



Stanford - South Africa

Biomedical Informatics Program



Technical issues for PharmGKB

Russ B. Altman, MD, PhD
Professor of Genetics, Bioengineering &
Medicine
Stanford University



Five technical informatics issues

1. Overall architecture
2. RDBMS vs. KB
3. XML for data exchange
4. UI considerations
5. Pathway representation issues (DONE)
6. Terminologies (genes, drugs, diseases)
7. Makeup of PharmGKB team



1. Overall Architecture

3-tier system

1. Database tier (Relational DBMS)--holds data
2. Middle layer (Java)--holds logic for retrieval & analysis
3. Presentation layer (Javascript)--presents data and dynamically creates most web pages



View of PharmGKB 3-tier

[Nice picture available on PharmGKB, under Resources tab]

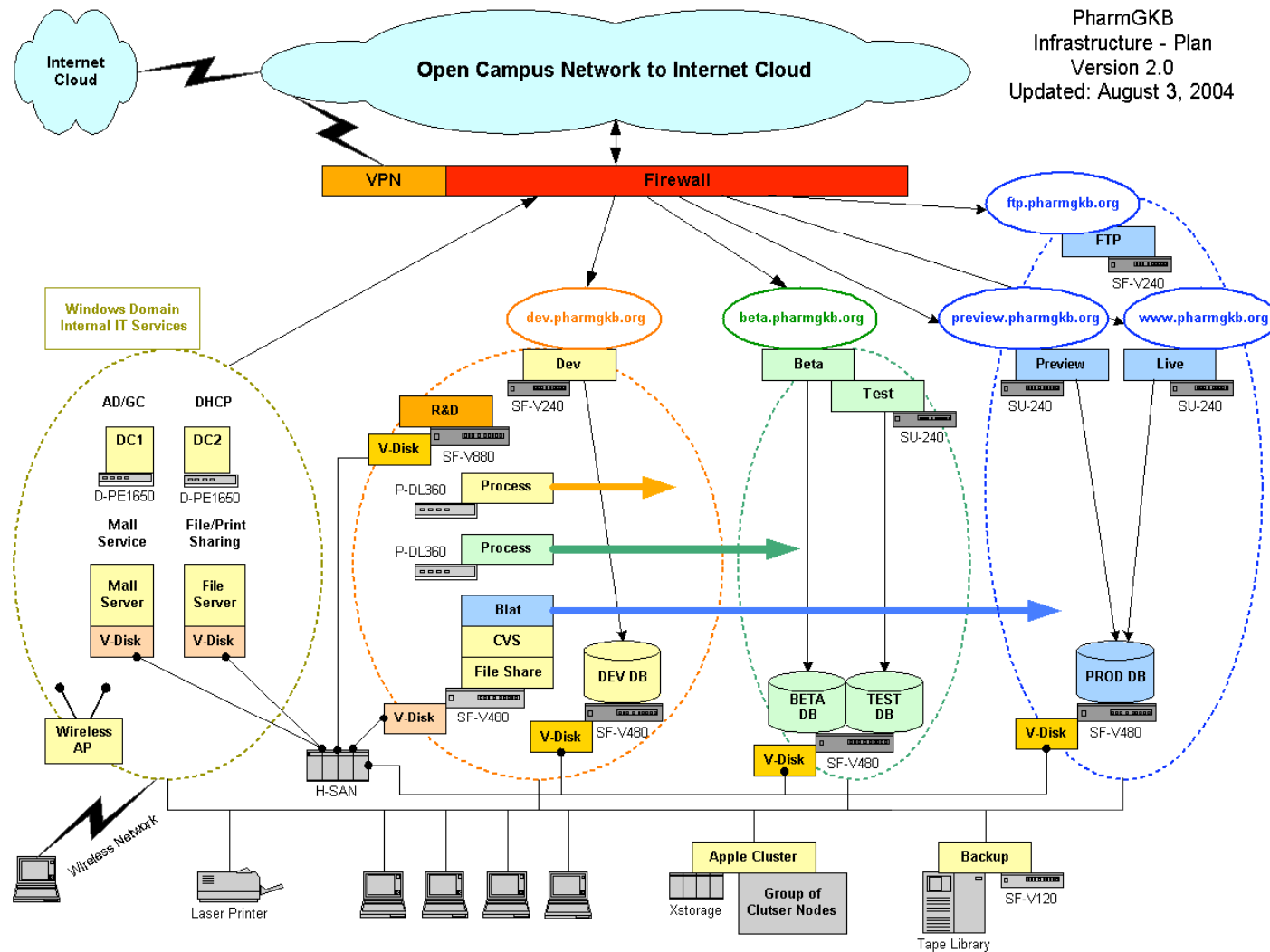


Tools used in building PharmGKB

Open Source	Acme GifEncoder (GIF generation), Ant (Java make system), Avalon (object management), AwStats (web stats), Axis (SOAP API), BLAT (genome alignment), CAS (authentication), Jakarta Commons Libraries (java utilities), CVS (version control), DOM4J (XML parser utility), EasyMock (mock object test library), FreeMarker (templating library), Hibernate (persistence library), Apache Httpd (web server), HttpUnit (testing library for web forms), Jetty (test web server), Junit (java testing framework), POI (excel parsing), Log4J (debug utility), Lucene (text indexing), McKoi DB (test relational DB), Oro (regular expression library), P6Spy (debug library for SQL), Subversion (will replace CVS), Struts (webform framework), Tomcat (JSP server), Trove (java libraries for collections), URL Rewrite (URL redirection), WebWork (will replace Struts), Xalan (XML formatting library), Xerces (XML parser)
Purchased	Acrobat Writer (PDF creation), Dreamweaver (fast web prototyping), GoLive (HTML editor), IDEA (Java development environment), Illustrator (graphics), InDesign (web page design tools), JIRA (bug tracking), Jive (discussion groups), MS Office Pro (word processing), Oracle (database), Photoshop (image editing), RoboHelp (help system management), TOAD (SQL editor), Visio Pro (block image creation)
Other	Java (programming)



PharmGKB Hardware infrastructure



2. RDBMS vs. KBS

- Initially, wanted to use a hierarchical knowledge-based system (e.g. Protégé)
- Performance drove us to RDBMS, with faster queries, but more difficult to change.
- Moving to use ontology and semantic web principles to disseminate PharmGKB.

