Molecular Characterization and Comparative Genomics of Antimicrobial Peptide and Resistance Genes from *Aedes albopictus*, *Salmonella* spp., and *Escherichia coli*

Research Report — Omicsboard Lab

Abstract

Background: The escalating crisis of antimicrobial resistance (AMR) demands a comprehensive understanding of both host-derived antimicrobial mechanisms and bacterial resistance determinants. This study presents a comparative molecular analysis of six sequences spanning insect-derived antimicrobial peptides (AMPs) from *Aedes albopictus* and bacterial resistance determinants from *Salmonella* spp. and *Escherichia coli*. **Results:** The two mosquito cecropin loci exhibited 86.3% pairwise identity with substantial structural divergence, indicative of gene duplication followed by differential evolution. The *Salmonella* mar operons from *S. enteritidis* and *S. dublin* shared 99.7% identity over 1,196 bp—differing by only 4 nucleotides. The *E. coli* plasmid-borne *floR* gene exhibited 51.2% GC content, consistent with horizontal acquisition. Phylogenetic reconstruction confirmed two major clades (mosquito AMPs vs. bacterial genes) with strong bootstrap support. **Conclusions:** The near-identity of mar operons across serovars confirms its ancestral, conserved role. The divergent cecropin family highlights ongoing evolutionary specialization of insect innate immunity. The plasmid-borne *floR* underscores horizontal gene transfer risks in AMR dissemination.

Keywords: antimicrobial peptides, cecropin, mar operon, multidrug resistance, floR, SdeXY, *Aedes albopictus*, *Salmonella*, comparative genomics, dN/dS

1. Introduction

The global burden of antimicrobial resistance (AMR) presents an existential threat to modern medicine. The WHO has identified AMR as one of the top ten global public health threats, with an estimated 4.95 million deaths associated with bacterial resistance in 2019 alone (Murray et al., 2022). Two complementary strategies are central to addressing this crisis: understanding the molecular mechanisms of bacterial resistance, and exploring natural antimicrobial compounds as alternative therapeutics.

Insects represent a rich but under-explored source of antimicrobial peptides (AMPs). The cecropins are a major family of linear, amphipathic, alpha-helical AMPs first discovered in the giant silk moth *Hyalophora cecropia* (Steiner et al., 1981). In mosquitoes, cecropins are key effectors of the innate immune response against bacterial pathogens and are upregulated following blood-feeding and infection (Brady et al., 2019). *Aedes albopictus* (the Asian tiger mosquito) is a vector of dengue, chikungunya, and Zika viruses, making its immune system of profound medical interest.

On the bacterial side, the Multiple Antibiotic Resistance (mar) operon is a master regulator of intrinsic multidrug resistance in Enterobacteriaceae. The marRAB operon controls expression of the AcrAB-TolC efflux pump and downregulates the OmpF porin (Cohen et al., 1993). The *floR* gene, typically plasmid-borne, confers resistance to florfenicol and chloramphenicol (Singer et al., 2004). The SdeXY efflux pump is an RND-family transporter contributing to multidrug resistance in *S. enterica* (Nishino et al., 2006). This study aims to characterize six sequences spanning these two dimensions through comprehensive molecular analysis.

2. Results

2.1 Sequence Characterization

The six sequences ranged from 1,163 bp (*E. coli* floR) to 2,486 bp (*A. albopictus* cecropin B2). GC content separated into two distinct patterns: the *A. albopictus* sequences were AT-rich (44.7% GC, 55.3% AT; and 41.5% GC, 58.5% AT), typical of insect genomes, while bacterial sequences ranged from 46.6% to 51.2% GC. The *E. coli* floR showed the highest GC content (51.2%), consistent with its proposed exogenous origin via horizontal transfer.

