

***Deliverable 5.3, December 2014***

# **Database of core information for the community**

## ***Work package 5 Community Building and Synergies***

*ISBE WP5 Coordinators:*  
James Sharpe and Jutta Steinkoetter

*Main responsible authors:*  
Joaquim Calbó, James Sharpe  
Centre for Genomic Regulation (CRG)

*Contact:*  
[joaquim.calbo@crg.eu](mailto:joaquim.calbo@crg.eu)

<b>Project ref. no.</b>	INFRA-2012-2.2.4: 312455
<b>Project title</b>	ISBE – Infrastructure for Systems Biology Europe
<b>Nature of Deliverable</b>	R= Report
<b>Contractual date of delivery</b>	Month 24
<b>Actual date of delivery</b>	Month 28
<b>Deliverable number</b>	D5.3
<b>Deliverable title</b>	Database of core information for the community
<b>Dissemination Level</b>	PU
<b>Status &amp; version</b>	Version 4
<b>Number of pages</b>	16
<b>WP relevant to deliverable</b>	WP5
<b>Lead Participant</b>	CRG, James Sharpe
<b>Author(s)</b>	Joaquim Calbó, James Sharpe
<b>Project coordinator</b>	Richard Kitney
<b>EC Project Officer</b>	Andrea de Candido

Dissemination level: PU = Public, RE = Restricted to a group specified by the Consortium (including Commission services), PP = Restricted to other programme participants (including Commission Services), CO= Confidential, only for members of the Consortium (including the Commission Services)

Nature of Deliverable: P= Prototype, R= Report, D=Demonstrator, O = Other.

## Table of Contents

---

<b>Summary:</b> .....	<b>4</b>
<b>Methods:</b> .....	<b>5</b>
Web searches .....	5
Community website .....	5
MERIL database and technology watch.....	6
Audit of Systems Biology provision .....	6
Data integration .....	7
<b>Community Website Database</b> .....	<b>8</b>
Community website – underlying database .....	13
<b>Conclusions and outlook</b> .....	<b>15</b>
Adding useful information to the community website .....	15
Improving the search and display functionalities on the website .....	16

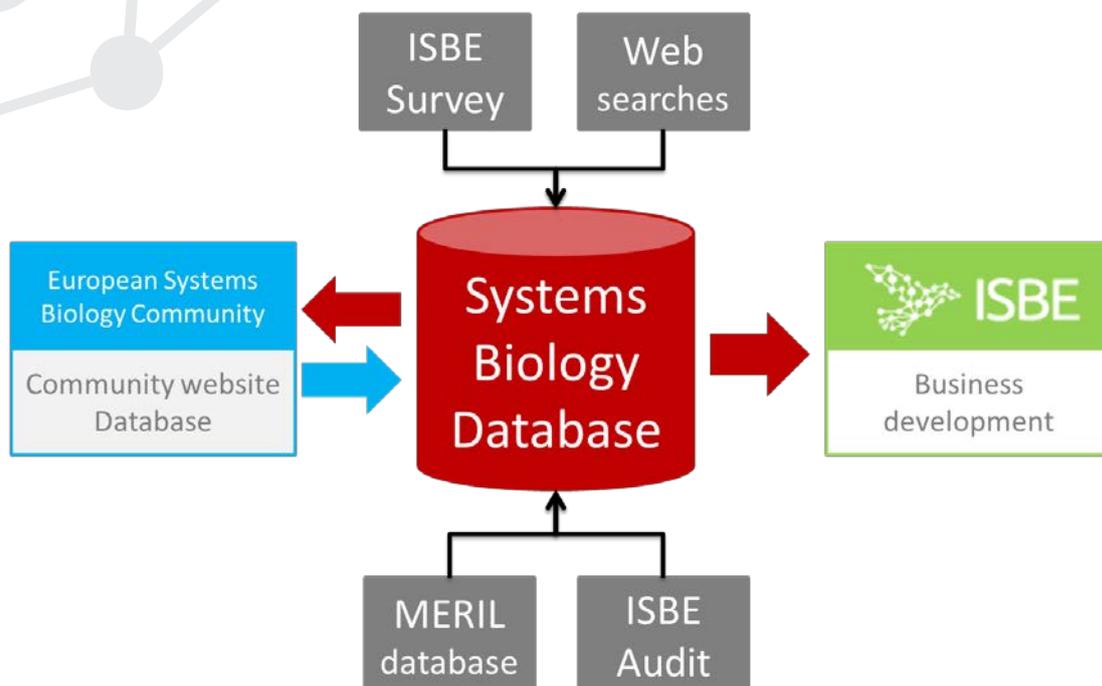
## Summary:

---

In the first 24 months of the preparatory phase, ISBE has made a substantial effort to gather information of relevance about and for the Systems Biology community (that is, the research community currently involved or interested in systems biology approaches, both from academia and from industry). The objectives of this activity are twofold:

- To be in a position to offer relevant information to the community
- To understand the systems biology landscape in Europe, the current service provision and user needs in order to make informed decisions in the design of ISBE business model and business plan

The main deliverable related to this effort is the database underlying the European Systems Biology Community website (<http://community.isbe.eu>), since this resource makes the data readily accessible to the community. However, it is important to emphasize that in its preparatory phase, ISBE has collected data from several sources (including the community website) which are integrated in the ISBE Systems Biology database. This (internal) database provides information to the ISBE consortium in order to drive the design of an infrastructure with and for the community. After internal use and verification, it is intended that much of this shall be made publically available through the portal.



