

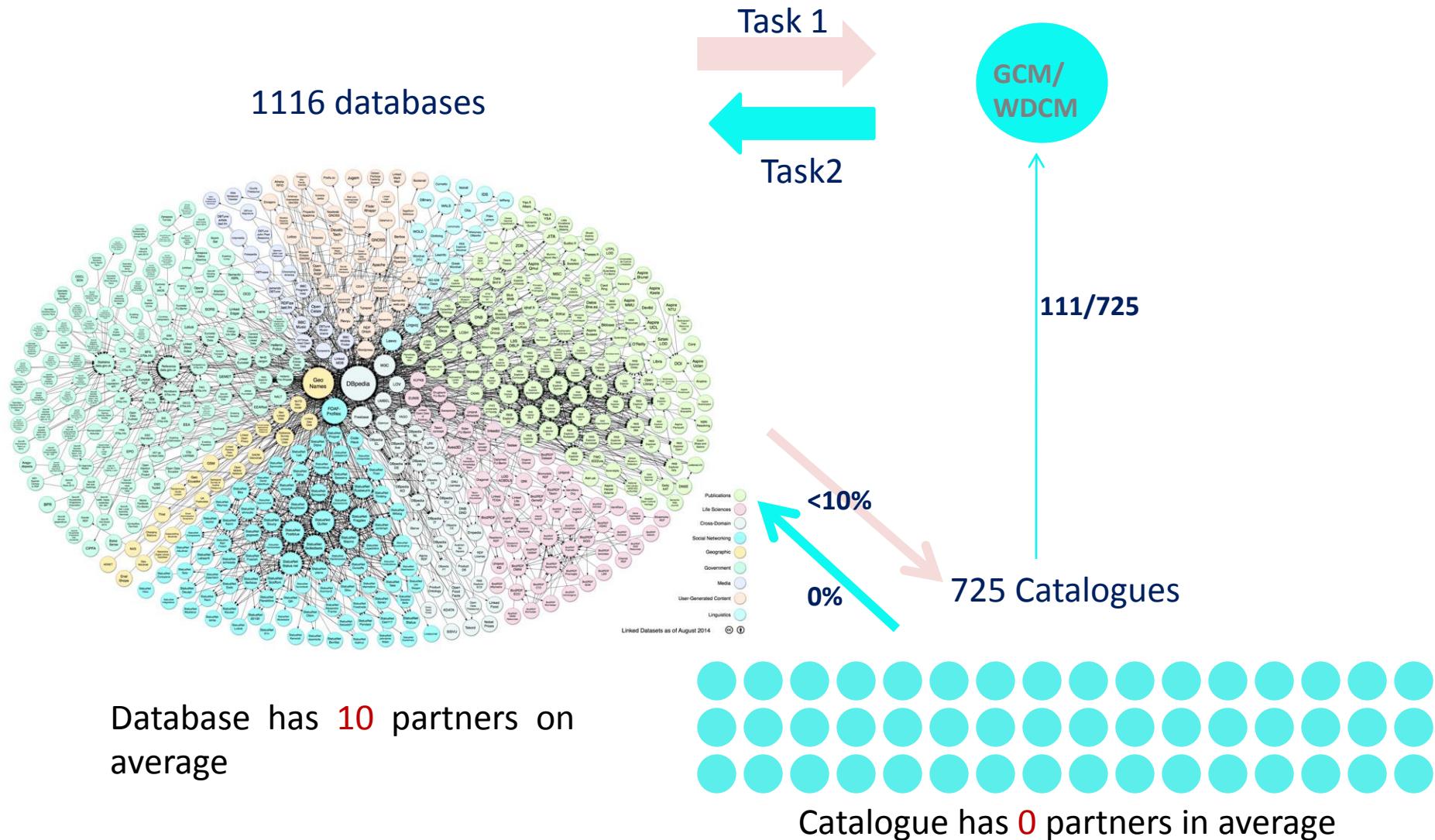
# Data integration with Life Science databases: the technical aspects

Alexander Vasilenko, Svetlana Ozerskaya, Oleg Stupar, Lujdmila Evtushenko ,  
Paolo Romano, Linhuan Wu, Juncai Ma

# Potential help for the Life Science world from microbial culture collections (mBRCs)

1. Assurance of repeatability of experimental data
2. Resolving nomenclatural issues related to microorganisms
3. Strain-specific characters

