


## Vias biológicas

By **Lissur Orsine**

21 de junho de 2021

### Vias biológicas

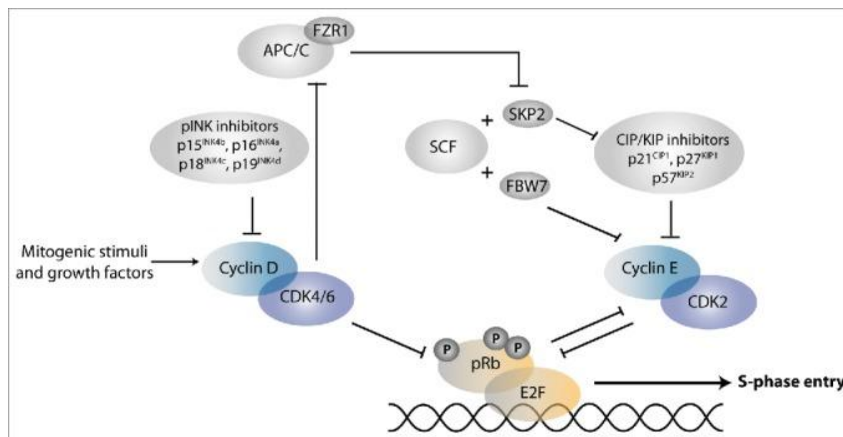
Lissur Orsine , Elisa Donnard 

Revisão: Diego Mariano 

BIOINFO – Revista Brasileira de Bioinformática. Edição #01. Julho, 2021.

DOI: 10.51780/978-6-599-275326-12

Uma via *biológica* é uma anotação de um conjunto de interações moleculares ocorrendo dentro de um sistema biológico. Sistema biológico pode se referir a uma célula, um tecido, um órgão ou, até mesmo, o organismo como um todo. As vias são compostas por entidades biológicas, que podem ser genes, enzimas, metabólitos, outras vias; e por interações entre as entidades biológicas, que podem ser positivas, negativas ou neutras. Por exemplo, a Figura 1 mostra a via de proliferação/diferenciação, isto é, o mecanismo molecular pelo qual uma célula precursora decide entre continuar em um estado indiferenciado ou se comprometer com um tipo celular específico. Neste caso, as entidades biológicas são representadas por elipses e as interações entre elas por setas ou barras em T.



**Figura 1.** Via de proliferação/diferenciação. Sob a influência de estímulos mitogênicos e fatores de crescimento, há a ativação dos complexos CDK-ciclina (elipses azuis) que são reguladores-mestre do processo de proliferação/diferenciação. Estes complexos são responsáveis por enfraquecer a ligação de pRb à E2F (elipses laranja) conduzindo a célula precursora à fase S do ciclo celular (ou seja, à proliferação). Em contrapartida, pINK e CIP/KIP (elipses cinzas) são inibidores dos complexos CDK-ciclina impedindo a progressão do ciclo celular e favorecendo a diferenciação. Já os complexos APC/C e SCF (elipses cinzas) podem atuar de formas distintas a depender de a quais fatores estão associados. Retirado de [1].

As vias biológicas mais popularmente conhecidas são as vias metabólicas, como aquelas presentes no mapa metabólico, o conjunto de todas as reações bioquímicas do metabolismo (veja um mapa metabólico interativo [aqui](#)) [2]. No entanto, existem outros tipos de vias biológicas, como as vias de regulação gênica, que representam a ativação ou repressão da transcrição de genes, e as vias de sinalização, que representam um fluxo de informação dentro ou entre células realizado através de interações entre proteínas. Ainda, as vias biológicas podem ser compostas por trechos de diferentes tipos, como no caso da via de proliferação/diferenciação mostrada acima.

## Vias e Redes

O conceito de via é associado a outro conceito conhecido na Biologia de Sistemas: as redes biológicas. Uma rede *biológica* também é uma representação de um conjunto de interações moleculares ocorrendo dentro de um sistema biológico. De fato, vias e redes são abordagens complementares dentro da Biologia de Sistemas. No entanto, elas apresentam particularidades que as tornam mais apropriadas a um ou outro tipo de estudo.

