

Rothamsted Repository Download

G - Articles in popular magazines and other technical publications

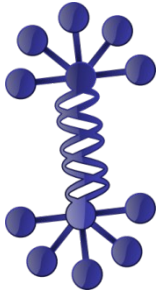
Canevet, C., Lysenko, A., Splendiani, A., Pocock, M. J. O. and Rawlings, C. J. 2010. *Analysis and visualisation of RDF resources in Oindex*. Springer Nature.

The publisher's version can be accessed at:

- <https://dx.doi.org/10.1038/npre.2010.5430>
- <https://arxiv.org/pdf/1012.1661>

The output can be accessed at: <https://repository.rothamsted.ac.uk/item/8v4x8>.

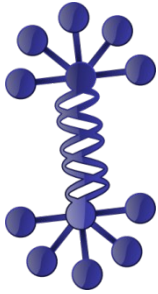
© Rothamsted Research. Licensed under the Creative Commons CC BY.



Analysis and visualisation of RDF resources in Oindex

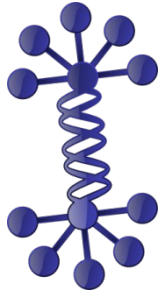
Catherine Canevet

catherine.canevet@bbsrc.ac.uk



Motivation

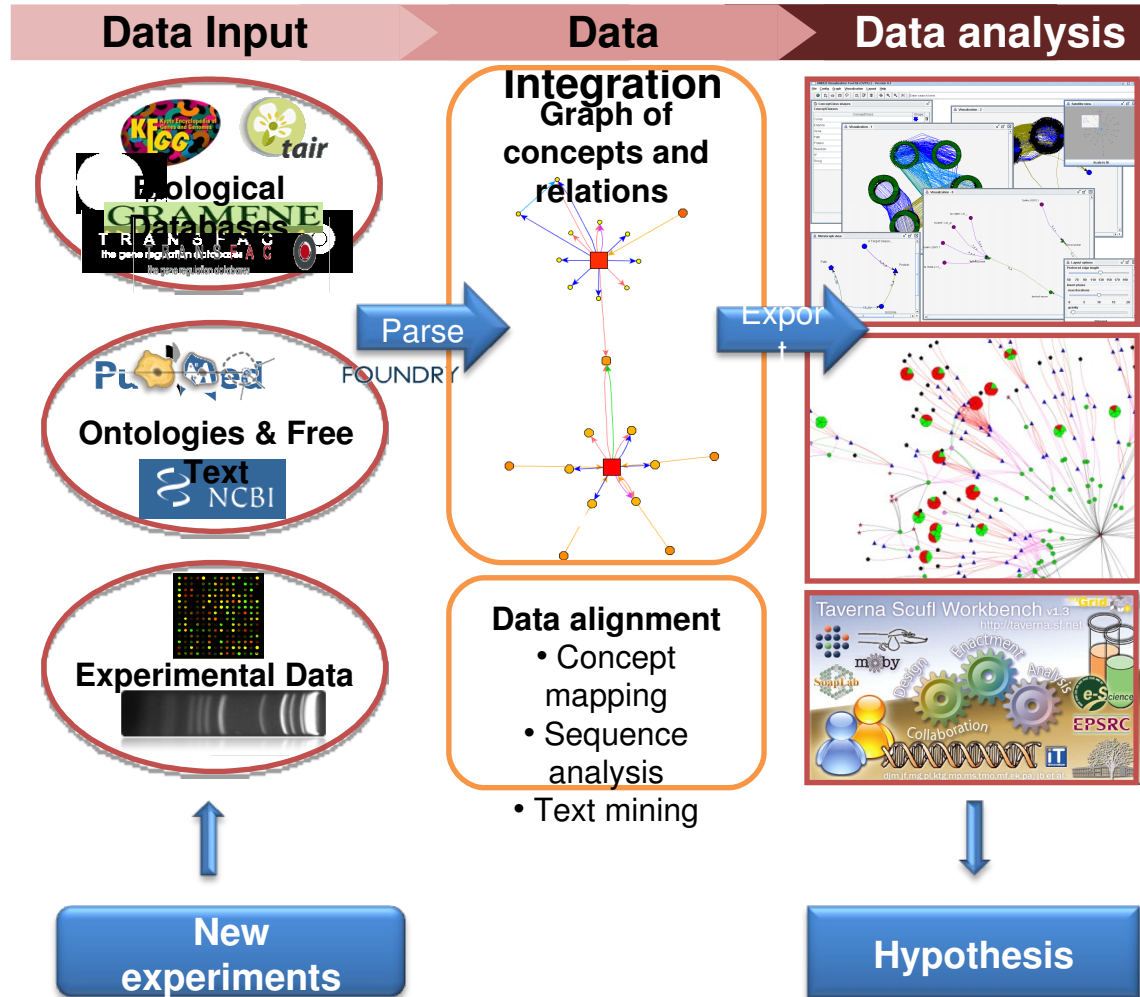
- Biomedical resources provide info on the Semantic Web
- Potential to advance biomedical research
- Need for tools
 - Data reduction
 - Data visualization
 - Interactive analysis capabilities



