Joint EUROGRAPHICS - IEEE TCVG Symposium on Visualization (2004)
O. Deussen, C. Hansen, D.A. Keim, D. Saupe (Editors)

Visual Triangulation of Network-Based Phylogenetic Trees

U. Brandes,[1,†] T. Dwyer,[2] and F. Schreiber[3,‡]

[1] Department of Computer & Information Science, University of Konstanz, Germany
[2] School of Information Technologies, University of Sydney, NSW, Australia
[3] Bioinformatics Center, Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany

Abstract

Phylogenetic trees are built by examining differences in the biological traits of a set of species. An example of such a trait is a biological network such as a metabolic pathway, common to all species but with subtle differences in each. Phylogenetic trees of metabolic pathways represent multiple aspects of similarity and hypothetical evolution in a single, yet complex structure that is difficult to understand and interpret. We present a visualization method that facilitates analysis of such structures by presenting multiple coordinated perspectives simultaneously. Each of these perspectives constitutes a useful visualization in its own right, but it is only together that they unfold their full explorative power.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation]: User Interfaces, J.3 [Computer Applications]: Life and Medical Sciences

1. Introduction

Phylogenetic analysis is an attempt to uncover the evolutionary relationships between organisms. It is an important tool in understanding evolutionary processes and in measuring genetic variations between species. Applications include the design of new drugs and reconstruction of the history of infectious diseases [HBN*96]. The result of phylogenetic analysis is a phylogenetic tree representing hypothetical ancestral relationships among a set of entities (see Fig. 1). Established methods for phylogenetic analysis are based on morphological attributes or nucleotide and protein sequences [Fel89, Fit77, AH92].

Recently, new methods using metabolic pathway data have been introduced, and phylogenetic trees based on such data are becoming increasingly important [FS01, HS03, LKT02]. These trees have a complex structure as each node of the tree is a network rather than a sequence as in simple phylogenetic trees. This development complicates visualization of phylogenetic trees as sequences can be seen as one-dimensional information whereas networks cannot. Therefore, network-based phylogenetic trees

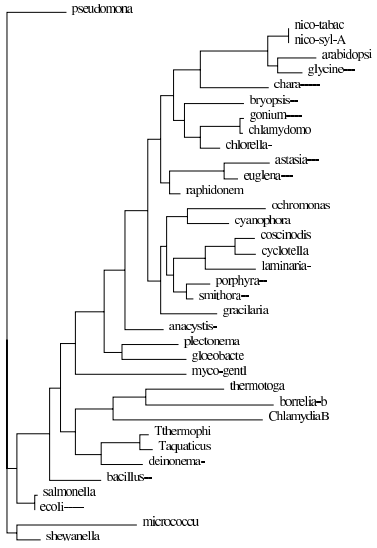

Figure 1: *Dendrogram representation of a (simple) phylogenetic tree in which leaves represent species and internal nodes (or branching points) represent hypothetical ancestors. Branch lengths give an indication of evolutionary time. (Data from [DKP95])*

† Partially supported by DFG under grant Br 2158/1-2.
‡ Supported by BMBF under grant 0312706A.

require more elaborate visualization methods that provide a general tree overview as well as detailed visualization of the networks represented by nodes in the tree.

A naïve approach would be to draw the metabolic pathways inside the leaf nodes of their phylogenetic tree. Such diagrams, however, fail to convey the essential information in the data because viewers are not able to easily compare the similarities and differences of pathways.

We propose a visualization approach based on the idea of triangulation [Jic79], i.e. employing multiple complementary tools to study an object that is not easily understood when applying a single method. In this approach different visualizations are shown for different aspects and a diagram of the phylogenetic tree is used as the main selection panel to determine the data shown in the other views.

The paper is organized as follows. In Sect. 2, we give some background on the type of phylogenetic trees considered and briefly review related approaches to visualizing phylogenetic trees. The proposed method of visualization is introduced in Sect. 3, and its application to typical real-world data demonstrated in Sect. 4. We conclude in Sect. 5.

2. Preliminaries

In this section we provide some background on the kind of data that we want to visualize.

2.1. Phylogenetic trees

A *phylogenetic tree* $T = (S, L)$ is a tree consisting of a set of nodes S (species) and a set of edges L (links). Leaf nodes of the tree (i.e. nodes having exactly one link) represent given species, sequences or similar entities; they are called *operational taxonomic units*. Internal nodes represent hypothetical ancestors generated from phylogenetic analysis; they are called *hypothetical taxonomic units*. See Fig. 1 for an example of a conventional drawing of a phylogenetic tree.

2.2. Metabolic networks

Metabolic reactions are transformations of chemical substances which occur in living beings and are usually catalyzed by enzymes. Important processes in organisms such as energy production and the synthesis of substances are based on reactions which form a large and complex network. A *metabolic pathway* is a subnetwork of the complete network of metabolic reactions (see Fig. 2). Such subnetworks may originate from biochemical textbooks (e.g., [Mic99]) or databases (e.g., KEGG [KG00]), or they can be defined by functional boundaries (e.g., the network between an initial and a final substance).