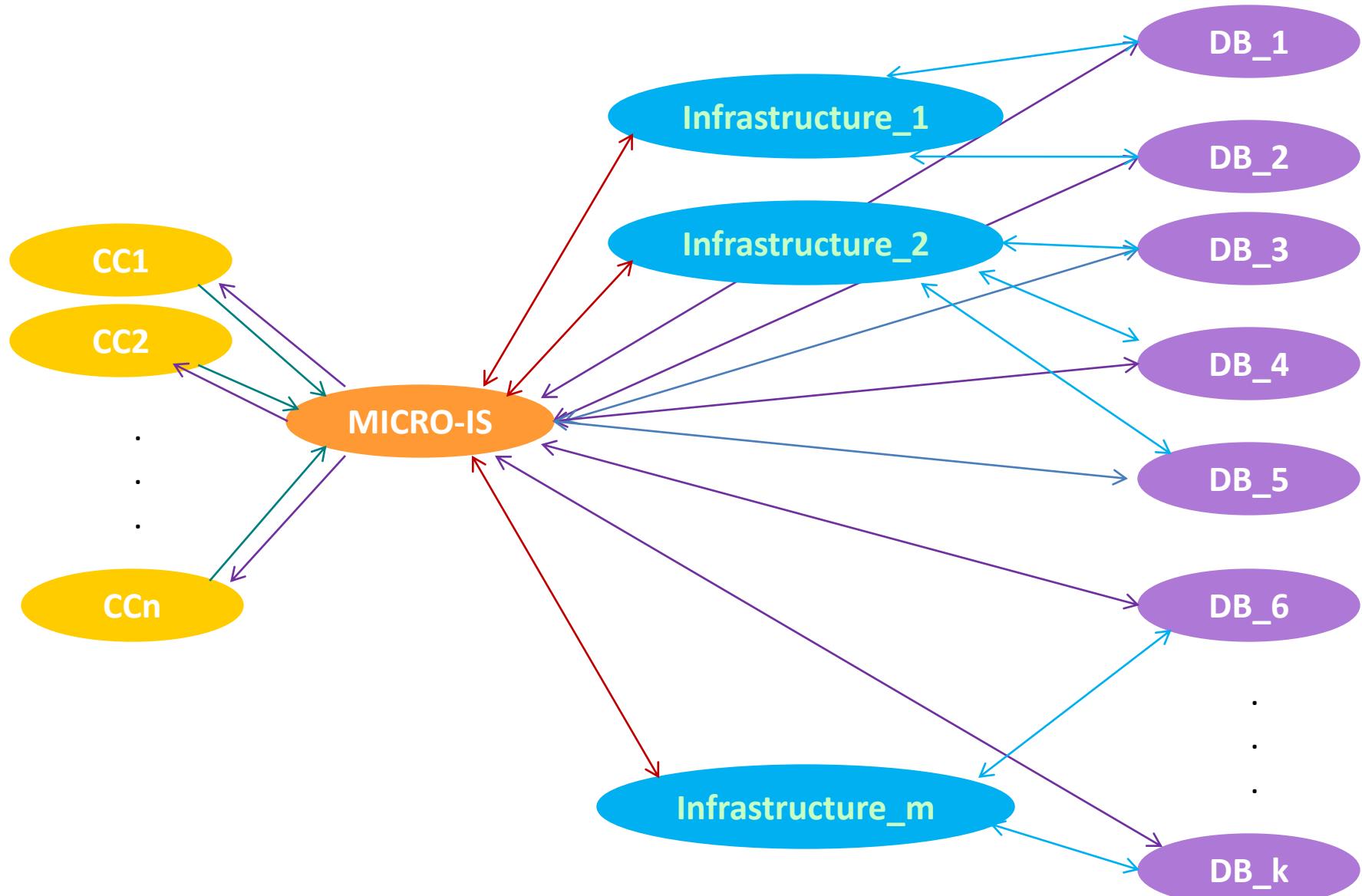
# Microbial databases – CC interconnection



Potentially this data integration could mean the tasks:

1. To make CC data visible and accessible from partner Life Science databases,
2. To make partner database records visible and accessible from CC aggregated catalogue,  
in the formats:
  - a. To give this data integration for human access,
  - b. To give this data integration for computer programs.

# General integration schema



# Life Science databases inspected

Total number of life science database names or references discovered in this study is more than 14 800. The total number of database references inspected manually is more than 5500 . The total number of life science databases collected visible online is 2660, the number of databases with microbial data collected is 1116 (plus 9318 bacterial databases in BioCyc).

The main sources:

- MB (1802 entries (<http://metadatabase.org>), 26.12.2015),
- Biosharing (724 databases, (<https://www.biosharing.org/>), 26.12.2015),
- BioMedBrigeds (814 Databases, 27.12.2015),  
(<http://wwwdev.ebi.ac.uk/fgpt/toolsui/>),
- Pathguide (363 database names, 2013), (<http://www.pathguide.org/>)
- ELIXIR list (579 entries, (<https://bio.tools/?q=database>), 28.1.2016)
- ExPASy (85 + 665 databases, ([http://www.expasy.org/old\\_links](http://www.expasy.org/old_links)), 12.2.2016)
- Bioinformatics Links Directory (621 databases)
- OBRC (<http://www.hsls.pitt.edu/obrc/>) 30.3.2017

