Towards a wiki system for human and animal cell lines

Dan Bolser & Paolo Romano

National Cancer Research Institute,
Genova, Italy
(paolo.romano@istge.it, skype: p.romano)

Outline

- Animal cell lines: aims and availability
- Wiki systems and Biological wikis
- Cell Lines Data Base and HyperCLDB
- Towards the CellLinesWiki

Cell lines information

- Animal cell lines are useful resources for research, representing an optimal model for many assays
- It is essential that cell lines are properly stored, characterized, maintained, distributed and used
- Biological Resources Centers offer an adequate infrastructure for these aims
- Their catalogues, including high-quality characterization data, are available on-line
- But…

Cell lines information

- But…
  a great wealth of knowledge, information, and data, is owned by researchers throughout the world:
  - developing new cell lines,
  - investigating cell lines properties,
  - using cell lines in their own experiments.
- And…
  this information is rarely published in literature and it usually gets lost.
- So… we need to find a way to capture it

Wiki systems

- Wiki systems are effective tools for the collaborative development of documentation, from small texts to encyclopedic sites
- Mediawiki, that is used for WikiPedia, is one of the most frequently used software for implementation of wiki systems
- It is a modular open source, many extensions are available, new features may be easily added
- Specific aims of wikis in biology can be envisaged
  - Development and sharing of documents
  - Annotation and extension of database contents

Annotation of db contents

- Motivation
  - maintenance of database contents is a heavy task, that also requires specialist expertise
- Objectives
  - to allow experts to contribute with their knowledge to the improvement of contents
  - to gather annotations and further information, e.g. experimental data
  - to point out inconsistencies and errors
- Database contents are preserved from uncontrolled changes

Many biological Wikis exist

- Gene Wiki
  - Specialized subsection on human genes in Wikipedia
  - Objective: build a high quality page for each human gene
  - http://en.wikipedia.org/wiki/Gene_Wiki
- WikiGenes
  - Able to manage and recognize contributors
  - Users' pages, rating by peers, …
  - http://www.wikigenes.org/
- WikiPathways
  - Community annotation of some pathway databases
  - Pathway editor available
  - http://www.wikipathways.org/

Cell Line Data Base (CLDB)

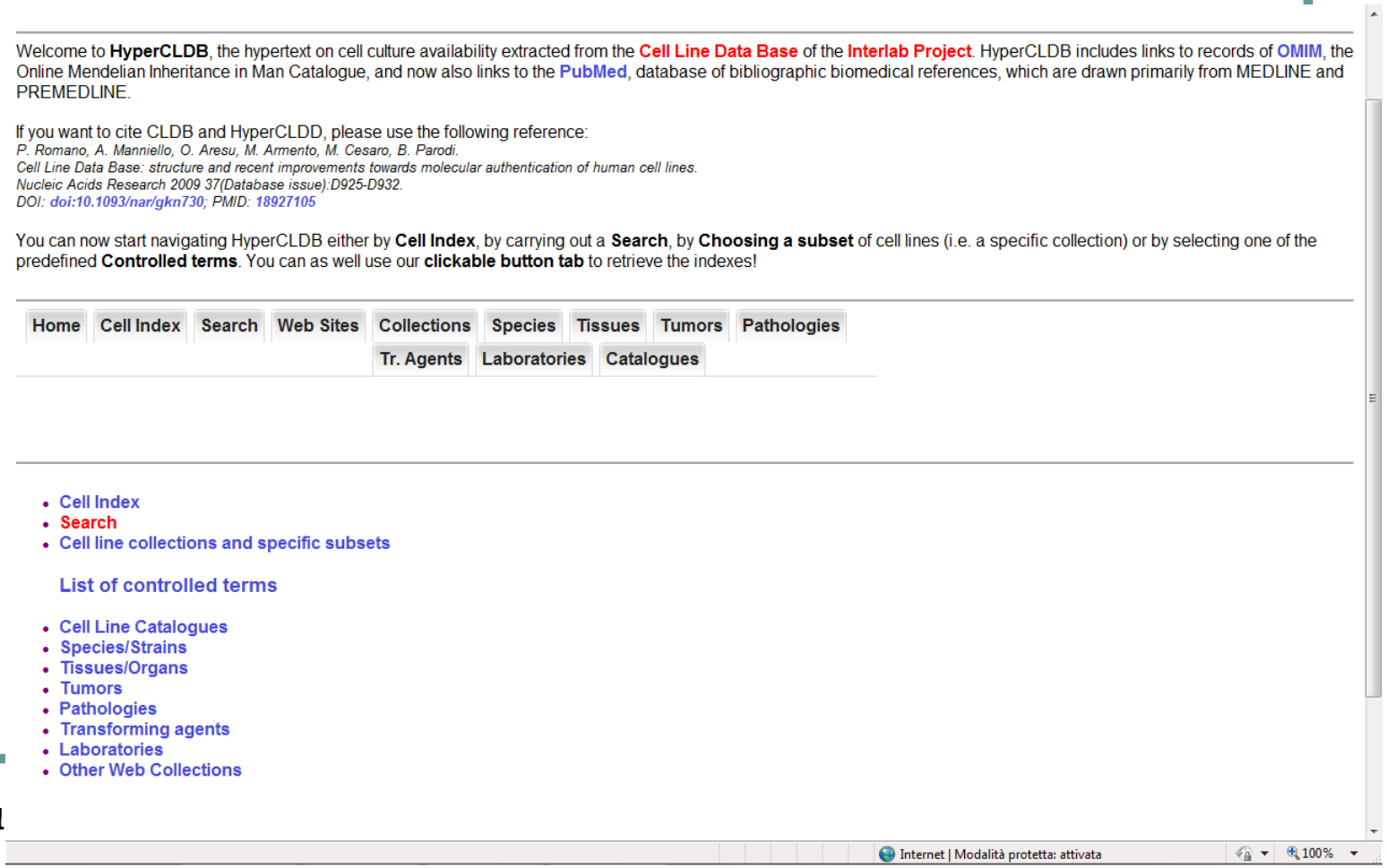
- Human and animal cell lines from 82 Italian laboratories and European collections
- The most relevant data refer to:
  - Identification (name, typology, karyology, morphology,...)
  - Origin (tissue, species, tumour, pathology,...)
  - Properties (known functions, products and applications)
  - Preservation and culture characteristics
  - Retrieval sources (bibliographic references, catalogue codes,...)
  - Quality control (identity and sterility)
  - Limited characterization (immunological profile, cytogenetic analysis...)

