of the enteropathogenic *Escherichia coli* type III secretion ATPase EscN. *Arch Biochem Biophys*. 2007;468:121–7.
43. Chatterjee A, Caballero-Franco C, Bakker D, Totten S, Jardim A. Pore-forming Activity of the *Escherichia coli* Type III Secretion System Protein EspD. *J Biol Chem*. 2015;290:25579–94.
44. Rancurel C, Legrand L, Danchin EGJ. Alienness. Rapid detection of candidate horizontal gene transfers across the tree of life. *Genes (Basel)*. 2017.
45. Cosentino S, Voldby Larsen M, Møller Aarestrup F, Lund O. PathogenFinder - Distinguishing Friend from Foe Using Bacterial Whole Genome Sequence Data. *PLoS One*. 2013.

46. Casjens S. Prophages and bacterial genomics: what have we learned so far? *Mol Microbiol.* 2003.
47. Gillings MR. Integrons: Past, Present, and Future. *Microbiol Mol Biol Rev.* 2014.
48. Hacker J, Kaper JB. Pathogenicity Islands and the Evolution of Microbes. *Annu Rev Microbiol.* 2002.
49. Wood JM, Bremer E, Csonka LN, Kraemer R, Poolman B, Van der Heide T, et al. Osmosensing and osmoregulatory compatible solute accumulation by bacteria. In: *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology.* 2001.
50. Wargo MJ. Homeostasis and catabolism of choline and glycine betaine: Lessons from *Pseudomonas aeruginosa.* *Applied and Environmental Microbiology.* 2013.

Tables

Table 1 - Details on origin and identification results of *M. morganii* isolates

| Flock | Strain Designation | Bird species | Age | Organ | Bacterial isolates |
|-------|--------------------|--------------------|----------|-------------------------------|---|
| A | PA17/10312 | Goose | 5 weeks | brain | <i>Morganella morganii</i> subsp. <i>sibonii</i> , <i>Escherichia coli</i> |
| B | PA18/15564 | Goose | 10 weeks | heart | <i>Morganella morganii</i> subsp. <i>morganii</i> , <i>Escherichia coli</i> , <i>Riemerella anatipestifer</i> |
| C | PA18/16407 | Goose | 40 weeks | heart, liver, intestine, lung | <i>Morganella morganii</i> subsp. <i>morganii</i> , <i>Escherichia coli</i> , <i>Clostridium perfringens</i> |
| D | PA18/25921 | Turkey | 1 day | heart | <i>Morganella morganii</i> subsp. <i>morganii</i> , <i>Escherichia coli</i> |
| E | PA19/9695 | Layer parent stock | 32 weeks | heart | <i>Morganella morganii</i> subsp. <i>morganii</i> , <i>Escherichia coli</i> |

Table 2 - Antibiotic susceptibility testing by disc diffusion method - R = Resistant, S = Susceptible, I = Intermediate.

| Antibiotics (μ g) | PA17/10312 | PA18/15564 | PA18/16407 | PA18/25921 | PA19/9695 |
|------------------------------------|------------|------------|------------|------------|-----------|
| Amoxicillin (10) | R | R | R | R | R |
| Ampicillin (10) | R | R | R | R | R |
| Colistin (10) | R | R | R | R | R |
| Doxycycline (30) | R | I | R | R | S |
| Enrofloxacin (5) | S | I | R | R | R |
| Neomycin (30) | I | I | I | S | S |
| Oxalic Acid (2) | S | S | S | R | R |
| Spectinomycin (100) | S | S | S | R | S |
| Tetracycline (30) | I | I | R | R | S |
| Tilmicosin (15) | R | R | R | R | R |
| Trimethoprim/Sulfamethoxazole (25) | S | S | S | S | S |
| Tylosin (150) | R | R | R | R | R |