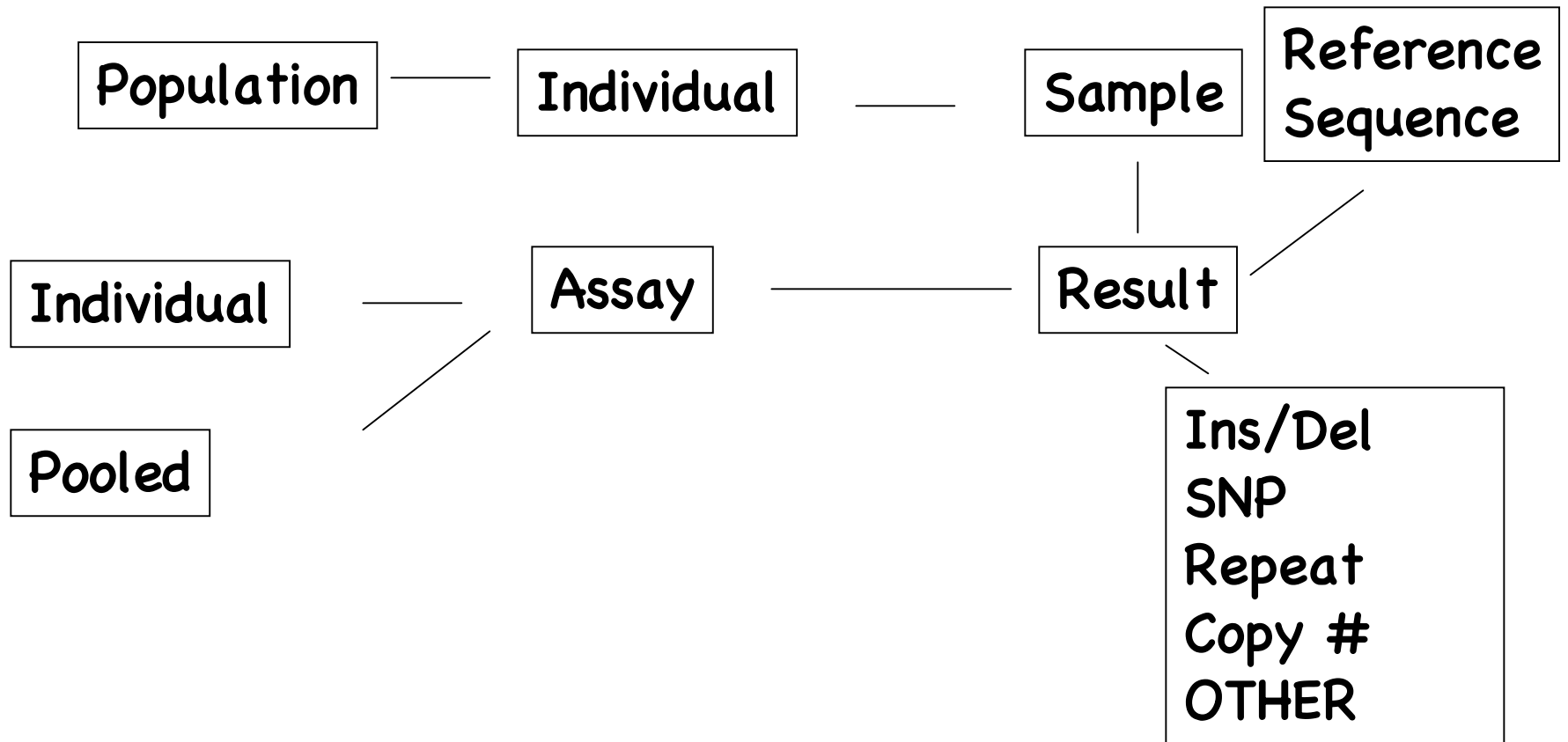


3. XML for data exchange

- We defined an XML syntax for transferring genotype data.
- Very useful for us to have all the correct data, verify it upon submission, use it for functionality
- Significant resistance from biology groups
 - Those without bioinformatics resources
 - Those with bioinformatics resources



Outline of Genotype Schema



PharmGKB XML Schema

<http://www.pharmgkb.org/schema/index.html>



4. UI considerations

- Don't innovate too much in basic web site
- Use Amazon/Google precedents
- Importance of attractive look
- Importance of consistency in interface
- Tough decisions about power vs. ease



6. Terminologies

- It is critical to adopt standard terminologies for exchange of PGx data.
- Key terminologies:
 - Genes (Human Genome Nomenclature Comm.)
 - SNP locations (UCSC Golden Path Browser)
 - Drugs (Micromedix drug information, cf. Also KEGG drug taxonomy http://www.genome.ad.jp/dbget-bin/get_htext?Drug+-n)
 - Diseases (Medical Entity Subject Headings)
 - Symptoms/Adverse events (MESH, WHO-ART, Meddra-
– www.meddramsso.com)
 - SNOMED



HGNC

{Search Google to see sample pages}



KEGG Drugs

{Search Google to see sample pages}



MESH/UMLS

{Search Google to see sample pages}



WHO-ART

{Search Google to see sample pages}



Related issue: version control

- As new releases come of the terminologies we need to update.
- Developed a “dictionary” system that manages terminologies and ensures that they are up to date and synchronized



Makeup of Team

(also see "Home > Team" tab at PharmGKB)

- Managers/Administrators (2)
- PhD-level scientific curators (4)
- Assistant curators (BS & MS) (2)
- Informatics specialists (2)
- Developers/programmers (3)
 - Database, Middle Layer, Presentation
- Graphic Design (1)
- Systems/Hardware (2)



Tools for tracking submissions

PharmGKB: View Submissions from PAAR-SJCRH