CLDB contents

- Total cell lines: 6,625
- 4,920 human cell lines of which
  - 1,994 from 299 pathologies,
  - 992 from 130 tumors, and
  - 193 transformed
- 1,380 animal cell lines, from 203 species, of which
  - 519 from 80 tumors
  - 537 transformed

HyperCLDB

HyperCLDB is a static hypertext allowing for the navigation of CLDB data



Welcome to **HyperCLDB**, the hypertext on cell culture availability extracted from the **Cell Line Data Base** of the **Interlab Project**. HyperCLDB includes links to records of **OMIM**, the Online Mendelian Inheritance in Man Catalogue, and now also links to the **PubMed**, database of bibliographic biomedical references, which are drawn primarily from MEDLINE and PREMEDLINE.

If you want to cite CLDB and HyperCLDD, please use the following reference:
P. Romano, A. Manniello, O. Aresu, M. Armento, M. Cesaro, B. Parodi.
Cell Line Data Base: structure and recent improvements towards molecular authentication of human cell lines.
Nucleic Acids Research 2009 37(Database issue):D925-D932.
DOI: doi:10.1093/nar/gkn730; PMID: 18927105

You can now start navigating HyperCLDB either by **Cell Index**, by carrying out a **Search**, by **Choosing a subset** of cell lines (i.e. a specific collection) or by selecting one of the predefined **Controlled terms**. You can as well use our **clickable button tab** to retrieve the indexes!