## Methods:

---

In order to populate the ISBE Systems Biology database and its public interface (the European Systems Biology Community website), ISBE WP5 members in collaboration with other ISBE partners have obtained information from different sources and using different strategies, integrating this in a central database as a collection of tables with linking fields. The main data types used to connect the different tables, through unique ID fields, are *Institution name* and *Researcher name*, since all information in the database can be linked to one of them. In addition, a relevant proportion of Researchers can be linked to Institutions through their affiliation information.

The following data sources have been exploited:

### Web searches

As described in Deliverable 5.1, an extensive web search for researchers (group leaders) working on systems biology was performed. Approximately 600 researchers were identified, affiliated to 200 institutions. Information on research interest and background was manually collected, and urls linking to group or institutional sites were recorded. The list of researchers and institutions was the seed for the ISBE Systems Biology Database.

### Community website

The European Systems Biology Community website was constructed to provide the community with a platform to help European systems biologists network and access relevant information about the community (Deliverable 5.2). Seed data for the underlying database was obtained through an unbiased literature search as described in Deliverable 5.1, resulting in a list of >7,000 researchers and >600 institutions. Automatically generated profiles can be claimed through the community website, and all information can be updated and improved by community members. Information linked to researchers includes keywords, affiliation, url, etc. Information linked to institutions includes address, affiliated researchers, url, etc.

The community website currently has 150 registrants and seeks to increase this number quite radically thanks to planned upgrades (outlined below) and to ISBE dissemination activities.

### MERIL database and technology watch

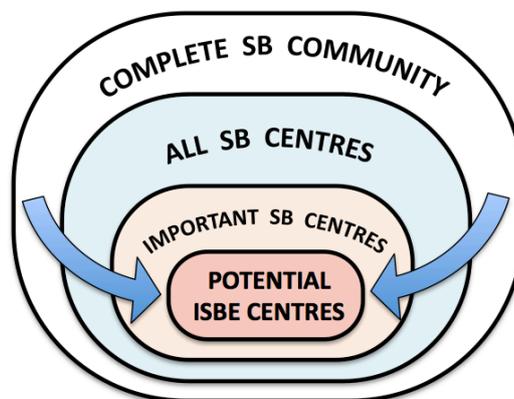
MERIL (Mapping of the European Research Infrastructure Landscape, <http://portal.meril.eu/>) provides a public database with information on the different types of research infrastructures of more-than-national relevance across all scientific domains. Searches have been performed in this database to retrieve a list of infrastructures of relevance to ISBE (including Systems Biology services, bioinformatics resources, and big data-generation centres in the fields of –omics, structural biology, and bioimaging). About 200 infrastructures have been identified, and about half of them mapped to the institutions listed in the ISBE Systems Biology database.

In addition, WP4 generated a list of centres providing systems-biology compliant data-generation services. Most of these centres have also been linked to research institutions in our database.

### Audit of Systems Biology provision

In order to delimit the current and future provision of systems biology services and resources in Europe, ISBE is currently running an Audit, which takes the form of direct consultation of several systems biology centres across Europe. Previous to this consultation, an extensive exploration has been done to find all institutions performing systems biology research in Europe (based on data from the ISBE systems biology database), and to identify at the national level the main systems biology actors (by curation of listed information by local experts in each of the ISBE participating countries). The process produced a country-by-country list of systems biology centres and recognized researchers that are being contacted in the consultation phase.

**SB Audit**



**Figure 1:** This scheme shows the focus of the ISBE Audit in the context of the Systems Biology community.

The Audit addresses the current and future SB service provision in Europe from two sides: qualitatively, by discussing with the experts their interest and wants in providing services to / using from a pan-European infrastructure; and quantitatively, by recording data about individual resources offered by the SB centres, its actual use, funding, and sustainability. Quantitative data will be integrated into the ISBE database and where possible made available to the community via the website. In addition, all information will be summarized in a report and used to shape the ISBE Business Plan according to the needs of SB service providers and users.

Of note, the ISBE Audit on systems biology provision is a fact-finding mission, and is not related to the selection process of the centres that will be part of the operational ISBE. An in-depth and formal process of selecting ISBE centres will be pursued over the next 12 months.

### Data integration

The many data-gathering activities listed above have resulted in heterogeneous datasets including different types of information and focusing on different organizational levels. However, all data tables refer to either individuals (researchers) or institutions. These two entities are normally the “primary key” for the tables in the ISBE systems biology database. For this reason, we have created unique ID for researchers and for institutions, and linked all information to these identifiers. As a result of the web-search, the unbiased literature search, and the manual curation of data through the community website, the ISBE database contains affiliation information for about 40% of the listed researchers, linking the researchers' unique identifiers with those of the institutions.