Nas vias, as interações moleculares apresentam caráter direcional, sendo possível delimitar um começo e um fim, uma entrada e uma saída, uma causa e um efeito. O próprio termo *via* (em inglês, *pathway*) remete à ideia de algo que conduz a algum lugar. O mesmo não ocorre com as redes, onde as interações moleculares não são obrigatoriamente sequenciais.

Além disso, nas vias as interações moleculares são atribuídas por especialistas, o que significa que elas representam conhecimento científico consolidado. Por outro lado, as redes são geradas a partir de experimentos em larga escala, como por exemplo ensaios de imunoprecipitação de proteínas que interagem entre si, que por sua natureza também incluem falso-positivos. Por isso, as redes geradas a partir deste tipo de dado podem conter interações moleculares ainda não descritas, assim como interações incorretas que ainda não foram identificadas como tal. Essa diferença na fonte dos dados faz com que, de um modo geral, a informação contida nas vias seja mais precisa, enquanto que, nas redes, mais abrangente.

### Para que serve uma via?

Talvez a aplicação mais conhecida das vias na Bioinformática seja em análises de enriquecimento. Neste contexto, o pesquisador dispõe de uma lista de genes/produtos gênicos a qual ele deseja atribuir significado, isto é, ele deseja responder a questões do tipo “*Que genes/produtos gênicos são estes?*”, “*Qual o papel molecular/celular destes genes/produtos gênicos?*” ou “*Como esses genes/produtos gênicos se relacionam?*”. Nas análises de enriquecimento, busca-se avaliar quais vias conhecidas contêm um número de genes/produtos gênicos da lista de interesse maior do que o esperado por acaso. Por exemplo, uma lista de genes com a transcrição aumentada

poucas horas após o reconhecimento de um patógeno por uma célula dendrítica humana do sistema imune inato mostra enriquecimento para genes presentes em vias de sinalização inflamatória [3]. Este tipo de análise depende da existência de bases de dados curadas contendo vias biológicas (discutidas abaixo) e permite uma rápida interpretação de resultados experimentais, ao associar genes no contexto de seus papéis funcionais conhecidos.

As vias também têm sido empregadas na visualização de dados ômicos e séries temporais, tendo como objetivo novamente facilitar a interpretação dos mesmos. Alguns exemplos são:

1. usando uma via relevante ao estudo, por exemplo a via de diferenciação/proliferação da Figura 1, os genes que foram identificados como superexpressos ou subexpressos em uma certa condição (como uma condição patológica [4] ou após tratamento com uma droga [5]) podem ser coloridos para revelar por exemplo uma perturbação biológica em um ponto específico de uma via, implicando um ou mais genes no controle da resposta observada;
2. pode-se assinalar todos os genes presentes em uma via que contém polimorfismos (SNPs) associados a uma ou mais doenças, revelando possíveis relações causais ou de dependência entre eles [6];
3. os genes de uma via podem ser coloridos de forma a sinalizar sua origem evolutiva, o que permite identificar ramos específicos que surgiram em um determinado ancestral ou que estão ausentes em certos organismos [7].

Vale lembrar que o propósito primordial das vias é retratar os mecanismos moleculares por trás de um fenômeno biológico. Assim, as vias ajudam a entender, do ponto de vista molecular, como uma dada estrutura morfológica ou um dado processo fisiológico, por exemplo, vieram a ser o que são.

## Bases de dados

Existem diversas bases de dados contendo vias. A mais popular delas é o [KEGG Pathway](#) [8, 9, 10] que abriga uma coleção de vias manualmente curadas, sendo considerada, por isso, uma base de dados de alta confiabilidade. O [Reactome](#) [11], por sua vez, se destaca por oferecer, além de um grande número de vias, um grande número de ferramentas de análise. Já o [WikiPathways](#) [12] é uma base de dados colaborativa (formato *wiki*) onde qualquer usuário registrado pode adicionar ou editar conteúdo.

