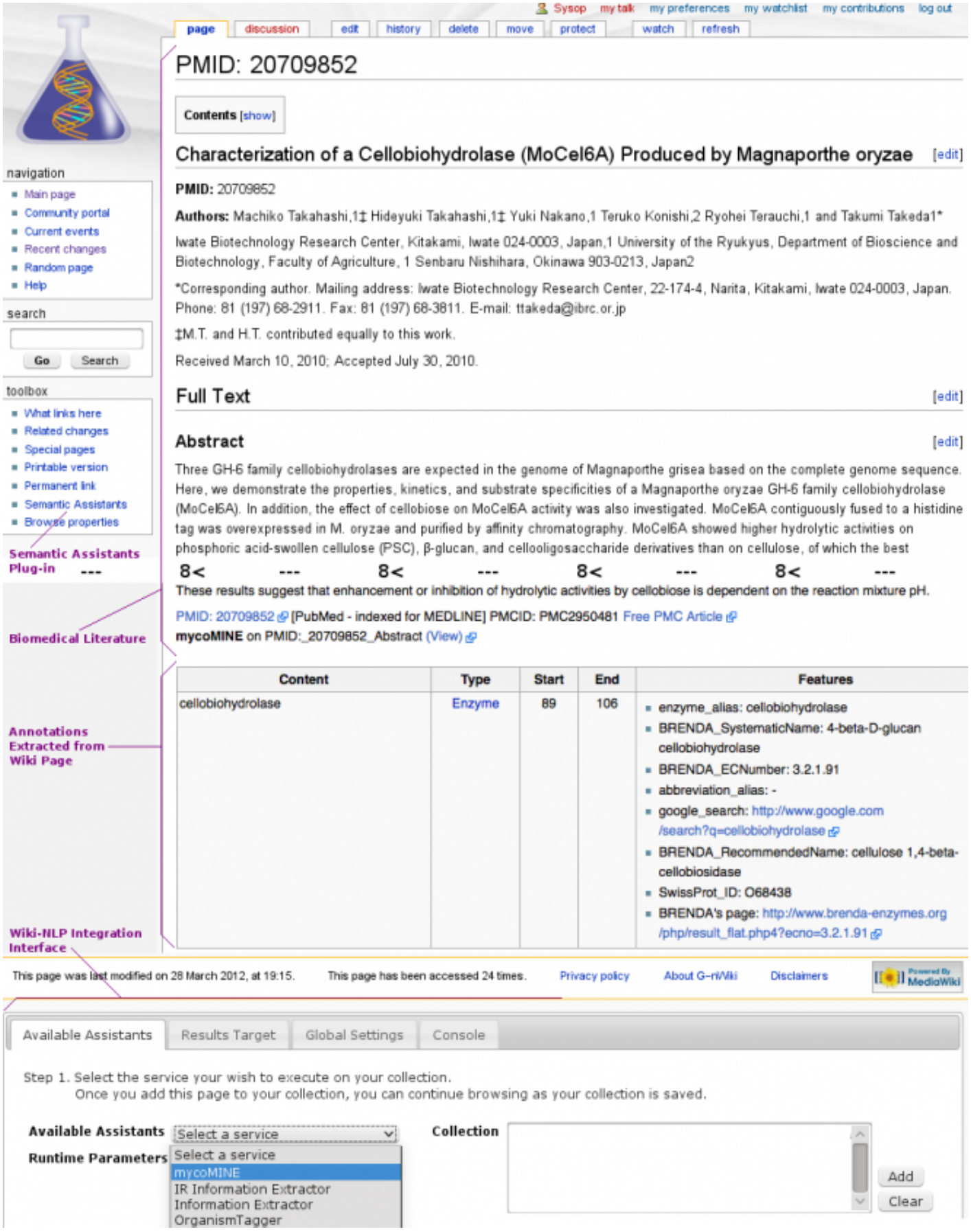
IntelliGenWiki: Intelligent Semantic Wikis for Life Sciences

- Genozymes
- Semantic Assistants
- Semantic Wiki

toc_collapse=0; Table of Contents

1. Overview

Researchers need to extract and manage critical knowledge from the massive amount of literature available in multiple and ever-growing repositories. The sheer volume of information makes the exhaustive analysis of literature a labor-intensive and time-consuming task, during which significant knowledge can be easily missed. We present *Intelli*GenWiki, a service-oriented solution that combines state-of-the-art techniques from the Natural Language Processing (NLP) and Semantic Web domains to support the knowledge discovery workflow in omics sciences. For a brief description of *Intelli*GenWiki, please see our paper [1], Sateli, B., M. - J. Meurs, G. Butler, J. Powlowski, A. Tsang, and R. Witte, "IntelliGenWiki: An Intelligent Semantic Wiki for Life Sciences", *NETTAB 2012*, vol. 18 (Supplement B), Como, Italy : EMBnet.journal, pp. 50–52, 11/2012.


