Journal of Computer Science & Systems Biology - Open Access

Research Article JCSB/Vol.2 May-June 2009

Damages in Metabolic Pathways: Computational Approaches

Nruthya Lakshminarasimhan and Somnath Tagore*

Department of Biotechnology & Bioinformatics, Dr D Y Patil University,
Plot 50, Sec 15, CBD Belapur, Navi Mumbai 400614, India

*Corresponding Author: Somnath Tagore, Department of Biotechnology & Bioinformatics, Dr D Y Patil University, Plot 50, Sec 15, CBD Belapur, Navi Mumbai 400614, India, E-mail: somnathtagore@yahoo.co.in

Received May 15, 2009; Accepted June 23, 2009; Published June 24, 2009

Citation: Lakshminarasimhan N, Tagore S (2009) Damages in Metabolic Pathways: Computational Approaches. J Comput Sci Syst Biol 2: 208-215. doi:10.4172/jcsb.1000033

Abstract

Damage analysis in metabolic pathways is one of the most highlighted fields in systems biology area. Metabolic errors are studied with respect to inherited enzyme deficiencies. The damages usually occur due to changes in a particular enzymatic reaction or deletion of a step. Some enzyme disorders can also cause the particular reaction not to occur causing accumulation or reduced presence of a metabolite. This leads to dysfunction and disease state. Furthermore, the deletion of a path may change the end product of a reaction and can give rise to disorders. Thus, if the pathway leading to these changes can be determined, it would be easier to determine the disorders. Various statistical methods have been employed to determine the damages in the metabolic pathway. One such method is graph analysis technique. Here a graph is represented in the form of nodes and edges and the various properties may be determined. The wet lab results may be interpreted using statistics. Other studies of metabolomics can be performed using flux based analysis, fourier -transformation, Lp and power laws.

Keywords: Enzymatic changes; Enzyme deficiency; Graph analysis; Metabolic blocks; Metabolic errors

Introduction

Systems biology is an emerging multi-disciplinary field in which the behavior of complex biological systems is studied by considering the interaction of all cellular and molecular constituents rather than using a reductionist approach (Kirschner, 2005; Liu, 2005). As a part of systems biology, we would concentrate on damages in metabolic pathways. Damage in the metabolic pathway is an occurrence of irregularity of the pathway which causes a particular physiological change. Garrod et al., (1902) was the first to propose the existence of metabolic errors, based primarily on his studies of alkaptonuria. The chemical conversions in the body may occurs as sequential chemical steps, called a metabolic pathway. For example, the breakdown of tyrosine in the body involves a number of discrete chemical steps. Garrod's studies indicated that only one step in this metabolic pathway is defective, leading to accumulation of substances that cannot get past this block. Enzymes are also responsible for the catalysis of the biochemical reactions in metabolic pathways (Herzyk et al., 2004; Hwang et al., 2007). Analogous enzymes are able to catalyze the same reactions, but they present no significant sequence similarity at the primary level, and possibly different tertiary structures as well (Miranda et al., 2008).

Some enzyme alterations do not allow the particular reaction to take place, which causes accumulation or reduced presence of a metabolite which leads to dysfunction and disease state. The deletion of a path may change the end product which would cause disorders (Guo, 1986; Stuart et al., 2008). Damages can be caused by genetic alterations and mutations reduced consumption of metabolite or excessive presence of them. With the subsequent discovery of the genetic role of DNA and its structure as well as greater understanding of protein structure, the genetic bases for inherited changes in protein function were clarified (Kelekar et al., 2007; Kravchuk et al., 2007; Doetsch et al., 2004). Furthermore, changes in the structure of a gene are

likely to change the structure of the enzyme (Speed et al., 2002; Cordwell, 1999). Most random changes are likely to make the enzyme function less well, and many mutations completely abolish the activity of the enzyme (Speed et al., 2002; Cordwell, 1999; Graber et al., 2009). Also, some activities are recessive, since only one functional copy of the gene is often sufficient to carry out the required enzyme activity (Speed et al., 2002; Herzyk et al., 2004; Lorch, 1999). Similarly, many proteins consist of two or more distinct polypeptide subunits, each coded by a different gene. Mutations in any one of the genes may interfere with function of the enzyme, producing similar or indistinguishable phenotypes (Matschinsky, 1996; Cordwell, 1999).

