

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROBIOLOGIA
AGRÍCOLA E DO AMBIENTE

**EPIDEMIOLOGIA MOLECULAR DE ISOLADOS DE *Escherichia coli*
PROVENIENTES DE DIFERENTES HOSPEDEIROS E O IMPACTO DOS
ANTIMICROBIANOS EM DIFERENTES CLONES PRODUTORES DE MCR-1.**

RAFAELA RAMALHO GUERRA

Orientador(a): Prof.^(a) Dr.^(a) Andreza Francisco Martins.

Porto Alegre
Novembro/2020

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM MICROBIOLOGIA
AGRÍCOLA E DO AMBIENTE

Rafaela Ramalho Guerra
Biotecnologista

Dissertação apresentada ao Programa de Pós-Graduação em Microbiologia Agrícola e do Ambiente do Instituto de Ciências Básicas da Saúde da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do título de mestre(a) em Microbiologia Agrícola e do Ambiente.

Orientador(a): Prof.^(a) Dr.^(a) Andreza Francisco Martins

CIP - Catalogação na Publicação

Guerra, Rafaela Ramalho
EPIDEMIOLOGIA MOLECULAR DE ISOLADOS DE *Escherichia coli* PROVENIENTES DE DIFERENTES HOSPEDEIROS E O IMPACTO DOS ANTIMICROBIANOS EM DIFERENTES CLONES PRODUTORES DE MCR-1. / Rafaela Ramalho Guerra. -- 2020.

119 f.

Orientador: Andreza Francisco Martins.

Dissertação (Mestrado) -- Universidade Federal do Rio Grande do Sul, Instituto de Ciências Básicas da Saúde, Programa de Pós-Graduação em Microbiologia Agrícola e do Ambiente, Porto Alegre, BR-RS, 2020.

1. Epidemiologia molecular. 2. Genotipagem de *E.coli*. 3. Resistência bacteriana. 4. Gene *mcr-1*. I. Martins, Andreza Francisco, orient. II. Título.

“Porque dele, e por meio dele, e para Ele são todas as coisas. A Ele seja a glória para sempre. Amém!”

Romanos 11:36

AGRADECIMENTOS

Primeiramente, gostaria de agradecer a Deus pela oportunidade de concluir este trabalho.

Agradeço também a minha orientadora Prof. Andreza Martins, que de forma incansável busca os melhores recursos para que a pesquisa seja desenvolvida e atinja o maior número de pessoas. Obrigada pelas palavras de apoio e exortação, que tanto contribuíram para o meu crescimento pessoal e profissional.

Obrigada aos meus colegas de laboratório Silvia Lentz, William Machado, Gabriela Simões, Otávio Lovison, Camila Muller e Thaianne Marques. A todos os alunos do Laboratório de Microbiologia Molecular e Saúde, meu muito obrigada! Agradeço também aos colegas que já fizeram parte desta equipe e que me apoiaram mesmo nas horas mais difíceis. Obrigada Roberta Taufer Boff, Gabrielle Saldanha, Franciele Adam, Daniela Signori e Ana Paula Rauber.

Agradeço a minha família que sempre esteve ao meu lado e foram essenciais para que eu continuasse na caminhada da academia. Em especial, minha Mãe, que sempre acreditou no meu potencial e ao meu irmão Eduardo, que me ajudou em todos os momentos de dificuldades.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa e ao PPGMAA por oportunizar o desenvolvimento dessa pesquisa.

**EPIDEMIOLOGIA MOLECULAR DE ISOLADOS DE *Escherichia coli*
PROVENIENTES DE DIFERENTES HOSPEDEIROS E O IMPACTO DOS
ANTIMICROBIANOS EM DIFERENTES CLONES PRODUTORES DE MCR-1.¹**

Autor: Rafaela Ramalho Guerra

Orientador(a): Prof^(a). Dr^(a). Andreza Francisco Martins

RESUMO

O uso de antimicrobianos na produção animal contribui com a pressão seletiva e colabora na expansão de cepas multirresistentes em diferentes ambientes. Já é relatado que os animais de produção atuam como potenciais reservatórios de microrganismos produtores de genes de resistência (ARGs). Dentro deste contexto, a resistência as polimixinas mediada por plasmídeos já foi bem reportada no mundo todo e a habilidade de disseminação do gene *mcr-1* precisa ser melhor compreendida. Assim, o objetivo deste trabalho foi avaliar a epidemiologia molecular de isolados de *Escherichia coli mcr-1* positivos e *mcr-1* negativos provenientes de diferentes hospedeiros através de três metodologias: MLST, CH Typing e SNPs (*Single Nucleotide Polymorphism*). Além disso, buscamos relacionar como o uso de antimicrobianos pode influenciar na expansão clonal de isolados produtores de *mcr-1* nos diferentes nichos ecológicos. Isolados de *E. coli mcr-1* positivos (n=6) recuperados de frangos de corte e suínos provenientes de granjas no Rio Grande do Sul tiveram seus genomas comparados com isolados recuperados de bancos de dados públicos (n = 37), sendo 19 produtores de *mcr-1* e 18 negativos para o gene, totalizando uma população de 43 isolados. A tipagem MLST apresentou 35 STs , sendo a ST-10 e a ST-101 as de maior frequência com 11,63% e 6,98%, respectivamente. Já a tipagem molecular por CH Typing, apresentou 32 CHs, sendo os tipos mais frequentes as CH 11-25 com 9,30% de frequência, CH 11-0 com 6,98% e CH 29-38 com 6,98%. A análise filogenética utilizando as três metodologias apresentou uma distribuição diversa e concordante ao longo das árvores, onde foi possível observar isolados produtores de *mcr-1* intimamente relacionados com isolados negativos para o gene. Além disso, isolados com uma diferença de tempo de 60 anos foram intimamente relacionados, podendo indicar uma disseminação não clonal entre as cepas, sugerindo que a pressão seletiva e eventos recentes de recombinação estão agindo sobre a população. Com base nos dados apresentados acima, e reforçando a ideia de que a permanência de características em uma população depende das forças evolutivas que atuam sobre ela, afirmamos que medidas integrativas apoiadas na abordagem *One Health* devem ser adotadas como estratégia para reduzir a disseminação de cepas multirresistentes.

Palavras-chave: *Escherichia coli*; *mcr-1*; epidemiologia molecular; expansão clonal; animais de produção.

¹Dissertação de Mestrado em Microbiologia Agrícola e do Ambiente – Instituto de Ciências Básicas da Saúde, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil. (xx p.) Novembro, 2020.

**MOLECULAR EPIDEMIOLOGY OF *Escherichia coli* ISOLATES FROM
DIFFERENT HOSTS AND THE IMPACT OF ANTIMICROBIAL ON DIFFERENT
CLONES PRODUCING MCR-1.¹**

Author: Rafaela Ramalho Guerra

Advisor: Prof^(a). Dr^(a). Andreza Francisco Martins

ABSTRACT

The use of antimicrobials in animal production can act as selective pressure and corroborate with the expansion of multidrug-resistant strains in different environments. It has already been reported that production animals are potential reservoirs of ARGs (antimicrobial resistance genes) and in spreading these genes by the food chain. In this context, plasmid-mediated colistin resistance has already been reported worldwide and the dispersal ability of the *mcr-1* gene must be better understood. So, the objective of this study was to determine the molecular epidemiology of *mcr-1* positive and negative *Escherichia coli* isolates from different hosts through three different methodologies: MLST, CH Typing, and SNPs (Single Nucleotide Polymorphism). Furthermore, it was observed how the use of antimicrobials could impact the clonal expansion of *mcr-1* producing isolates in different ecological niches. *mcr-1* positive *E. coli* isolates (n=6) recovered from broilers and pigs from farms in the Rio Grande do Sul had their genomes compared with isolates recovered from public databases (n = 37), being 19 *mcr-1* positive and 18 *mcr-1* negative totaling a population of 43 isolates. The MLST typing presented 35 single STs, being ST-10 and ST-101 the most frequent, with 11.63% and 6.98%, respectively. The molecular typing by CH Typing presented 32 unique CH, being CH 11-25 9.30%, CH 11-0; 6.98%, and CH 29-38; 6.98% the most frequent types among the 43 isolates. The phylogenetic analysis using the three methodologies showed a heterogeneous and concordant distribution along with the trees, where it was possible to observe *mcr-1* producing isolates closely related to negative isolates for the gene. Furthermore, the SNP tree demonstrated that the central genome of these isolates is well conserved, as isolates with a time difference of 60 years were closely related, which may indicate a non-clonal dispersion among the strains, suggesting that selective pressure and recent recombination events are acting on the population. Based on the data presented above, and reinforcing the idea that the permanence of characteristics in a population depends on the evolutionary forces that act on it, we highlight that integrative measures supported by One Health approaches should be adopted as a strategy to minimize the antimicrobial resistance spread.

Key words: *Escherichia coli*; *mcr-1*, molecular epidemiology; antimicrobials; clonal expansion; production animals.

¹Master of Science Thesis in Agricultural and Environmental Microbiology – Instituto de Ciências Básicas da Saúde, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil. (xx p.) November, 2020.

SUMÁRIO

1.	INTRODUÇÃO.....	12
2.	OBJETIVOS.....	15
2.1	OBJETIVO GERAL.....	15
2.2	OBJETIVOS ESPECÍFICOS.....	15
3.	REVISÃO DA LITERATURA.....	17
3.1	EPIDEMIOLOGIA MOLECULAR.....	17
3.2	TIPAGEM MOLECULAR.....	18
3.2.1	Eletroforese em gel em campo pulsado PFGE.....	18
3.2.2	Tipagem de Sequência Multilocus MLST.....	19
3.2.3	CH Typing.....	20
3.2.4	Tipagem de Sequência Multilocus do genoma completo e do genoma central wgMLST/cgMLST.....	21
3.2.5	Tipagem de Polimorfismo de Nucleotídeo Único (SNP).....	22
3.3	ANÁLISE FILOGENÉTICA.....	24
3.4	CARACTERÍSTICAS DE <i>E. COLI</i>	26
3.4.1	Plasticidade do Genoma e Variação Fenotípica.....	27
3.4.2	Transferência Horizontal Gênica e suas Consequências.....	28
3.5	RESISTÊNCIA MICROBIANA.....	29
3.5.1	Resistência às Polimixinas.....	31
3.5.2	Animais de Produção como Reservatórios de Genes de Resistência	33
4.	MATERIAIS E MÉTODOS.....	38
4.1	AMOSTRAGEM DO ESTUDO.....	38
4.1.1	Características dos Isolados.....	38
4.1.2	Sequenciamento Completo do Genoma.....	38
4.1.3	Isolados selecionados a partir de Banco de Dados Públicos.....	39
4.2	ANÁLISES DE BIOINFORMÁTICA.....	39
4.3	TIPAGEM MOLECULAR.....	41
4.4	ANÁLISES FILOGENÉTICAS.....	42
5.	RESULTADOS.....	44
6.	DISCUSSÃO.....	63

7.	CONCLUSÃO.....	74
8.	REFERÊNCIAS.....	76
9.	ARTIGO.....	88
10.	APÊNDICE.....	117

LISTA DE TABELAS

	Página
Tabela1- Diferença entre as metodologias baseadas em MLST.....	22
Tabela 2- Caracterização molecular das cepas de <i>Escherichia coli</i> isoladas de animais de produção, no Rio Grande do Sul.....	44
Tabela3- Características gerais de todos os isolados utilizados no estudo...	48
Tabela 1A- Porcentagem do genoma mapeado e número de variantes de nucleotídeo único SNV.....	117

LISTA DE FIGURAS

	Página
Figura 1-Definição <i>One Health</i> de disseminação da resistência microbiana.	31
Figura 2-Vias de disseminação de <i>E. coli</i> carreando o gene <i>mcr-1</i>	36
Figura 3-Fluxograma da metodologia utilizada no estudo.....	41
Figura 4-Árvore filogenética das sequências tipo STs.....	57
Figura 5-Árvore de abrangência mínima das sequências tipo STs.....	58
Figura 6-Árvore filogenética dos clonotipos CH.....	60
Figura 7-Árvore de abrangência mínima dos clonotipos CH.....	61
Figura 8-Árvore filogenética de SNPs.....	62
Figura 1A-Matriz utilizada na inferência da árvore filogenética de SNPs.....	119

LISTA DE ABREVIATURAS E SIGLAS

MLST	<i>Multilocus Sequence Typing</i>
UFDA	<i>Food and Drug Administration</i>
ST	<i>Sequence type</i>
CH	<i>CH Typing</i>
SNV	<i>Single Nucleotide Variant</i>
SNP	<i>Single Nucleotide Polymorphism</i>
WGS	<i>Whole Genome Sequencing</i>
PFGE	<i>Pulsed Field Gel Electrophoresis</i>
NGS	<i>Next Generation Sequencing</i>
GC	<i>Guanina/Citosina</i>
ESBL	β -lactamase de espectro estendido (<i>Extended-Spectrum β-Lactamase</i>)
ARGs	Genes de resistência a antibióticos (<i>Antibiotic Resistance Genes</i>)
MDR	Resistência a múltiplos medicamentos (<i>Multiple Drug Resistance</i>)
HGT	Transferência horizontal gênica (<i>Horizontal Gene Transfer</i>)
KPC	<i>Klebsiella Pneumoniae Carbapenemase</i>
NDM	<i>New Delhi metallo-β-lactamase</i>
SAM	<i>Sequence Alignment Map</i>
VCF	<i>Variant Calling Format</i>
AMR	Resistência antimicrobiana (<i>Antimicrobial resistance</i>)

1. INTRODUÇÃO

O tratamento de doenças infecciosas com antibióticos trouxe muitos benefícios clínicos desde a sua descoberta. Porém, o uso extensivo de antimicrobianos desencadeou o que pode ser uma das maiores ameaças à saúde humana, animal e ambiental no século XXI: a resistência microbiana (UDAONDO; HUERTAS, 2020).

Sabemos que os compostos antibióticos são essenciais e muito utilizados na terapia clínica humana, na profilaxia, e também como promotor de crescimento, na produção animal (SHEN *et al.*, 2020b). A exposição a quantidades substanciais desses compostos em animais de produção tornou os produtos pecuários potenciais reservatórios de genes de resistência a antimicrobianos (ARGs), aumentando assim as chances de disseminação entre diferentes nichos ecológicos através da transferência horizontal gênica (HGT). Quando falamos de ARGs é importante destacar a disseminação de genes que conferem resistência a antimicrobianos utilizados como último recurso no tratamento de infecções clínicas, como as polimixinas (colistina), utilizadas no tratamento de doenças causadas por microrganismos multirresistentes (MDR) (SUN *et al.*, 2018).

Desde 2016, com a descoberta do gene *mcr-1*, capaz de conferir resistência à colistina, sendo mediada por plasmídeo, muitos estudos têm relatado variantes do gene com perfil semelhante de resistência, sendo a variante *mcr-10* a mais recente descoberta, em 2020 (LIU *et al.*, 2016a; WANG *et al.*, 2020). O gene *mcr-1* já foi relatado em vários tipos de plasmídeos, como IncX4, IncI2 e IncHI2, expressando outras variantes do gene no mesmo plasmídeo e também conferindo uma resistência adicional a isolados que já expressavam genes relacionados a produção de betalactamases de espectro estendido ESBL, como *bla*_{CTXM-8}, *bla*_{CMY-2}, além de carbapenemases do tipo *bla*_{NDM} e *bla*_{KPC}, entre outras classes de antimicrobianos, como quinolonas e sulfonamidas (CASELLI *et al.*, 2018; IGWARAN; IWERIEBOR; OKOH, 2018; MARTINS-SORENSEN *et al.*, 2020; NANG; LI; VELKOV, 2019; REBELO *et al.*, 2018; WANG *et al.*, 2018a; YANG *et al.*, 2018a).

Esses dados são extremamente preocupante, uma vez que a disseminação de clones MDR são comumente disseminados através da cadeia alimentar, pelo contato direto com animais ou através do meio ambiente.

O uso de antibióticos como agente promotor do crescimento em animais de produção contribui para a pressão seletiva e, conseqüentemente, o surgimento de infecções bacterianas MDR na comunidade. O destino de todos os antibióticos administrados que não são absorvidos pelo animal são despejados no meio ambiente na forma de fezes e urina. Essa situação, juntamente com o tratamento insuficiente de resíduos industriais, residenciais e agrícolas, também contribui para a expansão da resistência microbiana ambiental (UDAONDO; HUERTAS, 2020).

Uma das espécies produtoras de *mcr-1* mais relatadas na literatura é a *Escherichia coli*, um microrganismo ubíquo da microbiota intestinal normal em humanos e animais, mas também encontrada como diversos patótipos associados a infecções graves (KAPER; NATARO; MOBLEY, 2004b). Apoiando a abordagem *One Health*, que reconhece que os humanos, os animais e a saúde ambiental estão intimamente ligados, consideramos esse microrganismo como um potencial marcador de resistência na comunidade. Sabendo que o conjunto de genes associados à resistência antimicrobiana em um ambiente (resistoma) não está relacionado apenas com o influxo de ARGs entre esses três setores, mas também na transferência horizontal desses genes entre bactérias comensais e cepas patogênicas (TÜMMLER, 2020), destacamos a importância do estudo epidemiológico de isolados MDR em diferentes tipos de hospedeiros. A epidemiologia molecular é uma ferramenta poderosa na prática de estudos deste tipo, pois a observação de variantes genômicas, como as variantes de nucleotídeo único SNVs (*Single Nucleotide Variants*) e suas relações na transmissão, evolução e expansão clonal de cepas emergentes conferem uma melhor compreensão da propagação da resistência (SABAT *et al.*, 2013; TÜMMLER, 2020).

Tendo em vista que os animais de produção podem ser reservatórios de ARGs com potencial disseminação entre humanos e ambiente, a genotipagem de cepas de *Escherichia coli* e o estudo genômico comparativo, através de análises filogenéticas, proporcionam um melhor entendimento da disseminação desses

isolados na população, possibilitando a criação de políticas públicas, medidas integrativas de inspeção, vigilância e implementação de abordagens *One Health*, com o intuito de minimizar a transmissão de cepas MDR e prever possíveis surtos e desequilíbrio econômico.

2. OBJETIVOS

2.1 OBJETIVO GERAL

Comparar genotipicamente cepas de *Escherichia coli*, produtoras de *mcr-1*, isoladas de frangos de corte e suínos, provenientes do Rio Grande do Sul com os dados de isolados disponíveis em bancos de dados públicos, a fim de identificar a diversidade genética e a disseminação das cepas em diferentes tipos de hospedeiros.

2.2 OBJETIVOS ESPECÍFICOS

2.2.1 Determinar o tipo de sequência (ST) e o complexo clonal (CC) dos isolados de *Escherichia coli mcr-1* positivos, provenientes de frangos de corte e suínos (deste estudo), utilizando ferramentas *online* de tipagem molecular, como MLST e CHTyper.

2.2.2 Caracterizar os isolados sequenciados neste estudo quanto a presença de genes de resistência, genes codificadores de fatores de virulência, sorotipo e tipo de plasmídeo.

2.2.3 Comparar filogeneticamente, os isolados sequenciados nesse estudo e os obtidos do banco de dados público, através da inferência de árvores filogenéticas a partir dos resultados do MLST.

2.2.4 Comparar filogeneticamente, os isolados sequenciados nesse estudo e os obtidos do banco de dados público, através da inferência de árvores filogenéticas a partir dos resultados do CHTyper.

2.2.5 Comparar filogeneticamente, os isolados sequenciados nesse estudo e os obtidos do banco de dados público, através da inferência de árvores filogenéticas baseada na análise de variantes de nucleotídeos únicos.

2.2.6 Construir rede de abrangência mínima a partir dos 7 genes *housekeeping*, bem como dos genes *fumC* e *fimH* para análise de microevolução em um modelo reticulado.

2.2.7 Comparar os resultados obtidos através dos 3 métodos de tipagem utilizados: *Multilocus Sequence Typing* MLST, CHTyping e chamada de variantes SNVs no genoma total.

3. REVISÃO DA LITERATURA

3.1 EPIDEMIOLOGIA MOLECULAR

Epidemiologia é o estudo da distribuição e prevalência de uma determinada doença, assim como os fatores de risco associados a ela. Essa disciplina apresenta um importante papel na manutenção da saúde pública e no controle de doenças, através do esclarecimento das interações entre patógeno-hospedeiro e a transmissibilidade da doença em uma determinada população (WANG; KING JORDAN; MAYER, 2014). Quando o estudo da etiologia da doença envolve fatores genéticos e ambientais, e como esses fatores se relacionam, esse estudo passa a ser chamado de epidemiologia molecular (KENRAD NELSON; MASTERS WILLIAMS, 2007).

A epidemiologia molecular se baseia na observação de marcadores moleculares ou genéticos, que se relacionam ao desenvolvimento, transmissão e a evolução de patógenos microbianos causadores de doenças em um grupo (RAVEN et al., 2019). A metodologia utilizada para tal observação é a tipagem molecular. Essa metodologia é capaz de medir diferenças genéticas entre cepas bacterianas e, em conjunto com as análises filogenéticas, essas diferenças podem ser classificadas em grupos ou complexos clonais, que são baseados na relação e diversidade genética de cada isolado (SCHÜRCH *et al.*, 2018; WANG; KING JORDAN; MAYER, 2014). Sabemos que a clonalidade de uma cepa está relacionada às forças evolutivas exercidas sobre o seu genoma, como a seleção natural e a recombinação (DIDELOT; MAIDEN, 2010). O efeito da recombinação sobre a clonalidade é direto, e quanto maior o número de recombinações menor a clonalidade. Já o efeito da seleção natural é mais complexo, mas pode ser definido como a força que favorece as expansões clonais de mutações adaptativas em uma população (SHAPIRO, 2016).

Desta forma, análises de tipagem molecular podem determinar a relação

genética de cepas de diferentes fontes de isolamento; diferentes locais geográficos e/ou diferentes períodos de tempo, inferindo assim a relação evolutiva desses dados.

3.2 TIPAGEM MOLECULAR

Como mencionado anteriormente, a tipagem molecular é a metodologia utilizada para discriminação de cepas a partir da sua variabilidade genética. A diversidade genética, dentro de um táxon bacteriano, é associada a variantes de nucleotídeo único, a composição variável do genoma acessório e a distribuição espacial de repetições e oligonucleotídeos dentro do genoma (SABAT *et al.*, 2013; TÜMMLER, 2020). Desta forma, essa metodologia pode ser empregada de diferentes maneiras, mas sempre seguindo o mesmo princípio, o de comparação a nível genético. Com o objetivo de fundamentar e facilitar o entendimento epidemiológico de um conjunto de amostras, algumas técnicas podem ser utilizadas, como a análise de padrões de fragmentos de tamanhos diferentes, gerados por enzimas de restrição, como a Macrorestrição do DNA seguida de eletroforese em gel de campo pulsado PFGE (*Pulsed Field Gel Electrophoresis*) (GROTHUESI; TÜMMLER, 1991); através da análise de sequências de um conjunto de genes específicos, como MLST (*Multilocus Sequence Typing*) (URWIN; MAIDEN, 2003) e CH Typing (WEISSMAN *et al.*, 2012), e também através da aplicação do Sequenciamento Completo do Genoma (*Whole Genome Sequencing – WGS*) o qual possibilita a análise de variantes com um poder discriminatório maior. Algumas das técnicas que utilizam esta abordagem são: a tipagem molecular do genoma completo e ou genoma central, como *whole genome* wgMLST e *core genome* cgMLST, onde a comparação é feita gene por gene em todo o genoma (SCHÜRCH *et al.*, 2018), além do mapeamento do genoma total contra um genoma de referência para identificação de variantes SNPs (*Single Nucleotide Polymorphism*) (TÜMMLER, 2020). Essas técnicas serão elucidadas abaixo.

3.2.1 Eletroforese em gel em campo pulsado PFGE

A eletroforese em gel de campo pulsado é globalmente aplicável a todas as bactérias e, portanto, tem sido o método de referência para tipagem de cepas em bacteriologia (TÜMMLER, 2020). Essa técnica se baseia na digestão do genoma bacteriano por endonucleases de restrição, onde a distribuição do tamanho dos fragmentos gerados pela digestão, representam a distribuição espacial dos sítios de clivagem no genoma (GROTHUESI; TÜMMLER, 1991).

Os fragmentos gerados são então separados por eletroforese em gel de campo pulsado (WANG; KING JORDAN; MAYER, 2014). Uma vez que os locais de reconhecimento da maioria das endonucleases de restrição são distribuídos aleatoriamente no cromossomo, os padrões de fragmentos representam impressões digitais genômicas imparciais (*fingerprinting genomic*) e podem, portanto, ser utilizados na atribuição de relações clonais entre cepas (GROTHUESI; TÜMMLER, 1991). Essas variações no padrão de fragmentos de restrição são indicativos de variações no genoma, podendo estar relacionados a inserções, deleções, rearranjos, incorporação ou perda de elementos do genoma acessório. (GINARD et al., 1997; WANG; KING JORDAN; MAYER, 2014).

A comparação dos padrões de PFGE de várias cepas bacterianas fornece informações para saber se essas cepas são ou não geneticamente relacionadas. As cepas com padrões idênticos são consideradas clonais e provavelmente originadas do mesmo ancestral. Por outro lado, as cepas com padrões distintos podem ou não ser geneticamente relacionadas, dependendo do número de bandas diferentes que apresentam (MARTAK et al., 2020).

Embora seja um método muito utilizado, é laborioso e depende bastante tempo. Sendo assim, os métodos baseados em sequência têm se tornado cada vez mais atrativos, pois permitem o depósito e armazenamento da informação em bancos de dados, contribuindo para a acessibilidade e estudo global da epidemiologia molecular (KOTŁOWSKI et al., 2020; MACIUCA et al., 2019).

3.2.2 Tipagem de Sequência Multilocus MLST

A tipagem de sequência multilocus é uma técnica capaz de caracterizar isolados de espécies bacterianas usando sequências de fragmentos internos de,

geralmente, sete genes de manutenção, chamados de *housekeeping genes* (MAIDEN *et al.*, 1998). A cada fragmento de gene, é atribuído um número inteiro correspondente ao alelo. Para cada cepa analisada, a combinação de alelos em cada locus define sua sequência tipo (ST) (WANG; KING JORDAN; MAYER, 2014). Logo, uma ST é definida pelo perfil alélico (combinação dos números inteiros) dos sete loci. Assim, as cepas que têm, pelo menos, quatro alelos em comum são geralmente definidas como o mesmo complexo clonal (FEIL *et al.*, 2004; TÜMMLER, 2020).

Anteriormente, os genes de manutenção de uma (ou mais) cepas eram amplificados por (*multiplex*) PCR e os amplicons de cada alelo eram sequenciados. Com o sequenciamento de nova geração se tornando um método cada vez mais acessível, o MLST atualmente pode ser facilmente extraído *in silico* através da comparação dos dados de sequenciamento do genoma completo contra um banco de dados *online*, como pubMLST (<https://pubmlst.org/>) onde milhares de combinações de loci, também conhecidos como esquemas, estão disponíveis de acordo com a espécie bacteriana (KIMURA, 2018).

Para *Escherichia coli*, podemos utilizar a EnteroBase (<https://enterobase.warwick.ac.uk/>) que dispõe gratuitamente os esquemas de alelos para esta espécie, os quais são compostos por 7 genes *housekeeping*: *adh* (adenilato quinase), *fumC* (fumarato hidratase), *gyrB* (DNA girase), *icd* (isocitrato / isopropilmalato desidrogenase), *mdh* (malato desidrogenase), *purA* (adenilossuccinato desidrogenase) e *recA* (ATP / GTP-binding motif) (https://enterobase.warwick.ac.uk/species/ecoli/allele_st_search).

3.2.3 CH Typing

Outra ferramenta, desenvolvida especialmente para tipagem desta espécie é o CH Typer, o qual utiliza o mesmo princípio de tipagem MLST (WIRTH *et al.*, 2006).

Em 2012, WEISSMAN *et al.*, desenvolveu a tipagem CH, um esquema de tipagem baseado em sequenciamento de dois loci gênicos, para uma determinação rápida dos grupos clonais de *E. coli* patogênica extraintestinal (WIRTH *et al.*, 2006).

A tipagem de CH é baseada na comparação de sequências de *fumC*, um dos genes *housekeeping* usados no esquema MLST, e em um fragmento interno do gene que codifica a adesina fimbrial do tipo 1, *fimH* (WEISSMAN *et al.*, 2012).

Em 2018, o *Center of Genomic Epidemiology* CGE (<https://cge.cbs.dtu.dk/services/chtyper/>) apresentou uma nova ferramenta online para este tipo de caracterização de cepas, chamada CH Typer, a qual permite a tipagem molecular de isolados a partir de sequências brutas do sequenciamento WGS (fastq) ou do genoma já montado (.fasta) (ROER *et al.*, 2018a). Desta maneira, o genoma de diversos isolados de *E. coli* foram recuperados da base de dados EnteroBase para estudo comparativo entre as técnicas de tipagem MLST e CH. Os resultados mostraram que a combinação de *fumC* e *fimH* teve um poder discriminatório ligeiramente maior do que o poder discriminatório obtido com o MLST (ROER *et al.*, 2018a). Consequentemente, o CH Typer é considerado uma ferramenta adequada, que pode atuar como uma alternativa rápida para a vigilância epidemiológica de surtos provocados por *E. coli*. (ROER *et al.*, 2018a; TCHESNOKOVA *et al.*, 2013).

3.2.4 Tipagem de Sequência Multilocus do genoma completo e do genoma central wgMLST/cgMLST

Uma abordagem alternativa para caracterizar cepas filogeneticamente é a tipagem de sequência multilocus do genoma total (*whole genome*) wgMLST ou a tipagem de sequência multilocus do genoma central (*core genome*) cgMLST. Essas metodologias possuem o mesmo princípio descrito para a tipagem de sequência multilocus tradicional (MLST) mas oferece uma resolução muito maior ao utilizar porções significativamente maiores do genoma (SCHÜRCH *et al.*, 2018; TÜMMLER, 2020). Uma comparação entre as técnicas pode ser visualizada na **Tabela1**.

As comparações genômicas de centenas a milhares de genes permite a atribuição de alelos quando estes são comparados a um conjunto de genes centrais conservados. Desta forma, regiões homólogas entre os genomas comparados, podem ser agrupadas em clusters de genes através de ferramentas de bioinformática (PEARCE *et al.*, 2018). Para *E. coli* são utilizados 2.513 loci de genes para análise wgMLST e 25.002 loci para cgMLST (EnteroBase, último acesso em: 26

de out. de 2020).

Embora seja uma metodologia com um alto poder discriminatório e muito utilizada na tipagem molecular baseada em sequências do genoma total, também possui suas limitações, pois regiões recombinantes com uma alta densidade de SNPs não poderão ser filtradas, podendo colapsar e sub-representar as variações alélicas do genoma (SCHÜRCH *et al.*, 2018).

Tabela 1: Diferença entre as metodologias baseadas em MLST (*multilocus sequence typing*).

MLST clássico	Core genome MLST (cgMLST)	Whole genome MLST (wgMLST)
7 loci	2.513 loci	25.002 loci
Genes conservados – <i>housekeeping genes</i>	Genes centrais	Qualquer sequência codificante
Regiões conservadas – baixo poder discriminatório	Regiões variáveis – alto poder discriminatório	Regiões altamente variáveis – altíssimo poder discriminatório
Esquemas diferentes para cada espécie/gênero	Esquemas diferentes para cada espécie/gênero	Esquemas diferentes para cada espécie/gênero

Fonte: EnteroBase. Disponível em: <https://enterobase.readthedocs.io/en/latest/enterobase-tutorials/deeper-lineages.html>. Acesso em: 26 de out. de 2020.

3.2.5 Tipagem de Polimorfismo de Nucleotídeo Único (SNP)

Através do sequenciamento de alto rendimento é possível explorar todas as variantes de nucleotídeo único, além da composição do genoma acessório (SCHÜRCH *et al.*, 2018). Essas abordagens mais abrangentes, são úteis para rastrear a diversidade intra-clonal a fim de compreender a microevolução de um clone ou para acompanhar a aquisição e perda de características de interesse, como virulência ou determinantes de resistência antimicrobiana (TÜMMLER, 2020).

O polimorfismo de nucleotídeo único provavelmente representa as variações genéticas mais abundantes no genoma. Logo, SNPs de todo o genoma

podem ser identificados e validados para discriminar cepas bacterianas (WANG; KING JORDAN; MAYER, 2014).

