

# Pathway-based Approaches to Pharmacogenomics

C.F. Thorn<sup>\*</sup>, M. Whirl-Carrillo, T.E. Klein and R.B. Altman

Stanford University School of Medicine, Department of Genetics, Stanford, CA 94305, USA

**Abstract:** Researchers are using pathway information for both pharmacogenomics study design and data analysis. Candidate gene approaches for the design of pharmacogenomic studies need reliable pathway information to choose the best candidate genes and SNPs, especially when using low to medium throughput genotyping technologies, to maximize the likelihood of success. With the mainstream use of high throughput gene expression microarrays and genome wide SNP assays, pathway-based approaches can also be valuable tool for data analysis. This review will discuss sources of pathway data and mechanisms for applying it to pharmacogenomics research.

## INTRODUCTION

Pathways are the representation of a series of reactions occurring between components in a process. Biological pathways tend to be complex networks of inter-relationships that describe interactions between molecules, macromolecules and cells building them up into physiological processes such as metabolism, signaling and transcriptional regulation. Pharmacogenomic pathways are biological pathways that involve genes and drugs and result in drug metabolism (pharmacokinetics) or drug response (pharmacodynamics).

Recent emphasis on systems biology and interdisciplinary research has brought together pharmacologists, geneticists, computational biologists, and physician scientists to examine drug action from both a whole organism outlook as well as a molecular perspective. Biological pathways provide an intermediate level view between the gene and the clinical outcome. There are a large number of pathways resources that have sprung up as a result of interdisciplinary projects (consisting of over 200 pathway websites and companies)<sup>1</sup>. These resources have a variety of goals and foci ranging from identifying gene functions in model organisms to providing tools for drug discovery. Pathway resources can aid pharmacogenomics researchers in bridging the knowledge gap from gene to whole organism, however knowing which resources to select from the many, and which one to use under certain situations, can be problematic. This review will evaluate which kinds of pathway tools to use for identifying different kinds of pharmacogenomic knowledge and highlight some of the key pathway resources.

Many of the pathways resources for pharmacogenomics are more readily useful to the drug-centered researcher who knows the background mechanism of action of a drug and is searching for candidate genes (phenotype to genotype approaches) rather than a gene-oriented researcher with novel gene variants searching for the drug whose actions they may influence (genotype to phenotype).

Pharmacogenomic pathways can be subdivided into pharmacodynamic (PD), that illustrate drug response, and

pharmacokinetic (PK) pathways, that depict drug metabolism. PD pathways have considerable overlap with existing metabolic and signaling pathways; a drug is often designed to act at a specific receptor and promote or block the known endogenous signaling pathway, or act on a specific enzyme and promote or block downstream metabolism. Thus the PD pathway often resembles a normal physiological pathway with drugs added. Nonetheless, PD pathways are not always as simple as knowing the target signaling or metabolic pathway as there may be additional previously unpredicted pathways effected by the drug which lead to side effects or adverse events. While PK pathways may follow existing metabolic pathways that are for elimination of exogenous compounds, predicting which genes will be involved *a priori* is far less obvious than for PD pathways and additional knowledge is needed.

This review introduces several pathway resources for study design and analysis and gives some use case scenarios that describe instances in which particular resources might be useful. This review will also discuss pathway database standards and exchange of data between pathway resources, both of which are of relevance to pharmacogenomic research because they facilitate aggregation of knowledge from multiple sources. Breaking down pathways into components is an important part of this process. As such, and reflecting the interdisciplinary nature of this field, pathways are often described using engineering terms since they can be represented by process diagrams consisting of nodes (boxes) and edges (arrows or connectors). The exchange of pathway data related to drug PD and PK may have a major impact on pharmacogenomics.

## PATHWAY RESOURCES FOR STUDY DESIGN AND ANALYSIS

With the expansion of single gene pharmacogenetics into pharmacogenomics, it is natural for study design to move from investigating single candidate genes to multiple candidates related by a common pathway. This expansion has tended to follow two routes: (1) to investigate novel genetic variation in well-connected genes in an established pathway or (2) to look for novel genes in related or inferred pathways. For example: Freimuth *et al.* resequenced 51 genes in pathways of chemotherapy drugs to capture the extent of variation and estimated frequencies in Caucasian, Asian and African DNA samples for the Coriell repository [Freimuth *et al.*, 2005]; studies of folate metabolism looked to find pathways

<sup>\*</sup>Address correspondence to this author at the Stanford University School of Medicine, Department of Genetics, 300 Pasteur Drive L337, Stanford, CA 94305; USA; Tel: 1-650-725-7013; E-mail: thorn@helix.stanford.edu

<sup>1</sup> A comprehensive list of biological pathways resources can be found at <http://www.pathguide.org/>