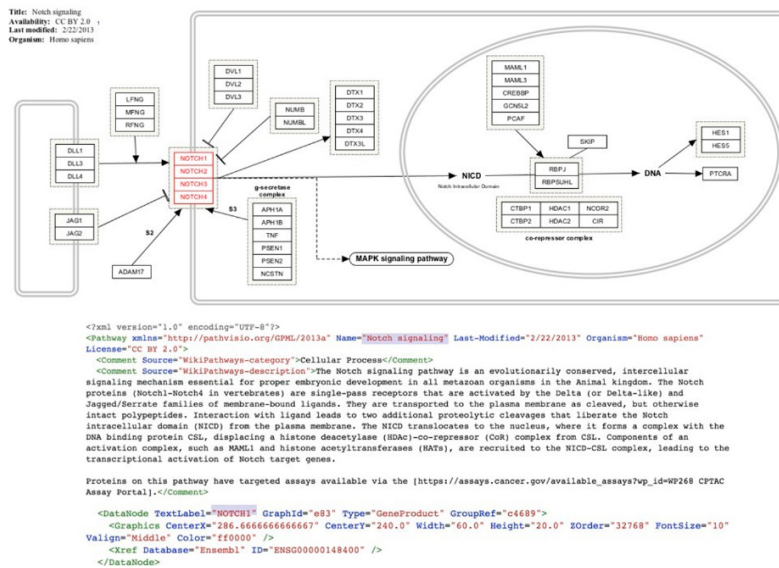
## Programas

Existem diversos programas para o desenho, a edição e a análise de vias. Embora inicialmente concebido para o estudo de redes, atualmente o [Cytoscape](#) [13] conta com a opção de visualizar e analisar vias advindas das bases de dados KEGG Pathway, Reactome e WikiPathways [14, 15, 16]. Com uma interface gráfica de fácil uso, o [PathVisio](#) [17, 18] é um programa mais voltado para o desenho de vias, sendo o programa de escolha da base de dados WikiPathways (tanto é possível fazer *upload* de vias criadas no

PathVisio para o WikiPathways quanto fazer *download* de vias disponíveis no WikiPathways para o PathVisio). Já o **CellDesigner** [19, 20] possui como diferencial a modelagem de vias.

