Published in final edited form as:

Curr Opin HIV AIDS. 2012 January ; 7(1): 10–16. doi:10.1097/COH.0b013e32834e390a.

An open-ended plea for the development of a global database of HIV vaccine responses

Peter Wilkinson^a, Abdelali Filali-Mouhim^a, Shuzhao Li^b, Jeffrey Ahlers^a, John Schatzle^a, Bali Pulendran^b, Rafick-Pierre Sekaly^a, and Mark J. Cameron^a

^aVaccine and Gene Therapy Institute of Florida, Port Saint Lucie, Florida

^bEmory Vaccine Center, Atlanta, Georgia, USA

Abstract

Purpose of review—The purpose of this review is to describe a critical need of the HIV research community for a globally accessible database of HIV vaccine responses that stores data from multiple assay platforms in the form of lists of correlates of immune protection and vaccine efficacy. This is not a detailed review but a first step toward developing a dialogue among investigators and funding organizations to build upon existing resources to efficiently develop a HIV vaccine response and correlates database. We also discuss examples of databases that complement our needs and could be integrated into our proposed database requirements.

Recent findings—Several vaccine-related databases that store information at the study level currently exist, however, at the present time, a correlates of immune protection database does not exist.

Summary—Here, we discuss the scientific climate surrounding HIV vaccine development with the evolution of systems biology approaches, the problems at hand for analyzing and harmonizing datasets generated from preclinical and clinical studies, and the curation and accessibility of useful information to model outcomes. We also compare key database requirements of a few existing globally accessible databases and provide several illustrative correlate database submission and utilization examples.

Keywords

correlates; database; HIV; immune responses; vaccine

INTRODUCTION

We live in an era that has seen impressive advances in the management of HIV pathogenesis in AIDS patients. Infection rates have been reduced in those at risk via blood donor and blood product screening, education, preventive measures that include condom use, circumcision, pre-exposure prophylaxis, as well as the initiation of antiretroviral therapy upon exposure [1]. In HIV-infected individuals, advances in antiretroviral therapy (ART) have reduced the suffering caused by the illness, boosted immunological status, and increased life expectancy dramatically, whether initiated during early stages of infection or

Copyright © Lippincott Williams & Wilkins.

Correspondence to Mark J. Cameron, Vaccine and Gene Therapy Institute of Florida, 11350 SW Village Parkway, Port Saint Lucie, FL 34987, USA. Tel: +1 772 345 4789; fax: +1 772 345 3675; mcameron@vgtifl.org.

Conflicts of interest

There are no conflicts of interest.

during progression to AIDS [2,3⁴]. HIV infection, however, is a global problem and access to ART is still lacking for many [1]. Moreover, ART resistance, side effects of prolonged drug usage, and progressive immune failure conspire to undermine treatment efficacy leading to AIDS-related complications such as neuropathy, cancers, and opportunistic infections. Therefore, development of a safe and effective vaccine against HIV remains the highest priority (see [4–9] for recent reviews).

Enormous amounts of resources and the efforts of diverse research disciplines have been dedicated to this goal, yet a preventive vaccine for HIV is still beyond reach. It is clear that the current model for vaccine development, in which only four human vaccine efficacy trials have been conducted in the past three decades, and in which vaccine candidates are evaluated and advanced mainly by immunological readouts that focus on quantifying a small number of parameters is insufficient to decipher the complexity of host responses required to provide protection against chronic HIV infection. More specifically, traditional immune-monitoring technologies and data treatments have not been able to capitalize on the large repositories of datasets from vaccine trials in humans and animal models that currently exist. Only by integrating the data by bioinformatics analyses can we discern the successes and failures that will enable researchers to more intelligently design future vaccines, especially given the complexities associated with generation of a long-lived immune response to an individual vaccine. Current evaluation of vaccine efficacy is limited by the data available for analysis and often relies on gross clinical endpoints. Although this may be sufficient when protective correlates are well defined, e.g., neutralizing antibodies, this approach falls short when the true protective mechanisms are not readily apparent or occur beyond a given time frame of analyses. In these situations, systems biology can provide an unbiased assessment of the global immune response to a vaccine and the ability to integrate correlate data obtained from multiple vaccine studies will provide a platform to systematically move forward.

