

Visual Comparison of Organism-Specific Metabolic Pathways

Katharina Unger*

Supervised by: Manuela Waldner†

Institute of Visual Computing & Human-Centered Technology
Vienna University of Technology
Vienna / Austria

Abstract

The Kyoto Encyclopaedia of Genes and Genomes (KEGG) resource is a combination of multiple databases, containing information about biochemical compounds, reactions, pathways, genes and much more. This database is one of the main resources for bioinformaticians and biologists to gain an understanding of molecular functionality inside organisms. The Orthology (KO) database from KEGG assigns pathways and genes with identical functionality to the same ortholog groups (KO entries). Therefore it is possible to map genes onto the pathway maps and obtain organism-specific visualizations. KEGG offers a web-based graph visualization to explore these pathways, however, the interaction possibilities are restricted and the rendering is inefficient. It is possible to visualize organism-specific pathways but a visual analysis tool to compare ortholog groups of multiple organisms is missing. In this work, we present an efficient interactive web application to compare ortholog groups of multiple organisms in the metabolic reference pathway. We introduce a graph overlay technique to mark the differences and similarities between multiple organisms and demonstrate it with two use cases. Additionally, we compare it against an existing point set membership visualization.

Keywords: Metabolic Pathways, KEGG: Kyoto Encyclopedia of Genes and Genomes, Functional Orthologs, Ortholog Groups, Organism-Specific Metabolic Pathways, HCL Color Space

1 Introduction

KEGG offers many possibilities for bioinformaticians and biologists to explore complex information about relations and dependencies of organisms at the molecular level. It provides multiple maps visualizing overviews and more detailed views of metabolic pathways in a static way. Biologists refer to these maps to gain a better understanding of reactions and compounds in a larger context, e.g. inside organisms. According to a collaborating domain expert, the preservation of the map layout is important as biologists

are used to it and can orientate fast. When new genes are identified, they are mapped to existing graph elements of the maps. Subsequently, the mapping between graphical elements and metabolic structures is not unique. Therefore interactive tools are desired, which allow exploring the complex layout and the structures mapped to it. Existing tools, however, use inefficient rendering which limits the interaction possibilities. Additionally, those tools do not offer an intuitive comparison of more than two organisms. However, this is a common task for biologists and is also of potential interest to the general public.

In this work, we contribute the design and implementation of an interactive web application. We introduce a new graph overlay technique to compare multiple organisms within the global reference pathways. Furthermore, we describe the implementation of our web application, which allows efficient rendering and therefore smooth interaction. We validate our visualizations by introducing use cases and report feedback from our domain expert.

2 KEGG

The goal of the Kyoto Encyclopaedia of Genes and Genomes (KEGG) [13, 15, 23] is to create a representation of biological processes consisting of four categories of databases: systems information (networks of reactions and relations), genomic information (genes and organisms), chemical information (molecules, metabolites, enzymes) and health information (diseases, drugs). To visualize relations and dependencies between metabolic compounds and reactions, KEGG introduced the pathway maps which were created manually.

A common task is to explore organism-specific relations in the context of the global reference pathway. Therefore, KEGG has a mapping procedure to map genes to the compounds and reactions on pathway maps. It also provides online tools to highlight a single selected organism within the reference pathway or to manually enter multiple genes and specify a color for them. According to our collaborating domain expert, biologists prefer graph overlays compared to showing an optimized sub-graph, as they are used to the layout and can orient fast in these maps.

The KEGG Orthology (KO) database, which is part of the genomic information, provides a characterization of

*e1325652@student.tuwien.ac.at

†waldner@cg.tuwien.ac.at

pathways and genes into KO entries. These KO entries are mapped to the graph layout. KO entries are manually defined by creating ortholog groups. Ortholog groups consist of pathways and genes with identical or similar functionality, which have a common ancestor gene. The mapping process of the KO entries consists of the KO assignment and the KEGG mapping. In the KO assignment, pathways and genes of the genome are assigned to KO entries. The KEGG mapping reconstructs the KEGG pathway maps with the newly assigned information.

The KEGG API [17] enables access to the databases and allows to create more complex queries linking data across these databases. Pathway maps can be accessed via this API as XML files defined by the KEGG Markup Language (KGML) [14]. These files contain the coordinates, type and color of the graph elements and basic information like names, type and a link leading to additional information of each entry. This gives the possibility to reconstruct and adapt pathways arbitrarily from these files.

To visualize shared KO entries of multiple organisms, we request all KO entries mapped to the global metabolic reference pathway (map ko01100). The KO entries which are part of single organisms are obtained by requesting the genes of an organism and then finding the corresponding KO entry for each gene. With the mapping of the KO entries to the map and the organism-specific KO entries, we can compute which entries are shared by which organisms and visualize them.