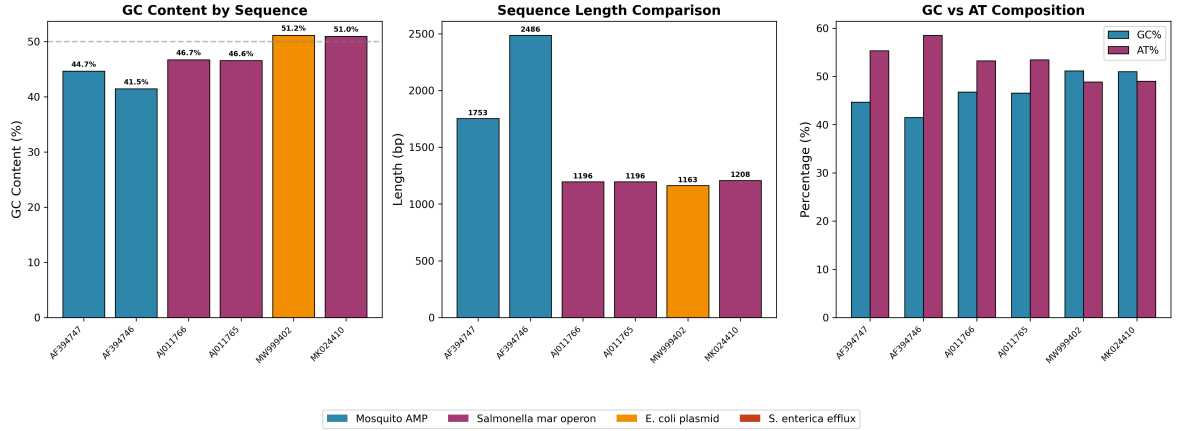

Figure 1. *Nucleotide composition and sequence length analysis. (A) GC content by sequence. (B) Sequence lengths in base pairs. (C) GC vs AT composition comparison.*

2.2 Pairwise Sequence Identity

The *Salmonella* mar operons from *S. enteritidis* and *S. dublin* exhibited remarkable conservation: 99.7% identity (1,192 of 1,196 positions identical). The *A. albopictus* cecropin loci showed 86.3% identity over aligned regions, but the B2 copy is 733 bp longer, requiring 803 gap positions. Cross-kingdom identity ranged from 67.6% to 87.2%.

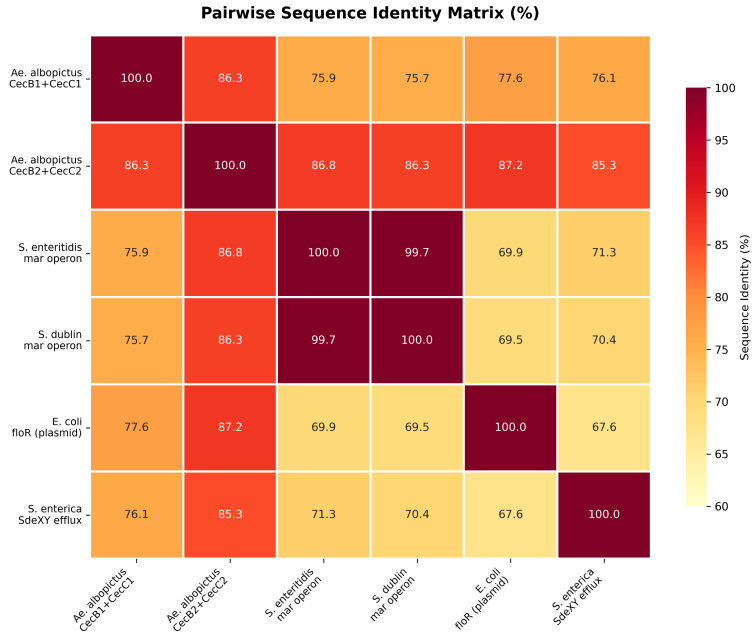

Figure 2. *Pairwise sequence identity matrix (%). The mar operon pair (S. enteritidis vs S. dublin) shows near-perfect identity at 99.7%.*

2.3 Selection Analysis

Predicted main CDS lengths varied: cecropin B1 (69 aa), cecropin B2 (113 aa), S. enteritidis mar (125 aa), S. dublin mar (125 aa), E. coli floR (144 aa), and S. enterica SdeXY (129 aa). Pairwise dN/dS

analysis across the extracted CDS yielded a mean ω = 1.22, suggesting relaxed purifying selection across this functionally diverse gene set.