A tipagem de cepas por sequenciamento do genoma completo, a partir de SNPs, pode ser executada através de mapeamento dos arquivos de leitura (*reads*) contra um genoma de referência ou a partir do genoma montado (*contigs*) (SCHÜRCH *et al.*, 2018). Sendo assim, se uma sequência de referência intimamente relacionada estiver disponível, uma abordagem de mapeamento de leitura pode ser usada para auxiliar na identificação das diferenças entre ela e a sequência recém-gerada. O mapeamento se refere ao alinhamento de cada leitura a uma posição em um genoma de referência. Para cada nucleotídeo de referência, várias leituras são normalmente alinhadas e a base correta para a posição é inferida do nucleotídeo de consenso derivado das leituras sobrepostas (LI; DURBIN, 2010). Os mapeadores registram a posição de cada leitura na referência e conferem uma pontuação de alinhamento para cada leitura; as sequências de leituras emparelhadas são frequentemente geradas no formato de arquivo *Sequence Alignment Map* (SAM) (LI *et al.*, 2009). O arquivo SAM pode então, ser processado para chamar os sítios de variantes de nucleotídeo único (SNVs), também chamados de SNPs. Os SNVs são geralmente relatados em um arquivo *Variant Calling Format* (VCF), que contém informações sobre SNVs e possíveis inserções / exclusões curtas (*Indels*) em comparação com a referência. É importante observar que a escolha da sequência de referência pode ter um grande impacto na variação relatada pelo software. Por exemplo, se as leituras de um organismo relacionado de maneira distante forem comparadas com uma referência, apenas as regiões semelhantes serão mapeadas; quaisquer dados de sequências exclusivos ou suficientemente divergentes não serão mapeados e, portanto, serão omitidos do relatório de variantes. Quanto mais distante a sequência de referência usada, mais regiões serão excluídas da análise (CARRIÇO *et al.*, 2018).

Uma vantagem das abordagens de mapeamento de leitura sobre as abordagens de gene por gene é a capacidade de analisar regiões não codificantes, que tem o potencial de revelar mudanças nas regiões reguladoras de genes que podem ser traduzidas em diferenças fenotípicas, embora isso seja limitado pelo

genoma de referência usado (THORPE *et al.*, 2017).

3.3 ANÁLISE FILOGENÉTICA

À medida que o sequenciamento de nova geração (NGS) se torna cada vez mais rápido e econômico, a tipagem molecular baseada no genoma total e a reconstrução filogenética se tornarão cada vez mais frequentes. A reconstrução da árvore filogenética é uma abordagem poderosa e visualmente intuitiva para inferir relações evolutivas entre sequências microbianas (FELSENSTEIN, 2004), sendo uma ferramenta extremamente útil na epidemiologia molecular (WANG; KING JORDAN; MAYER, 2014).

A reconstrução filogenética de todo o genoma consiste em alinhamentos de sequências do genoma inteiro, chamados de alinhamento de sequência múltipla (WANG; KING JORDAN; MAYER, 2014; YANG; RANNALA, 2012). Estes alinhamentos serão usados para inferir distâncias evolutivas entre os genomas das cepas analisadas e, a partir deles, uma árvore filogenética será inferida. As regiões de duplicação no genoma e a capacidade de transferência horizontal gênica entre cepas bacterianas, pode interferir na inferência das distâncias e, para evitar este problema, a reconstrução filogenética de todo o genoma é frequentemente conduzida usando alinhamentos concatenados dos chamados “genes centrais”, que estão presentes entre todas as cepas analisadas. Esse conjunto de genes, geralmente, se apresentam de forma mais conservada no genoma, sendo muito utilizados neste tipo de análise (WANG; KING JORDAN; MAYER, 2014).

A inferência de árvores filogenéticas pode ser descrita, basicamente, por três tipos de metodologias: fenética, cladística e probabilística (FELSENSTEIN, 1981; HALL, 1988). As árvores inferidas por métodos fenéticos se baseiam no cálculo das distâncias, de acordo com a similaridade entre as sequências observadas. Os cálculos das distâncias entre as sequências seguem vários modelos de substituição de nucleotídeos, que vão se basear na taxa de transição e transversão entre as bases nitrogenadas dos genomas e também em outros

parâmetros, como a quantidade GC ou a posição dos nucleotídeos no códon genético (YANG; RANNALA, 2012). Esses métodos atribuem valores às diferenças existentes entre duas sequências comparadas e geram uma matriz de distância a partir do alinhamento. A matriz então, será utilizada para a construção da árvore através de um algoritmo de agrupamento, como Neighbor Joining (SAITOU; NEI, 1987). Baseado no princípio de evolução mínima (CAVALLI-SFORZA; EDWARDS, 1967), o algoritmo Neighbor Joining identifica os vizinhos que sequencialmente minimizam o tamanho total da árvore. É um dos algoritmos mais usados, eficiente e rápido, mas não considera dados originais das sequências, apenas as distâncias.

Os métodos cladísticos são calculados levando em consideração todos os caminhos evolutivos possíveis, onde a metodologia de parcimônia é utilizada (FELSENSTEIN, 2004). Essa metodologia usa a informação do alinhamento como uma informação evolutiva para inferir a árvore com menor mudança evolutiva que explique os dados observados (FITCH, 1971; HARTIGAN, 1973). Uma das desvantagens é a formação de dados errôneos pela quantidade de homoplasias que podem ser encontradas no genoma, assim como substituições de bases em velocidades diferentes entre as cepas e o tempo computacional exigido para os cálculos (YANG; RANNALA, 2012).

Já os métodos probabilísticos, utilizam cálculos para inferir a árvore que melhor represente a variação de um determinado conjunto de sequências (FELSENSTEIN, 1981). A Máxima verossimilhança (Maximum Likelihood) é um método baseado na construção da árvore mais provável, estatisticamente, o qual utiliza estatística Bayesiana para o teste de hipóteses (RANNALA; YANG, 1996). Isso é uma vantagem em relação aos outros métodos, pois a probabilidade máxima permite muito mais flexibilidade do que outras abordagens, como taxas variáveis de evolução, entre locais e linhagens, e também é particularmente sensível a questões evolutivas, como recombinação, seleção e homoplasia. Ainda assim, algumas desvantagens, como a apresentação não adequada das taxas naturais de substituição e frequências, caso a amostragem não for significativa, e a dificuldade em avaliar todas as árvores possíveis, devido ao tempo computacional exigido, devem ser consideradas. (YANG; RANNALA, 2012).

Sabemos que a evolução microbiana frequentemente tende a se diversificar via recombinação e transferência horizontal gênica. Este tipo de evolução, em que sequências que foram separadas ao longo do tempo evolutivo são, posteriormente, reintroduzidas no mesmo fundo genético, é considerado reticulado, e existem métodos filogenéticos desenvolvidos para capturar e representar este processo evolutivo, como as redes filogenéticas. O uso de tais métodos reticulados pode render uma visão mais sutil e precisa das relações evolutivas entre um grupo de cepas microbianas (HUSON; BRYANT, 2006; WANG; KING; JORDAN; MAYER, 2014).

3.4 CARACTERÍSTICAS DE *E. COLI*

A caracterização de bactérias a nível de cepa pode se tornar um desafio quando estas apresentam uma alta variabilidade genética, principalmente quando apresentam diferenças importantes no potencial de virulência, mas também são comuns no ambiente, como a *Escherichia coli* e outras *Enterobacteriaceae* (KOTŁOWSKI *et al.*, 2020).

E. coli é um membro ubíquo da microbiota bacteriana intestinal permanente em humanos, outros animais de sangue quente e répteis. O ambiente natural dessas bactérias é a camada mucosa do ceco e cólon (KAPER; NATARO; MOBLEY, 2004; LEIMBACH; HACKER; DOBRINDT, 2013). No entanto, no curso da evolução, alguns clones de *E. coli* adquiriram elementos genéticos que os permitiram colonizar diferentes nichos ecológicos e expressar diversos mecanismos de virulência e patogenicidade (JAVADI; BOUZARI; OLOOMI, 2017).

Dependendo da localização da infecção, as cepas patogênicas de *E. coli* podem ser divididas em dois grupos básicos: *E. coli* patogênica intestinal (InPEC) e *E. coli* patogênica extraintestinal (ExPEC) (KAPER; NATARO; MOBLEY, 2004a; LEIMBACH; HACKER; DOBRINDT, 2013). Dentro do grupo do InPEC, os mais importantes são *E. coli* enteroagregativa (EAEC), *E. coli* entero-hemorrágica (EHEC), *E. coli* enteroinvasiva (EIEC), *E. coli* enteropatogênica (EPEC), *E. coli*

enterotoxigênica (ETEC), *E. coli* difusamente aderente (DAEC) e *E. coli* invasiva aderente (AIEC). *E. coli* uropatogênica (UPEC), *E. coli* associada à meningite (MNEC), *E. coli* associada à septicemia (SEPEC) e *E. coli* patogênica aviária (APEC) são os patótipos ExPEC mais comuns (CROXEN; FINLAY, 2010; KAPER; NATARO; MOBLEY, 2004a; KOTŁOWSKI *et al.*, 2020; LEIMBACH; HACKER; DOBRINDT, 2013).

Além disso, a classificação de *E. coli* pode ser feita através da divisão em subgrupos filogenéticos sendo, pelo menos, oito grupos filogenéticos reconhecidos dentro da espécie, denominados A, B1, B2, C, D, E, F e G. Os filogrupos diferem entre si a nível de nucleotídeo, sendo o WGS um grande aliado no estudo de estrutura e classificação populacional e também quanto a características fenotípicas, como tamanho do genoma e frequência e variedade de genes de virulência. (CLERMONT *et al.*, 2019).

3.4.1 Plasticidade do Genoma e Variação Fenotípica

O genoma de *E. coli* apresenta uma estrutura muito dinâmica, podendo ser dividida em dois tipos básicos: genoma central e genoma acessório. Definimos o genoma central como o conjunto de genes que estão presentes em todas as cepas de uma espécie bacteriana, como genes *housekeeping*, genes envolvidos na replicação, transcrição e tradução, ou seja, genes envolvidos em funções metabólicas básicas do microrganismo (LEIMBACH; HACKER; DOBRINDT, 2013; VERRAES *et al.*, 2013). No genoma central, eventos de inserção, deleção ou troca de nucleotídeos são observados, sendo a taxa de mutação, por nucleotídeo por geração entre 10^{-6} a 10^{-9} em bactérias. Por outro lado, o genoma acessório/flexível apresenta altas taxas de variabilidade da sequência de nucleotídeos, e é composto de genes que estão presentes apenas em algumas cepas ou que são únicos para determinados isolados (LEIMBACH; HACKER; DOBRINDT, 2013). Esse conjunto de genes é responsável por diversos fenótipos e adaptações a condições ambientais específicas em uma população ou espécie (MEDINI *et al.*, 2005). Exemplos para o *pool* genético flexível são os elementos genéticos móveis (MGE), como plasmídeos,

fagos, ilhas genômicas, transposons, elementos de inserção, bem como o mecanismo de recombinação de DNA exógeno no DNA do hospedeiro (DIXIT *et al.*, 2015). No contexto da patogenicidade, esse *pool* genético codifica fatores de virulência e resistência, que dão ao patógeno o potencial de colonizar o hospedeiro e causar doenças. Desta forma, a combinação do genoma central e do genoma flexível constitui o pangenoma (LEIMBACH; HACKER; DOBRINDT, 2013; MEDINI *et al.*, 2005).

Para cromossomos bacterianos, a recombinação preferencial entre genomas intimamente relacionados foi invocada como a principal força que mantém a coerência genômica, mas isso não garante um conjunto fixo de características fenotípicas (REDONDO-SALVO *et al.*, 2020a). Estudos relataram que as regiões sintênicas de DNA conservado, quando comparadas entre quaisquer duas cepas de *E. coli*, mostram apenas 3% de divergência de nucleotídeos. Resultados diferentes foram observados ao relatar uma diversidade intra espécie. Estima-se que apenas ~ 40% das proteínas de *E. coli* combinadas são conservadas entre todas as cepas, um conjunto de ~ 2.200 genes com alta homologia constituindo o genoma central (CHAUDHURI; HENDERSON, 2012; TENAILLON *et al.*, 2010; TOUCHON *et al.*, 2009). Os 60% residuais constituem o *pool* genético acessório de parálogos, alelos e singletons, frequentemente co-localizados em elementos móveis. Os resultados apresentados exemplificam a alta plasticidade do genoma de *E. coli*, que resulta nas diversas estratégias de adaptação de diferentes cepas (TENAILLON *et al.*, 2010).

3.4.2 Transferência Horizontal Gênica e suas Consequências

Como mencionado acima, a diversidade do genoma bacteriano e a capacidade de trocar material genético, possibilita a esses microrganismos a aquisição de genes importantes para sua sobrevivência ou dispensar genes acessórios que não estão sendo utilizados. Tanto a aquisição de genes através da transferência horizontal gênica, quanto a redução do genoma com o intuito de economizar energia, é refletido nos tamanhos variáveis do genoma de *E. coli*, que varia de ~ 4,6 a ~ 5,7 Mb (LEIMBACH; HACKER; DOBRINDT, 2013).

Mutações pontuais e rearranjos genéticos apenas levam ao desenvolvimento evolutivo, principalmente sem a criação de novos determinantes

genéticos, enquanto a HGT produz genomas extremamente dinâmicos. Assim, a troca de material genético pode alterar efetivamente o estilo de vida das espécies bacterianas (JAVADI; BOUZARI; OLOOMI, 2017) e, conseqüentemente, desempenhar um papel importante na propagação, de determinantes de virulência e resistência antimicrobiana, entre diferentes cepas e espécies bacterianas (LEIMBACH; HACKER; DOBRINDT, 2013).

Um estudo de vias de transferência horizontal de genes em bactérias, Santiago Redondo-Salvo demonstrou que, embora os plasmídeos tendem a uma recombinação preferencial, em relação a tipos de incompatibilidade, exibem uma diferença chave em relação ao cromossomo bacteriano, operando além da barreira de espécie/gênero do hospedeiro. A troca de plasmídeos entre a família *Enterobacteriaceae* é maior em relação às outras famílias bacterianas, e as espécies como, *Escherichia coli*, *Salmonella* e *Klebsiella* são capazes de trocar plasmídeos com a maioria das bactérias clinicamente relevantes da ordem *Enterobacterales* (REDONDO-SALVO *et al.*, 2020a) destacando ainda mais a importância dos MGEs na disseminação da resistência.

3.5 RESISTÊNCIA MICROBIANA

O uso extensivo de antimicrobianos desencadeou o que poderia ser uma das maiores ameaças à saúde humana, animal e ambiental no século 21: a resistência antimicrobiana (AMR) (OMS, 2015). Na verdade, algumas previsões estimam que as mortes atribuíveis à AMR podem aumentar de 700.000 para 10 milhões de vidas anualmente até 2050 (BRODRICK *et al.*, 2017; BROGAN; MOSSIALOS, 2016; UDAONDO; HUERTAS, 2020; VERRAES *et al.*, 2013).

A resistência antimicrobiana pode ser intrínseca ou adquirida. Onde a resistência intrínseca é característica inerente de uma espécie bacteriana ou gênero para um determinado antibiótico (GUARDABASSI; COURVALIN, 2006). Já a resistência adquirida é quando uma cepa suscetível tornou-se resistente, como consequência de uma alteração genética evolutiva.

A plasticidade do genoma bacteriano e a troca de material genético

através dos MGEs, como plasmídeos, é decorrente da pressão seletiva exercida sobre uma população, favorecendo a expansão clonal de características fenotípicas adaptativas, como a resistência aos antibióticos (SHAPIRO, 2016).

Infecções causadas por cepas MDR foram relatadas pela primeira vez em ambientes hospitalares, onde o uso de antimicrobianos era mais prevalente. No entanto, hoje em dia, os compostos antibióticos são muito utilizados não só para a terapia humana, mas também na terapia, profilaxia, metafilaxia e promoção do crescimento na produção animal (CASTANON, 2007; UDAONDO; HUERTAS, 2020). A exposição a quantidades substanciais de compostos antimicrobianos em animais de produção, como ilustrado na **Figura 1**, tornou os produtos de origem animal, assim como o ambiente das fazendas e abatedouros, potenciais reservatórios de ARGs (GARCÍA *et al.*, 2018; WANG *et al.*, 2019). Com isso, as chances de transmissão aos humanos através da cadeia alimentar, pelo contato direto com os animais e pelo meio ambiente (águas residuais), são aumentadas e contribuem para a disseminação da resistência aos antibióticos (IGWARAN; IWERIEBOR; OKOH, 2018; MAESTRE-CARBALLA *et al.*, 2019). Sabendo que o uso de antibióticos atua como pressão seletiva e contribui para a diversidade do genoma bacteriano, o monitoramento epidemiológico de genes que conferem resistência aos antibióticos, principalmente aqueles utilizados como último recurso terapêutico, como os carbapenêmicos e as polimixinas, é de extrema importância (SUN, 2018).

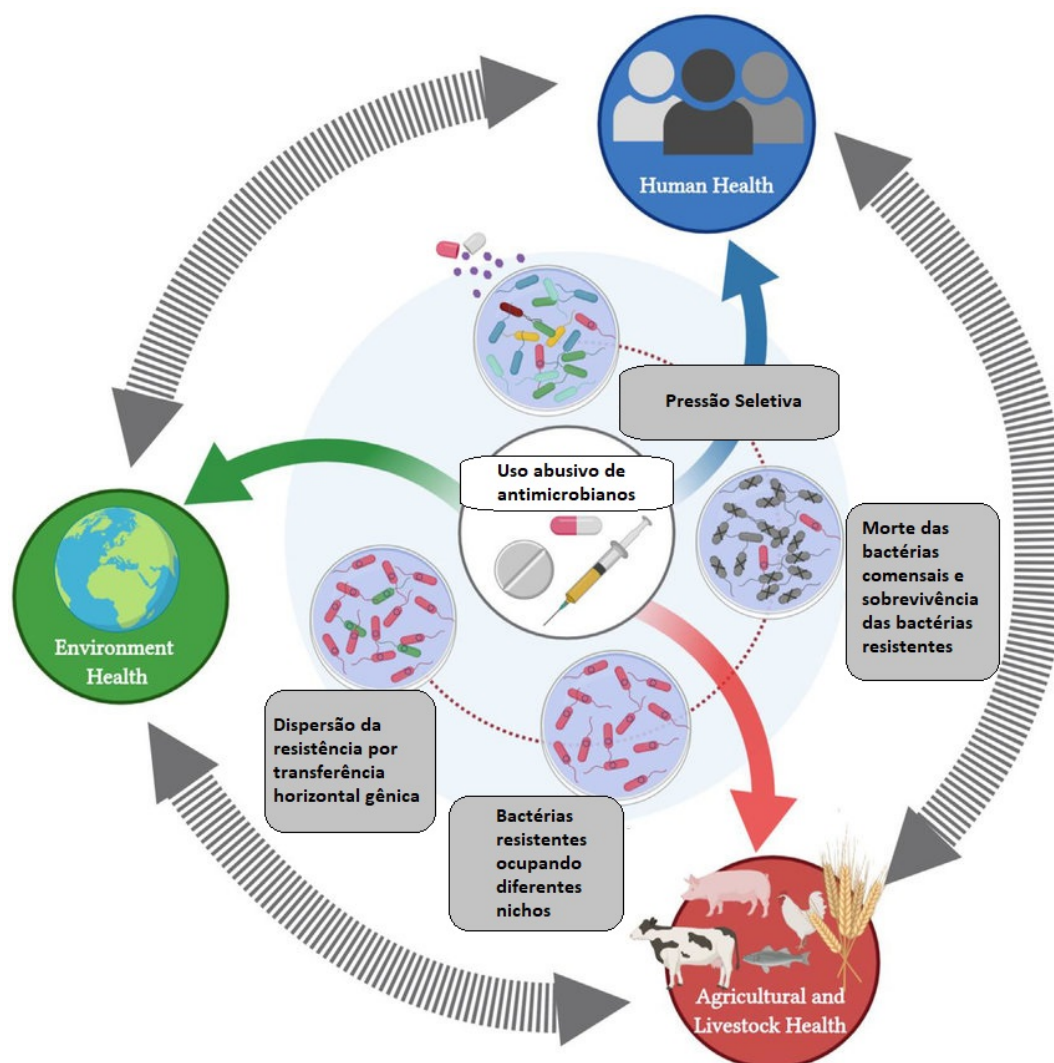


Figura 1: Definição *One Health* sobre a disseminação da resistência bacteriana nos três setores: saúde humana, ambiental e animal. Figura adaptada do artigo (Udaondo; Huertas., 2020).

3.5.1 Resistência às Polimixinas

As polimixinas (polimixina B e E) foram introduzidas no arsenal de antibióticos em 1950, porém, tendo em vista sua alta neurotoxicidade e nefrotoxicidade foram retiradas do mercado. Contudo, devido ao aumento de isolados resistentes aos carbapenêmicos na década de 90, elas tornaram-se o último recurso para o tratamento de infecções graves causadas por bacilos gram-

negativos, para os quais outras opções de tratamento são limitadas (SUN *et al.*, 2018). Já na medicina veterinária, as polimixinas, principalmente do tipo E (colistina) são largamente utilizadas para o controle da diarreia em suínos e aves de produção, além de funcionarem como promotor de crescimento e melhoramento animal (SHEN *et al.*, 2020a).

A resistência às polimixinas até novembro de 2015 era caracterizada por mutações cromossomais e não representava risco quanto à disseminação horizontal. Entretanto, a identificação de um mecanismo de resistência mediada por plasmídeo, através do gene *mcr-1*, entre isolados da família *Enterobacteriaceae*, proveniente de humanos e animais, causou preocupação (LIU *et al.*, 2016b).

MCR-1 é uma fosfoetanolamina transferase, que modifica o lipopolissacarídeo microbiano, adicionando fosfoetanolamina ao lipídeo A, levando a resistência às polimixinas (colistina), através da diminuição da afinidade do fármaco com a parede bacteriana (Liu *et al.*, 2016). Desde o relato do primeiro gene *mcr* (*mcr-1*) (LIU *et al.*, 2016b), algumas variantes do gene, incluindo *mcr-2* (XAVIER *et al.*, 2016), *mcr-3* (YIN *et al.*, 2017), *mcr-4* (CARATTOLI *et al.*, 2017), *mcr-5* (BOROWIAK *et al.*, 2017), *mcr-6* (ABUJOUN *et al.*, 2017), *mcr-7* (YANG *et al.*, 2018b), *mcr-8* (WANG *et al.*, 2018b), *mcr-9* (CARROLL *et al.*, 2019) e *mcr-10* (WANG *et al.*, 2020) têm sido descritas em isolados da família *Enterobacteriaceae*. Além disso, o gene *mcr-1* já foi relatado em mais de 70 países em cinco continentes, em várias espécies (CASELLI *et al.*, 2018; MARTINS-SORENSEN *et al.*, 2020; REBELO *et al.*, 2018), diferentes nichos ecológicos e em associação com diversos plasmídeos, como por exemplo, IncX4, IncI2 e IncHI2 (IGWARAN; IWERIEBOR; OKOH, 2018; NANG; LI; VELKOV, 2019; SHEN *et al.*, 2020b; YANG *et al.*, 2017).

A co-localização de variantes distintas do gene *mcr* em um único isolado já foi descrita (HERNÁNDEZ *et al.*, 2017; XIANG *et al.*, 2018) e plasmídeos conjugativos carreando *mcr-1* e genes de resistência do tipo *bla*_{CMY-2} (MACIUCA *et al.*, 2019), *bla*_{CTX-M-8} (SACRAMENTO *et al.*, 2018a) e carbapenemases do tipo *bla*_{NDM} (WANG *et al.*, 2018b) também já se tornou uma realidade.

No Brasil, vários estudos já relataram a presença de *mcr-1* em isolados de pacientes clínicos (AIRES *et al.*, 2017; CONCEIÇÃO-NETO *et al.*, 2017;

DALMOLIN *et al.*, 2017; FERNANDES *et al.*, 2016b; ZAMPARETTE *et al.*, 2020), isolados ambientais (SACRAMENTO *et al.*, 2018a) e isolados de animais (FERNANDES *et al.*, 2016a; LENTZ *et al.*, 2016; MONTE *et al.*, 2017; MORENO *et al.*, 2019; PALMEIRA *et al.*, 2018; SAIDENBERG *et al.*, 2020; SELLERA *et al.*, 2017). Esses achados ressaltam a capacidade de disseminação do gene *mcr-1* e, conseqüentemente, a escassez de alternativas de tratamento de infecções bacterianas MDR na comunidade.

A abordagem *One Health* reconhece que humanos, animais e saúde ambiental estão intimamente ligados e, portanto, o controle e a vigilância de ARGs precisam ser considerados em todos os três setores (JAVADI; BOUZARI; OLOOMI, 2017).

3.5.2 Animais de Produção como Reservatórios de Genes de Resistência

A Office International des Epizooties — OIE (Organização Mundial de Saúde Animal), responsável pelo Relatório Anual de Antimicrobianos para Uso em Animais, mostrou uma redução de 25% no uso de colistina como fator de crescimento, nos países participantes do relatório, durante o ano de 2018 a 2020. Embora alguns países estejam banindo o uso de antimicrobianos na produção animal (SHEN *et al.*, 2020a), antibióticos como a Colistina, Bacitracina e Polimixina B ainda representam 73% do uso profilático total de antibióticos utilizados nos países participantes do relatório (OIE-ARA., 2020).

No Brasil, o monitoramento da quantidade de antibióticos usados na pecuária ainda é precário. Esses dados estão disponíveis qualitativamente e com pouca vigilância. Em 2016, o Brasil deixou de usar a colistina como promotor de crescimento na pecuária, dando prazo para adequação a lei de 2 anos, entrando em vigor apenas em 2018. Contudo, a proibição não abrange o uso profilático da colistina na produção animal (IN-45 Diário Oficial da União., 2016).

Dados indicam que o uso global de antibióticos na produção de animais destinados a alimentação foi de 131.109 toneladas em 2013 e pode aumentar para 200.235 toneladas até 2030 (WANG *et al.*, 2019). Sendo os animais de produção, um dos principais reservatórios de genes que conferem resistência à colistina e que, tanto a transmissão clonal, quanto a horizontal, contribuem para a disseminação da

resistência (SHEN *et al.*, 2020a), o monitoramento do uso de antimicrobianos deste tipo se torna cada vez mais urgente.

No final de 2019, o Food and Drug Administration dos EUA, disponibilizou o relatório anual de distribuição de antimicrobianos com importância clínica humana, utilizados em animais de produção, destinados a alimentação, onde foi possível observar um aumento de 9% entre os anos de 2017 e 2018 (CVM; FDA; DS, 2018). Dados mais recentes serão disponibilizados no relatório de 2023, onde o FDA, junto ao CMV dos Estados Unidos (Centro de Medicina Veterinária), atuam em um plano de ação de cinco anos para monitoramento da administração de antimicrobianos em ambientes veterinários (<https://www.fda.gov/animal-veterinary/safety-health/antimicrobial-resistance>). Como já sabemos, a utilização de antimicrobianos neste setor, contribui para o surgimento de ARGs em bactérias, que transitam a cadeia alimentar, proporcionando a seleção de cepas resistentes, patogênicas e comensais, no trato intestinal de animais destinados à alimentação humana (BARLAAM *et al.*, 2019).

Além de animais de produção e humanos sintomáticos ou não, o gene *mcr-1* já foi encontrado em bactérias gram negativas, principalmente *E. coli*, em vários tipos de carnes disponíveis em supermercados e vegetais (BARLAAM *et al.*, 2019; GUENTHER *et al.*, 2017; HASMAN *et al.*, 2015; PERRIN-GUYOMARD *et al.*, 2016). Portanto, bactérias resistentes podem contaminar produtos alimentares e colonizar a microbiota humana através da cadeia alimentar pelo manuseio e / ou consumo de alimentos contaminados, como demonstrado na **Figura 2**. (BARLAAM *et al.*, 2019; HOELZER *et al.*, 2017; VAN DEN BOGAARD; STOBBERINGH, 1999; WITTE, 1998).

A microbiota intestinal tem um papel importante no combate à infecções por bactérias resistentes a antibióticos, pois atua como uma barreira natural envolvida na homeostase da saúde de humanos e animais (CLEMENTE *et al.*, 2012); entretanto, o uso de antibióticos de amplo espectro causa um desequilíbrio geral da microbiota residente (disbiose) que pode ter consequências para a saúde de ambas as espécies (CLEMENTE *et al.*, 2012; CULLY M., 2019). Uma dessas consequências é um aumento desproporcional na abundância de bactérias

específicas que podem carregar elementos móveis contendo ARGs (UDAONDO; HUERTAS, 2020).

O estudo de Dobrzanska et al., realizado em 2020, correlacionou mudanças na composição da microbiota intestinal, após antibioticoterapia preventiva, com ganho de peso e doença em bezerros. Neste trabalho, bezerros tratados com florfenicol tiveram a composição da microbiota intestinal alterada, observado-se um enriquecimento de Proteobactérias anaeróbias facultativas, principalmente da espécie *E. coli*. A presença dessa espécie mediando ARGs carregados por elementos móveis, forneceram a resistência a outras drogas não administradas aos bezerros. Os animais tratados com antibióticos também mostraram um aumento no peso corporal em comparação com o grupo não tratado, embora ambos os grupos seguissem a mesma dieta (DOBRZANSKA *et al.*, 2020). Esses resultados indicam que a administração de antibióticos em animais saudáveis leva a uma microbiota desequilibrada, que está relacionada a doenças e ao aumento de peso dos animais. Além disso, a disbiose promove o crescimento de microrganismos portadores de elementos móveis carregando ARGs clinicamente relevantes, e potencialmente aumentando o risco de transmissão de bactérias resistentes a antibióticos em diversos ambientes (DOBRZANSKA *et al.*, 2020; UDAONDO; HUERTAS, 2020).



Elisa Spinelli

Figura 2 : Principais vias de seleção e disseminação de *Escherichia coli* carregando o gene *mcr-1* em animais, humanos e o ambiente. Demonstração de como pode ocorrer a transmissão na cadeia alimentar: 1- fazenda; 2- matadouro; 3- águas residuais; 4- animais domésticos e silvestres; 5- intestino; 6- ambientes hospitalares. Figura adaptada (BARLAAM. et al.,2019).

As complexidades dos mecanismos pelos quais os microrganismos podem adquirir ARGs são impulsionadas não apenas por processos biológicos, mas também por fatores socioeconômicos, que devem ser levados em consideração na construção de estratégias para combater esta ameaça à saúde pública. Medidas integradas de inspeção, vigilância, sequenciamento de última geração e ferramentas de bioinformática avançadas, são necessárias para monitorar e controlar melhor o uso de antibióticos na produção animal e, conseqüentemente, reduzir a

contaminação e transmissão de cepas MDR ao longo da cadeia alimentar (JAVADI; BOUZARI; OLOOMI, 2017) .

4. MATERIAIS E MÉTODOS

4.1 AMOSTRAGEM DO ESTUDO

4.1.1 Características dos Isolados

Um total de 6 cepas de *E. coli* portadoras do gene *mcr-1* foram isoladas de swab cloacal de aves de corte (n = 3) e swab retal de suínos (n = 3) recuperados de fazendas de produção, em 2015 e 2018 no estado do Rio Grande do Sul. As coletas foram realizadas de acordo com os protocolos éticos previamente aprovados para os referidos estudos CEUA UFRGS 29445 e 33494.

Os isolados foram selecionados a partir do resultado positivo para o gene *mcr-1* que serviu como um critério de inclusão na escolha das amostras a serem sequenciadas.

4.1.2 Sequenciamento Completo do Genoma

O sequenciamento completo do genoma dos 6 isolados foi realizado no Laboratório de Resistência Bacteriana (LABRESIS-Hospital de Clínica de Porto Alegre). Para fins de descrição estes isolados serão chamados de RS-BR neste documento.

O DNA genômico foi extraído usando o Wizard® Genomic DNA Purification Kit Manual Técnico (Promega Itália, Milan, Italy) de acordo com as instruções do fabricante. As bibliotecas de sequenciamento foram preparadas com o Nextera XT DNA Sample Preparation Kit (Illumina) de acordo com o protocolo do fabricante. O sequenciamento das cepas foi realizado na plataforma Illumina MiSeq (Illumina, San Diego, CA, Estados Unidos) utilizando MiSeq Reagent Kit V2 (500 ciclos) e cobertura de 100x, produzindo leituras de extremidades emparelhadas de 150 bp.

4.1.3 Isolados selecionados a partir de Banco de Dados Públicos

As sequências do genoma completo dos isolados utilizados para fins comparativos no presente estudo, foram retiradas do banco de dados de Arquivos de Leituras de Sequências do NCBI (SRA) (<https://www.ncbi.nlm.nih.gov/sra>), digitando-se no campo de busca os seguintes termos: “*Escherichia coli* AND *mcr-1*”. A busca resultou em 417 isolados de *E. coli*, em maio de 2020. Os critérios utilizados para a escolha dos arquivos foram: arquivos de sequências referente ao genoma completo; submetidos a sequenciamento pela plataforma Miseq Illumina e provenientes de diferentes países e isolados de diferentes hospedeiros. Após a seleção, dezenove arquivos de sequência foram baixados no formato fastq e utilizados em análises posteriores como grupo de isolados *E. coli mcr-1* positivo.