Using graph analysis attempts have been made to study metabolic functions . (Herzyk et al., 2004; Valiente et al., 2004). The biological damages are generated in a metabolic network to experimentally determine the existence of an organism with damage. The enzyme might be altered in order to sum the importance of the enzyme in a particular metabolism (Schmitz-Peiffer, 2000). It was found that enzyme associated with high damage are involved in the production of compounds of small connectivity that connect important parts of the metabolism on the other hand, highly connected compounds tend to be redundant since they are produced by many reactions (Herzyk et al., 2004; Schmitz-Peiffer, 2000; Stuart et al., 2008). Using a graph analysis of its metabolism, the extent of the topological damage generated in the metabolic network by the deletion of an enzyme has been related to the experimentally determined viability of the organism in the absence of that enzyme. It is shown that the network is robust and that the extent of the damage relates to enzyme importance. It is also predicted that a large fraction (91%) of enzymes causes little damage when removed, while a small group (9%) can cause serious damage. Experimental results confirm that this group contains the majority of essential enzymes. The results may reveal a universal property of metabolic networks. (Lemke et al., 2004).

Types of Damages

Gene Expression Level

These include changes at the expression levels. Some genes are differentially expressed giving rise to faulty results. A very small change in the expression of a particular gene may have dramatic physiological consequences if the protein encoded by this gene plays a catalytic role in a specific cell function (Dudoit et al., 2002; Herzyk, 2004; Matschinsky, 1996; Krogh et al., 1996). From the biological prospective, even a very small change in the expression of a particular gene may have dramatic physiological consequences if the protein encoded by this gene plays a catalytic role in a specific cell function (Dudoit et al., 2002; Herzyk, 2004; Matschinsky, 1996; Schmitz-Peiffer 2000). Many other downstream genes may amplify the signal produced by a particular gene, increasing their chance of being selected by certain gene selection protocols. But, for a regulatory gene, however, the chance of being selected by such methods may diminish as one keeps hunting for downstream genes that tend to show much bigger changes in their expression (Speed et al., 2002; Herzyk et al., 2004; Gholson et al., 1982; Graber et al., 2009). As a result, the initial list of candidate genes may be enriched with many effector genes that do little to elucidate more fundamental mechanisms of biological processes. Furthermore, the characteristic of the regulatory genes is that their gene expression changes may be low, but they are highly correlated with the downstream highly expressed genes (Herzyk et al., 2004; Brown et al., 1997).

Mutations in the Gene

Mutations in the gene may be considered similar to gene expression changes, rather it can be considered as reason for differential gene expression. Genetic mutations cause changes in the gene that code for a particular protein participating in the pathway. If the gene participating in the functioning of the enzyme is mutated its role will change. The enzyme responsible for performing a particular catalytic reaction would fail, resulting in metabolite accumulation, causing disorders. (Cordwell, 1999; Doetsch et al., 2004). For example, autophagy is involved in cellular development and differentiation and may have a protective role against aging. It is also a form of cell death when allowed to proceed to completion and when cells unable to undergo apoptosis are triggered to die. It is often unclear whether it is directly involved in initiation and/or execution of cell death, or if it merely represents a failed or exhausted attempt to preserve cell viability (Abedin et al., 2007; Karantza-Wadsworth et al., 2007). Recent studies indicate that autophagy may play an active role in programmed cell death, but the conditions under which autophagy promotes cell death versus cell survival remain to be resolved. Defective autophagy has been implicated in tumorigenesis, as the essential autophagy regulator *beclin1* is monoallelically deleted in human breast, ovarian, and prostate cancers. *beclin1* is the mammalian ortholog of the yeast *atg6/vps30* gene, which is required for autophagosome formation. Beclin1 expression in human MCF7 breast cancer cells suppresses tumorigenesis. Recent studies revealed that autophagy enables tumor cell survival in vitro and in vivo when apoptosis is inactivated, as commonly occurs in human cancers (Kelekar et al., 2007; Kravchuk et al., 2007).There are many such

examples which can explain the gene mutation to be a reason for metabolic damage .