Table 3 – Genomic features of the five sequenced poultry strains

| | PA17/10312 | PA18/15564 | PA18/16407 | PA18/25921 | PA19/9695 |
|-----------------------|------------|------------|------------|------------|-----------|
| Number of contigs | 78 | 59 | 47 | 548 | 667 |
| Total length | 4,085,866 | 3,675,879 | 3,702,167 | 4,066,432 | 3,988,823 |
| N50 | 244,109 | 403,330 | 989,625 | 403,126 | 49,931 |
| GC content | 50.2% | 51.2% | 51.2% | 51.0% | 51.6% |
| Genes | 3840 | 3452 | 3490 | 3812 | 3756 |
| Coding genes | 3784 | 3393 | 3431 | 3718 | 3671 |
| Genome completeness | 100% | 100% | 100% | 100% | 100% |
| Contamination score* | 0.00 | 0.54 | 0.00 | 1.92 | 2.75 |
| Strain heterogeneity* | 0.00 | 0.00 | 0.00 | 37.50 | 50.00 |

* For a detailed description of these metrics, see the CheckM paper [31]

Table 4 – List of toxins and associated p-values for the positive-selection test in the PA17/10312 branch. Significant genes are highlighted in grey.

| Toxin | P-value |
|--------------|---------|
| RtxA | 1.0000 |
| XaxA | 0.9980 |
| XaxB | 0.9546 |
| HlyD-family | 1.0000 |
| tcaC / tcdB2 | 0.0200 |
| tccB3 | 1.0000 |
| XptA1 | 0.0003 |
| XptB1 | 0.0010 |
| XptC1 | 0.0036 |

Table 5 – List of novel virulence genes involved in specific to the M2 cluster

| Category | Genes |
|------------------|---|
| Adherence | <i>afaAB, eae, fdeC, pagN, papADF, pixACH, sfaG, sfpAC, stiC, stkB, ycbV</i> |
| Secretion system | <i>aatA, EC55989_3333, ECP_0235, ECVR50_0257, escNSTUV, espD, LF82_024, mxiH, O3M_04340, pipB, prgI, sipB</i> |
| Hypothetical | SG1030 |
| Toxins | <i>clbN, stx2A, stxA</i> |
| Iron uptake | <i>fyuA</i> |
| Others | STM0570 |

Additional Material

Additional File 1: Supplementary Table 1 – MICRONAUT antimicrobial substances and concentrations

Additional File 2: Supplementary Figure 1 – Genomic alignment computed with MAUVE between two of the sequenced poultry strains and the closest strain on the tree shown in figure 2: **(A)** PA18/15564 aligned to the DG56-16 strain, **(B)** PA17/10312 aligned to the INSRALV892 strain.

Additional File 3: Supplementary Figure 2 – Tree of the 52 *M. morganii* strains analyzed in this study derived by alignment of the core genome using the program *parsnp*. The KT strain was used as a reference for the multiple genome alignment.

Additional File 4: Supplementary Figure 3 – Multiple alignment of an internal region of the RtxA toxin in the five poultry strains showing high sequence divergence.

Additional File 5: Supplementary Table 2 – List of candidate genes in the PA17/10312 strain with evidence for horizontal gene transfer from closely-related *Morganellaceae* species

Additional File 6: Supplementary Table 3 – Annotation of prophages in the five sequenced poultry strains

Additional File 7: Supplementary Table 4 – Annotation of pathogenicity islands in the five sequenced poultry strains