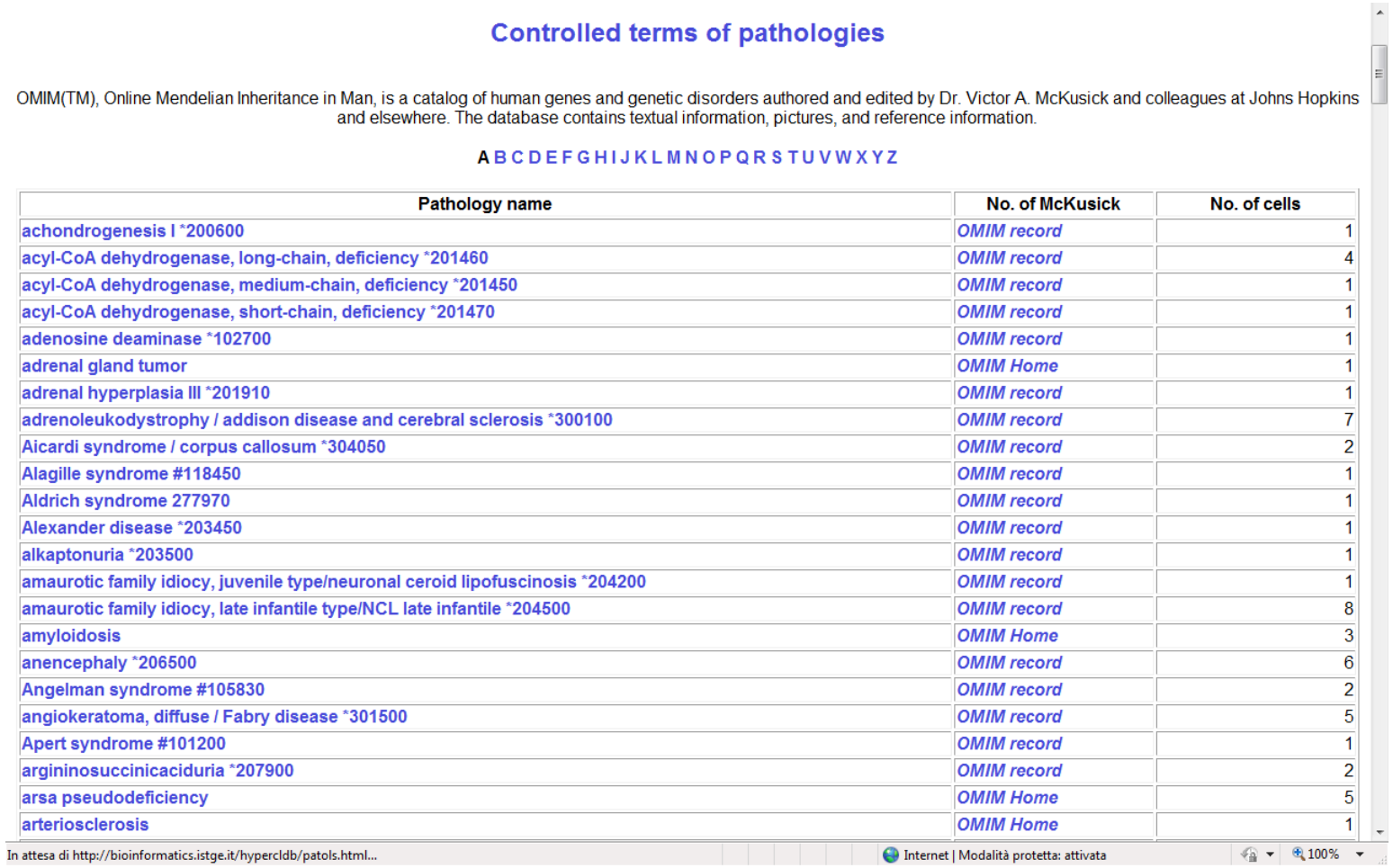
Home | Cell Index | Search | Web Sites | Collections | Species | Tissues | Tumors | Pathologies | Tr. Agents | Laboratories | Catalogues

- Cell Index
- Search
- Cell line collections and specific subsets

List of controlled terms

- Cell Line Catalogues
- Species/Strains
- Tissues/Organs
- Tumors
- Pathologies
- Transforming agents
- Laboratories
- Other Web Collections

