DOI: 10.1590/S0100-736X2014000200006

In silico phylogenetic and virulence gene profile analyses of avian pathogenic *Escherichia coli* genome sequences[1]

Thaís C.G. Rojas[2], Renato P. Maluta[2], Luciano V. Koenigkan[3] and Wanderley Dias da Silveira[2]*

ABSTRACT.- Rojas T.C.G., Maluta R.P., Koenigkan L.V. & Dias da Silveira W. 2014. ***In silico* phylogenetic and virulence gene profile analyses of avian pathogenic *Escherichia coli* genome sequences**. *Pesquisa Veterinária Brasileira 34(2):129-133*. Departamento de Genética, Evolução e Bioagentes, Instituto de Biologia, Universidade de Campinas, Cx. Postal 6109, Campinas, SP 13083-970, Brazil. E-mail: wds@unicamp.br

Avian pathogenic *Escherichia coli* (APEC) infections are responsible for significant losses in the poultry industry worldwide. A zoonotic risk has been attributed to APEC strains because they present similarities to extraintestinal pathogenic *E. coli* (ExPEC) associated with illness in humans, mainly urinary tract infections and neonatal meningitis. Here, we present *in silico* analyses with pathogenic *E. coli* genome sequences, including recently available APEC genomes. The phylogenetic tree, based on multi-locus sequence typing (MLST) of seven housekeeping genes, revealed high diversity in the allelic composition. Nevertheless, despite this diversity, the phylogenetic tree was able to cluster the different pathotypes together. An *in silico* virulence gene profile was also determined for each of these strains, through the presence or absence of 83 well-known virulence genes/traits described in pathogenic *E. coli* strains. The MLST phylogeny and the virulence gene profiles demonstrated a certain genetic similarity between Brazilian APEC strains, APEC isolated in the United States, UPEC (uropathogenic *E. coli*) and diarrheagenic strains isolated from humans. This correlation corroborates and reinforces the zoonotic potential hypothesis proposed to APEC.

INDEX TERMS: Avian pathogenic *Escherichia coli* (APEC), Multi-locus Sequence Typing (MLST), phylogenetic tree, virulence genes.

RESUMO.- [**Análises *in silico* da filogenia e do perfil de genes associados à virulência, dos genomas de linhagens de *Escherichia coli* de origem aviária.**] As infecções causadas por linhagens de *Escherichia coli* de origem aviária (APEC) são responsáveis por perdas significativas na indústria avícola em todo mundo. Risco zoonótico tem sido atribuído às linhagens APEC, devido às semelhanças existentes entre elas e linhagens de *E. coli* patogênicas extraintestinais (ExPEC) de origem humana, causadoras de infecções no trato urinário e meningite neonatal. Neste trabalho, apresentamos os resultados de análises *in silico* feitas a partir dos genomas de linhagens patogênicas de *E. coli*, incluindo genomas recentemente obtidos de linhagens APEC. Uma árvore filogenética foi obtida, com base na tipagem de sequência multilocus (MLST) de sete genes essenciais, revelando alta diversidade na composição de alelos, mas ainda assim possibilitando o agrupamento dos diferentes patótipos. Foi determinado também, para cada linhagem, o perfil gênico, por meio da presença ou ausência de 83 genes associados à virulência. A árvore filogenética e o perfil gênico demonstraram que existem semelhanças genéticas entre cepas APEC brasileiras, APEC isolada nos Estados Unidos, UPEC (uropathogenic *E. coli*) e linhagens produtoras de diarreia em humanos. Essa correlação corrobora e reforça a hipótese de que linhagens APEC apresentam potencial risco zoonótico.

TERMOS DE INDEXAÇÃO: *Escherichia coli* de origem aviária (APEC), tipagem de sequência multilocus (MLST), árvore filogenética, genes associados à virulência.

INTRODUCTION

Escherichia coli associated with extraintestinal diseases (ExPEC), include avian pathogenic *E. coli* (APEC), neonatal

[1] Received on September 27, 2013.

Accepted for publication on January 7, 2014.