Para fins de comparação genômica e validação dos métodos, um novo grupo de arquivos de sequências foi baixado do mesmo banco de dados citado acima. Um grupo de 18 sequências de isolados de *E.coli*, já utilizados como *outgroup* em estudos anteriores (JOENSEN *et al.*, 2015) foi escolhido pelos seguintes critérios de inclusão: arquivos de sequências referente ao genoma completo; submetidos a sequenciamento pela plataforma Miseq Illumina; provenientes de diferentes países; isolados de diferentes tipos de hospedeiros e com data de isolamento anterior ao ano 2000. Esses isolados não foram escolhidos com base no seu perfil de resistência e, de forma não proposital, nenhum dos 18 isolados, pertencentes a este grupo, foi positivo para o gene *mcr-1*.

Deste modo, 37 isolados foram selecionados dos bancos de dados públicos os quais, posteriormente, foram comparados a nível filogenético aos isolados RS-BR. As características gerais de todos os isolados utilizados neste estudo, bem como os números de acesso das respectivas sequências, estão representados na **Tabela 1A** do apêndice A..

4.2 ANÁLISES DE BIOINFORMÁTICA

Os arquivos de sequências dos quarenta e três (n=43) isolados, foram submetidos a análise *in silico*. A genotipagem dos 43 isolados, foi realizada

utilizando-se três metodologias diferentes: MLST, CHTyping e SNPs.

O tipo de arquivo fastq foi utilizado como arquivo de entrada em todas as metodologias utilizadas. As leituras brutas foram filtradas de acordo com a qualidade phred das sequências e submetidas a análise *in silico* através de ferramentas online e também através do software CLC Genomic Workbench 20.0.4 (<https://digitalinsights.qiagen.com>).

Para a identificação de ARGs no genoma, o banco de dados online ResFinder 4.0 (<https://cge.cbs.dtu.dk/services/ResFinder/>) e o banco de dados de resistência QIAGEN Microbial Insight-Antimicrobial (QMI-AR) foram utilizados.

Para melhor caracterização dos seis isolados sequenciados no presente estudo, a identificação do sorotipo, dos genes codificadores de fatores de virulência e os tipos de plasmídeos, foram analisados usando as seguintes ferramentas online: SerotypeFinder 2.0 (<https://cge.cbs.dtu.dk/services/SerotypeFinder/>), VirulenceFinder 2.0 (<https://cge.cbs.dtu.dk/services/VirulenceFinder/>) e PlasmidFinder 2.1 (<https://cge.cbs.dtu.dk/services/PlasmidFinder/>), respectivamente.

O *workflow* da metodologia utilizada está representado na **Figura 3**. As ferramentas e softwares, assim como os critérios adotados serão descritos nas próximas seções.

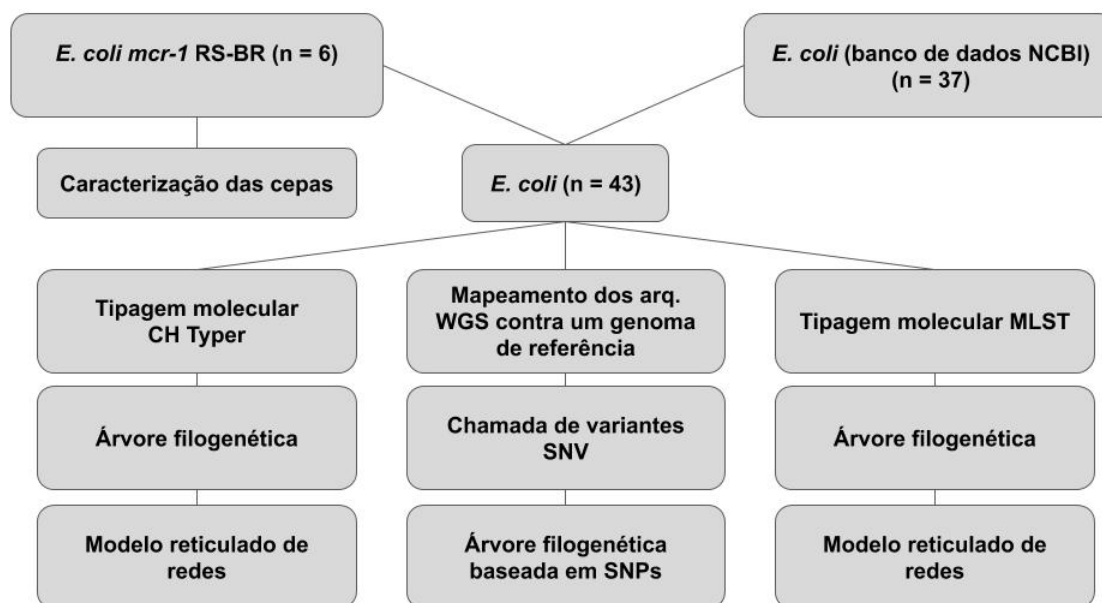


Figura 3: Fluxograma da metodologia utilizada no trabalho.

4.3 TIPAGEM MOLECULAR

A tipagem molecular MLST, dos 43 isolados, foi avaliada *in silico* usando o banco de dados para *E. coli* EnteroBase (<https://enterobase.warwick.ac.uk/species/index/ecoli>) seguindo o esquema de sete locus Achtman (CLERMONT; GORDON; DENAMUR, 2015; WIRTH *et al.*, 2006) e a plataforma MLST do *Center for Genomic Epidemiology* (<https://cge.cbs.dtu.dk/services/MLST/>). As sequências tipo (ST) referente aos 7 alelos dos genes *housekeeping* foram utilizadas para posterior análise filogenética.

Já a tipagem molecular baseada nas sequências internas de 469 nucleotídeos (nt) e 489-nt dos genes *fumC* e *fimH*, respectivamente, (WEISSMAN *et al.*, 2012) foi avaliada através da ferramenta CH Typer (<https://cge.cbs.dtu.dk/services/chtyper/>). As sequências do tipo CH também foram utilizadas para análises filogenéticas.

A tipagem baseada no sequenciamento do genoma total foi realizada no

software CLC Genomic Workbench 20.0.4 (<https://digitalinsights.qiagen.com>). As sequências de leitura (*reads*), dos 43 isolados, foram mapeadas contra um genoma de referência de *Escherichia coli* str. K-12 substr. MG1655 (nº de acesso GenBank NC_000913.3) para a chamada das variantes SNVs. Cada variação do tipo SNV identificada no genoma foi filtrada de acordo com a qualidade de base phred 20. Tanto o mapeamento das sequências, quanto os arquivos referentes as variantes, foram utilizados para a construção da árvore de SNPs. O número de SNVs e a porcentagem do genoma coberto pelo mapeamento serão apresentados na seção de resultados.

4.4 ANÁLISES FILOGENÉTICAS

Análises filogenéticas foram realizadas com o intuito de comparar os 43 isolados de *E. coli* quanto ao seu perfil clonal, distribuição das cepas *mcr-1* e a relação destas com o tipo de hospedeiro.

Todas as árvores foram construídas e visualizadas utilizando o software CLC Genomic Workbench 20.0.4 (<https://digitalinsights.qiagen.com>).

A árvore de SNPs foi inferida a partir dos arquivos de mapeamento das *reads* e da chamada de variantes de nucleotídeo único. Os parâmetros utilizados na construção seguiram o padrão recomendado pelo software, com um Z-score mínimo = 1,96. As SNVs que não obedeceram os critérios de filtro foram descartadas, totalizando 127.788 posições cobertas em todos os genomas analisados e utilizadas na construção da árvore filogenética. O método utilizado para a inferência da árvore foi Máxima Verossimilhança (ML) com 100 replicatas *bootstrap*. O modelo de substituição de nucleotídeos usado foi o Hasegawa-Kishino-Yano (HKY) (HASEGAWA; KISHINO; YANO, 1985) e a árvore foi enraizada no seu ponto médio.

A análise filogenética baseada nas STs consistiu no alinhamento múltiplo das sequências concatenadas dos 7 genes *housekeeping* (*adh*, *fumC*, *gyrB*, *icd*, *mdh*, *purA* e *recA*). A árvore não enraizada foi inferida com o algoritmo Neighbor Joining (SAITOU; NEI, 1987), com 1000 replicatas *bootstrap* e modelo evolutivo

Jukes-Cantor (Jukes e Cantor., 1969). Já a análise referente às sequências do tipo CH foi baseada no alinhamento múltiplo das sequências concatenadas dos dois genes *fumC* e *fimH*. A árvore não enraizada também foi inferida utilizando o algoritmo de agrupamento Neibor Joinning (SAITOU; NEI, 1987), com 1000 replicatas *bootstrap* e modelo evolutivo Jukes-Cantor (Jukes e Cantor., 1969).

Para fins de observação da distribuição evolutiva em um modelo reticulado, duas redes de abrangência mínima foram criadas com o software PopART (LEIGH; BRYANT, 2015), utilizando o alinhamento múltiplo das sequências tipo ST e das sequências tipo CH no formato .phylip.

5. RESULTADOS

Os 43 isolados utilizados neste estudo foram distribuídos da seguinte forma: 51,16 % (22/43) foram isolados de origem animal, 32,56% (14/43) de origem humana, seguidos de 4,65% de origem alimentícia (2/43), 4,65% (2/43) de origem ambiental e 6,98% (3/43) de origem desconhecida (Tabela 1A; apêndice A).

Os isolados RS-BR, apresentaram genes de resistência para as seguintes classes de antimicrobianos: tetraciclina, trimetoprim, fenicol, macrolídeos, quinolonas, aminoglicosídeos, sulfonamidas, colistina e beta-lactâmicos. Os genes relacionados a fatores de virulência mais observados foram: *astA* (*heat-stable enterotoxin 1*); *hly* (*hemolysin F*); *lpfA* (*fimbrial major protein*); *iroN* (*iron transport*); *iss* (*increased serum survival*); *gad* (*glutamate decarboxylase alpha*); *traT* (*complement resistance protein*) e *cma* (*colicin*). A descrição completa dos genes de resistência, fatores de virulência, sorotipo e tipos de plasmídeos identificados está disponibilizado na **Tabela 2**. O gene *mcr-1*, presente nos isolados RS-BR, foram localizados no plasmídeo do tipo IncX4. Não foi observado a presença de outros genes, relacionados a resistência microbiana, sendo expressos no mesmo plasmídeo que carrega o gene *mcr-1*.

Tabela 2: Caracterização molecular das cepas de *Escherichia coli* isoladas de animais de produção, no Rio Grande do Sul.

Isolado/ Hospedeiro	Genes de resistência	Genes de virulência	Sorotipo	Tipo de plasmídeo
215M02423/ Frango	<i>aadA1</i> ; <i>ant(3'')-IIa</i> ; <i>aph(3'')-Ib</i> ; <i>aph(3')-Ia</i> ; <i>aph(6)-Id</i> ; <i>ampC</i> ; <i>ampC1</i> ; <i>ampH</i> ; <i>bla_{TEM-1B}</i> ; <i>bla_{TEM-30}</i> ; <i>bla_{TEM-33}</i> ; <i>bla_{TEM-34}</i> ; <i>bla_{TEM-70}</i> ; <i>bla_{TEM-76}</i> ; <i>bla_{TEM-95}</i> ; <i>bla_{TEM-105}</i> ; <i>bla_{TEM-127}</i> ; <i>bla_{TEM-128}</i> ; <i>bla_{TEM-135}</i> ; <i>bla_{TEM-143}</i> ; <i>bla_{TEM-148}</i> ; <i>bla_{TEM-166}</i> ; <i>bla_{TEM-176}</i> ; <i>bla_{TEM-186}</i> ; <i>bla_{TEM-198}</i> ; <i>bla_{TEM-206-208}</i> ; <i>bla_{TEM-214}</i> ; <i>bla_{TEM-}</i>	<i>astA</i> ; <i>cma</i> ; <i>gad</i> ; <i>iha</i> ; <i>lpfA</i> .	-:H37	IIncB/O/K/Z; IncFIB; IncFII IncI1 IncX4.

	<i>217; bla_{CMY-59}; bla_{CMY-2}; baeR; baeS; sul2, dfrA1, dfrA14, tet(A), qnrB5; qnrB19; eptA; bacA; ermB; mcr-1.1; mdf(A).</i>			
249M02423/ Frango	<i>aadA2; aadA1; bla_{CMY-2}; mdf(A); sul3; tet(A); aadA2; ant(3'')-IIa; aph(3'')-IIa; aph(3'')-Ib; aph(6)-Ic; aph(6)-Id; ampC; ampC1; ampH; bla_{TEM-33}; bla_{TEM-34}; bla_{TEM-57}; bla_{TEM-104}; bla_{TEM-141}; bla_{TEM-164}; bla_{TEM-206}; bla_{TEM-209}; bla_{TEM-214}; bla_{TEM-216}; bla_{CMY-59}; baeR; baeS; cmlA1; sul2; sul3; dfrA1; dfrA12; tet(A); qnrB5; qnrB19; pmrF, mcr-1.1; eptA; bacA; ermB.</i>	<i>astA; gad.</i>	O4:H42	IInc11; Inc12; IncX4.
279M02423/ Frango	<i>dfrA1; dfrA12; ant(3'')-IIa; aadA1; aadA2; aph(3'')-Ib; aph(3'')-IIa; aph(6)-Ic; aph(6)-Id; cmlA1; tet(A); mdf(A); bla_{CMY-2}; bla_{TEM-1B}; sul2; sul3; mcr-1.1; ampC; ampC1; ampH; bla_{CMY-59}; floR; baeR; baeS; qnrB5; qnrB19; pmrF; eptA; bacA; ermB.</i>	<i>astA; celb; cma; gad; iss.</i>	O4:H42	ColpVC; IncB/O/K/Z; Inc11; Inc12; IncX4.
649M05549/ Suíno	<i>mph(B); mdf(A); aph(6)-Id; aph(3'')-Ib; ant(3'')-Ia; dfrA1. ant(3'')-IIa; ampC; ampC1; ampH; bla_{TEM-1A}; bla_{CMY-2}; bla_{TEM-54}; bla_{TEM-90}; bla_{TEM-116}; bla_{TEM-150}; bla_{TEM-156}; bla_{TEM-168}; bla_{TEM-171}; bla_{TEM-183}; bla_{TEM-201}; bla_{CMY-59}; linG; baeS; floR; pmrF; mcr-1.1; sul2; sul1; tet(A); eptA.</i>	<i>cib; cma; cvaC; gad; hlyF; iha; iron; iss; iucC; iutA; lpfA; ompT; sitA; terC; traT.</i>	-.H9	IncFIB; Inc11-I(Gamma); Inc12; IncX4; IncY.
651M05549/ Suíno	<i>mph(B); mdf(A); cmlA1; dfrA1; aadA2; ant(3'')-Ia; ant(3'')-IIa; aph(3'')-Ib; aph(6)-Id; ampC1; ampH; bla_{CTX-M-8}; bla_{TEM-1B}; bla_{TEM-30}; bla_{TEM-33}; bla_{TEM-34}; bla_{TEM-70}; bla_{TEM-76}; bla_{TEM-95}; bla_{TEM-105}; bla_{TEM-127-128}; bla_{TEM-135}; bla_{TEM-143}; bla_{TEM-148}; bla_{TEM-166}; bla_{TEM-176}; bla_{TEM-186}; bla_{TEM-198}; bla_{TEM-206-208}; bla_{TEM-214-215};</i>	<i>cia; gad; neuC; terC.</i>	O145:H 12	IncFIB(K); Inc11-I(Gamma); IncX4.

	<i>bla</i> _{TEM-217} ; <i>mcr-1.1</i> ; <i>floR</i> ; <i>baeR</i> ; <i>baeS</i> ; <i>sul2</i> ; <i>qnrB5</i> ; <i>qnrS1</i> ; <i>sul1</i> ; <i>sul3</i> ; <i>tet(A)</i> ; <i>pmrF</i> ; <i>eptA</i> ; <i>bacA</i> ; <i>ermB</i> .			
914M05549/ Suíno	<i>aadA2</i> ; <i>Inu(F)</i> ; <i>mph(B)</i> ; <i>ant(3'')-Ia</i> ; <i>ant(3'')-IIa</i> ; <i>aph(3'')-Ib</i> ; <i>aph(6)-Id</i> ; <i>mdf(A)</i> ; <i>dfrA1</i> ; <i>ampC</i> , <i>ampC1</i> ; <i>ampH</i> ; <i>bla</i> _{TEM-1B} ; <i>bla</i> _{TEM-30} ; <i>bla</i> _{TEM-33} ; <i>bla</i> _{TEM-34} ; <i>bla</i> _{TEM-70} ; <i>bla</i> _{TEM-76} ; <i>bla</i> _{TEM-95} ; <i>bla</i> _{TEM-105} ; <i>bla</i> _{TEM-127-128} ; <i>bla</i> _{TEM-135} ; <i>bla</i> _{TEM-143} ; <i>bla</i> _{TEM-148} ; <i>bla</i> _{TEM-166} ; <i>bla</i> _{TEM-176} ; <i>bla</i> _{TEM-186} ; <i>bla</i> _{TEM-198} ; <i>bla</i> _{TEM-206} ; <i>bla</i> _{TEM-207-208} ; <i>bla</i> _{TEM-214-215} ; <i>bla</i> _{TEM-217} ; <i>linG</i> ; <i>baeR</i> ; <i>baeS</i> ; <i>qnrB5</i> ; <i>qnrS1</i> ; <i>qnrB19</i> ; <i>mcr-1.1</i> ; <i>floR</i> ; <i>sul1</i> ; <i>sul2</i> ; <i>tet(A)</i> ; <i>pmrF</i> ; <i>eptA</i> .	<i>cea</i> ; <i>cia</i> ; <i>cma</i> ; <i>cvaC</i> ; <i>gad</i> ; <i>hlyF</i> ; <i>iroN</i> ; <i>iss</i> ; <i>iucC</i> ; <i>iutA</i> ; <i>lpfA</i> ; <i>ompT</i> ; <i>sitA</i> ; <i>terC</i> .	-:H21	Col(pHAD28); Col156; ColpVC; IncFIA; IncFIB(AP001918); IncFIC(FII); IncI1- I(Gamma); IncQ1; IncR; IncX4.

Dados obtidos *in silico* através das ferramentas ResFinder, VirulenceFinder, SeroTypeFinder e PlamidFinder, disponíveis no Center for Genomic Epidemiology (<https://cge.cbs.dtu.dk/services/>) e do software CLC Genomic Workbench 20.0.4 (<https://digitalinsights.qiagen.com>) utilizando o banco de dados CARD e QMI-AR para identificação dos genes de resistência.

Genes de Resistência: *aadA1* (streptomycin and spectinomycin resistance); *aph(6)-Id* (Streptomycin phosphotransferase, APH(6)-Id); *aadA2* (streptomycin and spectinomycin resistance); *aph(3'')-Ia* (Aminoglycoside 3'-phosphotransferase); *aph(6)-Ic* (Aminoglycoside O-phosphotransferase APH(6)-Ic); *aph(3'')-Ib* (Aminoglycoside 3'-phosphotransferase); *aph(3'')-IIa* (Aminoglycoside 3'-phosphotransferase); *ant(3'')-Ia* (Aminoglycoside (3'') (9) adenylyltransferase); *mph(B)* (Macrolide phosphotransferase); *ermB* (Macrolide-lincosamide-streptogramin B resistance protein); *mdf(A)* (Multidrug transporter MdfA); *dfrA14* (Dihydrofolate reductase); *cmlA1* (Bcr/CflA family efflux transporter); *dfrA12* (Dihydrofolate reductase); *dfrA1* (Dihydrofolate reductase); *Inu(F)* (lincosamide nucleotidyltransferase); *bla*_{CMY-2} (Beta-lactamase CMY-2); *bla*_{CMY-59} (Beta-lactamase CMY-59) *bla*_{TEM} (Beta-lactamase TEM-1-217); *bla*_{CTXM-8} (Beta-lactamase CTX-M); *ampC* (Beta-lactamase ampC); *ampC1* (Beta-lactamase ampC1); *ampH* (Beta-lactamase ampH); *tet(A)* (Tetracycline resistance protein, class A); *qnrB19* (Plasmid mediated quinolone resistance determining protein); *qnrB5* (Plasmid mediated quinolone resistance determining protein); *qnrS1* (Plasmid mediated quinolone resistance determining protein); *mcr-1.1* (MCR phosphoethanolamine transferase); *sul1* (sulfonamide resistant dihydropteroate synthase); *sul2* (sulfonamide resistant dihydropteroate synthase); *sul3* (sulfonamide resistant dihydropteroate synthase); *floR* (chloramphenicol exporter); *baeR* (transcriptional regulatory protein BaeR); *baeS* (signal transduction histidine-protein kinase BaeS); *linG* (lincosamide nucleotidyltransferase); *bacA* (undecaprenyl-diphosphatase); *pmrF* (polymyxin resistance glycosyltransferase transmembrane protein); *eptA* (phosphoethanolamine transferase EptA).

Genes de Virulência: *astA* (heat-stable enterotoxin 1); *hly* (hemolysin F); *lpfA* (fimbrial major protein); *iroN* (iron transport); *iss* (increased serum survival); *gad* (glutamate decarboxylase alpha); *traT* (complement resistance protein); *iss* (increased serum survival); *cma* (colicin); *celB* (Beta-galactosidase); *cib* (Colicin-Ib); *cvaC* (Colicin-V); *iha* (Bifunctional siderophore receptor/adhesin Iha); *iucC* (Aerobactin synthase); *iutA* (IutA ferric aerobactin siderophore receptor 1); *ampT* (Antimicrobial peptides); *sitA* (Iron transport protein, periplasmic binding protein); *terC* (tellurite resistance); *cia* (Colicin-Ia); *neuC* (UDP-N-acetylglucosamine 2-epimerase); *cea* (cell adhesion molecule).

A tipagem molecular por MLST dos 43 isolados resultou em 35 STs únicas, sendo a ST-10 e a ST-101 as de maior prevalência 11,63 % e 6,98 %, respectivamente. Já a tipagem molecular por CH Typing, apresentou 32 CH únicos,

sendo CH 11-25 9,30%, CH 11-0 6,98% e CH 29-38 6,98% os tipos mais abundantes entre os 43 isolados. Os resultados estão representados na **Tabela 3**.

Tabela 3: Características gerais de todos os isolados utilizados no estudo.

ID Run	Bio project	Bio Sample	Ano isolamento	País	Hospedeiro	ST	CC	CH	<i>mcr-1</i>	Antibióticos
215M02423*	PRJNA674879	SAMN16678977	2015	Brasil	Frango	58	155	4-32	x	Tetraciclina; Trimetoprim; Macrolídeo; Fluoroquinolona e Quinolona; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
249M02423*	PRJNA674879	SAMN16678978	2015	Brasil	Frango	2491	-	11-25	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídeo; Fluoroquinolona e Quinolona; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
279M02423*	PRJNA674879	SAMN16678979	2015	Brasil	Frango	2491	-	11-25	x	Tetraciclina; Trimetoprim; Macrolídeo; Fluoroquinolona e Quinolona; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina

649M05549*	PRJNA674879	SAMN16678980	2018	Brasil	Suíno	101	101	11-0	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídeo; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
651M05549*	PRJNA674879	SAMN16678981	2018	Brasil	Suíno	10	10	41-158	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídeo; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
914M05549*	PRJNA674879	SAMN16678982	2018	Brasil	Suíno	101	101	41-27	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídeo; Fluoroquinolona e Quinolona; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
DRR065948	PRJDB5007	SAMD00056132	2013	Japão	Bovino	457	-	88-145	x	Tetraciclina; Trimetoprim; Macrolídeo; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
DRR138478	PRJDB7066	SAMD00127158	2018	Tailândia	Humano	4014	101	41-86	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídeo; Fluoroquinolona e

										Quinolona; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
ERR1360256	PRJEB13576	SAMEA3931471	2016	Reino Unido	Suíno	6913	-	932-215	x	Tetraciclina; Macrolídeo; Fluoroquinolona e Quinolona; Betalactâmico; Colistina
ERR1407279	PRJEB13470	SAMEA3980468	2016	-	Retalho de alimento	88	23	4-39	x	Tetraciclina; Macrolídeo; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
ERR3479038	PRJEB34000	SAMEA5863154	2018	Bangladesh	Frango	1196	-	6-31	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídeo; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
ERR3479039	PRJEB34000	SAMEA5863155	2018	Bangladesh	Frango	1818	-	82-86	x	Macrolídeo; Fluoroquinolona e Quinolona; Aminoglicosídeo; Betalactâmico; Colistina
SRR10099927	PRJNA312475	SAMN12730721	2011	Espanha	Suíno	763	-	4-31	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídeo; Aminoglicosídeo;

										Sulfonamida; Betalactâmico; Colistina
SRR10099929	PRJNA312475	SAMN12730719	2007	Espanha	Suíno	224	-	4-61	x	Tetraciclina; Trimetoprim; Fenicol; Lincosamida; Colistina
SRR10099933	PRJNA312475	SAMN12730715	2006	Espanha	Suíno	10	10	11-54	x	Tetraciclina; Trimetoprim; Macrolídio; Aminoglicosídeo; Sulfonamida; Colistina
SRR10099940	PRJNA312475	SAMN12730726	2014	Espanha	Suíno	10	10	11-54	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídio; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
SRR10099942	PRJNA312475	SAMN12730724	2012	Espanha	Suíno	156	156	29-38	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídio; Fluoroquinolona e Quinolona; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
SRR10099945	PRJNA312475	SAMN12730713	2005	Espanha	Suíno	156	156	29-38	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídio; Aminoglicosídeo; Sulfonamida;

										Betalactâmico; Colistina
SRR10392137	PRJNA587479	SAMN13191328	2017	Egito	Carne crua	101	101	29-38	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídio; Fluoroquinolona e Quinolona; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
SRR10521259	PRJNA587093	SAMN13197356	2017	Brasil	Água	683	-	4-121	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídio; Aminoglicosídeo; Sulfonamida; Colistina
SRR5328644	PRJNA376494	SAMN06480071	2012	China	Suño	7314	-	11-521	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídio; Fluoroquinolona e Quinolona; Aminoglicosídeo; Rifampicina; Sulfonamida; Colistina
SRR5350657	PRJNA377578	SAMN06463354	2016	Canadá	Humano	3944	-	11-025	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídio; Fluoroquinolona e Quinolona; Aminoglicosídeo; Rifampicina; Sulfonamida;

										Betalactâmico; Colistina
SRR6750074	PRJNA434359	SAMN08554943	2012	China	Humano	6764	-	11-0	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídio; Fluoroquinolona e Quinolona; Aminoglicosídeo; Fosfomicina; Sulfonamida; Betalactâmico; Colistina
SRR6791684	PRJNA436232	SAMN08622661	2016	Estados Unidos	Humano	1485	648	31-58	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídio; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
SRR6791870	PRJNA436242	SAMN08622909	2017	Estados Unidos	Humano	10	10	11-24	x	Tetraciclina; Trimetoprim; Fenicol; Macrolídio; Fluoroquinolona e Quinolona; Aminoglicosídeo; Sulfonamida; Betalactâmico; Colistina
ERR1010266	PRJEB10700	SAMEA3529297	1988	Estados Unidos	Humano	216	-	11-23		Macrolídio
ERR1010278	PRJEB10700	SAMEA3529298	1973	Israel	Bovino	5768	-	746-453		-
ERR1010283	PRJEB10700	SAMEA3529299	1991	Dinamarca	Água	500	-	11-25		Tetraciclina; Trimetoprim;

										Fenicol; Macrolídio; Sulfonamida; Betalactâmico
ERR1010284	PRJEB10700	SAMEA3529306	1992	Espanha	Cabra	5775	-	22-48		Tetraciclina; Macrolídio; Aminoglicosídeo;
ERR1010246	PRJEB10700	SAMEA3529332	1990	Guiné- Bissau	Humano	1147	-	4-233		Tetraciclina; Macrolídio; Aminoglicosídeo; Betalactâmico
ERR1010252	PRJEB10700	SAMEA3529339	1992	Dinamarca	Humano	847	-	6-32		Macrolídio; Aminoglicosídeo
ERR1010255	PRJEB10700	SAMEA3529342	1990	Guiné- Bissau	Humano	181	168	11-0		Macrolídio
ERR1010277	PRJEB10700	SAMEA3529365	1986	Reino Unido	Humano	10	10	11-215		Macrolídio
ERR1010280	PRJEB10700	SAMEA3529378	1982	Coreia do Sul	Suíno	3057	-	54-0		Tetraciclina; Macrolídio; Aminoglicosídeo; Betalactâmico
ERR1010215	PRJEB10700	SAMEA3529382	1980	Estados Unidos	Humano	1434	-	11-53		Macrolídio
ERR1010216	PRJEB10700	SAMEA3529383	1984	Libéria	Humano	5782	-	748-0		Macrolídio
ERR1010285	PRJEB10700	SAMEA3529386	1999	Dinamarca	Humano	5777	-	14-57		Macrolídio
ERR1010184	PRJEB10700	SAMEA3529397	1950	-	Humano	583	122	24-21		Macrolídio
ERR1010185	PRJEB10700	SAMEA3529398	1955	Reino Unido	Humano	3	20	4-54		Macrolídio
ERR1010187	PRJEB10700	SAMEA3529400	1961	-	Suíno	5786	-	11-23		Macrolídio

ERR1010188	PRJEB10700	SAMEA3529401	1960	-	Suíno	1260	-	11-24		Macrolídio
ERR1010247	PRJEB10700	SAMEA3529333	1979	Suécia	Suíno	871	-	7-25		Macrolídio
ERR1010248	PRJEB10700	SAMEA3529334	1997	Dinamarca	-	1308	-	6-32		Macrolídio

Características gerais de todos os isolados utilizados no estudo. O tipo de sequencia ST, o complexo clonal CC e o clonotipo CH foram obtidos através na análise *in silico* pelas ferramenta MLSTFinder e CHtyper, disponíveis em :<http://www.genomicepidemiology.org/>. As classes de antimicrobianos é resultado de análise de resistência *in silico* pelo software CLC Genomic Workbench 20.0.4, disponível em : <https://digitalinsights.qiagen.com>. O ID de corrida, Bio Project, BioSample, ano, País e tipo de hospedeiro foram retirados do banco de dados público arquivos SRA NCBI, disponível em:<https://www.ncbi.nlm.nih.gov/sra>. As amostras sequenciadas neste estudo e provenientes do Laboratório de Microbiologia Molecular e Saúde estão representadas com *.

Os três isolados RS-BR de frango (215M02423, 249M02423 e 279M02423) foram distribuídos em duas STs diferentes sendo o isolado 215M02423 ST-58; CC 155 e os isolados 249M02423 e 279M02423 ST-2491 com complexo clonal desconhecido. Os 3 isolados RS-BR de suínos foram distribuídos em duas STs diferentes e dois complexos clonais, sendo o isolado 649M05549 e 914M05549 ST-101; CC 101 e o isolado 651M05549 ST-10; CC10.

A tipagem por CH Typing apresentou os seguintes CH para os isolados RS-BR de frango: 215M02423; CH 4-32, 249M02423 e 279M02423; CH 11-25 e as seguintes para os isolados RS-BR de suíno: 649M05549; CH 11-0, 651M05549; CH 41-158 e 914M05549; CH 41-27. Já os isolados provenientes do banco de dados público (n = 37) apresentaram 35 STs únicas e 30 clonotipos CH únicos. A sequência tipo mais predominante foi a ST-10 encontrada em quatro isolados (SRR10099933, SRR6791870, ERR1010277 e SRR6791870) e o clonotipo mais predominante foi CH 29-38, encontrado em três isolados (SRR10099942, SRR10099945 e SRR10392137).

A árvore filogenética de STs, construída com as sequências concatenadas dos 7 genes *housekeeping*, apresentou uma distribuição heterogênea das sequências tipo ao longo da árvore. Não foi possível observar uma tendência de agrupamento entre as cepas com o mesmo tipo de hospedeiro ou região geográfica. A distribuição das STs nos diferentes hospedeiros estão representadas na **Figura 4**.

Os resultados da rede de abrangência mínima das diferentes STs estão de acordo com o representado na árvore filogenética e estão representadas na **Figura 5**. O isolado RS-BR de frango 215M02423 ST-58 se mostrou relacionado com os isolados SRR10521259 ST-683 proveniente de águas residuais no Brasil e o isolado de origem alimentícia, ERR1407279 ST-88. Os isolados RS-BR de frango 249M02423 e 279M02423, de mesma ST-2491, foram intimamente relacionados com o isolado SRR6750074 ST-6764 de origem humana. O isolado RS-BR de suíno 651M05549 ST-10 se relacionou com isolados de mesma ST, sendo a maioria de origem animal (suínos). Os isolados RS-BR de suínos 649M05549 e 914M05549, ambos de ST-101 foram relacionados com um isolado de origem alimentícia de mesma ST.