DNA Structure Alteration

Compounds that form chemical bonds with DNA represent several important classes of anticancer drugs (Doetsch et al., 2004). Increased removal or replicative bypass of DNA adducts have been found to be among the major mechanisms by which cancer cells reverse the effects of intra- and inter strand cross-linking drugs (Lorch et al., 1999; Doetsch et al, 2004). Resistance of cancer cells to chemotherapeutic drugs is a major problem encountered in the treatment of tumors. The basis for resistance has been studied extensively and can be drug-specific, or nonspecific as in the case of alkylating and platinum drugs (Levine et al., 1999; Albertson, 2006; Lukas et al., 2005; Doetsch et al., 2004). Despite the existence of a large body of data characterizing DNA cross-link repair, the specific components of DNA damage processing pathways and their individual contributions to the repair of cross-links are poorly understood (Lorch et al., 1999; Kelekar et al., 2007; Doetsch et al., 2004). Exploring the influence of these pathways on cell survival and defining their potential relationships with each other are crucial for understanding the modes of action of DNA cross-linking anticancer drugs used currently in the clinic (de Andrade, 2006; Lang, 1999; Levine, 1999; Albertson, 2006; Lukas et al., 2005).

Mitochondrial Damages

Mitochondria primarily generate ATP, promoting the closure of KATP-channel (ATP-sensitive K+ channels) and, as a consequence, depolarization of the plasma membrane (de Andrade et al., 2006). Mitochondrial defects in β-cells perturb glucose stimulated insulin secretion and might be associated with diabetes (Matschinsky, 1996). At the clinical level, one of the most direct indications that mitochondrial dysfunction could impair insulin release comes from the association between mutated mitochondrial genome and maternally inherited diabetes (Rorsman, 1997; Lang 1999).

Transcription Level Changes

The damages related to transcription may be a result of the failure to detect the transcription binding sites or the wrong signal from the TF for the detection of upstream/ downstream regions (Caisheng et al., 2009). For example, let us consider a study on the physiological behaviour of *Lactococcus lactis* subsp. *cremoris* MG 1363 was characterized in continuous culture under various acidic conditions (pH 4·7–6·6) was performed. Biomass yield was diminished in cultures with low pH and the energy dedicated to maintenance increased due to organic acid inhibition and cytoplasmic acidification. Under such acidic conditions, the specific rate of glucose consumption by the bacterium increased, thereby enhancing energy supply. This acceleration of glycolysis was regulated by both an increase in the concentrations of glycolytic enzymes (hierarchical regulation) and the specific modulation of enzyme activities (metabolic regulation). However, when the inhibitory effect of intracellular pH on enzyme activity was taken into account in the model of regulation, metabolite regulation was shown to be the dominant factor controlling pathway flux. The changes in glycolytic enzyme concentrations were not correlated directly to modifications in transcript concentrations. Analyses of the relative contribution of the phenomena controlling enzyme synthesis indicated that translational regulation had a major influence compared to transcriptional regulation. An increase in the translation efficiency was accompanied by an important decrease of total cellular RNA concentrations, confirming that the translation apparatus of *L. lactis* was optimized under acid stress conditions.

Transcriptional damage may be caused due to the chromosomal changes and may lead to protein misfolding. This leads to protein structure damage and may not result in the formation of the particular protein. The transcriptional changes also lead to protein frameshift which would result in wrong metabolite metabolite formaition.

Enzymatic Dysfunction

Enzymes are proteins that catalyze specific metabolic steps. Many enzymes catalyze a single step; most catalyze only a few. Enzymes are highly specific with respect to the substrates on which they act (Herzyk et al., 2004; Cordwell 1999). Many metabolic pathways consist of a series of chemical reactions, each catalyzed by a separate enzyme. It is very necessary to detect those enzymes which may be deleted from the path because the metabolite to be formed in the reaction catalyzed by that particular enzyme may not be formed (Herzyk et al., 2004; Gholson et al., 1982; Guo, 1986). This blocks the further process. Sometimes low availability of the enzyme may cause the formation of an altogether new product (Graber et al., 2009). If they occurs enzymatic alterations the particular end compound would not form which may lead to major changes in the metabolism process depending on its occurrence in the pathway (Herzyk et al., 2004; Gholson et al., 1982; Schmitz-Peiffer, 2000). But sometimes important enzymes do not bring about a major change and an enzyme rarely occurring causes a major physiological change (Schmitz-Peiffer, 2000), Hwang et al., 2007). Hence it is important to identify such

enzymes and study their effects in metabolism.