PROBLEM

Although raw genomic data are widely available in public repositories such as the Gene Expression Omnibus (GEO) [10¹¹], an understanding of how those data actually correlate with direct clinical outcomes is often lacking, leading to a lack of integration and synergy between genomic discovery and vaccine development in the broader scientific community. A wide array of experimental tools is being employed in systems biology. These tools include microarrays, next generation sequencing, multiplex flow cytometry, mass spectrometry, multiplex cytokine/chemokine assays, high throughput quantitative PCR, siRNA screens, and increasingly sophisticated single cell and antibody assays. The use of transcriptional profiling as a tool to evaluate immune responses can provide a systems analysis of thousands of parameters simultaneously to characterize immune profiles. The distribution and archiving of data is a central pillar of a global scientific enterprise, and the large, complex datasets obtained postsystems assay pose a new challenge to the field of HIV vaccinology.

SOLUTION

A publicly available database to store and query the correlate outputs of systems biology approaches may help to identify more biomarkers and provide more precise and deeper insight into vaccine-induced protection mechanisms. This potential for novel insight and discovery of key regulators of immunity via a comprehensive systems approach, involving a publicly available integrated correlate database, serves as the basis for this commentary. This goal is included in the recommendations of the Global HIV Vaccine Enterprise and is the top priority of other clinical networks such as the Immune Tolerance Network,

ImmGenProject, NIH's Center for Human Immunology, Autoimmunity, and Inflammation (CHI), and Program in Systems Immunology and Infectious Disease Modeling [11–15].

In addition, with the current restructuring of the NIAID HIV Vaccine Trials Network with the potential to expand clinical trials and 'first in human' immunotherapy trials involving smaller numbers of volunteers, it is imperative to begin developing the infrastructure and resources necessary toward achieving this goal.

The studies outlined here validate the use of genomics and multiplatform approaches to measure and describe a large number of potential prognosticators of vaccine efficacy. The increasing amount of high volume datasets places a mandate for clinical networks to develop a computational resource infrastructure for summarizing signatures derived from data. Such a mandate has proven invaluable in the case of the Los Alamos National Laboratory (LANL) HIV [16] database, when the collection of viral sequences became an integral part of understanding diversity and viral evolution in relation to host immune responses and in designing vaccines that could provide global coverage of contemporaneous circulating virus. Now we have reached a similar critical point as systems biology joins forces with HIV vaccine research in an effort to identify correlates of immune protection. An integrated database and repository of information from preclinical and clinical trials will permit a more rapid and sensitive evaluation of vaccine strategies and pave the way for more informed design of next generation vaccine regimens. This approach is not without logistical challenges in that it relies on the sensitivity of individual laboratories and centers to share expression data, flow, protein, or other type of profiling information, and comparisons with previous data-sets made publicly available. The development of a comprehensive, publicly accessible database designed to develop models of protection and validate HIV vaccine design and delivery strategies, based upon signatures of protection, will be a major achievement attributable to a multidisciplinary HIV research community.

The main purpose of this database will be the comparative analysis of multiple vaccine studies with regards to clinical attributes and fully annotated and analyzed laboratory datasets. The database would not store or archive raw data. Rather, this database would contain datasets that are readily associated with specific clinical pictures such as:

1. description of the investigation,
2. vocabulary-controlled annotation of sample type,
3. de-identified vocabulary-controlled annotation of sample donor,
4. detailed description of the cohort,
5. statistics describing the effect/condition being measured, and
6. comprehensive lists of immune parameters with associated statistical data.

DATABASE DEPLOYMENT

The deployment of this database should be carried out in three stages:

1. Establish a repository of machine-readable data and annotation.
2. Develop a graphical user interface (GUI) that will facilitate usage by all researchers.
3. Develop integrative analysis tools.

Repository design

The globally accessible correlate database should be created using current Web 2.0+ based technologies, similarly to what other groups described below are incorporating today and apply the following criteria.

First, the central dogma should be firmly established and kept simple. Current databases from NCBI described recently in detail in 2011 [10¹¹] and European Bioinformatics Institute (EBI) for example, offer several public archival databases that have separate specific scopes, such as for sequence submissions, genomes, SNPs, proteins, enzymes, chemicals, and taxonomy.