HyperCLDB



Controlled terms of pathologies

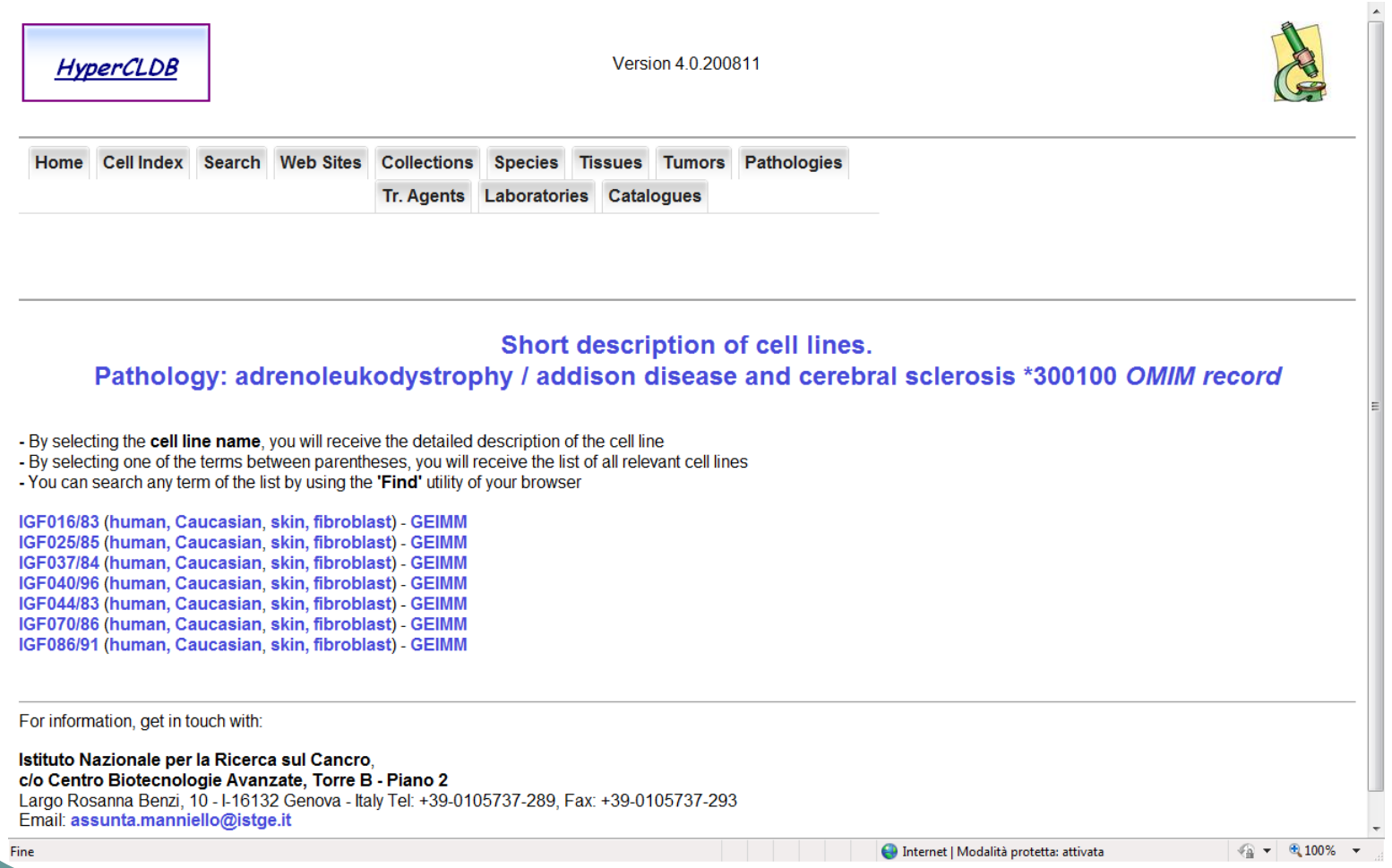
OMIM(TM), Online Mendelian Inheritance in Man, is a catalog of human genes and genetic disorders authored and edited by Dr. Victor A. McKusick and colleagues at Johns Hopkins and elsewhere. The database contains textual information, pictures, and reference information.

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Pathology name	No. of McKusick	No. of cells
achondrogenesis I *200600	OMIM record	1
acyl-CoA dehydrogenase, long-chain, deficiency *201460	OMIM record	4
acyl-CoA dehydrogenase, medium-chain, deficiency *201450	OMIM record	1
acyl-CoA dehydrogenase, short-chain, deficiency *201470	OMIM record	1
adenosine deaminase *102700	OMIM record	1
adrenal gland tumor	OMIM Home	1
adrenal hyperplasia III *201910	OMIM record	1
adrenoleukodystrophy / addison disease and cerebral sclerosis *300100	OMIM record	7
Aicardi syndrome / corpus callosum *304050	OMIM record	2
Alagille syndrome #118450	OMIM record	1
Aldrich syndrome 277970	OMIM record	1
Alexander disease *203450	OMIM record	1
alkaptonuria *203500	OMIM record	1
amaurotic family idiocy, juvenile type/neuronal ceroid lipofuscinosis *204200	OMIM record	1
amaurotic family idiocy, late infantile type/NCL late infantile *204500	OMIM record	8
amyloidosis	OMIM Home	3
anencephaly *206500	OMIM record	6
Angelman syndrome #105830	OMIM record	2
angiokeratoma, diffuse / Fabry disease *301500	OMIM record	5
Apert syndrome #101200	OMIM record	1
argininosuccinicaciduria *207900	OMIM record	2
arsa pseudodeficiency	OMIM Home	5
arteriosclerosis	OMIM Home	1