Current Strategies

To Detect Gene Expression Level Changes

Identifying genes and pathways associated with diseases such as cancer has been a subject of considerable research in recent years in the area of bioinformatics and computational biology (Speed et al., 2002; Herzyk et al., 2004). Many statistical learning techniques such as support vector machines, the relevance vector machines (RVM), LASSO, and sparse logistic regression have been applied to this problem (Krogh et al., 1996). Such algorithms fulfill two common goals, viz., to distinguish cancer and non-cancer patients with the highest possible accuracy and to identify a small subset of genes that are highly differentiated in different classes and to associate gene expression patterns with disease status (Breitling et al., 2004). In gene selection, when genes share the same biological pathway, the correlation between them can be high and those genes form a group (Breitling et al., 2004). The ideal gene selection methods eliminate the trivial genes and automatically include the whole group genes into the model once one gene among them is selected. Most importantly, almost all of the current methods are biased towards selecting those genes that display the most pronounced expression differences (Speed et al., 2002). Such methods select genes using purely statistical criteria (either rank score or classification accuracy) and this selection is thought to reflect their relative importance. Quite often, a certain number of genes with the smallest p-values or highest prediction accuracy are finally selected, while most biologists recognize that the magnitude of differential expression does not necessarily indicate biological significance (Breitling et al., 2004).

Although there is ongoing research to incorporate prior biological knowledge, such as partially known pathways in gene selection, to the best of our knowledge, there is no efficient method to hunt the upstream regulatory genes in gene selection and pathway discovery (Speed et al., 2002). Magnitude of differential expression does not necessarily indicate biological significance. Small changes can cause major physiological consequences if the protein encoded by the gene plays a catabolic role. Regulatory genes show low expression changes but they are highly correlated with the downstream highly expressed genes (Brown et al., 1997). No efficient method to hunt upstream regulatory genes in gene selection exists, therefore an algorithm was developed which uses Lp norm regularization which is equivalent to super Laplace prior over the model parameters (Herzyk et al., 2004). Both the generalization ability of the model and the scarcity achieved are critically dependent on the value of a regularized parameter, which has to be carefully tuned to the best performance. This best parameter can only be found through cross-validation and computationally intensive search (Breitling et al., 2004). Various methods have been proposed for identifying the small subsets of genes through Lp regularized Bayesian logistic regression and then define a novel similarity measure to identify the regularized genes that are highly correlated with each gene in the subset (Krogh et al., 1996).

To Detect Nucleosome Positioning

The positions of nucleosomes play important roles in diverse cellular processes that rely on access to genomic DNA, including DNA replication, recombination, repair, transcription, chromosome segregation, and cell division (Lorch et al., 1999). Recent studies have used DNA sequence features to predict genome-wide nucleosome positions with modest success, confirming that nucleosome positioning is partially encoded in the genomic DNA sequence. It has become clear that nucleosome positions are highly dynamic. Previous computational methods have predicted static nucleosome positions using DNA sequences with nucleosome formation or inhibition signals. However, more information besides the intrinsic DNA sequence is required to model nucleosome positioning dynamics. Till date, there has been no report on computational identification of dynamic nucleosome positions (Krogh et al., 1996). A novel computational approach was put forward for identifying dynamic nucleosome positioning at gene promoters on the base of dynamic transcriptional interaction and genomic sequence information. The predictions are in good agreement with experimentally determined nucleosome occupancy available in three cellular conditions. A recent study has used nucleosome occupancy information to assist identification of transcription factor (TF) binding sites (Lorch et al., 1999).

K-means clustering can also be used for detecting nucleosome positioning. Average nucleosome occupancy profile at promoters for each gene cluster can be found out and then pair-wise Euclidean distances can be calculated among these average profiles. The resulting distance reflects the degree of difference between the nucleosome occupancy profiles of two gene clusters (Caisheng et al., 2009). Based on dynamic transcriptional interaction and genomic sequence information, various computational approaches have been developed for identifying dynamic nucleosome positioning at promoters . Given gene expression data during one cellular condition, DNA sequences at gene promoters, and known position weight matrixes (PWMs) that correspond to candidate TFs, could be used to identify nucleosome positioning

dynamics during the condition.

Graph Theory and Analysis for Detecting Enzyme Alterations

A metabolic pathway may be represented computationally using a combination of statistical methods and programming methods (Speed et al., 2002). Data-sets are obtained from a central repository database which consists of metabolic pathway information, from which graph-based representations are plotted (Vilaprinyo et al, 2006; Niyogi et al, 2005). These may also be used to represent the metabolic pathway and the altered enzymes could be identified through this method (Agarwal et al., 2006; Aleman-Meza et al., 2006; Vilaprinyo et al., 2006). Nodes may be represented are the compounds or the metabolites of the reaction or enzymes and edges may act as reaction links. Each node would contain a dataset which would help us segregate the list of similar entities. Statistical calculations such as, p-value calculation, matrix methods could be performed to rank the nodes (Krogh et al., 1996; Curto et al., 1997).