Sysop my talk my preferences my watchlist my contributions log out
page discussion edit history delete move protect watch refresh
navigation
Main page
Community portal
Current events
Recent changes
Random page
Help
search
Go Search
toolbox
What links here
Related changes
Special pages
Printable version
Permanent link
Semantic Assistants
Browse properties
Semantic Assistants Plug-in
Biomedical Literature
Annotations Extracted from Wiki Page
Wiki-NLP Integration Interface
PMID: 20709852
Contents [show]
Characterization of a Cellobiohydrolase (MoCel6A) Produced by Magnaporthe oryzae [edit]
PMID: 20709852
Authors: Machiko Takahashi,1‡ Hideyuki Takahashi,1‡ Yuki Nakano,1 Teruko Konishi,2 Ryohei Terauchi,1 and Takumi Takeda1*
Iwate Biotechnology Research Center, Kitakami, Iwate 024-0003, Japan,1 University of the Ryukyus, Department of Bioscience and Biotechnology, Faculty of Agriculture, 1 Senbaru Nishihara, Okinawa 903-0213, Japan2
*Corresponding author. Mailing address: Iwate Biotechnology Research Center, 22-174-4, Narita, Kitakami, Iwate 024-0003, Japan. Phone: 81 (197) 68-2911. Fax: 81 (197) 68-3811. E-mail: ttakeda@ibrc.or.jp
‡M.T. and H.T. contributed equally to this work.
Received March 10, 2010; Accepted July 30, 2010.
Full Text [edit]
Abstract [edit]
Three GH-6 family cellobiohydrolases are expected in the genome of Magnaporthe grisea based on the complete genome sequence. Here, we demonstrate the properties, kinetics, and substrate specificities of a Magnaporthe oryzae GH-6 family cellobiohydrolase (MoCel6A). In addition, the effect of cellobiose on MoCel6A activity was also investigated. MoCel6A contiguously fused to a histidine tag was overexpressed in M. oryzae and purified by affinity chromatography. MoCel6A showed higher hydrolytic activities on phosphoric acid-swollen cellulose (PSC), β-glucan, and cellooligosaccharide derivatives than on cellulose, of which the best
8< --- 8< --- 8< --- 8< ---
These results suggest that enhancement or inhibition of hydrolytic activities by cellobiose is dependent on the reaction mixture pH.
PMID: 20709852 [PubMed - indexed for MEDLINE] PMCID: PMC2950481 Free PMC Article
mycoMINE on PMID:_20709852_Abstract (View)
Content Type Start End Features
cellobiohydrolase Enzyme 89 106
enzyme_alias: cellobiohydrolase
BRENDA_SystematicName: 4-beta-D-glucan cellobiohydrolase
BRENDA_ECNumber: 3.2.1.91
abbreviation_alias: -
google_search: http://www.google.com/search?q=cellobiohydrolase
BRENDA_RecommendedName: cellulose 1,4-beta-cellobiosidase
SwissProt_ID: O68438
BRENDA's page: http://www.brenda-enzymes.org/php/result_flat.php4?ecno=3.2.1.91
This page was last modified on 28 March 2012, at 19:15.
This page has been accessed 24 times.
Privacy policy
About G-riWiki
Disclaimers
Powered By MediaWiki
Available Assistants
Results Target
Global Settings
Console
Step 1. Select the service your wish to execute on your collection.
Once you add this page to your collection, you can continue browsing as your collection is saved.
Available Assistants
Select a service
Runtime Parameters
Select a service
mycoMINE
IR Information Extractor
Information Extractor
OrganismTagger
Collection
Add
Clear

IntelliGenWiki integration applied to MediaWiki

2. Features

*Intelli*GenWiki is a novel approach based entirely on open source and open standards that empowers a wiki-based literature curation environment by applying new Human-AI collaboration patterns through integrating text mining "assistants" that work collaboratively with humans on the literature. Such assistants aid curators to extract knowledge from text and produce machine-readable metadata that can be used in an Open Data context.

Rather than a new wiki system, *Intelli*GenWiki provides a cohesive, generic architecture that can be applied to various wiki engines. The resulting integration seamlessly provides the curators with NLP capabilities deployed in the GATE architecture and brokered through the Semantic Assistants framework as web services. This way, the complexity of executing NLP pipelines are hidden from the curators' point of view, rather, the available semantic assistants in the wiki work collaboratively with curators on wiki pages to automatically extract entities of interest from its textual content, such as organisms, genes or enzymes.

severely inhibited at pH 9.0. These results suggest that enhancement or inhibition of hydrolytic activities by cellobiose is dependent on the reaction mixture pH.

