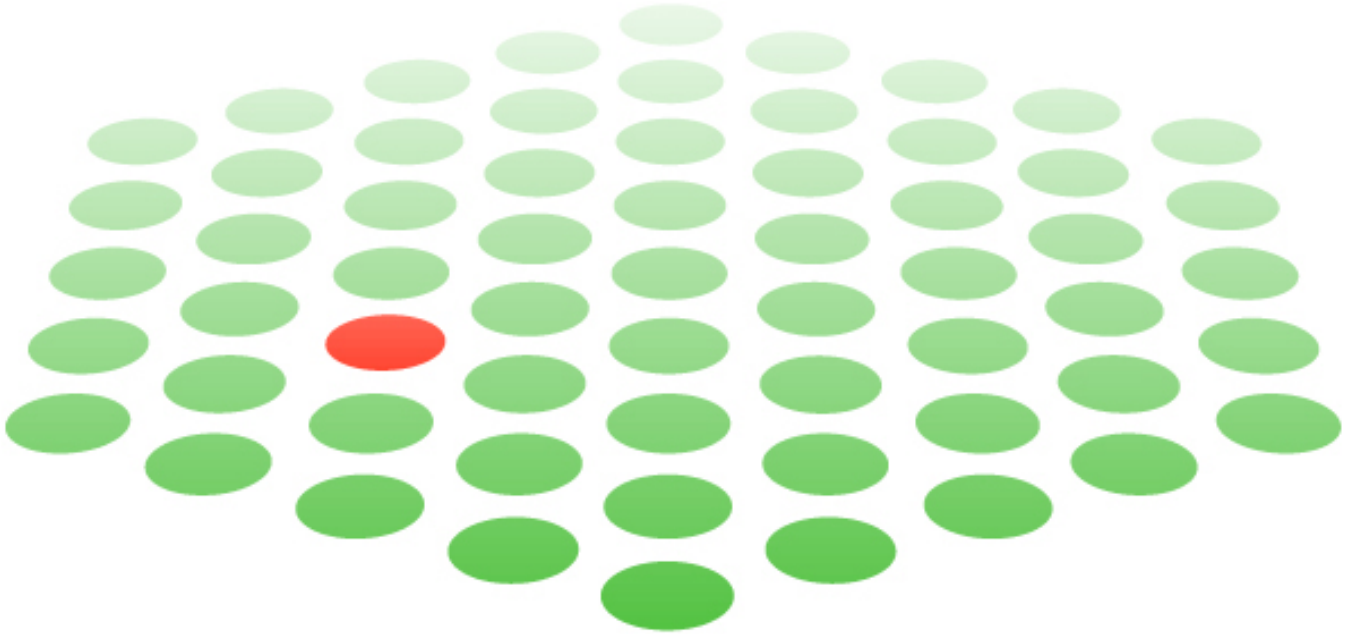
Biomedical literature mining

(and why we *really* need Open Access)



Lars Juhl Jensen
EMBL Heidelberg

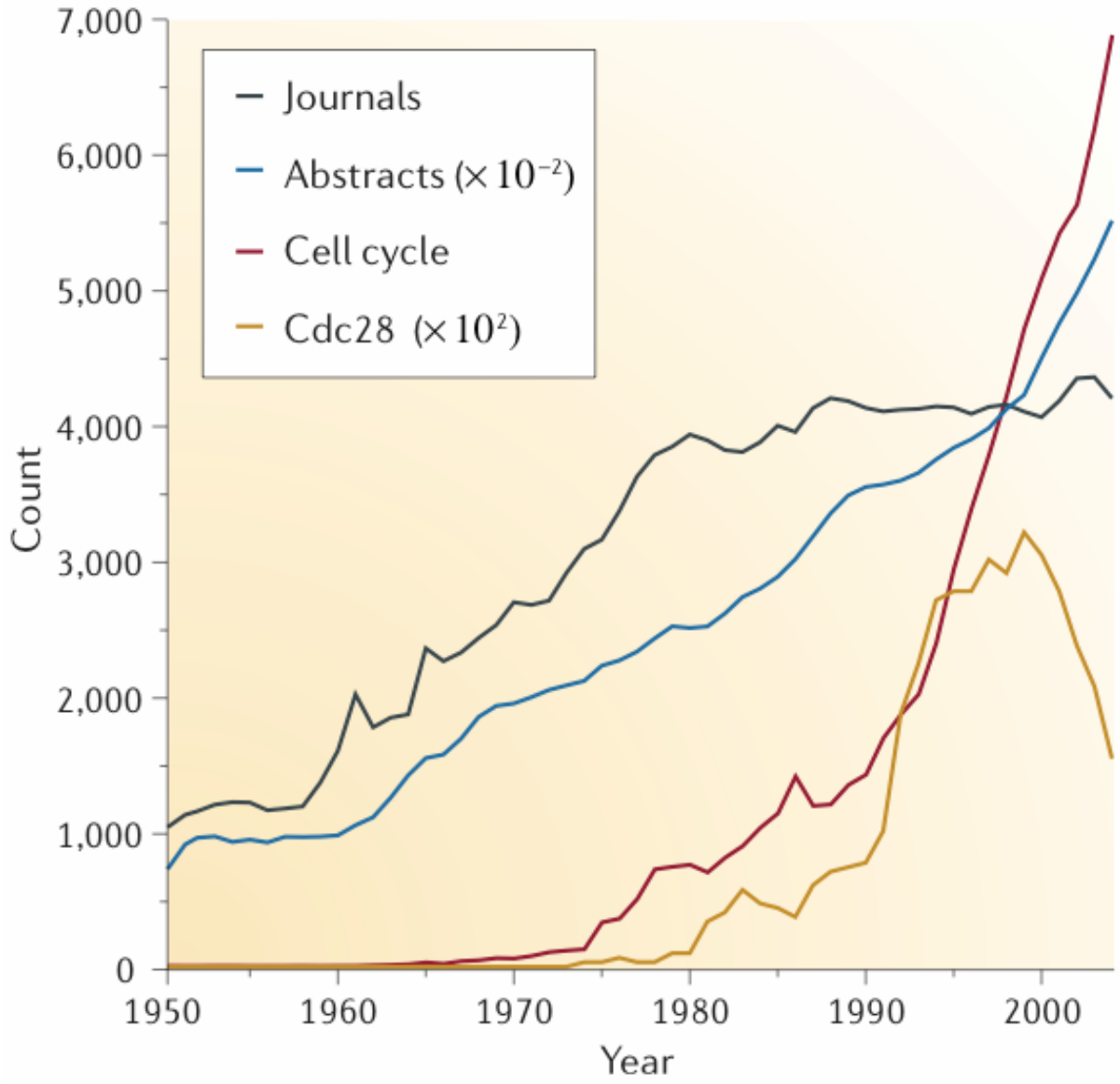
why biomedicine?

why literature mining?

why open access?