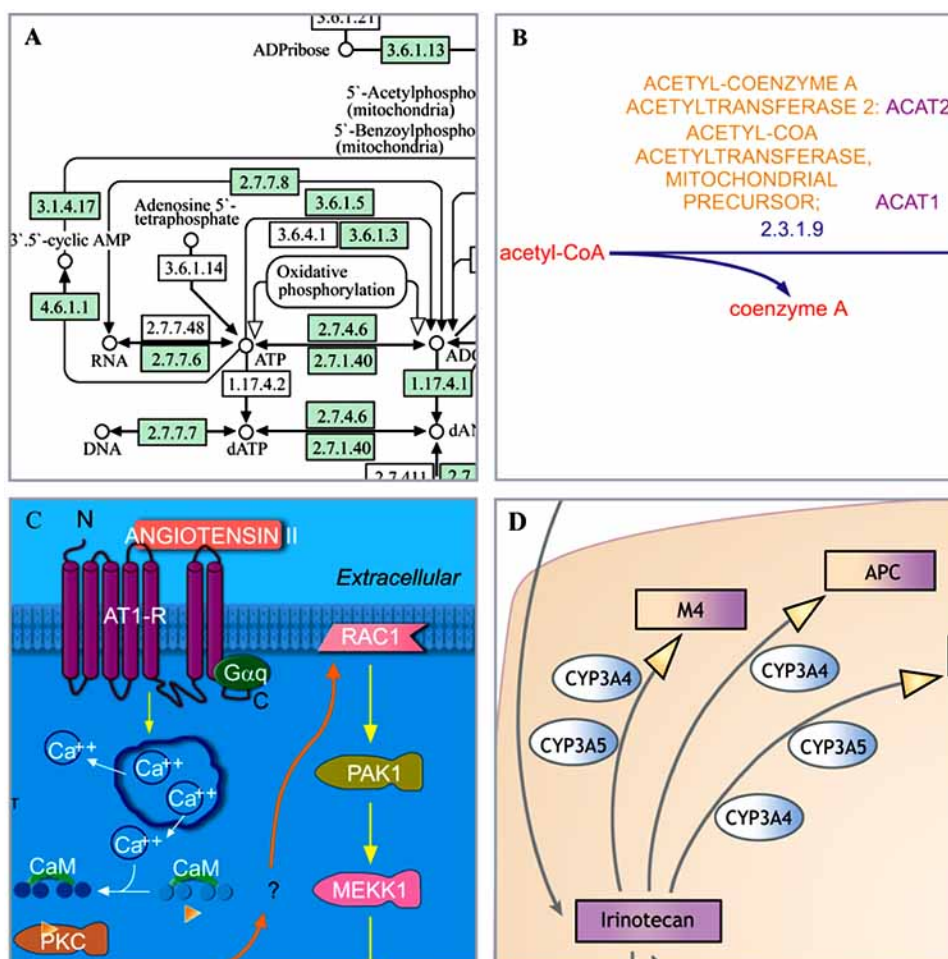
branching from well defined genes in the folate pathway to identify NOS3 and MCP-1 as new candidate genes associated with *Spina bifida* [Brown *et al.*, 2004; Jensen *et al.*, 2006].

High throughput gene expression microarrays generate such large quantities of data that analysis of the results can be difficult. Pathways offer one approach to expression array analysis that put the results in a biological perspective in a visual way that is easy to understand. A user may begin with a particular pathway and want to map expression array results onto that pathway to see what genes/proteins are affected. Alternatively, a user may wish to find pathways that incorporate genes with expression rates up- or down-regulated. Some resources can also be used to infer pathways or relatedness of genes through their common expression or comparison to data from other species. Whichever analysis style is used, pathways present a contextual view of the data.

There are several freely available resources for identifying candidate pathways and genes involved in PD. KEGG, the Kyoto Encyclopedia of Genes and Genomes pathways, has become a gold standard for metabolic pathways [Kanehisa *et al.*, 2006]. The pathways link enzymes, listed by their Nomenclature Committee of the International Union of Bio-

chemistry and Molecular Biology EC identifier, and substrates. The pathways are depicted as simple node and edge diagrams (Fig. 1A). Box shaped nodes represent the enzymes, contain the EC number, and link to the page describing the reaction, other pathways the enzymes is involved in, the genes encoding the enzyme, references for the reaction and other databases (Table 1). Circle nodes represent the substrates and link to structures, other reactions and pathways and other databases such as PubChem. Pathways are interconnected to each other to provide a whole metabolome view, both composite and colored organism specific, including human, mouse, rat, and all of the major model organisms. Although they have a drug database with structures and metabolites, their pathways are not indexed based on relevance to drugs. While the KEGG resource does not offer analysis tools on its website the pathways can be downloaded and used in other software packages to analyse microarray and other expression data.