Figures

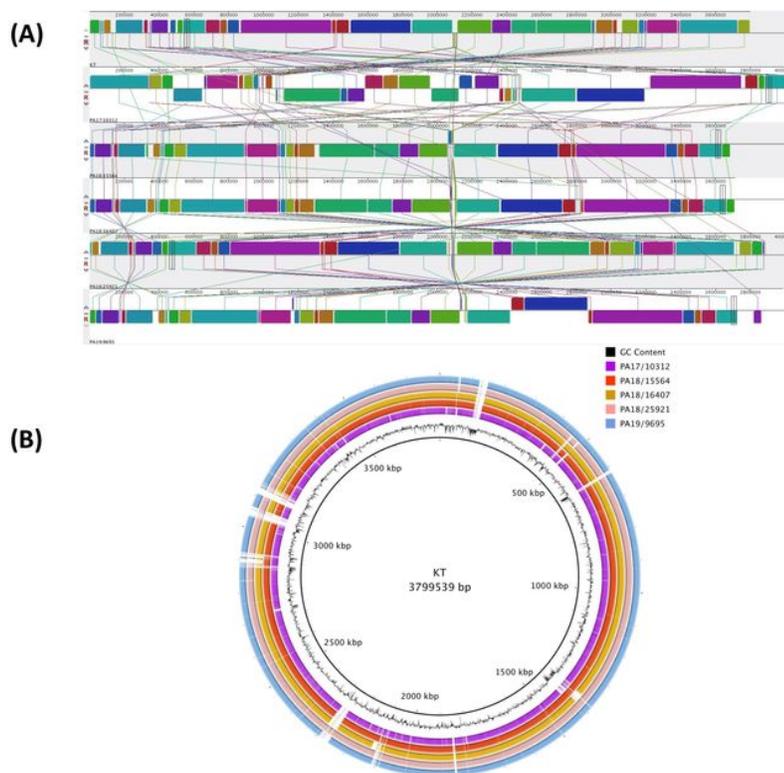


Figure 1

A) Multiple genome alignment of the five sequenced poultry strains and the KT reference strain B) Circular genomic map constructed using BRIG displaying a multiple genome alignment of the five poultry strains and the KT reference strain.

Figure 3

Phylogenetic distribution of resistance genes in the 52 analyzed *M. morganii* strains ordered according to the position in the tree from Figure 2. The five sequenced poultry strains are highlighted in yellow. Every cell contains the number of gene copies involved in a certain type of resistance for each strain. Cells are coloured using conditional formatting to facilitate readability.