[2] Departamento de Genética, Evolução e Bioagentes, Instituto de Biologia, Universidade de Campinas (Unicamp), Cx. Postal 6109, Campinas, SP 13083-970, Brazil. *Corresponding author: wds@unicamp.br

[3] Embrapa Informática Agropecuária, Avenida Dr. André Tosello 209, Barão Geraldo, Cx. Postal 6041, Campinas, SP 13083-886.

meningitis *E. coli* (NMEC) and uropathogenic *E. coli* (UPEC) (Dziva & Stevens 2008). APEC infections are collectively denominated colibacillosis. The infection may be local or systemic, and has been responsible for severe losses in the poultry industry worldwide. Losses are due to carcass condemnation, mortality and decrease in egg production (Dho--Moulin & Fairbrother 1999, Barnes et al. 2008).

Previous studies demonstrated that APEC, NMEC and UPEC strains share common virulence genes, overlapped O serogroups and phylogenetic groups (Rodriguez-Siek et al. 2005, Ewers et al. 2007, Moulin-Schouleur et al. 2007). These characteristics were corroborated by multilocus sequence typing (MLST) and genome comparison data (Johnson et al. 2007, Moulin-Schouleur et al. 2007) and led to a further hypothesis that APEC strains could act as UPEC or NMEC and, therefore, constitute a zoonotic risk. A recent work (Dziva et al. 2013) demonstrated that an APEC strain was more closely related to a human ST23 ETEC (enterotoxigenic *E. coli*) than to strain APEC O1. This suggests that the core genome of ST23 strains has the potential to generate variants able to cause disease on avian or human, depending of the accessory genome. For urinary tract infections and new-born meningitis, it has been hypothesized that APEC could access and colonize the human colon via the ingestion of contaminated poultry and then reach the urinary tract or the central nervous system (Rodriguez-Siek et al. 2005). Thereby, APEC could act as a cause of human ExPEC infections or as a virulence gene reservoir for human ExPEC strains, thus representing a possible zoonotic risk (Rodriguez-Siek et al. 2005, Moulin-Schouleur et al. 2007).

The availability of APEC genome sequences (Johnson et al. 2007, Rojas et al. 2012, 2013) permits the comparison of genome contents between APEC and human ExPEC strains, and helps to understand the evolutionary processes involved in shaping the phenotypes of different *E. coli* pathotypes. The phylogeny (based on MLST) and the cluster (based on the presence or absence of virulence genes) analyses, may be helpful to highlight the similarities between these different strains, and the potential nonspecific host relationships. In this work, the virulence gene profiles and the phylogenetic relationships of APEC strains were analyzed and compared to other previously published human pathogenic strains, to verify possible genetic similarities and the potential pathogenicity risk for birds and human beings.

MATERIALS AND METHODS

The MLST tree and the virulence gene profile analysis were performed with 14 pathogenic *Escherichia coli* genome sequences. The strains and their respective plasmids were chosen because they are representative of the pathotypes herein studied (Table 1).

MLST analysis, conducted *in silico*, were performed with all the allele sequences of the seven housekeeping genes *adk* (adenylate kinase), *fumC* (fumarate hydratase), *gyrB* (DNA gyrase), *icd* (isocitrate/isopropylmalate dehydrogenase), *mdh* (malate dehydrogenase), *purA* (adenylosuccinate dehydrogenase) and *recA* (ATP/GTP-binding motif) downloaded from the MLST Databases at the ERI, University College Cork (http://mlst.ucc.ie/mlst/dbs/Ecoli/Downloads_HTML). For each gene, all alleles were searched against the genomes sequences through BLAST tool, on a local platform against a local database, using the software BLAST 2.2.26 with the application blastn. The allele with the lowest e-value was then considered as the allele present in the strain. For each strain, the sequences of the correspondent alleles, for the seven housekeeping genes, were concatenated. All concatenated sequences were aligned with MUSCLE (Edgar 2004). The phylogenetic tree was constructed by the Maximum Likelihood method using MEGA 5.10.