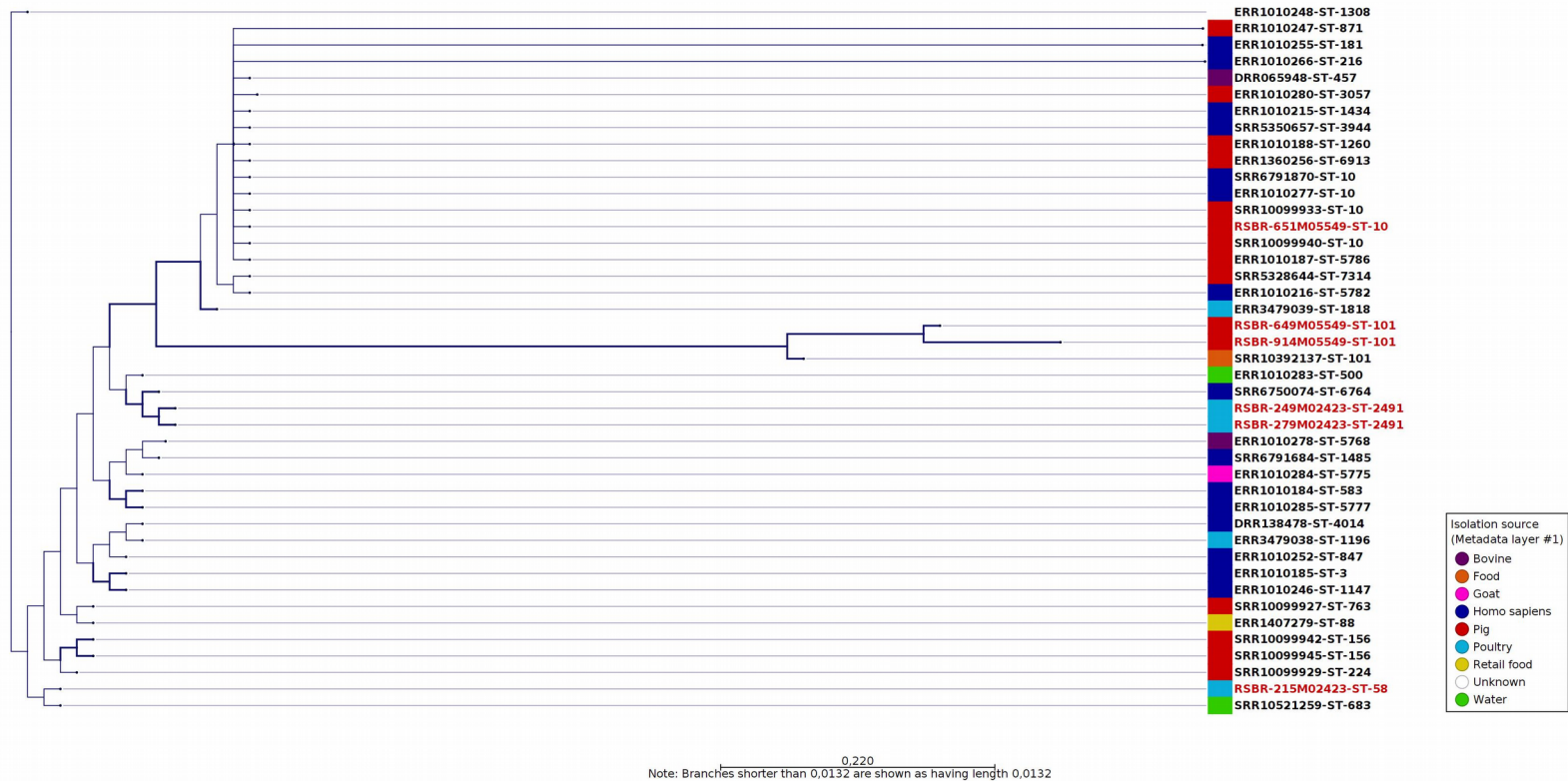


Figura 4: Árvore filogenética construída a partir das sequências do genes *housekeeping* concatenadas. A árvore foi inferida pelo algoritmo Neighbor Joining com 1000 replicatas *bootstrap* pelo software CLC Genomic Workbench 20.0.4 (<https://digitalinsights.qiagen.com>). As cores ao lado direito da árvore indicam os tipos de hospedeiros em que as cepas de *Escherichia coli* foram isoladas. Os ramos destacados representam um valor de *bootstrap* $\geq 90\%$. Os ramos que apresentaram uma distância igual a zero foram colapsados para melhor visualização. As amostras de frango e suíno provenientes do banco de amostras do nosso laboratório estão destacadas em vermelho.

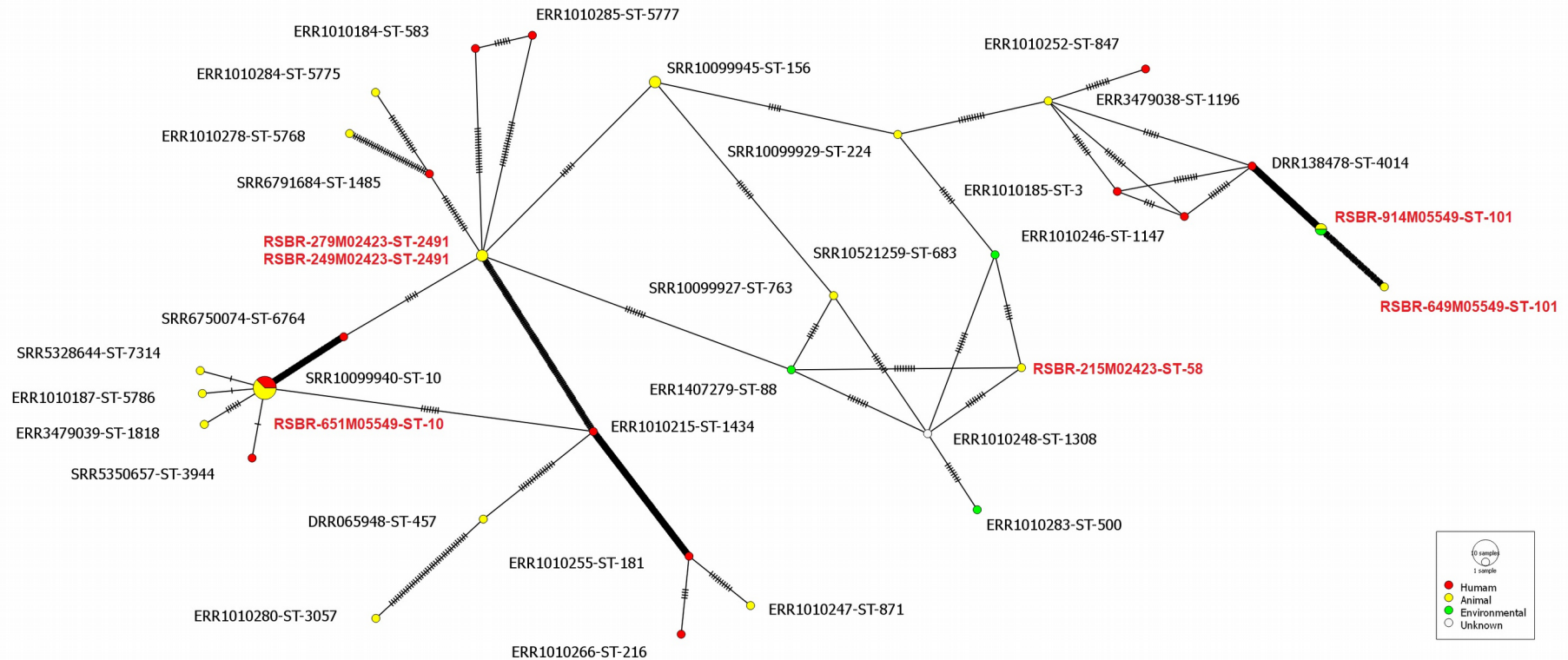


Figura 5: Árvore de abrangência mínima construída com a ferramenta online popART (<http://popart.otago.ac.nz/index.shtml>) a partir do alinhamento múltiplo dos genes *housekeeping*. Cada traço representa um evento de mutação. A diversidade de nucleotídeo foi $\pi = 0.213877$. As diferentes cores, representadas no canto inferior direito da figura representam a origem dos isolados, sendo: vermelho (origem humana); amarelo (origem animal) e verde (origem ambiental). Os isolados destacados na cor vermelha se referem aos isolados do presente estudo.

Utilizando o método de tipagem baseado na variabilidade de apenas dois genes podemos ver na **Figura 6** uma tendência maior a clusterização de forma menos discriminada. Ainda assim, não foi possível observar uma relação entre o tipo CH e o tipo de hospedeiros ou local geográfico.

A rede de abrangência mínima para as sequências do tipo CH apresentou 5 grupos com 2 isolados ou mais. Nessa rede, foi possível observar isolados de diferentes fontes se relacionando com poucos eventos de mutações entre os grupos e com uma diversidade de nucleotídeo ($\pi = 0.027$) menor do que a observada na rede de STs ($\pi = 0.213877$). Os grupos observados na árvore filogenética também foram visualizados na rede de abrangência mínima da **Figura 7**.

O resultado do mapeamento dos arquivos de leitura SRA dos 43 isolados, assim como a porcentagem de cobertura do genoma de referência, estão representados na **Tabela 1A** do apêndice A. Todos os SNVs foram filtrados, resultando em 127.788 posições comuns aos 43 isolados, as quais foram utilizadas na construção da árvore filogenética.

A comparação genômica baseada nos polimorfismos de nucleotídeo únicos encontrados no mapeamento do genoma total, gerou uma variabilidade de sequências distribuídas em uma árvore filogenética de SNPs, enraizada no ponto médio conforme apresentado na **Figura 8**. Os 43 isolados foram divididos em três sub-árvores, sendo uma delas, destacada pela cor vermelha, composta por isolados *mcr-1* positivos e, na sua maioria, isolados provenientes de suínos. Nas outras duas sub-árvores, podemos observar uma distribuição mais diversa entre os tipos de hospedeiros. A matriz de alinhamento de todas as amostras, utilizadas na construção da árvore de SNPs, está disponível na **Figura 1A** do apêndice A.

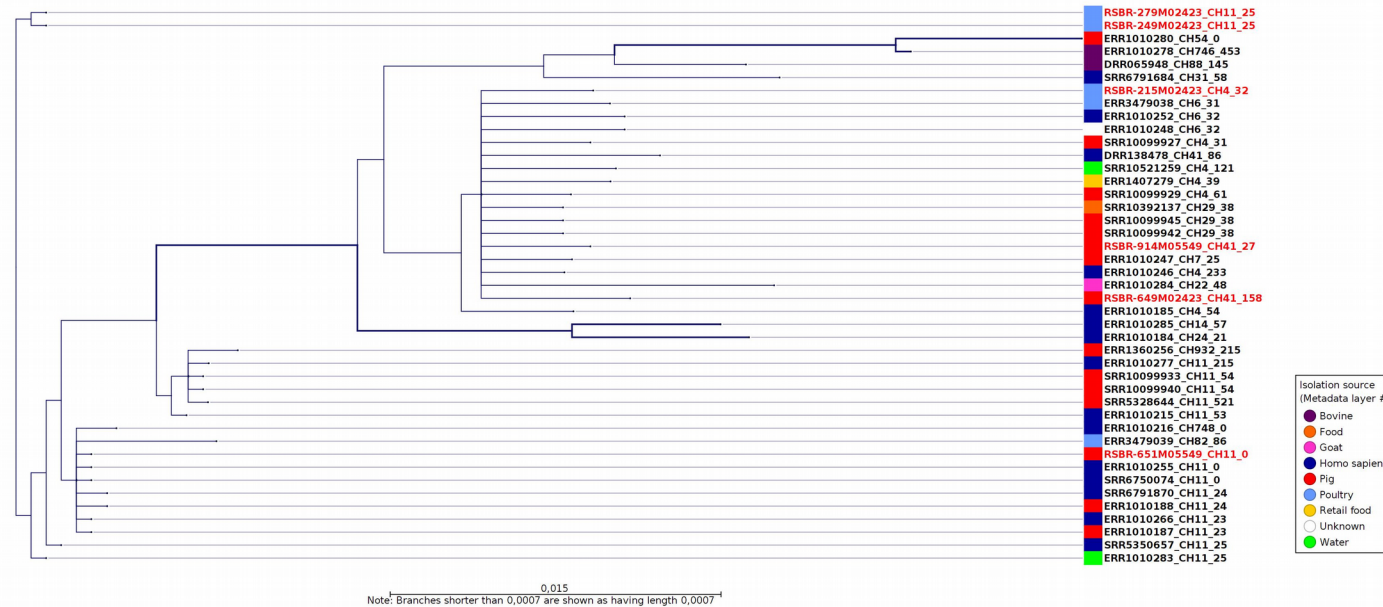


Figura 6: Árvore filogenética construída a partir das sequências do genes *fumC* e *fimH* concatenadas. A árvore foi inferida pelo algoritmo Neighbor Joining com 1000 replicatas *bootstrap* pelo software CLC Genomic Workbench 20.0.4 (<https://digitalinsights.qiagen.com>). As cores ao lado direito da árvore indicam os tipos de hospedeiros em que as cepas de *Escherichia coli* foram isoladas. Os ramos destacados representam um valor de *bootstrap* $\geq 90\%$. Os ramos que apresentaram uma distância igual a zero foram colapsados para melhor visualização. As amostras de frango e suíno provenientes do banco de amostras do nosso laboratório estão destacadas em vermelho.

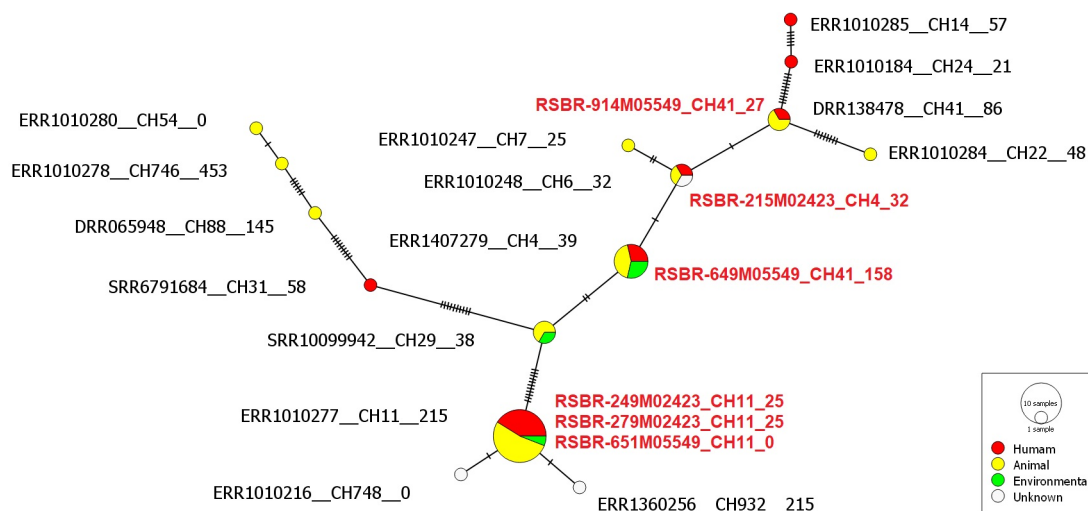


Figura 7: Árvore de abrangência mínima construída com a ferramenta *online* popART (<http://popart.otago.ac.nz/index.shtml>) a partir do alinhamento múltiplo dos genes *fumC* e *fimH*. Cada traço representa um evento de mutação. A diversidade de nucleotídeo foi $\pi = 0.027$. As diferentes cores, representadas no canto inferior direito da figura representam a origem dos isolados, sendo: vermelho (origem humana); amarelo (origem animal) e verde (origem ambiental). Os isolados destacados na cor vermelha se referem aos isolados do presente estudo.

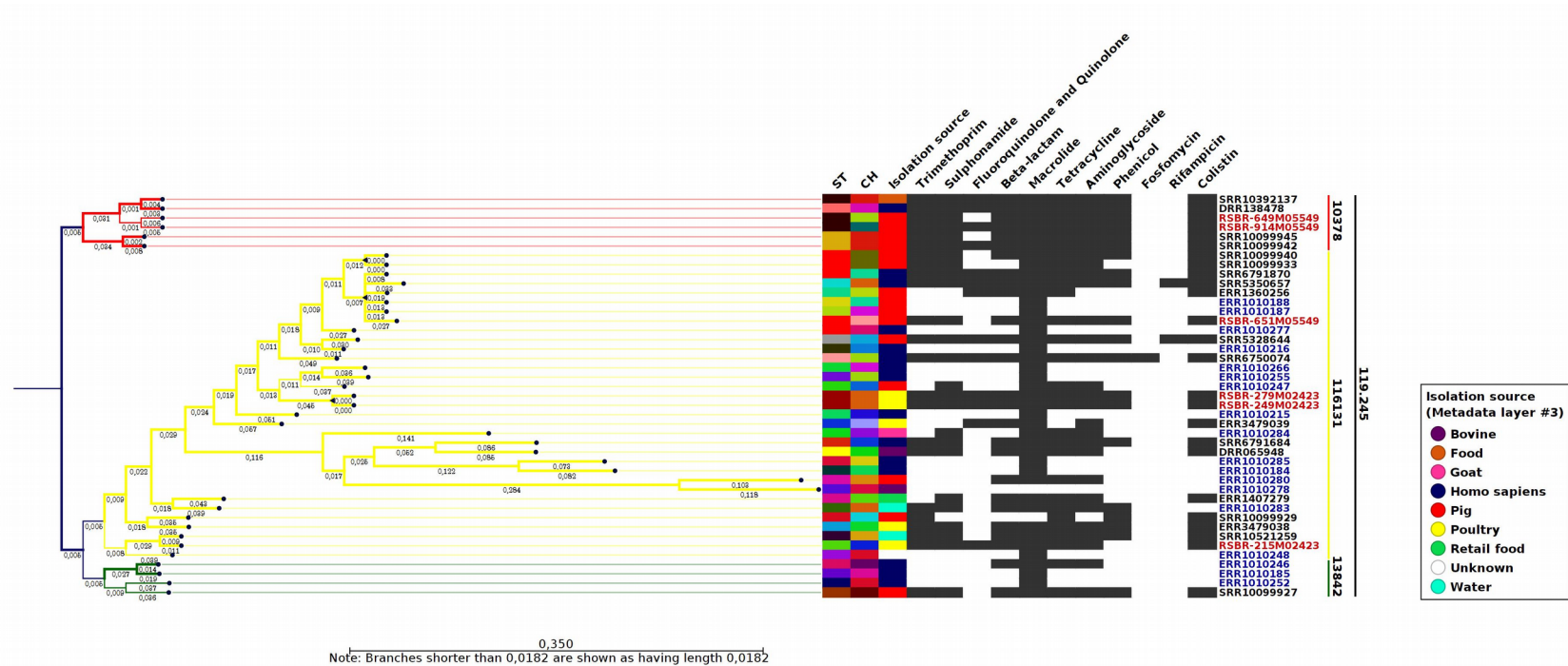


Figura 8: Árvore filogenética construída a partir das variantes de nucleotídeos únicos (SNVs) através do mapeamento do genoma total contra o genoma de referência (n° de acesso GenBank NC_000913.3). A árvore foi inferida utilizando o software CLC Genomic Workbench 20.0.4 (<https://digitalinsights.qiagen.com>) pelo método de Máxima Verossimilhança (ML) com 100 replicatas *bootstrap*. O modelo de substituição de nucleotídeos usado foi o Hasegawa-Kishino-Yano (HKY) e a árvore foi enraizada no seu ponto médio. As cores ao lado direito representam o tipo de hospedeiro em que as cepas de *Escherichia coli* foram isoladas. Os ramos destacados representam um valor de *bootstrap* $\geq 90\%$. As amostras de frangos e suínos provenientes do banco de amostras do nosso laboratório estão destacadas em vermelho. As amostras que foram isoladas antes do ano 2000 estão destacadas na cor azul. As diferentes cores ao lado direito da figura representam os diferentes tipos de sequência ST, diferentes clonotipos CH e a presença ou ausência de genes que conferem resistência a diferentes classes de antimicrobianos. Os números ao lado direito da árvore representam o n° de SNPs entre as sub-árvores e dentro de cada grupo.

6. DISCUSSÃO

Em relação a caracterização dos isolados RS-BR, podemos observar que cepas intimamente relacionadas exibiram perfis de virulência semelhantes, mas não idênticos, como no caso dos isolados de frango 249M02423 e 279M02423, ambos de ST-2491, e entre os isolados de suíno 249M05549 e 914M05549, ambos ST-101. Esses resultados estão de acordo com os resultados apresentados por Reid et. al., em que ao analisar os fatores de virulência para um grupo de isolados clonalmente relacionados observou perfis de virulência semelhantes, mas raramente idênticos (REID et. al., 2019). Observamos também que o gene *mcr-1* estava sendo carregado por um plasmídeo do tipo IncX4. Nenhum outro gene de resistência foi encontrado no mesmo plasmídeo, apoiando resultados já observados em outros estudos (NANG; LI; VELKOV, 2019; SHEN et al., 2020b). Além disso, é possível observar uma prevalência do gene *bla*_{CMY-2} presente nos 3 isolados de frango e do gene *floR* nos 3 isolados de suínos.

Maiuca et al., já haviam relatado isolados de *E. coli* provenientes de frangos, na Romênia, que além de transportarem o gene *mcr-1*, também transportavam o gene *bla*_{CMY-2}, capaz de conferir resistência a cefalosporinas de terceira geração. Além dos isolados de suínos, a grande maioria dos isolados recuperados do banco de dados público, apresentaram resistência a classe de antimicrobianos fenicol, **Tabela 3**. Resultados semelhantes já foram relatados por Matamoros et. al., onde foram detectados genes: *floR*, conferindo resistência a florfenicol e genes *baeR* e *baeS*, que conferem resistência a novobiocina, antimicrobianos utilizados apenas no setor pecuário.

Com base nesses achados, podemos explicar o fato de todos os isolados recuperados do banco de dados público, com data de isolamento inferior ao ano 2000, não apresentarem resistência ao fenicol. Podemos sugerir que a presença do gene *floR* é proveniente de uma transmissão mais recente, pois foi observada em diferentes tipos de hospedeiros, inclusive de origem humana. Esses dados reforçam a ideia de que os animais de produção servem como reservatórios de ARGs (MATAMOROS et al., 2017) e podem ser disseminados para outros nichos através

da cadeia alimentar ou contato direto com as fezes desses animais (BARLAAM *et al.*, 2019).

A partir das análises de tipagem molecular, baseadas nos 7 genes *housekeeping* MLST e nos genes *fumC* e *fimH* CHTyping, foi possível observar 35 STs e 32 clonotipos CH entre todos os isolados analisados. Sendo a ST-10 e o CH11-25 os mais prevalentes dentre os 43 isolados. Este resultado é consistente com achados em outros estudos, onde 688 cepas de *E.coli*, portadoras do gene *mcr-1*, foram atribuídas a 198 tipos distintos de STs, sendo a ST-10 e ST-101 o tipo de sequência mais representativa na amostra, principalmente relacionada a hospedeiros do tipo animal e ambiental (SHEN *et al.*, 2020b). Já os isolados ST-10 representados no presente estudo, foram mais prevalentes em hospedeiros de origem humana e animal. Matamoros *et al.*, com o intuito de entender a dinâmica por trás da disseminação mundial do gene *mcr-1*, propôs um estudo para determinar a estrutura da população de *E. coli* utilizando 312 isolados proveniente de 69 estudos recuperados em bancos de dados públicos. Neste estudo, foi possível observar 112 tipos de STs únicas, sendo a ST-10 a mais comum na população em estudo, compreendendo 12,8% dos isolados originários da África, Ásia, Europa e América do Sul. Podemos observar um comportamento semelhante em nossos resultados, onde 11,63% dos 43 isolados apresentaram ST-10, recuperados da Europa, América do Norte e América do Sul. A ausência de isolados com esse perfil clonal em outros continentes pode ser explicada pelo pequeno grupo depositado no banco de dados SRA NCBI que atendessem aos critérios de inclusão deste estudo.

Na árvore de abrangência mínima, inferida com as sequências dos 7 genes *housekeeping*, disponível na **Figura 5**, podemos observar isolados ST-10 de origem humana e animal formando um pequeno *cluster* com poucos eventos de mutações entre si. Isso sugere uma disseminação clonal em cepas representadas por esta ST. Sabemos que cepas ST-10, pertencentes ao complexo clonal CC10, são globalmente distribuídas e associadas a uma ampla variedade de ambientes (MACIUCA *et al.*, 2019; REID *et al.*, 2017; REID; DE MAERE; DJORDJEVIC, 2019). Embora isolados de *E. coli* pertencentes a este complexo, sejam atribuídos ao grupo de bactérias comensais presentes no intestino de suínos e outros animais de

produção, cepas com esse perfil clonal também são conhecidas pela sua promiscuidade entre os hospedeiros e pela sua capacidade de carregar genes de resistência e causando infecções intestinais (REID; DE MAERE; DJORDJEVIC, 2019). Essa característica foi observada no isolado RS-BR 651M05549 ST-10, recuperado de suínos em fazendas do Rio Grande do Sul que além do gene *mcr-1*, também apresentou gene de resistência *bla*_{CTX-M-8}. Isolados expressando beta-lactamases de espectro estendido ESBL do tipo *bla*_{CTX-M} já foram relacionados a cepas com o mesmo perfil clonal (FAM *et al.*, 2011; FOUNOU *et al.*, 2018; MANGES; JOHNSON, 2012; MATAMOROS *et al.*, 2017; OTEO *et al.*, 2009) fato que pode ser atribuído à capacidade de isolados ST-10 em adquirir genes ARGs, além da sua alta prevalência em humanos e animais de produção (MATAMOROS *et al.*, 2017). Resultados semelhantes foram observados na árvore filogenética baseada nos 7 genes *housekeeping* e representada na **Figura 4**, e na árvore filogenética baseada em SNPs da **Figura 8**, sugerindo uma disseminação clonal em isolados com esta sequência tipo (SHAPIRO, 2016).

Dois dos isolados RS-BR de origem suína, 649M05549 e 914M05549, apresentaram tipo de sequência ST-101 e complexo clonal 101, pelo método de tipagem MLST. *E. coli* do tipo ST-101 é considerado um clone patogênico, que foi recentemente associado a infecções do trato urinário e da corrente sanguínea, em humanos, no Brasil e na China (FERNANDES *et al.*, 2016b; ZHONG; LIU; ZHENG, 2019b), em aves migratórias, na Tailândia (AEKSIRI *et al.*, 2019), em amostra ambiental, incluindo matadouros, na China (SHEN *et al.*, 2020a) e em produtos alimentícios de origem animal, no Egito (SADEK *et al.*, 2020).

Essa característica promíscua em isolados deste tipo pode explicar a distribuição observada na árvore filogenética de STs da **Figura 4**. Onde os isolados 914M05549 e 649M05549 foram relacionados com o isolado SRR10392137, de mesma ST e de origem alimentícia, representando um grupo com maior distância genética em relação as outras STs. Um resultado distinto foi observado na metodologia CHTyping, onde as duas cepas foram separadas em dois clonotipos CH diferentes: 649M05549 CH41-158 e 914M05549 CH41-23. Na árvore filogenética de clonotipos CH, representada na **Figura 6**, podemos perceber que esses dois

isolados se relacionam com os isolados vizinhos de forma distinta da árvore de STs da **Figura 4**. O isolado 649M05549 se relacionou com o isolado ERR1010284 e o isolado 914M05549 com o isolado ERR1010247, ambos de origem animal e STs distintas. Essa diferença pode ser relacionada ao número de genes utilizados na tipagem e a variabilidade das sequências, embora o método CH Typing utilize apenas 2 genes, *fumC* e *fim H*, para a distinção clonal, essa metodologia tende a discriminar as cepas de mesma ST em sub-STs, pois o gene *fumC* tem uma variabilidade maior em relação aos outros 6 genes *housekeeping* utilizados na tipagem MLST (WIRTH *et al.*, 2006). Além disso, resultados semelhantes foram observados por Roer *et al.*, onde a tipagem de 243 isolados de *E. coli* resistentes a cefalosporina apresentou 18 STs únicas, pelo método MLST e 70 clonotipos CH, pelo método CHTyping (ROER *et al.*, 2018b).

Na árvore de SNPs, quando analisamos a diversidade no genoma total, percebemos a tendência de agrupamento entre os isolados 649M05549, 914M05549, SRR10392137 e DRR138478, todos de complexo clonal CC 101. Neste mesmo *cluster*, também observamos uma sub-árvore composta por isolados pertencentes ao CC 156, sendo este complexo já descrito em isolados de *E. coli* produtora de *bla*_{NDM-5}, provenientes de animais de produção na China (TANG *et al.*, 2019). Embora os isolados pertencentes a esse *cluster* sejam, na sua maioria, isolados de origem suína, não podemos afirmar que existe uma dispersão clonal entre as cepas de acordo com o tipo de hospedeiro, pois para a metodologia utilizada, apenas cepas que apresentam um número de variações SNPs ≤ 10 , são consideradas clonalmente relacionadas (DEKKER; FRANK, 2016; ROER *et al.*, 2017; SCHÜRCH *et al.*, 2018). Contudo, podemos considerar que isolados pertencentes a estes complexos clonais podem estar sendo disseminados entre os diferentes hospedeiros (humanos, suínos e de origem alimentícia) através da cadeia alimentar (VERRAES *et al.*, 2013).

Com base nesses resultados, podemos dizer que a recombinação tem um papel importante na disseminação de cepas deste tipo, garantindo variabilidade e promiscuidade entre diferentes hospedeiros, destacando-se assim a importância dos MGEs — profagos, ilhas genômicas, transposons e plasmídeos codificadores de

genes de virulência, genes de resistência e genes de DNA metil transferase (Mtase) – na disseminação clonal (ASHCROFT *et al.*, 2020). Observamos também, que estes isolados foram mais bem separados nesta metodologia em relação as outras metodologias utilizadas (MLST e CH Typing), indicando um genoma central conservado em cepas deste tipo. Isso reforça a ideia de que isolados deste tipo podem se disseminar de forma clonal, comumente dentro de um mesmo complexo clonal e mesma ST, mas também através de recombinação por MGEs (ASHCROFT *et al.*, 2020). Sabendo que a ST-101 e a ST-156 representam um dos principais clones emergentes de *E.coli* associada ao transporte de genes *bla*_{NDM-1} e *bla*_{NDM-5} causando resistência a carbapenem (ZHONG; LIU; ZHENG, 2019a; TANG *et al.*, 2019), resultados deste tipo servem como alerta do potencial risco de contaminação entre diferentes nichos ecológicos, incluindo a disseminação através da cadeia alimentar de cepas com esse perfil clonal.

Os isolados RS-BR provenientes de frangos, apresentaram dois tipos de STs únicas: 215M02423 ST-58 e complexo clonal CC155; 249M02423 e 279M02423, ambas de ST-2491 e complexo clonal desconhecido. Diferentemente da ST-2491, que não é bem documentada na literatura, a ST-58 já foi bem descrita e tem uma importância clínica e um perfil clonal que deve ser destacado. J. McKinnon *et al.*, identificaram um isolado de *E.coli* ST-58 resistente a múltiplas drogas (MDR), recuperado de urina e sangue, de um paciente com infecção do trato urinário e uroseps, na Austrália (MCKINNON; ROY CHOWDHURY; DJORDJEVIC, 2018). Além disso, cepas ST-58 foram isoladas de vegetais e solo na Tunísia (BEN SAID *et al.*, 2015), podendo ser um potencial transmissor de genes ARGs através da cadeia alimentar. No Brasil, isolados com mesmo perfil clonal foram relatados em humanos, aves, animais em extinção e em ecossistema de mangue poluído (BORGES; TARLTON; RILEY, 2019; DE CARVALHO *et al.*, 2020; FUENTES-CASTILLO *et al.*, 2020; SACRAMENTO *et al.*, 2018b). Como podemos observar, assim como a ST-10 e ST-101, isolada de suínos, a ST-58 CC 155 também tem um perfil zoonótico, podendo migrar entre diferentes nichos e transportar outros genes de resistência, pois cepas com esse perfil clonal compartilham padrões de resistência antimicrobiana idênticos em populações de animais e humanos (DE CARVALHO *et*

al., 2020). Tais evidências podem explicar significativamente o estabelecimento bem-sucedido dessa linhagem de forma internacional (BORGES; TARLTON; RILEY, 2019).

Em todas as árvores filogenéticas inferidas neste trabalho, foi possível observar o isolado 215M02423 ST-58 intimamente relacionado com o isolado SRR10521259 ST-683, positivo para o gene *mcr-1*, recuperado de água costeira no Brasil e resistente a várias classes de antimicrobianos, representado na **Tabela 3**. Na árvore de STs da **Figura 4**, além do isolado ambiental SRR10521259, o isolado 215M02423 se relaciona com outros de diferentes hospedeiros e diferentes STs. Esse comportamento também foi observado na rede de abrangência mínima de STs, na **Figura 5**, onde isolados de origem animal, ambiental e humana se relacionaram com poucos eventos de mutações. Com base nesses dados, podemos supor uma origem ou ancestral comum cuja relação evolutiva é mediada por recombinação entre esses isolados. Sabemos que a diversidade genotípica também é refletida pelo genoma acessório das cepas e a maior parte da divergência entre pares de genomas, dentro de um grupo evolucionário, se deve ao acúmulo de transferência recombinante de genomas divergentes, desde seu último ancestral comum (DIXIT *et al.*, 2015).