| Gene ID | PA18/16407 | PA18/15564 | PA19/8695 | PA18/25921 | PA17/10312 |
|--|------------|------------|-----------|------------|------------|
| Fimbriae, pili, and adhesion proteins | | | | | |
| misp1 | 1 | 1 | 1 | 1 | 1 |
| hofB | 1 | 1 | 1 | 1 | 1 |
| hofC | 1 | 1 | 1 | 1 | 1 |
| ppdA | 1 | 1 | 1 | 1 | 1 |
| ppdB | 1 | 1 | 1 | 1 | 1 |
| Motility and chemotaxis systems | | | | | |
| cheA | 1 | 1 | 1 | 1 | 1 |
| cheW | 1 | 1 | 1 | 1 | 1 |
| cheD | 1 | 1 | 1 | 1 | 1 |
| tap | 1 | 1 | 1 | 1 | 1 |
| cheR | 1 | 1 | 1 | 1 | 1 |
| cheY | 1 | 1 | 1 | 1 | 1 |
| cheZ | 1 | 1 | 1 | 1 | 1 |
| lysR | 1 | 1 | 1 | 1 | 1 |
| CrcB | 1 | 1 | 1 | 1 | 1 |
| SDR | 1 | 1 | 1 | 1 | 1 |
| umoA | 1 | 1 | 1 | 1 | 1 |
| umoB | 1 | 1 | 1 | 1 | 1 |
| umoC | 1 | 1 | 1 | 1 | 1 |
| umoD | 1 | 1 | 1 | 1 | 1 |
| rsaA | 1 | 1 | 1 | 1 | 1 |
| rsbB | 1 | 1 | 1 | 1 | 1 |
| resB | 1 | 1 | 1 | 1 | 1 |
| resC | 1 | 1 | 1 | 1 | 1 |
| resD | 1 | 1 | 1 | 1 | 1 |
| Iron acquisition | | | | | |
| hmuC | 1 | 1 | 1 | 1 | 1 |
| hmuT | 1 | 1 | 1 | 1 | 1 |
| hmuU | 1 | 1 | 1 | 1 | 1 |
| hmuV | 1 | 1 | 1 | 1 | 1 |
| afuA | 1 | 1 | 1 | 1 | 1 |
| afuB | 1 | 1 | 1 | 1 | 1 |
| afuC | 1 | 1 | 1 | 1 | 1 |
| feoA | 1 | 1 | 1 | 1 | 1 |
| feoB | 1 | 1 | 1 | 1 | 1 |
| btuC | 1 | 1 | 1 | 1 | 1 |
| btuB | 1 | 1 | 1 | 1 | 1 |
| btuD | 1 | 1 | 1 | 1 | 1 |
| yfeA | 1 | 1 | 1 | 1 | 1 |
| yfeB | 1 | 1 | 1 | 1 | 1 |
| yfeC | 1 | 1 | 1 | 1 | 1 |
| yfeD | 1 | 1 | 1 | 1 | 1 |
| ireA | 1 | 1 | 1 | 1 | 1 |
| Zinc acquisition | | | | | |
| znuA | 1 | 1 | 1 | 1 | 1 |
| znuB | 1 | 1 | 1 | 1 | 1 |
| znuC | 1 | 1 | 1 | 1 | 1 |
| Gene ID | | | | | |
| Type III secretion system | | | | | |
| YscA | 1 | 1 | 1 | 1 | 1 |
| YscC | 1 | 1 | 1 | 1 | 2 |
| IpaB | 1 | 1 | 0 | 1 | 1 |
| Two-component systems | | | | | |
| gseB | 1 | 1 | 1 | 1 | 1 |
| gseC | 1 | 1 | 1 | 1 | 1 |
| yedW | 1 | 1 | 1 | 1 | 1 |
| yedV | 1 | 1 | 1 | 1 | 1 |
| BarA | 1 | 1 | 1 | 1 | 1 |
| UhpA | 1 | 1 | 1 | 1 | 1 |
| Regulation | | | | | |
| rpoD | 1 | 1 | 1 | 1 | 1 |
| rpoH | 1 | 1 | 1 | 1 | 1 |
| rpoS | 1 | 1 | 1 | 1 | 1 |
| rpoE | 1 | 1 | 1 | 1 | 1 |
| rpoF | 1 | 1 | 1 | 1 | 1 |
| FecR | 1 | 1 | 1 | 1 | 1 |
| ryhB | 1 | 1 | 2 | 1 | 1 |