Unfortunately, the servers provided by KEGG are rather slow. To speed up the data acquisition process, we use server-side caching (see Section 5).

3 Related Work

As our work addresses the visualization of biochemical pathways, as well as visualization of overlapping sets, we will split the related work into these two sections. We will first cover existing tools in the context of metabolic pathways and related topics. Subsequently, we will introduce existing set membership visualization techniques.

3.1 Existing Tools

The KEGG Atlas [24], Pathway Projector [16] and iPath [8] are existing tools which offer visualizations of Pathways. All of them are web applications which use the KEGG global pathway map as a starting point for the visualization of queries based on the KEGG resource [23]. Every tool provides basic zooming and panning interaction to explore the global map and additional information of elements by clicking on nodes or edges. They allow to filter pathways per organism and to apply further, limited, filtering inside the selected organism.

KEGG provides some mapping tools for their application. With the *Search&Color Pathway* mapping tool or the *Reconstruct Pathway* tool it is possible to enter single KO

entries and specify colors for them, but it is not possible for whole organisms. The user has to filter the KO entries manually for differences and correspondences. For each KO entry, the user then has to define a color and enter this tuple into the tool to receive a visualization.

iPath is the most active web application in pathways visualization where the founders are still adding knowledge and interaction tools. The newest version is iPath3 [8], which is based on JavaScript and SVG, as opposed to the previous versions iPath1 [19] and iPath2 [31], which used a Flash-based framework. In the last version, the authors added a helper tool to compare two species and encode their differences and similarities in the overview map. However, this tool only provides a comparison of two organisms.

The Roche Pathways [25] is another tool handling metabolic pathways. The focus is on the visualization of the underlying data of the graph. They use the JavaScript library Leaflet [1] for their interactive map, providing semantic zooming. Another example in this domain for interactive visualization are Recon Maps [22]. This tool is implemented using the Google Maps API.

3.2 Set Membership Visualization

The pathway maps introduced by KEGG have a fixed layout. Domain experts are used to this layout, therefore it is important to keep the structure and visualize information on top of it. Subsequently, overlay techniques [30] are of most interest for us. These techniques are specified on the set membership of points. However, in the context of pathway maps, the entries we want to visualize in sets are not the points, but the edges of the graph.

The drawback of region-based approaches like Bubble Sets [7] or the approach by Byelas and Telea [5] is that a visualization of whole regions is not useful in our context. The elements in pathways are placed very compact and the data is linked in complex ways. The organisms we want to highlight on this graph, however, do not form regions. Additionally, our sets are highly overlapping, therefore graph visualizations like GMap [10], which only work with non-overlapping sets, are also not useful. Line-based approaches like Line Sets [2] and Kelp Diagrams [9] add line connections between nodes of common sets. However, our goal is to visualize set membership of edges. Combinations of region- and line-based techniques, like Kelp Fusion [21], have the same limitations. Hatching techniques, as used in Kelp Diagrams [9] can not be applied, because our graph is too dense, to see structures like this. However, we do adapt their presented overlay nesting technique for visualizing a small number of sets.

Visualizations which rearrange the data, like EulerView [28] or SetNet [26], are not useful as they do not preserve the well-known fixed layout of the pathway maps.

Apart from the visualization superimposed on the graph, overview visualizations of set relations are useful for the user, for better understanding. All kinds of diagrams based

on Euler and Venn-diagrams have the drawback that they do not scale well on more than three sets [3]. However, our goal is not to restrict the user in the number of organisms to compare. UpSet [20] or SEEM [11] scale better than Venn diagrams and allow for easier comparison of the set cardinalities, but they can still become rather large with a higher number of sets [20] or do not show all possible combinations of overlapping sets [11].

4 Comparative Graph Overlay

The goal of our graph overlay edge coloring technique is to visualize where multiple organisms overlap in the reference pathway and how they differ from each other. We use a technique that keeps the known graph structure, to facilitate orientation inside the graph for domain experts.

For the color encoding, we use the Hue-Chroma-Luminance color space (HCL). This color space is perceptually linear. This means that colors with equal distance inside this color space are perceived equally distant by humans [32]. The colors are coordinates in the 3D color space, where Hue and Chroma are polar coordinates in the color plane and the orthogonal Luminance corresponds to the brightness.

Every organism gets a base color assigned, where the hue between those colors is as distinct as possible, and the luminance and the chroma are constant. The base colors for the organisms are selected at luminance 75. At this luminance, the chroma can be maximized within the sRGB gamut. To calculate the hues, we divide the 360° representing all possible hues by the number of organisms. To avoid interpolation of opposite organisms into the same gray value, we always divide by an odd number. To maximize the chroma, we start with a value of 100 and iteratively reduce it until the value is valid (until we receive an RGB value, where no color component is clamped).