In attesa di http://bioinformatics.istge.it/hypercldb/patols.html... Internet | Modalità protetta: attivata 100%

HyperCLDB


HyperCLDB
Version 4.0.200811
Home
Cell Index
Search
Web Sites
Collections
Species
Tissues
Tumors
Pathologies
Tr. Agents
Laboratories
Catalogues
Short description of cell lines.
Pathology: adrenoleukodystrophy / addison disease and cerebral sclerosis *300100 OMIM record
- By selecting the cell line name, you will receive the detailed description of the cell line
- By selecting one of the terms between parentheses, you will receive the list of all relevant cell lines
- You can search any term of the list by using the 'Find' utility of your browser
IGF016/83 (human, Caucasian, skin, fibroblast) - GEIMM
IGF025/85 (human, Caucasian, skin, fibroblast) - GEIMM
IGF037/84 (human, Caucasian, skin, fibroblast) - GEIMM
IGF040/96 (human, Caucasian, skin, fibroblast) - GEIMM
IGF044/83 (human, Caucasian, skin, fibroblast) - GEIMM
IGF070/86 (human, Caucasian, skin, fibroblast) - GEIMM
IGF086/91 (human, Caucasian, skin, fibroblast) - GEIMM
For information, get in touch with:
Istituto Nazionale per la Ricerca sul Cancro,
c/o Centro Biotecnologie Avanzate, Torre B - Piano 2
Largo Rosanna Benzi, 10 - I-16132 Genova - Italy Tel: +39-0105737-289, Fax: +39-0105737-293
Email: assunta.manniello@istge.it
Fine
Internet | Modalità protetta: attivata
100%

HyperCLDB

This is the complete description of the cell line that you requested. By selecting a generic term (e.g. SPECIES) you will go to the corresponding list of controlled terms, while by selecting a specific term you will go to the list of corresponding cell line names.

IGF037/84 (human, Caucasian, skin, fibroblast, adrenoleukodystrophy / addison disease and cerebral sclerosis)

Primary culture, grown as monolayer
fibroblast morphology
SPECIES: **human, Caucasian** (male)
TISSUE/ORGAN: **skin, fibroblast**
PATHOLOGY: **adrenoleukodystrophy / addison disease and cerebral sclerosis***300100 ***(OMIM record)***
DEPOSITOR: Laboratorio Diagnosi pre-postnatale delle malattie metaboliche, Istituto Giannina Gaslini, Genova, Italy
PROPERTIES: biochemical study; genetic study; metabolism study

Available in the following **LABORATORY**:

- **Istituto Giannina Gaslini** (GEIMM, Genova)
  RPMI 1640 + 15-18% FBS + 1% Antibiotics + 2g/l NaHCO3 + 2mM L-Glutamine; Freezing medium: Culture medium + 25% FBS + 10% Glycerol; 37C, 5% CO2
  Cell line belongs to risk category Group 1
  Affected hemizygous; Biochemical diagnosis

Bibliographic references:

- Exc Med Inter Congr 1989;127

For information, get in touch with:

Istituto Nazionale per la Ricerca sul Cancro,
c/o Centro Biotecnologie Avanzate, Torre B - Piano 2
Largo Rosanna Benzi, 10 - I-16132 Genova - Italy Tel: +39-0105737-289, Fax: +39-0105737-293
Email: **assunta.manniello@istge.it**