A metabolic pathway is modeled by a directed graph $G = (V, E)$, where a node $v \in V$ represents either a substance or a reaction (including the associated enzyme), and an edge $(v, w) \in E$ indicates that substance v enters reaction w or that reaction v produces substance w.

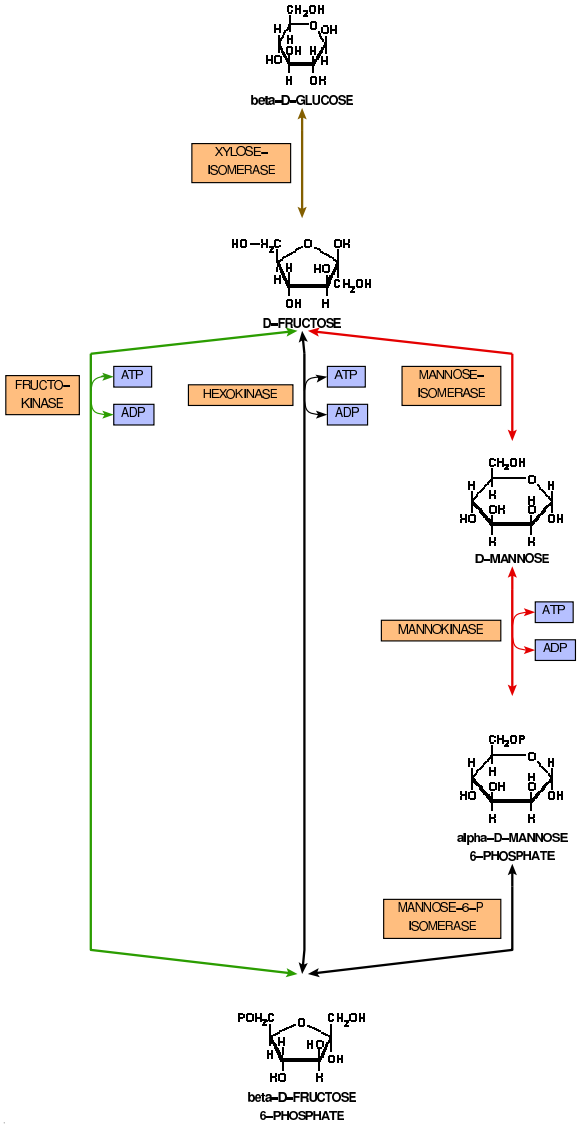

Figure 2: *Visualization of a metabolic pathway produced with BioPath [FPR*02]. White (containing structural formulas) and blue nodes represent chemical substances, orange nodes represent enzymes*

2.3. Complex phylogenetic trees

Not only morphological attributes and nucleotide sequences differ across organisms, but also metabolic pathways. Studies show significant variations even in central pathways such as glycolysis [DSS*99]. Such variations can be used for phylogenetic analysis.

To capture variations in metabolic networks, several similarity measures have been introduced. Most are characterized by the combination of structural information about the networks with additional data such as sequence information [FS01], the enzyme classification [HS03, TMH00], or information about the hierarchical clustering of reactions into pathways [LKT02]. A similarity measure which only depends on the structure of the pathways, i.e. on the presence or absence of nodes and edges, is discussed in [BDS03].

Phylogenetic trees can be computed from similarity matrices alone, and several of the above mentioned papers discuss phylogenetic trees generated from their particular similarity measure.

We call a *complex phylogenetic tree* (a phylogenetic tree of metabolic pathways) a tree $T = (\{G_1,\dots,G_k\} \cup \{H_1,\dots,H_l\}, L)$ defined on directed graphs. Leaf nodes $s \in \{G_1,\dots,G_k\}$ represent given metabolic pathways (i.e., the operational taxonomic units), whereas internal tree nodes $s \in \{H_1,\dots,H_l\}$ represent fictitious metabolic pathways (i.e., a hypothetical taxonomic unit) obtained from phylogenetic analysis. A tree edge indicates a propositional evolutionary relationship between the two incident pathways and is labeled with a numerical value that serves as in indicator for distance in evolutionary time.

3. Coordinated Visual Triangulation

Several methods have been proposed for visualizing phylogenetic trees in which leaf nodes represent one-dimensional information such as sequences [Fel89, RBB02, Pag96, Mak01]. Furthermore, that task of comparing a set of simple phylogenetic trees has been considered [KA02, MGT*03, SHB*01]. However, to the best of our knowledge, there is no software or method available that deals with the problem of visualizing complex phylogenetic trees, i.e. trees in which leaf nodes represent complex structures such as pathways.

Triangulation denotes the use of a combination of methodologies in the study of a particular phenomenon [Jic79]. In this paper we propose four perspectives, or "points of reference," to facilitate visual exploration of complex phylogenetic trees. They are depicted in Fig. 3 and used in a coordinated fashion, i.e. interaction with one perspective may affect others accordingly.

3.1. Phylogenetic tree

In the central diagram, only the phylogenetic tree itself is shown in a conventional way. Since the tree corresponds to an agglomerative clustering of its leaves, an appropriate form of visualizing it is a dendrogram in which edge lengths, if given, indicate dissimilarity according to the measure used (cf. Fig. 1).