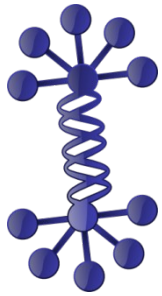
Ondex: a platform to exploit the Semantic Web

- 1) all information can be represented as a graph
 - 2) all elements of the graph can be annotated with ontologies
 - 3) data model conformant to Semantic Web framework, in particular to RDF
- ➔ Ondex ideally positioned

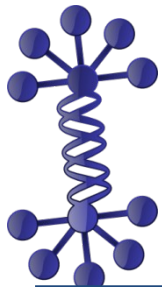
Data integration and visualisation in Ondex



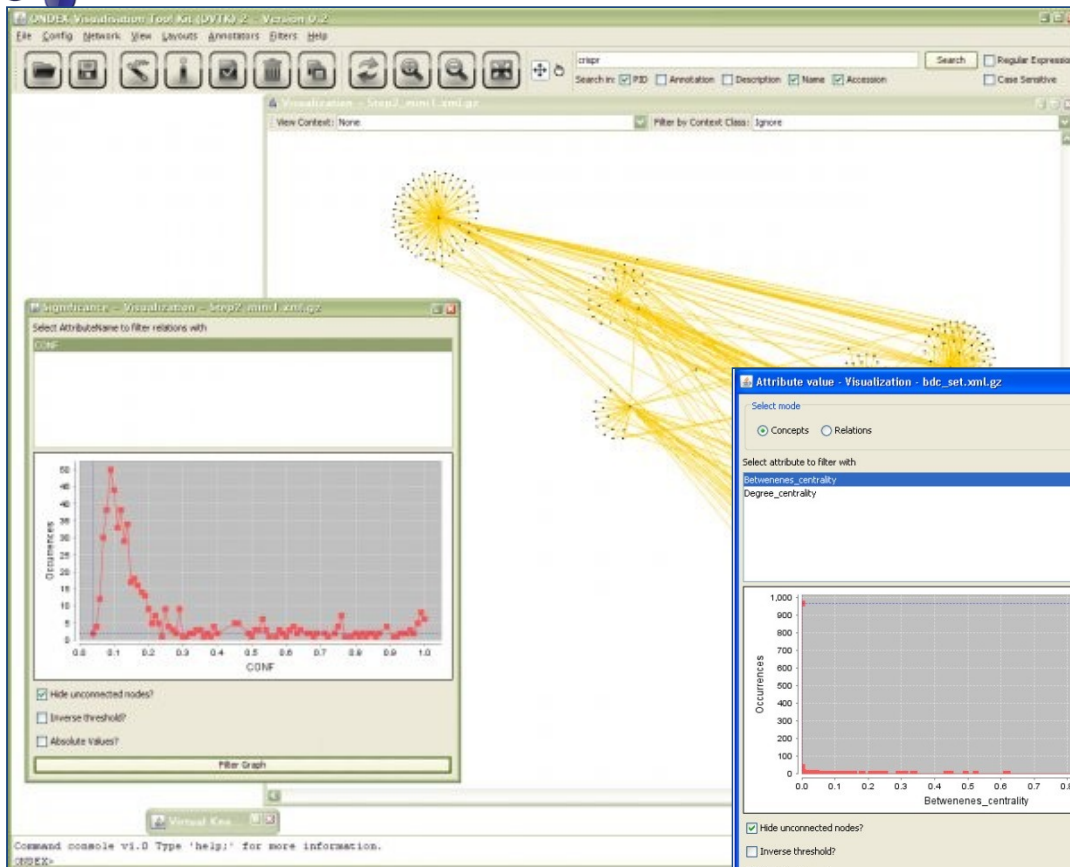
Nature Precedings : doi:10.1038/npre.2010.5430.1 : Posted 23 Dec 2010