# Databases parameters collected (an example)

- Unique identifier: BIODECORE-000438
- Database acronym: dbSNP
- Database name: The Database of Short Genetic Variation (single nucleotide polymorphism)
- Database URL: <http://www.ncbi.nlm.nih.gov/SNP/>
- Access level: Open
- Practical domain: health
- Microbial level: sp
- Year of the last correction: 2017
- Developer/Owner: USA, NCBI; USA, National Library of Medicine, National Institutes of Health
- Comment: Escherichia coli
- Orientation: -
- Properties: DNA, gene, genome, proteomics, publications, RNA
- Search by: -
- Ontologies list: SO
- Partner databases: Assembly, BioProject, BioSample, ClinVar, dbGaP, dbMHC, dbSTS, dbVar, Ensembl, GenBank, Homologene, IGSR, MapViewer, NCBI Gene, Nucleotide, OMIM, PMC, Protein, PubChem Substance, PubMed, RDP, RefSeq, UniGene, UniProtKB
- Program interface: ELIXIR WEB UI, Entrez Programming Utilities (E-Utils)

# Property keywords in databases

826	Gene	226	Interactome
716	Proteomics	219	Taxonomy
625	Publications	195	Drugs
517	Image	195	Peptide
389	RNA	183	Molecules
395	DNA	166	Metabolite
355	Genome	151	Pathogen
361	Enzyme	167	Immunology
316	Cell	156	Toxicology
297	Chemistry	132	Lipid
270	Pathways	45	Microbiome
263	Disease		

Average number of keywords assigned to a database = 6,55

# Databases with microbiome data

Biology Reference	HPMCD	PANGAEA
BioSamples	IMG	PLOS One
BioSystems	IMG/M	PMC
Bookshelf	IMG/VR	PSP
EMBL	MEDLINE	pubget
EMBL-EBI	MeSH	PubMed
ENA	Microbiome	PubMed Health
Espacenet	NARCIS	QIAGEN
Europe PMC	NCBI	RefSeq
ForestScience Current Database	NFSD	ScienceDirect
GO Database	NLM Catalog	SRA
GONUTS	Nowomics	TACONIC
GoPubMed	OMIM	UniProtKB
HGTree	OMIM (1)	VetMed Resource
HOMD	OReFiL	WikiGenes

# Biggest database producer: BESC

(BioEnergy Science Center)

**BioCyc** pathway/genome database: 9367 databases totally (<http://www.biocyc.org/biocyc-pgdb-list.shtml>)

Group 1 are 7 databases: **EcoCyc**, **MetaCyc**, **HumanCyc**, **AraCyc**, **YeastCyc**, **LeishCyc**, **TrypanoCyc**

Group 2 are 41 databases generated by program with curation done each is one strain:

*Agrobacterium fabrum C58*

*Anopheles gambiae*

*Aurantimonas manganoxydans SI85-9A1*

*Bacillus anthracis Ames*

*Bacillus subtilis 168*

*Bacteroides thetaiotaomicron VPI-5482*

*Candidatus Cardinium hertigii*

*Candidatus Evansia muelleri*

*Candidatus Portiera aleyrodidarum BT-QVLC*

*Caulobacter crescentus CB15*

*Caulobacter crescentus NA1000*

*Chlamydomonas reinhardtii*

*Clostridium saccharoperbutylacetonicum ATCC 27021*

*Cryptosporidium hominis TU502*

*Cryptosporidium parvum Iowa*

*Drosophila melanogaster*

*Escherichia coli B str REL606*

*Escherichia coli CFT073*

*Escherichia coli K-12 substr W3110*

*Escherichia coli O157:H7 str EDL933*

*Eubacterium rectale ATCC 33656*

*Helicobacter pylori 26695*

*Listeria monocytogenes 10403S*

*Methylosinus trichosporium OB3b*

*Mus musculus*

*Mycobacterium tuberculosis CDC1551*

*Mycobacterium tuberculosis H37Rv*

*Penicillium chrysogenum Wisconsin 54-1255*

*Peptoclostridium difficile 630*

*Plasmodium berghei ANKA*

*Plasmodium chabaudi*

*Plasmodium falciparum 3D7*

*Plasmodium vivax Sal-1*

*Plasmodium yoelii 17XNL*

*Schistosoma mansoni*

*Shigella flexneri 2a str 2457T*

*Streptomyces coelicolor A3(2)*

*Synechococcus elongatus PCC 7942*

*Thalassiosira pseudonana CCMP1335*

*Toxoplasma gondii ME49*

*Vibrio cholerae O1 biovar El Tor str N16961*

Group 3 are 9318 databases each database is one bacterial strain with no curation yet