Fine | Internet | Modalità protetta: attivata | 100%

CellLinesWiki

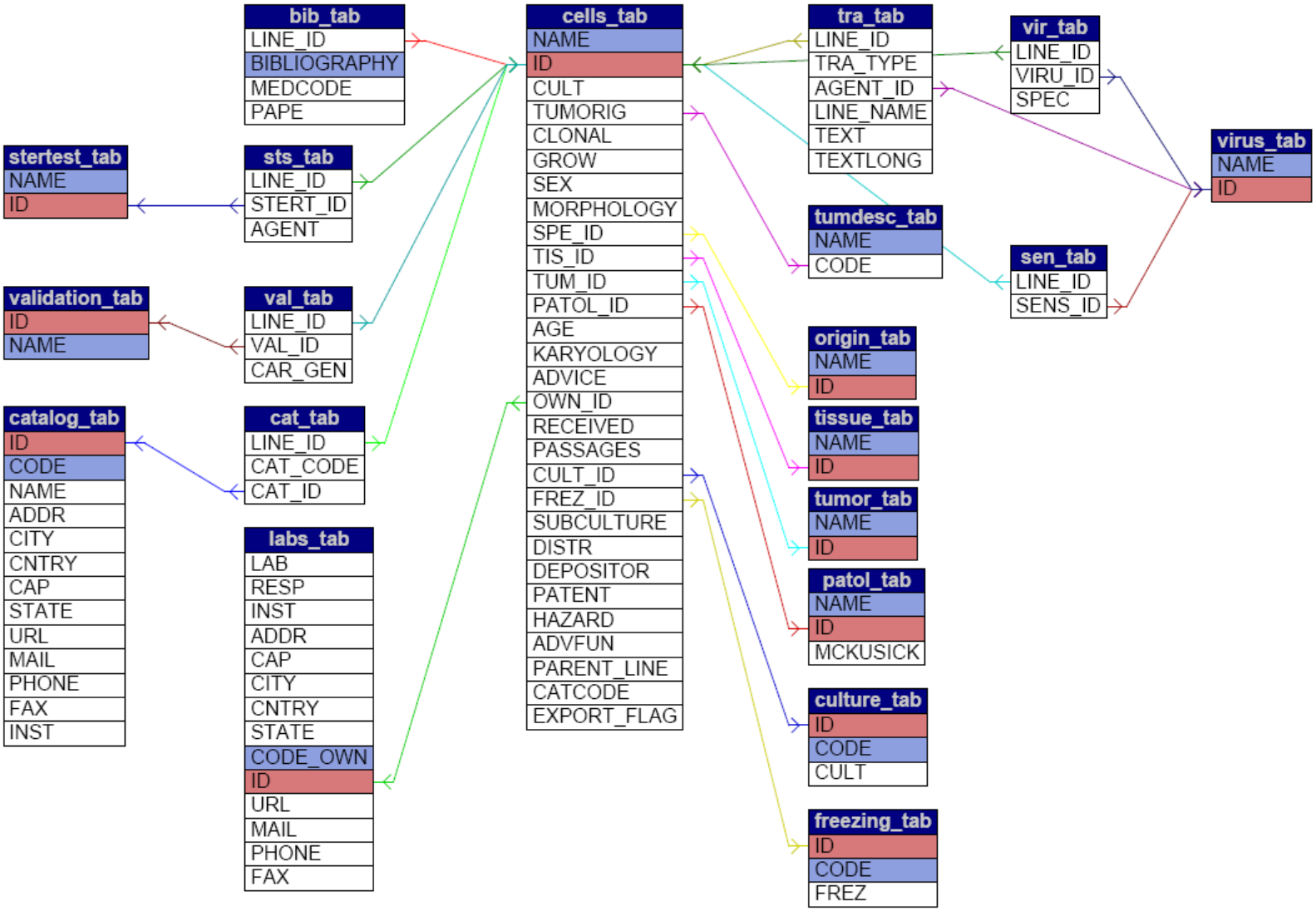
- The CellLinesWiki is a knowledge base on cell lines for Biological Resources Centers, collections and researchers (under development)

- CellLinesWiki consists of three data layers:
  - database authoritative information (from CLDB)
  - a curated set of extra data, maintained by a limited number of carefully nominated experts in the field
  - unstructured information provided by end users, authenticated, but not trusted at the above level

CellLinesWiki

- Using **MW** and some **extensions** we are designing a system whereby the data in CLDB can be browsed, queried and annotated
  - MediaWiki was chosen as the starting point for the development of CellLinesWiki
  - the Semantic MediaWiki (SMW) extension was also used as we want to store structured data
- How to convert the database schema in the wiki?

CLDB schema


bib_tab
LINE_ID
BIBLIOGRAPHY
MEDCODE
PAPE
stertest_tab
NAME
ID
sts_tab
LINE_ID
STERT_ID
AGENT
validation_tab
ID
NAME
val_tab
LINE_ID
VAL_ID
CAR_GEN
catalog_tab
ID
CODE
NAME
ADDR
CITY
CNTRY
CAP
STATE
URL
MAIL
PHONE
FAX
INST
cat_tab
LINE_ID
CAT_CODE
CAT_ID
labs_tab
LAB
RESP
INST
ADDR
CAP
CITY
CNTRY
STATE
CODE_OWN
ID
URL
MAIL
PHONE
FAX
cells_tab
NAME
ID
CULT
TUMORIG
CLONAL
GROW
SEX
MORPHOLOGY
SPE_ID
TIS_ID
TUM_ID
PATOL_ID
AGE
KARYOLOGY
ADVICE
OWN_ID
RECEIVED
PASSAGES
CULT_ID
FREZ_ID
SUBCULTURE
DISTR
DEPOSITOR
PATENT
HAZARD
ADVFUN
PARENT_LINE
CATCODE
EXPORT_FLAG
tra_tab
LINE_ID
TRA_TYPE
AGENT_ID
LINE_NAME
TEXT
TEXTLONG
vir_tab
LINE_ID
VIRU_ID
SPEC
virus_tab
NAME
ID
tumdesc_tab
NAME
CODE
sen_tab
LINE_ID
SENS_ID
origin_tab
NAME
ID
tissue_tab
NAME
ID
tumor_tab
NAME
ID
patol_tab
NAME
ID
MCKUSICK
culture_tab
ID
CODE
CULT
freezing_tab
ID
CODE
FREZ