Data reduction and representation methods

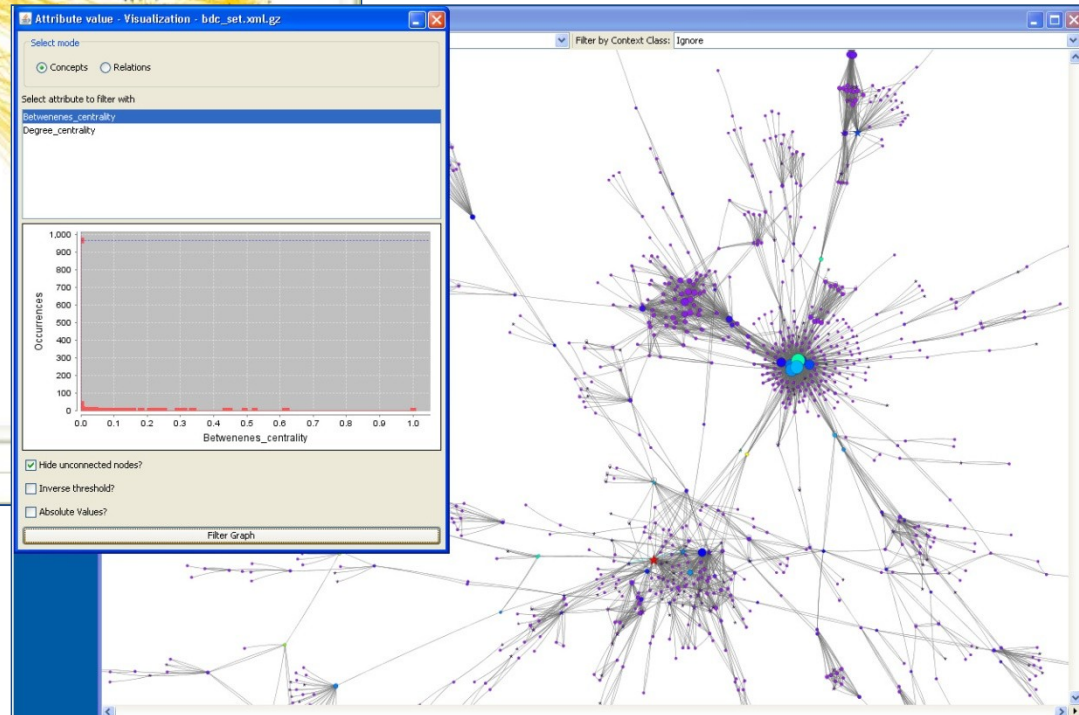


Nature Precedings : doi:10.1038/npre.2010.5430.1 : Posted 23 Dec 2010



Filter on confidence

Filter on centrality



- Export
 - Experimental
 - Discontinued
 - Stable
 - OXL Export
 - Report writer
- Filter
 - Discontinued
 - Stable
 - Experimental
 - Graph
 - Mapping
 - Stable
 - Concept accession-based mapping (Memory-efficient)
 - Inparanoid
 - Experimental
 - Parser
 - Stable
 - Aracyc2 (Release 6, October 14, 2009)
 - BCSDB (pre-release BCSDB-3 gamma version)
 - Brenda specific SBML Parser (BRENDA release online since 2nd July 2010)
 - EXPASY ENZYME (4/10/2009)
 - FASTA file parser
 - KEGG parser (Release 52.1, December 1, 2009)
 - KEGG parser, latest (Release 53.0, January 1, 2010)
 - OXL Parser
 - UniProt (UniProt release July2010)
 - Experimental
 - Statistics
 - Experimental
 - Transformer
 - Experimental
 - Stable
 - Relation Collapser

Arguments | Documentation | Required data | Comments

a subset of the TAIR9 database parsing genes, proteins, domains

Annotation If true also GO annotations, protein domain information and UniProt accessions will be parsed

InputDir Absolute path to input directory

Graph id Graph that will be operated on by this plugin.