10

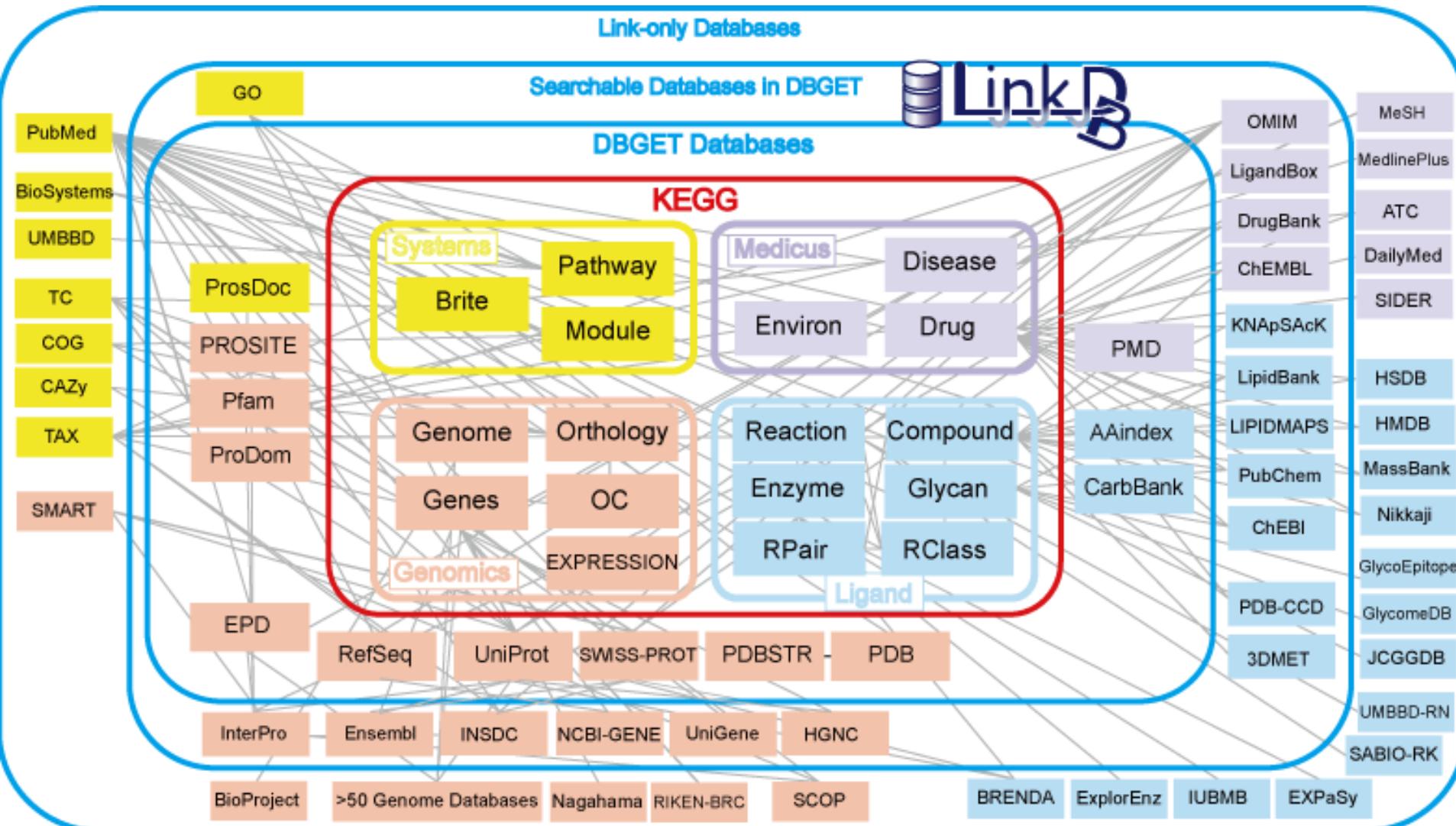
# EMBL-EBI (97<sub>dbs</sub>)

ArrayExpress	e!EnsemblChimpanzee	Ensembl	logRECOORD	PICR
ASD	ee	Enzyme Portal	MACiE	PomBase
ASTD	e!EnsemblCow	Enzyme Structures	MEROPS	PRIDE
ATD	e!EnsemblDog	EVA	MetaboLights	PROCOGNATE
BioModels	e!EnsemblFugu	Expression Atlas	Metal MACiE	Reactome
BioSamples	e!EnsemblFungi	FunTree	MicroCosm	RECOORD
Cellular Phenotype	e!EnsemblGenomes	GeneDB	MIRIAM collection	Rfam
Db	e!EnsemblGorilla	GWAS Catalog	MTBLS	RNAcentral
ChEBI	e!EnsemblHorse	HGNC	NRNL1	SAS
ChEMBL	e!EnsemblMetazoa	HipSci	NRNL2	SRS@EMBL-EBI
CluSTr	e!EnsemblMouse	IGSR	NRPL1	SureChEMBL
CSA	e!EnsemblPig	IMEx	NRPL2	TreeFam
DGVa	e!EnsemblPlants	IMGT/HLA	OLDERADO	UniChem
DNAtraffic	e!EnsemblProtists	IntAct	PANDIT	UniProt-GOA
DrugPort	e!EnsemblRabbit	IntEnz	PDBe	UniSave
e!Ensembl	e!EnsemblZebrafish	InterPro	PDBe EM Resources	VASCO
e!Ensembl S. cerevisiae	EGA	IPD	PDBeChem	VectorBase
e!Ensembl Bacteria	EMBL	IPD-ESTDAB	PDBsum	
e!Ensembl Cat	EMBL-EBI	IPD-HPA	Pfam	
e!Ensembl Chicken	EMDB	IPD-KIR	Pfam	
	ENA	IPD-MHC	PhenoDigm	