Ao observar os nossos resultados de maneira geral, levando em consideração os três métodos de tipagem utilizados: MLST, CHTyping e SNPs, podemos destacar que as metodologias foram concordantes entre si. As metodologias baseadas em sequências de apenas alguns genes, apresentam resultados discriminatórios e satisfatórios na diferenciação clonal e tipagem de cepas. Contudo, para uma análise mais robusta, envolvendo genes do genoma central e também do genoma acessório, a tipagem baseada em SNPs a partir de dados de sequenciamento WGS, apresentam resultados mais completos sobre a distribuição dessas cepas, não apenas referente ao seu perfil clonal, mas também referente a contribuição dos eventos de recombinação na população.

No modelo reticulado, para observação de microevolução baseado em sequências dos genes *fumC* e *fimH* na **Figura 6**, os isolados foram distribuídos entre os grupos centrais da rede com poucos eventos de mutação entre eles, indicando

que a tipagem por CH Typing é capaz de distinguir variações mínimas nas sequências, mesmo utilizando um número menor de genes, em relação aos 7 genes *housekeeping* utilizados para tipagem MLST. Portanto, a tipagem CH pode ser usada para estudar a diversidade clonal sub-ST ou como teste de triagem antes da seleção para sequenciamento WGS de cepas de interesse. É uma ferramenta altamente adequada que pode atuar como uma alternativa rápida para a vigilância MLST convencional e para a detecção de surtos.

Por outro lado, a árvore filogenética de SNPs se comportou de maneira um pouco diferente, onde a distribuição de isolados com diferentes tipos de hospedeiros, diferentes STs, diferentes clonotipos CH e diferentes complexos clonais se distribuíram na extensão da árvore. Isso ocorre porque, embora o número de SNVs entre o genoma dos isolados e o genoma de referência utilizado para a chamada de variantes, seja um número considerável (dados disponíveis na Tabela 3 do apêndice 9.3), a maioria do genoma central é conservado. Jain et al., mostraram que a conservação no cromossomo bacteriano é representada com uma identidade de nucleotídeo média ANI $\geq 90\%$ em isolados da mesma espécie, mas isso não garante um conjunto fixo de características fenotípicas (REDONDO-SALVO *et al.*, 2020b). Sabemos que bactérias mantêm a identidade genômica de uma espécie através de recombinação preferencial com genomas intimamente relacionados, garantindo assim a diferenciação entre as espécies (VARGHESE *et al.*, 2015), mas será que esse comportamento também é observado em plasmídeos? Redondo-Salvo et al., demonstraram que os plasmídeos também mantêm sua identidade por meio de recombinação preferencial (incompatibilidade de plasmídeos), mas apresentam uma diferença fundamental quando comparados ao DNA cromossomal das bactérias, pois são capazes de permear a barreira entre espécies / gêneros de hospedeiros diferentes. Essa característica foi observada na troca de plasmídeos entre bactérias da família *Enterobacteriaceae*, sendo em uma proporção maior do que em outras famílias de bactérias. Consequentemente, espécies como *Escherichia coli*, *Salmonella* e *Klebsiella* são capazes de trocar plasmídeos com a maioria das bactérias clinicamente relevantes da ordem *Enterobacterales*. Esse dado ressalta a importância de estudos epidemiológicos sobre a disseminação de

genes que conferem resistência a antimicrobianos utilizados como último recurso no tratamento de infecções.

Em nenhuma das metodologias utilizadas foi possível relacionar o tipo de sequência ST, clonotipos CH ou complexo clonal com um tipo de hospedeiro específico. Também não foi possível relacionar o perfil de resistência ao tipo de hospedeiro, mas observamos que isolados mais antigos apresentavam resistência a um número menor de antimicrobianos, como observado no isolado 651M05549 representado na **Figura 8**, que estava intimamente relacionado com isolados mais antigos e, na maioria das vezes, resistentes apenas a classe de macrolídeos, antimicrobiano este muito utilizado no tratamento de infecções comuns em vários animais de produção (PYÖRÄLÄ *et al.*, 2014). Esse tipo de resultado destaca a influência dos antimicrobianos atuando como pressão seletiva na disseminação de ARGs e clones resistentes, como foi observado na história evolutiva de isolados MRSA (*Staphylococcus aureus* resistente à meticilina). O MRSA foi observado pela primeira vez em 1960, um ano após a introdução dos β -lactâmicos de segunda geração, como a meticilina, na prática clínica. No entanto, a reconstrução filogenética mostrou que o MRSA realmente surgiu na década de 1940 via HGT do SCC *mec* (*chromosome staphylococcal cassette mec*), como consequência do uso inicial em massa de penicilina (HARKINS *et al.*, 2017). Isso sugere que além da pressão seletiva causada pelo uso dos antimicrobianos, a presença de elementos que conferem resistência, como os ARGs transportados por MGEs, possuem um papel importante na expansão clonal e são capazes de determinar quais clones serão estabelecidos em uma população (BAKER *et al.*, 2018).

Alguns ARGs móveis têm desempenhado esse papel em vários organismos e clones, por exemplo, o gene *bla*_{CTX-M-15} contribuiu na expansão clonal de *Escherichia coli* ST-131 e vários clones de *Klebsiella pneumoniae* (CG14 / 15, ST-101) (MATHERS; PEIRANO; PITOUT, 2015; WYRES; HOLT, 2016). Outra observação recorrente e muito comum é o acúmulo de mecanismos adicionais de resistência em um clone já estabelecido, como a resistência à fluoroquinolona em EMRSA-15 (*Epidemic methicillin-resistant Staphylococcus aureus 15*) (HOLDEN *et al.*, 2013). Este fenômeno é provavelmente impulsionado pelo uso crescente de

antimicrobianos para combater infecções, acompanhado de um aumento do tamanho da população efetiva do clone bem-sucedido (MATHERS; PEIRANO; PITOUT, 2015; WYRES; HOLT, 2016). Esse comportamento é particularmente comum em organismos que podem acumular múltiplos ARGs através de HGT, principalmente dentro da família *Enterobacteriaceae*, como discutido anteriormente por Redondo et. al. Dados deste tipo geram um alerta na disseminação de genes *mcr-1* em diferentes hospedeiros, pois além de conferir resistência mediada por plasmídeos, já foram observados juntamente com outros genes de resistência em um mesmo isolado (MACIUCA *et al.*, 2019; SACRAMENTO *et al.*, 2018a; WANG *et al.*, 2018a).

Isolados positivos para o gene *mcr-1* distribuídos ao longo da árvore filogenética e relacionados com isolados mais antigos indicam uma disseminação não clonal, onde a pressão seletiva e eventos recentes de recombinação estão agindo sobre a população. Um número maior de variações ou SNPs é um forte indicativo de que estão sob pressão seletiva. Com o passar do tempo, o número de SNPs tende a estabilizar e apenas aquelas mutações que conferem características essenciais para a bactéria permanecem no genoma (DIDELLOT; MAIDEN, 2010). Sendo assim, isolados com pouca diferença de SNPs são considerados clonais. Isso foi observado com os isolados de frangos 249M02423 e 279M02423, ambos ST-2491, que apresentaram um número de SNPs ≤ 10 no genoma central e, portanto, foram consideradas clonalmente relacionadas (SCHÜRCH *et al.*, 2018). A clonalidade dos dois isolados pode ser observada em todas as árvores filogenéticas inferidas no presente estudo.

Sabendo que a emergência e disseminação da resistência antimicrobiana ocorre quando clones bem-sucedidos emergem em resposta ao uso de antimicrobianos locais e, em seguida, expandem em disseminação pandêmica (BAKER *et al.*, 2018; BAYM *et al.*, 2016), podemos dizer que a exposição sustentada dos animais de produção aos antimicrobianos ao longo do tempo, pode seguir a mesma trajetória evolutiva para o gene *mcr-1*, como observado na árvore de SNPs da **Figura 8**.

Com base nos dados apresentados acima, e reforçando a ideia de que a

permanência de características em uma população depende das forças evolutivas que atuam sobre ela (DIDELLOT; MAIDEN, 2010), afirmamos que a utilização de antibióticos com importância clínica não devem ser utilizados em animais de produção, tendo como exemplo a lista de antibióticos liberados para esta prática, redigida pelo Food and Drug Administration FDA.

Um estudo investigativo observou as mudanças subsequentes na prevalência do gene *mcr-1* em isolados de origem animal, humana, alimentícia ambiental, após a proibição do uso da colistina na produção animal, em 2017, na China. Dados de 2017 e 2018 foram comparados com dados de 2016, antes da proibição do uso de colistina. Shen et al., observaram a diminuição da prevalência do gene *mcr-1*, sendo 45% das amostras em 2016, comparadas a 19% das amostras em 2018. Uma diminuição semelhante ocorreu em amostras recuperadas de fazendas, onde 19% das amostras de 2016 foram comparadas com 5%, em 2018. Além disso, SNPs observados em 63 genomas centrais e 30 genes acessórios também foram significativamente diferentes após a proibição. Esses dados são consistentes a adaptação rápida em resposta às mudanças nas pressões seletivas, principalmente a exercida pelo uso da colistina. Os pesquisadores também perceberam que isolados recuperados de pacientes com infecções e internados em hospitais eram geneticamente distintos dos isolados de outras fontes e foram menos afetados pela proibição do uso do antimicrobiano. Isso pode ser explicado pela grande proporção de genes *mcr-1*, em isolados clínicos, estar associada a plasmídeos do tipo IncI2 e IncHI2. A proporção de IncI2 foi significativamente maior após a intervenção, comparada a outros tipos de plasmídeos que comumente transportam o gene *mcr-1*. Esse resultado sugere que, talvez, plasmídeos do tipo IncI2 sejam mais estáveis na ausência de pressões seletivas exercidas pelo uso de colistina. No entanto, descobriram que a proporção de *mcr-1* localizado no cromossomo diminuiu pela metade após a proibição, sendo 4,8% em 2016 comparado a 2,3% em 2018 (SHEN et al., 2020b).

Estudos deste tipo encorajam a criação e prática de políticas públicas que, não apenas proíbam a utilização de antimicrobianos com importância clínica no setor pecuário como promotor de crescimento, mas também fiscalize a compra,

venda e distribuição dos mesmos na população.

7. CONCLUSÃO

- Os seis isolados RS-BR apresentaram quatro sequências tipo ST distintas (ST-10, ST-101, ST-58 e ST-2491) e cinco clonotipos CH distintos (CH 4-32, CH 11-25, CH 11-0, CH 41-158 e CH 41-27).
- Os isolados RS-BR apresentaram genes de resistência para as seguintes classes de antimicrobianos: tetraciclinas, trimetoprim, fenicol, macrolídeos, quinolonas, aminoglicosídeos, sulfonamidas, colistina e beta-lactâmicos. Os genes relacionados a fatores de virulência mais observados foram: *astA* (heat-stable enterotoxin 1); *hly* (hemolysin F); *lpfA* (fimbrial major protein); *iroN* (iron transport); *iss* (increased serum survival); *gad* (glutamate decarboxylase alpha); *traT* (complement resistance protein) e *cma* (colicin).
- O gene *mcr-1*, presente nos isolados RS-BR, estava sendo carregado em plasmídeos do tipo IncX4.
- A comparação filogenética utilizando a ferramenta CHTyper apresentou uma tendência maior a clusterização em relação a metodologia MLST e foi capaz de distinguir isolados de mesma ST em clonotipos diferentes.
- A árvore filogenética baseada em SNPs apresentou uma disseminação não clonal entre os isolados.
- Através da genômica comparativa, observou-se que *Escherichia coli* produzindo o gene *mcr-1* é altamente diversa e se apresentou em diferentes sequências tipo ST, clonotipos CH e hospedeiros. Não sendo possível relacionar a ST, o CH ou complexo clonal com um tipo de hospedeiro específico em nenhuma das metodologias de tipagem utilizadas. Também não foi possível relacionar o perfil de resistência ao tipo de hospedeiro, mas observamos que isolados mais antigos apresentavam resistência a um número menor de antimicrobianos, geralmente resistentes a macrolídeos. Esse tipo de resultado sugere a influência dos antimicrobianos atuando como pressão seletiva na disseminação de ARGs.
- As árvores de abrangência mínima foram concordantes às árvores

filogenéticas inferidas utilizando as metodologias MLST e CH Typing, apresentando poucos eventos de mutação entre isolados provenientes de diferentes hospedeiros, sugerindo a propagação da resistência em diferentes ambientes.

8. REFERÊNCIAS

ABUOUN, M. *et al.* **mcr-1 and mcr-2 variant genes identified in *Moraxella* species isolated from pigs in Great Britain from 2014 to 2015.** *Journal of Antimicrobial Chemotherapy*, [S. l.], v. 72, n. 10, p. 2745–2749, 2017. Disponível em: <https://doi.org/10.1093/jac/dkx286>. Acesso em: 29 set. 2020.

AEKSIRI, N. *et al.* **First Detection and Genomic Insight into mcr-1 Encoding Plasmid-Mediated Colistin-Resistance Gene in *Escherichia coli* ST101 Isolated from the Migratory Bird Species *Hirundo rustica* in Thailand.** *Microbial Drug Resistance*, [S. l.], v. 25, n. 10, p. 1437–1442, 2019. Disponível em: <https://doi.org/10.1089/mdr.2019.0020>. Acesso em: 5 out. 2020.

AIRES, C. A. M. *et al.* **Emergence of the plasmid-mediated mcr-1 gene in clinical KPC-2-producing *Klebsiella pneumoniae* sequence type 392 in Brazil.** *Antimicrobial Agents and Chemotherapy*, [S. l.], v. 61, n. 7, 2017. Disponível em: <https://doi.org/10.1128/AAC.00317-17>. Acesso em: 29 set. 2020.

ASHCROFT, M. M. *et al.* **Strain and lineage-level methylome heterogeneity in the multi-drug resistant pathogenic *Escherichia coli* ST101 clone.** *bioRxiv*, [S. l.], p. 2020.06.07.138552, 2020. Disponível em: <https://doi.org/10.1101/2020.06.07.138552>

BAKER, S. *et al.* **Genomic insights into the emergence and spread of antimicrobial-resistant bacterial pathogens.** [S. l.]: American Association for the Advancement of Science, 2018. Disponível em: <https://doi.org/10.1126/science.aar3777>. Acesso em: 29 out. 2020.

BARLAAM, A. *et al.* **Global Emergence of Colistin-Resistant *Escherichia coli* in Food Chains and Associated Food Safety Implications: A Review.** *Journal of Food Protection*, [S. l.], v. 82, n. 8, p. 1440–1448, 2019. Disponível em: <https://doi.org/10.4315/0362-028X.JFP-19-116>

BAYM, M. *et al.* **Spatiotemporal microbial evolution on antibiotic landscapes.** *Science*, [S. l.], v. 353, n. 6304, p. 1147–1151, 2016. Disponível em: <https://doi.org/10.1126/science.aag0822>. Acesso em: 29 out. 2020.

BEN SAID, L. *et al.* **Detection of extended-spectrum beta-lactamase (ESBL)-producing *Enterobacteriaceae* in vegetables, soil and water of the farm environment in Tunisia.** *International Journal of Food Microbiology*, [S. l.], v. 203, p. 86–92, 2015. Disponível em: <https://doi.org/10.1016/j.ijfoodmicro.2015.02.023>. Acesso em: 5 out. 2020.

BORGES, C. A.; TARLTON, N. J.; RILEY, L. W. ***Escherichia coli* from Commercial Broiler and Backyard Chickens Share Sequence Types, Antimicrobial Resistance Profiles, and Resistance Genes with Human Extraintestinal Pathogenic *Escherichia coli*.** *Foodborne Pathogens and Disease*, [S. l.], v. 16, n. 12, p. 813–822, 2019. Disponível em: <https://doi.org/10.1089/fpd.2019.2680>. Acesso em: 11 set. 2020.

BOROWIAK, M. *et al.* **Identification of a novel transposon-associated phosphoethanolamine transferase gene, mcr-5, conferring colistin resistance in d-tartrate fermenting *Salmonella enterica* subsp. *enterica* serovar Paratyphi B.** *Journal of Antimicrobial Chemotherapy*, [S. l.], v. 72, n. 12, p. 3317–3324, 2017. Disponível em: <https://doi.org/10.1093/jac/dkx327>. Acesso em: 29 set. 2020.

BRODRICK, H. J. *et al.* **Longitudinal genomic surveillance of multidrug-resistant *Escherichia coli* carriage in a long-term care facility in the United Kingdom.** *Genome Medicine*, [S. l.], v. 9, n. 1, p. 70, 2017. Disponível em: <https://doi.org/10.1186/s13073-017-0457-6>. Acesso em: 22 jul. 2020.

BROGAN, D. M.; MOSSIALOS, E. **A critical analysis of the review on antimicrobial resistance report and the infectious disease financing facility.** [S. l.]: BioMed Central Ltd., 2016. Disponível em: <https://doi.org/10.1186/s12992-016-0147-y>. Acesso em: 2 out. 2020.

CARATTOLI, A. *et al.* **Novel plasmid-mediated colistin resistance *mcr-4* gene in *Salmonella* and *Escherichia coli*, Italy 2013, Spain and Belgium, 2015 to 2016.** *Eurosurveillance*, [S. l.], v. 22, n. 31, 2017. Disponível em: <https://doi.org/10.2807/1560-7917.ES.2017.22.31.30589>. Acesso em: 29 set. 2020.

CARRIÇO, J. A. *et al.* **A primer on microbial bioinformatics for nonbioinformaticians.** *Clinical Microbiology and Infection*, [S. l.], v. 24, n. 4, p. 342–349, 2018. Disponível em: <https://doi.org/10.1016/j.cmi.2017.12.015>

CARROLL, L. M. *et al.* **Identification of novel mobilized colistin resistance gene *mcr-9* in a multidrug-resistant, colistin-susceptible *salmonella enterica* serotype typhimurium isolate.** *mBio*, [S. l.], v. 10, n. 3, 2019. Disponível em: <https://doi.org/10.1128/mBio.00853-19>. Acesso em: 29 set. 2020.

CASELLI, E. *et al.* **Spread of *mcr-1*–driven colistin resistance on hospital surfaces, Italy.** [S. l.]: Centers for Disease Control and Prevention (CDC), 2018. Disponível em: <https://doi.org/10.3201/eid2409.171386>. Acesso em: 29 set. 2020.

CASTANON, J. I. R. **History of the use of antibiotic as growth promoters in European poultry feeds.** [S. l.]: *Poult Sci*, 2007. Disponível em: <https://doi.org/10.3382/ps.2007-00249>. Acesso em: 2 out. 2020.

CAVALLI-SFORZA, L. L.; EDWARDS, A. W. **Phylogenetic analysis. Models and estimation procedures.** *American Journal of Human Genetics*, [S. l.], v. 19, n. 3, p. 233, 1967. Disponível em: <https://doi.org/10.2307/2406616>. Acesso em: 27 set. 2020.

CHAUDHURI, R. R.; HENDERSON, I. R. **The evolution of the *Escherichia coli* phylogeny.** [S. l.]: *Infect Genet Evol*, 2012. Disponível em: <https://doi.org/10.1016/j.meegid.2012.01.005>. Acesso em: 29 set. 2020.

CLEMENTE, J. C. *et al.* **The Impact of the Gut Microbiota on Human Health: An Integrative View.** *Cell*, [S. l.], v. 148, n. 6, p. 1258–1270, 2012. Disponível em: <https://doi.org/10.1016/j.cell.2012.01.035>. Acesso em: 4 out. 2020.

CLERMONT, O.; GORDON, D.; DENAMUR, E. **Guide to the various phylogenetic classification schemes for *Escherichia coli* and the correspondence among schemes.** *Microbiology (United Kingdom)*, [S. l.], v. 161, n. 5, p. 980–988, 2015. Disponível em: <https://doi.org/10.1099/mic.0.000063>. Acesso em: 30 set. 2020.

CLERMONT, O. *et al.* **Characterization and rapid identification of phylogroup G in *Escherichia coli*, a lineage with high virulence and antibiotic resistance potential.** *Environmental Microbiology*, [S. l.], v. 21, n. 8, p. 3107–3117, 2019. Disponível em: <https://doi.org/10.1111/1462-2920.14713>. Acesso em: 20 jan. 2021.

CONCEIÇÃO-NETO, O. C. *et al.* **Detection of the plasmid-mediated mcr-1 gene in clinical KPC-2-producing Escherichia coli isolates in Brazil.** [S. l.]: Elsevier B.V., 2017. Disponível em: <https://doi.org/10.1016/j.ijantimicag.2017.05.003>. Acesso em: 29 set. 2020.

CROXEN, M. A.; FINLAY, B. B. **Molecular mechanisms of Escherichia coli pathogenicity.** [S. l.]: Nat Rev Microbiol, 2010. Disponível em: <https://doi.org/10.1038/nrmicro2265>. Acesso em: 28 set. 2020.

CVM; FDA; DS. **2018 Summary Report On Antimicrobials Sold or Distributed for Use in Food-Producing Animals.** [S. l.: s. n.].

DALMOLIN, T. V. *et al.* **Co-occurrence of mcr-1 and blaKPC-2 in a clinical isolate of Escherichia coli in Brazil.** [S. l.]: Oxford University Press, 2017. Disponível em: <https://doi.org/10.1093/jac/dkx142>. Acesso em: 29 set. 2020.

CULLY, M. **Antibiotics alter the gut microbiome and host health.** Nature. Milestones in Human Microbiota Research 15, S19, 2019.

DE CARVALHO, M. P. N. *et al.* **International clones of extended-spectrum β -lactamase (CTX-M)-producing Escherichia coli in peri-urban wild animals, Brazil.** Transboundary and Emerging Diseases, [S. l.], 2020. Disponível em: <https://doi.org/10.1111/tbed.13558>. Acesso em: 11 set. 2020.

DIDELLOT, X.; MAIDEN, M. C. J. **Impact of recombination on bacterial evolution.** [S. l.]: Europe PMC Funders, 2010. Disponível em: <https://doi.org/10.1016/j.tim.2010.04.002>. Acesso em: 9 set. 2020.

DEKKER, J. P.; FRANK, K. M. **Next-generation epidemiology: Using real-time core genome multilocus sequence typing to support infection control policy.** Journal of Clinical Microbiology, [S. l.], v. 54, n. 12, p. 2850–2853, 2016. Disponível em: <https://doi.org/10.1128/JCM.01714-16>. Acesso em: 5 nov. 2020.

DIXIT, P. D. *et al.* **Recombinant transfer in the basic genome of Escherichia coli.** Proceedings of the National Academy of Sciences of the United States of America, [S. l.], v. 112, n. 29, p. 9070–9075, 2015. Disponível em: <https://doi.org/10.1073/pnas.1510839112>. Acesso em: 9 set. 2020.

DOBRZANSKA, D. A. *et al.* **Preventive antibiotic treatment of calves: emergence of dysbiosis causing propagation of obese state-associated and mobile multidrug resistance-carrying bacteria.** Microbial Biotechnology, [S. l.], v. 13, n. 3, p. 669–682, 2020. Disponível em: <https://doi.org/10.1111/1751-7915.13496>. Acesso em: 4 out. 2020.

FAM, N. *et al.* **CTX-M-15-Producing Escherichia coli Clinical Isolates in Cairo (Egypt), Including Isolates of Clonal Complex ST10 and Clones ST131, ST73, and ST405 in Both Community and Hospital Settings.** Microbial Drug Resistance, [S. l.], v. 17, n. 1, p. 67–73, 2011. Disponível em: <https://doi.org/10.1089/mdr.2010.0063>. Acesso em: 5 out. 2020.

FEIL, E. J. *et al.* **eBURST: Inferring Patterns of Evolutionary Descent among Clusters of Related Bacterial Genotypes from Multilocus Sequence Typing Data.** Journal of Bacteriology, [S. l.], v. 186, n. 5, p. 1518–1530, 2004. Disponível em: <https://doi.org/10.1128/JB.186.5.1518-1530.2004>. Acesso em: 23 set. 2020.

FELSENSTEIN, J. **Evolutionary trees from DNA sequences: A maximum likelihood**

approach. *Journal of Molecular Evolution*, [S. l.], v. 17, n. 6, p. 368–376, 1981. Disponível em: <https://doi.org/10.1007/BF01734359>. Acesso em: 2 out. 2020.

FELSENSTEIN, J. **Inferring Phylogenies.** [S. l.: s. n.].

FERNANDES, M. R. *et al.* **Silent dissemination of colistin-resistant *Escherichia coli* in South America could contribute to the global spread of the *mcr-1* gene.** *Eurosurveillance*, [S. l.], v. 21, n. 17, 2016 a. Disponível em: <https://doi.org/10.2807/1560-7917.ES.2016.21.17.30214>. Acesso em: 29 set. 2020.

FERNANDES, M. R. *et al.* **First report of the globally disseminated IncX4 plasmid carrying the *mcr-1* gene in a colistin-resistant *Escherichia coli* sequence type 101 isolate from a human infection in Brazil.** *Antimicrobial Agents and Chemotherapy*, [S. l.], v. 60, n. 10, p. 6415–6417, 2016 b. Disponível em: <https://doi.org/10.1128/AAC.01325-16>. Acesso em: 29 set. 2020.

FITCH, W. M. **Toward defining the course of evolution: Minimum change for a specific tree topology.** *Systematic Biology*, [S. l.], v. 20, n. 4, p. 406–416, 1971. Disponível em: <https://doi.org/10.1093/sysbio/20.4.406>. Acesso em: 26 set. 2020.

FOUNOU, L. L. *et al.* **Draft genome sequence of an extended-spectrum β -lactamase (CTX-M-15)-producing *Escherichia coli* ST10 isolated from a nasal sample of an abattoir worker in Cameroon.** *Journal of Global Antimicrobial Resistance*, [S. l.], v. 14, p. 68–69, 2018. Disponível em: <https://doi.org/10.1016/j.jgar.2018.06.004>. Acesso em: 5 out. 2020.

FUENTES-CASTILLO, D. *et al.* **Genomic characterization of multidrug-resistant ESBL-producing *Escherichia coli* ST58 causing fatal colibacillosis in critically endangered Brazilian merganser (*Mergus octosetaceus*).** *Transboundary and Emerging Diseases*, [S. l.], p. tbed.13686, 2020. Disponível em: <https://doi.org/10.1111/tbed.13686>. Acesso em: 13 ago. 2020.

GARCÍA, V. *et al.* **Co-occurrence of *mcr-1*, *mcr-4* and *mcr-5* genes in multidrug-resistant ST10 Enterotoxigenic and Shiga toxin-producing *Escherichia coli* in Spain (2006-2017).** *International Journal of Antimicrobial Agents*, [S. l.], v. 52, n. 1, 2018. Disponível em: <https://doi.org/10.1016/j.ijantimicag.2018.03.022>

GINARD, M. *et al.* **Genome organization of *Pseudomonas stutzeri* and resulting taxonomic and evolutionary considerations.** *International Journal of Systematic Bacteriology*, [S. l.], v. 47, n. 1, p. 132–143, 1997. Disponível em: <https://doi.org/10.1099/00207713-47-1-132>. Acesso em: 29 set. 2020.

GROTHUESI, D.; TÜMMLER, B. **New approaches in genome analysis by pulsed-field gel electrophoresis: application to the analysis of *Pseudomonas* species.** *Molecular Microbiology*, [S. l.], v. 5, n. 11, p. 2763–2776, 1991. Disponível em: <https://doi.org/10.1111/j.1365-2958.1991.tb01985.x>. Acesso em: 30 set. 2020.

GUARDABASSI, L.; COURVALIN, P. **Modes of antimicrobial action and mechanisms of bacterial resistance.** *In: Antimicrobial Resistance in Bacteria of Animal Origin.* [S. l.: s. n.]. *E-book*. Disponível em: <https://doi.org/10.1128/9781555817534.ch1>

GUENTHER, S. *et al.* **Environmental emission of multiresistant *Escherichia coli* carrying the colistin resistance gene *mcr-1* from German swine farms.** *Journal of Antimicrobial Chemotherapy*, [S. l.], v. 72, n. 5, p. 1289–1292, 2017. Disponível em: <https://doi.org/10.1093/jac/dkw585>. Acesso em: 4 out. 2020.

HALL, A. V. **A joint phenetic and cladistic approach for systematics**. Biological Journal of the Linnean Society, [S. l.], v. 33, n. 4, p. 367–382, 1988. Disponível em: <https://doi.org/10.1111/j.1095-8312.1988.tb00450.x>. Acesso em: 27 set. 2020.

HARKINS, C. P. *et al.* **Methicillin-resistant Staphylococcus aureus emerged long before the introduction of methicillin into clinical practice**. Genome Biology, [S. l.], v. 18, n. 1, 2017. Disponível em: <https://doi.org/10.1186/s13059-017-1252-9>. Acesso em: 29 out. 2020.

HARTIGAN, J. A. **Minimum Mutation Fits to a Given Tree**. Biometrics, [S. l.], v. 29, n. 1, p. 53, 1973. Disponível em: <https://doi.org/10.2307/2529676>

HASEGAWA, M.; KISHINO, H.; YANO, T. **Dating of the human-ape splitting by a molecular clock of mitochondrial DNA**. Journal of Molecular Evolution, [S. l.], v. 22, n. 2, p. 160–174, 1985. Disponível em: <https://doi.org/10.1007/BF02101694>. Acesso em: 1 set. 2020.

HASMAN, H. *et al.* **Detection of mcr-1 encoding plasmid-mediated colistin-resistant Escherichia coli isolates from human bloodstream infection and imported chicken meat, Denmark 2015**. Eurosurveillance, [S. l.], v. 20, n. 49, p. 30085, 2015. Disponível em: <https://doi.org/10.2807/1560-7917.ES.2015.20.49.30085>. Acesso em: 4 out. 2020.

HERNÁNDEZ, M. *et al.* **Co-occurrence of colistin-resistance genes mcr-1 and mcr-3 among multidrug-resistant Escherichia coli isolated from cattle, Spain, September 2015**. Eurosurveillance, [S. l.], 2017. Disponível em: <https://doi.org/10.2807/1560-7917.ES.2017.22.31.30586>

HOELZER, K. *et al.* **Antimicrobial drug use in food-producing animals and associated human health risks: what, and how strong, is the evidence?** BMC Veterinary Research, [S. l.], v. 13, n. 1, p. 211, 2017. Disponível em: <https://doi.org/10.1186/s12917-017-1131-3>. Acesso em: 4 out. 2020.

HOLDEN, M. T. G. *et al.* **A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant Staphylococcus aureus pandemic**. Genome Research, [S. l.], v. 23, n. 4, p. 653–664, 2013. Disponível em: <https://doi.org/10.1101/gr.147710.112>. Acesso em: 29 out. 2020.

HUSON, D. H.; BRYANT, D. **Application of phylogenetic networks in evolutionary studies**. [S. l.]: Oxford Academic, 2006. Disponível em: <https://doi.org/10.1093/molbev/msj030>. Acesso em: 27 set. 2020.

IGWARAN, A.; IWERIEBOR, B. C.; OKOH, A. I. **Molecular characterization and antimicrobial resistance pattern of Escherichia coli recovered from wastewater treatment plants in Eastern Cape South Africa**. International Journal of Environmental Research and Public Health, [S. l.], v. 15, n. 6, 2018. Disponível em: <https://doi.org/10.3390/ijerph15061237>. Acesso em: 12 set. 2020.

INSTRUÇÃO NORMATIVA Nº 45, DE 22 DE NOVEMBRO DE 2016 - Imprensa Nacional. [S. l.], [s. d.]. Disponível em: https://www.in.gov.br/material/-/asset_publisher/Kujrw0TZC2Mb/content/id/22078290/do1-2016-11-30-instrucao-normativa-n-45-de-22-de-novembro-de-2016-22078259. Acesso em: 12 set. 2020.

JAVADI, M.; BOUZARI, S.; OLOOMI, M. **Horizontal Gene Transfer and the Diversity of Escherichia coli**. In: *Escherichia coli - Recent Advances on Physiology, Pathogenesis and*

Biotechnological Applications. [S. l.]: InTech, 2017. *E-book*. Disponível em: <https://doi.org/10.5772/intechopen.68307>. Acesso em: 28 set. 2020.

JOENSEN, K. G. *et al.* **Rapid and easy in silico serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data.** *Journal of Clinical Microbiology*, [S. l.], v. 53, n. 8, p. 2410–2426, 2015. Disponível em: <https://doi.org/10.1128/JCM.00008-15>. Acesso em: 31 ago. 2020.

KAPER, J. B.; NATARO, J. P.; MOBLEY, H. L. **Pathogenic *Escherichia coli*.** *Nature reviews. Microbiology*, [S. l.], 2004 a.

KENRAD NELSON, B. E.; MASTERS WILLIAMS, C. F. **Infectious Disease Epidemiology: Theory and Practice**, Second Edition. [S. l.], 2007. Disponível em: <https://doi.org/10.1093/aje/kwn045>

KIMURA, B. **Will the emergence of core genome MLST end the role of in silico MLST?** *Food Microbiology*, [S. l.], v. 75, p. 28–36, 2018. Disponível em: <https://doi.org/10.1016/j.fm.2017.09.003>. Acesso em: 30 set. 2020.

