1    **Identification of genetic markers of resistance to macrolide class antibiotics in *Mannheimia***

2    ***haemolytica* isolates from a Saskatchewan feedlot**

3

4    Darien Deschner[a], Maarten J. Voordouw[a], Champika Fernando[a], John Campbell[b], Cheryl L.

5    Waldner[b], Janet E. Hill[a#]

6

7    [a] Department of Veterinary Microbiology, University of Saskatchewan, Saskatoon, Canada

8    [b] Department of Large Animal Clinical Sciences, University of Saskatchewan, Saskatoon,

9    Canada

10

11    Running head: Markers of macrolide resistance in *Mannheimia haemolytica*

12

13    [#]Correspondence: Janet.Hill@usask.ca

14

# Abstract

*Mannheimia haemolytica* is a major contributor to bovine respiratory disease (BRD), which causes substantial economic losses to the beef industry, and there is an urgent need for rapid and accurate diagnostic tests to provide evidence for treatment decisions and support antimicrobial stewardship. Diagnostic sequencing can provide information about antimicrobial resistance genes in *M. haemolytica* more rapidly than conventional diagnostics. Realizing the full potential of diagnostic sequencing requires a comprehensive understanding of the genetic markers of antimicrobial resistance. We identified genetic markers of resistance in *M. haemolytica* to macrolide class antibiotics commonly used for control of BRD. Genome sequences were determined for 99 *M. haemolytica* isolates with six different susceptibility phenotypes collected over two years from a feedlot in Saskatchewan, Canada. Known macrolide resistance genes *estT*, *msr*(E), and *mph*(E) were identified in most resistant isolates within predicted integrative and conjugative elements (ICEs). ICE sequences lacking antibiotic resistance genes were detected in 10/47 susceptible isolates. No resistance-associated polymorphisms were detected in ribosomal RNA genes, although previously unreported mutations in the L22 and L23 ribosomal proteins were identified in 12 and 27 resistant isolates, respectively. Pangenome analysis led to the identification of 79 genes associated with resistance to gamithromycin, of which 95% (75/79) had no functional annotation. Most of the observed phenotypic resistance was explained by previously identified antibiotic resistance genes, although resistance to the macrolides gamithromycin and tulathromycin was not explained in 39/47 isolates, demonstrating the need for continued surveillance for novel determinants of macrolide resistance.

## Importance

Bovine respiratory disease is the costliest disease of beef cattle in North America, and the most common reason for injectable antibiotic use in beef cattle. Metagenomic sequencing offers the potential to make economically significant reductions in turnaround time for diagnostic information for evidence-based selection of antibiotics for use in the feedlot. The success of diagnostic sequencing depends on a comprehensive catalog of antimicrobial resistance genes and other genome features associated with reduced susceptibility. We analysed the genome sequences of isolates of *Mannheimia haemolytica,* a major bovine respiratory disease pathogen, and identified both previously known and novel genes associated with reduced susceptibility to macrolide class antimicrobials. These findings reinforce the need for ongoing surveillance for markers of antimicrobial resistance to support improved diagnostics and antimicrobial stewardship.

## Introduction

Bovine respiratory disease (BRD) is a major concern in the feedlot industry where it is estimated to cause more than US$3 billion in losses each year globally (1). BRD is commonly controlled by treating high-risk groups of animals with antibiotics on arrival at the feedlot in a practice known as metaphylaxis (2, 3), which has been associated with selection of antimicrobial resistant bacteria (reviewed in (4)). Concerns about antimicrobial resistance (AMR) resulted in the passage of legislation increasing oversight of antimicrobial use in animal agriculture in Canada and the European Union (5–8). To meet the challenge of reducing the use of antibiotics and using antibiotics more prudently, veterinarians and producers need rapid and accurate diagnostic tests and sampling strategies to provide the best evidence for treatment decisions.

Diagnostic sequencing is the application of sequencing technologies in medical and veterinary diagnostics (9). The biggest benefit of diagnostic sequencing over traditional methods is the speed at which it can provide an answer, taking hours to get a result where traditional methods often take days. In the specific case of detecting AMR, the reduced turnaround time is due to bypassing the requirement for time-consuming culture and antimicrobial susceptibility testing of bacteria in the laboratory. Identifying genetic markers of resistance allows veterinarians to make evidence-based decisions about appropriate antibiotics, while avoiding prescribing ineffective antibiotics. Realizing the potential of diagnostic sequencing requires a comprehensive and informative database of all genetic markers of resistance (genes and mutations) in pathogens of interest.

*Mannheimia haemolytica* is the major bacterial agent associated with BRD in North America (10) and macrolide class antibiotics are often the choice for use in metaphylaxis or as

72    first-line therapeutics, which has resulted in the evolution of increased resistance to these drugs

73    (11–13). A number of macrolide resistance genes and mutations have been identified in *M.*

74    *haemolytica*, however, the recent characterization of a macrolide esterase encoded on a plasmid

75    carried by a *Sphingobacterium faecium* isolate recovered from a feedlot water trough, and widely

76    distributed among animal pathogens including *M. haemolytica*, highlights the need for further

77    studies to expand our knowledge of the genetic markers of macrolide resistance in cattle-

78    associated bacteria (14–16). Antimicrobial resistance genes (ARGs) in *M. haemolytica* are

79    frequently carried within mobile genetic elements known as integrative and conjugative elements

80    (ICEs) that are capable of spreading ARGs to related species such as *Pasteurella multocida* (17,

81    18). In addition to ARGs, single nucleotide polymorphisms (SNPs) in the genes encoding the

82    cellular targets of macrolides (ribosomal RNAs and ribosomal proteins) have also been known to

83    confer macrolide resistance in *M. haemolytica* and other species (19).

84          In the current study, we analysed the whole genome sequences of macrolide resistant and

85    susceptible *M. haemolytica* isolates collected over two years from a feedlot in Saskatchewan,

86    Canada, to identify the genetic markers of macrolide resistance. Our objective was to identify

87    known AMR genes and mutations, and to identify potentially novel markers associated with

88    macrolide resistance with the goal of enhancing *M. haemolytica* databases used in diagnostic

89    sequencing.

## Materials & Methods

90

### Selection of *M. haemolytica* isolates

91

92          Study isolates were selected from an existing collection of *M. haemolytica* isolates (20).

93    The collection was derived from deep nasopharyngeal swabs collected from 800 cattle at the

94    Livestock and Forage Centre of Excellence (LFCE) in the fall of 2020 and in the fall of 2021 at

95    three time points: on arrival, 13 days after arrival, and 36 days after arrival. Animals were

96    recently weaned steers of various beef breeds sourced from a local auction market and

97    originating from a combined total of 292 unique farming operations (20). Samples were cultured

98    to select for *M. haemolytica*, identification was confirmed by matrix-assisted laser

99    desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS; Bruker Daltonics,

100    Billerica, MA), and susceptibility of one *M. haemolytica* isolate per animal was determined using

101    the Sensititre system (Bovine/Porcine BOPO7F Vet AST Plate) at Prairie Diagnostic Services

102    (PDS, Saskatoon). Isolates were classified as either resistant, intermediate, or susceptible in

103    accordance with the breakpoints established by the Clinical Laboratory Standards Institute

104    (Table S1, (21)). Isolates classified as intermediate were categorized as resistant (non-wild type)

105    in all subsequent analyses.

106        We selected 52 macrolide resistant and 47 susceptible isolates for further study. Fifteen

107    isolates were selected in the fall 2020 collection period with one collected on arrival and the

108    remaining fourteen at 13 days post-arrival. Eighty-four isolates were selected in the fall 2021

109    collection period with three collected on arrival, 49 at 13 days post-arrival, and 21 at 36 days

110    post-arrival, two from sick animals, one from a dead animal, and eight from a collection of

111    isolates selectively cultured on 5% sheep blood agar plates supplemented with 16 µg/mL of

112    tulathromycin. Six of the selectively cultured isolates were derived from samples collected 36

113    days post-arrival, one from a sick animal and one from a dead animal. Antimicrobial

114    susceptibility testing of the selectively cultured isolates was performed as described above. To

115    maximize diversity, each of the selected isolates was from an animal sourced from a different

116    farm of origin and whenever possible we selected a susceptible isolate from the same time point

117 and pen for each resistant isolate. The 52 selected resistant isolates had a variety of phenotypes,

118 although all were resistant to at least one of the four tested macrolides: gamithromycin (GAM),

119 tulathromycin (TULA), tildipirosin (TILD), and tilmicosin (TILM). Many of the isolates were

120 also resistant to tetracycline (TET). We included the results for tetracycline resistance in our

121 analysis due to its common use in feedlots for the treatment and prevention of BRD and the

122 strong association of tetracycline resistance genes with ICEs (11).

123 Genomic DNA Extraction

124    Isolates were cultured aerobically on 5% sheep blood agar plates at 37°C for 24 hr, and

125 one colony was selected and grown aerobically in 7 mL of brain heart infusion (BHI) broth at

126 37°C for 24 hr with shaking at 200 rpm. Two 3 mL aliquots of each broth culture were pelleted,

127 and genomic DNA was then extracted using a salting-out protocol as previously described (22).

128 DNA concentration of extracts was determined using spectrophotometry (Nanodrop 2000) and

129 electrophoresis in 0.8% agarose was used to assess fragmentation and to detect RNA

130 contamination. Extracts that were found to have visible RNA contamination were then incubated

131 with 5 μL of RNase A (100 mg/mL) at 37°C for 1 hour, followed by the sodium acetate and

132 ethanol precipitation steps of the salting-out protocol.