Public databases also require a submission process that includes tools and format conventions to follow, as well as screening criteria for including data. Currently, database inclusion is largely limited to data that are a part of published findings, and most lack real statistical criteria for the submission. Recently, however, BioModels [17], an EBI database for quantitative kinetic models, includes a submission process that includes manual verification, as well as simulation tests.

The database should have several modes of accessibility and provide both a secure portal for experimentalist groups and public portals for general query on published material, as well as the means for storing and automatically updating queries when content changes. Currently, most databases offer web-based, as well as file-based, modes and, more recently, some have introduced portal features such as the recently revamped myNCBI [18] with reengineering of MeSH terms, streamlining views, as well as automated updates to saved queries. Concerning secured submissions, GEO offers by default up to 1-year after submission to secure submitted data, renewable until a manuscript is published using the accession number. A mechanism for querying the secured submissions by submitters against the existing published data will be available, however other submitters will not see secured datasets.

Finally, tooling will be provided in order to handle more complex queries. For example, ONCOMINE [19], a collection of expert-curated cancer gene expression profiles, provides integrated data input, data analysis, meta-analysis and visualization tools. Also, the *Molecular Signatures Database* (MSigDB) [20²¹] is a collection of annotated gene sets for use with Gene Set Enrichment Analysis (GSEA), a pathway analysis technique for interpreting gene expression data. The VIOLIN network [21], is an example of a complete vaccine site that includes many of the criteria described above, including accessibility, a portal, and automated notifications. However, a detailed database of correlate signatures that is needed by researchers examining the detailed mechanisms of vaccine-mediated protection is not available here, nor are the tools for crosscomparison between different studies, especially between unpublished and published results. Integrating existing groups such as VIOLIN, and the HIV database at LANL, with a correlate signature database would provide a significant increase in scientific value by orders of magnitude.

A putative design of the database is shown in Fig. 1. All descriptive words stored in the database should be controlled by a vocabulary. Associated meta-data will describe the investigations and cohorts. The database should be centered around an effect/condition or statistic between two groups of features such as samples, phenotypes or experimental conditions. This could be a result of a univariate statistical test (*t*-test, Wilcoxon *U*-test, ANOVA test), a canonical pathways analysis (GSEA or commercial pathway software), a class prediction analysis or any other custom analysis requiring bioinformatics expertise. For example, an effect could be the differential expression between two time points, for example, baseline and 12 h post-vaccination in humans, or perhaps a regression between

expression of a gene and a continuous clinical readout. The correlate lists submitted should not be limited to lists of probes/genes from transcription experiments, but should also be derived from pathway analysis, sequencing, flow cytometry, proteomics, metabolomics, antibody assays, etc. An effect/condition should have a name, cite the source investigation, have a type such as contrast or regression, and the number of samples in each group. The list of correlates should be described with a type (such as microarray expression, RNA-seq, pathway, cellular phenotype), a name (gene, pathway, or cell name), statistic (P -value, t -test), differential sign (p for up, n for down), a fold change, and an ontological reference wherein its name is derived (such as a gene from HUGO or a pathway defined by Biocarta or KEGG, or MSigDB, etc.).

The technical specifications of the server will need to be determined and should include access control from the consortia level through the web via a secure portal, file repositories, a database engine, web applications, load balancing and process queue to support multiple queries and tools, connectivity with existing public databases, and backups.

Correlate signature submission process

A submission process should be established to curate relevant candidate (preliminary data) and established (published and statistically tested) signatures and store them in the database. The following criteria are suggested. Submissions should be made on the part of a vaccine trial consortia or group, wherein secure access can be limited to those members until publication and the consortia or group can actively use tools for discovery. There should be a classification ontology for the quality of the signatures, such as ‘unpublished’, ‘published’, ‘candidate’ or the nature of the evidence such as ‘statistically modeled’, ‘internally validated’ (within cohort), or ‘externally validated’ (separate cohorts). There should be contextual descriptive ontologies established for illustrating the context (disease, patient or animal type, outcome) to which this signature is applicable.