Ranking of the nodes help in arranging the datasets according to the enzymes causing maximum change in the pathway (Agarwal, 2006; Kanehisa, 2000; Vilaprinyo et al., 2006). The graph can be modeled on the basis of the calculated values which help identify the dataset (enzymes) causing maximum pathway changes. The nodes in the graph are ranked in order to indentify the connections and the link in the pathway. A method of ranking nodes comprises the steps of analyzing each node with respect to at least one criterion, and assigning an intrinsic score to each node based on the analysis; identifying links between the nodes and generating a ranking score for each node based on the intrinsic scores of nodes linked therewith (Herzyk et al., 2004). Ranking of directed an undirected graphs would vary. According to the rank score the nodes are arranged according their particular rank. Weights are assigned to these ranked nodes by using matrix methods (Agarwal, 2006; Kanehisa et al., 2000; Niyogi et al., 2005).

A Bayesian method for identifying missing enzymes in predicted metabolic pathway databases

The Pathologic program constructs Pathway/Genome databases by using a genome's annotation to predict the set of metabolic pathways present in an organism. Pathologic determines the set of reactions composing those pathways from the enzymes annotated in the organism's genome. Most annotation efforts fail to assign function to 40–60% of sequences. In addition, large numbers of sequences may have non-specific annotations (e.g., thiolase family protein). *Pathway holes* occur when a genome appears to lack the enzymes needed to catalyze reactions in a pathway (Pagano Marcello et al., 1993). If a protein has not been assigned a specific function during the annotation process, any reaction catalyzed by that protein will appear as a missing enzyme or pathway hole in a Pathway/Genome database. As a result of this a method has been developed that efficiently combines homology and pathway-basedevidence to identify candidates for filling pathway holes in Pathway/Genome databases (Spang etal., 1998; Beutler, 1984). The program not only identifies potential candidate sequences for pathway holes, but combines data from multiple, heterogeneous sources to assess the likelihood that a candidate has the required function. The algorithm emulates the manual sequence annotation process, considering not only evidence from homology searches, but also considering evidence from genomic context (i.e., is the gene part of an operon?) and functional context (e.g., are there functionally-related genes nearby in the genome?) to determine the posterior belief that a candidate has the required function. The method can be applied across an entire metabolic pathway network and is generally applicable to any pathway database (Pearl, 1988). The program uses a set of sequences encoding the required activity in other genomes to identify candidate proteins in the genome of interest, and then evaluates each candidate by using a simple Bayesian classifier to determine the probability that the candidate has the desired function. 71% precision at a probability threshold of 0.9 during cross-validation was achieved using known reactions in computationally-predicted pathway databases. After applying this method to 513 pathway holes in 333 pathways from three Pathway/Genome databases, the number of complete pathways was increased by 42%. The pathway hole filler can be used not only to increase the utility of Pathway/Genome databases to both experimental and computational researchers, but also to improve predictions of protein function.

Influence of Metabolic Network Structure and Function on Enzyme Evolution

Most studies of molecular evolution are focused on individual genes and proteins. However, understanding the design principles and evolutionary properties of molecular networks requires a system-wide perspective. In the present work we connect molecular evolution on the gene level with system properties of a cellular metabolic network. In contrast to protein interaction networks, where several previous studies investigated the molecular evolution of proteins, metabolic networks have a relatively well-defined global function (Saher et al., 2009). The ability to consider fluxes in a metabolic network allows us to relate the functional role of each enzyme in a network to its rate of evolution.

The result of an experiment based on the yeast metabolic network, demonstrate that important evolutionary processes, such as the fixation of single nucleotide mutations, gene duplications, and gene deletions, are influenced by the structure and function of the network (Erb et al., 1999]. Specifically, central and highly connected enzymes evolve more slowly than less connected enzymes. Also, enzymes carrying high metabolic fluxes under natural biological conditions experience higher evolutionary constraints. Genes encoding enzymes with high connectivity and high metabolic flux have higher chances to retain duplicates in evolution (Watts et al., 1998; Jeong et al., 1999). In contrast to protein interaction networks, highly connected enzymes are no more likely to be essential compared to less connected enzymes. The presented analysis of evolutionary constraints, gene duplication, and essentiality demonstrates that the structure and function of a metabolic network shapes the evolution of its enzymes. The results underscore the need for systems-based approaches in studies of molecular evolution.

