Adaptive Evolution of Falcipain-Homologues in *Plasmodium*

Kary A.C.S. Ocaña, Daniel de Oliveira, Marta Mattoso

PESC/COPPE - Federal University of Rio de Janeiro, Rio de Janeiro, Brazil.

{kary, danielc, marta}@cos.ufrj.br

Four out of ten tropical neglected diseases target as research priorities by the Special Program of World Health Organization for Research and Training in Tropical diseases are caused by protozoan species (malaria, leishmaniasis, Chagas disease and African trypanosomiasis). One example is *Plasmodium falciparum*, which is a human parasite that causes the most severe malaria cases. There is a large number of antimalarial drugs, including quinine and chloroquine. However the misuse of these drugs has led to growing resistance of this protozoan. Due to that, new antimalarial inhibitors are needed and several drug discovery projects focus on researching cysteine protease family, which may provide useful new drug targets. Bioinformatics methodologies have demonstrated to be an efficient alternative for achieving this treatment goal. Evidence indicates that cysteine proteases play essential role in the medical treatment of malaria parasites. The best characterized *Plasmodium* cysteine proteases are called falcipains, which are considered as potential targets for antimalarial chemotherapy, since they are involved in important cellular functions in the parasite life cycle.

The main goal of this paper is to analyze the adaptive evolutionary history of falcipain (and their homologues) of *Plasmodium* species using bioinformatics programs, emphasizing the adaptive behavior of its active sites. This way, we used a phylogenetic-based methodology to (i) infer phylogenetic relationship between *Plasmodium* species, (ii) indicate the ratio of non-synonymous to synonymous substitutions (ω) to assess the positive Darwinian selection ($\omega>1$), and (iii) show the evolutionary behavior of the active sites that characterize falcipains in a Three-Dimensional (3D) structure. Also, to aid the management of this experiment, a bioinformatics scientific workflow was designed to be executed in parallel using cloud computing environments. The workflow activities are modeled as following: (i) falcipain's coding-sequences identification using the HMMER package and RefSeq database, (ii) alignments construction using the MAFFT program, (iii) evolutionary models election using ModelGenerator, (iv) phylogenetic tree construction with RAxML, (v) evolutionary analysis (M3-discrete and M8-beta&ω models) using codeml from PAML package, (vi) Likelihood Radio Test (LRT) calculation for comparing two evolutionary nested model pairs (M0*vs.*M3 and M7*vs.*M8), (vii) falcipain-homologues search in PDB database using BLAST and (viii) falcipain's active sites map onto the PDB crystal structure using Molsoft ICM browser for structural manipulations.

As initial results, M3 indicates that falcipains present significant variability in sequences and M8 indicates that they also present a large proportion of sites under diversifying positive selection: (i) *P. falciparum*, 27% ($1.06<\omega<15.96$); (ii) *P. knowlesi*, 8% ($1.14<\omega<250.06$) and (iii) *P. vivax*, 10% ($1.43<\omega< 530.97$). Moreover, the LRT verified the variability and positive selection in genes showing that statistics are in favor of presence of some predicted amino acid sites to be subject to positive selective pressure ($\omega>1$). Four sites of them were identified as having high posterior probability ($P>95$) evolving under positive selection, which were identified in a 3D falcipain structure (PDB file 1YVB). Interestingly none of these sites matched with the main four active sites that characterize the cysteine protease family (*(C) Cys-42, (H) His-174*, and *(N) Asn-204*).

Our results suggest that falcipains present evidence of variation in selective pressure among sites and that it has not significant evidence of positive selection in falcipain's active sites. The level of variability acting on falcipains in *Plasmodium* species indicates that adaptive molecular evolution plays an important role in the emergence of the novel functions, which probably contributes to the biological trait and adaptation of these parasites. Further evolutionary life hypotheses, functional characterization and proteomic studies are required to augment our current understanding of this enzyme and its functional role in *Plasmodium*.