## Formatos

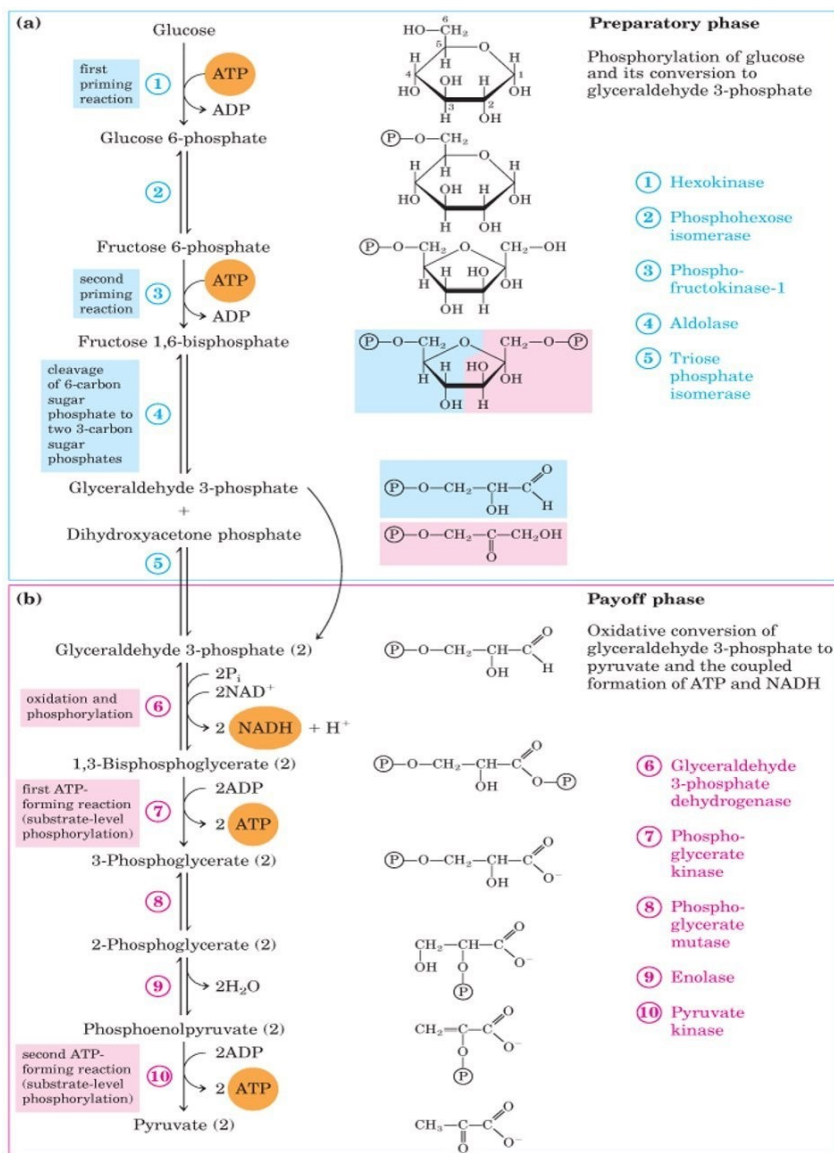
Além dos tradicionais formatos de imagem (PNG, SVG, PDF), vários formatos foram desenvolvidos para permitir a manipulação automática das vias [21]. Dentre eles, vale destacar os formatos *XML-based*, tais como o **KEGG Markup Language (KGML)** [22], implementado na base de dados KEGG Pathway; o **Graphical Pathway Markup Language (GPML)** [23], implementado na base de dados WikiPathways; o **Systems Biology Markup Language (SBML)** [24, 25]; e o **Biological Pathways eXchange (BioPAX)** [26]. A Figura 2 mostra um trecho do arquivo GPML correspondente à via de sinalização Notch. É possível notar que o arquivo GPML contém informações como o nome e a descrição da via, a espécie a que a via se refere, a anotação das entidades biológicas, dentre outras. Este tipo de padronização possibilita, por exemplo, encontrar em quais vias um dado gene/produto gênico está presente ou colorir as vias segundo um dado critério.



**Figura 2.** Diagrama e arquivo GPML para a via de sinalização Notch. Adaptado de [27].

## Simbologia

Há diversas formas de representar uma via. A Figura 3, por exemplo, mostra três versões da via glicolítica, via central no metabolismo de carboidratos. A primeira delas é a representação presente em um livro de bioquímica [28] (Figura 3), onde os substratos/produtos estão identificados por nome e estrutura química, as enzimas por números com os respectivos nomes exibidos à direita, e as reações por setas.



**Figura 3.** Representação da via glicolítica presente em um livro de bioquímica [28].

Logo após é exibida a representação da mesma via segundo KEGG Pathway (Figura 4). Neste caso, os círculos simbolizam os compostos químicos e os retângulos as enzimas (identificadas por seus *EC numbers* [29]), enquanto as setas correspondem às reações.