Reactome is another excellent source of metabolic pathways from multiple organisms including human and mouse and several pathogens [Joshi-Tope *et al.*, 2005]. The user begins with the highest level view depicting all of the reactions in the organism before selecting a focus area of me-



**Fig. (1).** Screenshots from Pathways Resources illustrating different styles of pathways and displays: (A) section from the KEGG purine metabolism pathway, (B) section from the HumanCyc mevalonate pathway, (C) section from the Biocarta Angiotensin II mediated activation of JNK Pathway *via* Pyk2 dependent signaling pathway, (D) section from the PharmGKB irinotecan pharmacokinetics pathway.

Table 1. Pathway Resources for Study Design and Data Analysis

Name (Website)	PD/PK	Display Style	Search by Pathway	Search by Drug	Search by Gene	Upload Format	Download Format	Provided Database	Microarray Data Overlay	External Links
Biocarta www.biocarta.com	PD	Pictorial	√		√	None				Genecard, KEGG, NCBI, SwissProt, Unigene
BioCyc http://biocyc.org	PD	Process Diagram	√		√	None	BioPax, SBML	√		IntEnz, IUBMB Enzyme Nomenclature, KEGG, NCBI, PRIAM, PUMA2, SwissProt, UniProt
Cytochrome P450 Drug Interaction Table http://medicine.iupui.edu/flockhart/table.htm	PK	List				None				NCBI
Cytoscape www.cytoscape.org		Process Diagram				SIF, GML, BioPax (via plugin)	GML		√	BIND, cPATH, GO, KEGG, TRANSFAC
GenMAPP http://www.genmapp.org		Process Diagram	√		√	KEGG	Images only: PDF, BMP, JPEG	√	√	Cytoscape, ENSEMBL, GO, KEGG, NCBI, NetPath, SOURCE, Uniprot, UCSC Golden Path
Ingenuity www.ingenuity.com		Process Diagram	√	√	√	Excel, tab-delimited text file	Images only: JPEG, WMF, TIFF	√	√	NCBI
KEGG http://www.genome.jp/kegg	PD	Process Diagram	√			None	XML, HTML	√	via GenMAPP	IUBMB Enzyme Nomenclature, NCBI, SwissProt, UniProt
PATIKA http://www.patika.org		Process Diagram				None	PDF, SVG, BioPax, SBML	√		GO, HPRD, InAct, NCBI, Reactome, UniProt
PharmGKB www.pharmgkb.org	PD PK	Pictorial	√	√	√	None	Adobe Illustrator, Excel	√		GDB, GenAtlas, Genecards, GO, MutDB, NCBI, PromoLign, Source, Swissprot, UCSC Golden Path
Reactome http://www.reactome.org	PD	Process Diagram, Pictorial	√			None	BioPax, SBML, PDF, Cytoscape, Protégé, SVG	√	√	ENSEMBL, KEGG, NCBI, UCSC Golden Path, UniProt
STKE Connections Map http://stke.sciencemag.org	PD	Pictorial	√			None	XML by request			NCBI

tabolism to study. Colors highlight experimentally confirmed reactions as compared to manually inferred and electronically inferred reactions. The user clicks down or zooms in to get to the specific reaction. The metabolic reactions are mostly process diagrams depicted with boxes and arrows, representing reactions and molecules including proteins, substrates and products. Signaling pathways and similar re

actions are depicted pictorially. Reactions are coded with the gene ontology terms that can aid in determining whether gene products are found in the same cellular component and thus direct protein-protein interaction may be possible. The proteins are listed by multiple names and clicking on them drills down to references and links to other resources (listed