LEIGH, J. W.; BRYANT, D. POPART: **Full-feature software for haplotype network construction.** *Methods in Ecology and Evolution*, [S. l.], v. 6, n. 9, p. 1110–1116, 2015. Disponível em: <https://doi.org/10.1111/2041-210X.12410>. Acesso em: 31 ago. 2020.

LEIMBACH, A.; HACKER, J.; DOBRINDT, U. ***E. coli* as an all-rounder: The thin line between commensalism and pathogenicity.** *Current Topics in Microbiology and Immunology*, [S. l.], v. 358, p. 3–32, 2013. Disponível em: https://doi.org/10.1007/82_2012_303. Acesso em: 28 set. 2020.

LENTZ, S. A. *et al.* **Letter to the editor: *Escherichia coli* harbouring *mcr*-1 gene isolated from poultry not exposed to polymyxins in Brazil.** *Eurosurveillance*, [S. l.], v. 21, n. 26, p. 30267, 2016. Disponível em: <https://doi.org/10.2807/1560-7917.ES.2016.21.26.30267>. Acesso em: 13 set. 2020.

LI, H. *et al.* **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics*, [S. l.], v. 25, n. 16, p. 2078–2079, 2009. Disponível em: <https://doi.org/10.1093/bioinformatics/btp352>. Acesso em: 26 set. 2020.

LI, H.; DURBIN, R. **Fast and accurate long-read alignment with Burrows–Wheeler transform.** *Bioinformatics*, [S. l.], v. 26, n. 5, p. 589–595, 2010. Disponível em: <https://doi.org/10.1093/bioinformatics/btp698>. Acesso em: 26 set. 2020.

LIU, Y. Y. *et al.* **Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: A microbiological and molecular biological study.** *The Lancet Infectious Diseases*, [S. l.], v. 16, n. 2, p. 161–168, 2016 b. Disponível em: [https://doi.org/10.1016/S1473-3099\(15\)00424-7](https://doi.org/10.1016/S1473-3099(15)00424-7)

MACIUCA, I. E. *et al.* **Genetic Features of *mcr*-1 Mediated Colistin Resistance in CMY-2-Producing *Escherichia coli* From Romanian Poultry.** *Frontiers in Microbiology*, [S. l.], v. 10, 2019. Disponível em: <https://doi.org/10.3389/fmicb.2019.02267>. Acesso em: 23 jul. 2020.

MAESTRE-CARBALLA, L. *et al.* **Insights into the antibiotic resistance dissemination in a wastewater effluent microbiome: bacteria, viruses and vesicles matter.** *Environmental Microbiology*, [S. l.], v. 21, n. 12, p. 4582–4596, 2019. Disponível em: <https://doi.org/10.1111/1462-2920.14758>. Acesso em: 2 out. 2020.

MAIDEN, M. C. J. *et al.* **Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms.** Proceedings of the National Academy of Sciences of the United States of America, [*S. l.*], v. 95, n. 6, p. 3140–3145, 1998. Disponível em: <https://doi.org/10.1073/pnas.95.6.3140>. Acesso em: 30 set. 2020.

MANGES, A. R.; JOHNSON, J. R. **Food-Borne Origins of Escherichia coli Causing Extraintestinal Infections.** Clinical Infectious Diseases, [*S. l.*], v. 55, n. 5, p. 712–719, 2012. Disponível em: <https://doi.org/10.1093/cid/cis502>. Acesso em: 5 out. 2020.

MARTAK, D. *et al.* **Comparison of pulsed-field gel electrophoresis and whole-genome-sequencing-based typing confirms the accuracy of pulsed-field gel electrophoresis for the investigation of local Pseudomonas aeruginosa outbreaks.** Journal of Hospital Infection, [*S. l.*], v. 105, n. 4, p. 643–647, 2020. Disponível em: <https://doi.org/10.1016/j.jhin.2020.06.013>

MARTINS-SORENSEN, N. *et al.* **A novel plasmid-encoded mcr-4.3 gene in a colistin-resistant Acinetobacter baumannii clinical strain.** Journal of Antimicrobial Chemotherapy, [*S. l.*], v. 75, n. 1, p. 60–64, 2020. Disponível em: <https://doi.org/10.1093/jac/dkz413>. Acesso em: 29 set. 2020.

MATAMOROS, S. *et al.* **Global phylogenetic analysis of Escherichia coli and plasmids carrying the mcr-1 gene indicates bacterial diversity but plasmid restriction.** Scientific Reports, [*S. l.*], v. 7, n. 1, p. 1–9, 2017. Disponível em: <https://doi.org/10.1038/s41598-017-15539-7>. Acesso em: 12 ago. 2020.

MATHERS, A. J.; PEIRANO, G.; PITOUT, J. D. D. **The role of epidemic resistance plasmids and international high- risk clones in the spread of multidrug-resistant Enterobacteriaceae.** Clinical Microbiology Reviews, [*S. l.*], v. 28, n. 3, p. 565–591, 2015. Disponível em: <https://doi.org/10.1128/CMR.00116-14>. Acesso em: 29 out. 2020.

MCKINNON, J.; ROY CHOWDHURY, P.; DJORDJEVIC, S. P. **Genomic analysis of multidrug-resistant Escherichia coli ST58 causing urosepsis.** International Journal of Antimicrobial Agents, [*S. l.*], v. 52, n. 3, p. 430–435, 2018. Disponível em: <https://doi.org/10.1016/j.ijantimicag.2018.06.017>

MEDINI, D. *et al.* **The microbial pan-genome.** [*S. l.*]: Curr Opin Genet Dev, 2005. Disponível em: <https://doi.org/10.1016/j.gde.2005.09.006>. Acesso em: 29 set. 2020.

MONTE, D. F. *et al.* **Chicken meat as a reservoir of colistin-resistant Escherichia coli strains carrying mcr-1 genes in South America.** Antimicrobial Agents and Chemotherapy, [*S. l.*], v. 61, n. 5, 2017. Disponível em: <https://doi.org/10.1128/AAC.02718-16>. Acesso em: 29 set. 2020.

MORENO, L. Z. *et al.* **First report of mcr-1-harboring Salmonella enterica serovar Schwarzengrund isolated from poultry meat in Brazil.** Diagnostic Microbiology and Infectious Disease, [*S. l.*], v. 93, n. 4, p. 376–379, 2019. Disponível em: <https://doi.org/10.1016/j.diagmicrobio.2018.10.016>. Acesso em: 29 set. 2020.

NANG, S. C.; LI, J.; VELKOV, T. **The rise and spread of mcr plasmid-mediated polymyxin resistance.** [*S. l.*]: Taylor and Francis Ltd, 2019. Disponível em: <https://doi.org/10.1080/1040841X.2018.1492902>. Acesso em: 4 set. 2020.

OTEO, J. *et al.* **Extended-spectrum β -lactamase-producing Escherichia coli in Spain belong to a large variety of multilocus sequence typing types, including ST10 complex/A, ST23 complex/A and ST131/B2.** International Journal of Antimicrobial Agents, [*S. l.*], v. 34, n. 2, p. 173–

176, 2009. Disponível em: <https://doi.org/10.1016/j.ijantimicag.2009.03.006>. Acesso em: 5 out. 2020.

PALMEIRA, J. D. *et al.* **Draft genome of a ST443 mcr-1- and blaCTX-M-2-carrying Escherichia coli from cattle in Brazil.** [S. l.]: Elsevier Ltd, 2018. Disponível em: <https://doi.org/10.1016/j.jgar.2018.05.010>. Acesso em: 29 set. 2020.

PEARCE, M. E. *et al.* **Comparative analysis of core genome MLST and SNP typing within a European Salmonella serovar Enteritidis outbreak.** International Journal of Food Microbiology, [S. l.], v. 274, p. 1–11, 2018. Disponível em: <https://doi.org/10.1016/j.ijfoodmicro.2018.02.023>

PERRIN-GUYOMARD, A. *et al.* **Prevalence of mcr-1 in commensal Escherichia coli from French livestock, 2007 to 2014.** Eurosurveillance, [S. l.], v. 21, n. 6, 2016. Disponível em: <https://doi.org/10.2807/1560-7917.ES.2016.21.6.30135>. Acesso em: 4 out. 2020.

PYÖRÄLÄ, S. *et al.* **Macrolides and lincosamides in cattle and pigs: Use and development of antimicrobial resistance.** Veterinary Journal, [S. l.], v. 200, n. 2, p. 230–239, 2014. Disponível em: <https://doi.org/10.1016/j.tvjl.2014.02.028>

RANNALA, B.; YANG, Z. **Probability distribution of molecular evolutionary trees: A new method of phylogenetic inference.** Journal of Molecular Evolution, [S. l.], v. 43, n. 3, p. 304–311, 1996. Disponível em: <https://doi.org/10.1007/BF02338839>. Acesso em: 26 set. 2020.

RAVEN, K. E. *et al.* **Genomic surveillance of Escherichia coli in municipal wastewater treatment plants as an indicator of clinically relevant pathogens and their resistance genes.** Microbial genomics, [S. l.], v. 5, n. 5, 2019. Disponível em: <https://doi.org/10.1099/mgen.0.000267>. Acesso em: 16 jul. 2020.

REBELO, A. R. *et al.* **Multiplex PCR for detection of plasmid-mediated colistin resistance determinants, mcr-1, mcr-2, mcr-3, mcr-4 and mcr-5 for surveillance purposes.** Eurosurveillance, [S. l.], v. 23, n. 6, 2018. Disponível em: <https://doi.org/10.2807/1560-7917.ES.2018.23.6.17-00672>. Acesso em: 10 set. 2020.

REDONDO-SALVO, S. *et al.* **Pathways for horizontal gene transfer in bacteria revealed by a global map of their plasmids.** Nature Communications, [S. l.], v. 11, n. 1, p. 1–13, 2020 a. Disponível em: <https://doi.org/10.1038/s41467-020-17278-2>. Acesso em: 9 set. 2020.

REID, C. J. *et al.* **Porcine commensal escherichia coli: A reservoir for class 1 integrons associated with IS26.** Microbial Genomics, [S. l.], v. 3, n. 12, 2017. Disponível em: <https://doi.org/10.1099/mgen.0.000143>. Acesso em: 11 set. 2020.

REID, C. J.; DE MAERE, M. Z.; DJORDJEVIC, S. P. **Australian porcine clonal complex 10 (CC10) Escherichia coli belong to multiple sublineages of a highly diverse global CC10 phylogeny.** Microbial Genomics, [S. l.], v. 5, n. 3, 2019. Disponível em: <https://doi.org/10.1099/mgen.0.000225>. Acesso em: 11 set. 2020.

ROER, L. *et al.* **WGS-based surveillance of third-generation cephalosporin-resistant Escherichia coli from bloodstream infections in Denmark.** Journal of Antimicrobial Chemotherapy, [S. l.], v. 72, n. 7, p. 1922–1929, 2017. Disponível em: <https://doi.org/10.1093/jac/dkx092>. Acesso em: 5 nov. 2020.

ROER, L. *et al.* **CHTyper, a web tool for subtyping of extraintestinal pathogenic Escherichia coli based on the fumC and fimH alleles.** [S. l.]: American Society for Microbiology,

2018 b. Disponível em: <https://doi.org/10.1128/JCM.00063-18>

SABAT, A. J. *et al.* **Overview of molecular typing methods for outbreak detection and epidemiological surveillance.** [S. l.]: European Centre for Disease Prevention and Control (ECDC), 2013. Disponível em: <https://doi.org/10.2807/ese.18.04.20380-en>. Acesso em: 28 set. 2020.

SACRAMENTO, A. G. *et al.* **Genomic analysis of MCR-1 and CTX-M-8 co-producing *Escherichia coli* ST58 isolated from a polluted mangrove ecosystem in Brazil.** [S. l.]: Elsevier Ltd, 2018 a. Disponível em: <https://doi.org/10.1016/j.jgar.2018.10.024>. Acesso em: 13 ago. 2020.

SADEK, M. *et al.* **Draft genome sequence of an mcr-1/Incl2-carrying multidrug-resistant *Escherichia coli* B1:ST101 isolated from meat and meat products in Egypt.** [S. l.]: Elsevier Ltd, 2020. Disponível em: <https://doi.org/10.1016/j.jgar.2019.11.015>. Acesso em: 10 set. 2020.

SAIDENBERG, A. B. S. *et al.* **mcr-Positive *Escherichia coli* ST131- H 22 from Poultry in Brazil.** *Emerging Infectious Diseases*, [S. l.], v. 26, n. 8, p. 1951–1954, 2020. Disponível em: <https://doi.org/10.3201/eid2608.191724>. Acesso em: 24 jul. 2020.

SAITOU, N.; NEI, M. **The neighbor-joining method: a new method for reconstructing phylogenetic trees.** *Molecular Biology and Evolution*, [S. l.], v. 4, n. 4, p. 406–425, 1987. Disponível em: <https://doi.org/10.1093/oxfordjournals.molbev.a040454>. Acesso em: 27 set. 2020.

SCHÜRCH, A. C. *et al.* **Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene-based approaches.** [S. l.]: Elsevier B.V., 2018. Disponível em: <https://doi.org/10.1016/j.cmi.2017.12.016>. Acesso em: 22 set. 2020.

SELLERA, F. P. *et al.* ***Escherichia coli* carrying IncX4 plasmid-mediated mcr-1 and blaCTX-M genes in infected migratory Magellanic penguins (*Spheniscus magellanicus*).** [S. l.]: Oxford University Press, 2017. Disponível em: <https://doi.org/10.1093/jac/dkw543>. Acesso em: 29 set. 2020.

SHAPIRO, B. J. **How clonal are bacteria over time?.** [S. l.]: Elsevier Ltd, 2016. Disponível em: <https://doi.org/10.1016/j.mib.2016.03.013>

SHEN, C. *et al.* **Dynamics of mcr-1 prevalence and mcr-1-positive *Escherichia coli* after the cessation of colistin use as a feed additive for animals in China: a prospective cross-sectional and whole genome sequencing-based molecular epidemiological study.** *The Lancet Microbe*, [S. l.], v. 1, n. 1, p. e34–e43, 2020 a. Disponível em: [https://doi.org/10.1016/s2666-5247\(20\)30005-7](https://doi.org/10.1016/s2666-5247(20)30005-7). Acesso em: 10 set. 2020.

SUN, D. Pull in and Push Out: **Mechanisms of Horizontal Gene Transfer in Bacteria.** *Frontiers in Microbiology*, [S. l.], v. 9, n. SEP, p. 2154, 2018. Disponível em: <https://doi.org/10.3389/fmicb.2018.02154>. Acesso em: 28 set. 2020.

SUN, J. *et al.* **Towards Understanding MCR-like Colistin Resistance.** [S. l.]: Elsevier Ltd, 2018. Disponível em: <https://doi.org/10.1016/j.tim.2018.02.006>

TANG, B. *et al.* **Characterization of an NDM-5 carbapenemase-producing *Escherichia coli* ST156 isolate from a poultry farm in Zhejiang, China.** *BMC Microbiology*, [S. l.], v. 19, n. 1, p. 82, 2019. Disponível em: <https://doi.org/10.1186/s12866-019-1454-2>. Acesso em: 5 nov. 2020.

TCHESNOKOVA, V. *et al.* **Predictive Diagnostics for Escherichia coli Infections Based on the Clonal Association of Antimicrobial Resistance and Clinical Outcome.** [S. l.], 2013. Disponível em: <https://doi.org/10.1128/JCM.00984-13>. Acesso em: 23 set. 2020.

TENAILLON, O. *et al.* **The population genetics of commensal Escherichia coli.** [S. l.]: Nat Rev Microbiol, 2010. Disponível em: <https://doi.org/10.1038/nrmicro2298>. Acesso em: 2 out. 2020.

THORPE, H. A. *et al.* **Comparative analyses of selection operating on nontranslated intergenic regions of diverse bacterial species.** Genetics, [S. l.], v. 206, n. 1, p. 363–376, 2017. Disponível em: <https://doi.org/10.1534/genetics.116.195784>. Acesso em: 26 set. 2020.

TOUCHON, M. *et al.* **Organised Genome Dynamics in the Escherichia coli Species Results in Highly Diverse Adaptive Paths.** PLoS Genetics, [S. l.], v. 5, n. 1, p. e1000344, 2009. Disponível em: <https://doi.org/10.1371/journal.pgen.1000344>. Acesso em: 2 out. 2020.

TÜMMLER, B. **Molecular epidemiology in current times.** Environmental Microbiology, [S. l.], v. 00, 2020. Disponível em: <https://doi.org/10.1111/1462-2920.15238>

UDAONDO, Z.; HUERTAS, M. J. **Fighting the enemy: one health approach against microbial resistance.** Microbial Biotechnology, [S. l.], v. 13, n. 4, p. 888–891, 2020. Disponível em: <https://doi.org/10.1111/1751-7915.13587>

URWIN, R.; MAIDEN, M. C. J. **Multi-locus sequence typing: A tool for global epidemiology.** [S. l.]: Elsevier Ltd, 2003. Disponível em: <https://doi.org/10.1016/j.tim.2003.08.006>. Acesso em: 23 set. 2020.

VAN DEN BOGAARD, A. E.; STOBBERINGH, E. E. **Antibiotic Usage in Animals.** Drugs, [S. l.], v. 58, n. 4, p. 589–607, 1999. Disponível em: <https://doi.org/10.2165/00003495-199958040-00002>. Acesso em: 4 out. 2020.

VARGHESE, N. J. *et al.* **Microbial species delineation using whole genome sequences.** Nucleic Acids Research, [S. l.], v. 43, n. 14, p. 6761–6771, 2015. Disponível em: <https://doi.org/10.1093/nar/gkv657>. Acesso em: 13 ago. 2020.

VERRAES, C. *et al.* **Antimicrobial resistance in the food chain: A review.** [S. l.: s. n.] Disponível em: <https://doi.org/10.3390/ijerph10072643>

WANG, C. *et al.* **Identification of novel mobile colistin resistance gene mcr-10.** Emerging Microbes and Infections, [S. l.], v. 9, n. 1, p. 508–516, 2020. Disponível em: <https://doi.org/10.1080/22221751.2020.1732231>. Acesso em: 10 set. 2020.

WANG, X. *et al.* **Emergence of a novel mobile colistin resistance gene, mcr-8, in NDM-producing Klebsiella pneumoniae article.** Emerging Microbes and Infections, [S. l.], 2018 a. Disponível em: <https://doi.org/10.1038/s41426-018-0124-z>

WANG, X.; KING JORDAN, I.; MAYER, L. W. **A Phylogenetic Perspective on Molecular Epidemiology.** [S. l.]: Elsevier Ltd, 2014. v. 1–3E-book. Disponível em: <https://doi.org/10.1016/B978-0-12-397169-2.00029-9>

WANG, Y. *et al.* **Antibiotic resistance gene reservoir in live poultry markets.** Journal of Infection, [S. l.], v. 78, n. 6, p. 445–453, 2019. Disponível em: <https://doi.org/10.1016/j.jinf.2019.03.012>

WEISSMAN, S. J. *et al.* **High-resolution two-locus clonal typing of extraintestinal pathogenic *Escherichia coli***. *Applied and Environmental Microbiology*, [S. l.], v. 78, n. 5, p. 1353–1360, 2012. Disponível em: <https://doi.org/10.1128/AEM.06663-11>

WIRTH, T. *et al.* **Sex and virulence in *Escherichia coli*: An evolutionary perspective**. *Molecular Microbiology*, [S. l.], v. 60, n. 5, p. 1136–1151, 2006. Disponível em: <https://doi.org/10.1111/j.1365-2958.2006.05172.x>. Acesso em: 21 jul. 2020.

WITTE, W. BIOMEDICINE: **Medical Consequences of Antibiotic Use in Agriculture**. *Science*, [S. l.], v. 279, n. 5353, p. 996–997, 1998. Disponível em: <https://doi.org/10.1126/science.279.5353.996>. Acesso em: 4 out. 2020.

WYRES, K. L.; HOLT, K. E. ***Klebsiella pneumoniae* Population Genomics and Antimicrobial-Resistant Clones**. [S. l.]: Elsevier Ltd, 2016. Disponível em: <https://doi.org/10.1016/j.tim.2016.09.007>. Acesso em: 29 out. 2020.

XAVIER, B. B. *et al.* **Identification of a novel plasmid-mediated colistin-resistance gene, *mcr-2*, in *Escherichia coli***, Belgium, June 2016. *Eurosurveillance*, [S. l.], v. 21, n. 27, p. 30280, 2016. Disponível em: <https://doi.org/10.2807/1560-7917.ES.2016.21.27.30280>. Acesso em: 29 set. 2020.

XIANG, R. *et al.* **Colocation of the polymyxin resistance gene *mcr-1* and a variant of *mcr-3* on a plasmid in an *Escherichia coli* isolate from a chicken farm**. *Antimicrobial Agents and Chemotherapy*, [S. l.], 2018. Disponível em: <https://doi.org/10.1128/AAC.00501-18>

YANG, Y. Q. *et al.* **Colistin resistance gene *mcr-1* and its variant in *Escherichia coli* isolates from chickens in China**. *Antimicrobial Agents and Chemotherapy*, [S. l.], v. 61, n. 5, 2017. Disponível em: <https://doi.org/10.1128/AAC.01204-16>. Acesso em: 27 jul. 2020.

YANG, Y. Q. *et al.* **Novel plasmid-mediated colistin resistance gene *mcr-7.1* in *Klebsiella pneumoniae***. *Journal of Antimicrobial Chemotherapy*, [S. l.], 2018 a. Disponível em: <https://doi.org/10.1093/jac/dky111>

YANG, Z.; RANNALA, B. **Molecular phylogenetics: Principles and practice**. [S. l.: s. n.] Disponível em: <https://doi.org/10.1038/nrg3186>

YIN, W. *et al.* **Novel plasmid-mediated colistin resistance gene *mcr-3* in *Escherichia coli***. *mBio*, [S. l.], v. 8, n. 3, 2017. Disponível em: <https://doi.org/10.1128/mBio.00543-17>. Acesso em: 29 set. 2020.

ZAMPARETTE, C. P. *et al.* **IncX4 Plasmid-Mediated *mcr-1.1* in Polymyxin-Resistant *Escherichia coli* from Outpatients in Santa Catarina, Southern Brazil**. *Microbial Drug Resistance*, [S. l.], v. 00, n. 00, p. 1–8, 2020. Disponível em: <https://doi.org/10.1089/mdr.2019.0203>

ZHONG, Y.-M.; LIU, W.-E.; ZHENG, Z.-F. **Epidemiology and molecular characterization of *mcr-1* in *Escherichia coli* recovered from patients with bloodstream infections in Changsha, central China**. *Infection and Drug Resistance*, [S. l.], v. Volume 12, p. 2069–2076, 2019 a. Disponível em: <https://doi.org/10.2147/IDR.S209877>. Acesso em: 5 out. 2020.

A_Fourth_Annual_Report_AMU.pdf. Acesso em: 13 set. 2020. Disponível em: https://www.oie.int/fileadmin/Home/eng/Our_scientific_expertise/docs/pdf/.

FAO. 2020. Meat Market Review, Overview of global meat market developments in 2019,

April 2020. Rome. Disponível em: <http://www.fao.org/3/ca8819en/CA8819EN.pdf>.

9. ARTIGO

Proposta de artigo submetido à revista International Journal of Antimicrobial Agents.

Comparative Genomic of *E. coli* harbouring *mcr-1* gene from different host

*Rafaela Ramalho Guerra*¹, *Silvia Adriana Mayer Lentz*¹, *Priscila Wink*², *Daiana de Lima-Morales*², *Afonso Luís Barth*², *Andreza Francisco Martins*^{1,2*}

1. Programa de Pós-graduação em Microbiologia Agrícola e do Ambiente, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
2. Laboratório de Pesquisa em Resistência Bacteriana (LABRESIS), Hospital de Clínicas de Porto Alegre, Rio Grande do Sul, Porto Alegre, Brazil

*Corresponding author: Instituto de Ciências Básicas da Saúde (ICBS), Universidade Federal do Rio Grande do Sul (UFRGS) Sarmiento Leite, 500/12101 Code 90050-170 Porto Alegre, Brazil. E-mail address: andrezafm20@gmail.com

ABSTRACT

As resistance to colistin mediated by the *mcr-1* gene has been reported worldwide, its evolutive history needs to be clarified. This study aimed to evaluate the phylogenomic relation between *mcr-1*-positive and *mcr-1*-negative *E. coli* isolates based on host origin. We also aimed to identify whether the resistance dissemination occurred by clonal expansion or by horizontal gene transfer. Six *mcr-1* positive *Escherichia coli* isolates from swine and poultry from farms in southern Brazil were compared to 37 *E. coli* isolates (19 *mcr-1*-positive and 18 *mcr-1*-negative) from a public database. The comparison of the genomes was performed by three different methods: MLST, Clonotyping (CH), and Whole Genome Sequencing to identify SNPs. A total of 35 and 32 STs and CHs were identified respectively. The SNP tree of the 43 isolates was created based on 127,788 SNPs in the core genome and its result was similar with MLST and CH, however, it was not possible to determine a relationship among *E. coli* clones with the host and geographic location. We identified 40 and 348 antimicrobial resistance genes (ARG) among isolates collected before (*mcr-1*-negative) and after (*mcr-1*-positive) the year 2000, respectively. Strains with different resistance profiles were closely related according to the SNP tree indicating that the *E. coli* core genome is very conserved. However, the higher number of ARG

found among the most recent isolates suggests that selective pressure by antimicrobial use is related to the spread of *mcr-1*.

Keywords: *Escherichia coli*, colistin resistance, whole-genome sequencing, *mcr-1* gene, MLST, CH typing, SNPs.

1 Introduction

Escherichia coli is a commensal microorganism in the gut microbiota of healthy humans and animals, as well as in soil, vegetables, and water. These characteristics make this microorganism an excellent biological marker for the dissemination of antimicrobial resistance genes (ARG) [1] such as with *mcr-1* gene. Colistin-resistant isolates, carrying *mcr-1* gene and its variants, have already been described in many environments around the world [2] and in different bacteria species [3–5]. We also know that, production animals are the main reservoir of *mcr-1* and that both clonal and horizontal [6] gene transfers (HGT) contribute to the dissemination of resistance associated with the use and overuse of antibiotics in food animal production in the last decades [7]. Although some authors suggest that the origin of *mcr-1* gene was from *Moraxella sp.* MSG13-CO3 and later it was mobilized for Enterobacteriaceae, its evolutive history still needs to be clarified [8,9].

Genomic analysis showed that bacteria species have a conserved core genome [10–12]; however, it does not confer a fixed set of genotypic characteristics. Genotype and the phenotype are not conservative homogeneously into taxa, being monomorphic in some species and others, such as *Escherichia coli*, showing considerable phenotypic variations [12–14]. The emergence of a new phenotype is associated with the genes transfer between and into bacterial population, promoting the acquisition of new genetic elements, that are inserted in the genome by recombination events. Therefore, genomic comparison can be used as an important tool in the spread surveillance of multidrug resistance (MDR) strains [15].

Thus, this study aimed to compare the genomic structure of *mcr-1*-positive *E. coli* isolates, recovered from different farms in southern Brazil, with isolates available in the public database to explore the phylogenetic relationships among these strains

in different hosts. Besides, the present study proposes to identify whether colistin resistance spreads through clonal expansion or acquisition by HGT.

2 Methods

We selected 6 *mcr-1*-positive *E. coli* isolates (3 from poultry and 3 from chicken) from previous studies of our laboratory. The isolates were collected from different batches and farms in Rio Grande do Sul, Brazil in 2015 and 2018 (Table 1). All isolates were positive for the *mcr-1* gene by PCR and had their whole genomes sequenced. These isolates were called RSBR in this paper.

2.1 Whole Genome Sequencing

The RSBR isolates were submitted to Whole Genome Sequencing by Illumina MiSeq™ (Illumina, San Diego, CA, United States). After broth incubation, we used the PureLink™ Genomic DNA mini kit (Invitrogen) for DNA extraction and purification. The Paired-end library was constructed with the Nextera™ XT DNA Library Prep Kit (Illumina) and run was performed with the MiSeq™ Reagent V2 kit (500 cycles) with coverage of 100x.

2.2 Isolate Sampling from Public Database

The sampling from the public database was screened from the NCBI Sequence Read Archive (SRA) and was searched for “*Escherichia coli* AND *mcr-1*”. In total, 417 *E. coli* genomes (last accessed in June 2020) were returned. We have included 19 SRA archives according to the following criteria: sequenced for MiSeq Illumina platform, belonging to different hosts (human, animals, and environment) and from different geographic locations (Asia, America, Africa, and Europe). This group was called as MCRP in this paper.

A second isolates group (MCRN), *mcr-1*-negative *E. coli*, was screened for validation and genomic comparison. The data of this group was recovered from the same public database and it has already been used as an external group in another study [16]. The 18 isolates belonging to MCRN were selected according to the same criteria

used for MCRP, except by the date of isolation (these isolates are from 1950 to 1999; Table 1).

Specific information for each strain and their access number can be found in Table 1. In total, 43 isolates were used for genotyping analyses (Figure S1).

2.3 Bioinformatic analysis

Genome characterization of the RSBR isolates was performed by *in silico* analysis. Raw reads were quality trimmed and assembled with CLC Genomic Workbench 20.0.4 (<https://digitalinsights.qiagen.com>). Sequence data have been deposited in the GenBank under individual accession numbers (Table 1).

Antimicrobial resistance genes were identified using ResFinder 4.0 [17], Comprehensive Antibiotic Resistance Database (CARD) [18], and QIAGEN Microbial Insight-Antimicrobial Resistance database (QMI-AR) (<https://digitalinsights.qiagen.com>). Multilocus sequence typing (MLST) was performed following the Achtman scheme [19–21] using pubMLST *E.coli* database (<https://enterobase.warwick.ac.uk/species/index/ecoli>). The clonotyping was based on the internal 469-nucleotide (nt) and 489-nt sequence of the *fumC* and *fimH* genes, respectively [22]. CH type was performed by CHtyper [23].

2.4 Phylogenetic analysis

All comparative analyses were performed using 43 *E. coli* isolates from multi-country and diverse host origins. SNPs were identified and used to construct a maximum likelihood tree with 100 bootstrap replicates [24]. These isolates were mapped to *E. coli* K12 *substrain* MG1655 (GenBank accession number NC_000913.3) used as the reference genome in this study.

Construction of phylogenetic tree MLST based was performed by concatenated sequences of seven housekeeping genes (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA*, and *recA*) and for the phylogenetic tree, CH based it was performed by concatenated sequences of 2 genes (*fumC* and *fimH*). The infer unroot tree was performed using

the Neighbor Joining algorithm with 1000 bootstrap replicate and evolutionary model Jukes-Cantor [25] for both MLST and CH.

Minimum spanning networks (MSN) of all STs and CHs were developed by PopART software [26] to infer the mutation events among the isolates.

The time of origin and MRCA (most recent common ancestor) of MCRP and RSBR isolates were evaluated by applying Bayesian phylogenetic analyses. The temporal signal was assessed by regression of root-to-tip, using Tempest software [27], where the outliers sequences were removed.

BactDating was used to estimate dated phylogenies in MCRP and RSBR isolates using the output of Gubbins. [28]. We ran the BactDating R package with MCMC chains of 1000000 generations using the strict gamma model. To test further the significance of the temporal signal in this dataset, the MCMC was run under the same parameters but with all sampling, dates assumed identical [29]. Finally, we assessed whether the effective sample size (ESS) on the first replicate model was greater than 200 using the effectiveSize function of the coda R package [30].

2.5 Antimicrobial Resistance Profile

To compare the difference in resistance profiles between the oldest isolates (MCRN) and recent isolates (MCRP and RSBR), all the 43 isolates were divided into two groups according to their collection year. Isolates with a collection date after the year 2000 belonged to the group A and isolates with a collection date before the year 2000 belonged to the group B. The number of antimicrobial resistance genes (ARG) was accounted for by the antimicrobial class. The difference in the number of ARG between groups A and B was evaluated by Kruskal-Wallis test. The frequency of the host type in groups A and B is demonstrated in Figure 1. The correlation between the number of ARG to different antimicrobials and the year of isolation was performed by Spearman's correlation.

3 Results

3.1 Molecular characterization of the isolates

From the 43 isolates analyzed in this work, 51.16% (22/43) were recovered from animals, 32.56% (14/43) from humans followed by 4.65% (2/43) from food-borne, 4.65% (2/43) from environmental and 6.98% (3/43) of unknown source.

A total of 35 STs were identified among the 43 isolates. The ST10 was the most frequently identified (5/43) followed by ST101 (3/43). ST10 was identified from both humans and animals, *mcr-1*-positive and *mcr-1*-negative isolates (Table 1). Meanwhile, clonotyping identified 32 *fumC-fimH* allele combinations (CH), being CH11-25 the most frequent (4/43). These isolates were identified from humans, animals, and environment, from the different geographic regions, and among *mcr-1*-positive and negative isolates (Table 1).

