Integrating mutation data of the TP53 human gene in the bioinformatics network environment

Paolo Romano
Bioinformatics and Structural Proteomics,
National Cancer Research Institute, Genova
(paolo.romano@istge.it)

Outline

- Motivation
  - Complexity of data integration in biology
  - A possible methodology for flexible systems
- Background
  - Functions of the human TP53 gene
  - IARC TP53 Mutation Database
- Outline of our work
  - Implementation of an SRS site for TP53
  - Implementation of Web Services for TP53
  - Development of workflows
- Conclusions

Information in biology: known facts

Biomedical research produces an increasing quantity of **new data** and **new data types**

- Genomics is producing an immense quantity of data
- Emerging domains, like mutation and variation analysis, polymorphisms, metabolism, as well as new high-throughput technologies, e.g., microarrays, will also contribute with huge amounts of data
- Analysis software must interoperate with databases
  - Databases as input for software
  - Results as new data to store and analyze

Information in biology: some figures

EMBL Data Library 88 (Sep 2006)
- Sequences: 80,591,891 (Bases: 146,595,277,574)
- Increase: +8,86% (+8,91%) since previous release
- Increase: +37,16% (+36,297%) since Sep 2005
- (http://www3.ebi.ac.uk/Services/DBStats/)

ArrayExpress (31/10/2006)
- Experiments 1,739 (ca 140 Gb)
- Increase +100% from October 2005 to October 2006
- (http://www.ebi.ac.uk/arrayexpress/Help/stats/index.html)

Nucleic Acids Research Supplement (2006)
- 858 molecular biology databases
- (http://nar.oxfordjournals.org/cgi/content/full/33/suppl_1/D5/DC1)

SRS public sites (17/11/2006)
- 1,300 libraries
- (http://downloads.lionbio.co.uk/publisrs.html)

Heterogeneicity of databanks

- A few dbs are managed in a homogenous way (nucleotide sequences at EBI, NCBI, DDBJ)
- Secondary databases are of the highest quality (good and extended annotation, quality control)
- Many databases are highly specialized, e.g. by gene, organism, disease, mutation
- Many databases are created by small groups or even by single researchers

- Databanks are distributed:
  - Different DBMS, data structures, query methods
  - Same information, different syntax and semantics

Goals of the integration

In this context, data integration and process automation are needed to:

- Automatically carry out analyses and/or searches involving more databases and software
- Effectively perform analyses involving large data sets
- Achieve a better and wider view of all available information
- Carry out a real data mining and discover new data

This can be done by computers, but…

Data integration longevity

- Integration needs stability
  - Standardization……
  - Good domain knowledge
  - Clearly defined data
  - Clearly identified goals
- Integration fears:
  - Heterogeneity of data and systems
  - Uncertain domain knowledge
  - Fast evolution of data
  - Highly specialized data
  - Lacking of predefined, clear goals
  - Originality, experimentalism ("let me see if this works")

Integration of biological information

In biology:

- A pre-analysis and reorganization of information is very difficult, because knowledge and related data change very quickly
- Complexity of information makes it difficult to design data models which are valid for different domains and over time
- Goals and needs of researchers evolve very quickly according to new theories and discoveries

Integration must therefore be carried out by using flexible systems that are easy to adapt and extend

A possible methodology

A methodology based on data standardization:

- XML schemas for the creation of the models of the information
- XML based languages for data representation and storage
- Web Services for data exchange and interoperation between software
- Computerised workflows for the definition and execution of analysis processes

Semantic information should be added when possible

XML and Web Services in bioinformatics

Models are lacking, XML and Web Services are increasingly being used

XML dialects

Sequences (BSML, Agave)
Proteins (SPML)

NCBI outputs (BlastXML)

Microarray (MAGE-ML)

Systems Biology Markup Language (SBML)
Biological Variation Markup Language (BVML)

Web Services

EMBOSS, XEMBL, Interpro (EBI)
eUtils (NCBI)
caBIO (NCICB)
KEGG API

GeneCruiser, Biosphere (microarray)
SIMAP (proteins)

CABRI (biological resources)
TP53 mutations (gene mutations)

bioMOBY (directory)
Soaplab (tools)

Workflow

"***A computerized facilitation or automation of a business process, in whole or part***". (Workflow Management Coalition)

Its main goal is the implementation of data analysis processes in standardized environments

Its main advantages relate to:

- **effectiveness**: being an automatic procedure, it frees bio-scientists from repetitive interactions with the web and it supports good practice,
- **reproducibility**: analysis can be replicated over time,
- **reusability**: intermediate results can be reused,
- **traceability**: the workflow is carried out in a transparent analysis environment where data provenance can be checked and/or controlled.

Workflow management systems

Software	Type	XML	Avail	URL
Taverna Workbench	**Stand-alone**	**XScufl**	**Open source**	**http://taverna.sourceforge.net/**
Biopipe	Library	Pipeline XML	Open source	http://www.gmod.org/biopipe/
ProGenGrid	Stand-alone	NA	NA	http://datadog.unile.it/progen
DiscoveryNet	Stand-alone	DPML	Commercial	http://www.discovery-on-the.net/
Kepler	Stand-alone	MoML	Open source	http://kepler-project.org/
GPipe	Web Interface, local services	GPipe XML	Open source	http://if-web1.imb.uq.edu.au/Pise/5.a/gpipe.html
EGene	Stand-alone	NA	Open source	http://www.lbm.fmvz.usp.br/egene/
BioWMS	Web Interface, remote services	XPDL	Public use	http://litbio.unicam.it:8080/biowms/
Biowep	**Portal**	**XScufl XPDL**	**Open source**	**http://bioinformatics.istge.it/biowep/**
BioWBI	Web Interface, local services	Proprietary	Commercial	http://www.alphaworks.ibm.com/tech/biowbi
Pegasys	Stand-alone	Pegasys DAG	Open source	http://bioinformatics.ubc.ca/pegasys/
Wildfire	Stand-alone	GEL	Open source	http://wildfire.bii.a-star.edu.sg/wildfire/
Triana	Stand-alone	Triana WL	Open source	http://www.trianacode.org/
Pipeline Pilot	Stand-alone	Proprietary	Commercial	http://www.scitegic.com/
FreeFluo	Library	WSFL & XScufl	Open source	http://freefluo.sourceforge.net/
Biomake	Library	NA	Open source	http://skam.sourceforge.net/