Table 1. *Escherichia coli* strains and plasmids analyzed by phylogenetic MLST tree and virulence gene profiles

Strain / Plasmid	Pathotype	GenBank accession number	Reference
SEPT362	APEC[a]	AOGL00000000.1	(Rojas et al. 2013)
S17	APEC	AOGN00000000.1	(Rojas et al. 2013)
O08	APEC	AOGM00000000.1	(Rojas et al. 2013)
SCI-07	APEC	AJFG00000000.1	(Rojas et al. 2012)
APEC O1	APEC	NC_008563.1	(Johnson et al. 2007)
pAPEC-O1-ColBM (plasmid)	APEC	NC_009837.1	(Johnson et al. 2006a)
pAPEC-O1-R (plasmid)	APEC	NC_009838.1	(Johnson et al. 2006b)
O157:H7 str. EDL933	EHEC[b]	NC_002655	(Perna et al. 2001)
O157:H7 str. Sakai	EHEC	NC_002695.1	(Bergholz et al. 2007)
O157:H7 str. Sakai pO157 (plasmid)	EHEC	NC_002128.1	(Makino et al. 1998)
O157:H7 str. Sakai pOSAK1 (plasmid)	EHEC	NC_002127.1	(Makino et al. 1998)
042	EAEC[c]	FN554766	(Chaudhuri et al. 2010)
042 pAA (plasmid)	EAEC	FN554767	(Chaudhuri et al. 2010)
H10407	ETEC[d]	FN649414	(Crossman et al. 2010)
H10407 p58 (plasmid)	ETEC	FN649416	(Crossman et al. 2010)
H10407 p666 (plasmid)	ETEC	FN649417	(Crossman et al. 2010)
H10407 p948 (plasmid)	ETEC	FN649418	(Crossman et al. 2010)
H10407 p52 (plasmid)	ETEC	FN649415	(Crossman et al. 2010)
LF82	AIEC[e]	CU651637	(Miquel et al. 2010)
0127:H6 E2348/69	EPEC[f]	FM180568	(Iguchi et al. 2009)
0127:H6 E2348/69 pE2348-2(plasmid)	EPEC	FM180570	(Iguchi et al. 2009)
0127:H6 E2348/69 pMAR2 (plasmid)	EPEC	FM180569	(Iguchi et al. 2009)
CFT073	UPEC[g]	AE014075	(Welch et al. 2002)
UTI89	UPEC	NC_007946.1	(Chen et al. 2006)
UTI89 pUTI89 (plasmid)	UPEC	CP000244	(Chen et al. 2006)
S88	NMEC[h]	CU928161	GeneBank
pECOS88 (plasmid)	NMEC	CU928146	(Peigne et al. 2009)

[a]Avian pathogenic *E. coli* (APEC), [b]Enterohemorragic *E. coli* (EHEC), [c]Enteroaggregative *E. coli* (EAEC), [d]Enterotoxigenic *E. coli* (ETEC), [e]Adherent--invasive *E. coli* (AIEC), [f]Enteropathogenic *E. coli* (EPEC), [g]Uropathogenic *E. coli* (UPEC), [h]Neonatal meningitis *E. coli* (NMEC).

Comparative analysis of virulence genes for several *E. coli* pathotypes was also performed with BLAST 2.2.26 software with the application tblastn that was executed with default parameters except for the e-value, which was adjusted to 1. A total of 83 proteins, which are representative of virulence associated genes, were searched against 14 chromosomal genomes and their plasmids sequences when available. The proteins were considered present in a genome when the identity percentage was greater than 90.

Gene Cluster 3.0 software grouped the genomes according to the presence or absence of all tested genes. A binary matrix was constructed to determine the distance, using Pearson correlation coefficient (centered), and clustering was performed by the single linkage. The distance tree was visualized through Java TreeView 1.1.6r2.

RESULTS

The *in silico* phylogenetic analysis conducted for APEC, human-associated ExPEC, and enteric pathogenic *E. coli*

strains defined which alleles of the seven housekeeping genes (*adk*, *fumC*, *gyrB, icd*, *mdh*, *purA* and *recA*) were present in each genome sequence. The seven MLST alleles were concatenated and used to generate a Maximum Likelihood tree (Fig.1).