The integration of the different datasets generates a complete picture of the systems biology landscape in Europe, where information of who is doing what in systems biology, where, when, and how can be retrieved. Importantly, most of the information is available to the community through the website, with the primary aim to facilitate the identification of systems biology expertise and resources, and to foster exchange and collaboration between researchers and institutions.

## Community Website Database

---

The main outcome of the work summarized here is the European Systems Biology community database and its public interface, the community website (CW).

CW offers easy access to information about researchers and institutions, and different options for data searching and visualization. Moreover, the CW has been created as a community tool, and consequently all information can be upgraded, improved and augmented by its members. Thus, data quantity and quality keeps growing as community members interact with the CW, in a positive feedback loop.

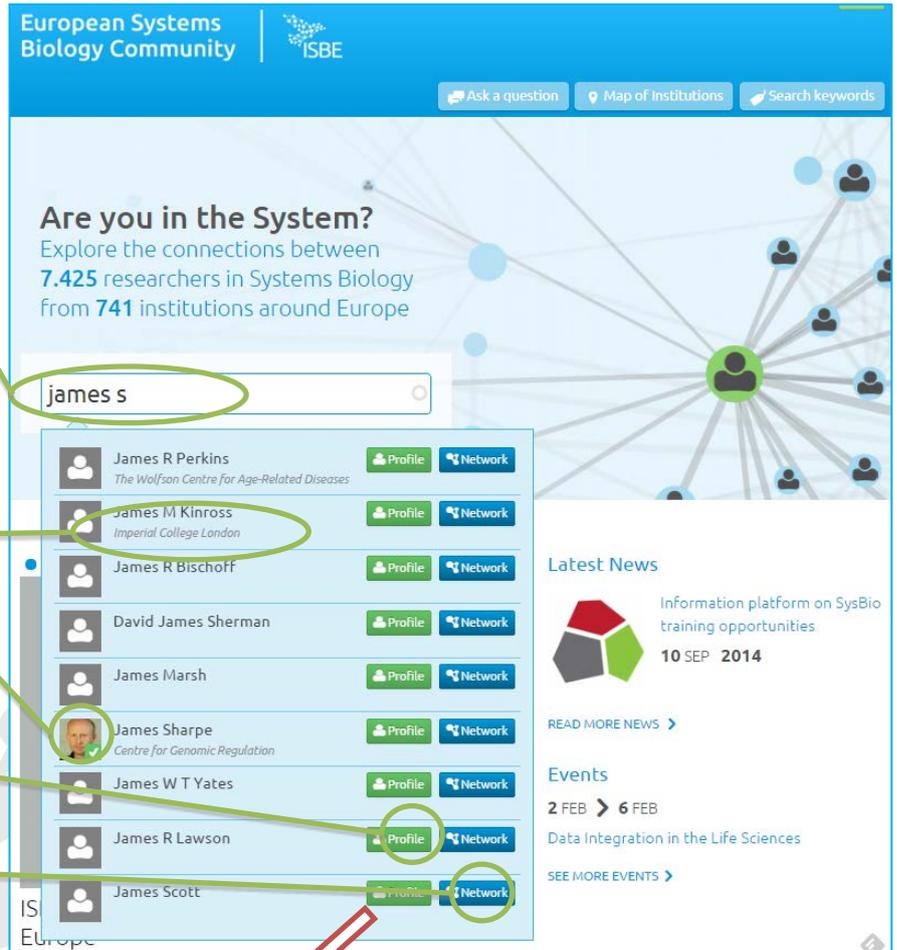
Data currently accessible through the CW are:

- Researcher profiles (see Fig 2)
  - o Keywords (research interests and expertise)
  - o Affiliation (current and past host institutions)
  - o SB-related publications (with links to PubMed records)
  - o Link to personal websites (URL)
  - o Picture
- Institution details (see Fig 3)
  - o Organization/department/group information
  - o Address
  - o Link to institutional website (URL)
  - o Links to parental / filial institutions
  - o Links to affiliated researchers

The main tools to search/retrieve information from the CW are:

- Search by name: researchers can be found by typing in their name (Fig 2)
- Search in the map: the map of institutions and its associated country filters facilitate the search of SB centres by exploring the map (Fig 3)
- Country lists of institutions: list of institutions in each European country, by alphabetical order, with links to institution page (Fig 3)
- Search by keyword: entering a keyword returns the list of researchers with that particular keyword linked to their profile. The list is ranked (community members with keyword in primary keyword list, first; community members with keyword in secondary keyword list, second; non-community members, last) (Fig 4)
- Network visualization tool: nodes in the network represent interconnected researchers, and link to their profile pages (Fig 5)

## Search name (home page)



European Systems Biology Community | ISBE

Ask a question | Map of Institutions | Search keywords

**Are you in the System?**  
Explore the connections between **7.425** researchers in Systems Biology from **741** institutions around Europe

