BMC
Bioinformatics

SLiMScape: a protein short linear motif analysis plugin for Cytoscape

Kevin T O'Brien[1,2,3], Niall J Haslam[1,2,3] and Denis C Shields[1,2,3]*

Abstract

Background: Computational protein short linear motif discovery can use protein interaction information to search for motifs among proteins which share a common interactor. Cytoscape provides a visual interface for protein networks but there is no streamlined way to rapidly visualize motifs in a network of proteins, or to integrate computational discovery with such visualizations.

Results: We present SLiMScape, a Cytoscape plugin, which enables both de novo motif discovery and searches for instances of known motifs. Data is presented using Cytoscape's visualization features thus providing an intuitive interface for interpreting results. The distribution of discovered or user-defined motifs may be selectively displayed and the distribution of protein domains may be viewed simultaneously. To facilitate this SLiMScape automatically retrieves domains for each protein.

Conclusion: SLiMScape provides a platform for performing short linear motif analyses of protein interaction networks by integrating motif discovery and search tools in a network visualization environment. This significantly aids in the discovery of novel short linear motifs and in visualizing the distribution of known motifs.

Background

High throughput experiments have greatly increased the number of known protein-protein interactions in the human interactome. However, most of these experiments do not indicate the mechanism of interaction. Many of these interactions occur between large globular domains but an estimated 15 - 40% are mediated by functional microdomains [1], of 3 - 10 amino acids in size called Short Linear Motifs (SLiMs). These often occur in intrinsically disordered regions and are involved in a number of functions such as binding, cleavage, subcellular targeting and post translational modifications.

Several predictive tools using many different methods have been developed to identify SLiMs [2-8]. We have developed SLiMScape, a plugin which uses two of these tools to add SLiM discovery and search functionality to Cytoscape [9]. Both novel and known motifs can be detected, providing a powerful tool for exploration and analysis of SLiMs within a protein interaction network. Cytoprophet [10], another Cytoscape plugin, predicts domain and protein interaction networks however SLiMScape is the first to focus on SLiMs.

Implementation

SLiMScape has been tested on Cytoscape version 2.8.3 and on all major platforms (Linux, Windows and MacOS). It is published under the GNU Lesser General Public License version 3.0. SLiMscape depends on several websites such as the SMART [11] domain database, Uniprot [12] protein database, DBFetch [13] and custom SLimFinder [2] and SLiMSearch [5] web services available at bioware.ucd.ie. For this reason web access is required. It is also possible to run SLiMFinder and SLiMSearch locally, however this requires installation of several other tools. A tutorial is available at bioware.ucd.ie/slimscape.

Results and discussion

SLiMscape consists of two primary analysis tools. The first, SLiMFinder [2], is capable of finding novel motifs by analyzing a protein interaction network. SLiMFinder has previously been demonstrated as a suitable tool for mining the human interactome for SLiM mediated interactions

*Correspondence: denis.shields@ucd.ie
[1]UCD Conway, Institute of Biomolecular and Biomedical Sciences, University College Dublin, Dublin, Ireland
[2]UCD Complex and Adaptive Systems Laboratory, University College Dublin, Dublin, Ireland
Full list of author information is available at the end of the article

[14]. The second, SLiMSearch [5] can find new instances of known or potential motifs within the same network. An important feature is the visualization of results within a protein interaction network and in the context of domains in nearby proteins.

Novel motif discovery

SLiMScape uses SLiMFinder to locate statistically over-represented sequences within sets of proteins, defined by selecting subsets of a protein interaction network. A variety of node selection methods are provided to automatically select protein sets and different types of motif can be targeted depending on this selection process. For example, SLiM-mediated interactions can be discovered using the "Batch Interactions" selection method. This will iteratively create protein sets containing each protein (hub) and its interactors (spokes). Figure 1 shows the results of this method on a network consisting of Proliferating Cell Nuclear Antigen (PCNA) and its interactors. Since, in a large network, this can be computationally demanding selection methods include carrying out a single search across a set of selected nodes, specified interactively through Cytoscape's interface. Another selection method named "Attribute" creates protein sets based on a user-defined attribute. A useful example is subcellular location, which can be used to find subcellular targeting motifs.

The "Batch Domains" selection method creates sets of proteins which contain a particular domain, while "Batch Domains Interactions" creates sets of proteins which interact with a particular domain. This is useful when searching for self interacting proteins and other sequences which bind to a domain. Protein sets can also be created manually, allowing for many custom strategies for finding motifs.