We demand the tree to be ordered such that the total dissimilarity of consecutive leaves is minimum. Such orderings

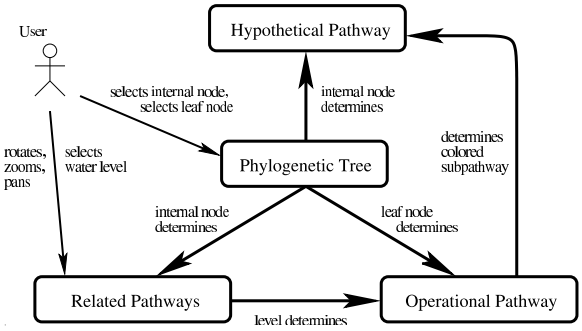

Figure 3: *Use-case diagram for triangulating a complex phylogenetic tree by four coordinated visualizations*

are known as *optimal leaf orderings* and can be computed efficiently for any similarity measure [BDW99, BJDG*02]. Note that in the general case, where the ordering is not restricted by a tree, optimal ordering is an $\mathcal{NP}$-hard problem (straightforward reduction from the Traveling Salesman Problem) and remains $\mathcal{NP}$-hard for typical similarity measures.

The tree diagram is interactive in that an internal node can be selected to focus on the subtree consisting of all its descendants. This selection causes the subtree to be highlighted and affects other visualizations as described below. Furthermore, a leaf node in the currently selected subtree may be checked.

3.2. Related pathways

An internal node in the phylogenetic tree is the hypothetical least common ancestor of all the leaves in its subtree. The pathways represented by these leaves are therefore considered related. To explore their particular relationships, they are shown in a single visualization designed with a $2\frac{1}{2}$-dimensional attitude [War01].

Two and a half-dimensional graph visualization is the stacking of a set of related graphs into the third dimension to support visual comparison [BDS03]. The order of the graphs in this stacking can be determined, either, based on the set of graphs currently shown, or by the order of the leaves within the phylogenetic tree. To maintain coherence of our different views we opt for the latter. Similarly, the spatial arrangement is kept consistent by computing a layout only for the union graph of all graphs representing leaves. This layout may also be tailored to the $2\frac{1}{2}$D situation in which the aesthetic requirements for the layout may be slightly different to the 2D situation (see [BDS03] for details).

The $2\frac{1}{2}$D structure is rendered in an interactive 3D perspective display with the usual navigational facilities. That is, the user can rotate, zoom, and pan the camera to get an informative view of the structure. An example is given in Fig. 8 in the next section.

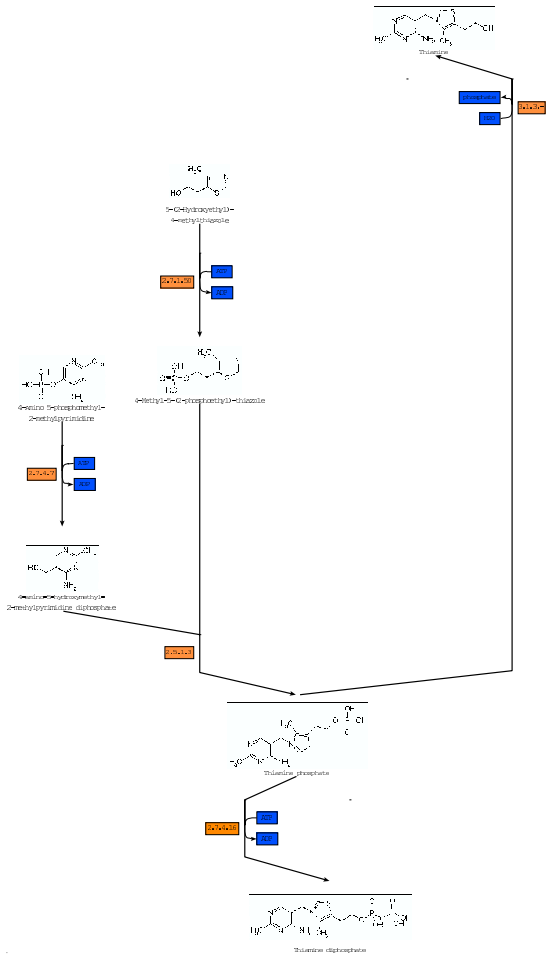

Figure 4: *Thiamine pathway of Haemophilus influenzae. The layout is determined by the layout in Fig. 5*

3.3. Operational pathway

An individual leaf can be singled out for detailed inspection by selecting it, either:

- directly from the phylogenetic tree;
- from a list of the leaves of the selected subtree;
- or from the $2\frac{1}{2}$D visualization of these leaves.

For the last option the user can move a semi-transparent "water-level" plane up and down through the $2\frac{1}{2}$D structure and the cross-section is shown in a 2D window. The water-level plane used in the selection process for our example is shown in Fig. 10.

While additional details on the selected operational pathway can be shown, it is important to maintain consistency across visualizations. Therefore, the layout of a pathway graph is determined from its plane in the $2\frac{1}{2}$D view. Figure 4 gives an example.