133    To confirm the identity of the genomic DNA as *M. haemolytica,* cpn60 PCR and

134 amplicon sequencing were performed using either universal cpn60 PCR primers (23) or *M.*

135 *haemolytica* specific cpn60 primers (Forward: GCG ATT GTA AAC GAG GGC TT; Reverse:

136 TGA CCG ACT GTT GAA TCT GAG). Sequencing was performed by the Sanger method and

137 the forward and reverse primer reads were combined using the pregap4 and gap4 software to

138 produce a consensus sequence that was aligned to the cpnDB database (24) using the FASTA-

7

139    based sequence comparison tool. Sequence identity of at least 90% over the length of the

140    amplicon sequence was the minimum threshold for identification as *M. haemolytica*.

141        As a final quality check, all extracts were assessed for DNA concentration immediately

142    prior to Illumina library preparation using fluorometry (Qubit dsDNA Broad Range Assay Kit,

143    Invitrogen). A minimum DNA concentration of 20 ng/µL was required for an extract to be

144    included in Illumina library preparation.

145    Whole Genome Sequencing

146        Illumina sequencing libraries were prepared using the Nextera XT Library Prep Kit

147    (Illumina) according to the manufacturer's instructions. Libraries were first diluted to a DNA

148    concentration of 0.2 ng/µL. Following indexing and purification, fragment size distribution in

149    prepared libraries was assessed (Bioanalyzer 2100, Agilent) and quantified using the Qubit ds

150    Broad Range assay kit (Invitrogen). Libraries were manually normalized to a concentration of 4

151    nM and then pooled and diluted to 8 pM with a final concentration of 1% PhiX DNA and

152    sequenced using the 500 cycle MiSeq Reagent Kit v2 (2 x 250 bp) on the Illumina MiSeq

153    platform. Up to 24 genomes were sequenced per flow cell.

154        Isolates were also sequenced on a GridION platform using the ligation sequencing kit

155    (SQK-LSK109) and the FLO-MIN106 flow cell (PDS, Saskatoon). MinKnow v 22.03.2 was the

156    operating system and Guppy v 6.06 was used for base calling. A Qscore cut-off threshold of 9

157    was utilized for all reads, with those below the cut-off being excluded from the sequencing

158    output.

## Whole Genome Sequence Assembly

159 Whole Genome Sequence Assembly

160    Raw, demultiplexed, Illumina reads were processed using Trimmomatic with a sliding

161 window of 4, quality score threshold of 20 and minimum sequence length of 36 bp (25). Raw

162 nanopore reads were processed using Filtlong v 0.2.1 with a minimum length of 1,000 bp.

163 Quality-filtered reads from both Illumina and Nanopore sequencing were utilized for hybrid *de*

164 *novo* whole genome assembly using Unicycler v 0.5.0 with the default settings (26). Whole

165 genome sequences were then assessed for quality using seqkit v 2.3.0 (27) and annotated using

166 Prokka v 1.14.5 (28).

## Identification of Antibiotic Resistance Genes

167 Identification of Antibiotic Resistance Genes

168    Whole genome assemblies were screened for known antibiotic resistance genes and

169 mutations using the Resistance Gene Identifier (RGI) algorithm provided by the Comprehensive

170 Antibiotic Resistance Database (CARD) using the default settings and database v 3.2.5 (29, 30).

171    The macrolide esterase *estT* was not present in the utilized version of the database, and an

172 alternative approach was therefore taken to examine the whole genome assemblies for the

173 presence of this gene. In Geneious Prime v 2023.2.1, a BLAST database was constructed

174 consisting of the whole genome assemblies of the study isolates. The reference sequence for *estT*

175 was then downloaded from the reference sequence for *Sphingobacterium faecium* strain WB1

176 plasmid pWB1 (CP094932.1) and compared against this database using the Geneious built-in

177 BLAST software to identify homologous regions of the genome assemblies.

## Identification of 23S rRNA, L4 and L22 mutations

178 Identification of 23S rRNA, L4 and L22 mutations

179    Barrnap v 0.9 (31) was used to extract all copies of 23S rRNA gene sequences from *M.*

180 *haemolytica* genome assemblies and then 23S rRNA gene sequences were aligned using snippy v

181    4.6.0 (32) to the reference sequence from *M. haemolytica* USDA-ARS-USMARC-183 (Genbank

182    Accession NR_103087) with nucleotide numbering according to this sequence.

183        For the L4 and L22 ribosomal proteins, two custom shell scripts were utilized that

184    extracted the L4 and 22 ribosomal protein amino acid sequences from the Prokka annotations

185    (Supplemental File 1). The ribosomal protein amino acid sequences were then aligned to the

186    reference sequences from *M. haemolytica* strain USMARC_2286 (L4: AGR7646 and L22:

187    AGR73803) using BLASTp and any amino acid differences were recorded manually.

## Pangenome Analysis

189        All whole genome sequences were submitted to Roary analysis software (v 1.7.7)

190    alongside a reference strain for the two *M. haemolytica* genotypes 1 and 2 (genotype 1=

191    CP017518; genotype 2= CP017519) (33). Gene clusters were sorted into one of four categories:

192    core, soft-core, cloud and shell, based on how many of the isolates in the examined dataset the

193    gene cluster appeared in. With a total input of 101 genomes (2 references and 99 study isolates),

194    core genes were those found in 99-101 isolates, soft core in 95-98 isolates, shell in 15-94 isolates

195    and cloud in less than 15 isolates. A multiFASTA alignment of core genes was produced using

196    MAFFT using the -e and -n flags. The roary_plots.py python script was then utilized to generate

197    three outputs: a pangenome matrix, a frequency plot, and a pie chart. A phylogenetic tree was

198    additionally generated using the core gene multiFASTA alignment as the input for FastTree (34).

199    The genotype of each isolate was inferred based on proximity to the genotype 1 and 2 reference

200    sequences in the phylogenetic tree.

## Whole Genome SNP Analysis

To reduce the impact of assembly errors in the identification of SNPs, all whole genome sequences were polished using Pilon v 1.24 (35). Each genome underwent a minimum of five rounds of pilon polishing prior to whole genome SNP analysis following the method described in (36). In brief, genomes were indexed using Hisat2 v 2.2.1 (37) to produce a SAM file, which was converted to a BAM file, and then sorted and indexed using Samtools v 1.17 (38) with the resulting BAM file being used as input to Pilon.

For whole genome SNP analysis, we utilized the snippy-multi function of snippy v 4.6.0 to generate a .VCF file for each analyzed genome containing all the identified SNPs (32). To reduce the number of SNPs that were related to genotype, we chose to analyze the two genotypes of *M. haemolytica* separately and used a susceptible isolate of either genotype as the reference genome (37787 S-06 for genotype 1 and 34455 S-12 for genotype 2). Finally, pyseer v 1.3.10 was used to identify SNPs that were associated with phenotypic resistance to macrolide class antibiotics for each of the individual genotypes using a p-value cut-off of 0.05, a maximum allele frequency equivalent to the proportion of phenotypically resistant isolates included in the analysis (0.5 for genotype 1 and for 0.54 for genotype 2), and the elastic net whole genome model (39, 40). The odds ratio of each SNP was then determined using Microsoft Excel for Microsoft 365 MSO (Version 2308 Build 16.016731.20052) 64-bit. SNPs that possessed an odds ratio lower than 0.8 or greater than 1.2 were considered as potentially explanatory for the observed phenotype.

## Association of Phenotypic Macrolide Resistance with Known AMR Genes and Candidate Genes

Generalized linear models (GLMs) with binomial errors were used to confirm the associations between known AMR genes and phenotypic resistance to each of the five antimicrobial drugs (GAM, TULA, TILD, TILM, TET) using the glm() function in R (41). Resistance to each of the five drugs was coded as a binomial variable as susceptible (0) or resistant (1). The known AMR genes *tet*(H), *estT*, APH(3')-1a, APH(3")-1b, and APH(6)-1d were present in the same 44 isolates. Similarly, the known AMR genes *msr*(E) and *mph*(E) were present in the same eight isolates. As statistical analyses cannot include identical explanatory factors, only *tet*(H) and *msr*(E) were included in the GLMs (i.e., the other 5 genes were not included in the analysis because they were redundant with *tet*(H) and *msr*(E)). Genotype was also included as an explanatory factor as it had previously been shown to be associated with phenotypic resistance (42). For these GLMs, a p-value $\leq 0.05$ was significant.

The same GLM approach was used to test the association between phenotypic resistance and the candidate genes that were identified in the pangenome analysis. The sign of the slope relating AMR to the presence/absence of a gene indicates whether the gene confers resistance (positive slope) or susceptibility (negative slope) to a given drug. For these GLMs, a Bonferroni-corrected p-value was used where 0.05 was divided by the number of candidate genes. Both unifactorial and multifactorial GLMs were performed. The unifactorial GLMs only included the candidate gene of interest, whereas the multifactorial GLMs also included genotype, *tet*(H), and *msr*(E).

### Protein structure predictions and alignments

Amino acid sequences were aligned with MUSCLE in Geneious Prime using default settings. Secondary structures were predicted with PSIPRED and DISOPRED3, and protein structure prediction from primary amino acid sequences was performed with DMPFold within the PSIPRED Protein Analysis Workbench (43). Structure visualization and alignments were generated with Pymol (Schrodinger, Inc.).