Statistics and power

Correlates defined by differential gene expression, pathway analysis (GSEA), or flow cytometry for example, should be manually reviewed by a curation committee composed of bioinformaticians and biological investigators that will decide whether the submitted data are to be included in the database or revised. There will be no limits placed on the correlate signature as it is uploaded, provided the appropriate contrast annotation and correlates are accompanied by a summary statistics and a P -value. However, what is uploaded and analyzed will likely be limited to either the top differential genes by fold change or P -value, or a selected subset that meets a certain statistical power based on the number of individuals per group. The lists of manually derived correlates should be limited in length to a small number, such as 25 or less, given the more anecdotal nature of the submission. An appropriate number will need to be derived from examining the current lists of validated signatures. It is important that submissions are supported by real evidence and include statistical modeling and functional validation in relevant ex-vivo and in-vivo settings [22].

User accessibility

The genomic correlate data acquired in this database should be aggregated via a web interface for easy access to the data by fellow researchers. A simple query interface should exist wherein terms are controlled by vocabularies.

Web services

The server should provide a Web service for accessing the data programmatically with similar functionality (described below) to the GUI. The published signatures from the

database should also be distributed in flat file format and the unpublished data should be secured as discussed above.

Consortia portal

Consortia should be able to register a named group with an identified individual to act as account manager and be able to add registered users to their group. The purpose of this would be to provide an active shared resource for use during trials and projects. Users from consortia should be required to create a free account and be able to save a history of their activities (such as submissions, queries and results of queries, in a similar fashion to how myNCBI is implemented).

Ad hoc queries

Queries should be able to be performed on a per gene (probe), pathway, vector, or immune response phase (innate, adaptive, acute, chronic, etc.) for casual users who are interested in retrieving, for example, documented measured effects wherein their chosen gene of interest is associated with an existing correlate signature. Users should also be able to submit gene lists and perform correlate enrichment using tools such as DAVID [23] with the result that their outputs would contain an overall functional assessment of the correlate signatures from the database and their query.

Complex queries

For a more comprehensive analysis, gene signatures and effects (conditions and states) will have to be submitted. Value-added tools, accessible only when the user formally uploads an entire signature with all necessary annotations and statistics, should be provided that can be used to query the correlate database and generate comprehensive comparative studies. Assembled correlates of protection and gene expression signatures will be linked to vaccine efficacy, and be usable as biological readouts. The head-to-head evaluation of vaccines in this manner will refine correlates, and provide information to the user allowing hypothesis generation, and a further understanding of mechanism of action of their vaccine regimens. Figure 2 contains an example of an inter-study comparison.

Tooling

Patterns of pathway deregulation across selected datasets will be provided following any query based on lists of genes or on a set of effect/condition contrasts. An example of such an analysis is shown in Figure 3 wherein each column represents a separate associated dataset contrast, and each row represents an upregulated (red) or downregulated (blue) pathway included in one or more contrasts.

The database, by design, may acquire pathway information associated with the correlate signature submitted by the users that will be controlled by existing vocabularies from different ontology sources (GO, KEGG, DAVID, Biocarta, or commercial pathway analysis software). For a submitted query based on a putative list of correlates, for example a list of genes, an enrichment analysis will be performed on existing correlate signatures, as well as the canonical pathways. Figure 4 exemplifies a top significant pathway analysis, as well as their member genes as would be generated.

Beyond genes and pathways

As lists of correlates may also be derived from other types of assays, tools should be provided that are able to generate reports and combine and contrast these signatures through the integration of the effects that are stored within the database. Tools that provide reports and enable patient/animal clinical profiling will also be very important.

CHALLENGES

It is clearly important to capture the results of correlation analysis and predictive modeling via a database server. The database, however, is an enabling infrastructure, with which anyone can perform secondary analysis and arrive at a consensus.