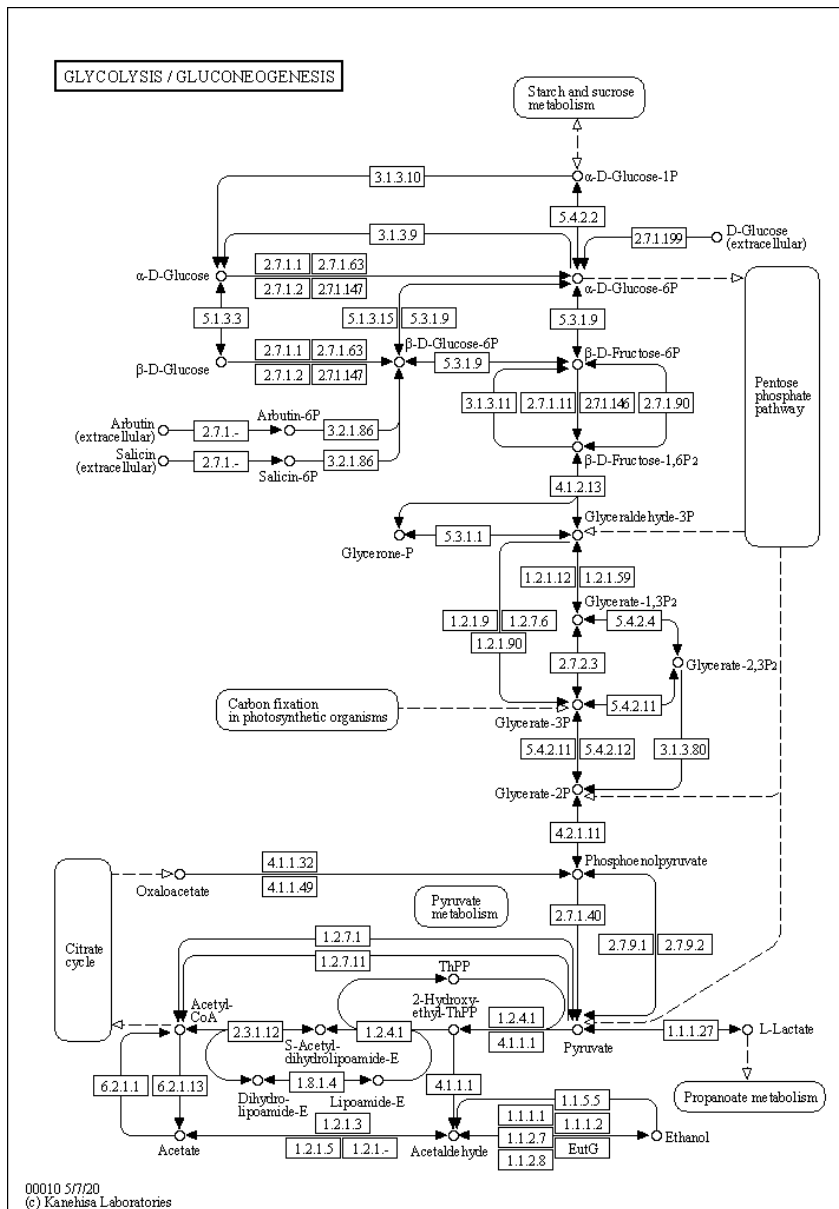


Figura 4. Representação da via glicolítica segundo KEGG Pathway [8, 9, 10].

Já a última versão é do *Systems Biology Graphical Notation* (SBGN) [30] (Figura 5), onde as moléculas são mostradas como círculos, as enzimas como retângulos arredondados associados às reações, e as reações em si como setas. O SBGN é uma iniciativa no sentido de estabelecer uma simbologia padronizada para a representação de vias biológicas, o que facilita, por sua vez, a integração de informações providas de fontes distintas. Outra convenção disponível é o *Molecular Interaction Map* (MIM) [31, 32].

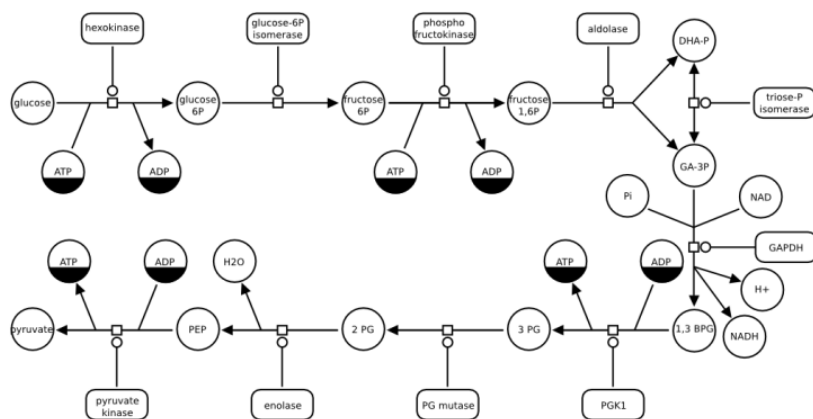


Figura 5. Representação da via glicolítica segundo SBGN [30].