### Mobile Genetic Element Analysis

All 99 genome sequences of the study isolates were submitted to the online web tool implementation of ICEfinder for identification of putative ICEs (44). The sequences of identified putative ICEs and integrative and mobilizable elements (IMEs) were downloaded and information such as the insertion site and direct repeat sequence were recorded in an Excel spreadsheet. Putative ICE sequences that were found to possess ≤ 10 type 4 secretion system (T4SS) components and/or lacked an identified insertion site were removed from further analysis. Manual curation and comparisons of the putative ICE sequences were done in Geneious Prime. To determine if the identified putative ICE sequences were novel or homologues of the previously identified ICE*Mh1* (45), the representative consensus sequence for each ICE was aligned to ICE*Mh1* from *M. haemolytica* 42548 using BLASTn (ICEberg database ID ICEO_0000728) (46).

## Results & Discussion

### Description of Isolates

Study isolates were initially characterized by their antimicrobial resistance phenotype resulting in six groups (Table 1, Table S2).

## Genome sequencing and assembly

The average length of the assembled genomes (n = 99) was 2.9 Mb (range 2.5 Mb – 5.2 Mb) with an average N50 score of 2.6 Mb (range 0.6 Mb – 2.7 Mb). The median number of contigs per assembly was 2 (range 1 – 26). The average estimated coverage for assemblies was 158× (median 131×, range 23× – 578×). Isolate 38241 S-35 was not included in the whole genome SNP analysis due to insufficient coverage with high quality Illumina reads. Based on phylogenetic analysis of core genomes identified in the pangenome analysis described below, most isolates (82/99) were identified as genotype 2, and the remainder (17/99) as genotype 1 (Figure S1). A detailed summary of the characteristics of the genome assemblies is provided in Table S3.

## Antibiotic Resistance Gene Identification

Antibiotic resistance genes were identified in each whole genome sequence utilizing the RGI software provided by CARD, where "perfect" hits were considered to represent the presence of a given gene, and "strict" hits were considered to represent the presence of a gene providing the resulting hit had a percent identity of at least 90% and a percent length between 99-101%, as calculated by dividing the length of the query protein by the length of the reference protein (Table 2). This analysis revealed two distinct genetic profiles that together accounted for 94% (49/52) of the phenotypically macrolide resistant isolates. One genetic profile consisted of the known macrolide resistance genes *msr*(E) and *mph*(E), which encode a ribosomal protection protein and a macrolide inactivating phosphotransferase, respectively (47, 48), and the sulfonamide resistance gene *sul2*. The other genetic profile included the tetracycline resistance gene *tet*(H) and its regulator *tetR*, the aminoglycoside resistance genes APH(3")-1b, APH(3')-1a, and APH(6)-1d, and *sul2*.

14

287     Resistance genes *msr*(E) and *mph*(E) were identified in only 8/52 isolates, all of which

288     had the GAM/TULA resistance phenotype (Table 2). Monomethyltransferase encoding gene

289     *erm42*, previously associated with macrolide resistance (16), was not detected in any isolate. In

290     all the remaining macrolide resistant isolates, the *tet*(H) gene was detected alongside the three

291     aminoglycoside resistance genes APH(3")-1b, APH(3')-1a, and APH(6)-1d (Table 2). The

292     presence of *tet*(H) has been strongly associated with the presence of ICEs, and the identified

293     aminoglycoside resistance genes have frequently been identified on ICE*Mh1* (17, 18, 49, 50).

294     The sulfonamide resistance gene *sul2* was identified in all but three phenotypically macrolide

295     resistant isolates, which may reflect *sul2* having been incidentally selected for alongside

296     macrolide resistance as part of the same mobile element. Regardless of the origin of the *sul2*

297     gene, CLSI presently lacks a breakpoint for sulfadimethoxine for *M. haemolytica* so no

298     determination on the resistance profile of the isolates was made. No known antibiotic resistance

299     genes were identified in the phenotypically macrolide susceptible isolates (Table 2).

300     Macrolide esterase gene *estT* was detected in all resistant isolates except for those with

301     the GAM/TULA resistance phenotype. *estT* was consistently annotated by Prokka as *rdmC*, a

302     methylesterase involved in the biosynthesis of the anthracycline antibiotic aclacinomycin (51).

303     The identification of *estT* offers a partial explanation for the observed phenotypic resistance,

304     however, it has been demonstrated biochemically that the enzyme encoded by *estT* has no

305     activity on gamithromycin (GAM) or tulathromycin (TULA) (14). Thus, its presence fails to

306     explain the observed resistance to these drugs in the GAM/TULA/TET/TILM/TILD,

307     TULA/TET/TILM/TILD, and GAM/TULA/TILM isolates. We used GLMs with binomial errors

308     to test whether genotype, *tet*(H) (identical presence/absence pattern to *estT*), and *msr*(E)

309     (identical presence/absence pattern to *mph*(E)) were associated with phenotypic resistance to

15

310   each of the five individual antimicrobial drugs (Table 3). The unifactorial GLM found that

311   genotype was significantly associated with resistance to TET, TILD, and TILM, and that *tet*(H)

312   and *msr*(E) were significantly associated with phenotypic resistance to all five tested

313   antimicrobials. The results changed when all three explanatory factors were included in the

314   multi-factorial GLM; *tet*(H) was significantly associated with phenotypic resistance to all five

315   tested antimicrobials, and *msr*(E) was significantly associated with phenotypic resistance to

316   GAM and TULA, consistent with the observations in Table 2. The association of *tet*(H) with

317   macrolide resistance is almost certainly due to the strong linkage of *tet*(H) with ICEs and not due

318   to any role of the *tet*(H) gene product in reduced susceptibility to macrolides. In the

319   multifactorial GLM, genotype was not associated with phenotypic resistance to any antimicrobial

320   drugs. The association between genotype and resistance to TET, TILM, and TILD in the

321   unifactorial GLM was probably driven by differences in the prevalence of *tet*(H) in genotypes 1

322   and 2 (0.0% (0/17) and 52% (44/84), respectively). The lack of a relationship between genotype

323   and resistance to any of the antibiotics tested somewhat contradicts previous reports that

324   genotype 2 isolates are more likely to carry known antibiotic resistance genes (42), however, the

325   limited number of genotype 1 isolates included in this study means this lack of association

326   should be interpreted with caution.

327   23S rRNA, L4 and L22 polymorphism identification

328          Mutations in the cellular targets of macrolides, 23S rRNA and ribosomal proteins L4 and

329   L22, have previously been associated with reduced susceptibility (47, 52). Six copies of the 23S

330   rRNA gene were identified in each of the study isolates and no resistance associated

331   polymorphisms were identified relative to a reference sequence (NR_103087) (Table S3). No

332   resistance associated differences were identified in the amino acid sequences of the L4 ribosomal

333 protein, however, in the L22 ribosomal protein, a D94A amino acid difference occurred in 23%

334 (12/52) of the macrolide resistant isolates and none of the susceptible isolates (Table S3). Of the

335 twelve isolates with the D94A L22 variant, all were genotype 2, and seven isolates had the

336 TULA/TET/TILM/TILD resistance phenotype, three isolates had the TET/TILM/TILD

337 phenotype, and two isolates had the GAM/TULA/TET/TILM/TILD phenotype.

338       Among the 12 isolates with the L22 ribosomal protein D94A sequence variant, the only

339 commonality was phenotypic resistance to the 16-membered lactone ring macrolides TILM and

340 TILD, which suggests this variant may alter the binding affinity of 16-membered lactone ring

341 macrolides but not the binding affinity of gamithroymycin (GAM, a 15-membered lactone ring

342 macrolide) or tulathromycin (TULA, a triamilide macrolide comprising a 90:10 mixture of 15-

343 membered lactone ring and 13-membered lactone ring regioisomers (53)).

344 Whole Genome SNP Analysis

345       To identify potentially novel sequence variants associated with reduced macrolide

346 susceptibility, we conducted a genome-wide SNP analysis of genotype 1 and genotype 2 isolates.

347 A total of 325 sequence variants (SNPs, insertions, or deletions) were initially identified as

348 having a significant association ($p \leq 0.05$) with observed phenotypic macrolide resistance (any of

349 the phenotypes in Table 1) in the genotype 1 isolates, however, all odds ratios were between 0.8

350 – 1.2. For the examined genotype 2 isolates, a total of 1,031 sequence variants were identified,

351 with all odds ratios falling between 0.8 – 1.2. These results suggest that SNPs are not an

352 explanation for the observed phenotypic macrolide resistance (Tables S4 and S5).

## Pangenome Analysis

A total of 101 isolates were included in the pangenome analysis, 99 study isolates and two reference strains representing genotype 1 and genotype 2. A total of 8,325 gene clusters were identified in the pangenome with each isolate possessing an average of 2,765 (median = 2,766) gene clusters (range = 2,486 – 5,181) (Table S6). Of the 8,325 gene clusters identified, 1,651 were identified as core genes (i.e., were present in 99-101 of the 101 examined isolates), which were used in a phylogenetic analysis to identify isolates as genotype 1 (17 isolates) or genotype 2 (82 isolates) based on their distance to and from the two reference genomes (Figure S1).