In Figure 1, we illustrate this color calculation with nine colors. To represent overlaps between sets, we interpolate the set colors and decrease the luminance with respect to the number of overlapping sets. For KO entries which are shared among multiple organisms, we transform the color into the cartesian LAB color space and interpolate the a and b values. Then, we compute the luminance with respect to the number of organisms that share this entry. We use an interpolation of the following form

$$l_m = 75 \left(1 - \left(\frac{m-1}{n-1} \right)^a \right) \quad (1)$$

where n is the total number of sets (organisms) and m is the number of overlapping sets. KO entries which are shared by all organisms will, therefore, be colored black. The more organisms, the darker the color. All graphical elements with a KO entry, which occurs in at least one organism, are colored with the calculated color, whereas all other elements are colored in a light gray, as it can be seen in Figure 2.

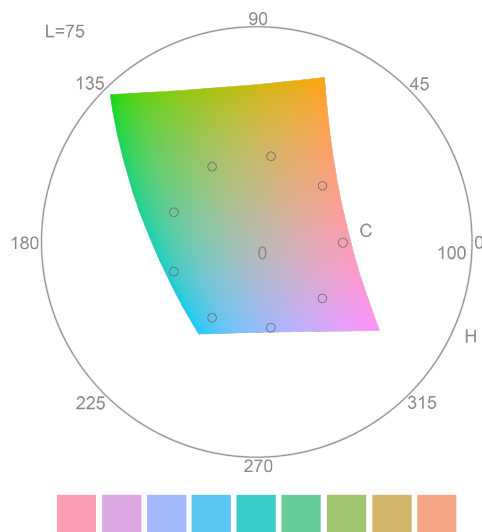


Figure 1: Selection of nine isoluminant colors in the HCL color space at luminance 75.

For each edge, we find all organisms that are associated with any KO assigned to the respective edge and compute the interpolated A and B color components and the corresponding luminance using Equation 1. We then render the edge using the LAB color. This way, the user can identify edges associated with a single organism by a bright edge color and edges shared by many organisms by a very dark color. Mouse-over tooltips show the exact associations.

To further support the user in the orientation of colors and the numbers of KO entries, we visualize the KO entries of the organisms in stacked bar charts. Each bar shows the ratio of KO entries of one organism with respect to the total amount of KO entries in the graph. The bars are subdivided according to the set overlaps, where the lowest bar segment shows the ratio of KO entries that is occupied only by the respective organism and the highest bar segment in black shows the ratio of KO entries shared by all organisms. Each of these bar segments is colored in the corresponding color, dependent on the organisms, which are part of it. An example of the stacked bar charts can be seen in Figure 3 in the lower right corner.

In our implementation, we do not restrict the number of organisms. However, as one can already see in Figure 1, it is hard to distinguish between nine isoluminant colors, particularly when the different colors are widely spread in the visualization. In Figure 2, we use our stacked bar charts to show how the color interpolation develops with respect to the number of organisms and its impact on the overlay. The bar charts are generated manually to show the different colors, and therefore do not represent any relations of KOs in the images below.

Additionally, to our edge color style, we implemented a visualization technique with a layer structure, inspired by Kelp Diagrams [9]. However, we do not use any type

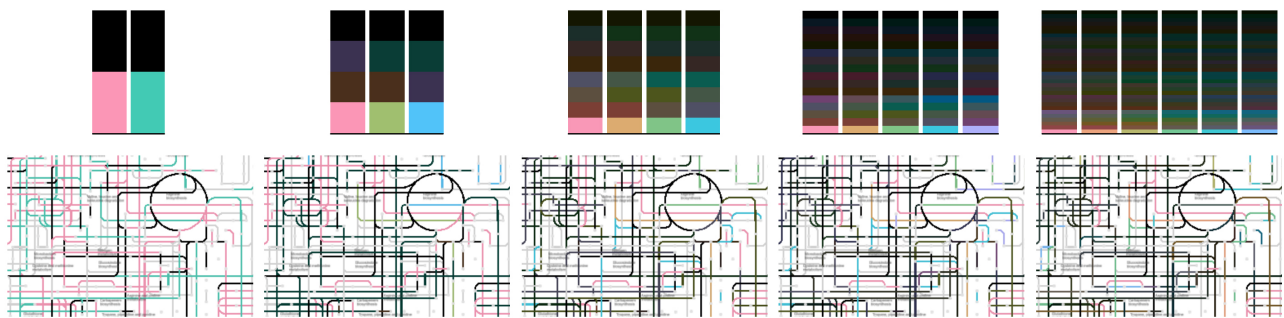


Figure 2: Colors indicating the individual organisms as well as all possible overlaps in the visualization, beginning with two organisms at the left column up to six organisms in the right column. The top row shows stacked bar charts with synthetic, equally sized overlap quantities to show the colors. The bottom row shows a small part of the global pathway map, with the colors illustrated in the charts above them.

or routing algorithm but just adapted their nested style to indicate overlapping sets.