3.4. Hypothetical pathway

The graph of the hypothetical pathway represented by an internal node of the phylogenetic tree consists of all elements present in any of the pathways represented in its subtree leaves. However, some elements may be present in a larger number of leaves than others. Thus, each element is shown with a brightness that is inversely proportional to its frequency of occurrence. For example, an element that appears in each leaf will appear in a solid color while an element that appears in just a single leaf will be very light.

Moreover, the currently selected operational pathway, which by necessity is embedded in the current hypothetical pathway, is highlighted. See Fig. 5 which clearly shows the context of the pathway from Fig. 4.

4. Application Example

To illustrate the usage of visual triangulation in the analysis of complex phylogenetic trees we study the thiamine pathway of twelve different species.

4.1. Data

To build the phylogenetic tree of the example pathways several steps have been applied. Firstly, all pathways were extracted from the KEGG database [KG00] and were transformed into directed graphs such that each pathway graph contained all metabolites of the thiamine reference pathway and those enzymes and connections occurring in a particular species.

Since the particular generation scheme for the complex phylogenetic tree is irrelevant in the context of this paper, we computed the following, purely structural, dissimilarity measure [BDS03]: for two metabolic pathway graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ consider the union graph $G = (V, E)$, where

$$V = V_1 \cup V_2 \quad \text{and} \quad E = E_1 \cup E_2$$

and a set of *relevant elements* $P \subseteq (V \cup E)$. The *Hamming distance* (with respect to P) between G_1 and G_2 is defined as the cardinality of the symmetric difference of the relevant elements present in either graph,

$$\delta_P(G_1, G_2) = |((V_1 \cup E_1) \triangle (V_2 \cup E_2)) \cap P)| \,.$$

Note that this definition yields a metric. Dissimilarity between two pathways can thus be defined in terms of missing edges, nodes, or both, by choosing P accordingly. Relevant elements can also be selected to define regions of interest or specific views of the pathway such as the enzyme graph model of [OFGM00], where enzymes are the nodes of a graph and substances are (hyper)-edges.

The second step to build the phylogenetic tree was the computation of the dissimilarity matrix for the selected

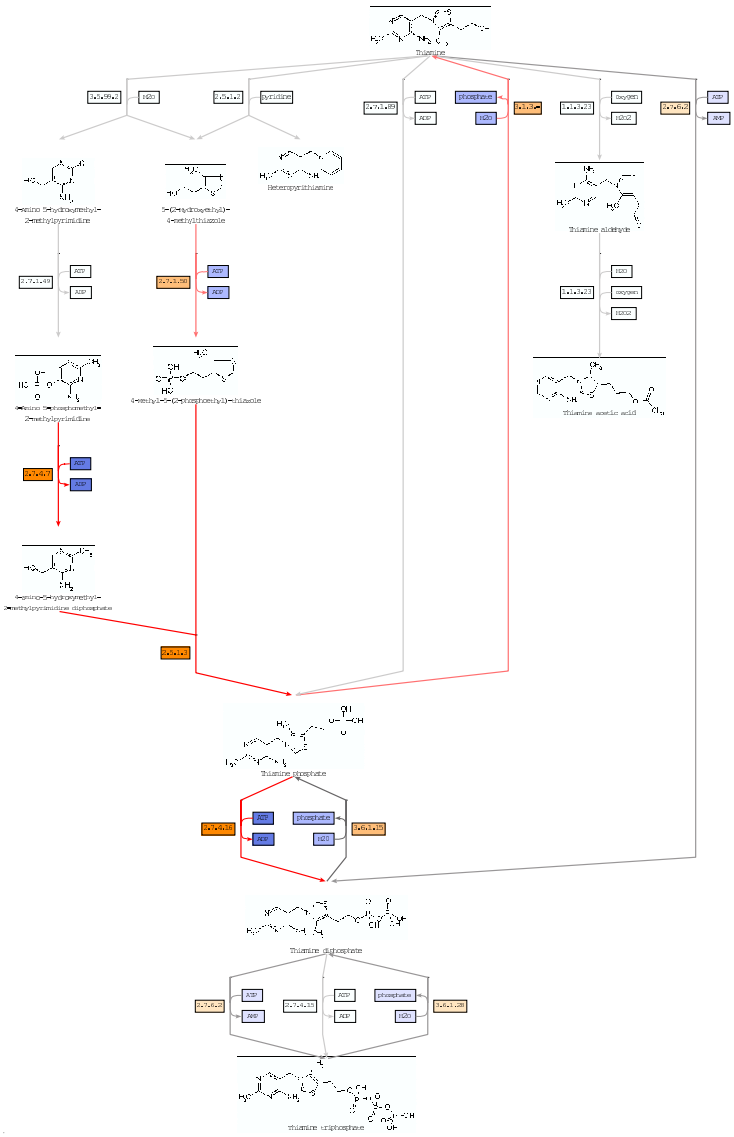

Figure 5: *Hypothetical thiamine pathway determined by all subtree leaves and the reference pathway; the brighter an element, the less likely its presence. The embedded operational pathway of Haemophilus influenzae is shown in red*

species. All nodes and edges were considered as relevant elements, i.e. $P = (V \cup E)$. The result of this step is shown in Fig. 6.