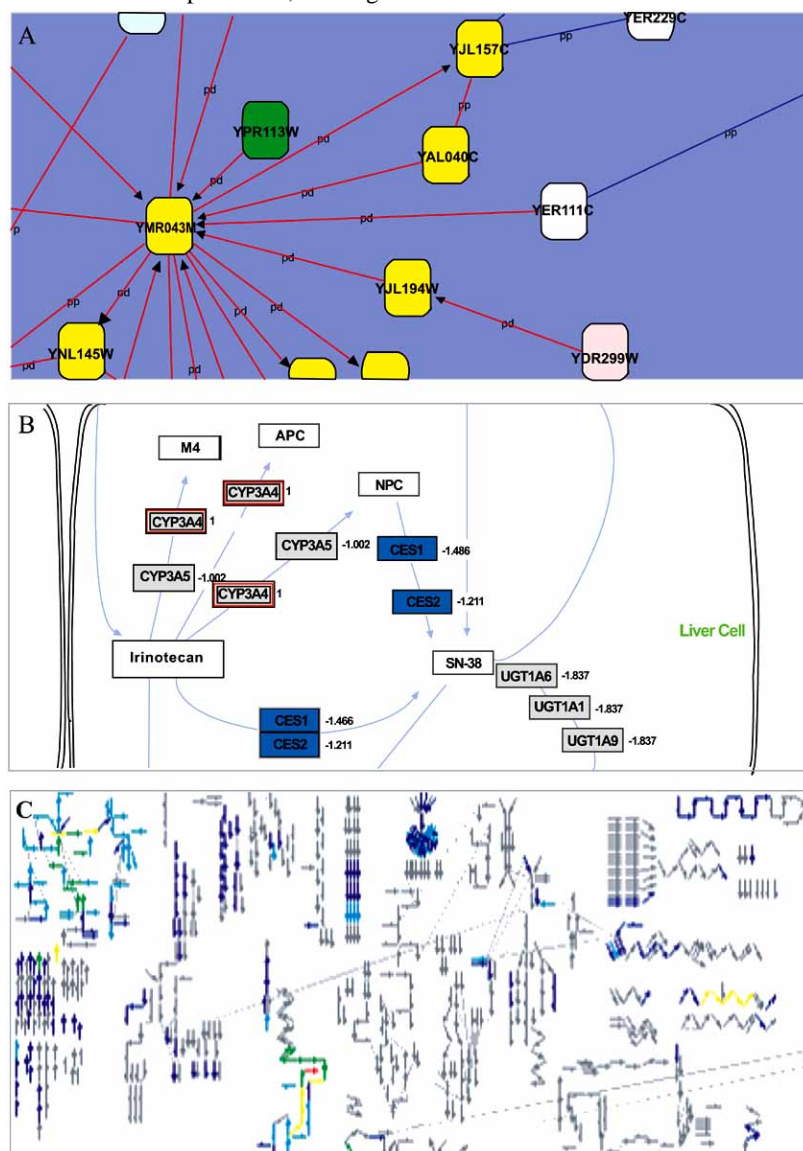
in Table 1) including Entrez gene where the corresponding gene name and symbol can be found. Searching by gene symbol or drug name in Reactome, as with KEGG, will not be successful as a metabolic pathway must be known to use this database. The Reactome resource offers the Skypainter tool for overlaying microarray and other data types onto its pathways (Fig. 2).

The BioCyc and HumanCyc pathway resources are part of the MetaCyc suite of databases developed at SRI International [Caspi *et al.*, 2006] (Fig. 1B). The databases were constructed using the Pathologic component of the Pathway Tools software that is available freely from the website [Karp *et al.*, 2002]. HumanCyc was generated by combining the human gene sets from the Ensembl and LocusLink databases and inferring the metabolic pathways [Romero *et al.*, 2004.]. It does not contain signaling pathways. Some data types, such as transport and cellular compartments, although

supported by the software, have yet to be added. The database can be queried by genes, proteins, pathways, EC reaction accession Ids and compounds, however this does not include drugs. The pathways can be downloaded *via* BioPAX and SBML (discussed in the section on Pathway Standards) as well as their own Pathway Tools format.

In addition to the metabolic pathway resources, there are a number of websites that provide more diagrammatic representations of pathways of relevance to pharmacogenomics.

The signal transduction knowledge environment at Science magazine, STKE Connections Map, has community submitted pathways that cover a broad range of cell biology signaling pathways. STKE Connections Map has 44 canonical pathways and 28 specific pathways that may be defined in a particular organism, cell type or condition. Pathways are depicted as colored nodes and edges. The nodes represent



**Fig. (2).** Screenshots from Pathways Resources illustrating different ways to analyze and display data using pathways: (A) section from a Cytoscape display of yeast genes and microarray data, (B) section from a GenMAPP display of the PharmGKB irinotecan pathway with superimposed expression data, (C) Reactome reaction map painted with data using Skypainter.

components, the color and shape of which represent subcellular localization and component type. Component types include proteins, complexes, DNA, metabolites and non-protein ligands. Gene information is often contained within the descriptive text and is not searchable. There are currently no drug pathways described as such amongst the specific pathways. However there are some pathways, such as the adrenergic receptor, pathway [Xiang and Kobilka, 2005] which are useful for PD for beta adrenergic agonists and beta blockers, if the already user knows the pathway mechanism of the drug of interest.

Biocarta pathways were originally developed as way to highlight commercial reagents for cell signaling research such as antibodies and inhibitors. They have colorful displays with a wide variety of node shapes to indicate different types of proteins. The layout is more like a figure from a paper and less like a process diagram, with cellular components and directional space (Fig. 1C). Biocarta provides generalized pathways and human and mouse specific views for a diverse array of disease and signaling pathways. Pathways are submitted by the community and moderated by scientists. Biocarta does not have any drug specific pathways although it does have a Multi-Drug Resistance Factors pathway that depicts some of the ABC transporters.