For each organism, we create a layer with all elements, where a KO entry of this organism is assigned. To render multiple organisms above each other, each of the organisms gets a different line width assigned. Due to the growing line width, this approach is restricted to the selection of three organisms. With more organisms, the increasing line width in this dense graph results in lines, which visually merge. On top of the organism-specific layers, we add an overlay, which is colored in gray values. This layer indicates how many organisms share KO entries on this element. If a KO entry of this edge is shared by all organisms, this overlay is black, if it is only shared by two, it is gray. When the KO entry only appears in one organism, no edge is drawn in this layer. This layer is not part of Kelp Diagrams, but it is our attempt to handle the ambiguous mapping of KOs on the map in this technique.

5 Implementation

For our implementation, we use a client-server infrastructure. The server handles the requests to the KEGG API and preprocesses the data to reduce the computation on the client-side and therefore improve performance. Additionally, the server caches resources from the KEGG API which are often used, to work independently from the API and reduce response times.

As backend, we use Python with a Flask [27] server. All requests from the frontend are sent to this server and, if necessary, are forwarded to the KEGG servers. We use server-side caching to further reduce response times.

As previously described, the existing tools mostly use Scalable Vector Graphics (SVG) for rendering. The global pathway map consists of more than 4000 entries of circles and lines. SVG rendering can become a bottleneck when visualizing thousands of items. This is why we decided to use WebGL [12]-hardware-accelerated rendering,

exposed through the HTML5 Canvas API. To render to the `<canvas>` element, we use Three.js [6], which is a JavaScript scenegraph library based on WebGL [12]. According to the WebGL specification, it is not possible to render line width other than 1.0. Therefore, the lines with a line width higher than 1.0 need to be created with triangle meshes. To improve performance, by minimizing draw calls, we created large meshes containing multiple lines instead of small meshes for each line.

For zooming and panning interactions, we use the event listeners from D3.js [4], which is a data-driven JavaScript library to manipulate documents. From the events, we receive the mouse coordinates, which we use to calculate the new distance and scale according to the mouse wheel delta and update the camera position. The hover and clicking interactions of our tool are obtained by raycasting the scene at the mouse position with the Three.js raycaster.

6 Results

Our web application is accessible online [29] and is tested with Google Chrome Version 71 and Firefox 64 on Windows 10. To view organism-specific pathways it is required to write the KEGG organism code into the search field and press enter. The organism codes consist of three to four letters and are listed online [18]. For visualizations of multiple organisms, the organism codes are concatenated with '+'. For example, 'hsa+eco+sty' visualizes the human organism, one type of e-coli and one type of salmonella. We show two use cases and compare the edge coloring overlay with the nested overlay visualization in the following paragraphs. After that, we compare the rendering performance with the KEGG Atlas [24], iPath3 [8] and the Roche Pathways [25]. Furthermore, we present the feedback by our domain expert and discuss the advantages and disadvantages of the different presentations.

