FungiDB: Using regular expressions to identify motifs

1. Identify oomycete-specific motif RxLR

Oomycetes infecting plants employ effector proteins to establish a successful infection. Effector proteins are commonly secreted and one of the largest group is the RxLR effectors, which is defined by the conserved N-terminal Arg-Xaa-Leu-Arg (RxLR) motif.

a) Navigate to Search for Genes panel, and click on the Sequence Analysis, and Protein Motif Pattern

b) Search for the RxLR motif, but think carefully as to how to you would represent this motif using regular expressions

   a. Write a regular expression that defines a protein sequence of RXLR as r.lr, where X (.) is any aminoacid.

Oomycetes have several effector proteins with known domains.

c) How many of the genes identified also have a transmembrane domain?

   a. Add Step, ... Protein Targeting and Localization, Transmembrane Domain Count