MEDLINE

17 million citations



Jensen et al., Nature Reviews Genetics, 2006

too much to read

literature mining

open access

information retrieval

finding the papers

ad hoc retrieval

About Entrez

Text Version

Entrez PubMed
Overview
Help | FAQ
Tutorial
New/Noteworthy
E-Utilities

PubMed Services
Journals Database
MeSH Database
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
Special Queries
LinkOut
My NCBI (Cubby)

Related Resources
Order Documents
NLM Mobile
NLM Catalog
NLM Gateway
TOXNET
Consumer Health
Clinical Alerts
ClinicalTrials.gov
PubMed Central

- To get started, enter one or more search terms.
- Search terms may be topics, authors or journals.

News flash: deliver PubMed search results directly to your desktop with an RSS feed.

To set up an RSS feed:

(1) Run your search in PubMed.

(2) Select *RSS Feed* from the *Send to* menu.

(3) Click *Create Feed* and copy the XML icon into your RSS Reader.

Read the PubMed Help to explore other options for automated e-mail updates using My NCBI.

PubMed is a service of the National Library of Medicine that includes over 15 million citations from MEDLINE and other life science journals for biomedical articles back to the 1950s. PubMed includes links to full text articles and other related resources.

users-specified query

“yeast AND cell cycle”

stemming

yeast / yeasts

dynamic query expansion

yeast / *S. cerevisiae*

About Entrez
Text Version

Entrez PubMed
Overview
Help | FAQ
Tutorial
New/Noteworthy
E-Utilities

PubMed Services
Journals Database
MeSH Database
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
Special Queries
LinkOut
My NCBI (Cubby)

Related Resources
Order Documents
NLM Mobile
NLM Catalog
NLM Gateway
TOXNET
Consumer Health
Clinical Alerts
ClinicalTrials.gov
PubMed Central

Items 1 - 20 of 13713

1: Homma MK, Wada I, Suzuki T, Yamaki J, Krebs EG, Homma Y. — Links

CK2 phosphorylation of eukaryotic translation initiation factor 5 potentiates cell cycle progression.
Proc Natl Acad Sci U S A. 2005 Oct 14; [Epub ahead of print]
PMID: 16227438 [PubMed - as supplied by publisher]

2: Leevers SJ, McNeill H. — Links

Controlling the size of organs and organisms.
Curr Opin Cell Biol. 2005 Oct 11; [Epub ahead of print]
PMID: 16226450 [PubMed - as supplied by publisher]

3: Wu JQ, Pollard TD. — Related Articles, Links

Counting cytokinesis proteins globally and locally in fission yeast.
Science. 2005 Oct 14;310(5746):310-4.
PMID: 16224022 [PubMed - in process]

4: David-Pfeuty T. — Related Articles, Links

The flexible evolutionary anchorage-dependent Pardee's restriction point of mammalian cells. How its deregulation may lead to cancer.
Biochim Biophys Acta. 2005 Sep 20; [Epub ahead of print]
PMID: 16219425 [PubMed - as supplied by publisher]

5: Anekonda TS, Reddy PH. — Related Articles, Links

Neuronal protection by sirtuins in Alzheimer's disease.
J Neurochem. 2005 Oct 7; [Epub ahead of print]
PMID: 16219030 [PubMed - as supplied by publisher]

6: Heinisch JJ. — Related Articles, Links

Baker's yeast as a tool for the development of antifungal kinase inhibitors-targeting protein kinase C and the cell integrity pathway.
Biochim Biophys Acta. 2005 Sep 12; [Epub ahead of print]

MEDLINE

abstracts

complete papers

Mitotic cyclin (Clb2)-bound Cdc28 (Cdk1 homolog) directly phosphorylated Swe1 and this modification served as a priming step to promote subsequent Cdc5-dependent Swe1 hyperphosphorylation and degradation

yeast?

cell cycle?

entity recognition

identifying the substance(s)

Mitotic cyclin (Clb2)-bound Cdc28 (Cdk1 homolog) directly phosphorylated Swe1 and this modification served as a priming step to promote subsequent Cdc5-dependent Swe1 hyperphosphorylation and degradation

Cdc28 $\Rightarrow$ yeast

Cdc28 $\Rightarrow$ cell cycle

good synonyms list

manual curation

orthographic variation

CDC28

Cdc28p

disambiguation

hairy

SDS

Cdc2

iHOP

information hyperlinked over proteins

Search Gene

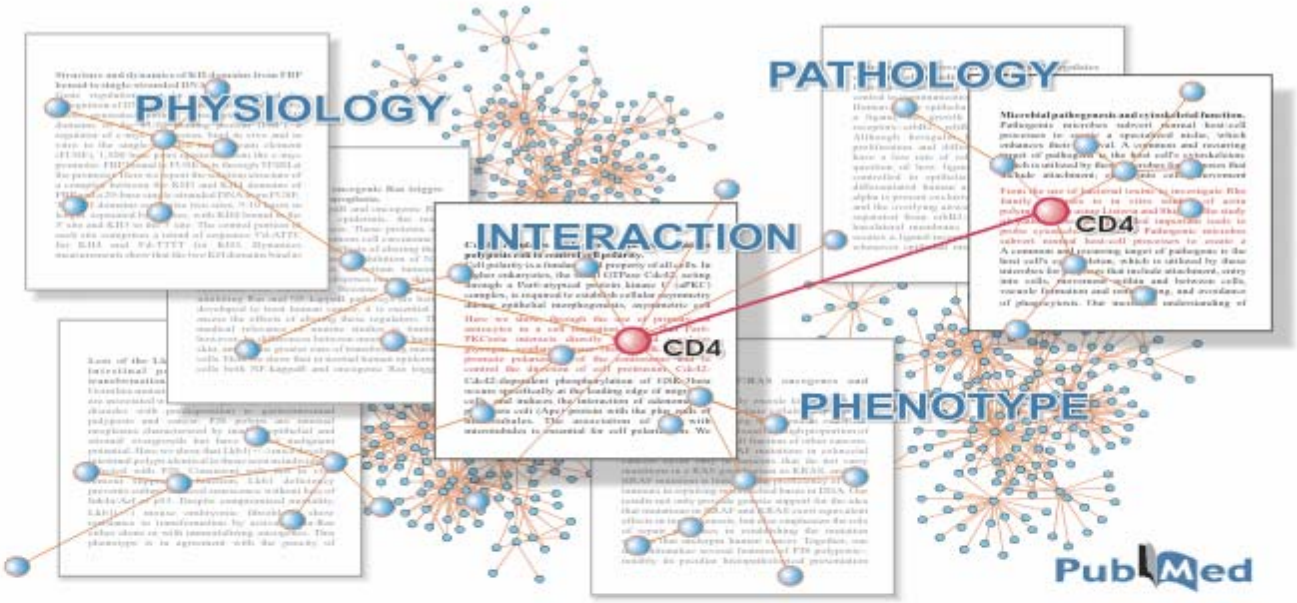
Gene Model
Developer's Zone new
Contact
Help

COCOON

Java

site powered by pdg

Concept & Implementation
by Robert Hoffmann



Hoffmann, R., Valencia, A. A Gene Network for Navigating the Literature. Nature Genetics 36, 664 (2004)

Search for a gene *synonym* or *accession number*... (Click here for an example: SNF1)

all fields in all organism

[SEARCH]

Symbol	Name	Synonyms	Organism
CDC28	Cell division control protein 28	CDK1, HSL5, SRM5, YBR1211, YBR160W	Saccharomyces cerevisiae

UniProt	P00546
IntAct	P00546
NCBI Gene	852457
NCBI RefSeq	NP_009718
NCBI Accession	CAA25065, CAA56509, CAA85119

Homologues of CDC28 ... new

Definitions for CDC28 ...