Bartocci et al, Proc. BITS Conf., 2006

Outline of our work

Aim: validate the methodology for biological data integration

Work:

4. Select a database of interest in our domain (oncology)
   → IARC TP53 Mutation Database

6. Implement it in a suitable environment
   → Sequence Retrieval System - SRS

8. Develop Web Services for a programmatic access to the database
   → SoapLab + Tomcat, Axis, AppLab

10. Develop workflows that effectively make use of it
    → Taverna Workbench

Why Study TP53 Mutations?

- TP53 somatic mutations are found in **most types of sporadic human cancers** at various frequencies (**from 5% to 70%**)
- TP53 mutations may also be inherited in **families with a predisposition to multiple cancers**, as in the Li-Fraumeni syndrome (LFS).
- Over **20,000 mutations** have been reported in the literature, with more than 2000 different missense mutations.
- In several cancers, the nature of TP53 mutations and their distribution along the coding sequence have allowed the identification of **tumor-specific mutation spectra**, revealing clues on the mechanisms that might have caused the mutation.
- The presence of a TP53 mutation may be **predictive** of the tumor **response to treatment** and **patient survival**.

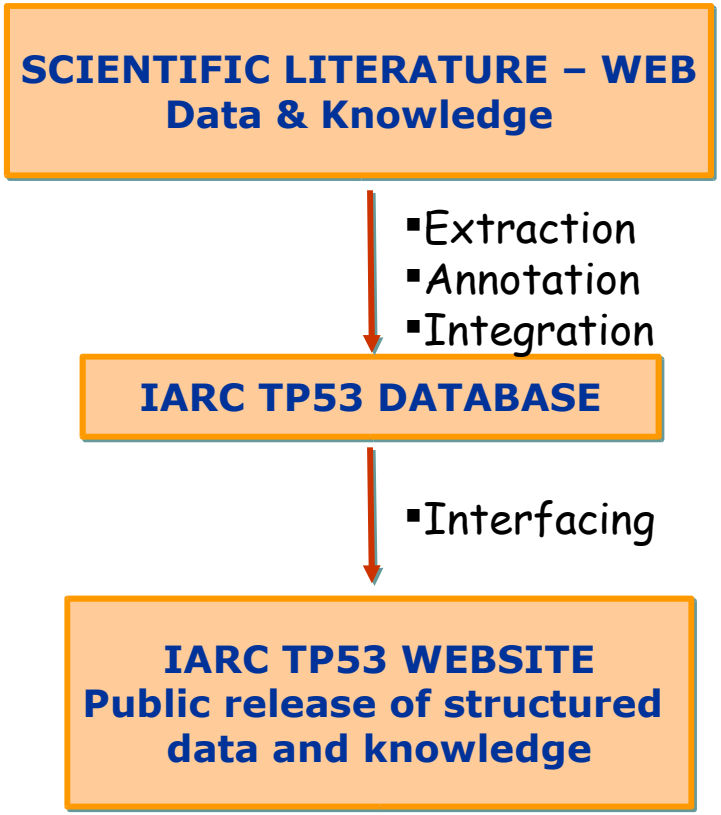
➡ TP53 mutations useful in:

- molecular epidemiology of cancer
- molecular genetics
- molecular pathology of cancer
- structural biology

P. Hainaut and M. Olivier. "TP53 Genetic Variations In Human Cancer"
slide show at http://www-p53.iarc.fr/

IARC TP53 Mutations Database

- Extract TP53 mutation data from publications
- Organize and annotate data into a format that allows easy retrieval and analysis
- Provide a web-based tool to analyse TP53 mutation patterns in cancers



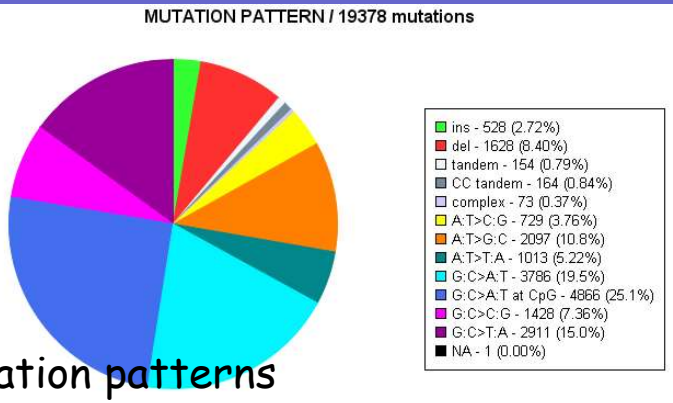
P. Hainaut and M. Olivier. "TP53 Genetic Variations In Human Cancer"
slide show at http://www-p53.iarc.fr/

IARC TP53 Mutation Database

IARC TP53 Mutation Database (http://www-p53.iarc.fr/)

- Release 11: 23,544 somatic mutations, 2,003 papers,
- Detailed information on:
  - molecular aletration, prognostic value, mutation prevalence,
  - biosource (morphology, topography, tumor's grade and stage),
  - patient's demographics and life-style,
  - germline mutations + family status,
  - bibliographic references,
  - cell-lines TP53 status,
  - known human polymorphisms,
  - functions' losses/preservations in mutated proteins.
- Vocabularies and standardized annotations:
  ICD-O for morphology, topography, stage, grade, …
  SNOMED
- On-line queries imply human interaction and are analysis oriented
- Download of data is limited to full data sets (tab delimited text files).