Quantitative, Scalable Discrete Event Simulation of Metabolic Pathways

DMSS(Discrete Metabolic Simulation System) seeks to create a scalable, object oriented, quantitative discrete event simulation framework based on a model of metabolic pathways as a graph (or network) of reactions. The system will facilitate the creation and evaluation of models through the simulation of experiments based on these models. Similar approaches have been used to model metabolic pathways, in particular the approach based on Petri nets (a type of graph) (Ehlde et al., 1995). This method has been suggested for modelling of pathways and for use in discrete{event simulation systems. Other approaches to modelling and simulating metabolic pathways include dfferential equation and Metabolic Control (Farza et al., 1992) Theory systems, for example. In addition, there have been a number of attempts to use knowledge {based and reasoning systems, for both qualitative and quantitative simulation. DMSS is designed to be used by bench biologists, allowing complex biologicalmodels to be expressed in biological terms (for example, as metabolite concentrations or chemical reactions), rather than as sets of obscure equation constants (Erb, 1999). That is, the biological relevance of the input model, simulation process and thus the results is imperative. The first step to achieving this goal is the ability to create biologically relevant models, hence a model must be defined in terms of biological constants (Pagano et al., 1993). The system should not require this knowledge to be described in mathematical terms, particularly not with respect to time. In contrast, the constants used in differential equation based mathematical models arise as a consequence of requiring the formulae to describe the reaction model for a range of reaction related parameters. Additionally, the stability and sensitivity of the equations for any given parameters need to be determined and accounted for, as these are often artifacts in such equation{based simulation systems. DMSS seeks to simulate non trivial systems, an important issue for metabolic simulation systems ones which are considered problematic for other techniques.The simulation of complex reaction systems can contribute to our understanding of biological systems and be an experimental platform for in silico biology.

Ranking on Graph Data

In ranking, one is given examples of order relationships among objects, and the goal is to learn from these examples a real-valued ranking function that induces a ranking or ordering over the object space (Agarwal, 2005). We consider the problem of learning such a ranking function when the data is represented as a graph, in which vertices correspond to objects and edges encode similarities between objects. Building on recent developments in regularization theory for graphs and corresponding Laplacian-based methods for classification, we develop an algorithmic framework for learning ranking functions on graph data (Belkin, 2004). 2004). We provide generalization guarantees for our algorithms via recent results based on the notion of algorithmic stability, and give experimental evidence of the potential benefits of our framework.

Application in Pathway Engineering

Graph theory and other statistical measures can be used to detect the enzyme alterations using computational and statistical methods. In order to detect the damages present in the pathway the graph theory plays a major role in representing the metabolic pathway and analyzing the connections between metabolites and their formation (Agarwal, 2006; Niyogi et al., 2005; Niyogi et al., 2004). This can be done from the pre-collected data repository. For detection of enzymatic changes the nodes and edges are connected and the degree of the enzymes and metabolites is calculated, which would further help find the rank of the same. From the ranks, one could find out occurrence of the enzymes and their positions in the pathway by comparing the ranks with the respective nodes. This study would help to categorize the enzymes that bring about maximum physiological change and hence detect the altered enzymes in a metabolic pathway. The detection of altered enzymes in a metabolic pathway may help find cure for many metabolic disorders.

Ranking may also help to analyze the differential gene expression from the microarray data available (Speed et al., 2002; Herzyk et al., 2004). Graph is plotted and the ranking is done based on highly differentially expressed genes. The ranking of a graph also depends on the degree of the graph. The degree of the graph is basically the number of sub-nodes a particular node is connected to (Herzyk et al., 2004; Breitling et al., 2004). The number of nodes having the same degree is cumulatively placed under the same rank.

Conclusion & Future Perspective

Damages in metabolic pathways can be identified by representing the metabolic pathway in the form of graphs. A graph based analysis which implements statistical methods help represent metabolic pathways and understand the function of a particular enzyme and the changes caused in the pathway due to the absence of the enzyme. Integration of statistical and computational methods has helped formulate an algorithm to determine the changes caused in a metabolic pathway. The detection of these damages can help understand the disorders caused due to these damages.