The presence or absence of 83 well-known virulence genes/traits (Kaper et al. 2004, Johnson et al. 2008b), resulted on a virulence gene profile for each strain (Fig.2). The sequence types (ST) identified in the APEC strains recently sequenced by our research group were: ST117 (strain SEPT362), ST363 (strain SCI-07), ST10 (strain S17) and ST453 (strain O08). To our knowledge, ST453 (strain O08 – from an omphalitis case) was not previously detected in other published APEC strains. In MLST database, ST453 was associated with human pathogenic *E. coli* strains (*E. coli* 28, U2183, HMMC097 and E25). ST363 (SCI-07 – isolated from a swollen head syndrome case) was detected before in two *E. coli* strains, a nonpathogenic (KK9/10) isolated from a human, and another pathogenic (IMT5119) isolated from a chicken. ST10 (strain S17 – isolated from a septicemia case) was detected in a large range of sources, such as human pathogenic and non-pathogenic strains, water, soil and animal-isolated strains (eg. camel, pig, bovine and wild rabbit). This ST was also identified in *E. coli* isolated from birds, such as chicken, wild pigeon and parrot. All strains isolated from wild pigeon and parrot were pathogenic, while there were both pathogenic and nonpathogenic strains isolated from chickens (MLST database). The ST117 (SEPT362 - septicemic) is frequently found in APEC strains, in fact, this ST was the most prevalent type detected in two different studies that determined ST for *E. coli* isolates from several Danish broiler breeders and layer farms (Gregersen et al. 2010, Olsen et al. 2011) and in the MLST database. This ST is also associated to human urinary tract infections (UTI) (Wu et al. 2012).

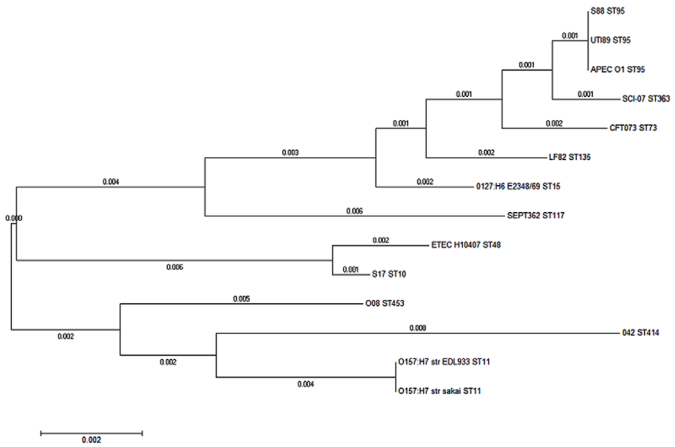

Fig.1. Phylogenetic Maximum Likelihood tree of the concatenated sequences of seven housekeeping genes. The sequence type (ST) is described beside the strain name.

In the MLST tree (Fig.1), strain O08 (ST453), together with 042 (enteroaggregative *E. coli* - EAEC), O157:H7 str. EDL933 and O157:H7 str. Sakai (both enterohemorragic *E. coli* – EHEC), formed a clade that is a sister group to the clade composed by all the other strains. On the other hand, strain APEC SCI-07 (ST363) presented a closer proximity to ST95 human ExPEC strains and APEC O1 (ST95). According to the phylogenetic tree, APEC S17 (ST10) presented a closer similarity to the H10407 ETEC strain (ST48). Strain SEPT362 (ST117) is represented in the phylogenetic tree as an isolated branch, which is a sister clade of that one that includes all human ExPEC strains, AIEC (adherent-invasive *E. coli*), EPEC (enteropathogenic *E. coli*) and APEC strains SCI-07 and APEC O1. These results demonstrate that strains isolated from different infectious processes, with a few exceptions, do not present 100% of identity. However, they could share genomic similarities in the seven housekeeping allelic compositions.

The cluster obtained by the virulence gene profiles (Fig.2) presented slight differences when compared to the MLST phylogenetic tree. Cluster was able to group the strains into two main classes: one composed only of intimin producing enteric strains (O157:H7 str. EDL933; O157:H7 str. Sakai and 0127:H6 E2348/69) and other composed of all APEC (APEC O1, SCI-07, SEPT362, S17 and O08), human ExPEC (UTI89, S88 and CFT073) and some enteric strains (ETEC H10407, EAEC 042 and AIEC LF82). In this second group, an ExPEC cluster formed by UTI89, S88, APEC O1 and CFT073 was well marked. A clearly defined similarity between strains S17 and ETEC H10407 and between SEPT362, SCI-07 and O08 was noticed, indicating a closer genetic identity, considering the genes herein analyzed.

