A Web Application for Efficient Analysis of Peptide Libraries

Eric Riemer Hare

January 14, 2014

1 Introduction

With the introduction of the R package *peptider*, analysis of the statistical properties of various peptide libraries is now possible. However, use of this package still requires familiarity with the R language, and more general programming concepts. In this paper, I introduce a new web interface called *PeLiCa* (or Peptide Library Calculator) which provides a useable and flexible front-end to peptider. PeLiCa is designed for biologists and others working with peptide libraries, and does not require programming knowledge to conduct an analysis.

2 Structure

PeLiCa is a Shiny application[REF]. Shiny is a framework for writing web-applications in the R language, requiring little-to-no javascript programming knowledge. PeLiCa uses this framework to provide interactivity. For instance, when the user of PeLiCa changes a property of the peptide library, such as the encoding scheme, the results, tables, and plots will instantly update to reflect the new library. PeLiCa is currently hosted on the Glimmer server provided by the RStudio team.

3 User Interface

Similar to other Shiny applications, PeLiCa consists of three primary UI components, the Configuration Panel, the Tab Panel, and the Results Panel, each illustrated in Figure 1. The Configuration Panel is located along the left-hand column. This panel allows for various parameters of peptide libraries to be adjusted, and some configuration options relating to PeLiCa to be changed depending on user preference. The top panel is the Tab Panel, which contains various tabs corresponding to different properties of peptide libraries that can be investigated. The bottom panel is the Results panel, which will contain the results of the analysis depending on the tab and configuration options selected.

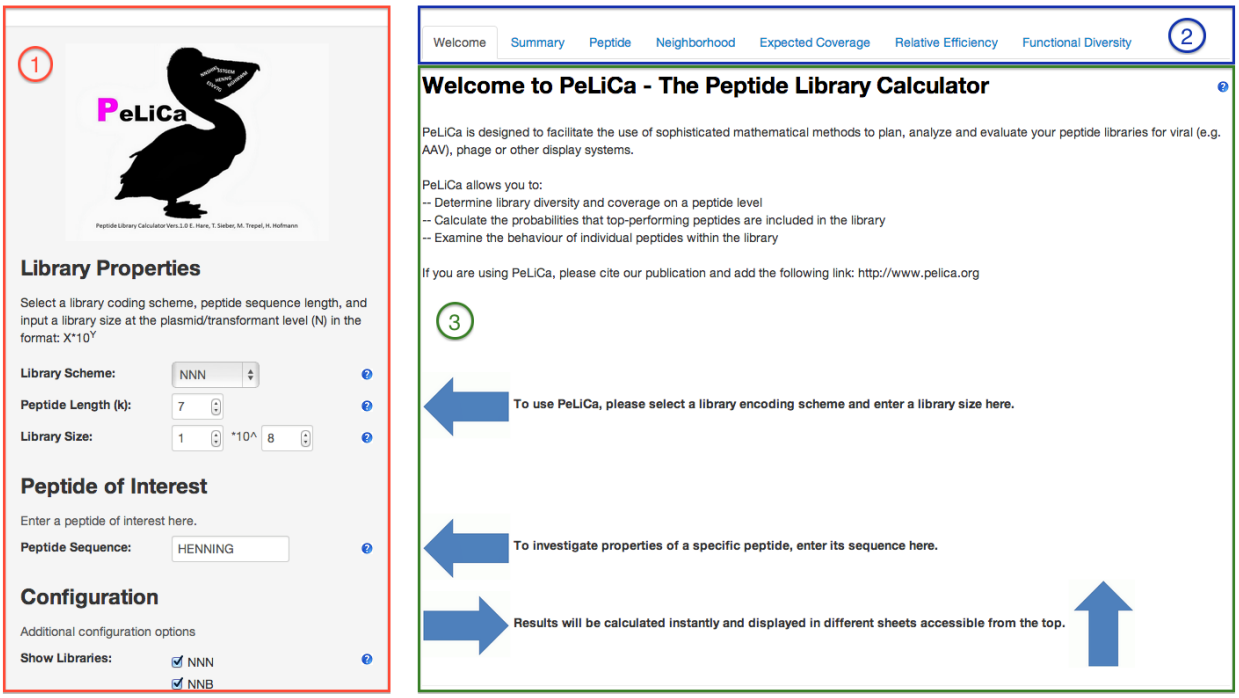

Figure 1: Screenshot of PeLiCa indicating the three primary UI components. (1) The Configuration Panel (2) The Tab Panel (3) The Results Panel.

Another important component of the user interface are the help tooltips. Throughout the application, blue question marks will be available. When the user moves their mouse cursor over these icons, helpful tooltips will appear to instruct the user on how to proceed, or provide more information about the library property being investigated.