References

1. Abedin MJ, Wang D, McDonnell MA, Lehmann U, Kelekar A (2007) Autophagy delays apoptotic death in breast cancer cells following DNA damage. Cell Death Differ 14: 500-510. » CrossRef » Pubmed » Google Scholar
2. Agarwal S (2006) Ranking on graph data. ICML pp25-32. » CrossRef » Google Scholar
3. Agarwal A, Chakrabarti S, Aggarwal S (2006) Learning to rank networked entities, SIGKDD Conference pp14-23. » CrossRef » Google Scholar
4. Agarwal S, Niyogi P (2005) Stability and generalization of bipartite ranking algorithms, Proceedings of the 18th Annual Conference on Learning Theory. » CrossRef » Pubmed » Google Scholar
5. Aita VM, Liang XH, Murty VV, Pincus DL, Yu W, et al. (1999) Cloning and genomic organization of beclin 1, a candidate tumor suppressor gene on chromosome 17q21. Genomics 59: 59-65. » CrossRef » Pubmed » Google Scholar
6. Albertson DG (2006) Gene amplification in cancer. Trends Genet 22: 447-55.» CrossRef » Pubmed » Google Scholar
7. Aleman-Meza B (2006) Searching and ranking documents based on semantic relationships, 22nd International Conference on Data Engineering Workshops (ICDEW'06), pp 5.
8. Bartkova J, Horejsí Z, Koed K, Krämer A, Tort F, et al. (2005) DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature 434: 864-70. » CrossRef » Pubmed » Google Scholar
9. Beljanski V, Marzilli LG, Doetsch PW (2004) DNA Damage-Processing Pathways Involved in the Eukaryotic Cellular Response to Anticancer DNA Cross-Linking Drugs. Mol Pharmacol 65: 1496-506. » CrossRef » Pubmed » Google Scholar
10. Belkin M, Matveeva I, Niyogi P (2004) Regularization and semi-supervised learning on large graphs, Proceedings of the 17th Annual Conference on Learning Theory.
11. Breitling R, Amtmann A, Herzyk P (2004) Graph-based Iterative Group Analysis (iGA): A simple tool to enhance sensitivity and facilitate interpretation of microarray experiments. BMC Bioinformatics 5: 100.» CrossRef » Pubmed » Google Scholar
12. Breitling R, Armengaud P, Amtmann A, Herzyk P (2004) Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett 573: 83-92. » CrossRef » Pubmed » Google Scholar
13. Cordwell SJ (1999) Microbial genomes and "missing" enzymes: redefining biochemical pathways. Archives Microbiology 172: 269-279.» CrossRef » Pubmed » Google Scholar
14. Curto R, Voit EO, Sorribas A, Cascante M (1997) Validation and steady-state analysis of a power-law model of purine metabolism in man. Biochem J 324: 761-75. » CrossRef » Pubmed » Google Scholar
15. Dai Z, Dai X, Xiang Q, Feng J, Deng Y, et al. (2009) Transcriptional interaction-assisted identification of dynamic nucleosome positioning, BMC Bioinformatics. 10 Suppl 1: S31. » CrossRef » Pubmed » Google Scholar
16. DeRisi JL, Iyer VR, Brown PO (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. Science. 278: 680-686.» CrossRef » Pubmed » Google Scholar
17. Guo Z (2008) Intramyocellular lipid kinetics and insulin resistance. Med. Hypotheses 70: 625-9.» CrossRef » Pubmed » Google Scholar
18. Herwig R, Aanstad P, Clark M, Lehrach H (2001) Statistical evaluation of differential expression on cDNA nylon arrays with replicated experiments. Nucleic Acids Res 29: E117. » CrossRef » Pubmed » Google Scholar
19. Huang W, Wang P, Liu Z, Zhang L (2009) Identifying disease associations via genome-wide association studies. BMC Bioinformatics 1: S68. » CrossRef » Pubmed
20. Hughey R, Krogh A (1996) Hidden Markov models for sequence analysis: extension and analysis of the basic method. Comput Appl Biosci 12: 95-107.» CrossRef » Pubmed » Google Scholar
21. Karantza-Wadsworth V, Patel S, Kravchuk O, Chen G,

Mathew R, et al. (2007) Autophagy mitigates metabolic stress and genome damage in mammary tumorigenesis. Genes Dev 21: 1621-35.» CrossRef » Pubmed » Google Scholar

22. Kirschner MW (2005) The Meaning of Systems Biology. Cell 121: 503-4.» CrossRef » Pubmed » Google Scholar

23. Kornberg RD, Lorch Y (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell 98: 285-294.» CrossRef » Pubmed » Google Scholar