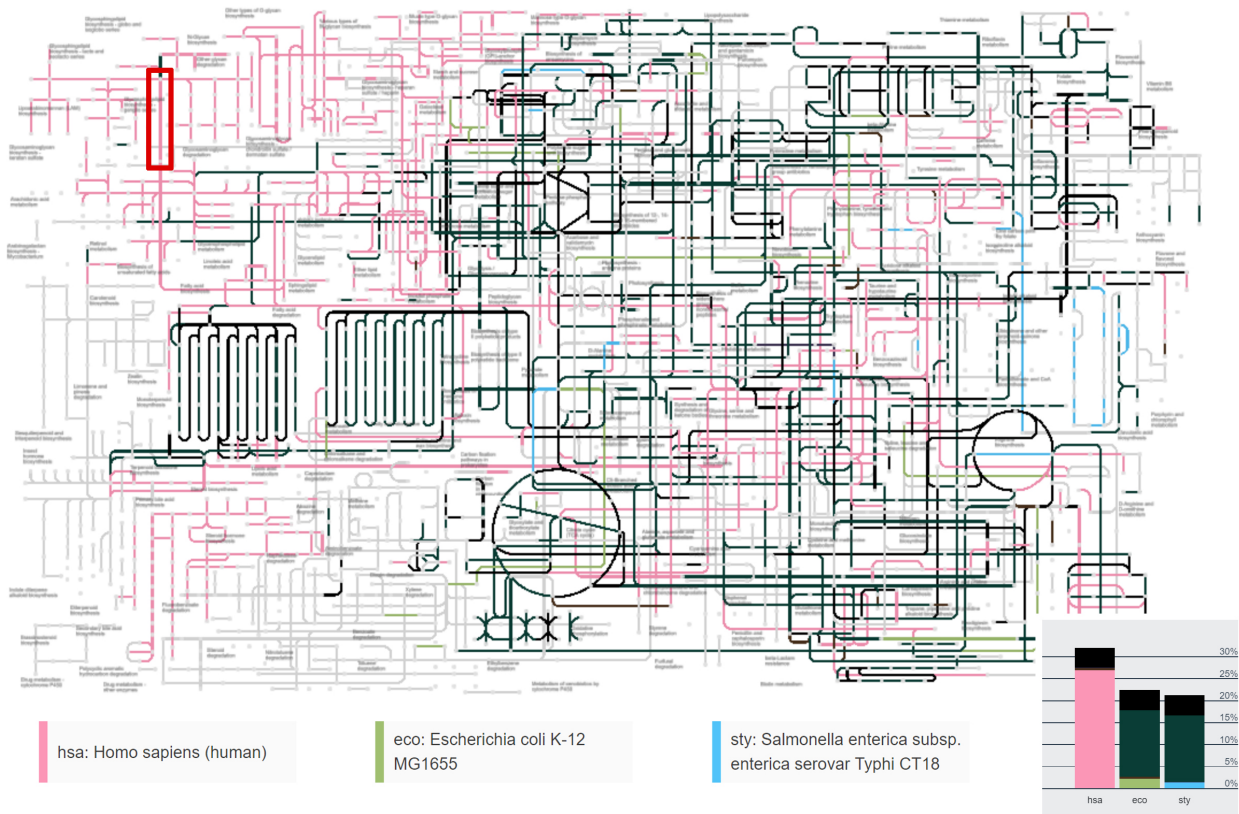


Figure 3: Our edge coloring overlay to compare the human organism with e-coli and salmonella.

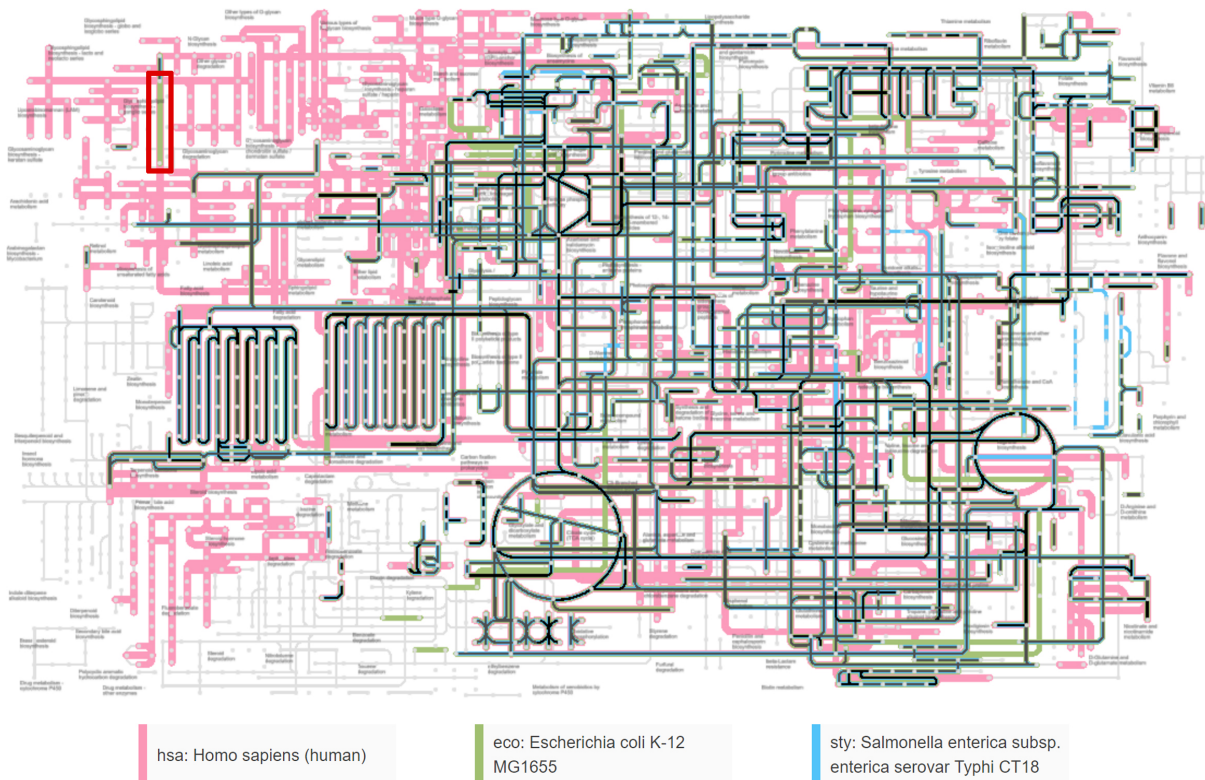


Figure 4: Kelp-inspired, nested overlay to compare the human organism with e-coli and salmonella.