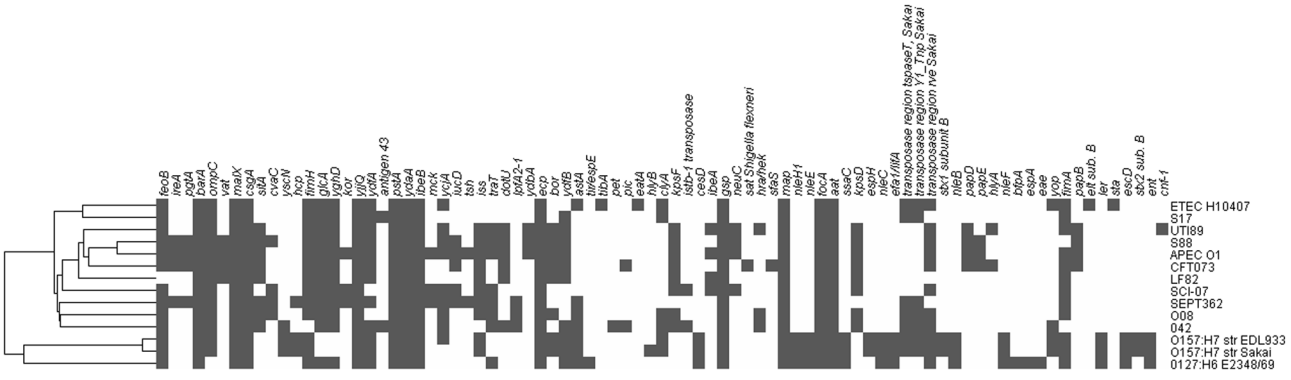

Fig.2. Gene cluster result for 14 pathogenic *Escherichia coli* strains. The presence (Gray Square) and the absence (White Square) of 83 virulence genes/traits are represented in the image. *E. coli* strains names are listed on the right. Gene names are listed above the image.

DISCUSSION

Several studies aimed to obtain virulence gene profiles in order to define the APEC pathotype. Most of these studies used multiplex PCR and focused on virulence genes localized mainly on the colicin V plasmid (Skyberg et al. 2003, Ewers et al. 2005, Johnson et al. 2008a, Schouler et al. 2012). However, the high diversity encountered in *E. coli* species difficult the use of a single set of virulence factors for diagnosis (Schouler et al. 2012). For Brazilian APEC strains, a recent study of our group, showed that the analysis of several biological characteristics, such as adhesion to eukaryotic cell cultures, pathogenicity levels according to the lethal dose (50%) assay, phylogenetic groups and virulence gene profiles, indicated that APEC strains could be organized into a structured set of subgroups associated with specific infectious syndromes (Maturana et al. 2011). These finds demonstrated that APEC strains are very diverse and that the detection of specific genes is not sufficient to define this group. Thus, a study involving MLST (Fig.1) and analysis of a large set of virulence genes, including those found in human ExPEC and enteric strains (Fig.2) were used to group the strains herein studied.

The phylogenetic tree, based on the seven housekeeping genes analyzed (Fig.1), demonstrated that, although the different *E. coli* pathotypes share some genomic similarities, they are very diverse. In addition, these results show that some strains from different origins have either identical or similar housekeeping genes background, notwithstanding their hosts. The virulence gene profiles (Fig.2), presented a higher discriminatory power to differentiate the different *E. coli* pathotypes than the phylogenetic tree, suggesting that gene content appears to be more important in identifying a determined pathotype than the evolutionary history. The phylogenetic analysis, also demonstrated, as indicated by the MLST study, that APEC strains are not a homogeneous group. Thus, the phylogenetic tree and virulence gene profiles revealed that the strains are very diverse, corroborating previous works (Maturana et al. 2011, Schouler et al. 2012). For strain S17, a septicaemic Brazilian APEC, both methodologies clearly showed that it has a closer genetic similarity with strain H10407, a human ETEC strain. In addition, Brazilian strains SEPT362 and SCI-07 shares genetic similarities with human UPEC strains and a previously published APEC strain, APEC O1.