## Considerações finais

As vias biológicas cumprem o importante papel de descrever, do ponto de vista molecular, os sistemas biológicos. Dessa forma, elas ajudam a entender como estes sistemas funcionam normalmente, e quais alterações ocorrem em caso de doença, ou até mesmo qual é o impacto de uma droga. Tradicionalmente, pesquisas envolvendo vias biológicas eram feitas manualmente e se concentravam em áreas como Bioquímica ou Biologia Molecular. O estudo das vias a partir da perspectiva da Bioinformática, em consequência da geração acelerada de dados em larga escala com novas técnicas moleculares, proporcionou uma melhor e mais eficiente aplicação do conhecimento biológico consolidado em vias, além da expansão das bases de dados existentes.

## Referências bibliográficas

- [1] Ruijtenberg, S., & van den Heuvel, S. (2016). Coordinating cell proliferation and differentiation: Antagonism between cell cycle regulators and cell type-specific gene expression. *Cell cycle (Georgetown, Tex.)*, 15(2), 196–212. <https://doi.org/10.1080/15384101.2015.1120925>
- [2] Michael, G.. (2014, Jan 1st). *Biochemical Pathways*. Roche. Retrieved from <http://biochemical-pathways.com/>
- [3] Amit, I., Garber, M., Chevrier, N., Leite, A. P., Donner, Y., Eisenhaure, T., Guttman, M., Grenier, J. K., Li, W., Zuk, O., Schubert, L. A., Birditt, B., Shay, T., Goren, A., Zhang, X., Smith, Z., Deering, R., McDonald, R. C., Cabili, M., Bernstein, B. E., ... Regev, A. (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. *Science (New York, N.Y.)*, 326(5950), 257–263. <https://doi.org/10.1126/science.1179050>

[4] Collino, A., Termanini, A., Nicoli, P., Diaferia, G., Polletti, S., Recordati, C., Castiglioni, V., Caruso, D., Mitro, N., Natoli, G., & Ghisletti, S. (2018). Sustained activation of detoxification pathways promotes liver carcinogenesis in response to chronic bile acid-mediated damage. *PLOS Genetics*, 14(5). <https://doi.org/10.1371/journal.pgen.1007380>

[5] Shah, K. B., Tripathy, S., Suganthi, H., & Rudraiah, M. (2014). Profiling of Luteal Transcriptome during Prostaglandin F<sub>2</sub>-Alpha Treatment in Buffalo Cows: Analysis of Signaling Pathways Associated with Luteolysis. *PLoS ONE*, 9(8). <https://doi.org/10.1371/journal.pone.0104127>

[6] Bronson, P. G., Chang, D., Bhangale, T., Seldin, M. F., Ortmann, W., Ferreira, R. C., Urcelay, E., Pereira, L. F., Martin, J., Plebani, A., Lougaris, V., Friman, V., Freiburger, T., Litzman, J., Thon, V., Pan-Hammarström, Q., Hammarström, L., Graham, R. R., & Behrens, T. W. (2016). Common variants at PVT1, ATG13-AMBRA1, AH11 and CLEC16A are associated with selective IgA deficiency. *Nature Genetics*, 48(11), 1425–1429. <https://doi.org/10.1038/ng.3675>

[7] Fani, R. (2012). The Origin and Evolution of Metabolic Pathways: Why and How did Primordial Cells Construct Metabolic Routes? *Evolution: Education and Outreach*, 5(3), 367–381. <https://doi.org/10.1007/s12052-012-0439-5>

[8] Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., & Tanabe, M. (2021). KEGG: integrating viruses and cellular organisms. *Nucleic acids research*, 49(D1), D545–D551. <https://doi.org/10.1093/nar/gkaa970>

