Evolutionary Analysis of Kisspeptin Proteins via Multiple Sequence Alignment and PAM250 Matrix

Authors: Kauan Divino Pous Mariano[1 2], Kamylla Sejane Pouso Freitas[3]

[1]Federal University of Goiás, Institute of Tropical Pathology and Public Health, Goiânia-GO, Brazil
[2]Federal University of Goiás, Institute of Informatics, Goiânia-GO, Brazil
[3]Centro Universitário Alfredo Nasser, Aparecida de Goiânia-GO, Brazil
** email: kauan@discente.ufg.br , kamyllasejane@unifan.edu.br*

Introduction: Kisspeptins (Kiss1 and Kiss2) are essential peptide hormones that regulate puberty onset and reproductive functions in vertebrates. Due to their biological importance and evolutionary divergence, comparative analysis of their sequences can shed light on their functional conservation. **Objectives:** This study aims to investigate the evolutionary relationships between Kiss1 and Kiss2 proteins from various vertebrates through multiple sequence alignment (MSA), supported by the PAM250 substitution matrix. **Methods:** Seven kisspeptin sequences from Homo sapiens, Mus musculus, Monodelphis domestica, Xenopus tropicalis (two isoforms), and Danio rerio (two isoforms) were collected from a public dataset (Kaggle:https://www.kaggle.com/datasets/samira1992/sequence-alignment-bioinformatics-dataset). The sequences, provided in FASTA format, were aligned using the MUSCLE algorithm implemented in a custom Python notebook developed on Google Colab. The alignment was performed with the PAM250 scoring matrix to highlight evolutionary substitutions. Pairwise identity values were computed for all sequence pairs, followed by visualization in a heatmap and generation of a hierarchical clustering dendrogram. All code and analyses are available at https://github.com/kauandivino/kisspeptin-evolutionary-analysis. **Results:** The alignment yielded a consensus of 214 amino acid positions, revealing significant conservation among mammalian Kiss1 proteins (identity ≥47%), while the frog and zebrafish Kiss2 isoforms showed lower identity (<35%) when compared to other sequences. Human, mouse, and opossum Kiss1 clustered tightly in both the identity heatmap and dendrogram analysis, indicating evolutionary conservation. In contrast, Kiss2 proteins formed distinct and more divergent clades, supporting their functional differentiation. Notably, frog-Kiss1, while part of the Kiss1 group, exhibited intermediate similarity and appeared more distant from the mammalian branch. Zebrafish-Kiss1 formed a separate cluster, suggesting species-specific divergence **Conclusion:** Our findings support that Kiss1 proteins are evolutionarily conserved in mammals, while Kiss2 exhibits greater sequence divergence, potentially reflecting functional diversification. The approach confirms the utility of bioinformatic tools like MUSCLE, substitution matrices (PAM250), and clustering algorithms for understanding molecular evolution in hormone-regulating protein families. This project contributes to the field of Golden Biotechnology by demonstrating the integration of computational biology and evolutionary genomics in the analysis of peptide signaling systems.

Keywords: kisspeptin; sequence alignment; evolutionary analysis; PAM250; bioinformatics; MSA; Golden Biotechnology

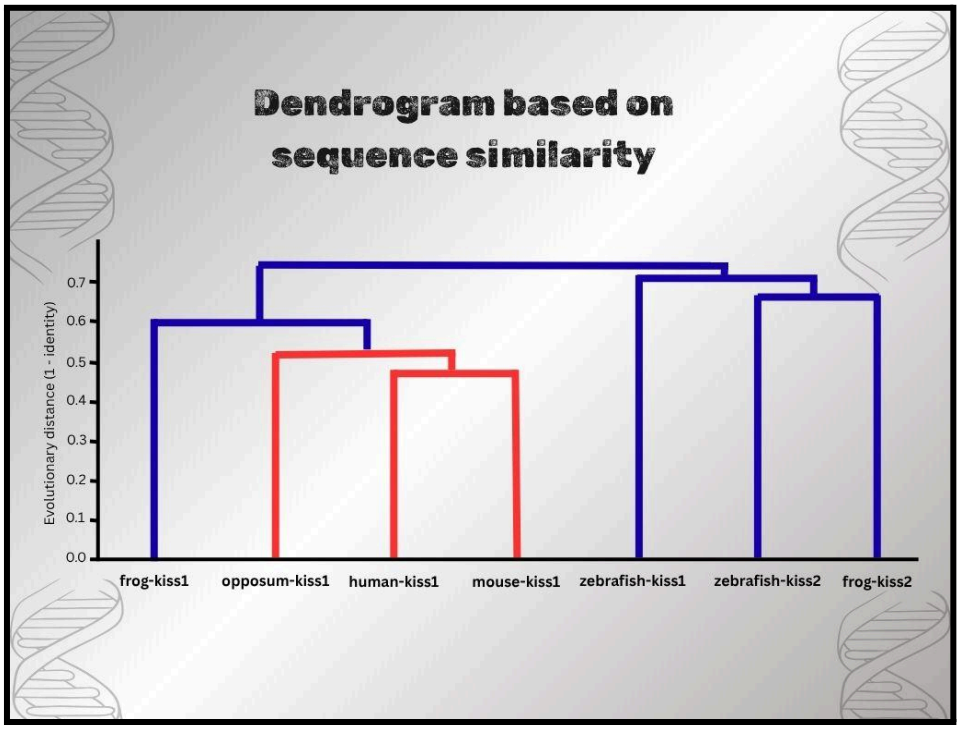

Figura 1. ***Hierarchical clustering of Kisspeptin protein sequences based on pairwise identity values. Created by the author.***

Data from Kaggle (2025). Code and analysis available at github.com/kauandivino/kisspeptin-evolutionary-analysis