To discover novel AMR genes, we used the same GLM approach for each of the five macrolide antimicrobials, as described previously. The unifactorial GLM only included the candidate gene of interest as the single explanatory factor. The multifactorial GLM included the candidate gene of interest, but also genotype, *tet*(H) and *msr*(E) as explanatory factors. Of the 8,135 genes identified in the pangenome, 1,413 were part of the core genome (i.e., they were present in all isolates), and these genes were not analysed for their association with AMR. Of the remaining 6,912 genes, 24 genes were excluded because they had the same presence/absence profile as *tet*(H), *msr*(E), or genotype, so that 6,888 candidate genes remained for GLM analysis (Table S7). A Bonferroni-corrected p-value of $0.05/6,888 = 7.26 \times 10^{-6}$ was used to determine whether a given candidate gene was significantly associated with phenotypic resistance to each of the five macrolide antimicrobials.

The unifactorial GLMs found 1,223 candidate genes that were significantly associated with phenotypic resistance to one or more of the tested macrolides (Table S8). In comparison, the multifactorial GLMs found only 150 candidate genes that were significantly associated with

375  resistance to GAM and no candidate genes that were significantly associated with resistance to

376  any of the other four examined macrolide antimicrobials (Table S9). Thus, most of the genes in

377  the unifactorial GLM were significant only because they were associated with either *tet*(H)

378  and/or *msr*(E) in the genome of *M. haemolytica*. Below we further investigate the 150 genes

379  identified in the multifactorial GLMs.

380      Of the 150 GAM-resistance associated genes, 79 were classified as putative resistance

381  genes (GLM slope was positive) and 71 as putative susceptibility genes (GLM slope was

382  negative). Of the 71 putative susceptibility genes, five had functional annotations: *clpP_2*

383  (protease subunit), *mutS* (DNA mismatch repair protein), *rplW* (ribosomal protein L23), *ssb_4*

384  (single-stranded DNA binding protein), and *tldD* (metalloprotease) (54–58), and the remainder

385  were annotated as hypothetical/putative proteins.

386      Of the 79 putative resistance genes, four had functional annotations: *glpX_1* (fructose-

387  1,6-bisphosphatase), *mnmE* (tRNA modification GTPase), *rusA* (endonuclease that corrects

388  defects during genetic recombination and some DNA repair), *ybcO* (putative nuclease) and the

389  remainder were annotated as hypothetical/putative proteins (59–63). One possible connection to

390  macrolide resistance is that the *mnmE*-encoded GTPase may introduce a tRNA modification that

391  prevents inhibition of protein translation by macrolide class antibiotics, however, as the motifs

392  that macrolide class antibiotics recognize are in mRNA, this seems unlikely (64). Similarly, it is

393  possible that the DNA repair activity of the *rusA*-encoded endonuclease may result in sequence

394  changes that alter or eliminate macrolide activating motifs (64).

395      Further examination of the candidate genes identified in the GLM analysis revealed that

396  the genes *mnmE*, *mutS*, *tldD*, *rplW* and *glpX_1* had a variant with the inverse association (i.e.,

397　associated with susceptibility). These variants were identified as distinct genes rather than

398　variants due to the nature of the pangenome tool used (Roary). In identifying genes shared across

399　multiple genomes, an all-vs-all comparison of all coding sequences is conducted and based on a

400　defined threshold for "same", coding sequences are grouped into genes and their distribution

401　among genomes is reported in the gene presence-absence table output from Roary. If multiple

402　variants of a gene are sufficiently different from each other so that they exceed the threshold for

403　clustering, they will be assigned to different gene groups. In this case, group_753 is a variant of

404　*mnmE*, group_508 for *mutS*, group_2294 for *tldD*, group_4230 for *rplW*, and group_452 for

405　*glpX_1*. Of particular interest is the *rplW* gene which encodes the L23 ribosomal protein, a

406　protein that is in close proximity to the nascent polypeptide exit tunnel of the ribosomal complex

407　(57). As mutations in other ribosomal proteins that surround the exit tunnel (L4 and L22) are

408　known to confer macrolide resistance, we further investigated the amino acid sequences of both

409　*rplW* variants and determined that the group_4230 variant has a 14 amino acid insertion relative

410　to the *rplW* sequence identified in susceptible isolates (Figure 1A). To determine the location of

411　the insertion within the three-dimensional structure of the protein, we aligned the *M. haemolytica*

412　*rplW* and group_4230 sequences with the L23 sequence of *Thermus thermophilus*, for which a

413　structure has been determined (65). The protein consists of a well-ordered part containing anti-

414　parallel beta-strands and a large loop that forms part of the peptide exit tunnel wall (65). The

415　predicted structure of the susceptibility-associated *rplW*-encoded L23 aligns well with *T.*

416　*thermophilus* L23 but the insertion in the group_4230 variant nearly doubles the size of the loop

417　(Figure 1B). This group_4230 variant of L23 was only identified in the 27 isolates with the

418　GAM/TULA/TET/TILM/TILD resistance phenotype, suggesting that alterations in the structure

419    of the peptide exit tunnel may contribute to the observed resistance to gamithromycin and/or

420    tulathromycin in these isolates.

421    Mobile Genetic Element Analysis

422         ICEfinder initially identified 169 putative ICEs among the examined genome sequences

423    (Table S10). After filtering the ICEfinder results as described in the methods, a total of 68

424    putative ICE sequences were identified across 66/99 (66%) of the isolates. At least one putative

425    ICE sequence was identified in 19/47 (40%) of the susceptible isolates and 47/52 (90%) of the

426    phenotypically resistant isolates. Manual examination of these sequences allowed us to

427    consolidate them into three putative ICEs (ICE*Mh1*-like01, ICE*Mh*GAMTULA, and

428    ICE*Mh*Susceptible) based on shared direct repeat sequences, and number and types of cargo

429    genes. For 37 of the putative ICEs, ICEfinder found an insertion site in a ribose-phosphate

430    pyrophosphokinase gene, however, manual inspection of the flanking regions of this gene

431    showed that these ICEs were in proximity to a tRNA-Leu gene, which is the typical insertion site

432    for ICEs in *M. haemolytica* (18, 66, 67). On this basis, the ICE sequence was extended for all 37

433    of these putative ICEs to include the section between the boundaries initially identified by

434    ICEfinder and the adjacent tRNA-Leu.

435         The sequences of all putative ICEs categorized as ICE*Mh1*-like01 were extracted and a

436    multiple sequence alignment of these sequences was performed to generate a consensus

437    sequence. This process was repeated for ICE*Mh*GAMTULA and ICE*Mh*Susceptible (Figure 2),

438    and the distribution of the ICEs among the study isolates was assessed (Table 4). The failure to

439    identify any putative ICEs in seven of the resistant isolates may be attributable to lower assembly

21

440   quality for these isolates as all the genome sequences for these isolates consisted of six or more

441   contigs.

442       ICE*Mh1*-like01 was identified in 40/52 of the macrolide resistant isolates (Table 4). This

443   ICE carried the known ARGs including APH(3')-1a, APH(3")-1b, APH(6)-1, *estT*, *sul2*, *tet*(H),

444   and a gene encoding the *tet*(H) regulator TetR, as well as the copper resistance gene *mco*.

445   ICE*Mh1*-like01 was nearly identical in gene arrangement and sequence of the ARGs to the

446   previously identified *M. haemolytica* ICE*Mh1* in the region containing the ARGs. The portion of

447   ICE*Mh1*-like01 that failed to match to ICE*Mh1* lacked any of the known ARGs, however this

448   region did include genes encoding a MobH relaxase, an integrase and two transposases, which

449   explains why this region was identified by ICEfinder as being part of a putative ICE (Figure 2,

450   Figure S2).

451       ICE*Mh*GAMTULA was found in 7/8 of the isolates resistant to gamithromycin and

452   tulathromycin and was found to carry the known macrolide ARGs *msr*(E) and *mph*(E) alongside

453   the sulfonamide resistance gene *sul2*. The *msr*(E) and *mph*(E) genes have previously been

454   identified in ICE*Mh1*-like ICEs, however, this has always been alongside other ARGs such as

455   *tet*(H) (17, 18, 67). Further, ICE*Mh*GAMTULA was found to share relatively little sequence

456   similarity or synteny with ICE*Mh1*, except for a region containing genes encoding a MobH

457   relaxase, an integrase, and several ICE-associated genes (Figure 2). The similarity of

458   ICE*Mh*GAMTULA to ICE*Mh1* in this region suggests that ICE*Mh*GAMTULA is related to

459   ICE*Mh1* but has since lost other genes carried by ICE*Mh1*. It is also possible that

460   ICE*Mh*GAMTULA is not an ICE but rather an integrated bacteriophage or an integrative and

461   mobilizable element (IME) (44, 68), a suggestion supported by the limited number of T4SS

462 components it carries compared to ICE*Mh1* (9 vs 14), which potentially inhibits the ability of

463 ICE*Mh*GAMTULA to undergo autonomous conjugative transfer.

464 ICE*Mh*Susceptible does not carry any known ARGs but does contain genes encoding a

465 MobH relaxase and an integrase, and some ICE-associated genes that are similar to the

466 corresponding genes in ICE*Mh1* in both arrangement and sequence identity (Figure 2).

467 ICE*Mh*Susceptible also included the same direct repeat sequence as ICE*Mh1*-like01 and the

468 insertion site of tRNA-Leucine. Due to the similarity in the "core" ICE genes to ICE*Mh1* and the

469 complete absence of ARGs, it is possible that ICE*Mh*Susceptible represents a degraded version

470 of ICE*Mh1,* a bacteriophage or an IME. This suggestion is supported by the observation that

471 ICE*Mh*GAMTULA contained only nine type four secretion system (T4SS) components

472 compared to the 14 T4SS components of ICE*Mh1*, which would potentially limit its ability to

473 undergo autonomous conjugative transfer.