6.1 Use Cases

In this section, we present some use cases. In the first example in Figure 3, we compare the human organism, against one type of e-coli and one type of salmonella. In Figure 4 we present the corresponding nested style visualization as a comparison.

In the overview visualizations in Figure 3 and Figure 4 it is easy to see which elements in the graph are KO entries that only appear in one organism (bright color) or in all organisms (black). Additionally, in Figure 3 the stacked bar charts show the ratio of KO entries per organisms with respect to the total number of KO entries in the map.

The information which entries are shared is encoded differently in the visualizations. In contrast to the edge color style, the nested style with the overlap layer explicitly shows whether KO entries associated with an edge are shared by different organisms. For instance, in the red box in Figure 3 and 4, there is an edge associated with multiple KO entries, which are not shared among human and e-coli. In the nested style this is visible as there is no overlap layer, but still two colors on this edge. In the edge color style, this property is not revealed. We can only see that there is at least one KO entry belonging to the human organism. However, in the color computation of one element, we give priority to the entries which are shared by multiple organisms.

The color blending approach in the edge color visualization already gives the information which organisms share entries. Dark green edges indicate that there are KO entries which are shared by e-coli and salmonella. There also are a few brown edges, where the human and e-coli share KO entries and also some violet edges with KO entries shared by the human and salmonella.

Overall, what we can see in this visualizations in Figure 3 and Figure 4 is that e-coli and salmonella are very similar and that there are only a few entries which differ. The human, however, has a lot of structures not in common with e-coli and salmonella. This are mostly Glycan structures, which are mapped in the upper left corner of the pathway, Lipide in the lower left corner and Amino Acids on the right side.

6.2 Performance

For our performance measurements, we used a PC with 8GB RAM, using an Intel i5 6300 CPU with a built-in Intel HD Graphics 520 GPU with 4GB on Windows 10. We used Google Chrome version 71 and the Chrome Devtools to record the performances when loading the pages. For all websites, client-side caching was disabled. Figure 5 illustrates the page loading times of the KEGG Atlas [24], iPath3 [8], Roche Pathways [25] and our approach. The KEGG Atlas took 8.72s, iPath3 5.19s, Roche Pathways 1.24s and our approach 0.98s. However, these measurements do not properly represent when the actual content is visible to the user.

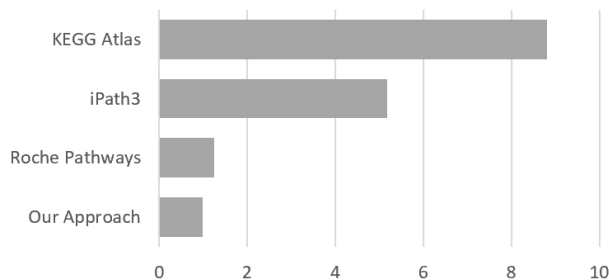


Figure 5: Comparison of loading times in seconds of KEGG Atlas [24], iPath3 [8], Roche Pathways [25] and our approach when loading the page without client-side caching.

Since our approach constantly executes draw calls on the GPU, the rendering performance in terms of FPS is important. Google Chrome limits the FPS to 60, therefore we know that our approach renders with at least 60 FPS.

Unfortunately, requests during runtime may take long, since we have to forward many of the requests to the KEGG servers. Requesting the organism-specific pathway of the human takes about 30s in the KEGG Atlas. Directly requesting it over the KEGG API takes about 9s, whereas loading the resource from our server only takes 0.5s.

6.3 Domain Expert Feedback

For evaluation and testing, our domain expert in genome evolution provided us with the KO entries of a genome retrieved from metagenomic studies, which is not available in the KEGG database. The KO entries of this genome in the context of the global metabolic pathways can be seen in Figure 6. In the standard workflow, our domain expert

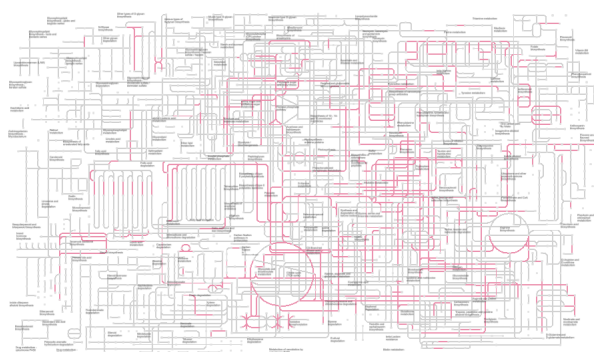


Figure 6: Genome retrieved from metagenomic studies from our domain expert.

compares such a genome with a closely related species to reconstruct its metabolic abilities and to determine the differences and innovations. Our expert performs a global comparison of prokaryotic organisms by gathering public available genomes. Functional annotations and compar-

isons are done with tools from KEGG and other organizations.