The virulence gene content analysis (Fig.2) displayed a clear discrimination between enteric and ExPEC strains. For ExPEC strains, several virulence factors were widespread (*ireA*, *cvaC*, *iucD*, *iss*, *traT*, *dotU* and *neuC*), and three genes could be found only in APEC strains (*kor*, *mck* and *tsh*). Human ExPEC strains and APEC O1 were clustered together considering both analyses, but these results were expected since it is known that APEC O1 is indeed similar to UPEC strains (Johnson et al. 2007). Moreover, genes present only in enteric strains (*yscN*, *cesD*, *nleH*1, *nleE*, *ssaC*, *nleC*, *efa*1*/lifA*, *nleB*, *nleF*, *yop*, *ler*, *escD* and *ent*) may have contributed to the formation of this ExPEC cluster.

The results demonstrated that APEC strains, with a closer genetic identity to human ExPEC strains, would represent potential zoonotic risks. Moreover, these data suggest the importance of core genomes and virulence genes analyses. These virulence genes are frequently present in plasmids, in Pathogenic islands (PAIs) and other mobile genetic elements. They are freely spread among bacterial genomes, such a way that, specific genes can be preserved or not on strains, depending on the challenges they face. These challenges include colonizing their different hosts, which may also share similar receptors for the bacterial genes products. In this way, evolution would work on two different directions: (1) favoring host-specific strains that could be eradicated in case of a host change and (2) selecting host unspecific strains, which would constitute reservoirs for important virulence genes that would be preserved in the environment. These last strains would have the potential to represent zoonotic risks.

Acknowledgements.- To FAPESP [Proc.2010/51421-8 and 2012/04931-6] and CAPES [Proc.23038.042588/2008-11]. CNPq provided a fellowship to the main author and to the corresponding author, and FAPESP provided a fellowship to R.P.M. [Proc.2012/05073-3].

REFERENCES

Barnes H., Nolan L. & Vaillancourt J. 2008. Colibacillosis, p.691-732. In: Saif Y. (Ed.), Diseases of Poultry. 12th ed. Blackwell Publishing, Ames. 1409p.

Bergholz T.M., Wick L.M., Qi W., Riordan J.T., Ouellette L.M. & Whittam T.S. 2007. Global transcriptional response of *Escherichia coli* O157:H7 to growth transitions in glucose minimal medium. BMC Microbiol. 7:97.

Chaudhuri R.R., Sebaihia M., Hobman J.L., Webber M.A., Leyton D.L., Goldberg M.D., Cunningham A.F., Scott-Tucker A., Ferguson P.R., Thomas C.M., Frankel G., Tang C.M., Dudley E.G., Roberts I.S., Rasko D.A., Pallen M.J., Parkhill J., Nataro J.P., Thomson N.R. & Henderson I.R. 2010. Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative *Escherichia coli* strain 042. PLoS One 5:e8801.

Chen S.L., Hung C.S., Xu J., Reigstad C.S., Magrini V., Sabo A., Blasiar D., Bieri T., Meyer R.R., Ozersky P., Armstrong J.R., Fulton R.S., Latreille J.P., Spieth J., Hooton T.M., Mardis E.R., Hultgren S.J. & Gordon J.I. 2006. Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. Proc. Natl Acad. Sci. USA 103:5977-5982.

Crossman L.C., Chaudhuri R.R., Beatson S.A., Wells T.J., Desvaux M., Cunningham A.F., Petty N.K., Mahon V., Brinkley C., Hobman J.L., Savarino S.J., Turner S.M., Pallen M.J., Penn C.W., Parkhill J., Turner A.K., Johnson T.J., Thomson N.R., Smith S.G. & Henderson I.R. 2010. A commensal gone bad: complete genome sequence of the prototypical enterotoxigenic *Escherichia coli* strain H10407. J. Bacteriol. 192:5822-5831.

Dho-Moulin M. & Fairbrother J.M. 1999. Avian pathogenic *Escherichia coli* (APEC). Vet. Res. 30:299-316.