Known motif search

New instances of a motif can be located using regular expression matching. This is especially useful when trying to find new instances of a SLiM found using SLiMFinder and can also be used to search for a user defined SLiM. Disorder and conservation masking are used to reduce the number of false positives. This functionality is implemented using SLiMSearch.

Visualization of results

SLiMScape presents an interactive view of results within a protein interaction network. SLiMs are first presented in a table and when activated all instances of that SLiM are highlighted, thus displaying the distribution of the SLiM. Additionally, domain annotations are provided to indicate possible SLiM-domain interactions and aid in the interpretation of results.
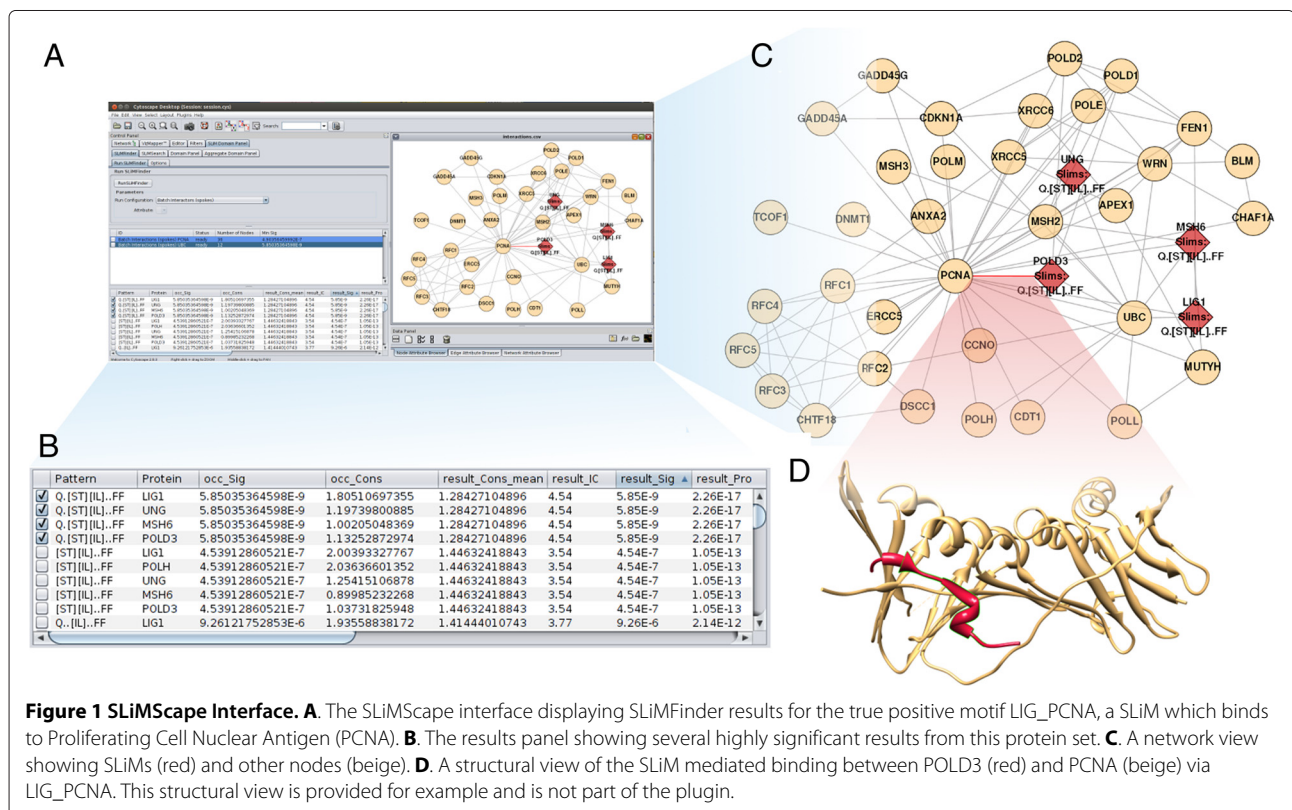

Figure 1 SLiMScape Interface. A. The SLiMScape interface displaying SLiMFinder results for the true positive motif LIG_PCNA, a SLiM which binds to Proliferating Cell Nuclear Antigen (PCNA). **B**. The results panel showing several highly significant results from this protein set. **C**. A network view showing SLiMs (red) and other nodes (beige). **D**. A structural view of the SLiM mediated binding between POLD3 (red) and PCNA (beige) via LIG_PCNA. This structural view is provided for example and is not part of the plugin.