Enhanced PubMed/Google query ... new

WARNING: Please keep in mind that gene detection is done automatically and can exhibit a certain error. Read more.

Find in this Page

Furthermore, SW14 associates with CLB2 protein and is a substrate for the CLB2-**associated** CDC28 kinase in vitro.

Furthermore, the Cks1 protein was shown to be physically **associated** with active forms of the Cdc28 protein kinase.

The cyclin-dependent kinase Cdc28p **associates** with the cyclin Clb2p to induce mitosis in the yeast Saccharomyces cerevisiae.

We find that G1 arrest in the cdc37-1 mutant is accompanied by a decrease in the Cdc28 activity **associated** with the G1 cyclin Cln2.

We found that Hct1 was **phosphorylated** in vivo at multiple CDK consensus sites during cell cycle stages when activity of the cyclin-dependent kinase Cdc28 is high and APC activity is low.

It is likely, therefore, that Cks1 mediates a more specialized **function** of the Cdc28 kinase such as its ability to form specific multimeric complexes or to localize properly in cellular compartments.

Cdc37 **promotes** the stability of protein kinases Cdc28 and Cak1.

In addition, Cdc37 **promotes** the production of Cak1, but not that of Cdc28, when coexpressed in insect cells.

The B-type cyclins Clb5 and Clb6 are the primary activators of the S phase **function** of the budding yeast CDK Cdc28.

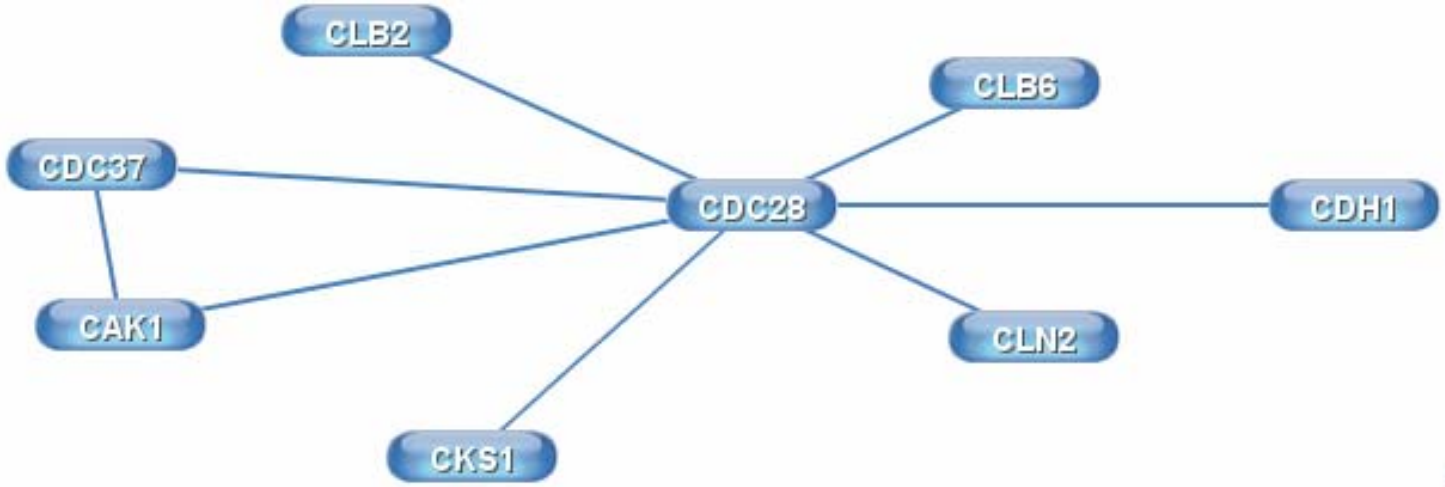
All three cak1 mutants displayed significant synthetic interactions with loss-of-**function** mutations in CDC28 and KIN28.

abstracts

complete papers

information extraction

formalizing the facts



[Redo graph layout]

The B-type cyclins Clb5 [CLB5] and Clb6 [CLB6] are the primary activators of the S phase **function** of the budding yeast CDK Cdc28 [CDC28].

Furthermore, SW14 associates with CLB2 [CLB2] protein and is a substrate for the CLB2-**associated** CDC28 [CDC28] kinase in vitro.

The cyclin-dependent kinase Cdc28 [CDC28]p **associates** with the cyclin Clb2 [CLB2]p to induce mitosis in the yeast Saccharomyces cerevisiae.

In addition, Cdc37 [CDC37] **promotes** the production of Cak1 [CAK1], but not that of Cdc28 [CDC28], when coexpressed in insect cells.

We found that Hct1 [CDH1] was **phosphorylated** in vivo at multiple CDK consensus sites during cell cycle stages when activity of the cyclin-dependent kinase Cdc28 [CDC28] is high and APC activity is low.

All three cak1 [CAK1] mutants displayed significant synthetic interactions with loss-of-**function** mutations in CDC28 [CDC28] and KIN28 [KIN28].

Cdc37 [CDC37] **promotes** the stability of protein kinases Cdc28 [CDC28] and Cak1 [CAK1].

Furthermore, the Cks1 [CKS1] protein was shown to be physically **associated** with active forms of the Cdc28 [CDC28] protein kinase.

We find that G1 arrest in the cdc37-1 mutant is accompanied by a decrease in the Cdc28 [CDC28] activity **associated** with the G1 cyclin Cln2 [CLN2].