Dziva F., Hauser H., Connor T.R., Van Diemen P.M., Prescott G., Langridge G.C., Eckert S., Chaudhuri R.R., Ewers C., Mellata M., Mukhopadhyay S., Curtiss R., Dougan G., Wieler L.H., Thomson N.R., Pickard D.J. & Stevens M.P. 2013. Sequencing and functional annotation of avian pathogenic *Escherichia coli* serogroup O78 strains reveal the evolution of *E. coli* lineages pathogenic for poultry via distinct mechanisms. Infect. Immun. 81:838-849.

Dziva F. & Stevens M. 2008. Colibacillosis in poultry: unravelling the molecular basis of virulence of avian pathogenic *Escherichia coli* in their natural hosts. Avian Pathol. 37:355-366.

Edgar R.C. 2004. Muscle: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113.

Ewers C., Janssen T., Kiessling S., Philipp H.C. & Wieler L.H. 2005. Rapid detection of virulence-associated genes in avian pathogenic *Escherichia coli* by multiplex polymerase chain reaction. Avian Dis. 49:269-273.

Ewers C., Li G., Wilking H., Kieling S., Alt K., Antão E.M., Laturnus C., Diehl

I., Glodde S., Homeier T., Böhnke U., Steinrück H., Philipp H.C. & Wieler L.H. 2007. Avian pathogenic, uropathogenic, and newborn meningitis--causing *Escherichia coli*: How closely related are they? Int. J. Med. Microbiol. 297:163-176.

Gregersen R.H., Christensen H., Ewers C. & Bisgaard M. 2010. Impact of *Escherichia coli* vaccine on parent stock mortality, first week mortality of broilers and population diversity of *E. coli* in vaccinated flocks. Avian Pathol. 39:287-295.

Iguchi A., Thomson N.R., Ogura Y., Saunders D., Ooka T., Henderson I.R., Harris D., Asadulghani M., Kurokawa K., Dean P., Kenny B., Quail M.A., Thurston S., Dougan G., Hayashi T., Parkhill J. & Frankel G. 2009. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. J. Bacteriol. 191:347-354.

Johnson T., Johnson S.J. & Nolan L.K. 2006a. Complete DNA sequence of a ColBM plasmid from avian pathogenic *Escherichia coli* suggests that it evolved from closely related ColV virulence plasmids. J. Bacteriol. 188:5975-5983.

Johnson T.J., Wannemeuhler Y.M., Scaccianoce J.A., Johnson S.J. & Nolan L.K. 2006b. Complete DNA sequence, comparative genomics, and prevalence of an IncHI2 plasmid occurring among extraintestinal pathogenic *Escherichia coli* isolates. Antimicrob. Agents Chemother. 50:3929-3933.

Johnson T.J., Kariyawasam S., Wannemuehler Y., Mangiamele P., Johnson S.J., Doetkott C., Skyberg J.A., Lynne A.M., Johnson J.R. & Nolan L.K. 2007. The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. J. Bacteriol. 189:3228-3236.

Johnson T.J., Wannemuehler Y., Doetkott C., Johnson S.J., Rosenberger S.C. & Nolan L.K. 2008a. Identification of minimal predictors of avian pathogenic *Escherichia coli* virulence for use as a rapid diagnostic tool. J. Clin. Microbiol. 46:3987-3996.

Johnson T.J., Wannemuehler Y., Johnson S.J., Stell A.L., Doetkott C., Johnson J.R., Kim K.S., Spanjaard L. & Nolan L.K. 2008b. Comparison of extraintestinal pathogenic *Escherichia coli* strains from human and avian sources reveals a mixed subset representing potential zoonotic pathogens. Appl. Environ. Microbiol. 74:7043-7050.

Kaper J.B., Nataro J.P. & Mobley H.L.T. 2004. Pathogenic *Escherichia coli*. Natl Rev. Microbiol. 2:123-140.

Makino K., Ishii K., Yasunaga T., Hattori M., Yokoyama K., Yutsudo C.H., Kubota Y., Yamaichi Y., Iida T., Yamamoto K., Honda T., Han C.G., Ohtsubo E., Kasamatsu M., Hayashi T., Kuhara S. & Shinagawa H. 1998. Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic *Escherichia coli* O157:H7 derived from Sakai outbreak. DNA Res. 5:1-9.