PMID: 20709852 [PubMed - indexed for MEDLINE] PMCID: PMC2950481 Free PMC Article
mycoMINE on PMID:_20709852_Abstract (View)

Content	Type	Start	End	Features
cellobiohydrolase	Enzyme	103	120	▪ enzyme_alias: cellobiohydrolase ▪ BRENDA_SystematicName: oligoxyloglucan reducing-end cellobiohydrolase ▪ BRENDA_ECNumber: 3.2.1.150 ▪ abbreviation_alias: - ▪ google_search: http://www.google.com/search?q=cellobiohydrolase ▪ BRENDA_RecommendedName: oligoxyloglucan reducing-end-specific cellobiohydrolase ▪ SwissProt_ID: - ▪ BRENDA's page: http://www.brenda-enzymes.org/php/result_flat.php4?ecno=3.2.1.150
Magnaporthe oryzae	Organism	143	161	▪ NCBI_Taxonomy_WebPage: http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=318829&mode=info ▪ organism_scientific_name: Magnaporthe oryzae ▪ organism_alias: Magnaporthe oryzae ▪ google_search: http://www.google.com/search?q=Magnaporthe+oryzae ▪ NCBI_Taxonomy_ID: 318829

Automatic entity extraction from the literature in *Intelli*GenWiki

For example, the *Intelli*GenWiki has been applied to a MediaWiki instance in the context of the Genozymes project, where the wiki is used as a distributed, collaborative literature curation platform for lignocellulose research [2]. As shown in the figure above, the NLP-enabled wiki system features the MediaWiki native interface as well as the Wiki-NLP integration user interface generated by our Semantic Assistants plug-in installed on the wiki. The figure shows the results from our mycoMINE pipeline that automatically extracts various entities from the given paper abstract and retrieves related additional information from external resources, e.g., their scientific name or commission number from open databases – a task that is typically done manually by curators.

In addition to automatic information extraction from wiki content, *Intelli*GenWiki implicitly produces semantic metadata that can be exploited in various ways, e.g., to be exported to external repositories or to provide semantic entity retrieval capabilities in the wiki, where applicable. In the figure below, we illustrate how curators of the *Intelli*GenWiki can find wiki pages containing specific entities of their interest based on their type.

```
{{#ask: [[hasType::Enzyme]]
   | ?Enzyme = Enzyme Entities Found
   | format = table
   | headers = plain
   | default = No pages found!
   | mainlabel = Page Name
}}
```

Property:Enzyme

Page Name	Enzyme Entities Found
PMID: 20709852	Cellobiohydrolase Cellulases endoglucanases β-glucosidases Invitrogen DNA polymerase

Semantic Query in IntelliGenWiki (top) and retrieved entities (bottom)

3. Download & Installation

In order to use *Intelli*GenWiki, you need to set up a MediaWiki version 1.16 or later and install our Semantic Assistants plug-in [3] . You will also need to import the Semantic Assistants templates in your wiki. These templates together with the plug-in will be published soon under an open source license.

If you also want to install your own Semantic Assistants server (offering NLP services), you should obtain the complete Semantic Assistants Architecture.

4. IntelliGenWiki Video Introduction

IntelliGenWiki was presented at the NETTAB 2012 Workshop in Como, Italy in November, 2013. The video of the presentation is available on http://www.nettab.org/2012/videos/BaharSateli.html.

5. Acknowledgments

Funding for the development of *Intelli*GenWiki has been generously provided by NSERC.

References

1. Sateli, B., M. - J. Meurs, G. Butler, J. Powlowski, A. Tsang, and R. Witte, "IntelliGenWiki: An Intelligent Semantic Wiki for Life Sciences", *NETTAB 2012*, vol. 18 (Supplement B), Como, Italy : EMBnet.journal, pp. 50–52, 11/2012.
2. Sateli, B., C. Murphy, R. Witte, M. - J. Meurs, and A. Tsang, "Text Mining Assistants in Wikis for Biocuration", *5th International Biocuration Conference*, Washington DC, USA : International Society for Biocuration, pp. 126, 04/2012.
3. Sateli, B., and R. Witte, "Supporting Wiki Users with Natural Language Processing", *The 8th International Symposium on*

Wikis and Open Collaboration (WikiSym 2012), Linz, Austria : ACM, 08/2012.