Web Analysis Tools

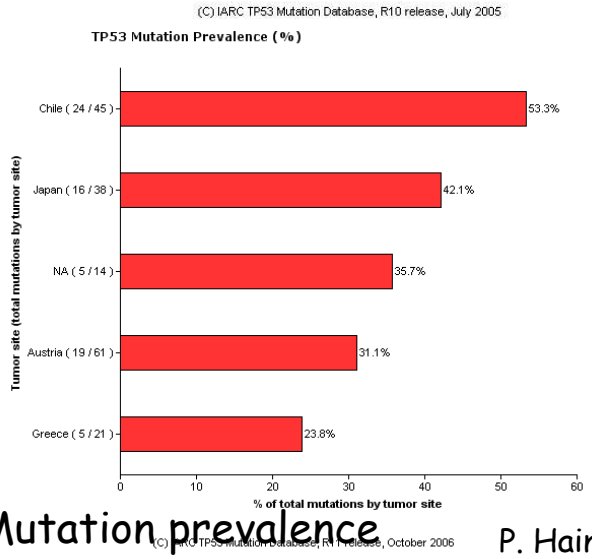


Mutation patterns

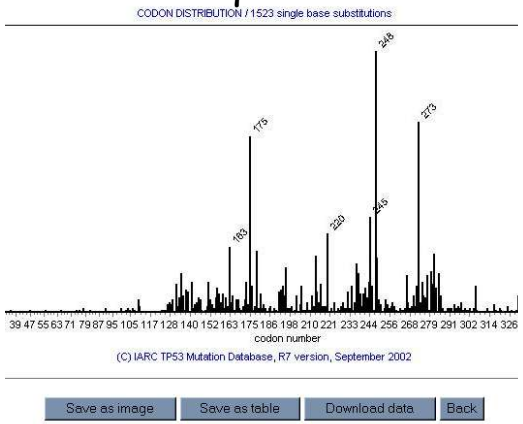
AA Change	Lost Function	Retained Function	Cell Lines	Ref ID (PubMed)	SwissProt
L22E/W23S	TA (BAX, MDM2, IGF-BP3, RGC)	TA (WAF1, p53CON, CCNG1, GADD45)	Yeast	35	no link
E62_W91del	TA (PIG3); APO	TA (MDM2)	H1299	22	no link
A63_A78del	TA (RGC)partial	NA	Yeast	6	no link
A63_A78del	TA (WAF1)	NA	Saos-2	6	no link
P85S	NA	TA (MDM2, PIG3)	H1299	22	no link
S96P	NA	TA (RGC)supertrans	Yeast	51	no link
S99F	TA (BAX, PIG3)	TA (WAF1)	Yeast	34	no link
S99Y	TA (BAX, PIG3)	TA (WAF1)	Yeast	34	no link
G108_L111delinsIQ	TA (WAF1, MDM2); GS	NA	H1299	65	no link
F113Y	TA (WAF1, MDM2, BAX, 14-3-3s, GADD45, AIP1, Noxa, P53R2)	NA	Yeast	46	no link
H115R	NA	TA (RGC)supertrans	Yeast	51	no link
S116A/S127A	NA	TA (RGC)	Saos-2	63	no link

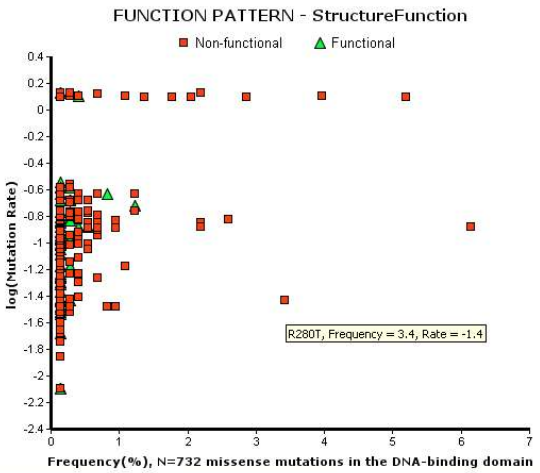
Functional properties of mutations

Frequencies/Distributions of mutations



Mutation prevalence





P. Hainaut and M. Olivier. "TP53 Genetic Variations In Human Cancer" slide show at http://www-p53.iarc.fr/

SRS Sequence Retrieval System

Reasons why SRS

- Manages heterogeneous databases
- Originally based on "flat file" format

→ Quick implementation of the mutation database

- Carries out integrated queries (more databases, databases and analysis tools)
- Internal, exlicit and implicit links, link operator <

→ Adapt to the many data sets that make up the IARC Mutation db

- Flexibility in creation of indexes (word, sentence, complex terms)

→ Support controlled vacabularies (ICD-O, SNOMED)

- Well known, researchers can proficiently query the system
- Many public sites and libraries
- Simple and effective query language and interface

→ Good example for improved integration of dbs

TP53 Mutation Database integration

SRS implementation of the TP53 Database
http://srs.o2i.it/srs71/

- Download of data sets from IARC
- Insertion in a purpose relational database (MySQL)
- Scripts for extracting data and creating flat files (perl)
- Definition of data and indexes formats (icarus)

List of Databanks - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://srs.o2i.it/srs71bin/cgi-bin/wgetz?-page+databanks

LION

Help Center

Quick Searches | Select Databanks | Query Form | Tools | Results | Projects | Custom Views | Information