Finally, from this matrix, the phylogenetic tree was constructed using the *neighbor-joining* algorithm of [SN87], a hierarchical clustering method implemented in the phylogenetic analysis package *phylip* [Fel89]. This tree is shown in Fig. 7.

Note that metabolic pathway databases are known to contain inconsistencies and errors [WB01], furthermore the species-specific pathway data in KEGG is derived from genome data but not experimentally proven. This lack of complete and reliable data has a direct negative influence on the quality of the phylogenetic tree.

4.2. Visualizations

We have implemented a prototype system supporting coordinated visual triangulation as described in Sect. 3 based on the WilmaScope 3D graph visualization system [DE03].

Figure 7 shows the phylogenetic tree with edge lengths set according to the computed values. On the right hand side of the same figure we display the hypothetical pathway view

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)	(l)	
(a)	0	13	9	7	9	3	12	15	3	19	3	10	*Yersina pestis*
(b)	13	0	10	12	10	16	13	22	10	18	16	3	*Methanobacterium thermoautotrophicum*
(c)	9	10	0	16	0	6	3	18	6	28	6	7	*Archaeoglobus fulgidus*
(d)	7	12	16	0	16	10	13	16	10	12	10	9	*Treponema pallidum*
(e)	9	10	0	16	0	6	3	18	6	28	6	7	*Pyrococcus horikoshii*
(f)	3	16	6	10	6	0	9	12	6	22	0	13	*Haemophilus influenzae*
(g)	12	13	3	13	3	9	0	15	9	25	9	10	*Arabidopsis thaliana*
(h)	15	22	18	16	18	12	15	0	18	10	12	25	*Saccharomyces cerevisiae*
(i)	3	10	6	10	6	6	9	18	0	22	6	7	*Mycobacterium leprae*
(j)	19	18	28	12	28	22	25	10	22	0	22	21	*Mus muculus*
(k)	3	16	6	10	6	0	9	12	6	22	0	13	*Bacillus subtillis*
(l)	10	3	7	9	7	13	10	25	7	21	13	0	*Aeropyrum pernix*

Figure 6: *Hamming-distance matrix of thiamine pathways*

consisting of all elements present in any of the leaf pathways. Any node of the phylogenetic tree may be selected by the user thus changing the information shown in the hypothetical pathway view. Currently no selection is made so the union graph of all elements present in any of the pathways is shown. Note that the lightest boxes in the hypothetical pathway view correspond to reactions which are found in the thiamine reference pathway of the database but not in any of the species-specific pathways. These have been included as a reference for the biologist. Also note that although our prototype system only produces very simple, unlabeled, gray-scale 2D diagrams a full system should feature more informative diagrams, such as the examples produced by the metabolic pathway layout algorithm [Sch02] shown in Fig. 4 and 5.

Fig. 8 shows our related pathways view, displaying the set of pathways from the phylogenetic tree view (Fig. 7) in a single $2\frac{1}{2}$-dimensional visualization. Note that the stacking order matches the order of leaves shown in the tree view.

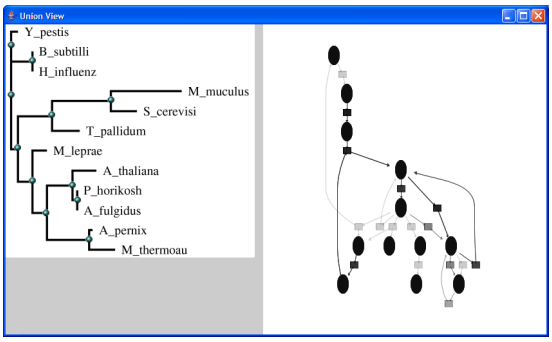

Figure 7: *Phylogenetic tree view and Hypothetical pathway view showing the phylogenetic tree built from the thiamine pathway of twelve species. Since no internal node of the tree has been selected by the user the hypothetical pathway corresponds to the root of the tree. The ultimate drawing of the sketched hypothetical pathway on the right is shown in Fig. 5*

Note also that the bottom layer shows the full set of reactions from the thiamine reference pathway as described above.

The process of browsing operational pathways (individual leaf pathways) is shown in Fig. 9 for the pathway of Haemophilus influenzae. Figure 10 shows how a pathway (Archaeoglobus fulgidus) was selected as a cross-section of the correspondence view by moving the transparent blue "water-level" plane.

