

## IJS Internal Report

# Deliverable D1.1B: BioMine update at M24

National Institute of Biology, Jožef Stefan Institute

Version 1 FINAL

**Abstract:** We present the latest update of the BioMine system for plant databases. The network of plant genes built from important public knowledge identified in deliverable 3.2A (D3.2A) was incorporated into the application. Additionally, the latest exports for plant species were used to update the database from the version of 1 July 2016 to version 18 January 2018.

Document administrative information	
Project acronym:	HinLife
Project number:	J7-7303
Deliverable number:	D1.1b
Deliverable full title:	BioMine update at M24
Document identifier:	HinLife D1.1B v1.FINAL
Lead partner short name:	NIB
Report version:	Version 1, FINAL
Report preparation date:	25/01/2018
Lead author:	Vid Podpečan, Živa Ramšak
Co-authors:	Nada Lavrač, Kristina Gruden, Dragana Miljković
Status:	Final

## Introduction

The Task 1.1 of WP1 is focused on data preprocessing and transformation into a heterogeneous information network. More specifically, we are focused on updating the Biomine system, its heterogeneous networks, and developing an application which enables the exploration of these networks. Within the scope of this deliverable we present the update of the Biomine system for plant species from GoMapMan web application ([www.gomapman.org](http://www.gomapman.org)) and additionally, the inclusion of the manually built network from deliverable D3.2A into the Biomine graph.

## Results

The plant organism set in Biomine contains the following plant species:

- arabidopsis (*Arabidopsis thaliana*)
- potato (*Solanum tuberosum*)
- rice (*Oryza sativa*)
- tomato (*Solanum lycopersicum*)
- tobacco (*Nicotiana tabacum*)
- beet (*Beta vulgaris*)
- cacao tree (*Theobroma cacao*)
- pearl millet (*Pennisetum glaucum*)
- bread wheat (*Triticum aestivum*)

The last two species (pearl millet, bread wheat) were added in the timeframe between the latest export in 2016 and now, and hence represent an upgrade of the knowledge contained in Biomine. This is reflected in an increase in the number of nodes and links between both versions of the plant dataset (Table 1).

The second part of the upgrade of the plant database includes the network generated in deliverable D3.2A from publically available data sources, including protein-protein interactions, transcription factor regulation and regulation via small RNA. Based on the connections and their assigned reliability levels an export file was generated on GoMapMan application, available at [http://www.gomapman.org/export/current/biomine/ath\\_cnn](http://www.gomapman.org/export/current/biomine/ath_cnn). All binding reactions were assigned the interaction type '*interacts\_with*', while the remainder (both transcription factor regulation and regulation via small RNA) were assigned the interaction type '*affects*' (Figure 1).

Table 1: Comparison between the current and latest previous version of Biomine, including the number of nodes and links in each version, and the specific subset of databases included. Note, that v2016 does not include the CNN database, for which a specific export was generated and included as part of this deliverable.

	<b>Biomine plants v2018-01-18</b>	<b>Biomine plants v2016-07-01</b>
<b>Nodes</b>	1,348,527	897,651
<b>Links</b>	3,154,841	2,368,585
<b>DB: TAXON</b>	1.11	1.11
<b>DB: GO</b>	2018-01-18_01:38	2016-06-28_03:38
<b>DB: EntrezGene</b>	2018-01-18	2016-06-29
<b>DB: GoMapMan</b>	2018-01-18	2016-06-29
<b>DB: CNN</b>	2018-01-18	N/A
<b>DB: PubMed</b>	2018-01-18	2016-07-01

Figure 1: (A) An example of a connection as written in the export file for CNN on GoMapMan and (B) representation of that connection within the larger graph context in Biomine.

```

Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT1G04400)
Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT1G22770)
Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT1G32900)
Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT1G65480)
Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT2G21660)
Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT2G40080)
Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT2G46790)
Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT3G46640)
Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT5G02810)
Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT5G15840)
Gene(GoMapMan/Gene:AT1G01060) affects[reliability=1] Gene(GoMapMan/Gene:AT5G61380)
Gene(GoMapMan/Gene:AT1G01060) interacts_with[reliability=1] Gene(GoMapMan/Gene:AT1G01060)
Gene(GoMapMan/Gene:AT1G01060) interacts_with[reliability=1] Gene(GoMapMan/Gene:AT1G22770)
Gene(GoMapMan/Gene:AT1G01060) interacts_with[reliability=1] Gene(GoMapMan/Gene:AT2G46830)
Gene(GoMapMan/Gene:AT1G01060) interacts_with[reliability=1] Gene(GoMapMan/Gene:AT3G22380)
Gene(GoMapMan/Gene:AT1G01060) interacts_with[reliability=1] Gene(GoMapMan/Gene:AT3G60250)
Gene(GoMapMan/Gene:AT1G01060) interacts_with[reliability=1] Gene(GoMapMan/Gene:AT5G02810)
Gene(GoMapMan/Gene:AT1G01060) interacts_with[reliability=1] Gene(GoMapMan/Gene:AT5G24470)
Gene(GoMapMan/Gene:AT1G01060) interacts_with[reliability=1] Gene(GoMapMan/Gene:AT5G61380)
Gene(GoMapMan/Gene:AT1G01060) interacts_with[reliability=1] Gene(GoMapMan/Gene:AT5G67380)
    
```