SRS

Display Options

List databanks:

by groups

List

Databank Information

Databank	Release	No. of Entries	Indexing Date	Group	Availability
TP53_SOMATIC	11	23544	07-Mar-2007	Somatic Mutations data sets	OK
TP53_PREVALENCE	11	1570	07-Mar-2007	Somatic Mutations data sets	OK
TP53_PROGNOSIS	11	142	07-Mar-2007	Somatic Mutations data sets	OK
TP53_MUTATIONS	11	5688	07-Mar-2007	Somatic Mutations data sets	OK
TP53_GERMLINE	11	1783	07-Mar-2007	Germline Mutations data sets	OK
TP53_POLYMORPHISM	11	42	07-Mar-2007	Polymorphisms data sets	OK
TP53_MUTATION_FUNCTION1	11	2505	07-Mar-2007	Mutant Functions data sets	OK
TP53_MUTATION_FUNCTION2	11	2312	07-Mar-2007	Mutant Functions data sets	OK
TP53_CELL_LINES	11	1754	07-Mar-2007	Cell-line Status data sets	OK
TP53_SOMATIC_REF	11	2003	07-Mar-2007	Bibliographic References data sets	OK
TP53_GERMLINE_REF	11	148	07-Mar-2007	Bibliographic References data sets	OK

Done

Databank Information - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://srs.o2i.it/srs71bin/cgi-bin/wgetz?-page+LibInfo+-id+2AMGU1UXvAU+-lib+TP53_SOMATIC

LION

Help Center

Quick Searches | Select Databanks | Query Form | Tools | Results | Projects | Custom Views | Information

SRS

Name

Somatic Mutation data set (IARC TP53 Mutation Database, Rel. 11)

Status

The current release has 23544 entries and was indexed 07-Mar-2007.

Description

This is the Somatic Mutations data set of the IARC TP53 Mutation Database (Rel. 11)

WWW

http://www-p53.iarc.fr/index.html

Contact

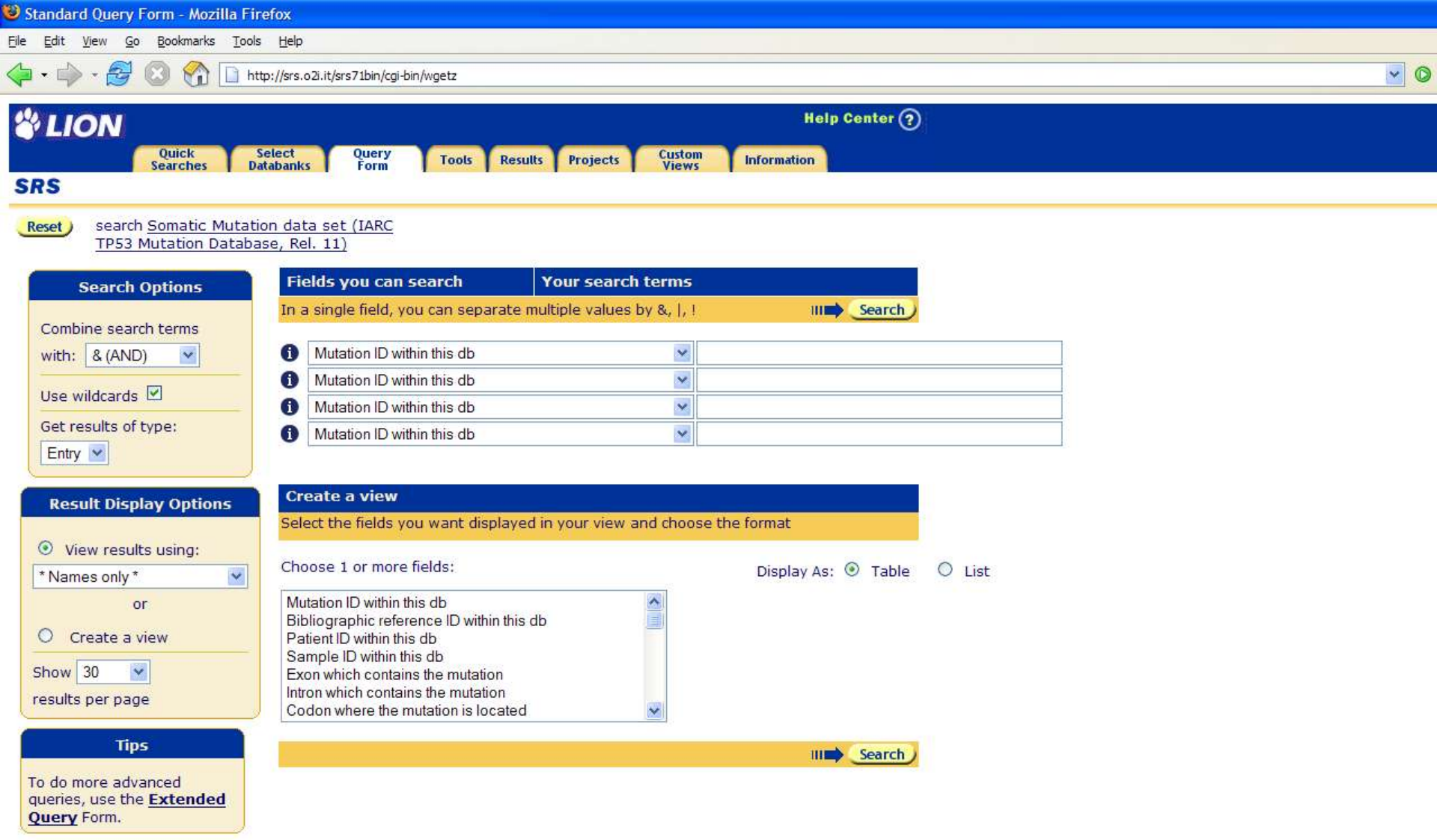
International Agency for Research on Cancer (IARC) TP53 database,
IARC,
150 Cours Albert Thomas,,
69372 Lyon CEDEX 08, France.