1. Given the great potential value of data integration from vaccine trials, a mandate of a central database that stores correlate information goes beyond the basic requirement of data robustness, availability and reproducibility. In the fields of genomics and functional genomics, the statistical power is not always clear at the time of initial study. It is thus important not only to accumulate published insights, but also to gain analytical strength by combining and contrasting data sets, from both humans and animal models. The correlate data will also need to be validated, either in other high throughput approaches or experimentally. The users will be able to provide these connections at upload if available or link externally to publications.
2. A correlate database may create concerns for users about intellectual property protection. The uploading of raw datasets prior to publication is now commonplace for GEO users. However, a correlate database user may be more reluctant to upload an intricately developed immunoprotective correlate signature than the raw data uploaded to GEO. Like GEO, providing secured services until the correlate signature is published as described above will mitigate these concerns.
3. Alignment with Institutional Review Boards and ethics guidelines aimed at protection of patient privacy will also be an important issue. As such, only de-identified annotations will be submitted, as well as annotations that cannot be used to derive the identity of the individual.
4. The database will grow out of a willing compilation of existing databases held by individual labs and consortia and/or GEO mining. The database need not be populated with only HIV-related information. Effective vaccines for other viruses (e.g., yellow fever) and bacteria will also be of a value in the foundation of the database. At the present, devotion of funds to this type of functional genomics initiative or bioinformatics endeavor is limited. Funding support needs to be secured realizing the long-term investment required and partnership with existing NIH-funded, publicly available database initiatives might be a workable option.

CONCLUSION

Although this commentary provides a basic road-map, an integrated database of HIV-vaccine-related correlate signatures will permit a more rapid and unbiased evaluation of licensed and experimental vaccines, and pave the way for more informed design of next generation vaccine regimens. To our knowledge, this approach is unique from existing databases that rely on the sensitivity of a particular laboratory's or consortium's correlate profiling and comparisons with previous datasets, and those datasets publicly available in the database. An integrated, accessible database that will be used by consortia during trials to uncover and/or validate gene signatures of protection will provide a fundamental tool for systematically evaluating vaccine candidates, as well as maximize resources that hasten the goal toward preventing HIV infection and disease.

Acknowledgments

Supported in part by the Bill and Melinda Gates Foundation Grant #38650 of The Collaboration for AIDS Vaccine Discovery (P.W., A.F.M., R.P.S., M.J.C.).

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

Additional references related to this topic can also be found in the Current World Literature section in this issue (pp. 86–87).

1. Girard MP, Osmanov S, Assossou OM, Kieny MP. Human immunodeficiency virus (HIV) immunopathogenesis and vaccine development: a review. *Vaccine*. 2011; 29:6191–6218. [PubMed: 21718747]
2. Schneider MF, Gange SJ, Williams CM, et al. Patterns of the hazard of death after aids through the evolution of antiretroviral therapy: 1984–2004. *AIDS*. 2005; 19:2009–2018. [PubMed: 16260908]
- 3▪▪. Palella FJ Jr, Delaney KM, Moorman AC, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV outpatient study investigators. *N Engl J Med*. 1998; 338:853–860. This paper describes an early highly cited study of the positive effects of antiretroviral therapy in HIV infected individuals. [PubMed: 9516219]
4. Ives NJ, Gazzard BG, Easterbrook PJ. The changing pattern of aids-defining illnesses with the introduction of highly active antiretroviral therapy (HAART) in a London clinic. *J Infect*. 2001; 42:134–139. [PubMed: 11531320]
5. Rollier CS, Reyes-Sandoval A, Cottingham MG, et al. Viral vectors as vaccine platforms: deployment in sight. *Curr Opin Immunol*. 2011; 23:377–382. [PubMed: 21514130]
6. Munier CM, Andersen CR, Kelleher AD. HIV vaccines: progress to date. *Drugs*. 2011; 71:387–414. [PubMed: 21395355]
7. Kim JH, Rerks-Ngarm S, Excler JL, Michael NL. HIV vaccines: lessons learned and the way forward. *Curr Opin HIV AIDS*. 2010; 5:428–434. [PubMed: 20978385]
8. Montagnier L. 25 years after HIV discovery: prospects for cure and vaccine. *Virology*. 2010; 397:248–254. [PubMed: 20152474]
9. Perrin H, Canderan G, Sekaly RP, Trautmann L. New approaches to design HIV-1 t-cell vaccines. *Curr Opin HIV AIDS*. 2010; 5:368–376. [PubMed: 20978376]
- 10▪▪. Sayers EW, Barrett T, Benson DA, et al. Database resources of the national center for biotechnology information. *Nucleic Acids Res*. 2011; 39:D38–D51. This study outlines the recent re-organization of NCBI resources and describes new data content, integration, and improvements to the NCBI portal. The extent of integration of content and the introduction of a portal continues to make the NCBI resource a critical resource for the research community. [PubMed: 21097890]
11. [[Accessed 9 November 2011]] Global HIV vaccine enterprise. <http://www.hivvaccineenterprise.org/scientific-strategic-plan>
12. [[Accessed 9 November 2011]] Immune tolerance network. <http://www.immunetolerance.org/professionals/research/lab-protocols>
13. Heng TS, Painter MW. The immunological genome project: networks of gene expression in immune cells. *Nat Immunol*. 2008; 9:1091–1094. [PubMed: 18800157]
14. Center for human immunology, autoimmunity and inflammation. [[Accessed 9 November 2011]] <http://www.nhlbi.nih.gov/resources/chi/>
15. [[Accessed 9 November 2011]] Laboratory of systems biology. <http://www.niaid.nih.gov/labsandresources/labs/aboutlabs/lbs/pages/default.aspx>
16. Lanl . [[Accessed 9 November 2011]] <http://www.hiv.lanl.gov>
17. Li C, Donizelli M, Rodriguez N, et al. Biomodels database: an enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Syst Biol*. 2010; 4:92. [PubMed: 20587024]