It is likely, therefore, that Cks1 [CKS1] mediates a more specialized **function** of the Cdc28 [CDC28] kinase such as its ability to form specific multimeric complexes or to localize properly in cellular compartments.

co-mentioning

statistical methods

NLP

Natural Language Processing

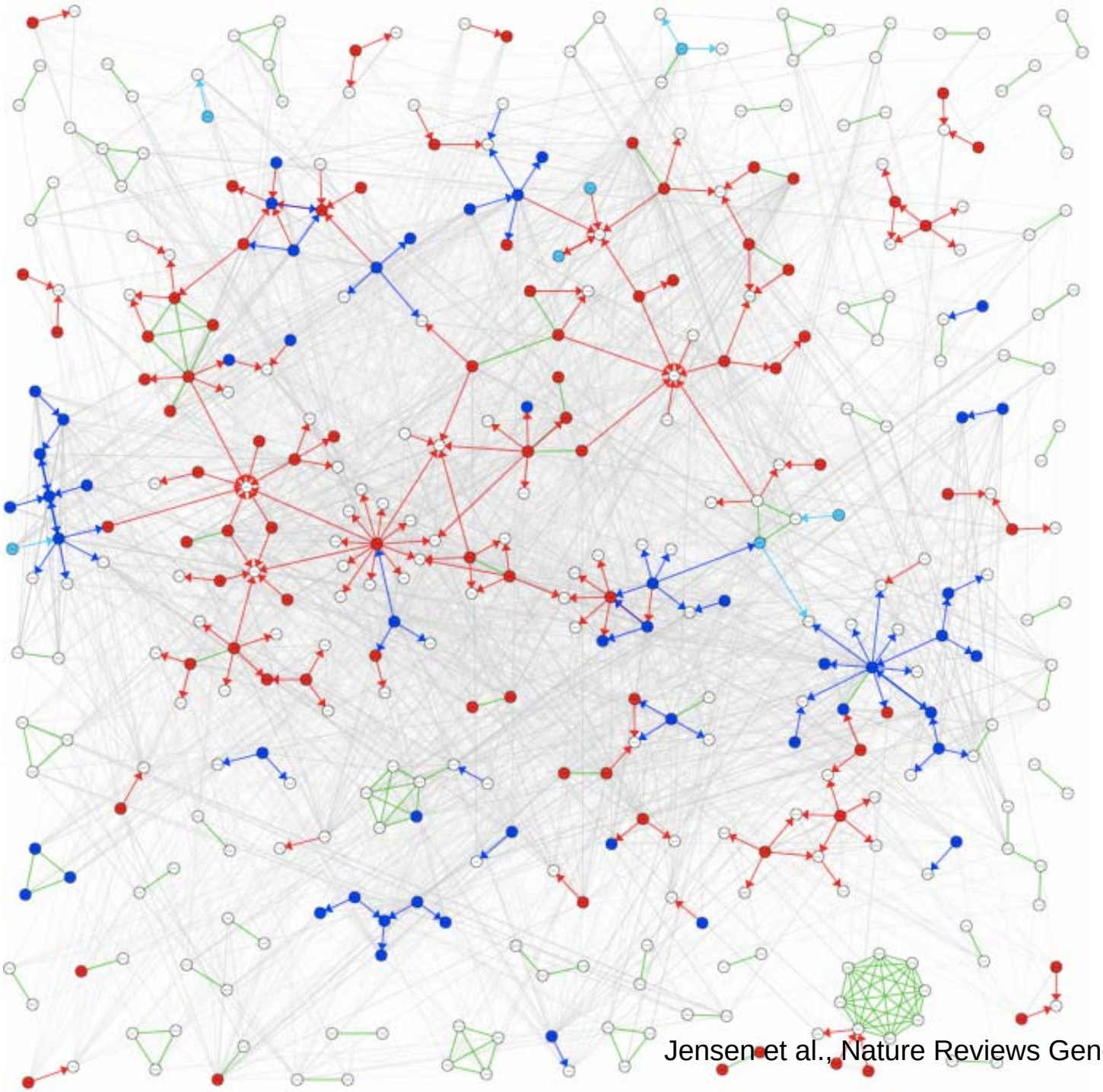
Gene and protein names

Cue words for entity recognition

Verbs for relation extraction

[nxexpr The expression of
[nxgene the cytochrome
genes
[nxpg CYC1 and CYC7]]]
is controlled by
[nxpg HAP1]

Mitotic cyclin (Clb2)-bound Cdc28 (Cdk1 homolog) directly phosphorylated Swe1 and this modification served as a priming step to promote subsequent Cdc5-dependent Swe1 hyperphosphorylation and degradation



Jensen et al., Nature Reviews Genetics, 2006

new discoveries

text mining

WELCOME: This site is open during construction.

Ready to use **Arrowsmith?**

Start

References

Start

Corner for Collaborative Informatics

Go to Website

Author-ity: A tool for identifying Medline articles written by a particular author