Modeling the RDB in the wiki

- RDB tables are implemented in the wiki as SMW 'classes', which are defined by MW categories.
- A page is an instance of the given class.
- It collects together a set of properties (one for each column in the table).
- Each page represents a row of data in the table.

- The 'form inputs' adapt to the type of property.
- Auto-completion allows easy input.
- Forms are created using special pages.
- Every data may be easily queried

e.g. The 'Laboratory' class

labs_tab
LAB
RESP
INST
ADDR
CAP
CITY
CNTRY
STATE
CODE_OWN
ID
URL
MAIL
PHONE
FAX

- Each 'laboratory' page carries a template to format and store the associated values

{{Template:Laboratory
| laboratory = Cancerologia
| institute = Istituto di Cancerologia
| responsible person = Sandro Grilli
| ...

Cell Lines Wiki is a prototype system devoted to information on cell lines. Currently, there are 6632 Cell lines in the wiki, from Catalogs with 82 associated Laboratories.

Search for a lab:[show]

About

[edit]

The wiki will include three distinct sections:

- HyperCLDB is a category for information from the Cell Line Data Base
- Talk, that is the usual discussion section, will be devoted to an open discussion on cell lines included in the Cell Line Data Base
- Annotation will finally be devoted to scientific annotations provided by users of the cell lines and related to new data and uses and to the various experiences

It currently contains 11,212 pages, edited by its users.

CellLinesWiki is brought to you by IST Bioinformatics group!

Cell line query

Name:

Species or strain:

Tissues or organs:

Tumors:

Pathology:

Search

Navigation

Main page
Recent changes
Random page
Help

index pages

Cell Collections and Banks
Species and Strains
Tissues and Organs
Tumors
Pathologies
Transforming Agents
Laboratories
Cell line catalogues

HyperCLDB

Home Page

Toolbox

What links here
Related changes
Upload file
Special pages
Printable version
Permanent link

Cell Lines Wiki is a prototype system devoted to information on cell lines. Currently, there are 6632 Cell lines in the wiki, from Catalogs with 82 associated Laboratories.

Search for a lab:[show]

About

[edit]

The wiki will include three distinct sections:

- HyperCLDB is a category for information from the Cell Line Data Base
- Talk, that is the usual discussion section, will be devoted to an open discussion on cell lines included in the Cell Line Data Base
- Annotation will finally be devoted to scientific annotations provided by users of the cell lines and related to new data and uses and to the various experiences

It currently contains 11,212 pages, edited by its users.

CellLinesWiki is brought to you by IST Bioinformatics group!

Cell line query

Name:

HeL

HEL
HEL 12469
HEL 299
HEL 92.1.7
HELA-S3
HeLA
HeLa

Search

Navigation

Main page
Recent changes
Random page
Help

index pages

Cell Collections and Banks
Species and Strains
Tissues and Organs
Tumors
Pathologies
Transforming Agents
Laboratories
Cell line catalogues

HyperCLDB

Home Page

Toolbox

What links here
Related changes
Upload file
Special pages
Printable version
Permanent link

This page was last modified on 5 October 2011, at 14:12.

This page has been accessed 320 times.

Run query: Cell line/Query

Found 2 cell lines.

