Title: **CSFGBro, a Browser Companion for Biocurators of Fungal Genomics Literature**

Authors: Vahé Chahinian, Marie-Jean Meurs, Erin McDonnell, Ingo Morgenstern, Greg Butler and Adrian Tsang

Institution: Centre for Structural and Functional Genomics, Concordia University, Montréal, QC, Canada

Abstract:

Discovery and development of effective fungal enzyme cocktails are cornerstones of the biorefinery industry because these cocktails can convert lignocellulose into fermentable sugars for biofuels production.
In this context, the manual curation of fungal genes encoding lignocellulose-active enzymes is an essential step for supporting further research and experiments, as it allows researchers to easily access reliable knowledge.

We present the ongoing development of CSFGBro, an augmented browsing tool supporting the manual curation of literature related to genomics-based lignocellulose research.

In Web pages, CSFGBro highlights information automatically retrieved by the GATE-based mycoMINE text mining system.
Developed in close collaboration with researchers working on the Genozymes [http://www.fungalgenomics.ca] project, mycoMINE annotates documents with tags such as enzyme, fungus, activity or pH/Temperature conditions.
Tag categories of interest have been selected according to those reported in the mycoCLAP database [http://mycoclap.fungalgenomics.ca] of fungal genes encoding biochemically characterized lignocellulose-active proteins.

CSFGBro provides researchers and biocurators with an overview of the document content.
A sidebar displays mentions retrieved in the document.
Clicking on a highlighted mention opens a popup, which shows the features of the mention, any available standard identifiers, and hyperlinks to external sources of knowledge.

For instance, an enzyme popup displays the enzyme EC Number, its recommended and systematic names reported on BRENDA, its SwissProt identifiers, and hyperlinks to related pages (BRENDA).

The CSFGBro backend is composed of a RESTful interface programmed in server side JavaScript [http://nodejs.org/].
The interface converts XML mycoMINE results into HTML, adds highlighting, creates a sidebar menu listing them and adds the on-click functionality.

The CSFGBro front end is a user script that sends requests to the interface to retrieve the data and then displays it. The user script is programmed in JavaScript and runs in the browser itself.

Two curators evaluated a prototype with limited features on the triage of 114 PubMed abstracts. Using the tool, the time needed for triage was reduced by 21%, showing the relevance of the approach.