474 Having established that the known macrolide ARGs *msr*(E), *mph*(E) and *estT*, and the

475 known tetracycline resistance gene *tet*(H) were all present within an ICE, we next investigated

476 whether any of the resistance-associated candidate genes identified in the multifactorial GLM

477 analysis of the pangenome were also within the boundaries of predicted ICEs. A representative of

478 each ICE was selected (ICE*Mh1*-like01 from isolate MH077, ICE*Mh*GAMTULA from isolate

479 MH017, and ICE*Mh*Susceptible from isolate 36267_S_17) and then inspected to determine

480 whether any of the candidate genes associated with phenotypic resistance (as shown by the

481 multifactorial GLM analysis) were located within the ICE sequence. None of the 150 candidate

482 genes associated with resistance/susceptibility to GAM were identified within the boundaries of

483 any of the putative ICEs. The lack of any putative susceptibility/resistance genes within ICEs

484    may be due to the inclusion of *tet*(H) as an explanatory factor in the multifactorial GLMs, as

485    *tet*(H) is known to be strongly associated with the presence of ICE in *M. haemolytica.* Thus,

486    including *tet*(H) in our multifactorial GLM would mean that any other genes also associated with

487    the presence of an ICE would fail to be identified. Since macrolide class antimicrobials function

488    by inhibiting translation (64), it is unlikely that the presence of any given gene would confer

489    susceptibility to them and we therefore suggest that future investigations focus on the 79 putative

490    GAM resistance genes identified using our multifactorial GLM analysis.

491        The identification of an ICE in susceptible isolates suggests that ICEs are widespread in

492    *M. haemolytica* and that ICEs may confer some beneficial trait to their bacterial hosts other than

493    AMR as, in most cases, ICEs are known to confer a fitness cost to their bacterial host. ICEs are

494    known to contain genes that encode other beneficial phenotypes such as nitrogen fixation,

495    aromatic compound degradation, and phage defense systems, however, it is not clear whether

496    ICE*Mh*Susceptible confers any traits to its host (69–72).

497        Identification of ICEs containing the macrolide resistance gene *estT* provides an

498    explanation for the observed resistance to the 16-membered macrolides tilmicosin (TILM) and

499    tildipirosin (TILD), however it fails to explain the observed resistance to gamithromycin (GAM)

500    and/or tulathromycin (TULA) in isolates containing ICE*Mh1*-like01 since the *estT*-encoded

501    esterase enzyme had no activity on gamithroymcin (GAM) or tulathromycin (TULA) in a

502    previous report (14). While the L23 ribosomal protein variant we identified partially explains the

503    observed resistance to GAM in the isolates with the GAM/TULA/TET/TILM/TILD phenotype,

504    we still lack an explanation for the observed GAM and TULA resistance in isolates with the

505    TULA/TET/TILM/TILD and GAM/TULA/TILM phenotypes. Further investigation of the 79

506 candidate resistance genes identified in the multifactorial GLM analysis may offer such an

507 explanation.

## Conclusions

509       Our results reinforce the strong association of ICEs with AMR in *M. haemolytica,* with

510 90% of the resistant isolates examined in this study containing a putative ICE. The presence of

511 *estT* explains the observed phenotypic resistance to the 16-membered lactone ring macrolides

512 TILM and TILD, however based on previous reports of biochemical activity, it fails to provide

513 an explanation for the observed resistance to GAM and/or TULA, in isolates with resistance to

514 all four macrolides or those that exhibited resistance to TULA, TILM and TILD. In the isolates

515 that are only resistant to GAM and TULA, the presence of *msr*(E) and *mph*(E) explained this

516 observed phenotypic resistance. By quantifying the relationships between the presence/absence

517 profiles of genes in the *M. haemolytica* pangenome and antimicrobial resistance phenotypes, we

518 identified candidate genes that may contribute to the observed resistance and that should be

519 studied further. We were also able to identify a GAM resistance-associated insertion in the L23

520 ribosomal protein gene and a macrolide resistance associated variant in the L22 ribosomal

521 protein, however neither of these associations has been functionally verified. Finally, we found

522 that ICEs containing no identifiable ARGs are common among susceptible isolates, which

523 suggests that these elements may confer other fitness benefits to *M. haemolytica*. Taken together

524 these results highlight that the success of diagnostic sequencing for detecting macrolide

525 resistance will rely not only on the detection of well-established ARGs but also ongoing

526 monitoring for novel mutations in the cellular targets of these antibiotics.

## Data Availability

Assembled genome sequences have been deposited to NCBI under the BioProject accession
PRJNA1088094.

## Acknowledgements

## References

1.	DeDonder KD, Apley MD. 2015. A literature review of antimicrobial resistance in
	pathogens associated with bovine respiratory disease. Anim Health Res Rev 16:125–134.

2.	Horton LM, Depenbusch BE, Dewsbury DM, McAtee TB, Betts NB, Renter DG. 2023.
	Comprehensive Outcomes Affected by Antimicrobial Metaphylaxis of Feedlot Calves at
	Medium-Risk for Bovine Respiratory Disease from a Randomized Controlled Trial. Vet Sci
	10:67.

3.	O'Connor AM, Hu D, Totton SC, Scott N, Winder CB, Wang B, Wang C, Glanville J, Wood
	H, White B, Larson R, Waldner C, Sargeant JM. 2019. A systematic review and network

548        meta-analysis of injectable antibiotic options for the control of bovine respiratory disease in

549        the first 45 days post arrival at the feedlot. Anim Health Res Rev 20:163–181.

550  4.    Andrés-Lasheras S, Jelinski M, Zaheer R, McAllister TA. 2022. Bovine Respiratory

551        Disease: Conventional to Culture-Independent Approaches to Studying Antimicrobial

552        Resistance in North America. 4. Antibiotics 11:487.

553  5.    Alexander TW, Yanke LJ, Topp E, Olson ME, Read RR, Morck DW, McAllister TA. 2008.

554        Effect of Subtherapeutic Administration of Antibiotics on the Prevalence of Antibiotic-

555        Resistant *Escherichia coli* Bacteria in Feedlot Cattle. Appl Environ Microbiol 74:4405–

556        4416.

557  6.    Government of Canada. 2017. Responsible use of Medically Important Antimicrobials in

558        Animals. https://www.canada.ca/en/public-health/services/antibiotic-antimicrobial-

559        resistance/animals/actions/responsible-use-antimicrobials.html.

560  7.    European Union. 2018. Regulation (EU) 2019/4 of the European Parliament and of the

561        Council on the manufacture, placing on the market and use of medicated feed, amending

562        Regulation (EC) No. 183/2005 of the European Parliament and of the Council and repealing

563        Council Directive 90/167/EEC.

564  8.    European Union. 2018. Regulation (EU) 2019/6 of the European Parliament and of the

565        Council of 11 December 2018 on veterinary medicinal products and repealing Directive

566        2001/82/EC.

567    9.    Hilt EE, Ferrieri P. 2022. Next Generation and Other Sequencing Technologies in

568           Diagnostic Microbiology and Infectious Diseases. Genes 13:1566.

569    10.  Highlander SK. 2001. Molecular genetic analysis of virulence in *Mannheimia (Pasteurella)*

570           *haemolytica*. 3. Front Biosci-Landmark 6:1128–1150.

571    11.  Brault SA, Hannon SJ, Gow SP, Warr BN, Withell J, Song J, Williams CM, Otto SJG,

572           Booker CW, Morley PS. 2019. Antimicrobial Use on 36 Beef Feedlots in Western Canada:

573           2008–2012. Front Vet Sci 6:329.

574    12.  Snyder E, Credille B. 2020. *Mannheimia haemolytica* and *Pasteurella multocida* in bovine

575           respiratory disease: How are they changing in response to efforts to control them? Vet Clin

576           North Am Food Anim Pract 36:253–268.

577    13.  Lubbers BV, Hanzlicek GA. 2013. Antimicrobial multidrug resistance and coresistance

578           patterns of *Mannheimia haemolytica* isolated from bovine respiratory disease cases—a

579           three-year (2009–2011) retrospective analysis: J Vet Diagn Invest 25:413–417.

580    14.  Dhindwal P, Thompson C, Kos D, Planedin K, Jain R, Jelinski M, Ruzzini A. 2023. A

581           neglected and emerging antimicrobial resistance gene encodes for a serine-dependent

582           macrolide esterase. Proc Natl Acad Sci 120:e2219827120.

583    15.  Rice JA, Carrasco-Medina L, Hodgins DC, Shewen PE. 2007. *Mannheimia haemolytica*

584           and bovine respiratory disease. Anim Health Res Rev 8:117–128.

585    16.   Desmolaize B, Rose S, Warrass R, Douthwaite S. 2011. A novel Erm

586        monomethyltransferase in antibiotic-resistant isolates of *Mannheimia haemolytica* and

587        *Pasteurella multocida*. Mol Microbiol 80:184–194.

588    17.   Michael GB, Kadlec K, Sweeney MT, Brzuszkiewicz E, Liesegang H, Daniel R, Murray

589        RW, Watts JL, Schwarz S. 2012. ICE*Pmu1*, an integrative conjugative element (ICE) of

590        *Pasteurella multocida*: structure and transfer. J Antimicrob Chemother 67:91–100.