4 Features

4.1 Configuration Panel

The first set of features available in PeLiCa involve the specification of properties of the library of interest. These features are displayed in Figure 2. Users can first select a library scheme. PeLiCa provides several built-in library schemes, the first of which is the NNN scheme, in which all four bases (Adenine, Guanine, Cytosine, and Thymine) can occur at all three positions in a particular codon. The second is the NNB scheme, where the first two positions are unrestricted, but the third position can only be three bases. The third is NNK/S, which covers both NNK and NNS schemes, with the third position restricted to two bases. Both NNK and NNS have identical statistical properties in the analysis available in PeLiCa[REF]. Finally, there are trimer-based libraries in which the codons are pre-defined. PeLiCa also includes variations of these four library types in which Cysteine is treated as a non-viable amino acid.

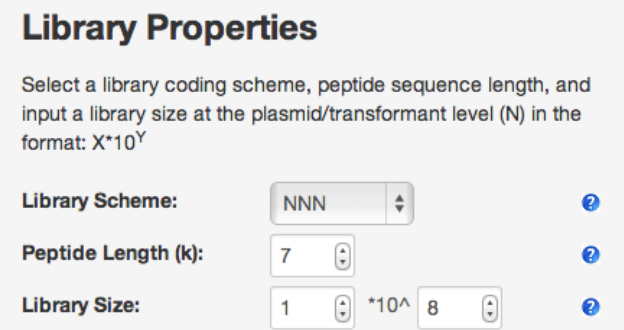

Figure 2: The Library Properties section of the Configuration Panel.

PeLiCa also has support for user-defined library schemes. If the user selects "Custom Scheme" for the library scheme, they will be presented with an upload dialog, along with a set of instructions for uploading a custom scheme. Using a custom scheme with PeLiCa requires a minimal use of programming and may be less suitable for those unfamiliar with R.

Users can also specify the peptide length and the library size. Peptide lengths can range from six to ten amino acids, with support for larger peptides coming soon. The library size is specified in scientific notation of the form $x \times 10^y$. The user will specify values for x and y in this equation. x can range from 1.0 to 9.9 in increments of 0.1, and y can range from six to 14 in increments of one, yielding a supported range of library sizes between 1.0×10^6 and 9.9×10^{14}. Support for large library sizes is also in progress.

Users can then specify a peptide of interest, as shown in Figure 3, and configure which other libraries to display in the Results Panel in Figure 4.

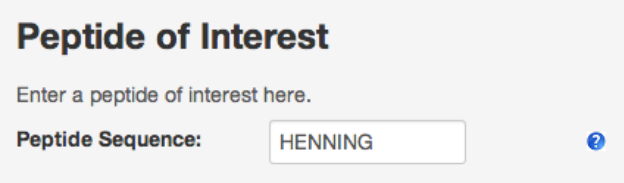

Figure 3: The Peptide of Interest section of the Configuration Panel.

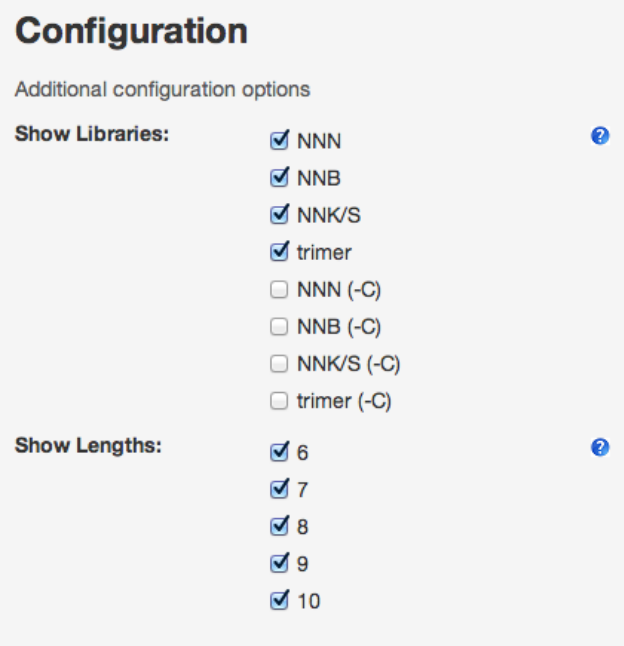

Figure 4: The Configuration section of the Configuration Panel.

4.2 Results Panel

The primary functionality for PeLiCa is available in the Results Panel, which are accessible through each of the tabs in the Tab Panel.

4.2.1 Welcome

PeLiCa begins on the Welcome tab. The Welcome tab provides information on the functionality of PeLiCa, and a quick guide for its use. The tooltip on this tab illustrates the system requirements.

4.2.2 Summary

The Summary tab contains most of the key information from the other tabs, condensed into an easy-to-digest format. First, information on your library is displayed. Some of this information includes the coverage, the peptide diversity, and the probabilities of peptide inclusion. Information on your selected peptide is displayed below this, summarizing the inclusion probability and the number of different DNA encodings of this particular peptide. Finally, a table displaying a randomly generated sample of peptides is shown at the bottom. This table includes the amino acids comprosing the peptide, the peptide class under the chosen encoding scheme, the number of DNA encodings, and the probability of inclusion in your library.