[9] Kanehisa M. (2019). Toward understanding the origin and evolution of cellular organisms. *Protein science : a publication of the Protein Society*, 28(11), 1947–1951. <https://doi.org/10.1002/pro.3715>

[10] Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1), 27–30. <https://doi.org/10.1093/nar/28.1.27>

[11] Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R., Loney, F., May, B., Milacic, M., Rothfels, K., Sevilla, C., Shamovsky, V., Shorsler, S., Varusai, T., Weiser, J., Wu, G., ... D'Eustachio, P. (2020). The reactome pathway knowledgebase. *Nucleic acids research*, 48(D1), D498–D503. <https://doi.org/10.1093/nar/gkz1031>

[12] Martens, M., Ammar, A., Riutta, A., Waagmeester, A., Slenter, D. N., Hanspers, K., A Miller, R., Digles, D., Lopes, E. N., Ehrhart, F., Dupuis, L. J., Winckers, L. A., Coort, S. L., Willighagen, E. L., Evelo, C. T., Pico, A. R., & Kutmon, M. (2021). WikiPathways: connecting communities. *Nucleic acids research*, 49(D1), D613–D621. <https://doi.org/10.1093/nar/gkaa1024>

[13] Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: a software environment for



integrated models of biomolecular interaction networks. *Genome research*, 13(11), 2498–2504. <https://doi.org/10.1101/gr.1239303>

[14] Nishida, K., Ono, K., Kanaya, S., & Takahashi, K. (2014). KEGGscape: a Cytoscape app for pathway data integration. *F1000Research*, 3, 144. <https://doi.org/10.12688/f1000research.4524.1>

[15] Wu, G., Feng, X., & Stein, L. (2010). A human functional protein interaction network and its application to cancer data analysis. *Genome biology*, 11(5), R53. <https://doi.org/10.1186/gb-2010-11-5-r53>

[16] Kutmon, M., Lotia, S., Evelo, C. T., & Pico, A. R. (2014). WikiPathways App for Cytoscape: Making biological pathways amenable to network analysis and visualization. *F1000Research*, 3, 152. <https://doi.org/10.12688/f1000research.4254.2>

[17] Kutmon, M., van Iersel, M. P., Bohler, A., Kelder, T., Nunes, N., Pico, A. R., & Evelo, C. T. (2015). PathVisio 3: an extendable pathway analysis toolbox. *PLoS computational biology*, 11(2), e1004085. <https://doi.org/10.1371/journal.pcbi.1004085>

[18] van Iersel, M. P., Kelder, T., Pico, A. R., Hanspers, K., Coort, S., Conklin, B. R., & Evelo, C. (2008). Presenting and exploring biological pathways with PathVisio. *BMC bioinformatics*, 9, 399. <https://doi.org/10.1186/1471-2105-9-399>

[19] Funahashi, A., Matsuoka, Y., Jouraku, A., Morohashi, M., Kikuchi, N., & Kitano, H. (2008). CellDesigner 3.5: A Versatile Modeling Tool for Biochemical Networks. *Proceedings of the IEEE*, 96(8), 1254–1265. <https://doi.org/10.1109/jproc.2008.925458>

[20] Funahashi, A., Morohashi, M., Kitano, H., & Tanimura, N. (2003). CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *BIOSILICO*, 1(5), 159–162. [https://doi.org/10.1016/s1478-5382\(03\)02370-9](https://doi.org/10.1016/s1478-5382(03)02370-9)

[21] Viswanathan, G. A., Seto, J., Patil, S., Nudelman, G., & Sealfon, S. C. (2008). Getting started in biological pathway construction and analysis. *PLoS computational biology*, 4(2), e16. <https://doi.org/10.1371/journal.pcbi.0040016>

[22] Kanehisa Laboratories. (2016, Aug 29) KGML (KEGG Markup Language). KEGG: Kyoto Encyclopedia of Genes and Genomes. Retrieved from <https://www.kegg.jp/kegg/xml/>