# NCBI (70<sub>dbs</sub>)

Assembly	Dengue virus database	MedGen	PubChem Compound
BioProject	ECRbase	MEDLINE	PubChem Substance
BioSample	Genbank	MeSH	PubMed
BioSystems	Gene	MMDB	PubMed Health
Bookshelf	Genetic Codes	NCBI	RefSeq
CCDS	Genome	NCBI taxonomy	RefSeqGene
CDD	GEO	NCBI Trace Archives	Retroviruses
ClinGen	GEO DataSets	NLM Catalog	SKY/M-FISH and CGH
ClinVar	GEO Profiles	Nucleotide	SPARCLE
Clone DB	GSS	OMIM	SpliceInfo
COGs	GTR	Organelle genomes	SRA
dbEST	Histone	PMC	Structure
dbGaP	HIV-1	PopSet	TPA
dbMHC	Homologene	Probe	UniGene
dbProbe	IBIS	Protein	UniVec
dbSNP	Influenza Virus Resource	Protein Clusters	Viral genomes
dbSTS	MapViewer	PubChem	Virus Variation
dbVar		PubChem BioAssay	

# Databases interconnection example



# Interconnection matrix

	3D Lectin	3D RIBOSC	5S RNA	Da	A.pernix	UniPROBE	UniProt-G	UniProtKB	UniRef	YMP	YPM	Zif-BASE	AN
2P2Idb	0	0	0	0	0	0	0	0	0	0	0	0	1
3D Genome Browser	0	0	0	0	0	0	0	0	0	0	0	0	1
3D Lectin	0	0	0	0	0	0	0	0	0	0	0	0	0
3D RIBOSOMAL MODIFICATIO	0	0	0	0	0	0	0	0	0	0	0	0	3
3DBIONOTES	0	0	0	0	0	0	0	0	0	0	0	0	0
3Dee	0	0	0	0	0	0	0	0	0	0	0	0	0
3DEM Benchmark	0	0	0	0	0	0	0	0	0	0	0	0	0
3D-footprint	0	0	0	0	0	0	0	0	0	0	0	0	1
3DID	0	0	0	0	0	0	0	0	0	0	0	0	6
PubChem	0	0	0	0	0	0	0	0	0	0	0	0	57
PubChem BioAssay	0	0	0	0	0	0	0	0	0	0	0	0	16
PubChem Compound	0	0	0	0	0	0	0	0	0	0	0	0	31
PubChem Substance	0	0	0	0	0	0	0	0	0	0	0	0	25
pubget	0	0	0	0	0	0	0	0	0	0	0	0	2
Public Data Portal	0	0	0	0	0	0	0	0	0	0	0	0	1
PubMed	0	0	0	0	1	0	0	1	0	0	0	1	565
PubMed Health	0	0	0	0	0	0	0	0	0	0	0	0	6
YPM	0	0	0	0	0	0	0	0	0	0	0	0	1
ZFIN	0	0	0	0	0	0	1	1	0	0	0	0	20
Zif-BASE	0	0	0	0	0	0	0	0	0	0	0	0	0
ZOBODAT Vespoidea	0	0	0	0	0	0	0	0	0	0	0	0	1
ZorapteraSF	0	0	0	0	0	0	0	0	0	0	0	0	1
ИИПС	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>CN</b>	5	8	4	10		6	148	161	15	2	3	5	

UniProtKB database has integration parameters CN=161, AN=501.

In CC catalogues both values are mostly 0.

# Databases with more than 26 partners (CN)

364	<a href="#">Pathguide</a>	55	<a href="#">dbProbe</a>	38	<a href="#">NCBI taxonomy</a>	31	<a href="#">PANDORA</a>
161	<a href="#">UniProtKB</a>	55	<a href="#">NCBI</a>	38	<a href="#">OpenHelix</a>	30	<a href="#">FungiDB</a>
159	<a href="#">iProClass</a>	53	<a href="#">PiroplasmsDB</a>	38	<a href="#">ViralZone</a>	30	<a href="#">JPGV</a>
153	<a href="#">COL</a>	52	<a href="#">e!EnsemblGenomes</a>	37	<a href="#">GPCRs</a>	30	<a href="#">Reactome</a>
148	<a href="#">UniProt-GOA</a>	49	<a href="#">EMBL</a>	37	<a href="#">T3DB</a>	30	<a href="#">STRING</a>
143	<a href="#">OReFiL</a>	49	<a href="#">ENA</a>	36	<a href="#">EcoCyc</a>	29	<a href="#">BioModels</a>
129	<a href="#">EcoliWiki</a>	49	<a href="#">SBKB</a>	36	<a href="#">Nucleotide</a>	29	<a href="#">CAZy</a>
128	<a href="#">GeneCards</a>	46	<a href="#">EcoGene</a>	35	<a href="#">BacMet</a>	29	<a href="#">ConsensusPathDB</a>
91	<a href="#">PIR</a>	46	<a href="#">SGD</a>	35	<a href="#">ChEBI</a>	29	<a href="#">EuPathDB</a>
73	<a href="#">UCD 2D-PAGE</a>	45	<a href="#">Gene</a>	35	<a href="#">PRODORIC</a>	29	<a href="#">MACiE</a>
63	<a href="#">Hits</a>	44	<a href="#">InterMitoBase</a>	34	<a href="#">ESTHER</a>	29	<a href="#">RefSeq</a>
61	<a href="#">SWISS-2DPAGE</a>	43	<a href="#">EMBL-EBI</a>	33	<a href="#">Genome</a>	28	<a href="#">EcoProDB</a>
57	<a href="#">PubChem</a>	43	<a href="#">OMIM</a>	33	<a href="#">gpmDB</a>	28	<a href="#">RNACentral</a>
57	<a href="#">PubChem BioAssay</a>	43	<a href="#">OMIM (1)</a>	32	<a href="#">Ebolavirus</a>	28	<a href="#">ThaleMine</a>
57	<a href="#">PubChem Compound</a>	42	<a href="#">MetaCyc</a>	32	<a href="#">HOGENOM</a>	28	<a href="#">YeastMine</a>
57	<a href="#">PubChem Substance</a>	40	<a href="#">Guide to Pharmacology</a>	32	<a href="#">STITCH</a>	27	<a href="#">BioSystems</a>
		40	<a href="#">MalaCards</a>	31	<a href="#">eNet</a>	27	<a href="#">OrthoDB</a>
				31	<a href="#">HPIDB</a>		