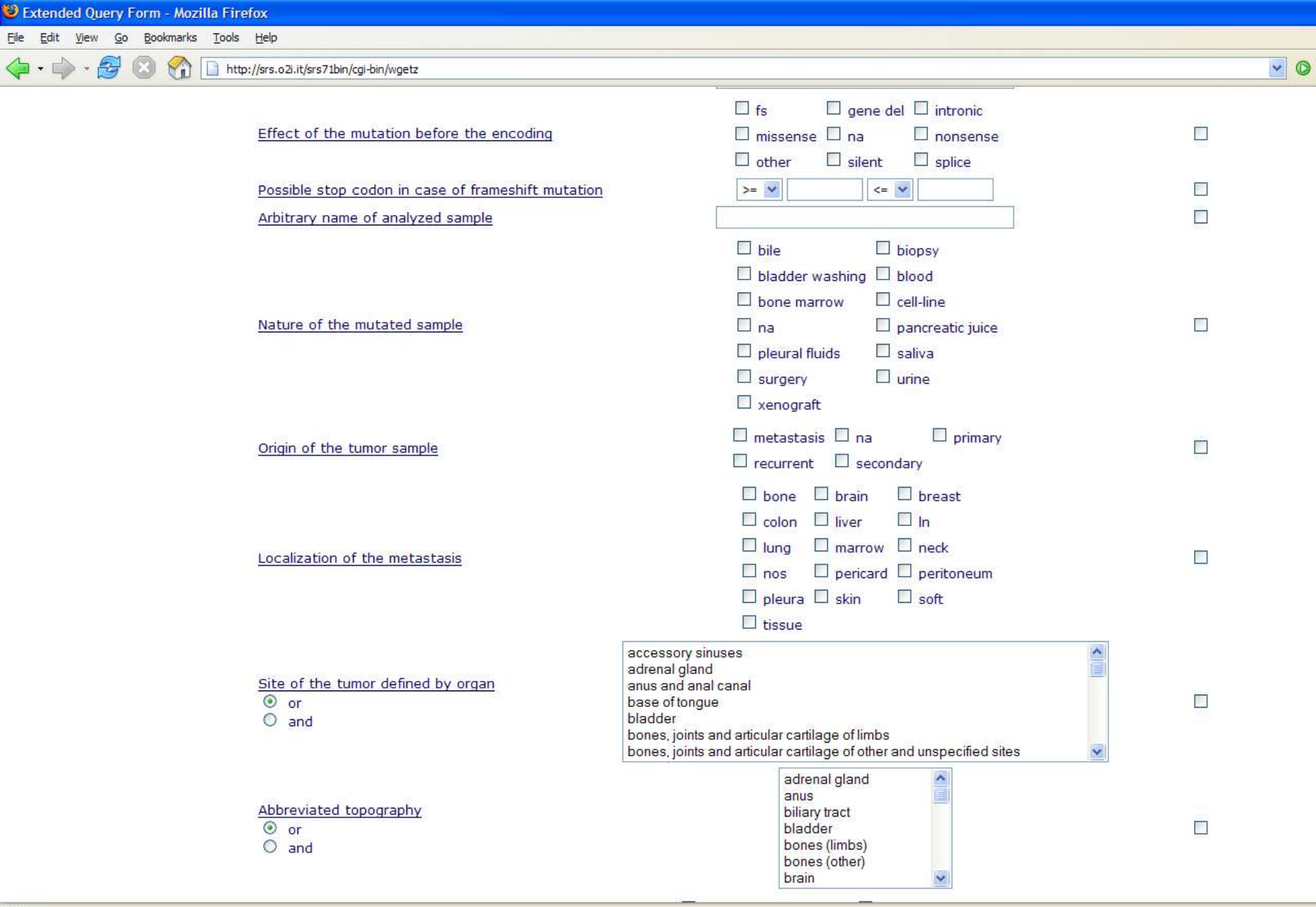
Phone: +33 (0)4 72 73 84 85
Fax : +33 (0)4 72 73 85 75

Data-fields in SRS

Field Name	Short Name	Type	No. of Keys	No. of Entry References	Indexing Date	Status
Mutation ID within this db	mid	id	23544	23544	07-Mar-2007	ok
Bibliographic reference ID within this db	rei	index	1995	23544	07-Mar-2007	ok
Patient ID within this db	ind	index	21181	23544	07-Mar-2007	ok
Sample ID within this db	sai	index	21537	23544	07-Mar-2007	ok
Exon which contains the mutation	exo	index	10	22905	07-Mar-2007	ok
Intron which contains the mutation	itr	index	10	618	07-Mar-2007	ok
Codon where the mutation is located	cno	num	364	23544	07-Mar-2007	ok
Genomic nucleotide position of mutation	nug	num	1078	23544	07-Mar-2007	ok
Occurance of mutation in a CpG site	cpg	index	2	23544	07-Mar-2007	ok
Occurance of mutation in a Splice site	sps	index	2	23544	07-Mar-2007	ok
Type/nature of mutation	typ	index	13	23544	07-Mar-2007	ok
Description and relevant information of mutation	des	index	305	23544	07-Mar-2007	ok
Base sequence of Wild Type codon	wtc	index	59	23544	07-Mar-2007	ok
Base sequence of Mutant codon	muc	index	115	23544	07-Mar-2007	ok
Wild Type Aminoacids	wta	index	22	23544	07-Mar-2007	ok
Mutated Aminoacids	mua	index	69	23544	07-Mar-2007	ok
Mutation description at the protein level	pde	index	1649	23544	07-Mar-2007	ok
Effect of the mutation before the encoding	eff	index	9	23544	07-Mar-2007	ok
Possible stop codon in case of frameshift mutation	pus	num	78	23544	07-Mar-2007	ok