| IgA protease | | | | | |
| zarpA | 1 | 1 | 1 | 1 | 1 |
| zarpB | 1 | 1 | 1 | 1 | 1 |
| zarpC | 1 | 1 | 1 | 1 | 1 |
| zarpD | 1 | 1 | 1 | 1 | 1 |
| Haemolysins | | | | | |
| hmpA | 1 | 1 | 1 | 1 | 1 |
| hmpB | 1 | 1 | 1 | 1 | 1 |
| Superoxide stress | | | | | |
| katA | 1 | 1 | 1 | 1 | 1 |
| sos5 | 1 | 1 | 1 | 1 | 1 |
| sosC | 1 | 1 | 1 | 1 | 0 |
| sodB | 1 | 1 | 1 | 1 | 1 |
| oxyR | 1 | 1 | 1 | 1 | 1 |
| sodA | 1 | 1 | 1 | 1 | 1 |
| Late-stationary phase stress | | | | | |
| pspA | 1 | 1 | 1 | 1 | 1 |
| pspB | 1 | 1 | 1 | 1 | 1 |
| pspC | 1 | 1 | 1 | 1 | 1 |
| Gene ID | | | | | |
| Urease hydrolysis and putrescine production | | | | | |
| ureA | 1 | 1 | 1 | 1 | 1 |
| ureB | 1 | 1 | 1 | 1 | 1 |
| ureC | 1 | 1 | 1 | 1 | 1 |
| ureD | 1 | 1 | 1 | 1 | 1 |
| ureE | 1 | 1 | 1 | 1 | 1 |
| ureF | 1 | 1 | 1 | 1 | 1 |
| ureG | 1 | 1 | 1 | 1 | 1 |
| ureH | 1 | 1 | 1 | 1 | 1 |
| speF | 2 | 2 | 2 | 2 | 2 |
| potE | 1 | 1 | 1 | 1 | 1 |
| carA | 1 | 1 | 1 | 1 | 1 |
| carB | 1 | 1 | 1 | 1 | 1 |
| argI | 1 | 1 | 1 | 1 | 1 |
| argG | 1 | 1 | 1 | 1 | 1 |
| argH | 1 | 1 | 1 | 1 | 1 |
| speA | 1 | 1 | 1 | 1 | 1 |
| speB | 1 | 1 | 1 | 1 | 1 |
| Toxins | | | | | |
| RtxA | 1 | 1 | 1 | 1 | 1 |
| XaaA | 1 | 1 | 1 | 1 | 1 |
| XaaB | 1 | 1 | 1 | 1 | 1 |
| HlyA | 0 | 0 | 0 | 1 | 0 |
| HlyB | 0 | 0 | 0 | 1 | 0 |
| HlyC | 0 | 0 | 0 | 1 | 1 |
| HlyD | 0 | 0 | 0 | 1 | 1 |
| HlyD-family | 1 | 1 | 1 | 1 | 1 |
| tccA | 0 | 1 | 0 | 0 | 1 |
| tccC / tccB2 | 1 | 2 | 1 | 2 | 1 |
| tccB3 | 1 | 1 | 1 | 1 | 1 |
| xptA1 | 2 | 2 | 2 | 2 | 3 |
| xptB1 | 2 | 2 | 1 | 3 | 1 |
| xptC1 | 1 | 2 | 1 | 2 | 1 |
| Ethanolamine utilisation system | | | | | |
| eusS | 1 | 1 | 1 | 1 | 1 |
| eusP | 1 | 1 | 1 | 1 | 1 |
| eusQ | 1 | 1 | 1 | 1 | 1 |
| eusT | 1 | 1 | 1 | 1 | 1 |
| eusD | 1 | 1 | 1 | 1 | 1 |
| eusM | 1 | 1 | 1 | 1 | 1 |
| eusN | 1 | 1 | 1 | 1 | 1 |
| eusE | 1 | 1 | 1 | 1 | 1 |
| eusJ | 1 | 1 | 1 | 1 | 1 |
| eusG | 1 | 1 | 1 | 1 | 1 |
| eusH | 1 | 1 | 1 | 1 | 1 |
| eusA | 1 | 1 | 1 | 1 | 1 |
| eusB | 1 | 1 | 1 | 1 | 1 |
| eusC | 1 | 1 | 1 | 1 | 1 |
| eusL | 1 | 1 | 1 | 1 | 1 |
| eusK | 1 | 1 | 1 | 1 | 1 |
| pduS | 1 | 1 | 1 | 1 | 1 |
| pduT | 1 | 1 | 1 | 1 | 1 |
| eusR | 1 | 1 | 1 | 1 | 1 |