Maturana V.G., Pace F., Carlos C., Mistretta Pires M., Campos T.A., Nakazato G., Guedes Stheling E., Logue C.M., Nolan L.K. & Dias da Silveira W. 2011. Subpathotypes of avian pathogenic *Escherichia coli* (APEC) exist as defined by their syndromes and virulence traits. Open Microbiol. J. 5:55-64.

Miquel S., Peyretaillade E., Claret L., Vallée A., Dossat C., Vacherie B., Zineb E.H., Segurens B., Barbe V., Sauvanet P., Neut C., Colombel J.F., Medigue C., Mojica F.J., Peyret P., Bonnet R. & Darfeuille-Michaud A. 2010. Complete genome sequence of Crohn's disease-associated adherent-invasive *E. coli* strain LF82. PLoS One 5:e12714.

Moulin-Schouleur M., Répérant M., Laurent S., Brée A., Mignon-Grasteau S., Germon P., Rasschaert D. & Schouler C. 2007. Extraintestinal pathogenic *Escherichia coli* strains of avian and human origin: link between phylogenetic relationships and common virulence patterns. J. Clin. Microbiol. 45:3366-3376.

Olsen R.H., Stockholm N.M., Permin A., Christensen J.P., Christensen H. & Bisgaard M. 2011. Multi-locus sequence typing and plasmid profile characterization of avian pathogenic *Escherichia coli* associated with increased mortality in free-range layer flocks. Avian Pathol. 40:437-444.

Peigne C., Bidet P., Mahjoub-Messai F., Plainvert C., Barbe V., Médigue C., Frapy E., Nassif X., Denamur E., Bingen E. & Bonacorsi S. 2009. The plasmid of *Escherichia coli* strain S88 (O45:K1:H7) that causes neonatal meningitis is closely related to avian pathogenic *E. coli* plasmids and is associated with high-level bacteremia in a neonatal rat meningitis model. Infect. Immun. 77:2272-2284.

Perna N.T., Plunkett G., Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J. & Kirkpatrick H.A. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature 409:529-533.

Rodriguez-Siek K.E., Giddings C.W., Doetkott C., Johnson T.J., Fakhr M.K. & Nolan L.K. 2005. Comparison of *Escherichia coli* isolates implicated in human urinary tract infection and avian colibacillosis. Microbiol. 151:2097-2110.

Rojas T.C.G., Maluta R.P., Parizzi L.P., Koenigkan L.V., Yang J., Yu J., Pereira G.A. & Dias da Silveira W. 2013. Genome sequences of avian pathogenic *Escherichia coli* strains isolated from Brazilian commercial poultry. Genome Announc. 1:e0011013.

Rojas T.C.G., Parizzi L.P., Tiba M.R., Chen L., Pereira G.A., Sangal V., Yang J., Yu J. & Dias da Silveira W. 2012. Draft genome of a Brazilian avian-pathogenic *Escherichia coli* strain and in silico characterization of virulence--related genes. J. Bacteriol. 194:3023.

Schouler C., Schaeffer B., Brée A., Mora A., Dahbi G., Biet F., Oswald E., Mainil J., Blanco J. & Moulin-Schouleur M. 2012. Diagnostic strategy for identifying avian pathogenic *Escherichia coli* based on four patterns of virulence genes. J. Clin. Microbiol. 50:1673-1678.

Skyberg J.A., Horne S.M., Giddings C.W., Wooley R.E., Gibbs P.S. & Nolan L.K. 2003. Characterizing avian *Escherichia coli* isolates with multiplex polymerase chain reaction. Avian Dis. 47:1441-1447.

Welch R.A., Burland V., Plunkett G., Redford P., Roesch P., Rasko D., Buckles E.L., Liou S.R., Boutin A., Hackett J., Stroud D., Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T., Mobley H.L., Donnenberg M.S. & Blattner F.R. 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. Proc. Natl. Acad. Sci. USA 99:17020-17024.

Wu G., Ehricht R., Mafura M., Stokes M., Smith N., Pritchard G.C. & Woodward M.J. 2012. *Escherichia coli* isolates from extraintestinal organs of livestock animals harbour diverse virulence genes and belong to multiple genetic lineages. Vet. Microbiol. 160:197-206.