591    18.   Eidam C, Poehlein A, Leimbach A, Michael GB, Kadlec K, Liesegang H, Daniel R,

592        Sweeney MT, Murray RW, Watts JL, Schwarz S. 2015. Analysis and comparative genomics

593        of ICE*Mh1*, a novel integrative and conjugative element (ICE) of *Mannheimia haemolytica*.

594        J Antimicrob Chemother 70:93–97.

595    19.   Olsen AS, Warrass R, Douthwaite S. 2015. Macrolide resistance conferred by rRNA

596        mutations in field isolates of *Mannheimia haemolytica* and *Pasteurella multocida*. J

597        Antimicrob Chemother 70:420–423.

598    20.   Younes JA, Ramsay DE, Lacoste S, Deschner D, Hill JE, Campbell J, Waldner CL. 2022.

599        Changes in the phenotypic susceptibility of *Mannheimia haemolytica* isolates to macrolide

600        antimicrobials during the early feeding period following metaphylactic tulathromycin use in

601        western Canadian feedlot calves. Can Vet J 63:920–928.

602    21.   CLSI. 2020. VET01: Performance Standards for Antimicrobial Disk and Dilution

603        Susceptibility Tests for Bacteria Isolated From Animals, Fifth Edition. Wayne,

604        Pennsylvania, USA.

605    22.  Martin-Platero AM, Valdivia E, Maqueda M, Martinez-Bueno M. 2007. Fast, convenient,

606          and economical method for isolating genomic DNA from lactic acid bacteria using a

607          modification of the protein "salting-out" procedure. Anal Biochem 366:102–4.

608    23.  Links MG, Dumonceaux TJ, Hemmingsen SM, Hill JE. 2012. The chaperonin-60 universal

609          target is a barcode for bacteria that enables *de novo* assembly of metagenomic sequence

610          data. PLoS ONE 7:e49755.

611    24.  Vancuren SJ, Hill JE. 2019. Update on cpnDB: a reference database of chaperonin

612          sequences. Database 2019:doi:10.1093/database/baz033.

613    25.  Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina

614          sequence data. Bioinformatics 30:2114–20.

615    26.  Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome

616          assemblies from short and long sequencing reads. PLOS Comput Biol 13:e1005595.

617    27.  Shen W, Le S, Li Y, Hu F. 2016. SeqKit: A Cross-Platform and Ultrafast Toolkit for

618          FASTA/Q File Manipulation. PLOS ONE 11:e0163962.

619    28.  Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–

620          2069.

621    29.  Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W,

622          Nguyen A-LV, Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran H-K, Werfalli RE, Nasir

623          JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A,

624          Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F,

625 Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar GV, McArthur AG. 2020.

626 CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance

627 database. Nucleic Acids Res 48:D517–D525.

628 30. Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR, Wlodarski MA, Edalatmand A,

629 Petkau A, Syed SA, Tsang KK, Baker SJC, Dave M, McCarthy MC, Mukiri KM, Nasir JA,

630 Golbon B, Imtiaz H, Jiang X, Kaur K, Kwong M, Liang ZC, Niu KC, Shan P, Yang JYJ,

631 Gray KL, Hoad GR, Jia B, Bhando T, Carfrae LA, Farha MA, French S, Gordzevich R,

632 Rachwalski K, Tu MM, Bordeleau E, Dooley D, Griffiths E, Zubyk HL, Brown ED,

633 Maguire F, Beiko RG, Hsiao WWL, Brinkman FSL, Van Domselaar G, McArthur AG.

634 2023. CARD 2023: expanded curation, support for machine learning, and resistome

635 prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res

636 51:D690–D699.

637 31. Seemann T. 2018. barrnap 0.9: rapid ribosomal RNA prediction.

638 32. Seemann T. 2015. snippy: fast bacterial variant calling from NGS reads.

639 33. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D,

640 Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis.

641 Bioinformatics 31:3691–3693.

642 34. Price MN, Dehal PS, Arkin AP. 2009. FastTree: Computing Large Minimum Evolution

643 Trees with Profiles instead of a Distance Matrix. Mol Biol Evol 26:1641–1650.

644    35.    Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q,

645           Wortman J, Young SK, Earl AM. 2014. Pilon: An Integrated Tool for Comprehensive

646           Microbial Variant Detection and Genome Assembly Improvement. PLOS ONE 9:e112963.

647    36.    Masonbrink R. 2021. Multiple rounds of Pilon polishing/gap-filling on a genome assembly.

648           Bioinforma Workb.

649           https://bioinformaticsworkbook.org/dataAnalysis/GenomeAssembly/Iterate_Pilon_Genome

650           _Polishing.html. Retrieved 6 September 2023.

651    37.    Zhang Y, Park C, Bennett C, Thornton M, Kim D. 2021. Rapid and accurate alignment of

652           nucleotide conversion sequencing reads with HISAT-3N. Genome Res 31:1290–1295.

653    38.    Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T,

654           McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools.

655           GigaScience 10:giab008.

656    39.    Lees JA, Galardini M, Bentley SD, Weiser JN, Corander J. 2018. pyseer: a comprehensive

657           tool for microbial pangenome-wide association studies. Bioinformatics 34:4310–4312.

658    40.    Lees JA, Mai TT, Galardini M, Wheeler NE, Corander J. 2019. Improved inference and

659           prediction of bacterial genotype-phenotype associations using pangenome-spanning

660           regressions. bioRxiv https://doi.org/10.1101/852426.

661    41.    R Core Team. 2021. R: A Language and Environment for Statistical Computing. Vienna,

662           Austria.

663   42.   Clawson ML, Murray RW, Sweeney MT, Apley MD, DeDonder KD, Capik SF, Larson RL,

664         Lubbers BV, White BJ, Kalbfleisch TS, Schuller G, Dickey AM, Harhay GP, Heaton MP,

665         Chitko-McKown CG, Brichta-Harhay DM, Bono JL, Smith TPL. 2016. Genomic signatures

666         of *Mannheimia haemolytica* that associate with the lungs of cattle with respiratory disease,

667         an integrative conjugative element, and antibiotic resistance genes. BMC Genomics 17:982.

668   43.   Buchan DWA, Jones DT. 2019. The PSIPRED Protein Analysis Workbench: 20 years on.

669         Nucleic Acids Res 47:W402–W407.

670   44.   Liu M, Li X, Xie Y, Bi D, Sun J, Li J, Tai C, Deng Z, Ou H-Y. 2019. ICEberg 2.0: an

671         updated database of bacterial integrative and conjugative elements. Nucleic Acids Res

672         47:D660–D665.

673   45.   Klima CL, Zaheer R, Cook SR, Booker CW, Hendrick S, Alexander TW, McAllister TA.

674         2014. Pathogens of Bovine Respiratory Disease in North American Feedlots Conferring

675         Multidrug Resistance via Integrative Conjugative Elements. J Clin Microbiol 52:438–448.

676   46.   Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search

677         tool. J Mol Biol 215:403–410.

678   47.   Desmolaize B, Rose S, Wilhelm C, Warrass R, Douthwaite S. 2011. Combinations of

679         macrolide resistance determinants in field isolates of *Mannheimia haemolytica* and

680         *Pasteurella multocida*. Antimicrob Agents Chemother 55:4128–4133.

681    48.   Su W, Kumar V, Ding Y, Ero R, Serra A, Lee BST, Wong ASW, Shi J, Sze SK, Yang L, Gao

682         Y-G. 2018. Ribosome protection by antibiotic resistance ATP-binding cassette protein. Proc

683         Natl Acad Sci 115:5157–5162.

684    49.   Conrad CC, Daher RK, Stanford K, Amoako KK, Boissinot M, Bergeron MG, Alexander T,

685         Cook S, Ralston B, Zaheer R, Niu YD, McAllister T. 2020. A sensitive and accurate

686         recombinase polymerase amplification assay for detection of the primary bacterial

687         pathogens causing bovine respiratory disease. Front Vet Sci 7:208.

688    50.   Stanford K, Zaheer R, Klima C, McAllister T, Peters D, Niu YD, Ralston B. 2020.

689         Antimicrobial Resistance in Members of the Bacterial Bovine Respiratory Disease

690         Complex Isolated from Lung Tissue of Cattle Mortalities Managed with or without the Use

691         of Antimicrobials. 2. Microorganisms 8:288.

692    51.   Wang Y, Niemi J, Airas K, Ylihonko K, Hakala J, Mäntsälä P. 2000. Modifications of

693         aclacinomycin T by aclacinomycin methyl esterase (RdmC) and aclacinomycin-10-

694         hydroxylase (RdmB) from *Streptomyces purpurascens*. Biochim Biophys Acta BBA -

695         Protein Struct Mol Enzymol 1480:191–200.

696    52.   Zaman S, Fitzpatrick M, Lindahl L, Zengel J. 2007. Novel mutations in ribosomal proteins

697         L4 and L22 that confer erythromycin resistance in *Escherichia coli*. Mol Microbiol

698         66:1039–1050.

699    53.   Nowakowski M, Inskeep P, Risk J, Skogerboe T, Benchaoui H, Meinert T, Sherington J,

700         Sunderland S. 2004. Pharmacokinetics and lung tissue concentrations of tulathromycin, a

701         new triamilide antibiotic, in cattle. Vet Ther 5:60–74.