Pathway resources for PK are less common. Individual papers in the literature can provide highly reliable relationships and elegant diagrams regarding specific drugs. However, it is necessary for the user to find these in the myriad of articles within the published literature, which is not easy given that one must examine the full text of the article to locate, view and assess the pathway. Even obtaining a list of the superset of all of the genes commonly involved in PK, i.e. well known drug-metabolizing enzymes, phase 1 and phase 2 and transporters is not trivial.

Currently the only free resource for drug PK pathways is the Pharmacogenetics and Pharmacogenomics Knowledge Base, PharmGKB [Klein *et al.*, 2001]. PharmGKB has both PK and PD pathways that are drug centered and gene based. These pathways are proposed by panels of experts from the Pharmacogenomics Research Network (PGRN) who discuss and submit the data, which is then hand-curated and formed into the final pathway. The graphic representations are full color figures that are highly interactive (Fig. 1D). The nodes are drugs and genes that link to drug and gene pages in the PharmGKB. The arrows are processes that can be annotated with pop-up graphs of gene variation plotted against phenotypes to support the relationship. The specific literature references for each arrow can be downloaded as an Excel file. Compiling the pathways is an intensive manual process and this resource is currently limited to 22 pathways. However the representation and creation of more PK and PD pathways is an ongoing process. The gene centered searching provided by PharmGKB makes it easy to find candidate genes and SNPs for both PK and PD, with links from the homepage listing PD genes and PD variants, or PK genes and PK variants based on their associated literature. The Annotated Pgx gene variant project, to highlight well characterized variants of pharmacogenetic importance, will aid those searching for functional SNPs in genes that are unfamiliar to them. PharmGKB also provides annotations of drugs that have a

documented relationship with a gene, and the link to the publication documenting the details. This allows genecentered users to assess drugs of potential pharmacogenomic significance to plan a genotype to phenotype study design.

Some additional resources that may be useful for selecting candidate PK genes for drug studies are the Cytochrome P450 Drug-Interaction Table provided by Flockhart *et al.*, which lists commonly prescribed drugs and the predominant CYP gene involved in their metabolism. Variation in the CYP genes is comprehensively covered by the Human Cytochrome P450 (CYP) Allele Nomenclature Committee website [Ingelman-Sundberg, 2002], in the NAT family of PK genes at the Arylamine N-Acetyltransferase (NAT) Nomenclature website [Vatsis *et al.*, 1995] and in drug transporter genes at the UCSF Pharmacogenetics of Membrane Transporters website [Stryke *et al.*, 2003].

There are several software products that can be used to integrate microarray and expression data with pathway information. Cytoscape is a commonly used, freely available program for pathways based analysis of microarray data. It allows for integration of multiple pathways and hierarchical organization of pathways [Shannon *et al.*, 2003]. A user can modify pathways by dragging and dropping nodes and edges with the mouse, and can even build pathways in this manner. However, it is quicker to create a ready-made pathway *via* file upload (see Table 1 for details). Cytoscape enables users to visualize microarray expression levels over a pathway, and the use of the data as a filter to select a subset of genes in the pathway. First, a pathway, or network, needs to be loaded into the Cytoscape program. From there, expression data can be loaded onto the network *via* a text file. The format of the text file is straightforward, and requires one gene name and the corresponding expression values (or ratios) for each experiment to be listed on a single line, delimited by tabs or spaces. P-values or other measure of significance can be provided for each value if the user wishes. Different expression levels can then be represented as different attributes for the genes/nodes. Genes can be colored or shaped according to the expression level (Fig. 2A). They can even be colored along a gradient from red to green to coincide with array grids, allowing the user to see at a glance which genes in the pathway are being expressed and at what level.

GenMAPP is another popular free resource for drawing pathways and mapping microarray data onto pathways [Dahlquist *et al.*, 2002] (Fig. 2B). It offers several functions in addition to the creation of pathway images and can help the user find global associations involved in an expression dataset. Pathways are constructed through a graphical interface or can be selected from a set of community shared MAPPs. Using the provided "Gene Databases", information can be found about connections between a gene of interest and other genes, which can then be used to create pathway drawings. An advantage of the GenMAPP application is that it offers a step-by-step web-based tutorial for the first-time user. However, the application is only available for use with Windows platforms. GenMAPP offers KEGG pathways and other pathways for visualization, as well as allowing the user to create their own pathway image. GenMAPP contains a module specifically designed for integration of expression data. The data is represented in the graphic by specific color-

ing of the nodes. The gene IDs used for the expression data must match the ID in the Gene Database the user downloaded. Gene Databases can be UniGene, Affymetrix, GenBank, etc. The user needs to determine the code for each gene ID system (Affymetrix, Unigene, SwissProt, etc.), also derived from the Gene Database. The file format for expression data requires that each row contain the gene ID and the system code. Optional fields include the control average and standard deviation, the expression average and standard deviation, p-value and quality. The user controls the colors used to visualize the expression data on the pathway image by setting multiple parameters in a "wizard" window. GenMAPP allows for very customized views of the array data.