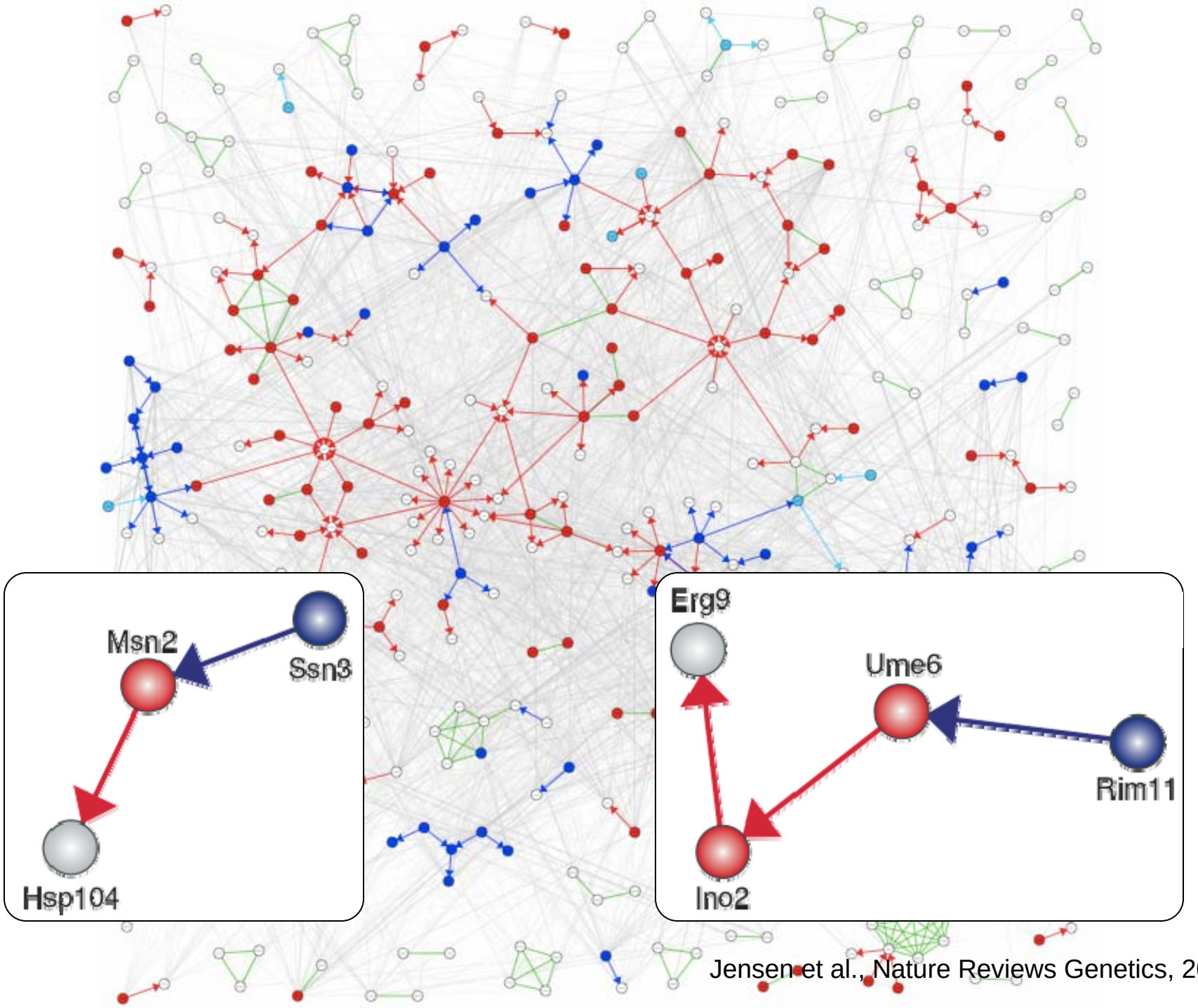
Start

Compendium of Biomedical Text Mining Tools

Start

Anne O'Tate: A tool for summarizing the results of a PubMed query.