702    54.    Sauer RT, Baker TA. 2011. AAA+ Proteases: ATP-Fueled Machines of Protein Destruction.

703          Annu Rev Biochem 80:587–612.

704    55.    Acharya S, Foster PL, Brooks P, Fishel R. 2003. The Coordinated Functions of the *E. coli*

705          MutS and MutL Proteins in Mismatch Repair. Mol Cell 12:233–246.

706    56.    Sancar A, Williams KR, Chase JW, Rupp WD. 1981. Sequences of the ssb gene and protein.

707          Proc Natl Acad Sci 78:4274–4278.

708    57.    Kramer G, Rauch T, Rist W, Vorderwülbecke S, Patzelt H, Schulze-Specking A, Ban N,

709          Deuerling E, Bukau B. 2002. L23 protein functions as a chaperone docking site on the

710          ribosome. Nature 419:171–174.

711    58.    Allali N, Afif H, Couturier M, Van Melderen L. 2002. The Highly Conserved TldD and

712          TldE Proteins of *Escherichia coli* Are Involved in Microcin B17 Processing and in CcdA

713          Degradation. J Bacteriol 184:3224–3231.

714    59.    Kaur G, Subramanian S. 2016. Classification of the treble clef zinc finger: noteworthy

715          lessons for structure and function evolution. Sci Rep 6:32070.

716    60.    Donahue JL, Bownas JL, Niehaus WG, Larson TJ. 2000. Purification and Characterization

717          of *glpX*-Encoded Fructose 1,6-Bisphosphatase, a New Enzyme of the Glycerol 3-Phosphate

718          Regulon of *Escherichia coli*. J Bacteriol 182:5624–5627.

719    61.    Cabedo H, Macián F, Villarroya M, Escudero JC, Martínez-Vicente M, Knecht E,

720          Armengod ME. 1999. The *Escherichia coli trmE* (*mnmE*) gene, involved in tRNA

721      modification, codes for an evolutionarily conserved GTPase with unusual biochemical

722      properties. EMBO J 18:7063–7076.

723    62.   Mahdi AA, Sharples GJ, Mandal TN, Lloyd RG. 1996. Holliday Junction Resolvases

724      Encoded by Homologous *rusA* Genes in *Escherichia coli* K-12 and Phage 82. J Mol Biol

725      257:561–573.

726    63.   González-Pastor JE, Hobbs EC, Losick R. 2003. Cannibalism by Sporulating Bacteria.

727      Science 301:510–513.

728    64.   Vázquez-Laslop N, Mankin AS. 2018. How Macrolide Antibiotics Work. Trends Biochem

729      Sci 43:668–684.

730    65.   Öhman A, Rak A, Dontsova M, Garber MB, Härd T. 2003. NMR structure of the ribosomal

731      protein L23 from *Thermus thermophilus*. J Biomol NMR 26:131–137.

732    66.   Beker M, Rose S, Lykkebo CA, Douthwaite S. 2018. Integrative and Conjugative Elements

733      (ICEs) in Pasteurellaceae Species and Their Detection by Multiplex PCR. Front Microbiol

734      9:1329.

735    67.   Cameron A, Zaheer R, McAllister TA. 2019. Emerging Variants of the Integrative and

736      Conjugant Element ICEMh1 in Livestock Pathogens: Structural Insights, Potential Host

737      Range, and Implications for Bacterial Fitness and Antimicrobial Therapy. Front Microbiol

738      10:2608.

739    68.   Johnson CM, Grossman AD. 2015. Integrative and Conjugative Elements (ICEs): What

740      They Do and How They Work. Annu Rev Genet 49:577–601.

741    69.  Wang M, Liu G, Liu M, Tai C, Deng Z, Song J, Ou H-Y. 2024. ICEberg 3.0: functional

742         categorization and analysis of the integrative and conjugative elements in bacteria. Nucleic

743         Acids Res 52:D732–D737.


744    70.  Hirose J. 2023. Diversity and Evolution of Integrative and Conjugative Elements Involved

745         in Bacterial Aromatic Compound Degradation and Their Utility in Environmental

746         Remediation. Microorganisms 11:438.


747    71.  Johnson CM, Harden MM, Grossman AD. 2022. Interactions between mobile genetic

748         elements: An anti-phage gene in an integrative and conjugative element protects host cells

749         from predation by a temperate bacteriophage. PLoS Genet 18:e1010065.


750    72.  Haskett TL, Ramsay JP, Bekuma AA, Sullivan JT, O'Hara GW, Terpolilli JJ. 2017.

751         Evolutionary persistence of tripartite integrative and conjugative elements. Plasmid 92:30–

752         36.


753

754

**Table 1.** Antibiotic resistance phenotypes of study isolates

| Phenotype[1] | Number of Isolates |
| --- | --- |
| Susceptible | 47 |
| GAM/TULA/TET/TILM/TILD | 27 |
| TULA/TET/TILM/TILD | 11 |
| GAM/TULA | 8 |
| TET/TILM/TILD | 5 |
| GAM/TULA/TILM | 1 |
| Total | 99 |

[1]GAM = gamithromycin, TULA = tulathromycin, TILM = tilmicosin, TILD = tildipirosin, TET = tetracycline. GAM, TULA, TILM and TILD are all macrolide class antibiotics. TET is a tetracycline class antibiotic.

761

| Phenotype[1] | No. of isolates | *msr*(E) | *mph*(E) | *erm42* | *sul2* | *tet*(H) | APH(3")-1b | APH(3')-1a | APH(6)-1d | *estT* |
|---|---|---|---|---|---|---|---|---|---|---|
| GAM/TULA | 8 | 8/8 | 8/8 | 0/8 | 7/8 | 0/8 | 0/8 | 0/8 | 0/8 | 0/8 |
| GAM/TULA/TET/TILM/TILD | 27 | 0/27 | 0/27 | 0/27 | 25/27 | 27/27 | 27/27 | 27/27 | 27/27 | 27/27 |
| TULA/TET/TILM/TILD | 11 | 0/11 | 0/11 | 0/11 | 11/11 | 11/11 | 11/11 | 11/11 | 11/11 | 11/11 |
| TET/TILM/TILD | 5 | 0/5 | 0/5 | 0/5 | 5/5 | 5/5 | 5/5 | 5/5 | 5/5 | 5/5 |
| GAM/TULA/TILM | 1 | 0/1 | 0/1 | 0/1 | 1/1 | 1/1 | 1/1 | 1/1 | 1/1 | 1/1 |
| Susceptible | 47 | 0/47 | 0/47 | 0/47 | 0/47 | 0/47 | 0/47 | 0/47 | 0/47 | 0/47 |

762 **Table 2.** Antibiotic resistance genes detected in study isolates.

763 [1]GAM = gamithromycin, TULA = tulathromycin, TET = tetracycline, TILM = tilmicosin, TILD = tildipirosin.

764

765  **Table 3.** Results of unifactorial and multifactorial GLMs for
766  each antibiotic. Shown are the p-values for the three
767  explanatory factors: genotype, *tet*(H), and *msr*(E).

| Analysis | Drug | genotype | *tet*(H) | *msr*(E) |
|---|---|---|---|---|
| Unifactorial | GAM | 0.33 | 4.9e-07 | 3.2e-05 |
| Unifactorial | TULA | 0.97 | 1.67e-14 | 0.0004 |
| Unifactorial | TET | 2.6e-06 | < 2.2e-16 | 0.002 |
| Unifactorial | TILM | 1.8e-06 | < 2.2e-16 | 0.002 |
| Unifactorial | TILD | 2.6e-06 | < 2.2e-16 | 0.002 |
| | | | | |
| Multifactorial | GAM | 0.52 | 1.4e-10 | 1.2e-06 |
| Multifactorial | TULA | 0.52 | < 2.2e-16 | 1.2e-06 |
| Multifactorial | TET | 0.52 | < 2e-16 | > 0.99 |
| Multifactorial | TILM | 0.52 | < 2e-16 | > 0.99 |
| Multifactorial | TILD | 0.52 | < 2e-16 | > 0.99 |

768

769

770 **Table 4**. Distribution of identified putative ICEs in *M. haemolytica* isolates in the present study.