# Databases with big attraction number (AN)

565 PubMed	76 PROSITE	54 COGs	38 CATH	16
501 UniProtKB	71 CAS	53 Genome	38 MetaCyc	
275 NCBI taxonomy	71 IntAct	51 DIP	36 CDD	
255 RCSB PDB	70 Reactome	51 STRING	36 SUPERFAMILY	
239 Genbank	69 ChEBI	50 BioCyc	35 dbSNP	
229 Gene	63 FlyBase	50 ENZYME	35 EcoCyc	
199 KEGG	61 MEDLINE	49 KEGG PATHWAY	34 ChEMBL	
193 RefSeq	61 UniGene	48 HPRD	34 PDBsum	
187 EC	59 BioGRID	48 WormBase	33 PLOS One	
182 Pfam	57 PubChem	47 BioProject	33 PRINTS	
160 InterPro	56 GEO	47 BRENDA	31 Homologene	
157 Protein	56 NCBI	47 DDBJ	31 ProDom	
152 Ensembl	56 SMART	47 TAIR	31 PubChem	
125 Nucleotide	55 DrugBank	47 TIGRFAMS	Compound	
109 OMIM	55 MGI	46 MINT	31 wwPDB	
99 SGD	55 PIR	45 MeSH	30 PDBe	
80 ENA	55 PMC	43 PANTHER	30 UCSC Archaeal	
80 HGNC	55 SCOP	40 GeneCards	Genome Browser	

# Integration candidates

Producer	Databases	Total attraction	% of maximal
1 NCBI	70	2909	33
2 EMBL-EBI	97	1209	14
3 SIB	37	762	9
4 Kioto University	19	348	4
5 Instute Paster	18	148	2
6 BioCyc	9378	133	2
7 InterMine	16	20	0
1+2+3+4	133	5228	59
1+3+4	92	4019	45
2+3+4	78	2319	26
Total	1115	8870	100

# MICRO-IS database partners in efficient solution

ArrayExpress, Assembly, BioModels, BioProject, BioSample, BioSamples, BioSystems, Bookshelf, CDD, Cellosaurus, ChEBI, ChEMBL, COGs, CSA, dbEST, dbProbe, dbSNP, Dengue virus database, DNATraffic, DrugPort, e!Ensembl, e!Ensembl Saccharomyces cerevisiae, e!EnsemblBacteria, e!EnsemblFungi, e!EnsemblGenomes, e!EnsemblProtists, EMBL, EMBL-EBI, EMDB, ENA, Ensembl, ENZYME, Enzyme Structures, EPD, EVA, Expression Atlas, Genbank, Gene, GeneDB, Genetic Codes, Genome, GEO, GEO DataSets, GEO Profiles, GSS, HAMAP, Hits, HIV-1, Homologene, IMEx, Influenza Virus Resource, IntAct, InterPro, KEGG, KEGG BRITE, KEGG DISEASE, KEGG GENES, KEGG GENOME, KEGG GLYCAN, KEGG LIGAND, KEGG MEDICUS, KEGG MODULE, KEGG Organisms, KEGG ORTHOLOGY, KEGG PATHWAY, MACiE, MapViewer, MedGen, MEDLINE, MEROPS, MeSH, MetaboLights, MIAPEGelDB, MMDB, MTBLS, NCBI, NCBI taxonomy, NCBI Trace Archives, neXtProt, NLM Catalog, Nucleotide, OMA, OMIM, OpenFlu, Organelle genomes, PathComp, PathPred, PathSearch, PaxDB, PDBe, PDBe EM Resources, PDBsum, Pfam, PICR, PMC, PMP, PomBase, PopSet, PRIDE, Probe, PROSITE, Protein, Protein Clusters, Protein Spotlight, Proteomes, PubChem, PubChem BioAssay, PubChem Compound, PubChem Substance, PubMed, PubMed Health, Reactome, RefSeq, RefSeqGene, Retroviruses, Rfam, Rhea, RNAcentral, SPARCLE, SpliceInfo, SRA, Structure, SugarBind, SWISS-2DPAGE, SWISS-MODEL, SwissVar, UniGene, UniProt-GOA, UniProtKB, UniRef, Viral genomes, ViralZone, Virus Variation