Search term:

	James R Perkins <i>The Wolfson Centre for Age-Related Diseases</i>	Profile	Network
	James M Kinross <i>Imperial College London</i>	Profile	Network
	James R Bischoff	Profile	Network
	David James Sherman	Profile	Network
	James Marsh	Profile	Network
	James Sharpe <i>Centre for Genomic Regulation</i>	Profile	Network
	James W T Yates	Profile	Network
	James R Lawson	Profile	Network
	James Scott	Profile	Network

Latest News  
Information platform on SysBio training opportunities  
10 SEP 2014  
READ MORE NEWS >

Events  
2 FEB > 6 FEB  
Data Integration in the Life Sciences  
SEE MORE EVENTS >

Search term

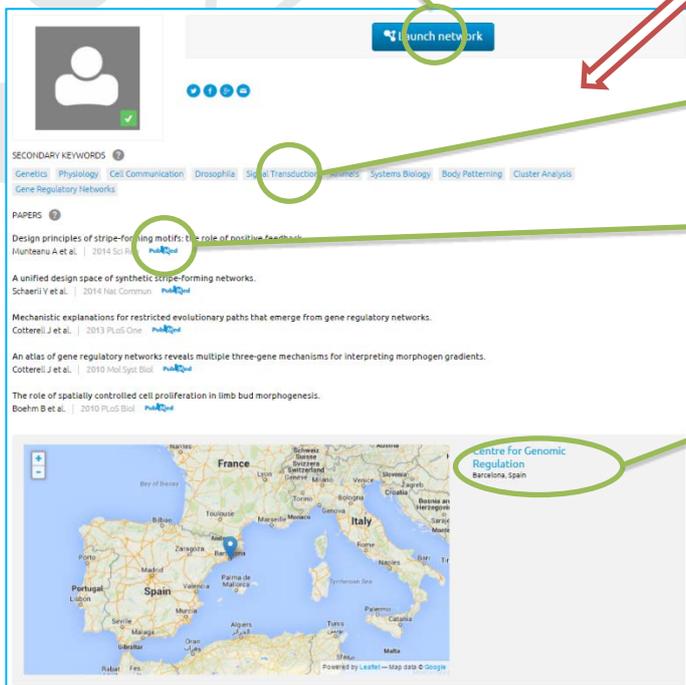
Affiliation

Registered member

Link to profile

Link to network

## Profile page



Launch network

SECONDARY KEYWORDS  
Genetics | Physiology | Cell Communication | Drosophila | Signal Transduction | Genomes | Systems Biology | Body Patterning | Cluster Analysis  
Gene Regulatory Networks

PAPERS  
Design principles of stripe-forming motifs: the role of positive feedback  
Muntesanu A et al. | 2014 Sci Rep  
A unified design space of synthetic stripe-forming networks.  
Schaerli Y et al. | 2014 Nat Commun  
Mechanistic explanations for restricted evolutionary paths that emerge from gene regulatory networks.  
Cottarelli J et al. | 2013 PLoS One  
An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients.  
Cottarelli J et al. | 2010 Mol Syst Biol  
The role of spatially controlled cell proliferation in limb bud morphogenesis.  
Boehm B et al. | 2010 PLoS Biol

Centre for Genomic Regulation  
barcelona, Spain

**Keywords:** Link to researchers list

**Papers:**  
Link to PubMed

**Affiliation:**  
Link to institution page

**Figure 2:** Search by name in the Community website database, and profile view.

## Map of Institutions

**Countries**

Click on one of the country-names below to see the list of institutions from that country (the number between brackets indicates the number of institutions listed).

- Austria (17)
- Belgium (21)
- Bulgaria (1)
- Croatia (5)
- Czech Republic (10)
- Denmark (17)
- Finland (10)
- France (53)
- Germany (179)
- Greece (13)
- Hungary (5)
- Iceland (1)
- Ireland (11)
- Italy (40)
- Luxembourg (1)
- Netherlands (45)
- Norway (6)

Filter by country: Ireland

## List of Institutions per country

**Institutions Ireland**

- Centre For Scientific Computing and Complex Systems Modelling
- Conway Institute
- Cork Institute of Technology
- Hamilton Institute
- National Institute for Bioprocessing Research and Training (NIBRT)
- National Institute for Cellular Biotechnology
- National University of Ireland
- Royal College of Surgeons in Ireland
- Systems Biology Ireland (SBI)
- Teagasc Food Research Centre
- University College Dublin (UCD)

[Back to country overview](#)