In Fig. 10 all windows in the system showing all the coordinated views are presented in a single screenshot. In the tree view window on the bottom left an internal node of the phylogenetic tree has been selected highlighting the left and right subtrees beneath it. The hypothetical pathway corresponding to this internal node is shown in the right panel of the tree view window. Note that many of the elements of the pathway are a lighter gray indicating that they occur in fewer of the pathways represented by the leaves. Selecting a subtree in this view also affects the other views. The larger window in the background shows the correspondence view for the selected subtree and the list of navigable leaves is

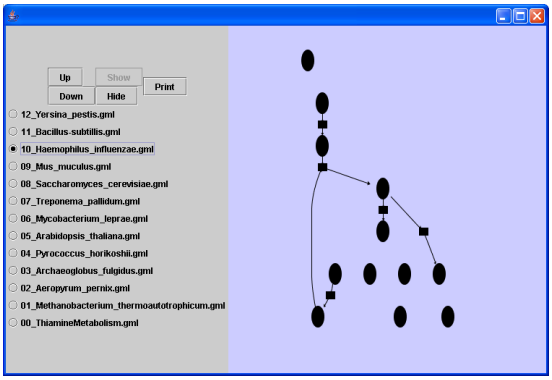

Figure 9: *An operational pathway can be selected either directly, using the radio buttons shown, or from the $2\frac{1}{2}D$ view for more careful analysis. Its final drawing is shown in Fig. 4*

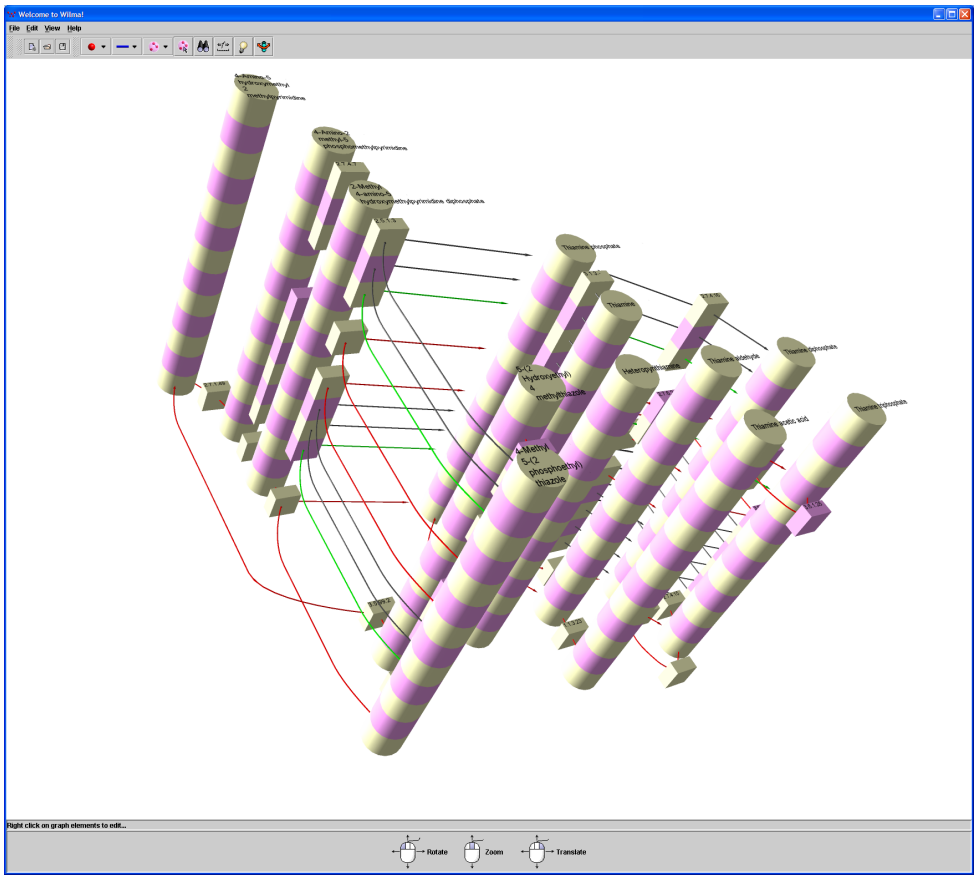

Figure 8: *Related pathways view showing pathways from Fig. 7 stacked in leaf order with the thiamine reference pathway located at the bottom of the stacking*

also reduced in the operational pathway browser on the bottom right. Possibilities for navigation through the complex phylogenetic tree and operations on the various views (e.g., zooming) are also shown in Fig. 3.

5. Discussion

An early version of the system featuring only the $2\frac{1}{2}$D related pathway view and the operational pathway view was recently evaluated in a user study with biologists. The ideas presented in this paper are designed to address issues raised in this study. The study involved a "cognitive walk-through" methodology in an interview type setting. That is, the domain experts were asked to use the system to explore a set of metabolic pathways, asked to think aloud as they proceeded with the exploration, and their feedback recorded.

In general the feedback regarding the use of the $2\frac{1}{2}$D stacked view for comparing related pathways was very positive. However, most users felt that the number of pathways which could be successfully visualized simultaneously in this way was limited. Perhaps to around six to ten pathways at a time. A method for interactively selecting which of the available pathways are included in the stack is therefore essential. Thus, the phylogenetic tree view not only provides this facility, but does so in a way that is meaningful and intuitive for the biologists.

A further logical extension supporting this mode of interaction is to allow for the selection of arbitrary multiple internal nodes and leaves from the phylogenetic tree for inclusion in the stack. It is also planned to further develop the components visualizing operational and hypothetical pathways to directly produce graphics as in Figs. 4 and 5 and thus eliminate cross-implementation data exchange.