Our domain expert was able to gain some interesting findings with our web application. The expert stated that it is clear that in terms of central carbon metabolism there are some similarities but, overall, the metabolic abilities of the genome added to our framework is different to a related genome she compared it against. These are, for instance, lipid and cofactor biosynthesis. Overall, the domain expert was very impressed by the interactive framerates supported by our framework.

6.4 Discussion

In the nested style visualization, it can be seen that we are losing some of the underlying structure of the graph in dense areas. Due to the increasing line width, some of the lines cannot be distinguished from each other. This is the reason why we had to restrict it to three organisms. In the edge color visualization, however, we maintain the original structure of the graph and just recolor the elements which allow us to compare more organisms. However, with both styles, we are losing some of the information.

One drawback of the nested style is that the user does not see which organisms share an entry if the overlay layer indicates so. For both visualizations, we provide additional information like this as a tooltip.

Another challenge of this complex graph layout are overlapping edges throughout the layout. The coloring between our visualizations can be different, due to a changed rendering order of the elements. Static representations like our images can, therefore, be misleading. The highlighting of the elements on hover in our tool shows the complete edge in the foreground without any overlaps of other edges to overcome this issue. But still, the first look can communicate wrong information.

One more thing to mention is that the nested visualization introduces some kind of prioritising of the organisms depending on the order. In the provided example, a lot of emphasis lies on the human organism, not just because it has much more isolated KO entries in the graph, but also due to the selection as the first organism with the highest line width. This kind of emphasis can be desired or not, dependent on the use case.

As mentioned in the Evaluation we added a genome generated from our domain expert to our framework. A problem was that about two-thirds of the KO entries of this genome were not yet mapped to the global KO map provided by KEGG.

7 Conclusion and Future Work

Our pathway visualization tool allows interaction in real time and is more efficient than other existing web applications in this context. We gain this efficiency by rendering to a canvas using WebGL instead of using SVGs as done

by related frameworks. The introduced edge color technique allows the user to identify which related genes are present in the selected organisms. Due to the dense graph layout, it was necessary to come up with a new set overlay visualization. The ambiguous mapping of KO entries to graph elements requires real-time interactive exploration methods to retrieve precise mapping information.

For future work, we consider a more powerful visualization with linked views, allowing to explore different aspects of the relations between multiple organisms. Extended interaction, may be useful, like highlighting the elements of one organism in the visualization when hovering its label. Another interesting extension may be semantic zooming, similar to that of the Roche Pathways [25]. At some level we could, for instance, render the chemical structure of the compounds to support better exploration of the pathway. With our efficient rendering, such dynamic representation changes are seamlessly possible.

According to our collaborating domain expert, a useful future extension might be the possibility to enter the KO entries of genomes, which are not present in the KEGG databases. Additionally, reverse searches, for instance, if given the name of a compound, a highlighting of the KO entries related with this compound, are of interest.

References

- [1] V. Agafonkin. Leaflet: an open-source JavaScript library for mobile-friendly interactive maps. <https://leafletjs.com/>, accessed 26.01.2019.
- [2] B. Alper, N. Riche, G. Ramos, and M. Czerwinski. Design Study of LineSets, a Novel Set Visualization Technique. *IEEE Trans. on Vis. and Comp. Graph.*, (12):2259–2267, 2011.
- [3] B. Alsallakh, L. Micallief, W. Aigner, H. Hauser, S. Miksch, and P. Rodgers. Visualizing Sets and Set-typed Data: State-of-the-Art and Future Challenges. In *Eurographics Conference on Visualization-STARs*, pages 1–21. The Eurographics Association, 2014.
- [4] M. Bostock, V. Ogievetsky, and J. Heer. D³ data-driven documents. *IEEE Trans. on Vis. and Comp. Graph.*, (12):2301–2309, 2011.
- [5] H. Byelas and A. Telea. Visualization of Areas of Interest in Software Architecture Diagrams. In *Proceedings of the 2006 ACM Symposium on Software Visualization*, pages 105–114. ACM, 2006.
- [6] R. Cabello. JavaScript 3D library. <https://threejs.org/>, accessed 27.01.2019.
- [7] C. Collins, G. Penn, and S. Carpendale. Bubble Sets: Revealing Set Relations with Isocontours over Existing Visualizations. *IEEE Trans. on Vis. and Comp. Graph.*, (6):1009–1016, 2009.