Link to profile

## Institution page

**University College Dublin**

ACRONYM: UCD

ADDRESS: Belfield  
Dublin 4  
Ireland

LOCATION

DEPARTMENTS

Systems Biology Ireland

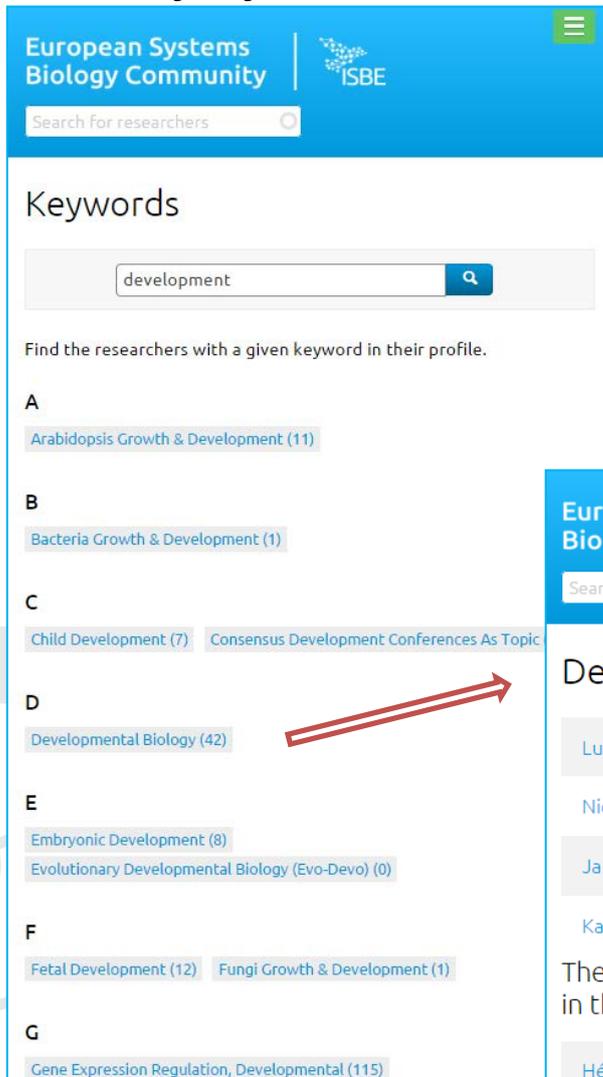
RESEARCHERS

- Lan K Nguyen
- Eamonn Gormley
- Walter Kolch
- Gerard Cagney
- Maico van Attilkum
- Boris N Kholodenko
- William Fitzmaurice
- Miguel A. S. Cavadas
- David Matallanas