PATIKA, Pathway Analysis Tools for Integration and Knowledge Acquisition, also has its own database of information focusing on human pathways integrated from a variety of sources including Entrez, UniProt, PubChem, GO, InAct and Reactome. Users can query and access data from a web interface, and build pathway images as a result of the queries. PATIKA exports pathway information in BioPAX and SBML formats. It allows for printing of pathway drawings, and saves drawings as JPEG or SVG images.

The Ingenuity Pathways Analysis software, although proprietary, is worth noting for two reasons (1) the knowledge base that it is built on and (2) it is currently the only system to support incorporation of high throughput SNP data. The knowledge base was built from curated information from the literature. They developed an ontology that represents over 300,000 classes of biological objects including genes, proteins, drugs, cellular processes and diseases. This allows for aggregation of data into a powerful set of dynamically generated pathways. As with the other resources, expression data can be mapped onto the pathways. In addition, Ingenuity also supports analysis of Affymetrix Gene Chip arrays to show the pathways for the genes for which those SNPs map to.

## EXAMPLES OF SITUATIONS WHERE PATHWAYS RESOURCES CAN BE USED TO AID PHARMACOGENOMICS STUDIES

### Use Case #1

An investigator has observed a variable drug response in a population for a drug with no previous reports of pharmacogenomic significance. The investigator knows the endogenous signaling pathway that the drug acts on but has limited information on the PK of the drug. They wish to identify a set of reliable candidate genes and SNPs to test in their population.

#### *Suggested Resources*

Reactome, KEGG, HumanCyc, Cytochrome P450 Drug Interaction Table and PharmGKB. To find candidate PD genes, the investigator could perform a search on Reactome, KEGG or HumanCyc for the endogenous pathway of the drug action. This would give lists of protein names and external resources such as Entrez gene at NCBI that have the HGNC gene symbol needed to definitively identify the gene. To find candidate PK genes, the investigator might look at the Cytochrome P450 drug interactions table to examine if it has a listed interaction with a CYP gene or if one can be in-

ferred by a known interaction with another drug. The investigator would then go to PharmGKB with their candidate genes and look at Annotated PGx Gene Information or variants listed for those genes, clicking down to view assay and primer information and frequencies in different populations. Thus they can compile a shortlist of SNPs that occur at observable frequencies in their population and have known effects on function in other contexts.

### Use Case #2

An investigator has microarray data from drug treated cells. They know of several genes and SNPs already implicated in the pharmacogenomic action of this drug but wish to examine all possible genes involved and identify new candidates.

#### *Suggested Resources*

Reactome, GenMAPP, Cytoscape, KEGG, PharmGKB and Ingenuity. The investigator could use Reactome to get a global view of all of the metabolic pathways touched by their microarray data using the Skypainter tool. It is also possible to drill down and look at each pathway individually. They could also import all of the PD related pathways that include the known genes from KEGG, or the PK pathway for their drug if available from PharmGKB, into Cytoscape or GenMAPP and visualize the important genes in their dataset. Ingenuity also can provide this kind of analysis for a subscription fee.

### Use Case #3

An investigator is interested in a particular gene and has identified variation in different populations. They wish to find out what significance the variation may have, if any, on drug response.

#### *Suggested Resources*

PharmGKB, KEGG, Reactome and HumanCyc. The investigator might start at PharmGKB to see if there are any phenotype files that examine their gene of interest in the context of drug treatment or literature annotations linking their gene of interest to a drug or drug pathways associated with that gene. If this was unsuccessful they could check for any pathways in KEGG, Reactome or HumanCyc that contain the protein coded by their gene of interest that might be targeted by a drug (this would require some knowledge of pharmacology). They might take related proteins from pathways in KEGG, Reactome or HumanCyc and go back to PharmGKB and find the gene pages and see if those have associations to drugs or diseases. This may give them a short list of drugs to test with their variants in an expression system.

## PATHWAY STANDARDS

As discussed, there are many pathway databases in existence, each with its own unique representation of the data. The many different data representations can make data sharing of knowledge between applications and databases tedious and cumbersome. Often, it requires human intervention. There are several active initiatives to address the problem of data transfer. Each one aims to provide a set of standard terminology and formatting in order that pre-written computer

programs can transfer the information from the standard format into the local application format for use, and vice versa. The majority of standards being developed use XML, eXtensible markup language, format. XML uses element tags that surround the data and are organized in a hierarchy. The format allows for the easy addition of new data types making it a good format for capturing diverse types of pathway information and can allow standards to use modules from each other. Formats of note with respect to pharmacogenomics pathways are PSI-MI, CellML, SBML and BioPAX, with SBML and BioPAX being the most commonly used of this group. An excellent review by Stromback and Lambrix compares the components and relative strengths and weaknesses of PSI-ML, SBML and BioPAX from a bioinformatics perspective [Stromback and Lambrix, 2005].