http://preview.pharmgkb.org/views/submissions.jsp?pld=10

PharmGKB The Pharmacogenetics and Pharmacogenomics Knowledge Base

Search PharmGKB: Go

Home Search Submit Resources PGRN Contributors My PharmGKB sign out help feedback

Submission PS204273 View Submissions from [PAAR-SJCRH](#)
(Pharmacogenetics of Anticancer Agents - St. Jude Children's Research Hospital)

Submitter: [PA](#)
Contents: [Ge](#)
Download: [\[X](#) View: All Submissions

Next Action: [PA](#)
Options: [\[A](#)

Jump To:
[Submissions By Project](#)
[Submissions By User](#)
[Browse Index](#)

Genotype Submissions (45)

[ABCB1](#) - [view all variant positions](#)

alternate symbols: ABC20, GP170, MDR1, NM_000927.1, P-GP, P-gp, PGY1

Submission Id	Last Action	Next Action
PS200427	Uploaded to www.pharmgkb.org (05/19/2003)	
PS200737	Uploaded to www.pharmgkb.org (05/19/2003)	
PS204196	Uploaded to preview.pharmgkb.org (07/30/2004)	PAAR-SJCRH's PI to review submission and approve if acceptable.
PS204272	Uploaded to preview.pharmgkb.org (07/30/2004)	PAAR-SJCRH's PI to review submission and approve if acceptable.
PS204273	Uploaded to preview.pharmgkb.org (07/30/2004)	PAAR-SJCRH's PI to review submission and approve if acceptable.

Variant Positions

Legend
- deletion
* feature derived from NCE

Golden Path Position [G](#)

[chr7:86736872](#) [AF](#)
[chr7:86738845](#) [AF](#)

[[Supporting Data](#)]

History

- 07/30/2004: Subm
File name: paar_subr
Note: PAAR-SJCRH; P
- 07/30/2004: Passe
- 07/30/2004: Passe
- 07/30/2004: Uploa

[ABCG2](#) - [view all variant positions](#)

alternate symbols: ABC15, ABCP, BCRP, BCRP1, BMDP, EST157481, MRX, MXR, MXR1

Submission Id	Last Action	Next Action
PS203469	Uploaded to www.pharmgkb.org (02/26/2004)	

[ADRB2](#) - [view all variant positions](#)

alternate symbols: ADRB2R, ADRBR, B2AR, BAR

Submission Id	Last Action	Next Action
---------------	-------------	-------------

Add a Comment