Done


Standard Query Form - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://srs.o2i.it/srs71bin/cgi-bin/wgetz
LION
Help Center
Quick Searches
Select Databanks
Query Form
Tools
Results
Projects
Custom Views
Information
SRS
Reset
search Somatic Mutation data set (IARC TP53 Mutation Database, Rel. 11)
Search Options
Combine search terms with: & (AND)
Use wildcards
Get results of type: Entry
Result Display Options
View results using: * Names only *
or
Create a view
Show 30 results per page
Tips
To do more advanced queries, use the Extended Query Form.
Fields you can search
Your search terms
In a single field, you can separate multiple values by &, |, !
Search
Mutation ID within this db
Mutation ID within this db
Mutation ID within this db
Mutation ID within this db
Create a view
Select the fields you want displayed in your view and choose the format
Choose 1 or more fields:
Display As: Table List
Mutation ID within this db
Bibliographic reference ID within this db
Patient ID within this db
Sample ID within this db
Exon which contains the mutation
Intron which contains the mutation
Codon where the mutation is located
Search
SRS Release 7.1.2 Copyright © 1997-2003 LION bioscience AG. All Rights Reserved. Terms of Use Feedback
Done


Extended Query Form - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://srs.o2i.it/srs71bin/cgi-bin/wgetz
fs
gene del
intronic
Effect of the mutation before the encoding
missense
na
nonsense
other
silent
splice
Possible stop codon in case of frameshift mutation
>=
<=
Arbitrary name of analyzed sample
bile
biopsy
bladder washing
blood
bone marrow
cell-line
Nature of the mutated sample
na
pancreatic juice
pleural fluids
saliva
surgery
urine
xenograft
Origin of the tumor sample
metastasis
na
primary
recurrent
secondary
bone
brain
breast
colon
liver
ln
Localization of the metastasis
lung
marrow
neck
nos
pericard
peritoneum
pleura
skin
soft
tissue
Site of the tumor defined by organ
or
and
accessory sinuses
adrenal gland
anus and anal canal
base of tongue
bladder
bones, joints and articular cartilage of limbs
bones, joints and articular cartilage of other and unspecified sites
Abbreviated topography
or
and
adrenal gland
anus
biliary tract
bladder
bones (limbs)
bones (other)
brain
Done

TP53 Mutations Web Services

Implementing web services that allow:

- The retrieval of information from TP53 databases by using remote calls to SRS
- The possibility of including such services in complex workflows

Reproducing current behaviour:

- Search by interesting properties (TP53)
- Combine results
- Integrate data with other sources by using IDs/common terms

Two types of services:

- Search for a specific feature and return ID
- Search for an ID and return full record (or predefined sections)

Soaplab: SOAP-based Analysis Web Service

"Soaplab is a set of Web Services providing a programatic access to some applications on remote computers. It is often referred to as an **Analysis (Web) Service**" (Martin Senger, EBI).

It allows for the implementation of Web Services offering access to:

- local command-line applications
- contents of ordinary web pages (GowLab)
- EMBOSS

Requirements

- Apache Tomcat servlet engine and Axis SOAP toolkit, Java
- perl, mySQL

Graphically…

1. ACD (analysis description)
2. acd2xml (convertor)
3. XML (analysis description)

4. Tomcat (or any servlet engine)
   - Apache Axis (SOAP Toolkit)
   - 5 Soaplab services
   - MySQL with results
   - Files: IOR/analyses/...
   - AppLab server
   - Filesystem with analysis results
   - Command-line driven analysis tools

6. Provided ready clients — Soaplab
7. Your own client programs
   - Talisman
   - BioPerl modules (Bio::Tools::Run::Analysis)
   - myGrid clients

1. Create analysis description
2. Convert to XML
3. Start AppLab server
4. Start Tomcat server
5. Deploy Web Services
6. Use existing client, or/and
7. write your own client

This is Soaplab

Important:
- how to write ACD
- what API to use by the clients

Downoad: soaplab-clients-*version*

Downoad: applab-cembalo-soaplab-*version* (for EMBOSS)

Downoad: analysis-interfaces-*version*

Binary distribution

http://industry.ebi.ac.uk/soaplab/dist/

Soaplab: getP53MutationIdsByExon

SRS can be accessed by using a properly formatted URL:
http://srs.o2i.it/srs71bin/wgetz?+[{tp53_somatic}-Exon:'1']+-ascii+-lv+2000000

```
appl: getP53MutationIdsByExon [
  documentation: "Get TP53 mutation IDs by exon from IARC TP53 Mutation
     Database (SRS implementation, see srs.o2i.it/srs71/)"
  groups: "O2I"
  nonemboss: "y"
  comment: "launcher get"
  supplier: "http://srs.o2i.it/srs71bin/cgi-bin/wgetz"
  comment: "method [{$libs}-Exon:'$exon'] -lv 2000000 -ascii"
]

string: libs [ parameter: "Y" ]
string: exon [ parameter: "Y" ]
outfile: result [ ]
```

Web Services TP53

Web Service Name	Input	Output
getP53MutationsByProperty	lib, text	Full record
getP53MutationsByIds	id	Full record
getP53MutationIdsByType	lib, mutation type	id(s)
getP53MutationIdsByEffect	lib, effect	id(s)
getP53MutationIdsByExon	lib, exon number	id(s)
getP53MutationIdsByIntron	lib, intron number	id(s)
getP53MutationIdsByCodonNumber	lib, codon number	id(s)
getP53MutationIdsByCpgSite	lib, cpg site (true/false)	id(s)
getP53MutationIdsBySpliceSite	lib, splice site (true/false)	id(s)
getP53MutationIdsByMetastasisLocalization	lib, metastasis localization (organ)	id(s)
getP53MutationIdsByTumorOrigin	lib, origin (primary, secondary, …)	id(s)