PSI-MI, Proteomics Standards Initiative, Molecular Interaction ([psidev.sourceforge.net/mi/xml/doc/user/](http://psidev.sourceforge.net/mi/xml/doc/user/)), is designed specifically to describe protein-protein interactions [Orchard *et al.*, 2004]. It has strengths in capturing experimental data, particularly for mass spectrometry. Genes, drugs and DNA are not currently represented making it unsuitable for pharmacogenomic pathways at this point. However, with those additions, and the expansion of use of proteomics and metabonomics in pharmacogenomics research, this format may become more important.

CellML ([www.cellml.org](http://www.cellml.org)) is an XML-based format, developed primarily by the Bioengineering Institute at the University of Auckland [Lloyd *et al.*, 2004]. The format was designed for describing mathematical models of cellular and sub-cellular processes.

SBML, Systems Biology Mark-Up Language ([sbml.org](http://sbml.org)), is another XML-based format. Its focus is on quantitative biochemical network models and it is intended to be able to describe any chemical reaction. It is now an effort by the systems biology community in general. There are many applications that can accept and export data in this format, though it is generally geared toward simulation/analysis programs for biochemical networks. A recent paper by Oda *et al.*, describes the use of SBML to capture the complex and important signaling pathway of the epidermal growth factor receptor, EGFR [Oda *et al.*, 2005]. This is of note, not only because of the targeting of new anti-cancer therapies to this pathway, but also because of the way that it includes detailed information about key residues within the proteins in the pathway and their phosphorylation status.

The BioPAX Initiative ([www.biopax.org](http://www.biopax.org)) was created specifically to enable the exchange of pathway information between databases. BioPAX Level 2 enables molecular interaction representation, hierarchical pathways and basic experimental descriptions, and it incorporates PSI-MI. It is quickly being adopted by other pathway databases, including MetaCyc and Reactome, and promises to be backwards compatible. It is being developed to become the primary pathway exchange format. However, as with the other standards described above, its use for pharmacogenomics at present is still limited due to the lack of terms for genes, SNPs and drugs, although it recognizes the need to develop the term gene in later releases.

## FUTURE DIRECTIONS

Increasing quantities of data will become available as the cost of high throughput methods decreases. This will populate the databases more fully but also add to the challenges of integrating data and incorporating new methods and data types. In order to do this comprehensively for pharmacogenomics there needs to be a greater effort to agree on and use controlled vocabularies, and ways of describing the data, so that data can be aggregated and integrated and remain useful. It also increases the demand for novel ways to interpret data and pathways resources aim to provide this.

The pathways resources represent a relatively new area of interdisciplinary research and as such there are limitations with what they can currently provide as discussed in this review and shown in Table 1. However, the resolution and adoption of standards will facilitate better exchange between the different sources of data. This will allow researchers interested in a particular disease or discipline to find pathways of interest, connected *via* the component genes or phenotypes, that were developed by researchers from a completely different focus or therapeutic area. In order for this to be of value for pharmacogenomics the standards needs to be further developed for describing, storing, displaying and exchanging data about gene variation and drugs in a pathway context.

## ACKNOWLEDGEMENTS

This work was supported by U-01 GM-61374, funded by the National Institute of General Medicine Sciences.

The Angiotensin II mediated activation of JNK Pathway *via* Pyk2 dependent signaling pathway, shown in part in Fig. (1C) was submitted to Biocarta by Michael Shih, PhD. The PharmGKB Irinotecan PK pathway shown in part in Fig. (1D) was authored by C. F. Thorn, M.W. Carrillo, J. Ramirez, S. Marsh, E.G. Schuetz, M.E. Dolan, F. Innocenti, M.V. Relling, H.L. McLeod and M.J. Ratain.