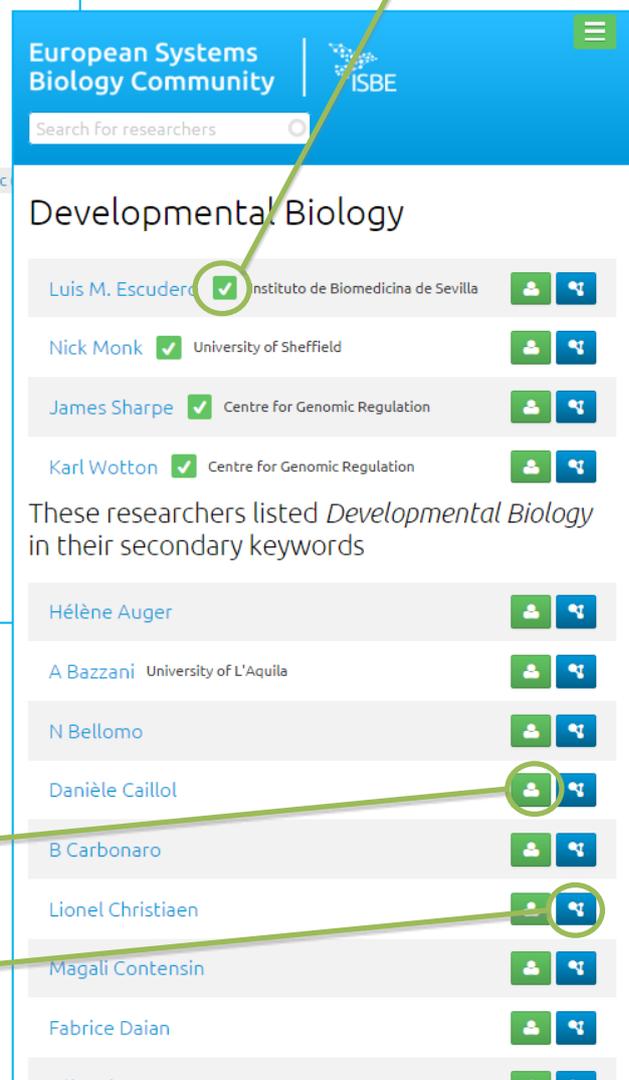
Link to department

**Figure 3:** Search institutions via map or list.

## Search by Keyword



## Researchers



Registered member

Developmental Biology

Luis M. Escudero  Instituto de Biomedicina de Sevilla  

Nick Monk  University of Sheffield  

James Sharpe  Centre for Genomic Regulation  

Karl Wotton  Centre for Genomic Regulation  

These researchers listed *Developmental Biology* in their secondary keywords

Hélène Auger  

A Bazzani University of L'Aquila  

N Bellomo  

Danièle Caillol  

B Carbonaro  

Lionel Christiaen  

Magali Contensin  

Fabrice Daian  

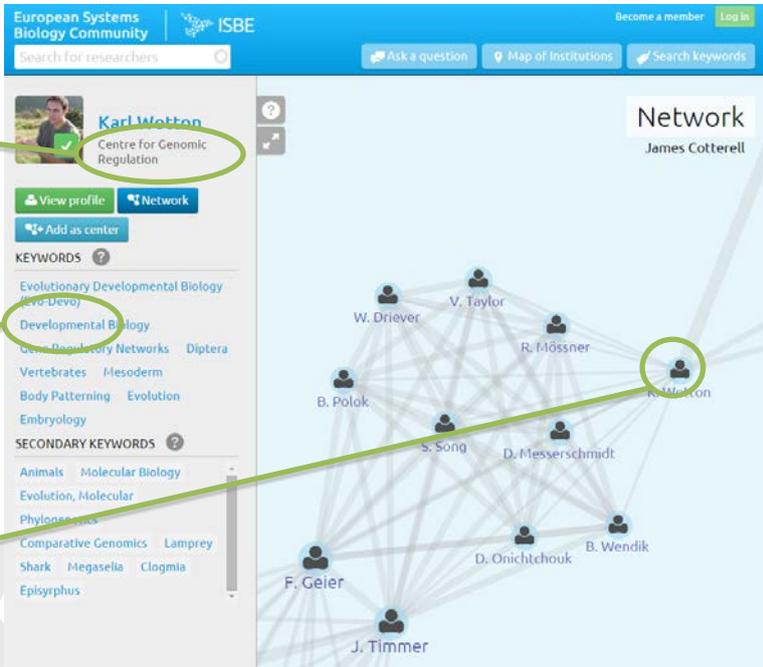
Stéphanie Drapeau  

Link to profile

Link to network

**Figure 4:** Search by keyword and list of search results.

## Network visualization



The screenshot shows the ISBE network visualization interface. On the left, a profile for Karl Wotton is displayed, including his affiliation (Centre for Genomic Regulation) and a list of keywords. On the right, a network graph shows connections between various researchers, with Karl Wotton highlighted as a central node. Annotations with green lines point to these specific elements.

**Affiliation** (points to Karl Wotton's affiliation: Centre for Genomic Regulation)

**Keywords: Link to researchers list** (points to the 'Developmental Biology' keyword)

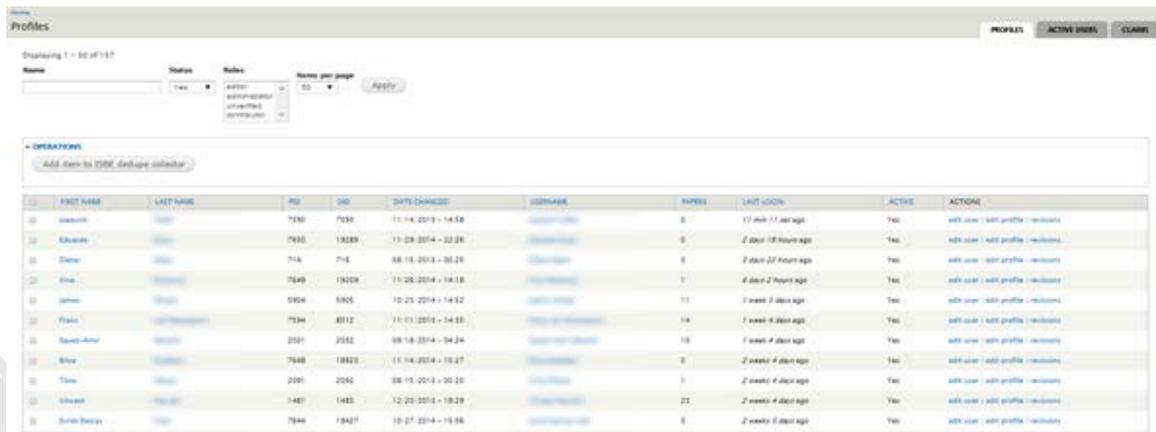
**Nodes link to profile preview** (points to the Karl Wotton node in the network graph)