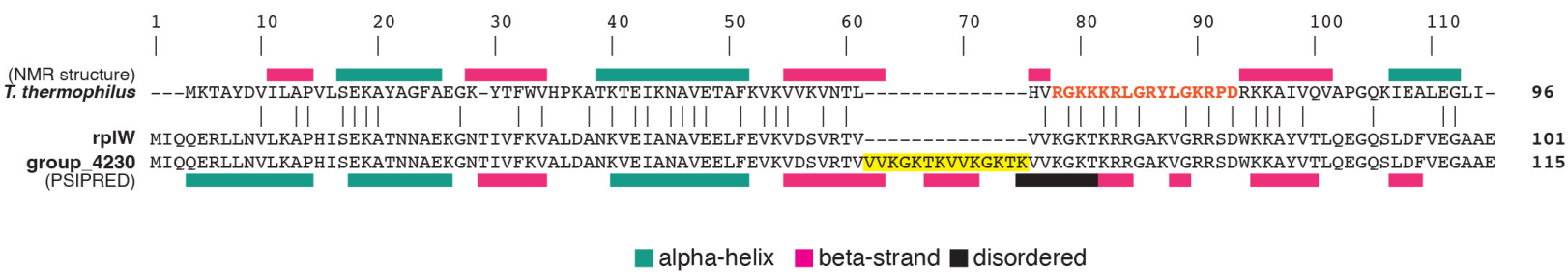| Phenotype | No. of isolates | ICE*Mh1*-like01 | ICE*Mh*GAMTULA | ICE*Mh*Susceptible |
|---|---|---|---|---|
| Susceptible | 47 | 0 | 0 | 19 |
| GAM/TULA/TET/TILM/TILD | 27 | 24 | 0 | 0 |
| TULA/TET/TILM/TILD | 11 | 10 | 0 | 0 |
| TET/TILM/TILD | 5 | 5 | 0 | 0 |
| GAM/TULA | 8 | 0 | 7 | 0 |
| GAM/TULA/TILM | 1 | 1 | 0 | 0 |
| Direct Repeat Sequence | | cgtgtcggttcgagtccgacc | aaataataatgaaaa | cgtgtcggttcgagtccgacc |
| Total | 99 | 40 | 7 | 19 |

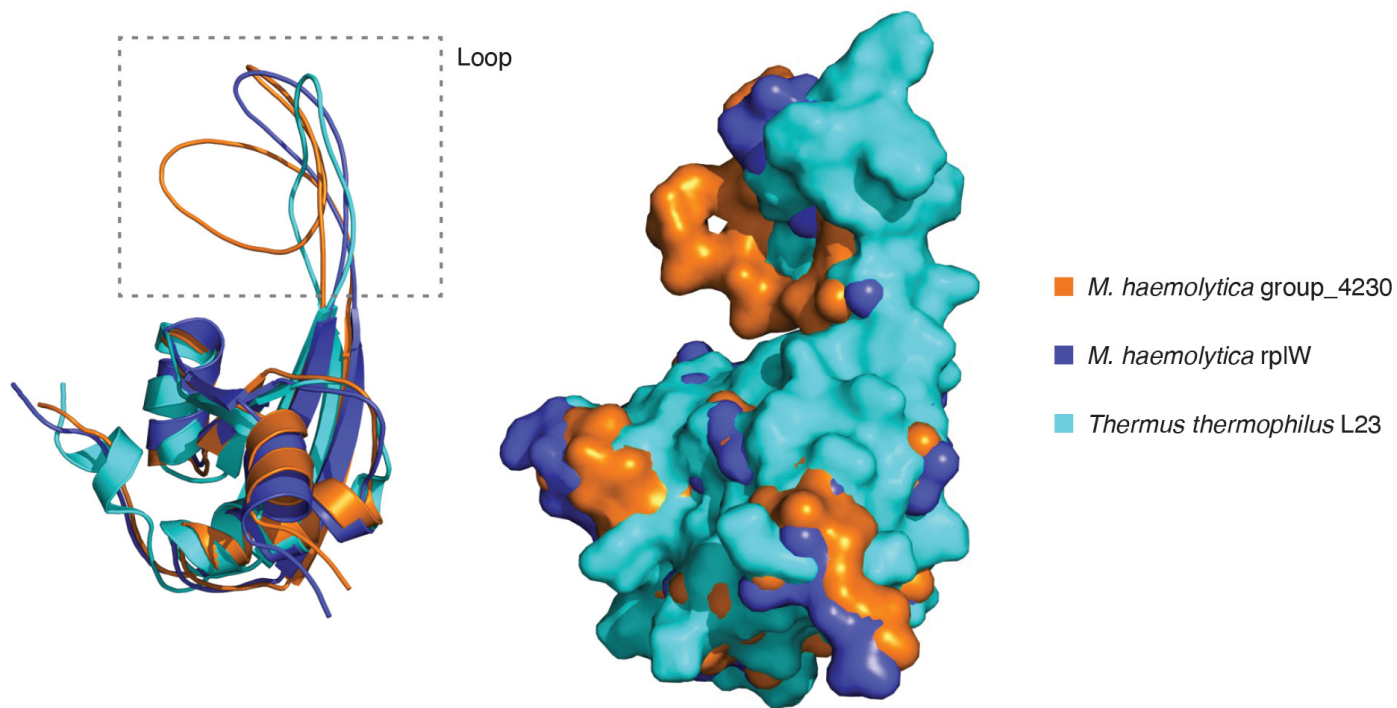771

772

**Figure Captions**

**Figure 1.** (A) Alignment of the amino acid sequences of L23 ribosomal proteins from *T. thermophilus* (PDB accession 1N88) and *M. haemolytica* (rplW and group_4230). A 14-amino acid insertion in group_4230 is indicated by yellow highlighting, and the flexible loop of L23 by orange text. Secondary structure features (determined by NMR for *T. thermophilus* or predicted by PSIPRED and DISOPRED3 for *M. haemolytica* sequences) are indicated above and below the alignment according to the legend. Sequences were aligned using MUSCLE pairwise alignment in Geneious Prime with default settings. (B) Ribbon (left) and space-filling models (right) of the alignment of the NMR structure of *T. thermophilus* L23 with predicted structures of *M. haemolytica* rplW and group_4230. The loop structure that is extended by the insertion in group_4230 is indicated on the ribbon diagram.

**Figure 2.** Putative ICE sequences identified in *M. haemolytica* isolates in the present study, ICE*Mh1*-like01, ICE*Mh*GAMTULA and ICE*Mh*Susceptible. ICE*Mh1* (top) is included for reference. Regions containing known ARGs are expanded to show detailed gene organization. ARGs, ICE-associated genes (encoding relaxases, transposases, integrases, and other functions) and genes encoding hypothetical proteins are indicated by colour according to the legend. Scale bar indicates 20 kilobase pairs.
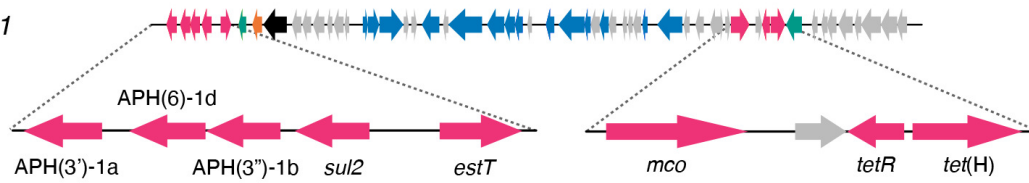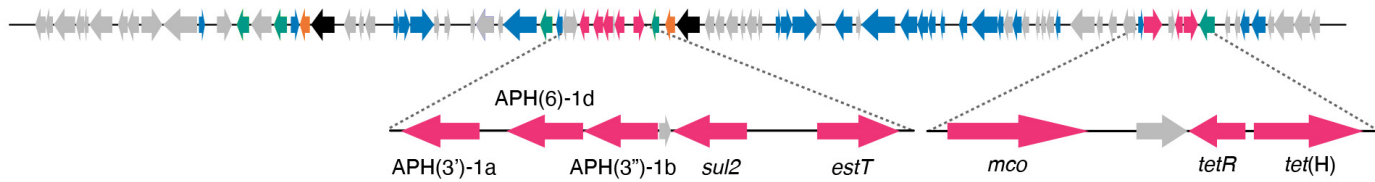
**A.**

```
                1         10        20        30        40        50        60        70        80        90        100       110
                |         |         |         |         |         |         |         |         |         |         |         |
(NMR structure)
T. thermophilus ---MKTAYDVILAPVLSEKAYAGFAEGK-YTFWVHPKATKTEIKNAVETAFKVKVKVNTL--------------HVRGKKKRLGRYLGKRPDRKKAIVQVAPGQKIEALEGLI-   96
                      | | |   ||||   |  | |  |  || ||| || |  | |   |  |                       | |  ||| || || || |  | | | |  |
rplW            MIQQERLLNVLKAPHISEKATNNAEKGNTIVFKVALDANKVEIANAVEELFEVKVDSVRTV--------------VVKGKTKRRGAKVGRRSDWKKAYVTLQEGQSLDFVEGAAE   101
group_4230      MIQQERLLNVLKAPHISEKATNNAEKGNTIVFKVALDANKVEIANAVEELFEVKVDSVRTVVVKGKTKVVKGKTKVVKGKTKRRGAKVGRRSDWKKAYVTLQEGQSLDFVEGAAE   115
(PSIPRED)
```

■ alpha-helix  ■ beta-strand  ■ disordered

**B.**



Loop

■ *M. haemolytica* group_4230
■ *M. haemolytica* rplW
■ *Thermus thermophilus* L23

ICE*Mh1*

APH(6)-1d

APH(3')-1a    APH(3")-1b   *sul2*    *estT*

*mco*    *tetR*    *tet*(H)

ICE*Mh1*–like01

APH(6)-1d

APH(3')-1a    APH(3")-1b   *sul2*    *estT*
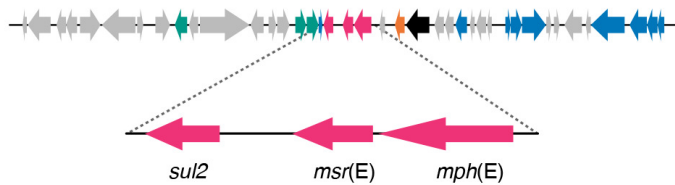
*mco*    *tetR*    *tet*(H)

ICE*Mh*GAMTULA

*sul2*    *msr*(E)    *mph*(E)

ICE*Mh*Susceptible

- AMR gene
- Relaxase
- Transposase
- Integrase
- Other ICE related function
- Hypothetical protein

20 kb