## REFERENCES

- Brown, K. S.; Cook, M.; Hoess, K.; Whitehead, A. S. and Mitchell, L. E. (2004) Evidence that the risk of spina bifida is influenced by genetic variation at the NOS3 locus. *Birth Defects Res. A Clin. Mol. Teratol.* **70**(3), 101-106.
- Caspi, R.; Foerster, H.; Fulcher, C. A.; Hopkinson, R.; Ingraham, J.; Kaipa, P.; Krummenacker, M.; Paley, S.; Pick, J.; Rhee, S. Y.; Tissier, C.; Zhang, P. and Karp, P. D. (2006) MetaCyc: A multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Res.* **34**, D511-D516.
- Dahlquist, K. D.; Salomonis, N.; Vranizan, K.; Lawlor, S. C. and Conklin, B. R. (2002) GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways. *Nat. Genet.* **31**, 19-20.
- Dogrusoz, U.; Erson, E. Z.; Giral, E.; Demir, E.; Babur, O.; Cetintas, A. and Colak, R. (2003) PATIKAwEB: a Web interface for analyzing biological pathways through advanced querying and visualization. *Bioinformatics* **22**(3), 374-375.
- Freimuth, R. R.; Xiao, M.; Marsh, S.; Minton, M.; Addleman, N.; Van Booven, D. J.; McLeod, H. L. and Kwok, P. Y. (2005) Polymorphism discovery in 51 chemotherapy pathway genes. *Hum. Mol. Genet.* **14**(23), 3595-3603.
- Ingelman-Sundberg, M. (2002) Polymorphism of cytochrome P450 and xenobiotic toxicity. *Toxicology* **181-182**, 447-452.
- Jensen, L. E.; Etheredge, A. J.; Brown, K. S.; Mitchell, L. E. and Whitehead, A. S. (2006) Maternal genotype for the monocyte chemoattractant protein 1 A(-2518)G promoter polymorphism is associated

- with the risk of spina bifida in offspring. *Am. J. Med. Genet. A* **140**(10), 1114-1118.
- Joshi-Tope, G.; Gillespie, M.; Vastrik, I.; D'Eustachio, P.; Schmidt, E.; de Bono, B.; Jassal, B.; Gopinath, G. R.; Wu, G. R.; Matthews, L.; Lewis, S.; Birney, E. and Stein, L. (2005) Reactome: a knowledge-base of biological pathways. *Nucleic Acids Res.* **33**, D428-D432.
- Kanehisa, M.; Goto, S.; Hattori, M.; Aoki-Kinoshita, K. F.; Itoh, M.; Kawashima, S.; Katayama, T.; Araki, M. and Hirakawa, M. (2006) From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res.* **34**, D354-D357.
- Karp, P. D.; Paley, S. and Romero, P. (2002) The Pathway Tools Software. *Bioinformatics* **18**, S225-S232.
- Klein, T. E.; Chang, J. T.; Cho, M. K.; Easton, K. L.; Fergerson, R.; Hewett, M.; Lin, Z.; Liu, Y.; Liu, S.; Oliver, D. E.; Rubin, D. L.; Shafa, F.; Stuart, J. M. and Altman, R. B. (2001) Integrating genotype and phenotype information: an overview of the pharmGKB project. *Pharmacogenomics J.* **1**, 167-170.
- Lloyd, C. M.; Halstead, M. D. B. and Neilsen, P. F. (2004) CellML: its future, present and past. *Prog. Biophys. Mol. Biol.* **85**(2-3), 433-450.
- Oda, K.; Matsuoka, Y.; Funahashi, A. and Kitano, H. (2005) A comprehensive pathway map of epidermal growth factor receptor signaling. *Mol. Syst. Biol.* **1**, E1-E17.
- Orchard, S.; Taylor, C. F.; Hermjakob, H.; Weimin-Zhu; Julian R. K. Jr. and Apweiler, R. (2004) Advances in the development of common interchange standards for proteomic data. *Proteomics* **4**(8), 2363-2365.
- Romero, P.; Wagg, J.; Green, M. L.; Kaiser, D.; Krummenacker, M. and Karp, P. D. (2004) Computational prediction of human metabolic pathways from the complete human genome. *Genome Biol.* **6**, 1-17.
- Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Nada, A.; Schwikowski, B. and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498-2504.
- Stromback, L. and Lambix, P. (2005) Representations of molecular pathways; an evaluation of SBML, PSI MI and BioPAX. *Bioinformatics* **21**(24), 4401-4407.
- Stryke, D.; Huang, C. C.; Kawamoto, M.; Johns, S. J.; Carlson, E. J.; Deyoung, J. A.; Leabman, M. K.; Herskowitz, I.; Giacomini, K. M. and Ferrin, T. E. (2003) SNP analysis and presentation in the Pharmacogenetics of Membrane Transporters Project. *Pac. Symp. Biocomput.* 535-547.
- Vatsis, K. P.; Weber, W. W.; Bell, D. A.; Dupret, J.-M.; Price Evans, D. A.; Grant, D. M.; Hein, D. W.; Lin, H. J.; Meyer, U. A.; Relling, M. V.; Sim, E.; Suzuki, T. and Yamazoe, Y. Nomenclature for N-acetyltransferases. *Pharmacogenetics* **5**, 1-17.
- <http://www.louisville.edu/medschool/pharmacology/NAT.html>
- Xiang, Y. and Kobilka, B. K. Adrenergic Pathway. *Sci. STKE* (Connections Map), [http://stke.sciencemag.org/cgi/cm/stkecm;CMP\\_8762](http://stke.sciencemag.org/cgi/cm/stkecm;CMP_8762).

---

Received: 28 June, 2006

Revised: 14 September, 2006

Accepted: 18 September, 2006