[23] Finterly, Riutta, A., Kumar, S., Summer-Kutmon, M., Willighagen, E., & Pico, A.. (2017, Jan 7) PathVisio/GPML. GitHub. Retrieved from <https://github.com/PathVisio/GPML>

[24] Keating, S. M., Waltemath, D., König, M., Zhang, F., Dräger, A., Chaouiya, C., Bergmann, F. T., Finney, A., Gillespie, C. S., Helikar, T., Hoops, S., Malik-Sheriff, R. S., Moodie, S. L., Moraru, I. I., Myers, C. J., Naldi, A., Olivier, B. G., Sahle, S., Schaff, J. C., Smith, L. P., ... SBML Level 3 Community members (2020). SBML Level 3: an extensible format for the exchange and reuse of biological models. *Molecular systems biology*, 16(8), e9110. <https://doi.org/10.15252/msb.20199110>

[25] Hucka, M., Finney, A., Sauro, H. M., Bolouri, H., Doyle, J. C., Kitano, H., Arkin, A. P., Bornstein, B. J., Bray, D., Cornish-Bowden, A., Cuellar, A. A., Dronov, S., Gilles, E. D., Ginkel, M., Gor, V., Goryanin, I. I., Hedley, W. J., Hodgman, T. C., Hofmeyr, J. H., Hunter, P. J., ... SBML Forum (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics (Oxford, England)*, 19(4), 524–531. <https://doi.org/10.1093/bioinformatics/btg015>

[26] Demir, E., Cary, M. P., Paley, S., Fukuda, K., Lemer, C., Vastrik, I., Wu, G., D'Eustachio, P., Schaefer, C., Luciano, J., Schacherer, F., Martinez-Flores, I., Hu, Z., Jimenez-Jacinto, V., Joshi-Tope, G., Kandasamy, K., Lopez-Fuentes, A. C., Mi, H., Pichler, E., Rodchenkov, I., ... Bader, G. D. (2010). The BioPAX community standard for pathway data sharing. *Nature biotechnology*, 28(9), 935–942. <https://doi.org/10.1038/nbt.1666>

[27] Hanspers, K., Kelder, T., Pico, A., Salomonis, N., Willighagen, E., Weitz, E., & Ehrhart, F.. (2007, May 22). Notch signaling (Homo sapiens). WikiPathways. Retrieved from <https://www.wikipathways.org/index.php/Pathway:WP268>

[28] Nelson, D. L., Cox, M. M.. (2004). *Lehninger Principles of Biochemistry* (4th ed.). New York, NY: W. H. Freeman. ISBN 978-0716743392

[29] Webb, E. C. (1992). *Enzyme nomenclature 1992: recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzymes*. Published for the International Union of Biochemistry and Molecular Biology by Academic Press. ISBN 978-0-12-227164-9

[30] Le Novère, N., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A., Demir, E., Wegner, K., Aladjem, M. I., Wimalaratne, S. M., Bergman, F. T., Gauges, R., Ghazal, P., Kawaji, H., Li, L., Matsuoka, Y., Villéger, A., Boyd, S. E., Calzone, L., Courtot, M., ... Kitano, H. (2009). The Systems Biology Graphical Notation. *Nature biotechnology*, 27(8), 735–741. <https://doi.org/10.1038/nbt.1558>

[31] Luna, A., Karac, E. I., Sunshine, M., Chang, L., Nussinov, R., Aladjem, M. I., & Kohn, K. W. (2011). A formal MIM specification and tools for the common exchange of MIM diagrams: an XML-Based format, an API, and a validation method. *BMC bioinformatics*, 12, 167. <https://doi.org/10.1186/1471-2105-12-167>

[32] Kohn, K. W., Aladjem, M. I., Weinstein, J. N., & Pommier, Y. (2006). Molecular interaction maps of bioregulatory networks: a general rubric for systems

biology. *Molecular biology of the cell*, 17(1), 1–13.  
<https://doi.org/10.1091/mbc.e05-09-0824>