Figure 4

Comparison of known virulence genes in the five sequenced poultry strains using the KT strain as reference.

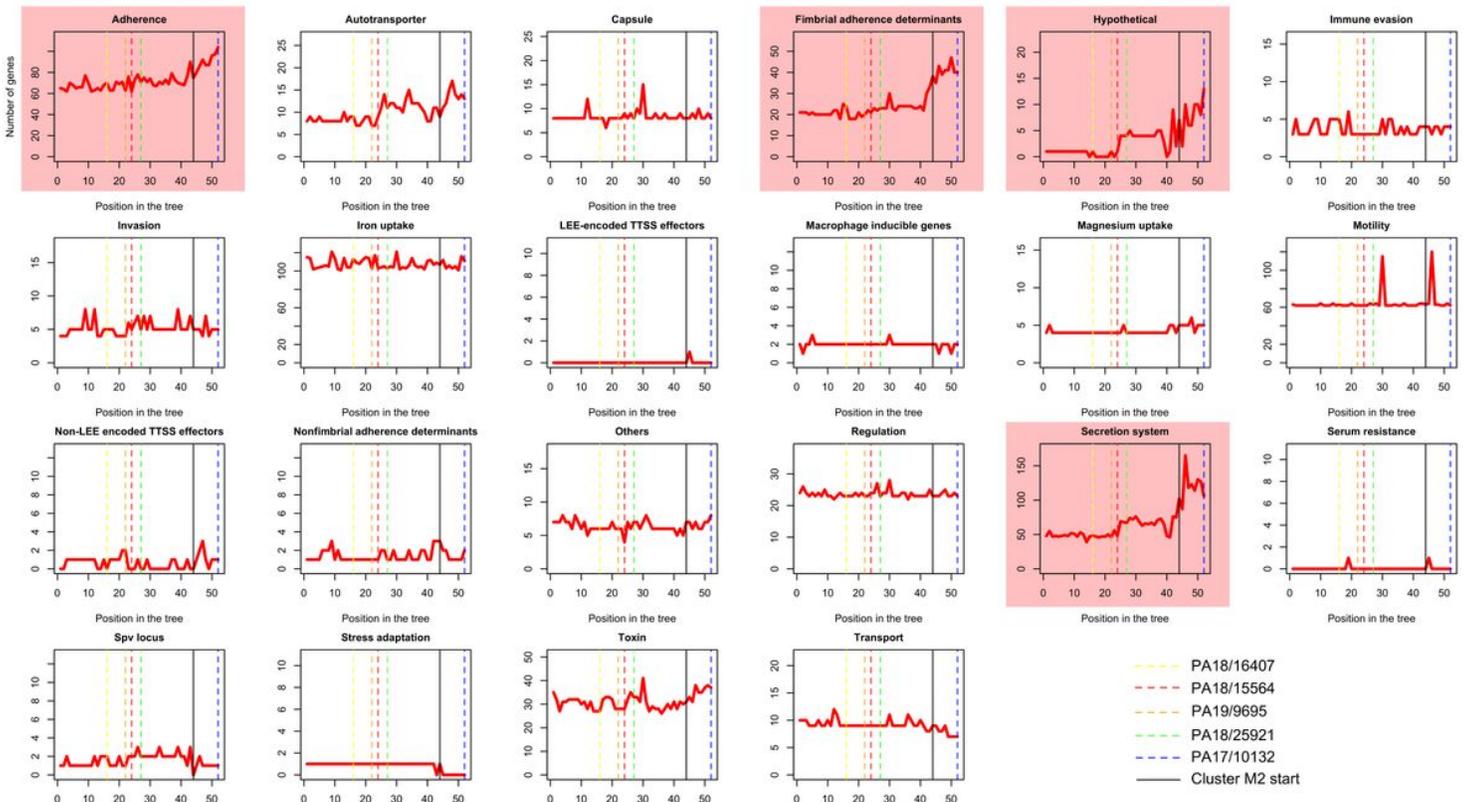


Figure 5

Phylogenetic distribution of virulence genes of different virulence categories as defined from the VFDB – each graph shows the number of genes per strain for every virulence category. Axes labels are shown in the top-left graph only – strains are sorted according to the phylogenetic position on the tree in Figure 2. The position of the five avian strains in each graph is highlighted by vertical dashed lines of different colours. Graphs showing variation between clusters M1 and M2 are highlighted by a pink background.

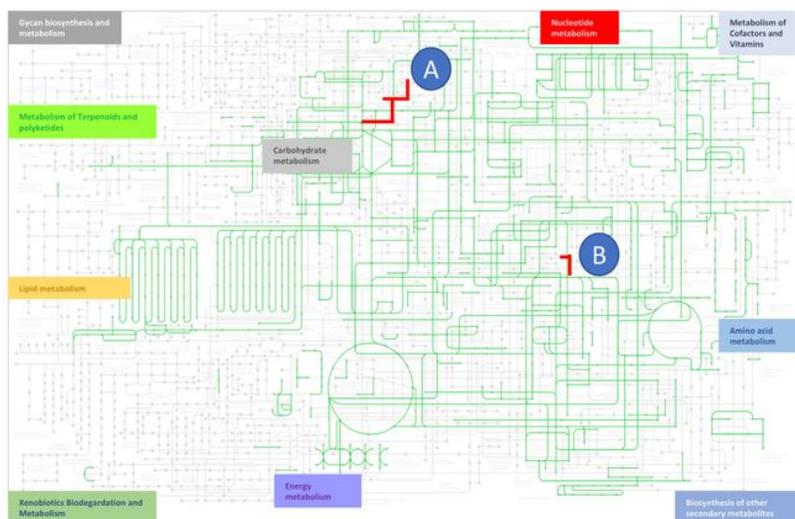


Figure 6

Metabolic network of *M. morganii* strain PA17/10312 showing two novel pathways specific to this strain: (A) Sucrose catabolism, sucrose => glucose, (B) Betaine biosynthesis, betaine => choline.

Supplementary Files

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

- [AdditionalFile7SupplementaryTable4.XLSX](#)
- [AdditionalFile2SupplementaryFigure1.JPEG](#)
- [AdditionalFile1SupplementaryTable1.DOCX](#)
- [AdditionalFile5SupplementaryTable2.DOCX](#)
- [AdditionalFile4SupplementaryFigure3.JPEG](#)
- [AdditionalFile3SupplementaryFigure2.JPEG](#)
- [AdditionalFile6SupplementaryTable3.DOCX](#)