Start

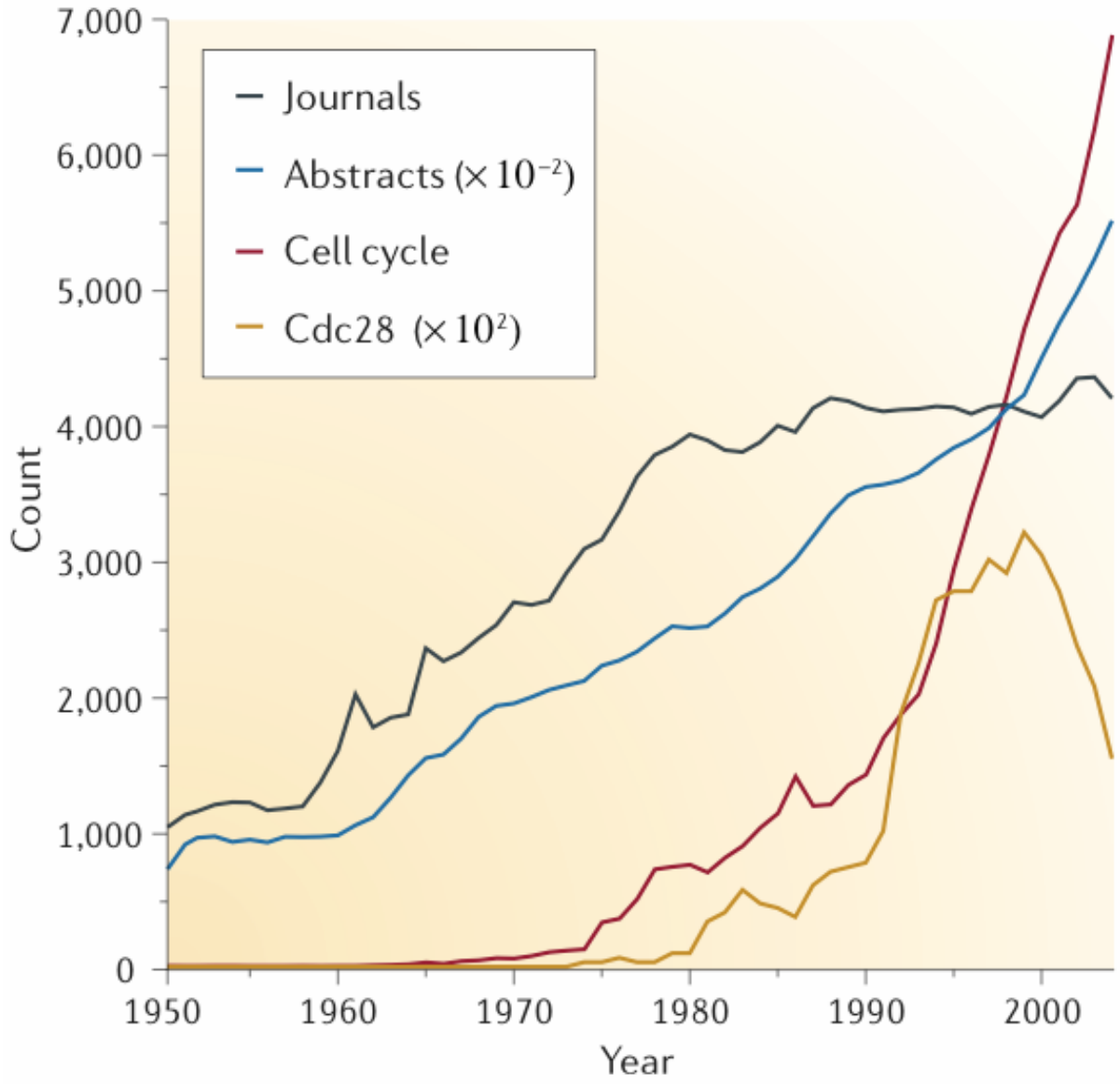


Jensen et al., Nature Reviews Genetics, 2006

abstracts

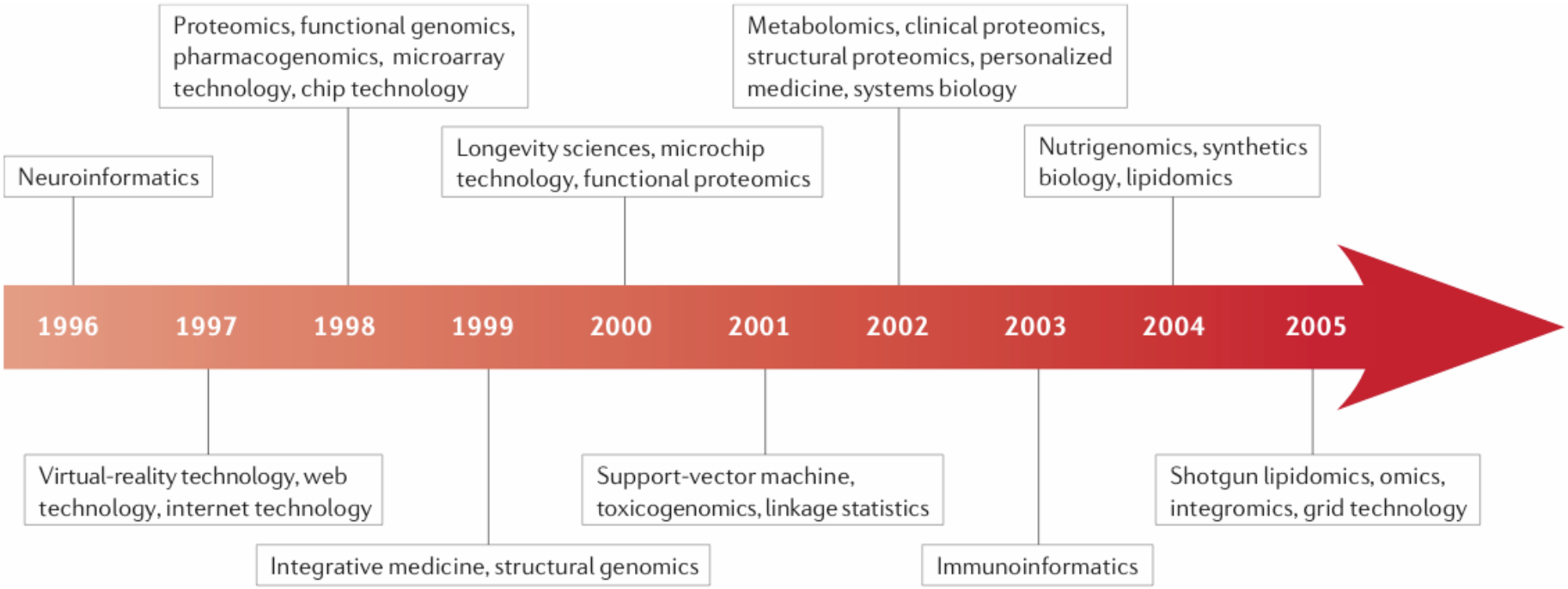
complete papers

temporal trends



Jensen et al., Nature Reviews Genetics, 2006

buzzwords



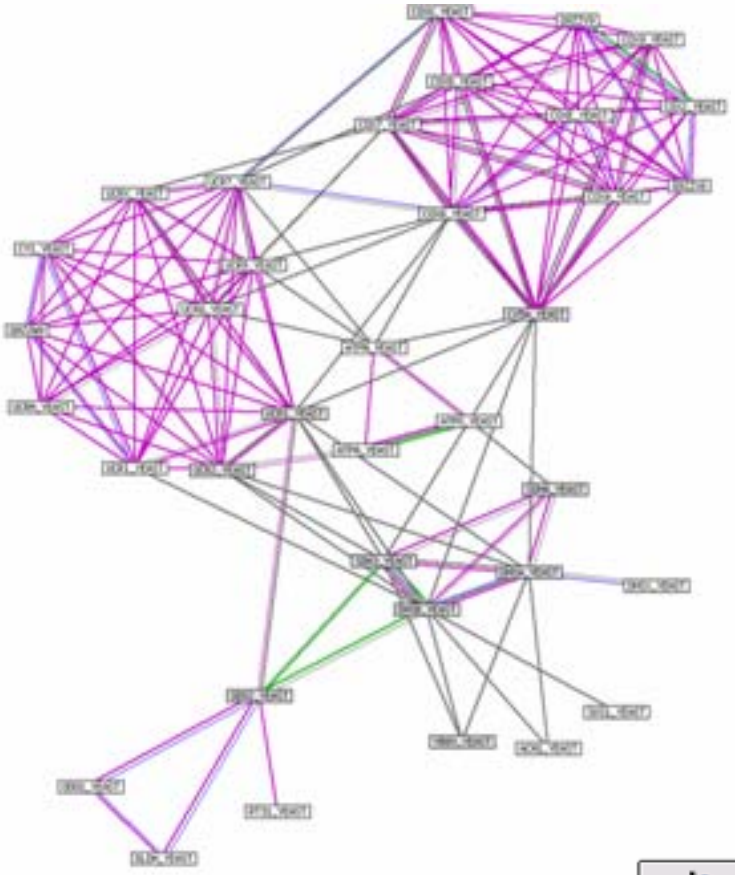
Jensen et al., Nature Reviews Genetics, 2006