4.2.3 Peptide

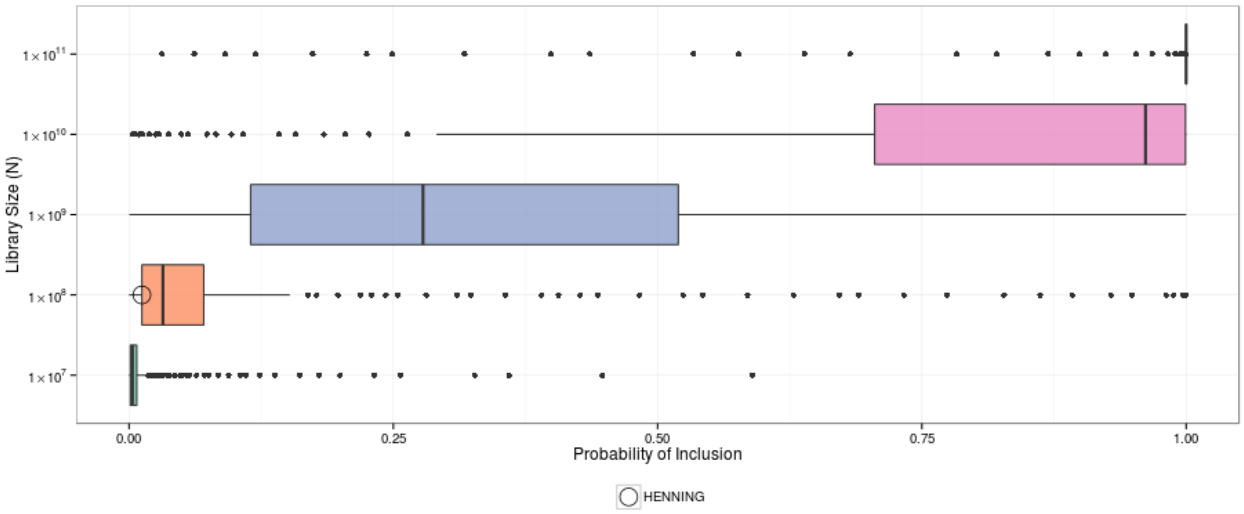

Figure 5: Boxplots displayed in PeLiCa representing the inclusion probabilities for an NNN library with peptide length seven and a library size of 1×10^8. The inclusion probability of HENNING is displayed in the plot as a circle.

4.2.4 Neighborhood

4.2.5 Expected Coverage

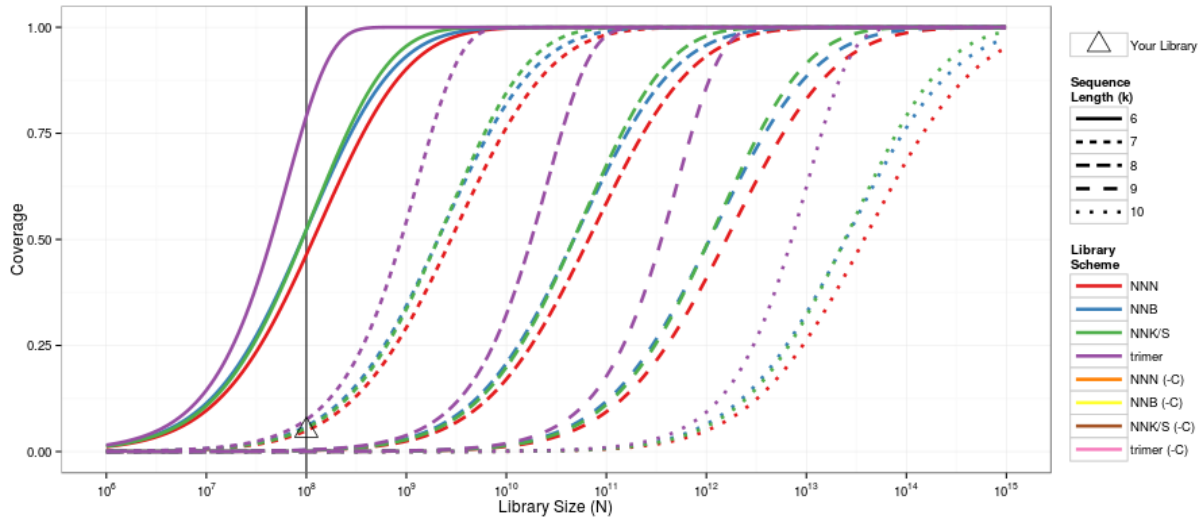

Figure 6: Plots of expected coverage for an NNN library with peptide length seven and a library size of 1×10^8

4.2.6 Relative Efficiency

4.2.7 Functional Diversity

5 Further Work

A new version of PeLiCa is in currently in progress. The new version supports lower resolution monitors, includes a new framework for tooltips, and supports a wider range of peptide lengths and library sizes. This new version is currently deployed on ShinyApps at `http://erichare.shinyapps.io/pelica`.

6 Conclusion