# Task 1a: Strains algorithm

The image shows two side-by-side search results pages. On the left is the NCBI Probe search results page for the query "Probe (Penicillium) AND 'Penicillium bialowiezense'". It displays 15 results, with the first one highlighted and circled in orange. The result is for a microarray element probe named Pen\_COX1\_11g, with accession ID Pr010275422. The target organism is listed as *Penicillium bialowiezense*. The right side shows the StrainInfo search results page for the query "taxon = 'Penicillium bialowiezense'". It displays 4 search results, each showing the species name and strain numbers. The first three results are for *Penicillium bialowiezense*, and the fourth is for *Penicillium brevicompactum*. A red circle highlights the first result in the StrainInfo list, which corresponds to the circled result in the NCBI list. Below the search results, there are four boxes containing WDCM codes and their respective collections:

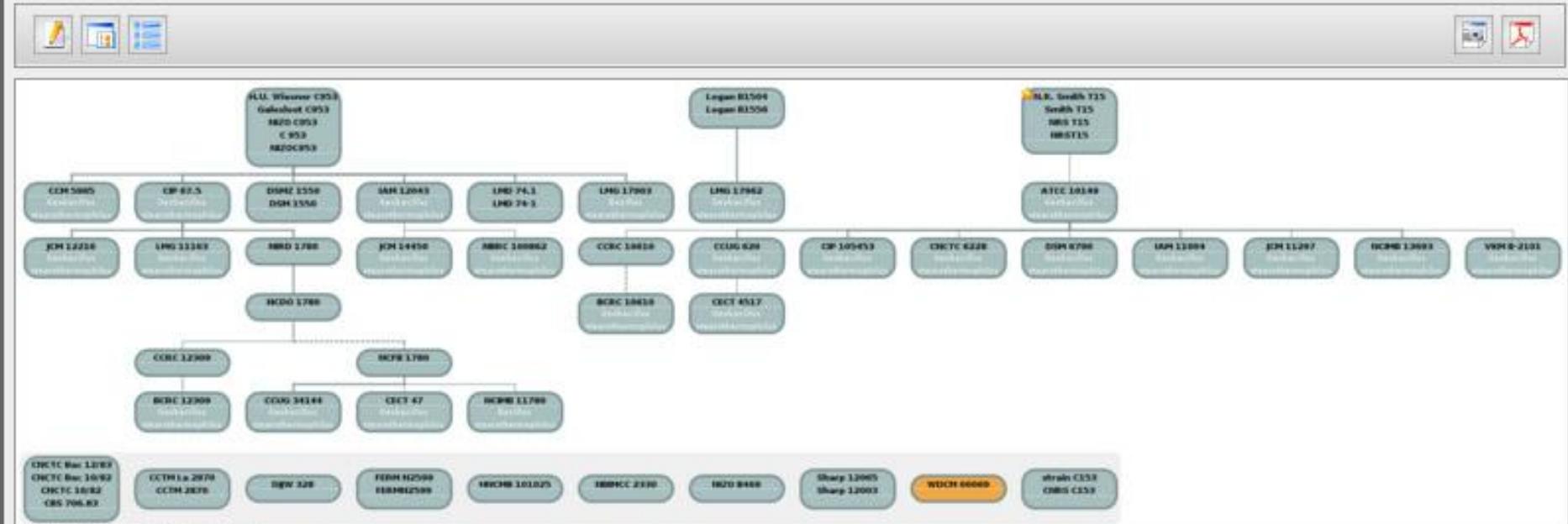
WDCM 133 Centraalbureau voor Schimmelcultures Filamentous fungi and Yeast Collection, Netherlands	WDCM 18, Food, Science, Australia, Ryde	WDCM 758, IBT, Culture Collection of Fungi, Denmark	WDCM 214, CABI, Genetic Resource Collection, UK
---	---	---	---

# StrainInfo strain exchange tree

## Strain Passport

**WDCM 00069 *Geobacillus stearothermophilus***

history



bioRxiv preprint doi: https://doi.org/10.1101/2011.04.12.250001; this version posted April 12, 2011. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-NC-ND 4.0 International license](#).

This Histri was manually curated (last saved on 2011-04-12).

**CCM 5965**  
*Geobacillus*  
*stearothermophilus*

**CIP 67.5**  
*Geobacillus*  
*stearothermophilus*

**DSMZ 1550**  
**DSM 1550**

**IAM 12043**  
*Geobacillus*  
*stearothermophilus*

# Solutions: task 1a: Name processing \*

## *Penicillium cyaneofulvum*

Summary:

*Penicillium cyaneofulvum* Biourge, La Cellule 33: 174 (1923)

Synonymy:

=*Penicillium brunneorubrum* Dierckx, Annales de la Société Scientifique de Bruxelles 25 (1): 88 (1901)

=*Penicillium griseoroseum* Dierckx, Annales de la Société Scientifique de Bruxelles 25 (1): 86 (1901)

=*Penicillium chrysogenum* Thom, U.S.D.A. Bureau of Animal Industry Bulletin 118: 58 (1910)

=*Penicillium baculatum* Westling, Svensk Botanisk Tidskrift 4: 139 (1910)

=*Penicillium notatum* Westling, Arkiv før Botanik 11 (1): 95 (1911)

...