The same idea is also applicable to phylogenetic trees built from other complex structures such as protein-protein interaction networks or signal transduction pathways, and to multi-level structures in general. Future work will investigate utilizing these other data sources and application areas in our system, as well as modifications to address scalability issues.

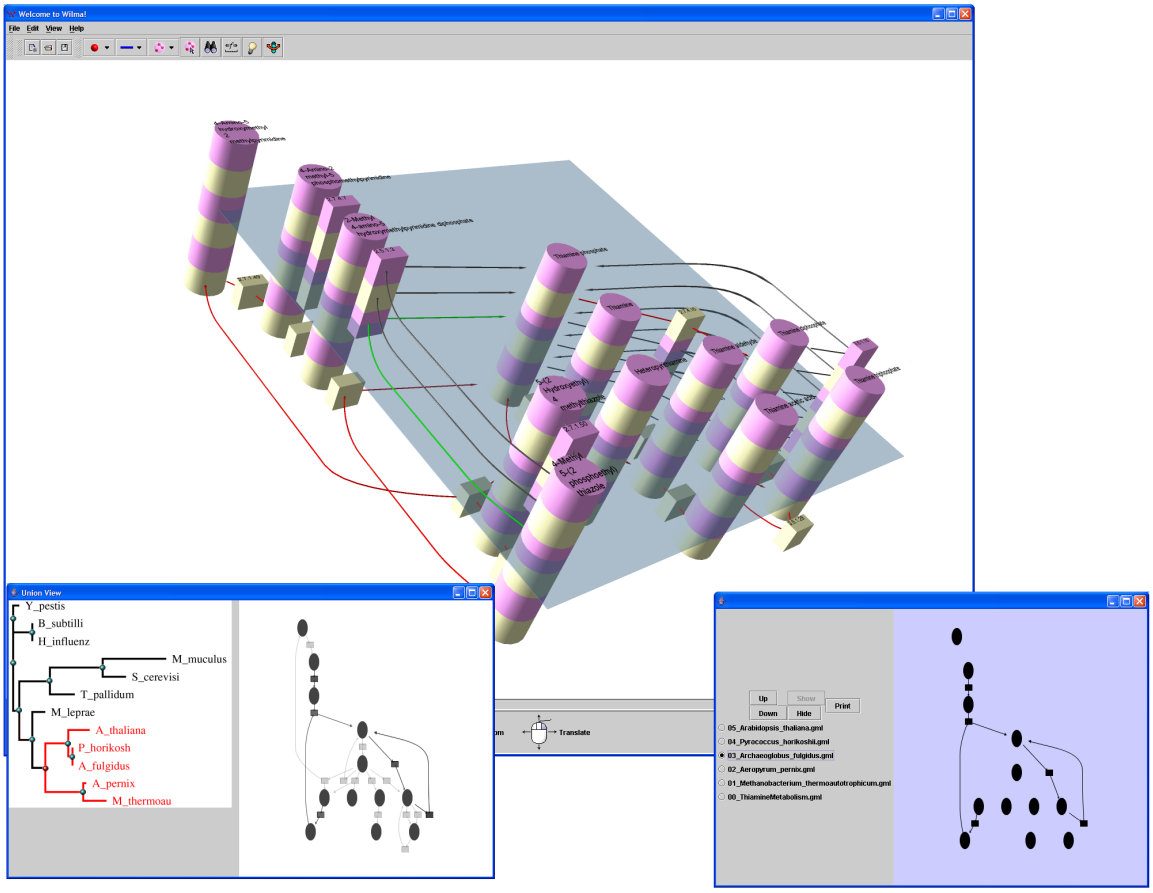

Figure 10: *A screen shot showing all the coordinated views. Again the phylogenetic tree window from Fig. 7 is shown. However, here an internal node has been selected by the user highlighting a subtree. Only elements which exist in leaves of this subtree and the reference pathway are shown in the hypothetical pathway on the right side of this window. The selection also applies to the leaf pathways shown in the* $2\frac{1}{2}D$ *stack (background window) and the individual pathways available for browsing in the slice view (lower right window)*

References

[AH92] ADACHI J., HASEGAWA M.: *PROTML: Maximum likelihood inference of protein phylogeny*. No. 27 in Computer Science Monographs. Institute of Statistical Mathematics, Tokyo, 1992, ch. MOLPHY: Programs for molecular phylogenetics, pp. 1–77. 1

[BDS03] BRANDES U., DWYER T., SCHREIBER F.: Visualizing related metabolic pathways in two and a half dimensions. In *Proc. Intl. Symp. Graph Drawing (GD 2003)*, Springer LNCS 2912 (2004), pp. 111–122. 3, 4

[BDW99] BURKHARD R. E., DEINEKO V. G., WOEGINGER G. J.: The traveling salesman and the *PQ*-tree. *Mathematics of Operations Research 24* (1999), 262–272. 3

[BJDG*02] BAR-JOSEPH Z., DEMAINE E. D., GIFFORD D. K., HAMEL A. M., JAAKKOLA T. S., SREBRO N.: *K*-ary clustering with optimal leaf ordering for gene expression data. In *Proc. Intl. Workshop Algorithms in Bioinformatics (WABI 2002)*, Springer LNCS 2452 (2002), pp. 506–520. 3