Three of the 43 isolates were *fimH*-negative and were associated with ST101, ST6764, and ST181.

Table 1: Characteristics of *E.coli* isolates used in comparative genomics analysis

ID Run	Bio project	Bio sample	Collection date	Country	Isolation source	CH	ST	CC	Antimicrobial Resistance
RSBR samples submitted to WGS									
215M02423 ^a	PRJNA674879	SAMN16678977	2015	Brazil	Poultry	4-32	58	155	Tetracycline; Trimethoprim; Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
249M02423 ^a	PRJNA674879	SAMN16678978	2015	Brazil	Poultry	11-25	2491	-	Tetracycline; Trimethoprim; Phenicol; Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
279M02423 ^a	PRJNA674879	SAMN16678979	2015	Brazil	Poultry	11-25	2491	-	Tetracycline; Trimethoprim; Phenicol; Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
649M05549 ^a	PRJNA674879	SAMN16678980	2018	Brazil	Pig	11-0	101	101	Tetracycline; Trimethoprim; Phenicol; Macrolide; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
651M05549 ^a	PRJNA674879	SAMN16678981	2018	Brazil	Pig	41-158	10	10	Tetracycline; Trimethoprim; Phenicol; Macrolide; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
914M05549 ^a	PRJNA674879	SAMN16678982	2018	Brazil	Pig	41-27	101	101	Tetracycline; Trimethoprim; Phenicol; Macrolide; Lincosamide; Fluoroquinolone and Quinolone; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin

MCRP samples recovered to SRA NCBI database

SRR10099945	PRJNA312475	SAMN12730713	2005	Spain	Pig	29-38	156	156	Tetracycline; Trimethoprim; Phenicol; Macrolide; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
SRR10099933	PRJNA312475	SAMN12730715	2006	Spain	Pig	11-54	10	10	Tetracycline; Trimethoprim; Macrolide; Aminoglycoside; Sulphonamide; Colistin
SRR10099929	PRJNA312475	SAMN12730719	2007	Spain	Pig	4-61	224	-	Tetracycline; Trimethoprim; Phenicol; Lincosamide; Colistin
SRR10099927	PRJNA312475	SAMN12730721	2011	Spain	Pig	4-31	763	-	Tetracycline; Trimethoprim; Phenicol; Macrolide; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
SRR10099942	PRJNA312475	SAMN12730724	2012	Spain	Pig	29-38	156	156	Tetracycline; Trimethoprim; Phenicol; Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
SRR5328644	PRJNA376494	SAMN06480071	2012	China	Pig	11-521	7314	-	Tetracycline; Trimethoprim; Phenicol; Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Rifampicin; Sulphonamide; Colistin
SRR6750074	PRJNA434359	SAMN08554943	2012	China	Homo sapiens	11-0	6764	-	Tetracycline; Trimethoprim; Phenicol; Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Fosfomycin; Sulphonamide; Beta-lactam
DRR065948	PRJDB5007	SAMD00056132	2013	Japan	Bovine	88-145	457	-	Tetracycline; Trimethoprim; Macrolide; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
SRR10099940	PRJNA312475	SAMN12730726	2014	Spain	Pig	11-54	10	10	Tetracycline; Trimethoprim; Phenicol; Macrolide; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin

SRR5350657	PRJNA377578	SAMN06463354	2016	Canada	Homo sapiens	11-25	3944	-	Tetracycline; Trimethoprim; Phenicol; Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Rifampicin; Sulphonamide; Beta-lactam; Colistin
ERR1360256	PRJEB13576	SAMEA3931471	2016	United Kingdom	Pig	932-215	6913	-	Tetracycline; Macrolide; Fluoroquinolone and Quinolone; Beta-lactam; Colistin
ERR1407279	PRJEB13470	SAMEA3980468	2016	-	Retail food	4-39	88	23	Tetracycline; Macrolide; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
SRR6791684	PRJNA436232	SAMN08622661	2016	USA	Homo sapiens	31-58	1485	648	Tetracycline; Trimethoprim; Phenicol; Macrolide; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
SRR6791870	PRJNA436242	SAMN08622909	2017	USA	Homo sapiens	11-24	10	10	Tetracycline; Trimethoprim; Phenicol; Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
SRR10392137	PRJNA587479	SAMN13191328	2017	Egypt	Food	29-38	101	101	Tetracycline; Trimethoprim; Phenicol; Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
SRR10521259	PRJNA587093	SAMN13197356	2017	Brazil	Water	4-121	683	-	Tetracycline; Trimethoprim; Phenicol; Macrolide; Aminoglycoside; Sulphonamide; Colistin
DRR138478	PRJDB7066	SAMD00127158	2018	Thailand	Homo sapiens	41-86	4014	101	Tetracycline; Trimethoprim; Phenicol; Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Sulphonamide; Beta-lactam; Colistin
ERR3479038	PRJEB34000	SAMEA5863154	2018	Bangladesh	Poultry	6-31	1196	-	Tetracycline; Trimethoprim; Phenicol; Macrolide; Aminoglycoside; Sulphonamide;

ERR3479039	PRJEB34000	SAMEA5863155	2018	Bangladesh	Poultry	82-86	1818	-	Beta-lactam; Colistin Macrolide; Fluoroquinolone and Quinolone; Aminoglycoside; Beta-lactam; Colistin
MCRN samples recovered to SRA NCBI database									
ERR1010184	PRJEB10700	SAMEA3529397	1950	not provided	Homo sapiens	24-21	583	122	Macrolide
ERR1010185	PRJEB10700	SAMEA3529398	1955	United Kingdom	Homo sapiens	4-54	3	20	Macrolide
ERR1010188	PRJEB10700	SAMEA3529401	1960	not provided	Pig	11-24	1260	-	Macrolide
ERR1010187	PRJEB10700	SAMEA3529400	1961	not provided	Pig	11-23	5786	-	Macrolide
ERR1010278	PRJEB10700	SAMEA3529298	1973	Israel	Bovine	746-453	5768	-	
ERR1010247	PRJEB10700	SAMEA3529333	1979	Sweden	Pig	7-25	871	-	Tetracycline; Macrolide; Aminoglycoside; Sulphonamide
ERR1010215	PRJEB10700	SAMEA3529382	1980	USA	Homo sapiens	11-53	1434	-	Macrolide
ERR1010280	PRJEB10700	SAMEA3529378	1982	South Korea	Pig	54-0	3057	-	Tetracycline; Macrolide; Aminoglycoside; Beta-lactam
ERR1010216	PRJEB10700	SAMEA3529383	1984	Liberia	Homo sapiens	748-0	5782	-	Macrolide
ERR1010277	PRJEB10700	SAMEA3529365	1986	United Kingdom	Homo sapiens	11-215	10	10	Macrolide
ERR1010266	PRJEB10700	SAMEA3529297	1988	USA	Homo sapiens	11-23	216	-	Macrolide
ERR1010246	PRJEB10700	SAMEA3529332	1990	Guinea-Bissau	Homo sapiens	4-233	1147	-	Tetracycline; Macrolide; Aminoglycoside; Beta-lactam
ERR1010255	PRJEB10700	SAMEA3529342	1990	Guinea-Bissau	Homo sapiens	11-0	181	168	Macrolide
ERR1010283	PRJEB10700	SAMEA3529299	1991	Denmark	Water	11-25	500	-	Tetracycline; Trimethoprim; Phenicol;

ERR1010284	PRJEB10700	SAMEA3529306	1992	Spain	Goat	22-48	5775	-	Macrolide; Sulphonamide; Beta-lactam Tetracycline; Macrolide; Aminoglycoside;
ERR1010252	PRJEB10700	SAMEA3529339	1992	Denmark	Homo sapiens	6-32	847	-	Macrolide; Aminoglycoside
ERR1010248	PRJEB10700	SAMEA3529334	1997	Denmark	-	6-32	1308	-	-
ERR1010285	PRJEB10700	SAMEA3529386	1999	Denmark	Homo sapiens	14-57	5777	-	Macrolide

a

RSBR isolates mcr-1 positive.

MCRP: isolates mcr-1 positive.

MCRN: isolates mcr-1 negative.

SRA NCBI database: Sequence Read Archive data from National Center for Biotechnology Information.

3.2 Resistance analysis

In total, 388 genes conferring resistance to different antimicrobials were identified through different databases. The most common mechanisms of resistance in the 43 isolates were associated with macrolides, tetracyclines, and aminoglycosides. All the isolates harboring the *mcr-1* gene presented genes conferring resistance to more than 3 antimicrobial classes. We observed that the number of ARG tends to increase according to the collection year of the isolates (Spearman's ρ 0.688 $p < 0.001$; Table S1). Furthermore, we observed that the number of ARG in recent isolates (348/388; group A) was significantly higher than in older isolates (40/388; group B; $p < 0.001$) (Table S2). Colistin, Trimethoprim, and Beta-lactams presented a greater difference between groups A and B (Figure 1; $p < 0.001$). Only ARG conferring resistance to rifampicin and fosfomicin presented no significant difference between the two groups.

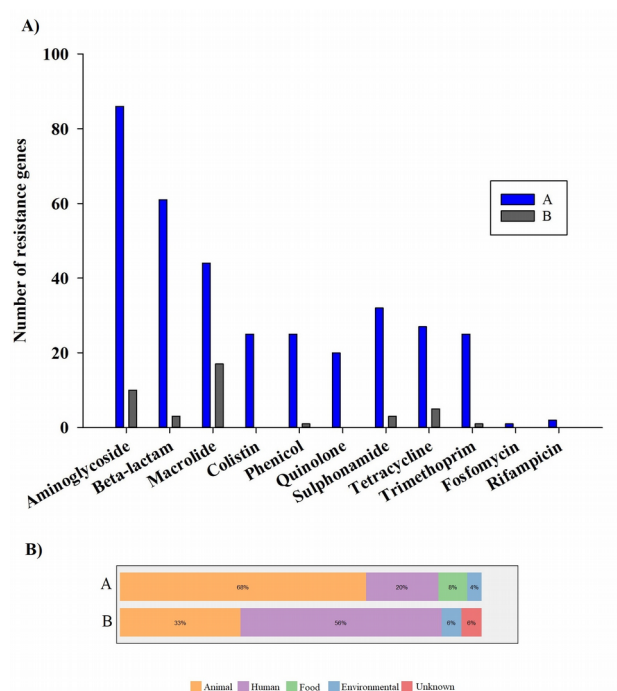


Figure 1: A) Number of resistance genes distributed in different antimicrobial class according to the collection date. Group A represents the isolates with a collection

date after the year 2000s and group B represents the isolates with a collection date before the year 2000s. B) Frequency of host type distributed in group A and group B.

3.3 Phylogenetic analysis

A maximum-likelihood tree of the 43 *E. coli* isolates was created based on 127,788 SNPs in the core genome (Figure 2). The isolates were grouped in 3 major clusters and in one of them all isolates are *mcr-1*-positive (Figure 2; cluster A). The RSBR isolates were distributed in two of the three clusters (Figure 2; clusters A and B), closely related with MCRP and MCRN isolates from humans and the environment.

The isolates RSBR 649M05549 and 914M05549 (ST101; swine) were related to isolates DRR13878 (ST101) from foodborne and SRR10392137 (ST4014) from the human. These strains differed from each other by an average of 1600 SNPs (Figure S3).

The isolate RSBR 215M02423 (ST58; poultry) was related with the isolate SRR10521259 (ST683; environmental). These isolates differed by 2572 SNPs and presented a similarity of 97.99%. The isolates RSBR 651M05549 (ST10; swine), ERR1010187 (ST5786; human; MCRN), and ERR1010188 (ST1260; human; MCRN) were closely related and differed from each other by an average of 2800 SNPs (Figure S3).

The isolates RSBR 249M02423 and 279M02423 (ST2491; poultry) have high similarity (just 3 SNPs different) and were closely related with isolate ERR1010247 (ST871; pig) with an average 9860 SNPs among them (Figure S2).

Also, the RSBR 651M05549, 279M02423, and 249M02423 isolates were closely related to MCRN isolates and most of them were resistant just to aminoglycosides. However, isolates RSBR 649M05549, 914M05549 and 215M02423 were related to MCRP isolates, most of them harboring resistance genes for different classes of antimicrobial (Figure 2).

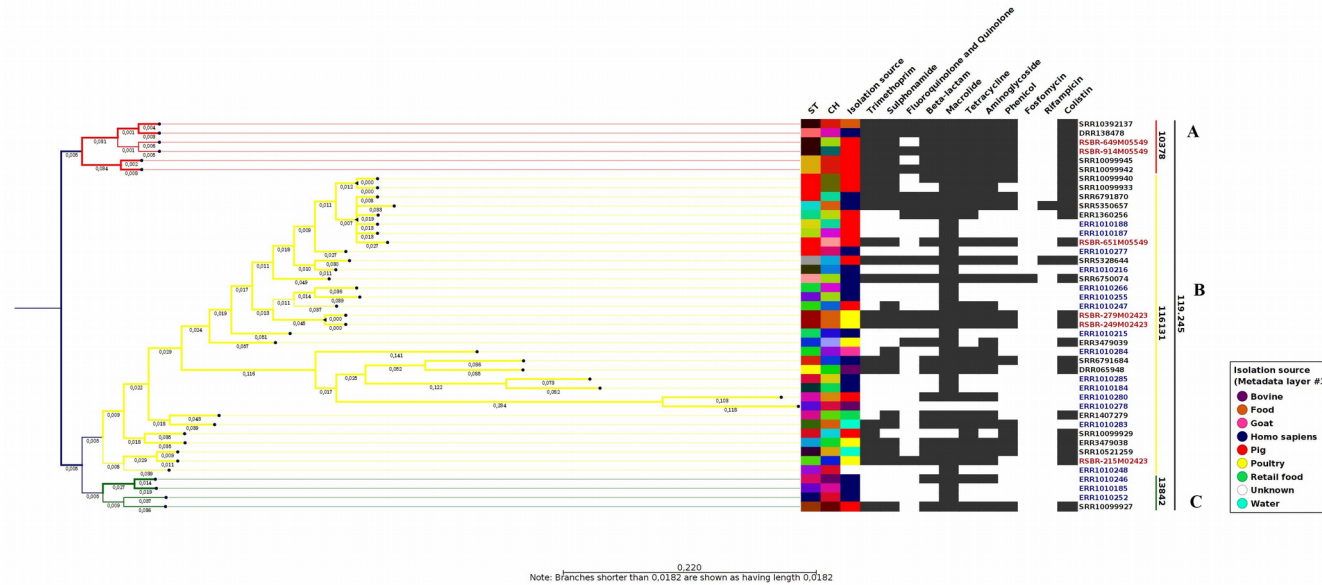


Figure 2: Phylogenetic tree built from single nucleotide polymorphisms (SNPs) by mapping the total genome against the reference genome (GenBank access number NC_000913.3). The tree was inferred using the CLC software Genomic Workbench 20.0.4 (<https://digitalinsights.qiagen.com>) by the Maximum Likelihood (ML) method with 100 bootstrap replicates. The highlighted branches represent a value of bootstrap $\geq 90\%$ and the colors indicate different sub-trees (Cluster A in red, Cluster B in yellow, Cluster C in green). The colors on the right side (legend) represent the isolation source from each isolate. The resistance profile is represented by black square according to antimicrobial class. The RSBR isolates are highlighted in red, the MCRN are in blue and the MCRP are in black. The number of SNPs is represented within each cluster and among clusters.

According to Minimum Spanning Tree generated from STs (Figure 3), the isolates belonging to ST 10 were grouped into the cluster from animals (SRR10099940 and SRR10099933) and humans (SRR6791870 and ERR1010277) host. Most isolates were not grouped into clusters but some isolates from animals (RSBR 215M02423 and SRR10099929) belonging to ST 58 and ST 224, and isolates from environmental (SRR105221259 and ERR1407279) belonging to ST 683 and ST 88, which were closely related and spread with few mutational events (Figure 3). The Minimum Spanning Tree obtained from CHs, grouped isolates in five clusters with up to two mutational events and one of them with more than ten isolates from different hosts and clonotypes (Figure S3).

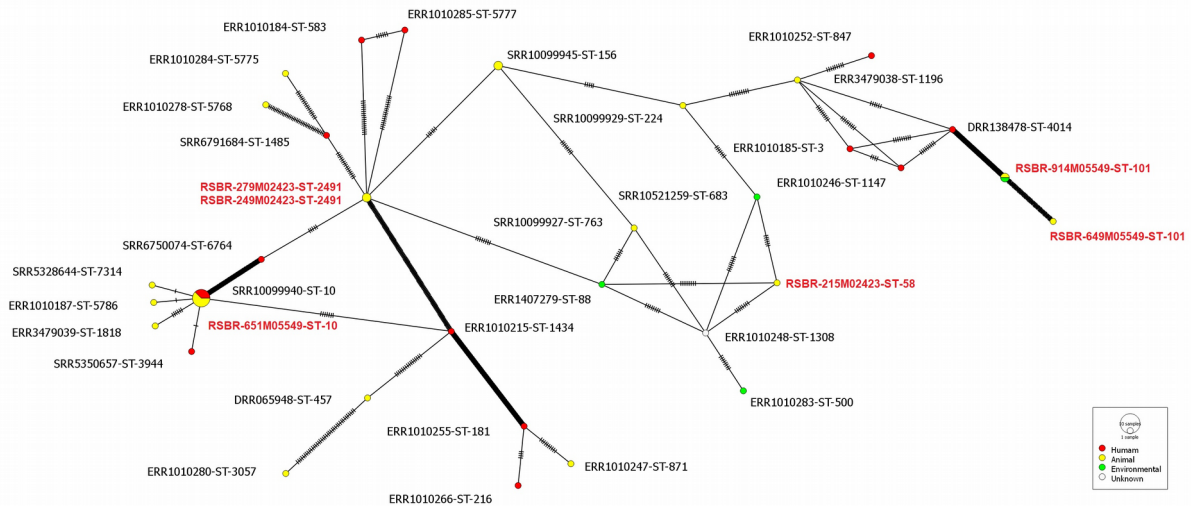


Figure 3: Minimum Spanning tree built with the online tool popART (<http://popart.otago.ac.nz/index.shtml>) from the multiple alignment of *housekeeping* genes. Each trace represents a mutation event. The nucleotide diversity was $\pi = 0.2139$. The different colors, represented in the lower right corner of the figure represent the origin of the isolates, being red (human origin); yellow (animal origin) and green (environmental origin).

3.4 The time origin analysis of *mcr-1* positive isolates

The deviance information criterion DIC [31] was 352.99 compared to 225.54 when the correct dates were used, which indicates conclusively that the temporal signal is significant with $R^2 = 0.97$ (Figure S4; A).

A strict gamma model was inferred with a mean rate μ 9.94 [8.66-11.3] substitutions per genome per year. The root date for 23 isolates *mcr-1* positive (RSBR and MCRP) is estimated to be 1996 with a 95% credible interval, from 1993 to 1997 (Figure S4; B). Furthermore, we observed that substitution rates are lower on longer time scales compared to recent time scales (Figure S4; C).

4 Discussion

According to molecular typing results, we found a high diversity among the 43 isolates with more than 30 STs and clonotypes, and ST10 and CH11-25 were the most frequent. Similar results were observed in studies where the population structure of *mcr-1* carrying *E. coli* was highly diverse and usually dominated by strains related to ST10 widely distributed among the continents [6,32,33]. Our results confirm these observations, showing that 11.63% (5/43) of the isolates belonging to ST10 and were isolated in Europe, North America, and South America.

It is important to consider that RSBR isolates belonged to a variety of STs clinically important, as ST10, ST101, and ST58 (Table 1). These results suggest the potential of environmental spread and promiscuity of these strains among different hosts [34–36]. This is also observed in the Minimum Spanning tree (Figure 3) where the ST10 isolates from human and animal hosts were grouped into the same cluster, suggesting a clonal dispersion of this strain [34]. Additionally, the RSBR isolates were closely related to other isolates from different hosts (Figure 2) and, consequently, it did not find a phylogenetic relationship between host source and ST or CH, which is supported by previous studies [6,37]. This result suggests that different *E. coli* strains maintain an infection or colonization mechanism without strict host restriction.

Although the CH typing separated strains belonging to the same ST, in this study, MLST and CH were largely concordant with the phylogenomic results observed in the SNP tree (Figure 2). Phylogenetic inferences demonstrated the performance of linkage disequilibrium in the population, a feature already observed in the genomic study of *E. coli* [38]. For this reason, it was possible to observe isolates closely related but differing in terms of genetic repertoires, such as positive and negative *mcr-1* isolates. In the same way, isolates of the same ST may show varying and heterogeneous levels of core genome divergence between them, as reported previously [38,39].

Interestingly, Cluster A and C are formed by isolates with a lower number of SNPs, similar resistance profiles into the cluster, and one host type, mostly (animals in cluster A and humans in cluster C). On the opposite, cluster B is the largest cluster in the SNP tree and the most diverse, with different STs, CHs, and hosts. The STs observed in this cluster are associated with both commensal strains widely identified in animals and humans and pathogenic strains, related to infections disease as ST10.

Regardless of the resistance profile, our study revealed that all the *mcr-1*-positive *E. coli* isolates carry multiple ARGs (Table 1) mainly the resistance genes to macrolides, tetracyclines, and aminoglycosides. Besides that, the beta-lactamase genes as *bla*_{CTX-M8} and *bla*_{CMY-2} also were observed in the RSBR isolates. In the same way, other studies also revealed that *mcr-1*-positive isolates are usually multidrug-resistant [40–42].

It was also possible to observed that the number of genes and contigs associated with ARG (Figure 1, Table S2) conferring resistance to most antimicrobial classes was positively related to the collection date of the isolates, (Spearman's rho = 0.688, $p < 0.001$ and rho = 0.950, $p < 0.001$, respectively; Table S1), confirming the increase in the resistance over the years between isolates in this study. Our results are consistent with results observed in a controlled evolution study in which the minimum

inhibitory concentration (MIC) of the population exposed to polymyxin B increased 3000 times compared to MIC of wild strains [43].

Despite the limited number of SRA files of *mcr-1*-positive *E.coli* isolates before 2015 in public databases, (Figure S5), our estimate of its evolution is consistent with literature results, which report the *mcr-1* gene in isolates collected around the 2000s [46]. Phylogenetic dating estimated the MRCA to be 1996 and rate substitutions per genome per year of 9.94 with a tendency to decrease over time. This pattern has been attributed to the gradual elimination of deleterious mutations over time [45,46] and this phenomenon is observed in different bacterial genera [39].

The *mcr-1*-positive isolates recently recovered tend to have a higher substitution rate per genome per year, as lineages undergoing adaptive evolution are expected to accumulate substitutions more rapidly than those subjects to negative selection [47]. This further emphasizes the idea that the *mcr-1* gene spread is not a recent event and can remain in the population due to selective pressure by antimicrobials use, mainly in the livestock. However, it is important to note that replacement rates are determined by various factors, such as the direction and strength of natural selection, as well as generation time and population size [48], so these results should be interpreted with care.

Furthermore, Ahmed S. et al. demonstrated non-clonal and promiscuous dispersal of the *mcr-1* gene among isolates suggesting efficient horizontal gene transfer [40]. Our results are consistent with these findings, where isolates with different resistance profiles, dates, and isolation sources were closely related. This highlights the fact that a conserved genome does not guarantee a fixed phenotypic trait set [10,12] and probably, the selective pressure by antimicrobial use can directly act on recombination events ensuring dissemination of *mcr-1*-positive strains into different hosts.

The main limitation in this study was the few SRA files available in the public database, considering different host origins, geographical location, and collection date. In addition, the divergence time may be biased due to the number of samples analyzed and the absence of older *mcr-1*-positive isolates in the SRA database.

5 Conclusion

The comparative genomic demonstrated that *mcr-1*-positive *Escherichia coli* isolates are highly diverse phylogenetically and are spread in different sequence type ST, clonal complex CC, and hosts.

In general, the results presented by the 3 typing methods were concordant. It was not possible to establish the relationship among genotype and hosts, geographic location, and collect data. However, it was possible to relate the resistance profile over time and to estimate a common ancestor among the *mcr-1*-positive isolates. Besides that, we were able to relate the dispersion of the *mcr-1* gene, its divergence time, and the number of substitutions in the genome with the selective pressure imposed by the use of antimicrobials, over time.

Funding: This work was supported by Brazilian Federal Foundation for Support and Evaluation of Graduate Education (CAPES), National Council for Scientific and Technological Development (CNPq), and National Institute of Antimicrobial Resistance Research - INPRA (MCTI/CNPq/CAPES/FAPs nº 16/2014).

Declarations of interest: None.

References

- [1] Poolman JT, Wacker M. Extraintestinal pathogenic *Escherichia coli*, a common human pathogen: challenges for vaccine development and progress in the field. *J Infect Dis* 2016;213:6–13. <https://doi.org/10.1093/infdis/jiv429>.
- [2] Rebelo AR, Bortolaia V, Kjeldgaard JS, Pedersen SK, Leekitcharoenphon P, Hansen IM, et al. Multiplex PCR for detection of plasmid-mediated colistin resistance determinants, *mcr-1*, *mcr-2*, *mcr-3*, *mcr-4* and *mcr-5* for surveillance purposes. *Eurosurveillance* 2018;23. <https://doi.org/10.2807/1560-7917.ES.2018.23.6.17-00672>.
- [3] Wang X, Wang Y, Zhou Y, Li J, Yin W, Wang S, et al. Emergence of a novel mobile colistin resistance gene, *mcr-8*, in NDM-producing *Klebsiella*

pneumoniae article. *Emerg Microbes Infect* 2018;7.
<https://doi.org/10.1038/s41426-018-0124-z>.

- [4] Yang YQ, Li YX, Lei CW, Zhang AY, Wang HN. Novel plasmid-mediated colistin resistance gene *mcr-7.1* in *Klebsiella pneumoniae*. *J Antimicrob Chemother* 2018;73:1791–5. <https://doi.org/10.1093/jac/dky111>.
- [5] Li R, Zhang P, Yang X, Wang Z, Fanning S, Wang J, et al. Identification of a novel hybrid plasmid coproducing MCR-1 and MCR-3 variant from an *Escherichia coli* strain. *J Antimicrob Chemother* 2019.
<https://doi.org/10.1093/jac/dkz058>.
- [6] Shen C, Zhong L-L, Yang Y, Doi Y, Paterson DL, Stoesser N, et al. Dynamics of *mcr-1* prevalence and *mcr-1*-positive *Escherichia coli* after the cessation of colistin use as a feed additive for animals in China: a prospective cross-sectional and whole genome sequencing-based molecular epidemiological study. *The Lancet Microbe* 2020;1:e34–43. [https://doi.org/10.1016/s2666-5247\(20\)30005-7](https://doi.org/10.1016/s2666-5247(20)30005-7).
- [7] Wang Y, Hu Y, Cao J, Bi Y, Lv N, Liu F, et al. Antibiotic resistance gene reservoir in live poultry markets. *J Infect* 2019;78:445–53.
<https://doi.org/10.1016/j.jinf.2019.03.012>.
- [8] Snedrud E, McGann P, Chandler M. The birth and demise of the IS*Apl1*-*mcr-1*-IS*Apl1* composite transposon: The vehicle for transferable colistin resistance. *MBio* 2018;9. <https://doi.org/10.1128/mBio.02381-17>.
- [9] Poirel L, Kieffer N, Nordmann P. In vitro study of IS*Apl1*-mediated mobilization of the colistin resistance gene *mcr-1*. *Antimicrob Agents Chemother* 2017;61.
<https://doi.org/10.1128/AAC.00127-17>.
- [10] Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;9. <https://doi.org/10.1038/s41467-018-07641-9>.
- [11] Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, Kyrpides NC, et al. Microbial species delineation using whole genome sequences. *Nucleic Acids Res* 2015;43:6761–71.
<https://doi.org/10.1093/nar/gkv657>.
- [12] Redondo-Salvo S, Fernández-López R, Ruiz R, Vielva L, de Toro M, Rocha EPC, et al. Pathways for horizontal gene transfer in bacteria revealed by a global map of their plasmids. *Nat Commun* 2020;11:1–13.
<https://doi.org/10.1038/s41467-020-17278-2>.

- [13] Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, et al. Sex and virulence in *Escherichia coli*: An evolutionary perspective. *Mol Microbiol* 2006;60:1136–51. <https://doi.org/10.1111/j.1365-2958.2006.05172.x>.
- [14] Shapiro BJ. How clonal are bacteria over time? *Curr Opin Microbiol* 2016;31:116–23. <https://doi.org/10.1016/j.mib.2016.03.013>.
- [15] Leimbach A, Hacker J, Dobrindt U. *E. coli* as an all-rounder: The thin line between commensalism and pathogenicity. *Curr Top Microbiol Immunol* 2013;358:3–32. https://doi.org/10.1007/82_2012_303.
- [16] Joensen KG, Tetzschner AMM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and easy in silico serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. *J Clin Microbiol* 2015;53:2410–26. <https://doi.org/10.1128/JCM.00008-15>.
- [17] Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother* 2020;75:3491–500. <https://doi.org/10.1093/jac/dkaa345>.
- [18] Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: Antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res* 2020;48:D517–25. <https://doi.org/10.1093/nar/gkz935>.
- [19] Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, et al. Sex and virulence in *Escherichia coli*: An evolutionary perspective. *Mol Microbiol* 2006;60:1136–51. <https://doi.org/10.1111/j.1365-2958.2006.05172.x>.
- [20] Clermont O, Gordon D, Denamur E. Guide to the various phylogenetic classification schemes for *Escherichia coli* and the correspondence among schemes. *Microbiol (United Kingdom)* 2015;161:980–8. <https://doi.org/10.1099/mic.0.000063>.
- [21] Larsen M V., Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, et al. Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol* 2012;50:1355–61. <https://doi.org/10.1128/JCM.06094-11>.
- [22] Weissman SJ, Johnson JR, Tchesnokova V, Billig M, Dykhuizen D, Riddell K, et al. High-resolution two-locus clonal typing of extraintestinal pathogenic *Escherichia coli*. *Appl Environ Microbiol* 2012;78:1353–60. <https://doi.org/10.1128/AEM.06663-11>.
- [23] Roer L, Johannesen TB, Hansen F, Stegger M, Tchesnokova V, Sokurenko E, et al. CHTyper, a web tool for subtyping of extraintestinal pathogenic

- Escherichia coli based on the *fumC* and *fimH* alleles. *J Clin Microbiol* 2018;56. <https://doi.org/10.1128/JCM.00063-18>.
- [24] Hasegawa M, Kishino H, Yano T aki. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 1985;22:160–74. <https://doi.org/10.1007/BF02101694>.
- [25] Jukes TH CC. Evolution of protein molecules. *Mammalian protein Metabolism*. New York: Academic Press; 1969.
- [26] Leigh JW, Bryant D. `popart` : full-feature software for haplotype network construction. *Methods Ecol Evol* 2015;6:1110–6. <https://doi.org/10.1111/2041-210X.12410>.
- [27] Rambaut A, Lam TT, Carvalho LM, Pybus OG. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* 2016;2. <https://doi.org/10.1093/ve/vew007>.
- [28] Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015;43:e15–e15. <https://doi.org/10.1093/nar/gku1196>.
- [29] Didelot X, Croucher NJ, Bentley SD, Harris SR, Wilson DJ. Bayesian inference of ancestral dates on bacterial phylogenetic trees. *Nucleic Acids Res* 2018;46:e134. <https://doi.org/10.1093/nar/gky783>.
- [30] CODA: convergence diagnosis and output analysis for MCMC - Open Research Online n.d. <http://oro.open.ac.uk/22547/> (accessed February 16, 2021).
- [31] Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B (Statistical Methodol)* 2002;64:583–639. <https://doi.org/10.1111/1467-9868.00353>.
- [32] Matamoros S, Van Hattem JM, Arcilla MS, Willemsse N, Melles DC, Penders J, et al. Global phylogenetic analysis of Escherichia coli and plasmids carrying the *mcr-1* gene indicates bacterial diversity but plasmid restriction. *Sci Rep* 2017;7:1–9. <https://doi.org/10.1038/s41598-017-15539-7>.
- [33] Ahmed S, Olsen JE, Herrero-Fresno A. The genetic diversity of commensal Escherichia coli strains isolated from non-antimicrobial treated pigs varies according to age group. *PLoS One* 2017;12:e0178623. <https://doi.org/10.1371/journal.pone.0178623>.