http://www.o2i.it:8080/axis/services

Taverna: Workflow development tool

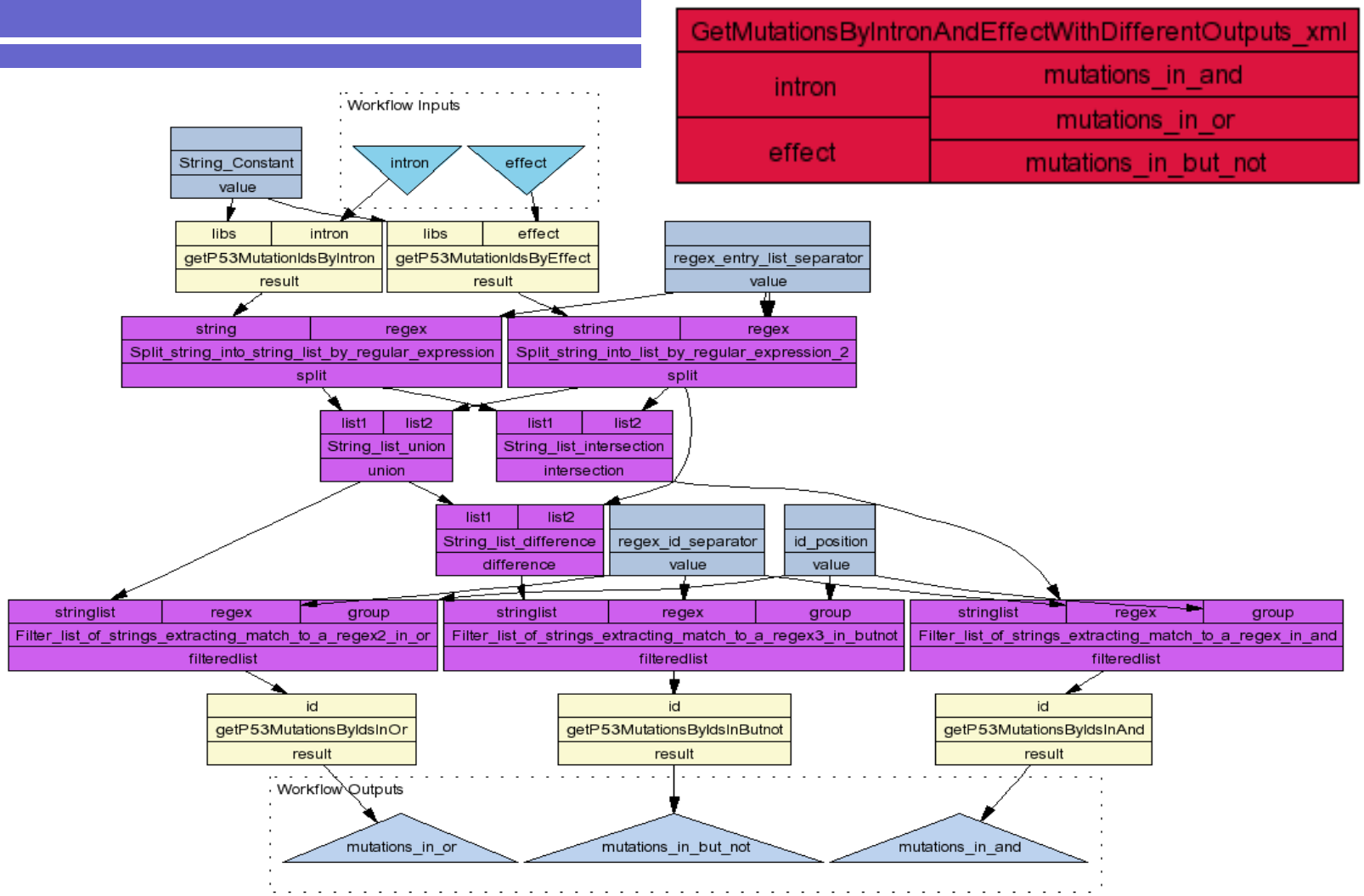
Goal: set up working workflows now

Taverna Workbench

- constructs complex analysis workflows
- access both remote and local processors
- defines alternative processors
- runs workflows
- visualizes the results by using various formats
- includes a bioinformatics data ontology

Requirements: java, Windows or Linux
Open source: http://taverna.sourceforge.net/
Version: 1.5.1 (stable, next version 2.0?)

Workflow for TP53 mutations' analysis



Conclusions

Moving from an interactive to an automated approach is needed in order to achieve a real data integration, but this requires new technologies and tools.

A possible methodology is based on:

- XML schemas for the creation of the models of the information,
- XML based languages for data representation and exchange,
- Web Services for the interoperability of software
- Computerised workflows for the definition and execution of analysis processes

We showed how:

- Databases can effectively be uploaded/imported to/in powerful integration tools, like SRS, with little effort
- Web Services can also be effectively implemented by using existing tools like SoapLab,
- Workflows can be created and effectively used for biological information management/mining
- The proposed methodology offer a good reference for the development of new systems and the refinement of existing ones

More conclusions

Only a few data models are available for biological information.

-> Efforts should be done for further design

Semantic information is still lacking, both with reference to terminologies (ontologies) and to information sources (metadata)

-> Semantic tools should be developed and widely used for data description and interoperation

Many information sources are not yet available through programmatic interfaces, such as Web Services.

-> Tools for making them available through WS should be developed

All proposed workflow management software for bioinformatics require knowledge of the WSs and skills and time for development and maintenance of workflows.

-> Portals can be the answer

SRSbyWS

SRSbyWS is a development aiming at allowing access to ALL databases that are available in public SRS sites.