Line name	Lab.	Cat.	Origin	Tissue	Tumor	Date
HeLa 382	GEIBM		Human, Black	Cervix	Carcinoma	
HeLa 382	ICLC	HTL95015	Human, Black	Cervix	Carcinoma, Epitheloid	30 August 1995

Additional query

Cell line query

Name: HeLa 382

Species or strain:

Tissues or organs:

Tumors:

Pathology:

Search

Navigation

Main page
Recent changes
Random page
Help

index pages

Cell Collections and Banks
Species and Strains
Tissues and Organs
Tumors
Pathologies
Transforming Agents
Laboratories
Cell line catalogues

HyperCLDB

Home Page

Toolbox

Upload file
Special pages

CLDB1607 - HeLa 382

Navigation

Main page
Recent changes
Random page
Help

index pages

Cell Collections and Banks
Species and Strains
Tissues and Organs
Tumors
Pathologies
Transforming Agents
Laboratories
Cell line catalogues

HyperCLDB

Home Page

Toolbox

What links here
Related changes
Upload file
Special pages
Printable version
Permanent link

Advice

Quality control and characterisation performed at ICLC

How to subculture

37C, 5% CO2

Lab	Contact	Inst	City	Country
Servizio di Biologia Molecolare	Giovanni Levi		Genova	Italy

Code	Culture conditions
CC:DM09	DMEM + 10% FBS + 2mM L-Glutamine

Code	Freezing conditions
FC:AM01	90% FBS + 10% DMSO

Transformation	
Transformation type	Transfection
Parent cell line (name)	HeLa
Summary	see advices

Identification data	
Laboratory code	GEIBM
Culture type	Continuous
Clonality	
Growth type	Monolayer
Tumorigenicity	
Morphology	epithelial-like
Origin data	
Origin	human, Black
Sex	F
Age	31 years old
Tissue type	cervix
Tumor type	carcinoma
Pathology	
Parent line	
Karyology	
Culture data	
Received date	
Passages	
Patent flag	

Category: Cell line

Edit Cell line: CLDB1607 - HeLa 382

Identification data | Origin data | Culture data | Transformation data

Laboratory code: GEIBM
Catalog code:
Culture type: Continuous
Clonality:
Growth type: Monolayer
Tumorigenicity:
Morphology: epithelial-like

Free text:

Edit summary:

This is a minor edit Watch this page

Navigation
Main page
Recent changes
Random page
Help

index pages
Cell Collections and Banks
Species and Strains
Tissues and Organs
Tumors
Pathologies
Transforming Agents
Laboratories
Cell line catalogues

HyperCLDB
Home Page

Toolbox
What links here
Related changes
Upload file
Special pages
Browse properties

Categories

The following categories contain pages or media. Unused categories are not shown here. Also see wanted categories.

Categories

Display categories starting at: Go

- Car (1 member)
- Cell line (6,593 members)
- Culture conditions (559 members)
- Freezing conditions (67 members)
- Laboratory (22 members)
- Origin (71 members)
- Pathology (331 members)
- Reference (2,009 members)
- Tissue type (73 members)
- Tumor type (77 members)
- Tumorigenicity (12 members)

Navigation

Main page
Recent changes
Random page
Help

index pages

Cell Collections and Banks
Species and Strains
Tissues and Organs
Tumors
Pathologies
Transforming Agents
Laboratories
Cell line catalogues

HyperCLDB

Home Page

Toolbox

Upload file
Special pages

Properties

The following properties are used in the wiki.

Showing below up to **50** results starting with **#1**.
View (previous 50 | next 50) (20 | 50 | 100 | 250 | 500)

1. *Covers property* of type Special:Types/ (10)
2. *Has display parameters* of type String (1)
3. *Has filter* of type Page (9)
4. *Requires filter* of type Page (2)
5. *Creates pages with form* of type Special:Types/ (8)
6. *Email* of type Email (54)
7. *URL* of type URL (24)
8. 46,X,dic dup(Y)(pter-q11 of type Page (1)
9. Address of type Text (82)
10. Cell line advice of type String (4126)
11. Cell line age of type String (1133)
12. Cell line catalog code of type String (4358)
13. Cell line clonality of type String (514)
14. Cell line culture type of type String (5179)
15. Cell line depositor of type String (2802)
16. Cell line distribution flag of type String (5328)
17. Cell line export flag of type String (6632)
18. Cell line function note of type String (3360)
19. Cell line growth type of type String (5190)
20. Cell line hazard of type String (2343)
21. Cell line id of type Number (6632)
22. Cell line karyology of type String (825)
23. Cell line morphology of type String (4857)
24. Cell line name of type String (6632)

Navigation

Main page
Recent changes
Random page
Help

index pages

Cell Collections and Banks
Species and Strains
Tissues and Organs
Tumors
Pathologies
Transforming Agents
Laboratories
Cell line catalogues

HyperCLDB

Home Page

Toolbox

Upload file
Special pages
Browse properties

Many thanks for the attention!