New

New memory graph

GraphName default
Graph id default

TAIR 9

Annotation
InputDir
Graph id default

UniProt

InputDir
TaxId
DbRefAcc
Accessions
AccessionsFile
ContextInformation
HideLargeScaleRef
GoFile
Graph id default

Concept accession-based mapping (Memory-efficient)

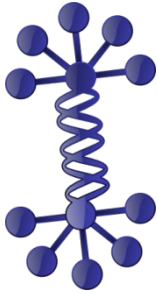
EquivalentCC
EquivalentCV
GDSEqualsConstraint
IgnoreAmbiguity
RelationType equ
WithinCVMMapping
ConceptClassRestriction
CVRrestriction
Graph id default

Relation Collapser

RelationType
ConceptClassRestriction
CVRrestriction
CloneGDS
CreateGDSSets
Graph id default

OXL Export

ExcludeConceptsOfConceptClass
ExcludeRelationsOfRelationType
ExcludeGDSWithAttribute
IncludeOnlyGDSAttribute
IncludeOnlyConceptClass
IncludeOnlyRelationType
Pretty
ExportIsolatedConcepts
GZip
FormatFile

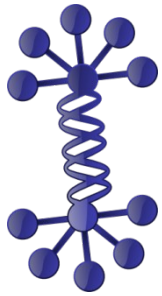


Demo

- Interactive importer based on SPARQL
- Query-driven construction of datasets
- Bring information from different RDF data resources into Oindex

