

EndoNet: an information resource about endocrine networks

Anatolij Potapov^{1,*}, Ines Liebich^{1,2}, Jürgen Dönitz¹, Knut Schwarzer¹, Nicole Sasse¹,
Torsten Schoeps¹, Torsten Crass¹ and Edgar Wingender^{1,2}

¹Department of Bioinformatics, University of Göttingen/Medical School, Goldschmidtstrasse 1, D-37077 Göttingen, Germany and ²BIOBASE GmbH, Halchtersche Strasse 33, D-38304 Wolfenbüttel, Germany

Received August 15, 2005; Revised and Accepted October 20, 2005

ABSTRACT

EndoNet is a new database that provides information about the components of endocrine networks and their relations. It focuses on the endocrine cell-to-cell communication and enables the analysis of intercellular regulatory pathways in humans. In the EndoNet data model, two classes of components span a bipartite directed graph. One class represents the hormones (in the broadest sense) secreted by defined donor cells. The other class consists of the acceptor or target cells expressing the corresponding hormone receptors. The identity and anatomical environment of cell types, tissues and organs is defined through references to the CYTOMER[®] ontology. With the EndoNet user interface, it is possible to query the database for hormones, receptors or tissues and to combine several items from different search rounds in one complex result set, from which a network can be reconstructed and visualized. For each entity, a detailed characteristics page is available. Some well-established endocrine pathways are offered as showcases in the form of predefined result sets. These sets can be used as a starting point for a more complex query or for obtaining a quick overview. The EndoNet database is accessible at <http://endonet.bioinf.med.uni-goettingen.de/>.

INTRODUCTION

Theoretical analyses in the post-sequencing era, in particular in the context of systems biology approaches, increasingly investigate the properties of all kinds of pathways and networks such as metabolic and signaling pathways. It is commonly accepted that we need formal descriptions of these networks to make systematic use of the overwhelming body

of facts gathered over decades of laboratory work both in narrow or global scale.

Corresponding databases have been created and are available now for metabolic networks (KEGG) (1,2), protein interaction networks (BIND and DIP) (3,4) and signaling pathways (CSNDB, Patika and TRANSPATH[®]) (5–8), just to name a few. So far, however, their main focus is on intracellular processes. Intercellular signaling is addressed only insofar as usually the pathways modeled start with extracellular ligands, and as there exist catalogs and databases about secreted proteins, or the ‘secretome’, of certain systems (9–12).

This shortcoming is part of the more comprehensive problem, the genotype–phenotype gap: from a certain genotype, we are able to infer a ‘molecular phenotype’, but for correlating it with a more complex phenotype such as a biological process, or even a certain disease and its clinical appearance, we still depend largely on the mere description of an observed correlation. There is no way to infer such a phenotype through all the different layers of increasing complexity between genomic DNA sequences and the physiological function of whole organs and their interplay within an organism.

There may be principal barriers preventing such an inference across different complexity levels, but even to explore these limits, we have to make attempts to bridge the genotype–phenotype gap. We have to do the next step towards modeling intercellular networks that are inextricably linked to the physiology of multicellular organisms (13). Being one of the most complex constructs in the body, the endocrine system comprises numerous cells and tissues that secrete hormones which pass through the body, activate specific receptors of target cells and initiate there multiple intracellular signaling pathways.

Here, we present a new database, EndoNet, which provides information about the components of endocrine networks and their relations, and enables the analysis of intercellular regulatory pathways in humans. The EndoNet database is accessible at <http://endonet.bioinf.med.uni-goettingen.de/>.

*To whom correspondence should be addressed. Tel: +49 0 551 39 14917; Fax: +49 0 551 39 14914; Email: anatolij.potapov@bioinf.med.uni-goettingen.de

RESULTS

The biological schema of endocrine actions

Development and function of different organs as well as the response of a whole multicellular organism to its environment is coordinated through a complex communication system between specialized cells that are part of its organs. This communication is mostly mediated by hormones. In a broader sense, this functional class of biomolecules also comprises growth factors, cytokines, chemokines and other signal transmitters. This generic view is supported, for instance, by the definition given for 'hormone activity' by Gene Ontology (GO) (14): 'Any substance formed in very small amounts in one specialized organ or group of cells and carried (sometimes in the bloodstream) to another organ or group of cells, in the same organism, upon which it has a specific regulatory action'. This definition is also broad enough to include modes of hormonal actions as diverse as endocrine, paracrine and autocrine effects. Accordingly, with the term 'hormone' one might refer to any extracellular substance that induces specific responses in target cell and helps to coordinate growth, differentiation, gene expression and metabolic activities of various cells, tissues and organs in multicellular organisms (15).

