Supplementary Material SM8

**UNIVERSIDAD DE INVESTIGACIÓN DE TECNOLOGÍA EXPERIMENTAL YACHAY**

**Escuela de Ciencias Biológicas e Ingeniería**

**TÍTULO: In Silico Discovery of Hemolytic Peptides Through a Novel Approach Based on Network Science and Similarity Searching Methods**

Trabajo de integración curricular presentado como requisito para la obtención del título de Biólogo

**Autor*:***

Castillo Mendieta Kevin Alexis

**Tutora:**

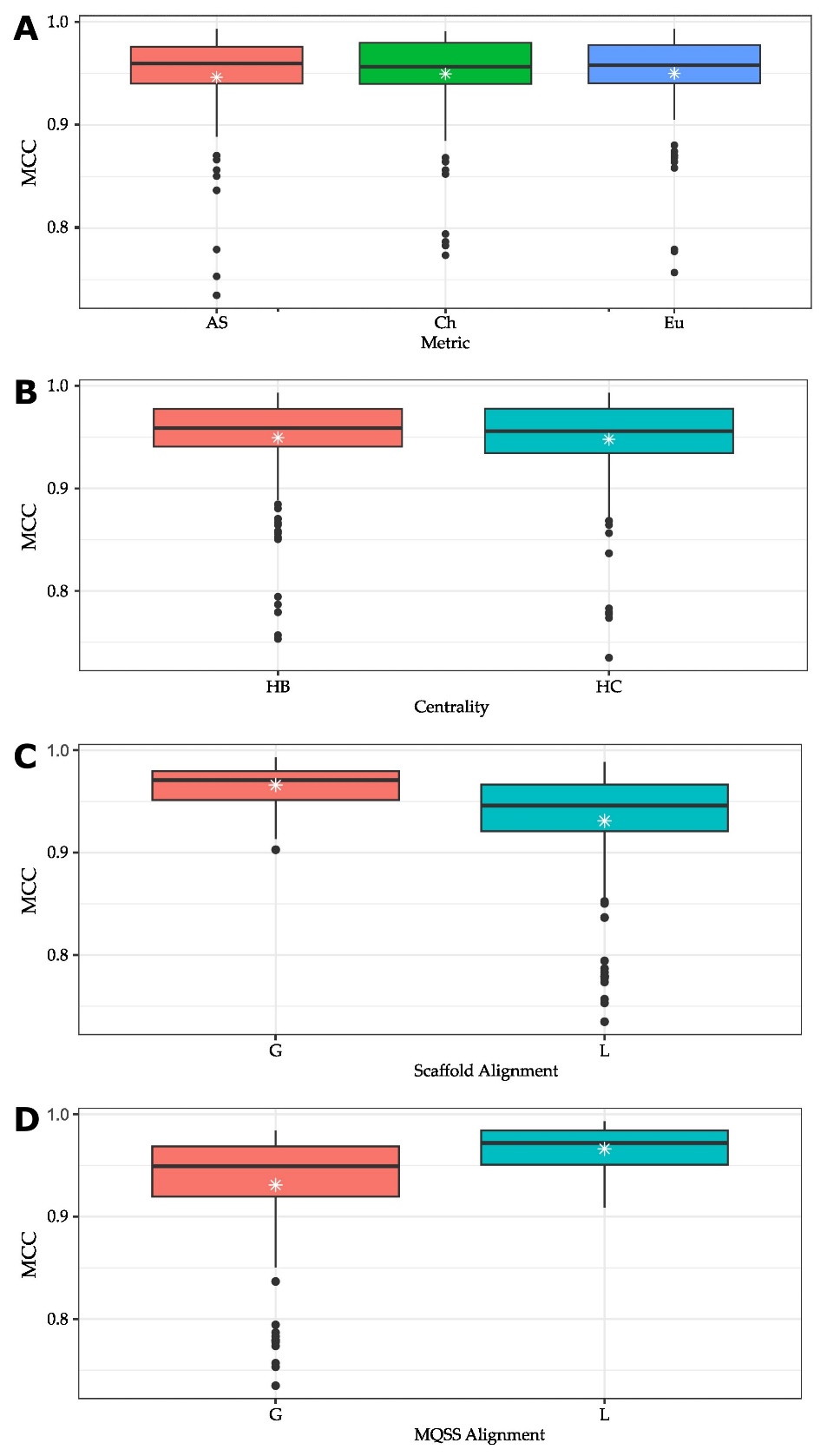
M.Sc. Pérez-Cárdenas Mariela

**Co-tutor*:***

Ph.D. Marrero-Ponce Yovani

Urcuquí, mayo 2023

**SM8.1**



**Figure 1.** Boxplots showing the relation betweenthe Matthews correlation coefficient (MCC) and scaffold/model parameters: **(A)** metric, **(B)** centrality, **(C)** scaffold alignment and (**D**) MQSS alignment. MCC was calculated after evaluating models on *HemoPI-1 Main* dataset. This figure was created with ggplot2 R package [1] and edited with Inkscape [2].

**SM8.2**

**Table 1.** Table showing the relation between the cutoff *s*, the average Matthews correlation coefficient (MCC) and the average number of sequences per scaffold.

|  |  |  |
| --- | --- | --- |
| Cutoff *s* | Average MCC | Average Number of Sequences per Scaffold |
| 0.4 | 0.550 | 272.1 |
| 0.5 | 0.716 | 386.2 |
| 0.6 | 0.851 | 518.1 |
| 0.7 | 0.933 | 659.6 |
| 0.8 | 0.964 | 841.6 |
| 0.9 | 0.974 | 1177.2 |

**References:**

1. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Use R!; 1st ed.; Springer New York, NY, 2009; ISBN 978-0-387-98141-3.

2. Inkscape. Inkscape Project 2023.