- Welcome feedback on our prototype

Data used in demo integrated by Keywan Hassani-Pak



Input

45778 predicted poplar genes

Data Integration

PoplarCyc, Poplar TFDB

Comparative Genomics

BLAST against UniProt Plants

HMMer against Pfam

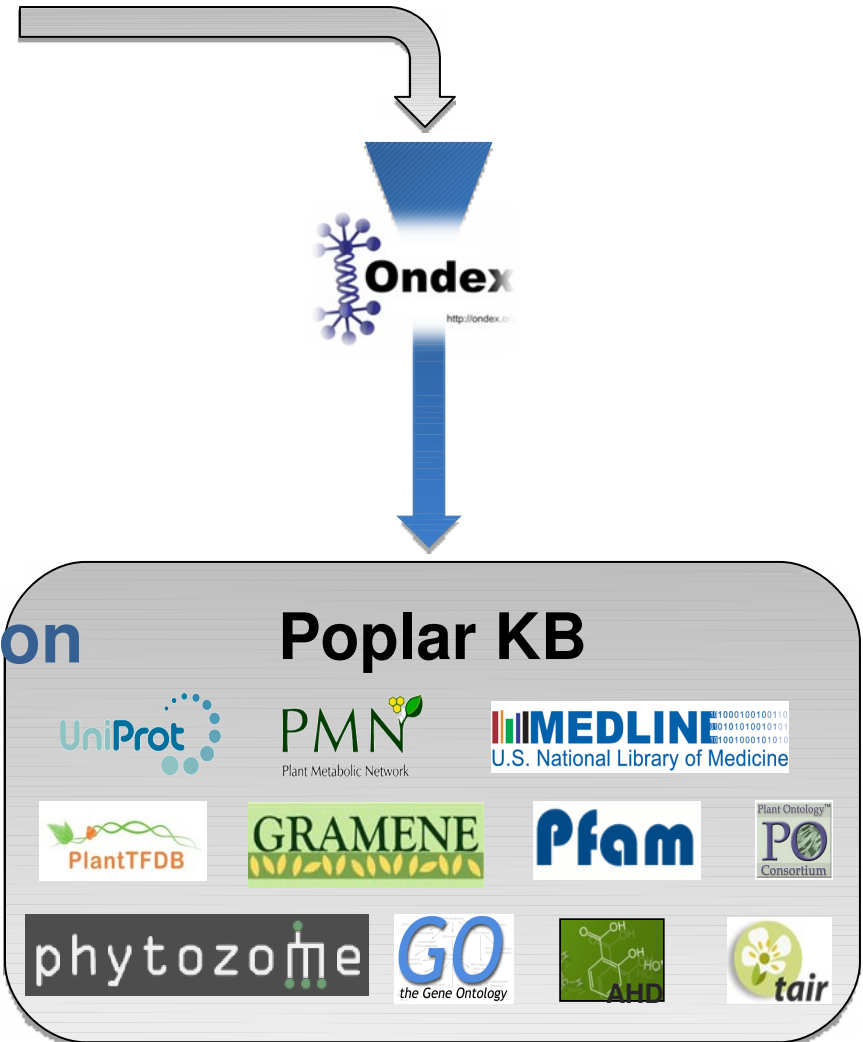
Inparanoid to Arabidopsis

Autom. Function Prediction

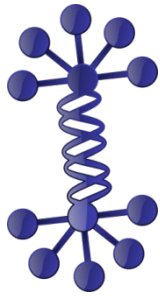
Gene, Plant and Trait Ontology

Text Mining

Co-occurrence of genes and traits in Medline abstracts



Connecting and querying



Nature Precedings : doi:10.1038/npre.2010.5430.1 : Posted 23 Dec 2010

The screenshot shows the Ondex software interface. The main window is titled "Ondex - Version 0.3 (build null)". The menu bar includes File, Edit, View, Appearance, Tools, and Help. The toolbar contains various icons for file operations and search. The search bar contains the text "Enter search here...". Below the search bar, there are dropdown menus for "Restrict by:" (set to "Concept Class"), "Data Source", and "Tag". There are also checkboxes for "Regular Expr." and "Case Sensitive".

The "Item Information" window is open, displaying the following details:

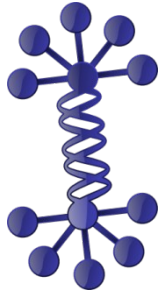
- General Information**
 - Name: <http://www.ondex.org/concept/313>
 - Parser ID: <http://www.ondex.org/concept/313>
 - Type: TO
 - DataSource: <http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql>
 - Evidence: IMPD
- Tags**
- Synonyms**
 - <http://www.ondex.org/concept/313>
 - shoot branching
- Accessions**
- URI**
 - <http://www.ondex.org/concept/313>

The "Visualization - 0" window is empty, showing a blue triangle cursor.

The "Console" window shows the following command history:

```
Command console v1.0 Type 'help;' for more information.
ONDEX>sparql_mode
Changing configuration - please wait...
SPARQL interpreter loaded successfully
ONDEX>connect("http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql")
ONDEX>construct (?x ?y "shoot branching")
where {graph ?g (?x ?y "shoot branching")}
ONDEX>
```

Right-click pre-defined queries



Nature Precedings : doi:10.1038/npre.2010.5430.1 : Posted 23 Dec 2010

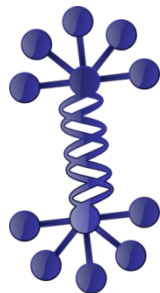
The screenshot displays the Ondex software interface, version 0.3 (build null). The main window is titled "Ondex - Version 0.3 (build null)" and contains several panes:

- Item Information**: A pane on the left showing details for a concept. It includes sections for General Information (Name, Parser ID, Type, DataSource, Evidence), Tags, Synonyms, Accessions, and URI.
- Visualization - 0**: A large central area for visualizing data. A right-click context menu is open over this area, showing options: Link, Hide, Show, Show / Edit Properties, and Query. The Query option is expanded, showing sub-options: Find relations subjects of, Find relations objects of, Find subjects of, Fill attributes, Find objects of, and Resolve URL.
- Console**: A pane at the bottom showing the command history and output of the SPARQL interpreter.