**Figure 5:** Searching connection in network visualization.

## Community website – underlying database

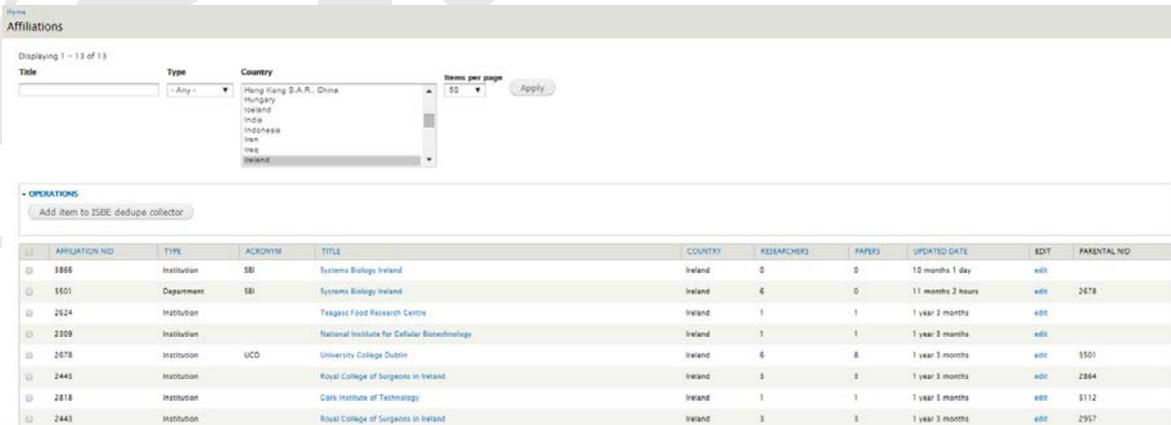
The underlying database, managed with Drupal (open source content management platform, [www.drupal.org](http://www.drupal.org)), is organized in the following categories (see screen captures), which are interlinked as described previously. Listed information can be downloaded by editors and administrators as xls files.

### Profiles / users



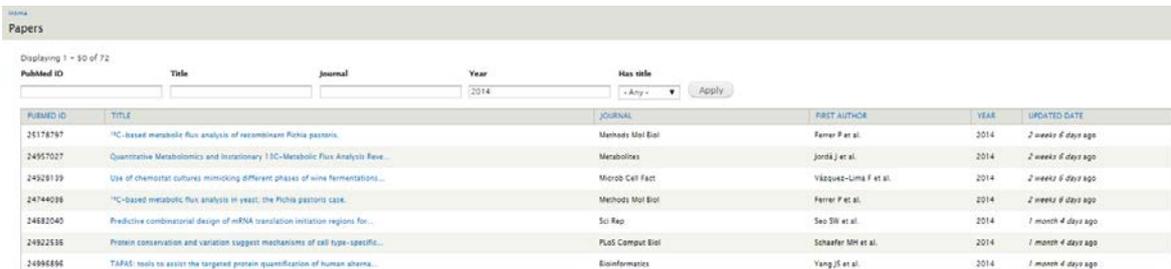
SI	FIRST NAME	LAST NAME	PID	DID	DATE CHANGED	USERNAME	PAPERS	LAST LOGIN	ACTIVE	ACTIONS
11	Steinhilber	Stefan	7930	7930	11-14-2013 - 14:58	steinhilber	0	17 min 11 sec ago	Yes	edit user   edit profile   reassign
12	Edwards	Stefan	7930	18289	11-28-2014 - 22:26	stefan.edwards	0	2 days 18 hours ago	Yes	edit user   edit profile   reassign
13	Zeman	Stefan	714	714	08-18-2013 - 20:26	stefan.zeman	0	2 days 22 hours ago	Yes	edit user   edit profile   reassign
14	Irak	Stefan	7849	13029	11-28-2014 - 14:18	stefan.irak	1	2 days 2 hours ago	Yes	edit user   edit profile   reassign
15	Jones	Stefan	8904	8905	10-23-2014 - 14:12	stefan.jones	11	1 week 3 days ago	Yes	edit user   edit profile   reassign
16	Plani	Stefan	7934	8012	11-11-2013 - 14:33	stefan.plani	14	1 week 4 days ago	Yes	edit user   edit profile   reassign
17	David-Abner	Stefan	2091	2092	09-18-2014 - 24:24	stefan.david-abner	10	1 week 1 day ago	Yes	edit user   edit profile   reassign
18	Repp	Stefan	7848	18820	11-14-2014 - 19:27	stefan.repp	0	2 weeks 4 days ago	Yes	edit user   edit profile   reassign
19	Tsai	Stefan	2091	2092	08-18-2013 - 20:26	stefan.tsai	1	2 weeks 4 days ago	Yes	edit user   edit profile   reassign
20	Wissner	Stefan	1481	1482	12-20-2012 - 18:29	stefan.wissner	22	2 weeks 4 days ago	Yes	edit user   edit profile   reassign
21	Burns-Berry	Stefan	7944	18427	10-27-2014 - 18:56	stefan.burns-berry	0	2 weeks 5 days ago	Yes	edit user   edit profile   reassign

### Institutions



SI	AFFILIATION ID	TYPE	ACRONYM	TITLE	COUNTRY	RESEARCHERS	PAPERS	UPDATED DATE	EDIT	PARENTAL ID
11	3866	Institution	SB	Systems Biology Ireland	Ireland	0	0	10 months 1 day	edit	
12	3501	Department	SB	Systems Biology Ireland	Ireland	6	0	11 months 2 hours	edit	2478
13	2624	Institution		Teagasc Food Research Centre	Ireland	1	1	1 year 2 months	edit	
14	2309	Institution		National Institute for Cellular Biotechnology	Ireland	1	1	1 year 3 months	edit	
15	2678	Institution	UCD	University College Dublin	Ireland	6	8	1 year 3 months	edit	3501
16	2443	Institution		Royal College of Surgeons in Ireland	Ireland	3	3	1 year 3 months	edit	2864
17	2818	Institution		Cork Institute of Technology	Ireland	1	1	1 year 3 months	edit	3112
18	2443	Institution		Royal College of Surgeons in Ireland	Ireland	3	3	1 year 3 months	edit	2957