Conclusion

SLiMScape integrates SLiM discovery tools with Cytoscape, a visualization program popular in the field of bioinformatics. This provides a platform for SLiM discovery which allows discovery of novel and known motifs within a protein interaction network. This increases the speed at which SLiM analyses can be performed and aids in the interpretation of results.

Availability and requirements

- **Project name:** SLiMScape
- **Project home page:**
  http://sourceforge.net/projects/slimscape/
- **Operating system(s):** Platform independent
- **Programming language:** Java, Python
- **Other requirements:** Cytoscape 2.8.3, Java 1.6 or higher, python 2.7 (optional), Slimsuite (optional), Blast (optional), IUPred (optional), Muscle (optional) and Clustal (optional).
- **License:** GNU LGPLv3
- **Any restrictions to use by non-academics:** see license

Author details
[1] UCD Conway, Institute of Biomolecular and Biomedical Sciences, University College Dublin, Dublin, Ireland. [2] UCD Complex and Adaptive Systems Laboratory, University College Dublin, Dublin, Ireland. [3] School of Medicine and Medical Science, University College Dublin, Dublin, Ireland.

References

1. Neduva V, Russell RB: **Peptides mediating interaction networks: new leads at last.** *Current Opinion in Biotechnol* 2006, **17**(5):465–71.
2. Edwards RJ, Davey NE, Shields DC: **SLiMFinder: a probabilistic method for identifying over-represented, convergently evolved, short linear motifs in proteins.** *PLoS one* 2007, **2**(10):e967.
3. Neduva V, Linding R, Su-Angrand I, Stark A, de Masi F, Gibson TJ, Lewis J, Serrano L, Russell RB: **Systematic discovery of new recognition peptides mediating protein interaction networks.** *PLoS Biology* 2005, **3**(12):e405.
4. Dinkel H, Michael S, Weatheritt RJ, Davey NE, Van RoeyK, Altenberg B, Toedt G, Uyar B, Seiler M, Budd A, Jödicke L, Dammert MA, Schroeter C, Hammer M, Schmidt T, Jehl P, McGuigan C, Dymecka M, Chica C, Luck K, Via A, Chatr-Aryamontri A, Haslam N, Grebnev G, Edwards RJ, Steinmetz MO, Meiselbach H, Diella F, Gibson TJ: **ELM–the database of eukaryotic linear motifs.** *Nucleic Acids Res* 2012, **40**(Database issue):D242—D251.
5. Davey NE, Haslam NJ, Shields DC, Edwards RJ: **SLiMSearch 2.0: biological context for short linear motifs in proteins.** *Nucleic Acids Res* 2011, **39**(Web Server issue):W56–W60.
6. Davey NE, Cowan JL, Shields DC, Gibson TJ, Coldwell MJ, Edwards RJ: **SLiMPrints: conservation-based discovery of functional motif fingerprints in intrinsically disordered protein regions.** *Nucleic Acids Res* 2012, **40**(21):10628–10641.
7. Trabuco LG, Lise S, Petsalaki E, Russell RB: **PepSite: prediction of peptide-binding sites from protein surfaces.** *Nucleic acids research* 2012, **40**(Web Server issue):W423–W427.
8. Mooney C, Pollastri G, Shields D C, Haslam NJ: **Prediction of Short Linear Protein Binding Regions.** *J Mol Biol* 2011.
9. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: a software environment for integrated models of biomolecular interaction networks.** *Genome Res* 2003, **13**(11):2498–2504.
10. Morcos F, Lamanna C, Sikora M, Izaguirre J: **Cytoprophet: a Cytoscape plug-in for protein and domain interaction networks inference.** *Bioinformatics (Oxford, England)* 2008, **24**(19):2265–2266.
11. Letunic I, Doerks T, Bork P: **SMART 7: recent updates to the protein domain annotation resource.** *Nucleic Acids Res* 2012, **40**(Database issue):D302–D305.
12. The UniProt Consortium: **Reorganizing the protein space at the Universal Protein Resource (UniProt).** *Nucleic Acids Res* 2012, **40**(Database issue):D71–D75.
13. Lopez R, Duggan K, Harte N, Kibria A: **Public services from the European Bioinformatics Institute.** *Briefings in bioinformatics* 2003, **4**(4):332–340.
14. Edwards RJ, Davey NE, O'Brien K, Shields DC: **Interactome-wide prediction of short, disordered protein interaction motifs in humans.** *Mol bioSystems* 2012, **8**:282–295.