Enrichment of repeats in proteins of protozoan parasites relates to the ability to invade host cells and evade host immune response

Tiago Antônio de Oliveira Mendes, Francisco Pereira Lobo, Thiago Souza Rodrigues, Gabriela Flávia Rodrigues Luiz, Wanderson Duarte da Rocha, Ricardo Toshio Fujiwara, Santuza Maria Ribeiro Teixeira, Daniella Castanheira Bartholomeu
UFMG, EMBRAPA, CEFET-MG, UFMG, UFPR, UFMG, UFMG, UFMG

Background: Proteins containing repetitive amino acid domains are widespread in all life forms. In parasitic organisms, proteins containing repeats play important roles, such as cell adhesion and invasion and immune evasion. Therefore, extracellular and intracellular parasites are expected to be under different selective pressure regarding the repetitive content in their genomes. Here, we investigated whether there is a bias in the repetitive content found in the predicted proteomes of six exclusively extracellular, and seventeen obligate intracellular protozoan parasites, as well as four free-living protists and tried to correlate the results with the distinct ecological niches they occupy and with distinct protein functions. Results: We found that intracellular parasites have higher repetitive content in their proteomes than do extracellular parasites (t-test with Bonferroni correction, $p<0.0001$) and free-living protists (t-test with Bonferroni correction, $p=0.0014$). In intracellular parasites, these repetitive proteins are located mainly at the parasite surface or are secreted (t-test with Bonferroni correction, $P<0.0001$) and are enriched by amino acids known to be part of N- and O-glycosylation sites (t-test with Bonferroni correction, $p=0.0018$). Furthermore, intracellular parasites have higher content of proteins with perfect repeats that are expressed by the developmental stages capable to invade mammalian host cells (Chi-square test, $p=0.019$), in comparison to the other stages, and these proteins have molecular functions associated with cell invasion. On the other hand, in extracellular parasites, degenerate repetitive motifs are enriched in proteins that are likely to play a role in evading host immune response (Chi-square, $p<0.0001$). Conclusions: Taken together, our results support the hypothesis that both the ability to invade host cells and to escape the host immune response may have shaped the expansion and maintenance of perfect and degenerate repeats in the genomes of intra- and extracellular parasites.

Keywords: adaptive evolution, repeat proteins, protozoan parasites, genome evolution, host-parasite interaction
Concentration area: Genomics Evolution