### Papers



PUBMED ID	TITLE	JOURNAL	FIRST AUTHOR	YEAR	UPDATED DATE
25174767	<sup>13</sup> C-based metabolic flux analysis of recombinant <i>Pichia pastoris</i> .	Methods Mol Biol	Ferrer P et al.	2014	2 weeks 6 days ago
24957027	Quantitative Metabolomics and Isotopically <sup>13</sup> C-Metabolic Flux Analysis Reveals...	Metabolites	Jordá J et al.	2014	2 weeks 6 days ago
24928139	Use of chemostat cultures mimicking different phases of wine fermentations...	Microb Cell Fact	Vázquez-Lima F et al.	2014	2 weeks 6 days ago
24744098	<sup>13</sup> C-based metabolic flux analysis in yeast: the <i>Pichia pastoris</i> case.	Methods Mol Biol	Ferrer P et al.	2014	2 weeks 6 days ago
24682040	Predictive combinatorial design of mRNA translation initiation regions for...	Sci Rep	Seo SH et al.	2014	1 month 4 days ago
24922536	Protein conservation and variation suggest mechanisms of cell type-specific...	PLoS Comput Biol	Schaefer MH et al.	2014	1 month 4 days ago
24996896	TAPAS: tools to assist the targeted protein quantification of human stems...	Bioinformatics	Yang JS et al.	2014	1 month 4 days ago

## Taxonomies (keywords)

Home

### Taxonomies

Displaying 1 - 19 of 19

**Keyword**  
development  items per page: 50

**OPERATIONS**  
- Choose an operation -

<input type="checkbox"/>	NAME	COUNT
<input type="checkbox"/>	growth & development	216
<input type="checkbox"/>	Gene Expression Regulation, Developmental	116
<input type="checkbox"/>	Developmental Biology	46
<input type="checkbox"/>	Plant Development	19
<input type="checkbox"/>	Fetal Development	12
<input type="checkbox"/>	arabidopsis growth & development	11



## Conclusions and outlook

---

Several of the ISBE WP5 community building activities are related with gathering and providing information from and to the Systems Biology community. The produced datasets are focused on individuals and institutions in Europe that are actively involved in Systems Biology research. This document summarizes the different data-gathering activities, the integration of the different datasets into a restricted-access ISBE Systems Biology Database, and the publication of relevant content of the database through an open interface, the European Systems Biology Community website. Thus, the accumulated and curated information has a dual use: on the one hand it provides relevant information for the community that can facilitate resource finding, knowledge sharing, and collaboration; on the other, it provides the ISBE consortium with in-depth information on the European systems biology landscape essential for the design of ISBE infrastructure (which will be drafted in the ISBE Business Plan document).

The future steps for the expansion and maintenance of the database aim to increase community contribution and involve a series of upgrades on the community website that have as main objectives:

1. Add useful information to the community website, in a linked and searchable way
2. Improve the search and display functionalities on the website enlargement

These upgrades will be initiated in January 2015 and launched in June 2015, and they should contribute to make the CW a useful tool provided to the community by ISBE during its preparatory phase, but functioning and expanding through the subsequent building and operation phases.

### **Adding useful information to the community website**

Three priority areas for additional information were identified at the ISBE-WP5 meeting held in Dublin, February 2014 (ranked in descending order of importance):

1. SB (training) events
2. SB tools-resources
3. Collaborative projects / programmes

Information should be community-related:

- Relevant to the community
- Linked to people, institutions, SB terms
- Can be contributed by the community

Taking this into account, we plan an initial population of the SB community database with data provided by the involved WPs (WP10 for training and WP2 for tools), or by establishing collaboration with other actors (mainly ERASysAPP, EMTRAIN, CASyM...) but then allowing contribution/editing by community members.

However, the CW and the underlying database need a technical upgrade in order to accommodate additional data types and new relations between them (i.e. “events” as new data type and “person x organized event y” as relation).

In addition to the information described above as new categories for the CW and database, other SB-relevant information will be provided through the CW but not included or interlinked with the database. In particular, the CW will include display pages for:

- iAnn events list/calendar (<http://iann.pro/> )
- ERASysAPP training platform (<https://www.erasysapp.eu/training-and-exchange/graduate-study-programs> )

iAnn will provide a comprehensive list of events, not linked to the web database.

ERASysAPP has developed some training-related tools that can be integrated in our website via IFRAME.

### **Improving the search and display functionalities on the website**

One major upgrade for the CW is the implementation of a more powerful search engine which allows simultaneous search of all data types (profiles, institutions, events...) and across multiple fields (name, description, address, keywords...).

Search results will also be displayed in more easily interpretable formats (table, map, network visualization), some of which should be downloadable as xls files.