- [8] Y. Darzi, T. Yamada, I. Letunic, and P. Bork. iPath3.0: interactive pathways explorer v3. *Nucleic Acids Research*, 46(W1):W510–W513, 04 2018.
- [9] K. Dinkla, M.J. van Kreveld, B. Speckmann, and M.A. Westenberg. Kelp Diagrams: Point Set Membership Visualization. In *Comp. Graph. Forum*, volume 31, pages 875–884. Wiley Online Library, 2012.
- [10] E.R. Gansner, Y. Hu, and S.G. Kobourov. GMap: Drawing Graphs as Maps. In *International Symposium on Graph Drawing*, pages 405–407. Springer, 2009.
- [11] R. Gove, J. Saxe, S. Gold, A. Long, and G. Bergamo. SEEM: a scalable visualization for comparing multiple large sets of attributes for malware analysis. In *Proceedings of the Eleventh Workshop on Visualization for Cyber Security*, pages 72–79. ACM, 2014.
- [12] Khronos WebGL Working Group. OpenGL ES for the Web. <https://www.khronos.org/webgl/>, accessed 27.01.2019.
- [13] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Research*, 45(D1):D353–D361, 2016.
- [14] M. Kanehisa, S. Goto, S. Kawashima, Y. Okuno, and M. Hattori. The KEGG Resource for Deciphering the Genome. *Nucleic Acids Research*, 32(suppl_1):D277–D280, 2004.
- [15] M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research*, 44(D1):D457–D462, 2015.
- [16] N. Kono, K. Arakawa, R. Ogawa, N. Kido, K. Oshita, K. Ikegami, S. Tamaki, and M. Tomita. Pathway Projector: Web-Based Zoomable Pathway Browser Using KEGG Atlas and Google Maps API. *PLOS One*, 4(11):e7710, 2009.
- [17] Kanehisa Laboratories. KEGG API. <https://www.kegg.jp/kegg/rest/keggapi.html>, accessed 16.10.2018.
- [18] Kanehisa Laboratories. KEGG Organism List. <http://rest.kegg.jp/list/organism>, accessed 04.02.2019.
- [19] I. Letunic, T. Yamada, M. Kanehisa, and P. Bork. iPath: interactive exploration of biochemical pathways and networks. *Trends in Biochemical Sciences*, 33(3):101103, March 2008.
- [20] A. Lex, N. Gehlenborg, H. Strobel, R. Vuillemot, and H. Pfister. UpSet: Visualization of Intersecting Sets. *IEEE Trans. on Vis. and Comp. Graph.*, 20(12):1983–1992, 2014.
- [21] W. Meulemans, N. Riche, B. Speckmann, B. Alper, and T. Dwyer. KelpFusion: A hybrid set visualization technique. *IEEE Trans. on Vis. and Comp. Graph.*, 19(11):1846–1858, 2013.
- [22] A. Noronha, A.D. Daniélsdóttir, P. Gawron, F. Jóhannsson, S. Jónsdóttir, S. Jarlsson, J.P. Gunnarsson, S. Brynjólfsson, R. Schneider, I. Thiele, et al. ReconMap: an interactive visualization of human metabolism. *Bioinformatics*, 33(4):605–607, 2016.
- [23] H. Ogata, S. Goto, K. Sato, W. Fujibuchi, H. Bono, and M. Kanehisa. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 27(1):29–34, 01 1999.
- [24] S. Okuda, T. Yamada, M. Hamajima, M. Itoh, T. Katayama, P. Bork, S. Goto, and M. Kanehisa. KEGG Atlas mapping for global analysis of metabolic pathways. *Nucleic Acids Research*, 36(suppl_2):W423–W426, 2008.
- [25] F. Hoffmann-La Roche. Roche Biochemical Pathways. <http://biochemical-pathways.com/#/map/1>, accessed 23.01.2019.
- [26] P. Rodgers, G. Stapleton, B. Alsallakh, L. Michalief, R. Baker, and S. Thompson. A task-based evaluation of combined set and network visualization. *Information Sciences*, 367:58–79, 2016.
- [27] A. Ronacher. Flask. Web Development, one drop at a time. <http://flask.pocoo.org/>, accessed 27.01.2019.
- [28] P. Simonetto. *Visualisation of overlapping sets and clusters with Euler diagrams*. PhD thesis, Université Sciences et Technologies-Bordeaux I, 2012.
- [29] K. Unger. Visual Comparison of Organism-Specific Metabolic Pathways. <https://users.cg.tuwien.ac.at/kunger/pr-pathway-exploration/>, accessed 04.02.2019.
- [30] C. Vehlow, F. Beck, and D. Weiskopf. The State of the Art in Visualizing Group Structures in Graphs. In *Eurographics Conference on Visualization-STARs*, volume 2. The Eurographics Association, 2015.
- [31] T. Yamada, I. Letunic, S. Okuda, M. Kanehisa, and P. Bork. iPath2.0: interactive pathway explorer. *Nucleic Acids Research*, 39(suppl_2):W412–W415, 2011.
- [32] A. Zeileis, K. Hornik, and P. Murrell. Escaping RGBland: Selecting Colors for Statistical Graphics. *Computational Statistics & Data Analysis*, 53(9):3259–3270, 2009.