It is implemented by using:

- The list of publicly available SRS sites
- A local relational db for sites, databases, analysis tools and implementations (MySQL, kept up-to-date by a cron process)
- SoapLab for implementing the WS
- Scripts for identifying the "best" site to search and for retrieving data (perl/php)

Implemented WS include:

- Querying a database (the system identifies the best site)
- Querying a site for the list of implemented databases
- Querying a database for the list of available implementations
- Querying by ID, free text, any field
- Retrieving IDs, complete entries, any field
- Submission of "complex" queries involving many databases

Biowep: features (i)

We designed biowep, a workflow enactment portal for bioinformatics that:

- allows for the carrying out of predefined workflows
- supports workflows annotation through a simple ontology for bioinformatics processors (domain, task, i/o)
- implements search and selection of workflows on the basis of their annotation
- supports retrieval of workflows on the basis of users' profiling
- allows storing and retrieval of workflows' executions and related results

Biowep: features (ii)

The system:

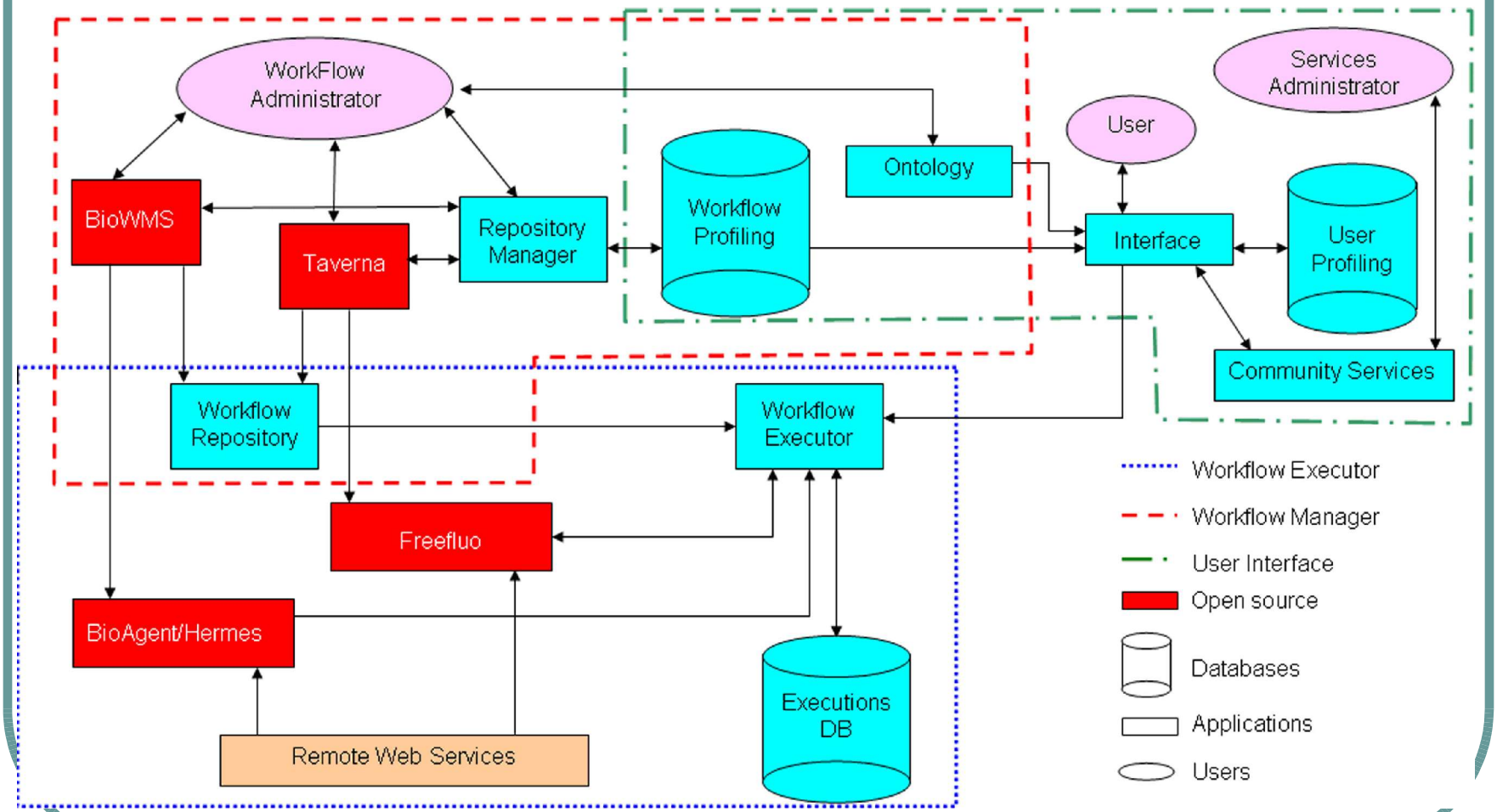
- uses open source sw (Taverna WB and mySQL)
- makes access to all services supported by Taverna
- stores workflows in Scufl and/or XPDL formats
- is available under LGPL license
- requirements: Java SDK + Tomcat with Axis

Available on-line:

http://bioinformatics.istge.it/biowep/ (reference site)
http://bioinformatics.istge.it:8080/biowep/ (portal)

Romano P et al, Biowep: a workflow enactment portal for bioinformatics applications, BMC Bioinformatics 2007, 8 (Suppl 1):S19

Biowep: architecture



Some acknoledgements…

Special thanks to my colleagues:

- Domenico Marra (SRS and Web Services),
- Chiara Rasi (workflows and biowep portal)

This work has partially been supported by the Italian Ministry for Education, University and Research (MIUR), projects:

- "Oncology over Internet – O_2I" (2002 – 2005)
- "Laboratory for Interdisciplinary Technologies in Bioinformatics - LITBIO" (2006 – 2009)

…and an announcement!

Workshop NETTAB 2007

http://www.nettab.org/2007/

A Semantic Web for Bioinformatics:

Goals, Tools, Systems, Applications

June 12 - 15, 2007

University of Pisa, Italy

Deadline for oral communications:
March 16, 2007