grant applications

integration of text and data

Home · Download · Help/Info

STRING



Genomic neighborhood

Species co-occurrence

Gene fusions

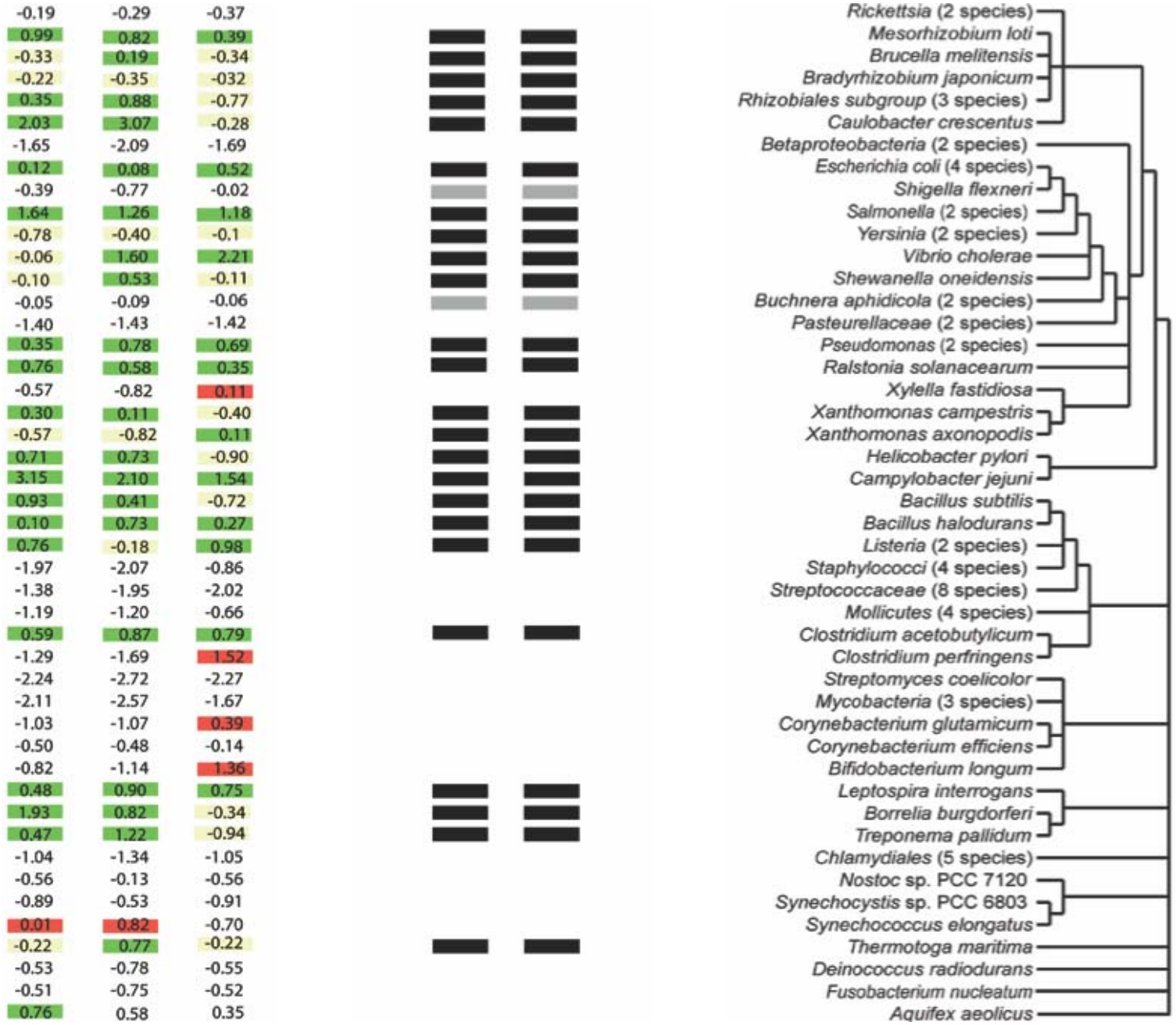
Experimental interaction data

Microarray expression data

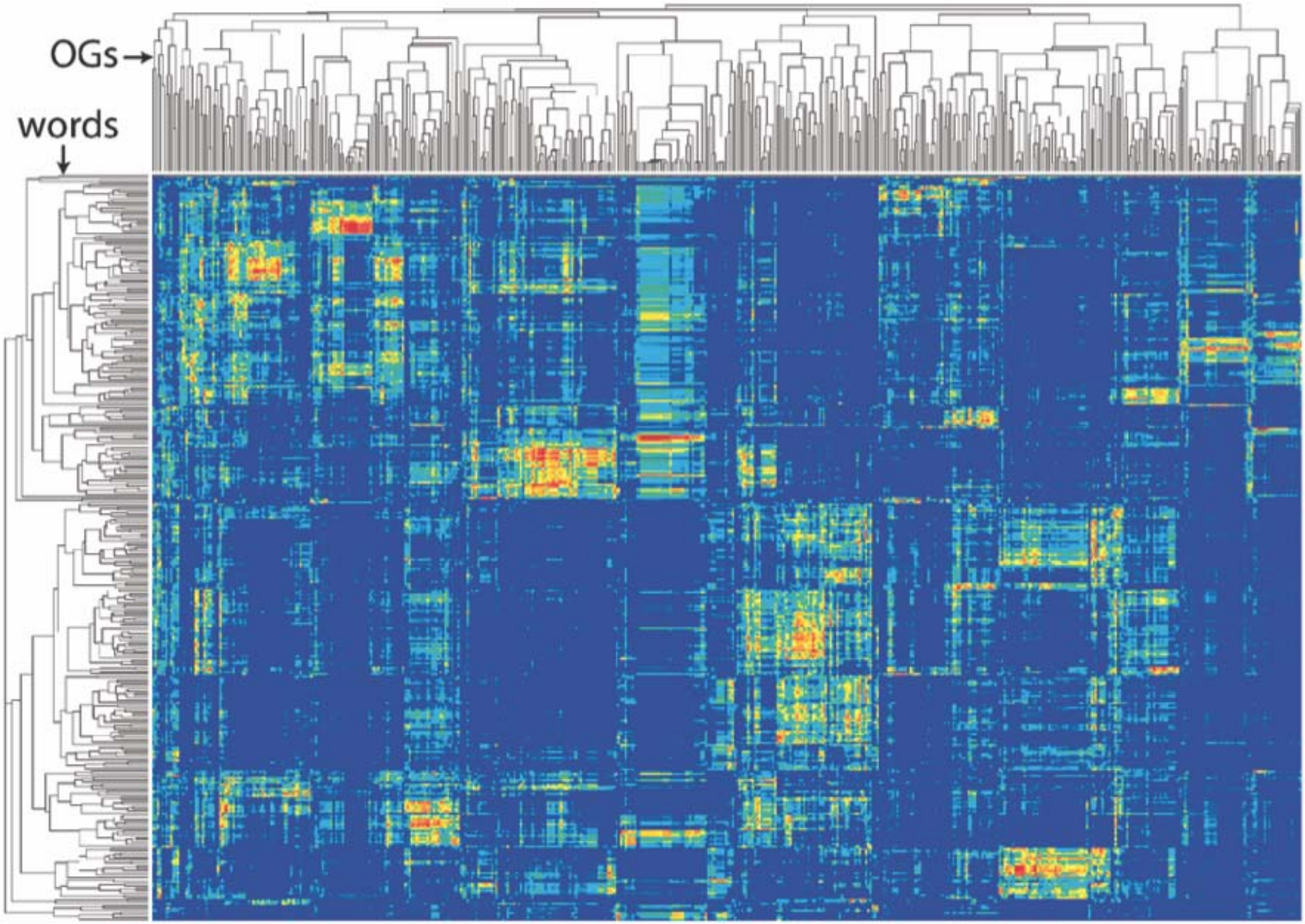
Database imports

Literature mining

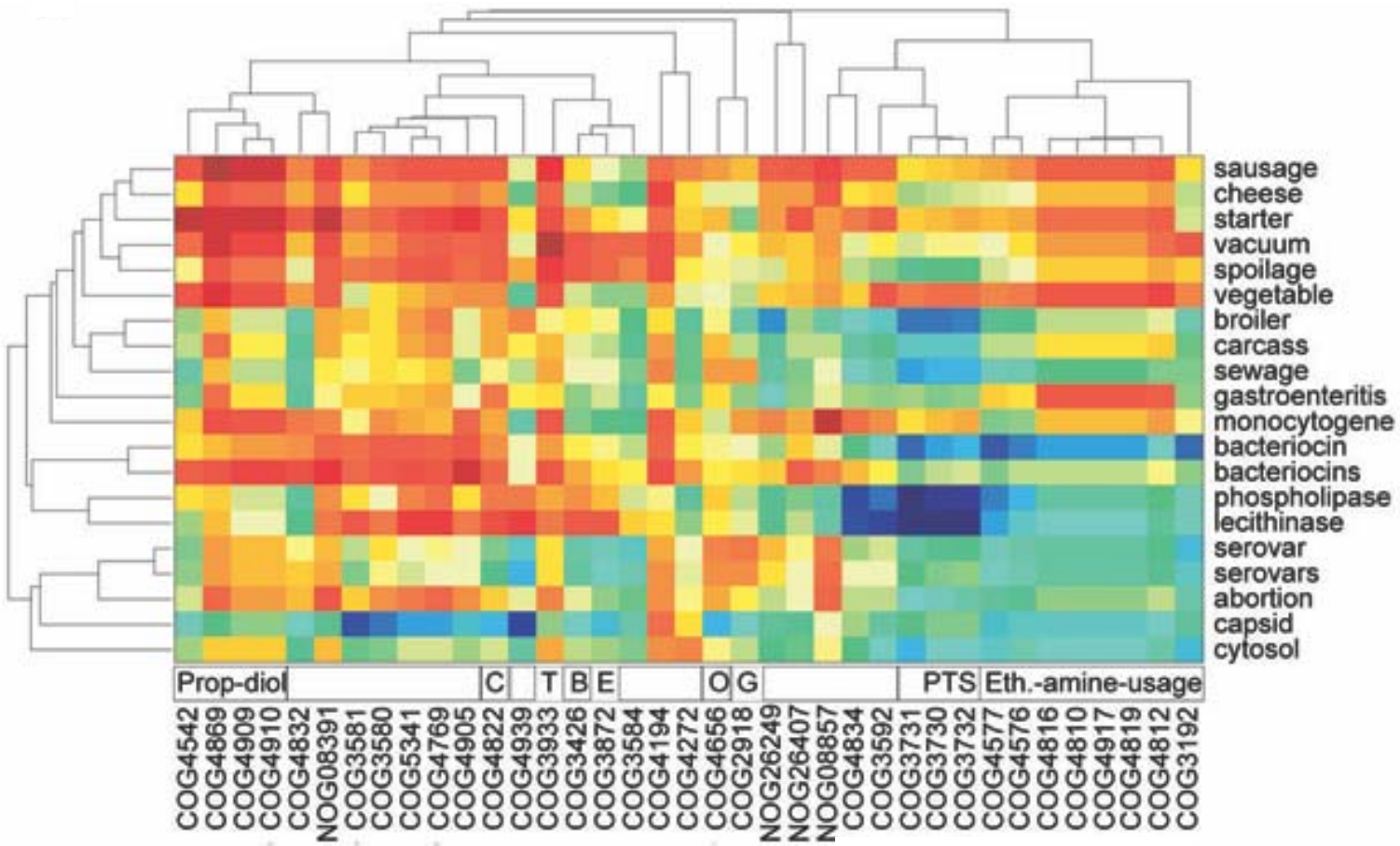
genotype to phenotype



Korbel et al., PLoS Biology, 2005