The Console pane contains the following text:

```
Command console v1.0 Type 'help;' for more information.
ONDEX>sparql_mode
Changing configuration - please wait...
SPARQL interpreter loaded successfully
ONDEX>connect ("http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql")
ONDEX>construct {?x ?y "shoot branching"}
where {graph ?g {?x ?y "shoot branching"}}
ONDEX>
```

Filling attributes



Ondex - Version 0.3 (build null)

File Edit View Appearance Tools Help

Enter search here... Regular Expr. Case Sensitive

Restrict by: Concept Class Data Source Tag

Item Information

View

General Information

Name: <http://www.ondex.org/concept/292>
Parser ID: 13130077
Type: Publication
DataSource: <http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql>
Evidence: IMPD

Annotation

[NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, MUTAGENESIS OF ALA-49 AND ARG-50]

Tags

Synonyms

<http://www.ondex.org/concept/292>
PMID:13130077

Accessions

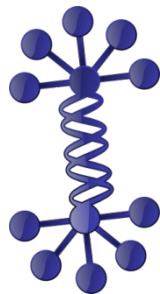
<http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql:10.1073/PNAS.1932414100> 13130077

Publication

LAX and SPA: major regulators of shoot branching in rice.
Keishi Komatsu, Masahiko Maekawa, Shin Ujiie, Yuzuki Satake, Ikuyo Furutani, Hironobu Okamoto, Ko Shimamoto, Junko Kyoizuka
Year: 2003
Published in: Proc. Natl. Acad. Sci. U.S.A.
The aerial architecture of plants is determined primarily by the pattern of shoot branching. Although shoot apical meristem initiation during embryogenesis has been extensively studied by molecular genetic approaches using Arabidopsis, little is known about the genetic mechanisms controlling axillary meristem initiation, mainly because of the insufficient number of mutants that specifically alter it. We identified the LAX PANICLE (LAX) and SMALL PANICLE (SPA) genes as the main regulators of axillary meristem formation in rice. LAX encodes a basic helix-loop-helix transcription factor and is expressed in the boundary between the shoot apical meristem and the region of new meristem formation. This pattern of LAX expression was repeatedly observed in every axillary meristem, consistent with our observation that LAX is involved in the formation of all types of axillary meristems throughout the ontogeny of a rice plant. Ectopic LAX expression in rice caused pleiotropic effects, including dwarfing, an altered pattern of stem elongation, darker color, bending of the lamina joint, absence of the midribs of leaves, and severe sterility.

Visualization - 0

Proteins related to shoot branching



Nature Precedings : doi:10.1038/npre.2010.5430.1 : Posted 23 Dec 2010

Oindex - Version 0.3 (build null)

File Edit View Appearance Tools Help

Enter search here... Regular Expr.

Restrict by: Data Source Tag Case Sensitive

Item Information

View

General Information

Name: <http://www.ondex.org/concept/159>
Parser ID: LAX
Type: Protein
DataSource: <http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql>
Evidence: IMPD

Tags

Synonyms

<http://www.ondex.org/concept/159>
LAX
LAX1
LOC_Os01g61480
Os01g0831000
P0446G04.27
Transcription factor LAX PANICLE

Accessions

<http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql>: AP003252
G7XAG6 PS50888 IPR001092 AB115668 4327431
LOC_Os01g0831000 OSA:4327431 OS:38423 PF00010
NP_001044699.1 Os01g0831000

Taxonomy

[39947](#)

Aminoacid sequence (AA):

MDPFGFP I HPQYHLNPT AGLGGEMR
GGGRFRGAK LSTDPQVAA RERFRISDR
F PULRSLVP GGSMDTVSM LEQALHYWKT
LKAQVTLHQA ALVQNEEGCQ HADVAARSA
AD ADLALQL MRRHGAGDD DAGHTLEMA
PHEQAVGYGD GPAAQHMQA LDFAGQLMGC

Visualization - 0

Metadata Legend

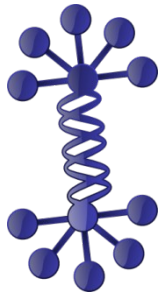
Concept Classes | Data Sources | RelationTypes | EvidenceTypes

		Publication	12 (12)
		TO	7 (7)
		Protein	8 (8)
			Total Concepts 27 (27)

Console

```
Command console v1.0 Type 'help;' for more information.
ONDEX>sparql_mode
Changing configuration - please wait...
SPARQL interpreter loaded successfully
ONDEX>connect("http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql")
ONDEX>construct {?x ?y "shoot branching"}
where {graph ?g {?x ?y "shoot branching"}}
ONDEX>
```

Colour proteins by taxonomy ID



Nature Precedings : doi:10.1038/npre.2010.5430.1 : Posted 23 Dec 2010

Ondex - Version 0.3 (build null)

File Edit View Appearance Tools Help

Enter search here... Regular Expr.

