

# A Multi-objective Optimization Framework for Multiple Sequence Alignment with Metaheuristics

Cristian Zambrano-Vega<sup>1</sup>, Antonio J. Nebro<sup>2</sup>(✉), José García-Nieto<sup>2</sup>,  
and José F. Aldana-Montes<sup>2</sup>

<sup>1</sup> Facultad de Ciencias de la Ingeniería, Universidad Técnica Estatal de Quevedo,  
Quevedo, Los Ríos, Ecuador

czambrano@uteq.edu.ec

<sup>2</sup> Edificio de Investigación Ada Byron, University of Málaga, Málaga, Spain

{antonio,jnieto,jfam}@lcc.uma.es

<http://www.uteq.edu.ec/>

<http://khaos.uma.es/>

**Abstract.** The alignment of more than two biological sequences is a widely used technique in a number of areas of computational biology. However, finding an optimal alignment has been shown to be an NP-complete optimization problem. Furthermore, Multiple Sequence Alignment (MSA) can be formulated according to more than one score function, leading to multi-objective formulations of this problem. Due to these reasons, metaheuristics have been proposed to deal with MSA problems. In this paper, we present jMetalMSA, an Open Source software tool for solving MSA problems with multi-objective metaheuristics. Our motivation here is to offer to the scientific community in computational biology, a platform including state-of-the-art optimization algorithms aimed at solving different formulations of the MSA. We describe the main features of jMetalMSA, including the metaheuristics and scores that are currently available. In addition, we show a working example for illustration purposes.

**Keywords:** Multiple Sequence Alignment · Computational biology · Multi-objective optimization · Metaheuristics

## 1 Introduction

A Multiple Sequence Alignment (MSA) is an alignment of three or more biological sequences (ADN, ARN, protein) with the goal of identifying regions of similarity that may be a consequence of functional, structural, or evolution relationships among the sequences. MSA has many applications in field of computational biology, such as phylogenetic tree estimation, secondary structure prediction, and critical residue identification.

Finding an MSA can be defined as an optimization problem where a scoring function is to be maximized (or minimized). MSA is an NP-complete optimization problem [1], where the time complexity of finding an optimal alignment raises

© Springer International Publishing AG 2017

I. Rojas and F. Ortuño (Eds.): IWBBIO 2017, Part II, LNBI 10209, pp. 245–256, 2017.

DOI: 10.1007/978-3-319-56154-7\_23

In Springer Link

[https://link.springer.com/chapter/10.1007/978-3-319-56154-7\\_23](https://link.springer.com/chapter/10.1007/978-3-319-56154-7_23)