24. Lang J (1999) Molecular mechanisms and regulation of insulin exocytosis as a paradigm of endocrine secretion. Eur J Biochem 259: 3-17. » CrossRef » Pubmed » Google Scholar

25. Lemke N, Herédia F, Barcellos CK, Dos Reis AN, Mombach JC (2004) Essentiality and damage in metabolic networks. Bioinformatics 20: 115-9.» CrossRef » Pubmed » Google Scholar

26. Liu ET (2005) Systems biology, integrative biology, predictive biology. Cell 121: 505-6. » CrossRef » Pubmed » Google Scholar

27. Liu WC, Lin WH, Davis AJ, Jordán F, Yang HT, et al. (2007) A network perspective on the topological importance of enzymes and their phylogenetic conservation. BMC Bioinformatics 8: 121.» CrossRef » Pubmed » Google Scholar

28. Maechler P, de Andrade PB (2006) Mitochondrial damages and the regulation of insulin secretion, Biochem Soc Trans 34: 824-7.» CrossRef » Pubmed » Google Scholar

29. Matschinsky FM (1996) Banting Lecture 1995. A lesson in metabolic regulation inspired by the glucokinase glucose sensor paradigm. Diabetes 45: 223-241.» CrossRef » Pubmed » Google Scholar

30. Modre-Osprian R, Osprian I, Tilg B, Schreier G, et al. (2009) Dynamic simulations on the mitochondrial fatty acid beta-oxidation network. BMC Syst Biol 3: 2.» CrossRef » Pubmed » Google Scholar

31. Nardai G, Korcsmáros T, Papp E, Csermely P (2003) Reduction of the endoplasmic reticulum accompanies the oxidative damage of diabetes mellitus. Biofactors 17: 259-67.» CrossRef » Pubmed » Google Scholar

32. Nasu S, Wicks FD, Gholson RK (1982) L-Aspartate oxidase, a newly discovered enzyme of Escherichia coli, is the B protein of quinolinate synthetase. J Biol Chem 257: 626-632. » CrossRef » Pubmed » Google Scholar

33. Ogata H, Fujibuchi W, Goto S, Kanehisa M (2000) A heuristic graph comparison algorithm and its application to detect functionally related enzyme clusters. Nucleic Acids Res 28: 4021-8.» CrossRef » Pubmed » Google Scholar

34. Otto TD, Guimarães AC, Degrave WM, de Miranda AB (2008) AnEnPi: Identification and annotation of analogous enzymes. BMC Bioinformatics 9: 544. » CrossRef » Pubmed » Google Scholar

35. Pan DA, Lillioja S, Kriketos AD, Milner MR, Baur LA, et al. (1997) Skeletal muscle triglyceride levels areinversely related to insulin action. Diabetes 46: 983-988. » CrossRef » Pubmed » Google Scholar

36. Rorsman P (1997) The pancreatic -cell as a fuel sensor: an electrophysiologist's opinion. Diabetologia 40: 487495. » CrossRef » Pubmed » Google Scholar

37. Rossell´o F, Valiente G (2004) Analysis of Metabolic Pathways by Graph Transformation, Proceedings 2nd International Conference of Graph Transformation (ICGT'04), Springer LNCS 3256 pp70-82.

38. Seressiotis A, Bailey JE (1986) MPS - an algorithm and database for metabolic pathway synthesis. Biotechnol Lett 8: 837-842.» CrossRef » Google Scholar

39. Schmitz-Peiffer C (2000) Signalling aspects of insulin resistance in skeletal muscle: mechanisms induced by lipid oversupply. Cell Signal 12: 583-94.» CrossRef » Pubmed » Google Scholar

40. Vera J, Balsa-Canto E, Wellstead P, Banga JR, Wolkenhauer O (2007) Power-Law Models of Signal Transduction Pathways. Cell Signal 19: 1531-41.» CrossRef » Pubmed » Google Scholar

41. Vilaprinyo E, Alves R, Sorribas A (2006) Use of physiological constraints to identify quantitative design principles for gene expression in yeast adaptation to heat shock. BMC Bioinformatics 7: 184. » CrossRef » Pubmed » Google Scholar

42. Weirauch MT, Wong CK, Byrne AB, Stuart JM (2008) Information-based methods for predicting gene function from systematic gene knock-downs. BMC Bioinformatics 9: 463.» CrossRef » Pubmed » Google Scholar