Restrict by: Concept Class Data Source Tag Case Sensitive

Colour Concepts by Gene...

Select attribute to annotate concepts with

- DOI
- JOURNAL_REF
- JOURNAL_REF:0
- MeSH
- PUB_TYPE
- TAXID**
- URI
- YEAR

Set blanks to white?

Exclusive use of attribute?

Annotate Graph

Legend Colour Con...

Color
3694
39947

LOC_Os01g01400
Os01g0831000
P0446G04.27
Transcription factor LAX PANICLE

Accessions
<http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql>: AP003252
G7XAQ6 PS50888 IPR001092 AB115668 4327431
LOC_Os01g0831000 OSA:4327431 OS:38423 PF00010
NP_001044699.1 Os01g0831000

Taxonomy
[39947](#)

Aminoacid sequence (AA):

```

MNDPRGTF I HPQYHLNPT AGGLGGRMR
GGGRFRPGAK LSTDPQSVAA RERFRISDR
F PULRSLVP GGSMDTVSM LEQILHYVKT
LKAQVTLHQA ALVQNEEGCQ HADVAATPSA
AD ADLALEL MRRHGAGDD DAGHTLEMA
PMQEAUGYGD GPARQHQQA LDFAGQLMGC
    
```

Metadata Legend

Concept Classes | Data Sources | RelationTypes | EvidenceTypes

		Publication	12 (12)
		TO	7 (7)
		Protein	8 (8)
			Total Concepts 27 (27)

Copy to clipboard

Refresh Metadata

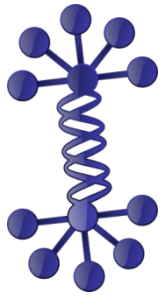
Visualization - 0

Console

```

Command console v1.0 Type 'help;' for more information.
ONDEX>sparql_mode
Changing configuration - please wait...
SPARQL interpreter loaded successfully
ONDEX>connect("http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql")
ONDEX>construct {?x ?y "shoot branching"}
where {graph ?g {?x ?y "shoot branching"}}
ONDEX>
    
```

Scale relation based on BLAST



Nature Precedings : doi:10.1038/npre.2010.5430.1 : Posted 23 Dec 2010

Scale/Colour Relations by N...

resize relations

Min Relation Size: 1
Max Relation Size: 10

inverse scaling

colour edges

Min Colour: [Blue]
Max Colour: [Red]

No attribute colour: [Black]
 Colour on rainbow scale
 Hide non-matching
 Hide unconnected nodes

Select an attribute to annotate relations with

BITSCORE
BLEV
COVERAGE
OVERLAP
QUERY_LENGTH
TARGET_LENGTH
TRANSLATION_FRAME

Annotate Graph

Legend Colour Con...

Color	Color
3694	[Purple]
39947	[Green]

Visualization - 0

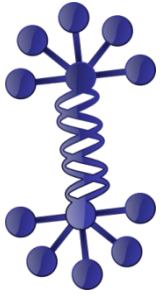
Command console v1.0 Type 'help;' for more information.

```
ONDEX>sparql_mode
Changing configuration - please wait...
SPARQL interpreter loaded successfully
ONDEX>connect ("http://babvs03.rothamsted.bbsrc.ac.uk:8050/sparql")
ONDEX>construct { ?x ?y "shoot branching"
where { graph ?g { ?x ?y "shoot branching" } }
ONDEX>
ONDEX>
ONDEX>
```

Metadata Legend

Concept Classes	Data Sources	Relation Types	Evidence Types
[Orange]	[Blue]	Publication	12 (12)
[Blue]	[Blue]	TO	7 (7)
[Red]	[Blue]	Protein	8 (8)

Total Concepts 27 (27)
Copy to clipboard
Refresh Metadata



Acknowledgements

- Artem Lysenko
- Andrea Splendiani
- Matthew Pocock
- Chris Rawlings