18. Giglia E. Pubmed in progress: latest changes in mesh and myNCBI. *Eur J Phys Rehabil Med.* 2011; 47:525–528. [PubMed: 21946409]
19. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, et al. Oncomine 3. 0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia.* 2007; 9:166–180. [PubMed: 17356713]
20. Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures data base (MSigDB) 3.0. *Bioinformatics.* 2011; 27:1739–1740. This manuscript describes the version 3 release of MSigDB that includes an increase in the number of available annotations sets over version 2. This collection is an excellent example of the importance of having well annotated gene sets that are publically accessible and are leveraged through tools like the GSEA application. [PubMed: 21546393]
21. Xiang Z, Todd T, Ku KP, et al. Violin: vaccine investigation and online information network. *Nucleic Acids Res.* 2008; 36:D923–D928. [PubMed: 18025042]
22. Trautmann L, Sekaly RP. Solving vaccine mysteries: a systems biology perspective. *Nat Immunol.* 2011; 12:729–731. [PubMed: 21772284]
23. David. [[Accessed 9 November 2011]] <http://david.abcc.ncifcrf.gov/>

KEY POINTS

- The development of a safe and effective vaccine against HIV remains the highest priority.
- Systems biology can provide an unbiased assessment of the global immune response to a vaccine.
- There is a lack of understanding of how widely available public data actually correlate with direct clinical outcomes, leading to a lack of integration and synergy between genomic discovery and vaccine development in the broader scientific community.
- The ability to integrate correlate data obtained from multiple vaccine studies will provide a platform to systematically move HIV vaccine development forward.
- A publicly available database to store and query the correlate outputs of systems biology approaches may help identify more biomarkers and provide more precise and deeper insight into vaccine-induced protection mechanisms.

FIGURE 1.

Content needed for the HIV database. Investigations (1), samples (2), donors (3), and cohort (4) descriptions are required. The six types of data needed are centered on the effects/conditions that are being studied (5). Descriptive meta-data for, as well as the correlate (6) lists with statistical data, are comprised of genes, pathways, cell type, and so on.

FIGURE 2.

Genes by study, comparative transcriptional profiling. This is a mock-up visualization using a type of heat-map for head-to-head comparison between studies, showing filtered interferon response genes (IRGs) that significantly correlate with a particular condition/effect. The color scale represents direction of gene expression regulation.

FIGURE 3.

Samples by pathways. This is a mock-up visualization of samples assayed over time from two studies mapped against canonical pathways, a heat color scale is used to denote the proportion of genes differentially regulated in a pathway versus the total number of potential genes in each pathway. Pathways could also be substituted for compartments or other functional groupings related to other data types.

FIGURE 4.

Genes by pathways. This is a mock-up visualization of top transcriptionally regulated genes at day 1 following vaccination mapped by top enriched pathways by $-\log P$ -value (pathway enrichment statistic). Each row represents an upregulated pathway in innate immunity for example, and each column represents a gene's expression direction. An upregulated (red) or downregulated (blue) gene may be included in one or more regulated pathway(s). This type of representation is not limited to transcription, but could map any list of correlates that can be associated with pathways, or any functional or compartmental groupings associated with those correlates.