- [34] Reid CJ, De Maere MZ, Djordjevic SP. Australian porcine clonal complex 10 (CC10) *Escherichia coli* belong to multiple sublineages of a highly diverse global CC10 phylogeny. *Microb Genomics* 2019;5. <https://doi.org/10.1099/mgen.0.000225>.
- [35] Zhong YM, Liu WE, Zheng ZF. Epidemiology and molecular characterization of *mcr-1* in *Escherichia coli* recovered from patients with bloodstream infections in Changsha, central China. *Infect Drug Resist* 2019;12:2069–76. <https://doi.org/10.2147/IDR.S209877>.
- [36] McKinnon J, Roy Chowdhury P, Djordjevic SP. Genomic analysis of multidrug-resistant *Escherichia coli* ST58 causing urosepsis. *Int J Antimicrob Agents* 2018;52:430–5. <https://doi.org/10.1016/j.ijantimicag.2018.06.017>.
- [37] Tang F, Wang J, Li D, Gao S, Ren J, Ma L, et al. Comparative genomic analysis of 127 *Escherichia coli* strains isolated from domestic animals with diarrhea in China. *BMC Genomics* 2019;20:212. <https://doi.org/10.1186/s12864-019-5588-2>.
- [38] Denamur E, Clermont O, Bonacorsi S, Gordon D. The population genetics of pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2021;19:37–54. <https://doi.org/10.1038/s41579-020-0416-x>.
- [39] McNally A, Oren Y, Kelly D, Pascoe B, Dunn S, Sreecharan T, et al. Combined Analysis of Variation in Core, Accessory and Regulatory Genome Regions Provides a Super-Resolution View into the Evolution of Bacterial Populations. *PLOS Genet* 2016;12:e1006280. <https://doi.org/10.1371/journal.pgen.1006280>.
- [40] Ahmed S, Das T, Islam MZ, Herrero-Fresno A, Biswas PK, Olsen JE. High prevalence of *mcr-1*-encoded colistin resistance in commensal *Escherichia coli* from broiler chicken in Bangladesh. *Sci Rep* 2020;10:1–13. <https://doi.org/10.1038/s41598-020-75608-2>.
- [41] Maciuca IE, Cummins ML, Cozma AP, Rimbu CM, Guguianu E, Panzaru C, et al. Genetic Features of *mcr-1* Mediated Colistin Resistance in CMY-2-Producing *Escherichia coli* From Romanian Poultry. *Front Microbiol* 2019;10. <https://doi.org/10.3389/fmicb.2019.02267>.
- [42] Ben Said L, Jouini A, Klibi N, Dziri R, Alonso CA, Boudabous A, et al. Detection of extended-spectrum beta-lactamase (ESBL)-producing Enterobacteriaceae in vegetables, soil and water of the farm environment in Tunisia. *Int J Food Microbiol* 2015;203:86–92. <https://doi.org/10.1016/j.ijfoodmicro.2015.02.023>.

- [43] Spohn R, Daruka L, Lázár V, Martins A, Vidovics F, Grézal G, et al. Integrated evolutionary analysis reveals antimicrobial peptides with limited resistance. *Nat Commun* 2019;10:1–13. <https://doi.org/10.1038/s41467-019-12364-6>.
- [44] Wang R, Van Dorp L, Shaw LP, Bradley P, Wang Q, Wang X, et al. The global distribution and spread of the mobilized colistin resistance gene *mcr-1*. *Nat Commun* 2018;9. <https://doi.org/10.1038/s41467-018-03205-z>.
- [45] Penny D. Evolutionary biology: Relativity for molecular clocks. *Nature* 2005;436:183–4. <https://doi.org/10.1038/436183a>.
- [46] Ho SYW, Duchêne S, Molak M, Shapiro B. Time-dependent estimates of molecular evolutionary rates: Evidence and causes. *Mol Ecol* 2015;24:6007–12. <https://doi.org/10.1111/mec.13450>.
- [47] Eyre-Walker A, Keightley PD. The distribution of fitness effects of new mutations. *Nat Rev Genet* 2007;8:610–8. <https://doi.org/10.1038/nrg2146>.
- [48] Duchêne S, Holt KE, Weill FX, Le Hello S, Hawkey J, Edwards DJ, et al. Genome-scale rates of evolutionary change in bacteria. *Microb Genomics* 2016;2:e000094. <https://doi.org/10.1099/mgen.0.000094>.

Supplementary Materials

Table S1

Spearman's correlation matrix														
Resistance genes														
Resistance contigs	0.950***													
Aminoglycoside	0.892***	0.841***												
Beta-lactam	0.842***	0.815***	0.688***											
Macrolide	0.664***	0.635***	0.557***	0.455**										
Colistin	0.829***	0.843***	0.721***	0.748***	0.564***									
Phenicol	0.763***	0.831***	0.632***	0.635***	0.587***	0.649***								
Quinolone	0.586***	0.496***	0.489***	0.462**	0.176	0.551***	0.424**							
Sulphonamide	0.851***	0.837***	0.797***	0.654***	0.655***	0.714***	0.741***	0.369*						
Tetracycline	0.769***	0.791***	0.648***	0.658***	0.399**	0.710***	0.596***	0.364*	0.707***					
Trimethoprim	0.866***	0.889***	0.775***	0.711***	0.545***	0.753***	0.824***	0.486***	0.803***	0.707***				
Fosfomycin	0.252	0.157	0.166	0.250	-0.090	0.131	0.125	0.308*	0.220	0.088	0.117			
Rifampicin	0.252	0.271	0.105	0.156	0.066	0.184	0.267	0.442**	0.079	0.276	0.276	-0.035		
Collection date	0.688***	0.708***	0.604***	0.684***	0.530***	0.858***	0.597***	0.527***	0.666***	0.606***	0.647***	0.826***	0.025	0.093

Note: ** $p < 0.01$, *** $p < 0.001$.

Table S2

Antimicrobial Class	N° resistance genes		
	A	B	H
Aminoglycoside	86	10	21.836*
Beta-lactam	61	3	23.509*
Macrolide	44	17	13.359*
Colistin	25	0	42.000*
Phenicol	25	1	17.677*
Quinolone	20	0	12.739*
Sulphonamide	32	3	21.405*
Tetracycline	27	5	21.197*
Trimethoprim	25	1	23.846*
Fosfomicin	1	0	0.72
Rifampicin	2	0	1.39
Total (n=388)	348/388	40/388	28.831*

Note: * $p < 0.001$

H: Kruskal-Wallis test.

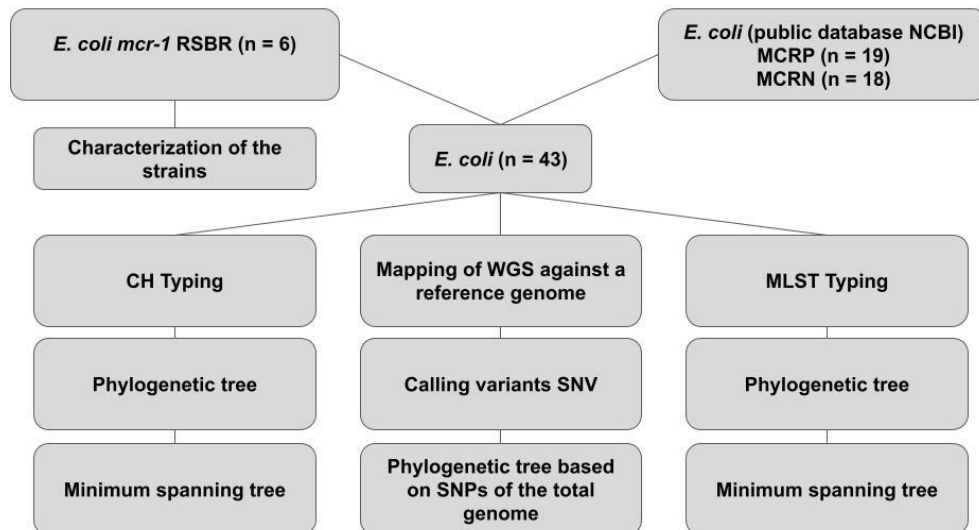


Figure S1: Methodology fluxogram.

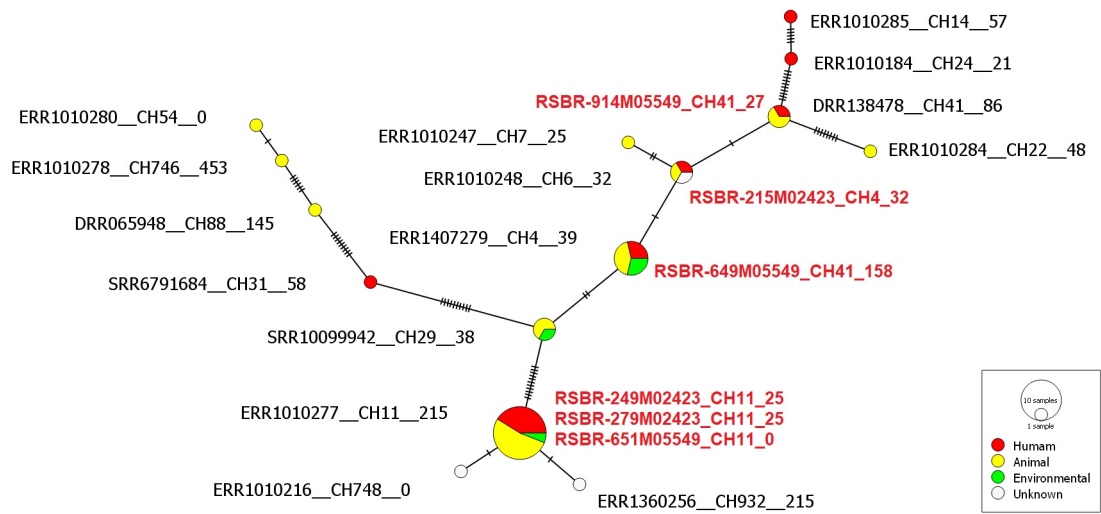
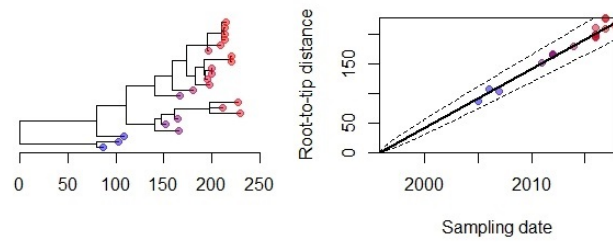


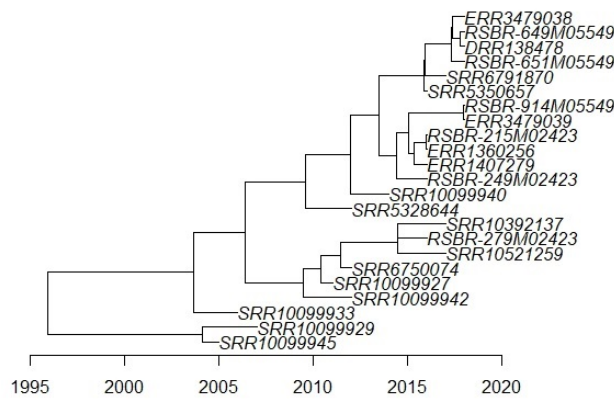
Figure S2: Minimum spanning tree built with the online tool popART (<http://popart.otago.ac.nz/index.shtml>) from the multiple alignment of the genes *fumC* and *fimH*. Each trace represents a mutation event. The nucleotide diversity was $\pi = 0.027$. The different colors, represented in the lower right corner of the figure represent the origin of the isolates, being: red (human origin); yellow (animal origin) and green (environmental origin).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	
215M02243	1	85.77	85.77	92.69	85.11	92.92	73.60	92.88	70.27	91.95	85.64	85.63	86.11	85.91	92.27	86.22	92.83	92.82	85.94	85.93	85.72	99.83	60.87	90.24	74.63	70.58	85.62	90.13	91.74	86.24	92.56	91.69	85.57	85.55	92.15	92.62	92.77	97.99	86.79	85.45	85.44	72.94	85.73	
249M02423	2	18183	100.00	85.97	89.93	86.07	73.40	86.02	70.01	85.91	90.86	90.84	89.20	91.71	86.27	92.28	86.48	86.11	91.10	91.39	91.41	99.58	60.79	86.60	74.32	70.31	90.59	86.66	85.95	88.53	86.20	86.19	90.75	90.74	85.55	86.07	85.95	86.04	90.50	90.39	91.41	72.82	91.11	
279M02423	3	18185	3	85.96	89.92	86.06	73.40	86.01	70.01	85.91	90.86	90.84	89.20	91.71	86.27	92.28	86.48	86.11	91.10	91.38	91.41	99.57	60.78	86.60	74.32	70.30	90.59	86.66	85.95	88.53	86.20	86.19	90.75	90.74	85.55	86.07	85.95	86.04	90.50	90.39	91.41	72.82	91.11	
649M05549	4	9336	17933	17935		85.46	98.87	73.76	99.01	70.34	92.14	85.93	85.92	86.42	86.11	92.54	92.67	92.41	92.73	86.25	86.64	85.89	59.92	60.99	90.26	74.70	70.65	85.86	89.87	91.94	86.52	92.66	91.74	85.74	85.74	92.80	93.26	98.90	92.74	87.15	85.77	85.80	73.12	86.13
651M05549	5	19010	12872	12875	18585		85.42	73.53	85.43	70.11	85.18	97.79	97.77	90.78	94.37	85.44	89.87	85.49	85.34	89.87	86.56	93.77	59.76	60.97	85.75	75.05	70.41	96.91	85.66	85.36	87.81	85.35	85.52	96.22	96.21	84.90	85.36	85.48	85.34	92.62	95.52	91.63	72.97	97.12
914M05549	6	9045	17805	17807	1438	18629		73.76	99.10	70.36	92.20	85.94	85.93	86.43	86.17	92.59	92.83	92.88	86.26	86.09	85.85	59.94	60.97	90.32	74.72	70.66	86.01	90.04	92.07	86.53	92.79	91.87	85.70	85.69	92.87	93.33	99.06	93.11	87.10	85.90	85.78	73.16	86.11	
08M06548	7	33711	33990	33991	33532	33830	33531		73.70	74.53	73.59	73.66	73.63	73.63	73.77	73.40	73.77	73.74	73.58	73.51	73.61	62.89	63.23	73.64	76.96	75.03	73.69	73.69	73.75	73.43	73.68	73.78	73.46	73.44	73.32	73.78	73.73	73.69	73.64	73.53	73.41	84.99	73.66	
DM118478	8	9101	17869	17871	1271	18618	1454	33607		70.36	92.23	85.97	85.96	86.34	86.11	92.71	86.50	92.56	92.86	86.26	86.08	85.79	59.89	60.96	90.36	74.69	70.67	85.82	89.90	92.13	85.49	92.89	91.89	85.72	85.71	92.89	93.42	98.28	92.96	87.05	85.79	85.80	73.09	86.15
ERR1010184	9	37993	38317	38319	37902	38202	37874	32551	37874		70.25	70.12	70.11	70.21	70.10	70.30	70.28	70.37	70.11	69.96	69.99	59.89	58.58	59.60	70.11	70.92	86.23	70.15	70.25	70.41	70.11	70.30	69.93	69.93	69.93	70.39	70.40	70.27	70.13	70.05	70.30	74.88	70.07	
ERR1010185	10	10289	16000	18003	10040	18918	9964	33751	9934	38012		85.73	85.71	86.17	86.02	96.76	76.08	76.20	92.18	92.61	85.98	85.85	59.81	60.84	90.15	74.58	70.49	85.77	89.46	91.20	86.25	92.60	91.21	85.64	85.63	91.86	92.41	92.07	92.27	86.90	85.55	85.59	72.94	85.91
ERR1010187	11	18350	13676	13679	17974	2830	17969	33678	17931	38185	18238		99.86	91.42	95.26	86.03	90.73	86.15	85.99	90.83	90.53	94.85	59.65	60.89	86.38	74.42	70.38	98.20	86.42	85.82	88.45	85.99	86.06	97.29	97.28	85.50	85.99	85.99	85.97	93.48	96.46	92.59	72.86	98.49
ERR1010188	12	18364	13699	13702	17991	2851	17986	33696	17948	38198	18255	182		91.40	95.24	86.02	90.72	86.13	85.98	90.82	90.51	94.84	59.64	60.89	86.37	74.41	70.37	98.18	86.40	85.81	88.43	85.98	86.04	97.28	97.26	85.50	85.98	85.98	93.47	96.44	92.58	72.85	98.47	
ERR1010215	13	17745	13800	13803	17346	11781	17338	33694	17452	38063	17675	10968	10989		91.76	86.33	85.97	86.60	86.23	90.13	89.01	91.36	59.86	61.05	86.69	74.64	70.50	91.04	86.73	86.28	89.36	86.34	86.43	91.13	91.12	85.80	86.23	86.33	86.31	91.02	91.08	89.54	73.11	91.58
ERR1010216	14	17999	10589	10592	17748	7196	17678	33691	17747	38203	17867	6060	6081	10533		86.19	91.24	86.36	86.06	91.52	91.08	91.39	59.59	61.03	86.57	74.41	70.33	94.89	86.58	85.88	88.91	86.12	86.21	95.47	95.46	85.66	86.17	86.00	86.09	96.05	93.79	92.21	72.88	95.55
ERR1010246	15	9871	17540	17543	9738	18609	9474	33515	9309	37839	4142	17849	17865	17467	17653		86.63	92.60	93.11	86.28	86.20	85.95	59.94	60.94	90.37	74.69	70.64	85.94	89.80	91.70	86.59	92.94	91.63	85.84	85.83	92.16	92.71	92.56	92.62	87.12	85.77	85.86	73.11	86.20
ERR1010247	16	17613	9862	9865	17029	12944	17216	33988	17254	38330	17531	11848	13864	13255	11192	17082		86.86	86.33	92.31	92.89	91.00	59.66	60.86	86.94	74.28	70.28	90.65	86.81	86.38	86.64	86.63	90.30	90.29	85.99	86.47	86.42	86.56	90.22	90.11	90.96	72.72	90.92	
ERR1010248	17	9157	17280	17283	9697	18542	9456	33520	9051	37973	9994	11702	17718	17223	17433	9457	16787		92.62	86.47	86.39	86.13	59.82	61.10	90.45	74.67	70.58	86.19	89.88	91.75	86.63	92.84	91.70	85.90	85.88	92.18	92.64	92.52	92.81	87.32	85.93	86.02	73.17	85.27
ERR1010252	18	9169	17753	17755	9286	18734	9092	33553	8994	37854	9437	17903	17920	17598	17813	8798	17464	9427		86.20	86.14	85.90	59.82	60.98	90.39	74.67	70.64	85.86	89.78	91.63	86.38	93.08	91.65	85.70	85.68	92.33	92.87	92.83	92.99	87.14	85.78	85.83	73.15	86.14
ERR1010254	19	17962	13369	13372	12586	12949	17559	33764	17555	38192	17915	11714	11734	12607	10832	17512	9821	17294	17640		92.85	85.92	91.17	61.03	86.89	74.36	70.30	90.48	86.38	88.70	86.41	86.31	90.50	90.49	85.61	86.14	86.18	86.25	90.49	85.80	85.83	75.03	97.03	
ERR1010266	20	17983	13006	13009	17841	13339	17779	33851	17785	38380	18012	12107	12126	13926	11392	17630	9093	17394	17716	9130		90.76	59.64	61.05	86.92	74.33	70.23	90.18	86.71	86.02	89.68	86.24	86.30	90.20	90.19	85.65	86.10	85.99	86.18	89.99	89.81	89.97	72.83	90.58
ERR1010277	21	18252	10977	10980	18024	7964	18084	33717	18161	38350	18214	6575	6598	11045	6501	17956	11497	17724	18015	11343	11813		59.70	61.09	86.23	74.58	70.25	94.90	86.28	85.83	88.38	85.94	86.05	95.84	95.83	85.18	85.86	85.71	86.98	93.50	93.68	91.71	73.04	95.27
ERR1010278	22	51325	51655	51658	51212	51421	51189	47415	51249	52932	51357	51556	51573	51291	51643	51190	51545	51216	51217	51489	51574	51000		81.21	59.85	63.11	59.15	59.69	59.81	59.98	59.66	59.86	59.81	59.58	59.56	59.93	59.89	59.85	59.71	59.72	59.52	63.11	59.63	
ERR1010280	23	50005	50109	50112	49854	49869	49872	49887	49883	51618	50043	49973	49989	49778	49803	49917	50015	49713	49861	49793	49767	49590	24012		61.07	61.07	63.11	60.20	60.99	60.92	61.07	60.86	60.87	60.99	60.95	60.93	60.49	60.95	60.96	60.86	60.76	60.88	63.38	60.83
ERR1010283	24	12472	17122	17125	12446	18204	13684	12317	37941	12591	17402	17418	17007	17156	12308	16687	12198	12282	16749	16715	17591	51310	49746		74.57	70.62	86.15	92.27	90.02	87.08	90.63	89.82	86.03	86.02	86.66	90.15	90.28	90.34	87.30	86.18	86.14	73.02	86.43	
ERR1010284	25	32418	32809	32813	31884	32305	29417	32341	37160	32474	32689	32700	32409	32697	32343	32871	32370</																											

A) Rate=9.92e+00,MRCA=1995.71,R2=0.97,p<1.00e-04



B)



C)

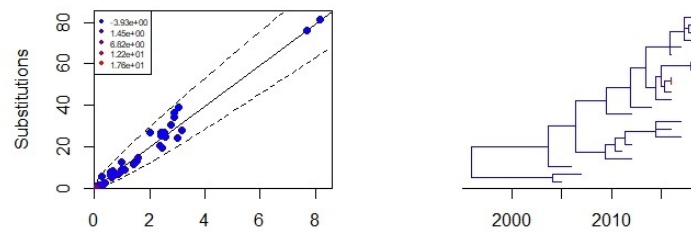


Figure S4: The estimated dates across the isolates *mcr-1* positive. The inference was performed with BactDating R package with MCMC chains of 1000000 generations using the strict gamma model. A) Linear regression of root-to-tip (y-axis) versus sampling dates (x-axis). B) Estimated dated genealogy. C) Branch-by-branch comparison of duration in years (x-axis) and number of observed substitutions (y-axis).

10. APÊNDICE

Tabela 1A: Porcentagem do genoma mapeado e número de variantes de nucleotídeo único SNV.

Amostra	Genoma mapeado (%)	N° de variantes detectadas
215M02423*	71	52.921
249M02423*	78	34.504
279M02423*	79	33.295
649M05549*	85	55.017
651M05549*	81	14.569
914M05549*	82	56.147
DRR065948	70	95.860
DRR138478	75	54.995
ERR1360256	77	13.872
ERR1407279	75	54.671
ERR3479038	78	52.921
ERR3479039	79	43.091
SRR10099927	82	56.346
SRR10099929	81	55.753
SRR10099933	69	14.739
SRR10099940	73	14.452
SRR10099942	71	58.492
SRR10099945	80	55.708
SRR10392137	75	54.003
SRR10521259	79	54.256
SRR5328644	81	28.010
SRR5350657	90	17.624
SRR6750074	77	31.851
SRR6791684	70	100.531
SRR6791870	86	11.776
ERR1010266	86	38.566
ERR1010278	82	144.842
ERR1010283	80	55.970
ERR1010284	79	95.118
ERR1010246	90	51.176
ERR1010252	88	54.282
ERR1010255	93	36.961
ERR1010277	86	22.427
ERR1010280	83	140.551
ERR1010215	90	33.608
ERR1010216	93	20.271
ERR1010285	89	103.764

ERR1010184	89	104.305
ERR1010185	80	56.597
ERR1010187	80	14.307
ERR1010188	81	13.906
ERR1010247	88	35.076
ERR1010248	90	52.312

Dados retirados da análise de mapeamento realizada no software CLC Genomic Workbench, disponível em: <https://digitalinsights.qiagen.com>. Os isolados sequenciados neste estudo e provenientes do Laboratório de Microbiologia Molecular e Saúde estão destacados com *.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43						
215M02423	1		85.77	85.77	92.69	85.11	92.92	73.60	92.88	70.27	91.95	85.64	85.63	86.11	85.91	92.27	86.22	92.83	92.82	85.94	85.93	85.72	59.83	60.87	90.24	74.63	70.58	85.62	90.13	91.74	86.24	92.56	91.69	85.57	85.55	92.15	92.62	92.77	97.99	86.79	85.45	85.44	72.94	85.73					
249M02423	2	18183		100.00	85.97	89.93	86.07	73.40	86.02	70.01	85.91	90.86	90.84	89.20	91.71	86.27	92.28	86.48	86.11	91.10	91.39	91.41	59.58	60.79	86.60	74.32	70.31	90.59	86.66	85.95	88.53	86.20	86.19	90.75	90.74	85.55	86.07	85.95	86.04	90.50	90.39	91.41	72.82	91.11					
279M02423	3	18185	3		85.96	89.92	86.06	73.40	86.01	70.01	85.91	90.86	90.84	89.20	91.71	86.27	92.28	86.47	86.11	91.10	91.38	91.41	59.57	60.78	86.60	74.32	70.30	90.59	86.66	85.94	88.53	86.20	86.19	90.75	90.74	85.55	86.07	85.95	86.04	90.50	90.39	91.41	72.82	91.11					
649M05549	4	9336	17933	17935		85.46	98.87	73.76	99.01	70.34	92.14	85.93	85.92	86.42	86.11	92.54	86.67	92.41	92.73	86.25	86.04	85.89	59.92	60.99	90.26	74.70	70.65	85.86	89.87	91.94	86.52	92.66	91.74	85.74	85.74	92.80	93.26	98.90	92.74	91.65	85.77	85.80	82.92	86.13					
651M05549	5	19010	12872	12875	18585		85.42	73.53	85.43	70.11	85.18	97.79	97.77	90.78	94.37	85.44	89.87	85.49	85.34	89.27	89.56	93.77	59.76	60.97	85.75	75.05	70.41	96.91	85.66	85.36	87.81	85.35	85.52	96.22	96.21	84.90	85.36	85.48	85.34	92.62	95.52	91.63	72.97	97.12					
914M05549	6	9048	17805	17807	1498	18629		73.76	99.10	70.36	92.20	85.94	85.93	86.43	86.17	92.59	86.53	92.80	92.88	86.26	86.09	85.85	59.94	60.97	90.32	74.72	70.66	86.01	90.04	92.07	86.53	92.79	91.87	85.70	85.69	92.87	93.32	99.06	93.11	87.10	85.90	85.78	73.16	86.11					
00R065948	7	33373	33990	33991	33552	33930	33531		73.70	74.53	73.59	73.65	73.63	73.63	73.77	73.40	73.77	73.74	73.58	73.51	73.91	62.89	63.23	73.64	76.98	70.67	85.82	89.80	92.13	86.49	92.89	91.89	85.72	85.71	92.89	93.42	99.28	92.96	87.05	85.79	85.80	73.99	85.15						
08R138478	8	9301	17869	17871	3274	18616	1154	33607		70.38	92.23	85.97	85.95	86.34	86.11	92.73	86.50	92.56	92.86	86.26	86.08	85.79	59.89	60.96	90.38	74.69	70.67	85.82	89.80	92.13	86.49	92.89	91.89	85.72	85.71	92.89	93.42	99.28	92.96	87.05	85.79	85.80	73.99	85.15					
ER1010184	9	37993	38317	38319	37902	38202	37876	32551	37876		70.25	70.12	70.11	70.21	70.10	70.39	70.00	70.28	70.37	70.11	69.96	69.99	58.58	59.60	70.31	70.92	86.23	70.15	70.25	70.41	70.11	70.26	70.30	69.93	69.91	69.93	70.39	70.40	70.27	70.13	70.05	70.10	74.88	70.07					
ER1010185	10	10289	18000	18003	10040	18938	9964	33751	9934	38012		85.73	85.73	86.17	86.02	96.76	86.28	92.18	92.61	85.98	85.90	85.75	59.81	60.84	90.15	74.58	70.49	85.77	89.46	91.20	86.25	92.60	91.21	85.64	85.63	91.86	92.41	92.07	92.27	86.90	85.55	85.59	72.94	85.91					
ER1010187	11	18350	11676	11679	17974	2830	17969	33678	17931	38185	18238		99.86	91.42	95.26	86.03	90.73	86.15	85.99	90.83	90.53	94.85	59.65	60.89	86.38	74.42	70.38	98.20	86.42	85.82	88.45	85.99	86.04	97.29	97.28	85.50	85.99	85.99	85.97	93.48	86.46	92.59	72.86	98.48					
ER1010188	12	18364	11699	11702	17991	2851	17986	33696	17948	38198	18255	182		91.40	95.24	86.02	90.72	86.13	85.98	90.82	90.51	94.84	59.64	60.88	86.37	74.41	70.37	98.20	86.40	85.81	88.43	85.98	86.04	97.28	97.26	85.50	85.98	85.98	85.95	93.47	86.44	92.58	72.85	98.47					
ER1010215	13	17745	13800	13803	17348	11781	17338	33694	17452	38063	17675	10968	10989		91.76	86.33	89.56	86.60	86.23	90.13	91.81	91.36	59.86	61.05	86.69	74.64	70.50	91.04	86.73	86.24	89.36	86.34	86.43	91.13	91.12	85.80	86.23	86.33	86.31	91.02	91.08	89.54	72.81	91.58					
ER1010216	14	17999	10589	10592	17748	7196	17678	33691	17747	38203	17667	6060	6081	10533		86.19	91.24	86.66	86.06	91.52	91.08	94.91	59.59	61.03	86.57	74.41	70.33	94.89	86.58	85.88	88.12	86.12	95.47	95.46	85.86	86.17	86.00	86.09	96.05	93.79	92.21	72.88	95.55						
ER1010246	15	9871	17540	17543	9538	18609	9474	33515	9309	37839	4342	17849	18965	17467	17653		86.63	92.60	93.11	86.28	86.20	85.95	59.94	60.94	90.37	74.69	70.64	85.94	89.80	91.70	86.59	92.94	91.63	85.84	85.83	92.16	92.71	92.56	92.62	87.12	85.77	85.86	73.11	86.20					
ER1010247	16	17613	9862	9865	17029	12944	17216	33988	17254	38330	17531	11843	11864	13325	11192	17082		86.86	86.33	92.31	92.89	91.00	59.66	60.86	86.94	74.28	70.28	90.65	86.81	86.38	89.16	86.64	86.30	90.29	85.99	86.47	86.42	86.56	90.22	90.11	90.96	72.72	90.92						
ER1010248	17	9157	17280	17283	9697	18542	9456	33520	9501	37973	9994	17702	17718	17123	17433	9457	16787		92.62	86.47	86.39	86.13	59.92	61.10	90.45	74.67	70.58	86.19	89.88	91.75	86.63	92.84	91.70	85.90	85.88	92.10	92.64	92.52	92.81	87.32	85.93	86.02	73.17	86.27					
ER1010252	18	9169	17753	17755	9748	17968	9092	33553	8994	37858	9437	17903	17920	17598	17813	9489	17644	9427		86.20	86.49	85.90	59.92	60.98	90.39	74.67	70.64	86.19	86.86	89.78	91.63	86.68	92.88	91.65	85.70	85.68	92.33	92.87	92.83	92.99	87.14	86.78	85.83	73.15	86.14				
ER1010255	19	17962	11369	11372	17566	12949	17559	33764	17555	38192	17915	11714	11734	12607	10832	17532	9821	17294	17640		92.85	91.12	59.71	61.03	86.89	74.36	70.30	90.48	86.73	86.08	89.70	86.41	86.31	90.50	90.49	85.61	86.14	86.18	86.25	90.49	90.38	90.50	72.87	91.03					
ER1010266	20	17983	11006	11009	17841	13339	17851	17785	38380	18012	12107	12126	13026	11392	17630	9083	17394	17716	9130		90.76	59.64	61.05	86.92	74.33	70.23	90.18	86.71	86.02	89.68	86.24	86.30	90.20	90.19	85.65	86.10	85.99	86.18	89.99	89.81	89.97	72.83	90.58						
ER1010277	21	18252	10977	10980	18024	7964	18084	33717	18161	38350	18214	6575	6598	11045	6501	17956	11497	17724	18015	11343	11813		59.70	61.19	86.23	74.58	70.25	94.90	86.28	85.83	88.38	85.94	86.05	95.84	95.83	85.38	85.86	85.71	85.98	93.50	93.68	91.71	73.04	95.27					
ER1010278	22	51325	51655	51658	51212	51421	51189	47415	51249	52932	51357	51556	51573	51291	51643	51190	51545	51217	51489	51574	51500		61.21	59.85	61.11	59.15	59.69	59.81	59.88	59.56	59.96	59.56	59.96	59.58	59.56	59.96	59.58	59.71	59.72	59.52	63.11	59.63							
ER1010280	23	50005	50109	50112	49854	49869	49872	46987	49883	51618	50043	49977	49989	49778	49803	49917	50015	49713	49861	49793	49574	49590	24012		61.07	63.74	60.20	60.99	60.92	61.07	60.86	60.87	60.99	60.95	60.95	60.96	60.95	60.96	60.85	60.86	60.76	60.88	60.86	63.38	60.83				
ER1010283	24	12472	17122	17125	12446	18204	12364	33884	12317	37941	12591	17402	17418	17007	17156	12308	16687	12283	16749	32801	32769	32343	32871	32370	32361	32769	32801	32488	47139	46332	32489		71.75	74.49	74.56	74.69	74.47	74.05	74.69	74.44	74.43	74.23	74.66	74.66	74.41	74.48	74.27	77.01	74.44
ER1010284	25	32418	32809	32810	32324	31884																																											