Hormones can be classified based on their chemical nature, solubility, the distance over which the signal acts and so on (15,16). From the viewpoint of genome-phenotype relations, it is reasonable to distinguish between polypeptide, thus, genome-encoded hormones, on one hand, and those low-molecular weight hormones such as steroids, with only the machinery of their synthesis being genome-encoded, on the other hand. Another classification of hormones, which seems to be overlapping with the previous one refers to the intracellular location of their receptors and, thus, how the subsequent signal is further transduced: membrane-bound receptors usually trigger more or less complex signaling cascades towards the nucleus, whereas nuclear receptors, mainly bound by low-molecular weight hormones, have a very short signaling pathway downstream since they act as transcription factors themselves (15,16).

In intercellular communication, we can basically differentiate between two kinds of cells: donor cells which synthesize and secrete a hormone, and acceptor cells which express a hormone receptor (Figure 1a). Donor cells become active under the influence of an external, mostly environmental, stimulus. In the acceptor cell, binding of the hormone to a receptor triggers an intracellular signal transduction cascade with different kinds of end nodes and effects: transcription factors affecting the gene expression program of the acceptor cell, metabolic enzymes controlling the cell's metabolism, structural components which define the acceptor's morphological features, or components of the secretory apparatus regulating the release of other extracellular molecules. If synthesis and secretion of another hormone is among the effects exerted by receptor activation, the acceptor is turned into a donor cell, thus becoming an internal node of the organism's endocrine network. Acceptor cells which do not become producers of another hormone are called 'terminal target cells' of the endocrine network, but finally constitute the overall physiological effect of the respective hormonal pathway, or simply the phenotype (Figure 1a).

The EndoNet data model

In the EndoNet data model, two classes of entities—hormones (in the broadest sense) and their acceptor or target cells expressing the corresponding receptors—span a bipartite directed graph. Since one and the same hormone may be secreted by multiple cell types (donor cells), each such secretion event is represented by a hormone node on its own. Similarly, each cell type known to express a hormone receptor (acceptor cell) leads to an individual node. The graph's edges represent hormone transport and binding to a receptor (intercellular edges), on one hand, and triggering or inhibition of hormone secretion by a receptor activated by hormone binding (intracellular edges), on the other hand. Optionally, an edge representing the transport of a hormone can be subdivided by introducing the transport medium (usually blood) as an additional, intermediary node.

Thus, in the conceptual schema of the EndoNet database (Figure 1b), the links between cells/organs and hormone define donor cells ('D'), those between cells/organs and receptors acceptor cells ('A'). If an acceptor cell synthesizes another hormone in response to an incoming signal, it becomes an internal node in the emerging hormonal network.

In EndoNet, the pathway between a hormone receptor expressed in an acceptor cell and a hormone synthesized in the same cell (intracellular edge) is handled as a black box. In case of genome-encoded peptide hormones, cross-references to entries in the TRANSPATH® database, which describe the signaling cascade starting from the hormone's receptor and ending at the hormone's gene, are provided, if available. Datasets on non-peptide hormones will in future be enriched by a specific metabolic add-on which will include references to the databases KEGG (1) and BRENDA (17), allowing for further characterization of the steps performed during the hormone's synthesis and the regulation of both activity and expression of the enzymes involved.

By now, the EndoNet structure already allows for including descriptions of the physiological effects induced by hormone binding (see below, Future Developments).

The contents of EndoNet

In the present version of EndoNet, and as a first approach, we consider the endocrine (hormonal) network of the human body. For each molecule (hormone or receptor), a primary name and synonyms are given. In case of peptide hormones, the sequence of the processed polypeptide, rather than that of the protein precursor, is specified. For a multimeric protein hormone, the subunit composition as well as the sequences of all subunits are stored; the same holds true for hormone receptors.

Additionally, all peptide hormone and receptor datasets have links to HumanPSD™ (18) and to the Swiss-Prot database. The structures of non-peptide hormones can be accessed through the corresponding hyperlinks to the KEGG COMPOUND section. Finally, all molecules may have links to the TRANSPATH® database.