[DE03] DWYER T., ECKERSLEY P.: Wilmascope – a 3d graph visualization system. In *Graph Drawing Software*, Jünger M., Mutzel P., (Eds.). Springer, 2003, pp. 55–76. http://www.wilmascope.org. 5

[DKP95] DELWICHE C., KUHSEL M., PALMER J.: Phylogenetic analysis of tufA sequences indicates a cyanobacterial origin of all plastids. *Molecular Phylogenetics and Evolution 4*, 2 (1995), 110–

128. 1

[DSS*99] DANDEKAR T., SCHUSTER S., SNEL B., HUYNEN M., BORK P.: Pathway alignment: application to the comparative analysis of glycolytic enzymes. *Biochemical Journal 343* (1999), 115–124. 2

[Fel89] FELSENSTEIN J.: Phylip – phylogenetic inference package. *Cladistics 5* (1989), 164–166. http://evolution.genetics.washington.edu/phylip.html. 1, 3, 5

[Fit77] FITCH W. M.: On the problem of discovering the most parsimonious tree. *The American Naturalist 111* (1977), 223–257. 1

[FPR*02] FORSTER M., PICK A., RAITNER M., SCHREIBER F., BRANDENBURG F. J.: The system architecture of the BioPath system. *In Silico Biology 2*, 3 (2002), 415–426. 2

[FS01] FORST C. V., SCHULTEN K.: Phylogenetic analysis of metabolic pathways. *Journal Molecular Evolution 52* (2001), 471–489. 1, 3

[HBN*96] HOLMES E. C., BOLLYKY P. L., NEE S., RAMBAUT A., GARNETT G., HARVEY P. H.: *New Uses for New Phylogenies*. Oxford University Press, 1996, Ch. Using phylogenetic trees to reconstruct the history of infectious disease epidemics, pp. 169–186. 1

[HS03] HEYMANS M., SINGH A. K.: Deriving phylogenetic trees from the similarity analysis of metabolic pathways. *Bioinformatics 19*, Suppl. 1 (2003), 138–146. 1, 3

[Jic79] JICK T. D.: Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly 24* (1979), 602–611. 2, 3

[KA02] KLINGNER J., AMENTA N.: Case study: Visualizing sets of evolutionary trees. In *Proc. IEEE Information Visualization* (2002), pp. 71–74. 3

[KG00] KANEHISA M., GOTO S.: KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acid Research 28*, 1 (2000), 27–30. http://www.genome.ad.jp/. 2, 4

[LKT02] LIAO L., KIM S., TOMB J.-F.: Genome comparisons based on profiles of metabolic pathways. In *Proc. Intl. Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES '02)* (2002), pp. 469–476. 1, 3

[Mak01] MAKARENKOV V.: T-REX: reconstructing and visualizing phylogenetic trees and reticulation networks. *Bioinformatics 17*, 7 (2001), 664–668. 3

[MGT*03] MUNZNER T., GUIMBRETIÉRE F., TASIRAN S., ZHANG L., ZHOU Y.: Treejuxtaposer: Scalable tree comparison using focus+context with guaranteed visibility. In *Proc. SIGGRAPH* (2003). 3

[Mic99] MICHAL G.: *Biochemical Pathways*. Spektrum Akademischer Verlag, Heidelberg, 1999. 2

[OFGM00] OGATA H., FUGIBUCHI W., GOTO S., MKANEHISA: A heuristic graph comparison algorithm and its application to detect functionally related enzyme clusters. *Nucleic Acid Research 28*, 20 (2000), 4021–4028. 4

[Pag96] PAGE R. D. M.: Treeview: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences 12*, 4 (1996), 357–358. 3

[RBB02] ROST U., BAUER-BORNBERG E.: Treewiz: interactive exploration of huge trees. *Bioinformatics 18*, 1 (2002), 109–114. 3

[Sch02] SCHREIBER F.: High quality visualization of biochemical pathways in BioPath. *In Silico Biology 2*, 2 (2002), 59–73. 6

[SHB*01] STEWART C., HART D., BERRY D., OLSON G., WERNERT E., FISCHER W.: Parallel implementation and performance of fastDNAml: a program for maximum likelihood phylogenetic inference. In *Proc. 2001 ACM/IEEE Conference on Supercomputing (CDROM)* (2001), ACM, pp. 20–20. 3

[SN87] SAITOU N., NEI M.: The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution 4*, 4 (1987), 406–425. 5

[TMH00] TOHSATO Y., MATSUDA H., HASHIMOTO A.: A multiple alignment algorithm for metabolic pathway analysis using enzyme hierarchy. In *Proc. Intl. Conf. Intelligent Systems for Molecular Biology (ISMB 2000)* (2000), pp. 376–383. 3

[War01] WARE C.: Designing with a $2\frac{1}{2}$D attitude. *Information Design Journal 10*, 3 (2001), 255–262. 3

[WB01] WITTIG U., BEUCKELAER A. D.: Analysis and comparison of metabolic pathway databases. *Briefings in Bioinformatic 2*, 2 (2001), 126–142. 5