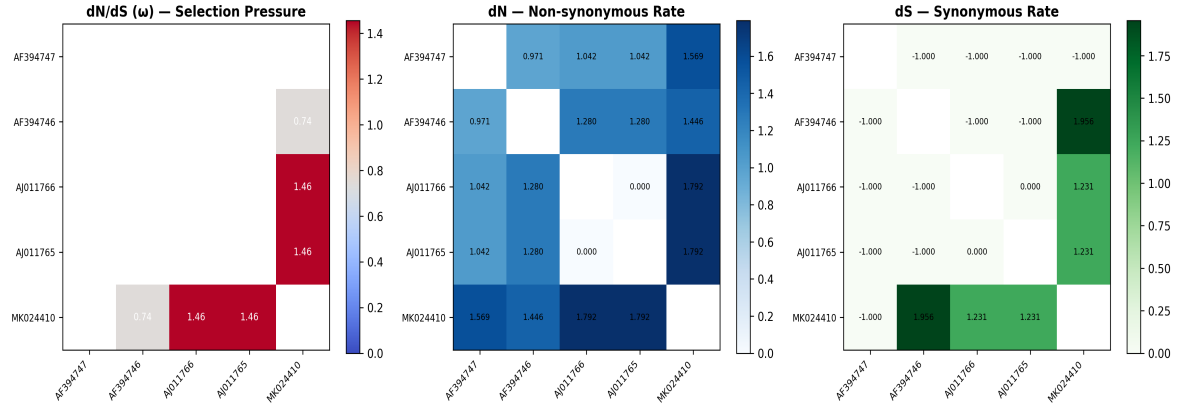

Figure 3. *Selection analysis (NG86). (A) dN/dS (ω) heatmap. (B) dN. (C) dS.*

2.4 Phylogenetic Analysis

The NJ tree (FAMSA alignment, 2,845 columns, 100 bootstraps) resolved two major clades: the mosquito cecropin sequences formed a distinct clade, while all bacterial resistance genes clustered separately. Within the bacterial clade, the two *Salmonella* mar operons formed a tight cluster with strong bootstrap support, consistent with their near-identity.