=*Penicillium fluorescens* Laxa, Zentralblatt für Bakteriologie und Parasitenkunde Abteilung 2 86 (5-7):160-165 (1932)

=*Penicillium camerunense* R. Heim, Bull. Acad. R. Belg. Cl. Sci.: 42 (1949)

=*Penicillium aromaticum* f. *microsporum* Romankova, Uchenn. Zap. Leningr. Univ. Zhadanov: 102 (1955)

=*Penicillium harmonense* Baghd., Novosti Sistemmatiki Nizshikh Rastenii 5: 102 (1968)

Current name:

*Penicillium chrysogenum* Thom, U.S.D.A. Bureau of Animal Industry Bulletin 118: 58 (1910)

Classification:

Fungi, Ascomycota, Pezizomycotina, Eurotiomycetes, Eurotiomycetidae, Eurotiales, Trichocomaceae, *Penicillium*

Facultative or heterotypic synonyms:

1. *Penicillium aromaticum* f. *microsporum* Romankova, Uchenn. Zap. Leningr. Univ. Zhadanov: 102 (1955)
2. *Penicillium baculatum* Westling, Svensk Botanisk Tidskrift 4: 139 (1910)
3. *Penicillium brunneorubrum* Dierckx, Annales de la Société Scientifique de Bruxelles 25 (1): 88 (1901)
4. *Penicillium camerunense* R. Heim, Bull. Acad. R. Belg. Cl. Sci.: 42 (1949)
5. *Penicillium chlorophaeum* Biourge, La Cellule 33: 271 (1923)
6. *Penicillium chrysogenum* Thom, U.S.D.A. Bureau of Animal Industry Bulletin 118: 58 (1910)

...

# TAXONOMY IN MICROBIAL DATABASES

Taxonomy	*	**
NCBI	115	267
GBIF	16	16
IF	7	9
COL	6	6
LPSN	2	2
Mycobank	2	5

\* References in databases with taxonomical data

\*\* In all the databases with microbial data

229 - Total number of databases with taxonomical data

94 - No reference to specific taxonomy database

# Task 2: solution tools

Producer	Interface	
NCBI	Entrez Programming Utilities (E-Utils)	
EMBL-EBI	RESTful Web Services interface, Semantic WEB, RDF, SPARQL endpoint	
SIB	RESTful Web Services interface, Semantic WEB	
Kioto University	KEGG API, LinkDB	

The diagram illustrates the mBRC tasks in BMB information structures, specifically focusing on MIRRI-IS specific areas of activities. It highlights several key components:

- mBRC tasks in BMB information structures**: A box containing a flowchart of tasks (T1-T4) and their relationships.
- Possible mBRC contributions for Bio-Medicine**: A list of three items: 1. Resolving nomenclatural issues related to microorganisms, 2. Analysis and assurance of reproducibility of experimental data, 3. Strain-specific characters.
- MIRRI-IS specific areas of activities**: A central section listing "mBRC regular tasks" and a task titled "Content search for microorganisms based on their properties".
- BMB+ tools & structures of particular interest for MIRRI**: A box listing internal solutions of MIRRI-IS, BMB ontologies, RDF, N3, SPARQL, BMB databases, and Bio-Medicine-related resources like MOGENIS, BMB ontologies, RDF, N3, SPARQL, MIRIAM Registry, MMRB, bioSharing.org, identifiers.org, EDAM, REST API, Tools registry, Standards registry, and Bio-industry, agriculture, ...
- Specific mBRC/MIRRI-IS future contributions to BMB+ tools & structures**: A list of six items: (1) mNames ontology, (2) mBRC catalogues, StrainInfo, WDCM in MIRRIAM data collections list and in bioSharing.org, (3) mBRC catalogues in Metabase (MB), (4) MIRRI-IS in Mammuth LCO Cloud Diagram, (5) Metabio vocabulary, (6) microbial phenotypic input into BMB (3-1) Online Dictionary of control identifiers.
- MIRRI-IS Data standard**: A circular diagram showing the relationship between MIRRI-IS and various data standards like mRDS, mRDS2, mRDS3, mRDS4, mRDS5, mRDS6, mRDS7, mRDS8, mRDS9, mRDS10, mRDS11, and Euro-BioImaging, BBMRI, EBI-PRIDE, PRIDEGREEN, ECRIN, AnIEET, ERINATCH, ERINAWA, Life science, BMB, INFRAFRONTIER, INSTRUCT.
- BMB infrastructures**: A box listing various BMB infrastructure projects.

Logos for the European Union, H2020, and the MIRRI project are present at the bottom left. Social media links and a QR code are at the bottom right.

# Thank you