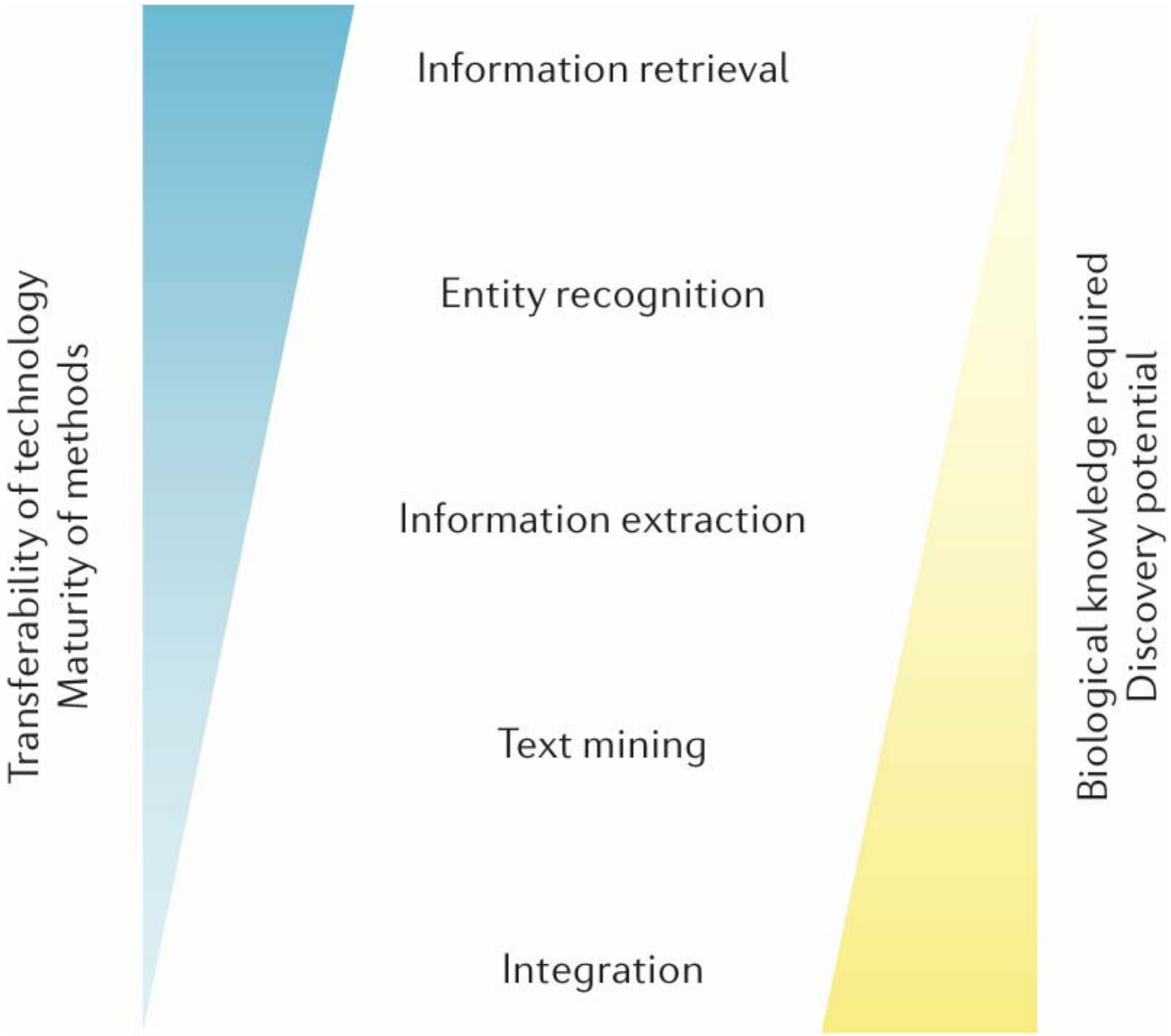


Korbel et al., PLoS Biology, 2005



Korbel et al., PLoS Biology, 2005

where are we now?



Jensen et al., Nature Reviews Genetics, 2006

abstracts

complete papers

restricted access

open access

the tools are there

now we need the text!

Acknowledgments

Jasmin Saric
Rossitza Ouzounova
Michael Kuhn
Jan Korbel
Tobias Doerks
Isabel Rojas
Miguel Andrade
Peer Bork