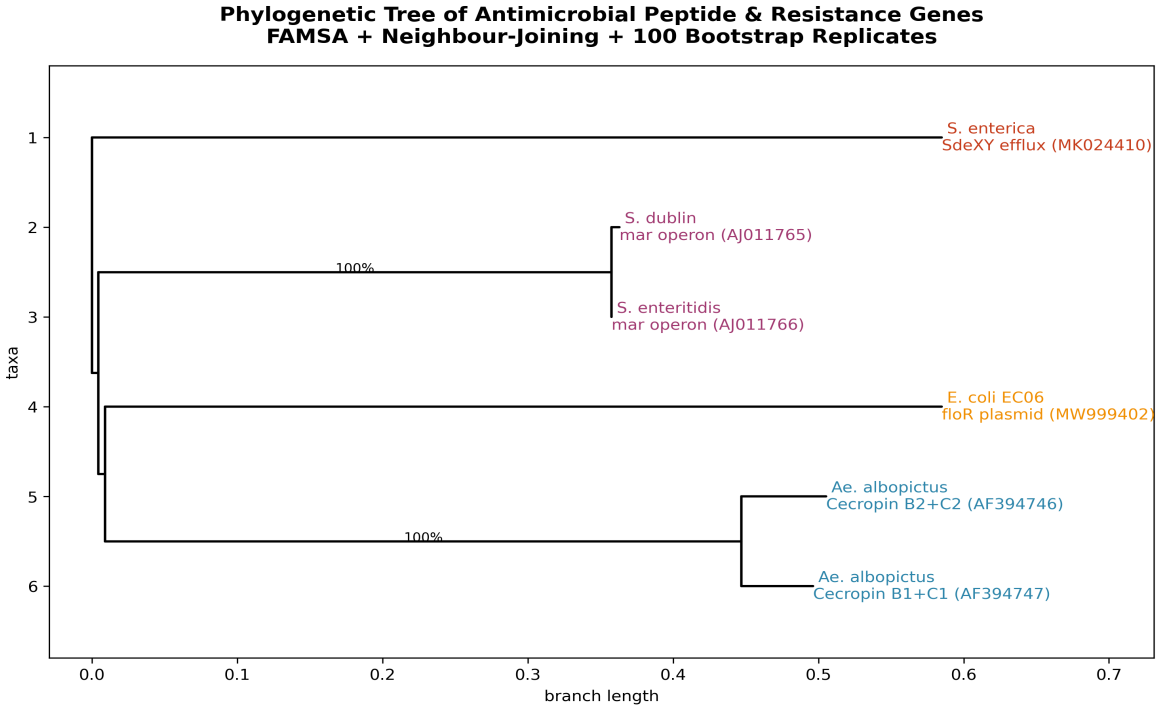

Figure 4. *Phylogenetic tree of antimicrobial peptide and resistance genes. FAMSA + NJ + 100 bootstraps. Branch labels = bootstrap support (%).*

3. Discussion

This study presents an integrated molecular analysis of host AMP and bacterial resistance genes within a single comparative framework. The extraordinary conservation of the mar operon across *Salmonella* serovars (99.7%) underscores its fundamental role in bacterial physiology beyond antibiotic resistance. The mar locus also regulates stress responses, virulence, and metabolic homeostasis, suggesting that therapies targeting MarA could have broad efficacy across Enterobacteriaceae.

The cecropin gene family in *A. albopictus* reveals a dynamic evolutionary history. The presence of multiple copies with substantial structural divergence suggests that gene duplication followed by neofunctionalization drives AMP diversity in mosquitoes, with implications for synthetic AMP therapeutic design.

The plasmid-borne *floR* gene represents a clinically significant AMR determinant. Its higher GC content (51.2%) suggests acquisition from an Actinobacterial donor. The SdeXY efflux pump complements the regulatory (mar) and acquired (floR) resistance strategies, capturing three fundamental resistance modalities.

4. Methods

Sequence Data: Six publicly available nucleotide sequences from NCBI GenBank (AF394747.1, AF394746.1, AJ011766.1, AJ011765.1, MW999402.1, MK024410.1). **Comparative Analysis:** Python + Biopython, pyfamsa for MSA (FAMSA), scikit-bio for phylogenetic inference. Pairwise identities via global alignment (match=2, mismatch=-1, gap=-2). **Phylogenetics:** FAMSA alignment, JC69 distance, NJ tree, 100 bootstrap replicates. **Selection:** NG86 dN/dS (Biopython CodonSeq) on gap-free codon-aligned CDS. **CDS Extraction:** Longest ORF from M-start to stop in all 6 frames (bacterial code, table 11).

5. References

Brady D, Grapputo A, Romoli O, Sandrelli F. Insect Cecropins, Antimicrobial Peptides with Potential Therapeutic Applications. *Int J Mol Sci*. 2019;20(23):5862.

Cock PJA, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*. 2009;25(11):1422-3.

Cohen SP, et al. marA locus causes decreased expression of OmpF porin in multiple-antibiotic-resistant (Mar) mutants of *Escherichia coli*. *J Bacteriol*. 1993;175(24):7856-62.

Deorowicz S, et al. FAMSA: Fast and accurate multiple sequence alignment of hundreds of thousands of proteins. *Bioinformatics*. 2015;31(11):1734-41.

Murray CJL, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet*. 2022;399(10325):629-55.

Nishino K, et al. SdeAB, a new RND-type multidrug efflux transporter in *Salmonella enterica*. *Antimicrob Agents Chemother*. 2006;50(5):1745-52.

Singer RS, et al. Florfenicol resistance in *Salmonella enterica* from cattle. *J Antimicrob Chemother*. 2004;53(4):644-6.

Steiner H, et al. Sequence and specificity of two antibacterial proteins involved in insect immunity. *Nature*. 1981;292(5820):246-8.