As described, EndoNet utilizes data about the tissues from which hormones are secreted and in which receptors are expressed to define donor and acceptor cells, respectively. The identity and anatomical environment of cell types, tissues and organs is defined through references to the CYTOMER®

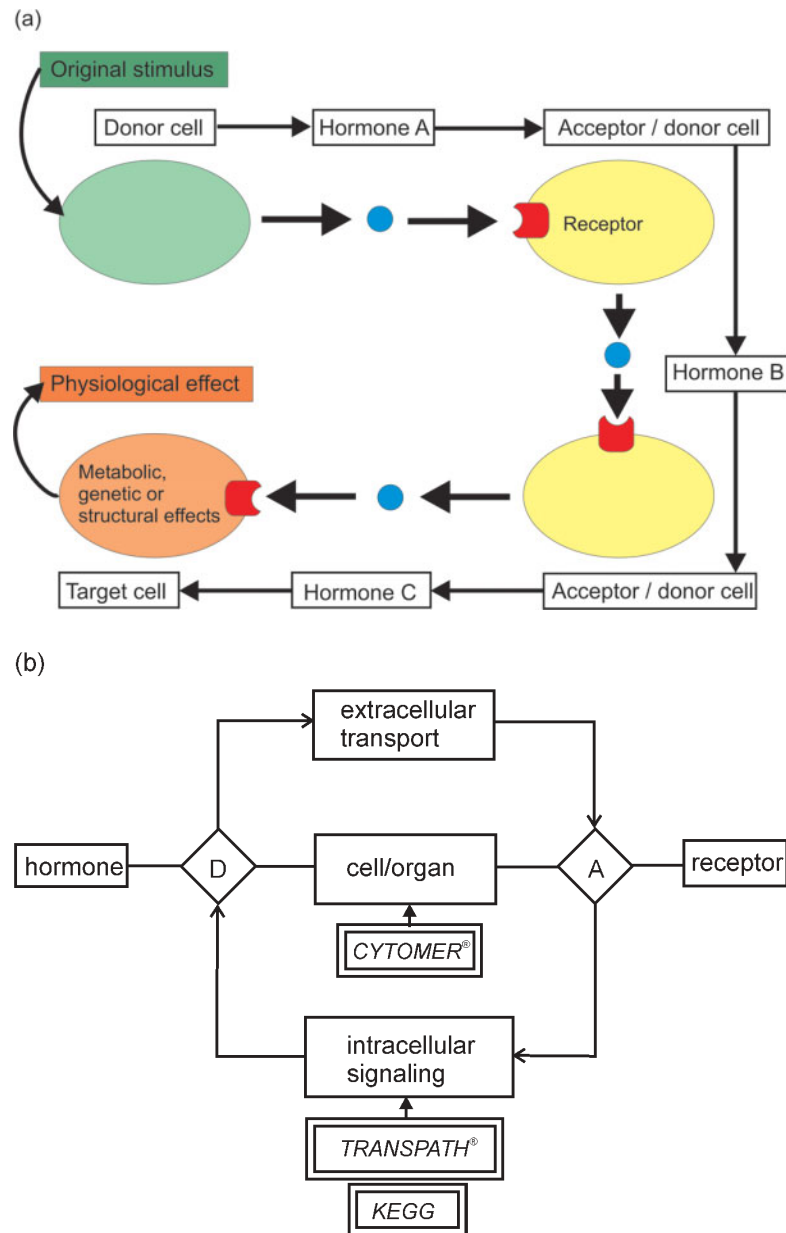


Figure 1. Structure of the hormonal network modeled in EndoNet. (a) Schematic structure of the relational EndoNet database. (b) The rhombs designated D and A are the linking tables between a tissue (cell, organ) and a hormone or a receptor, respectively, and represent the corresponding hormone donor or acceptor cell.

ontology (19,20); in the numerous cases where the receptors are ubiquitously expressed, just the root term 'human body' is linked.

Data on whether a hormone's synthesis is triggered or inhibited by another hormone through its respective receptor in a particular cell was obtained by manual selection from textbooks [e.g. (16), <http://endotext.com/>], monographies (21), original literature, the EST library information and the linked databases (TRANSPATH[®], HumanPSD[™] and Swiss-Prot). The contents of EndoNet are summarized in Table 1.

Web interface, queries and visualization

EndoNet can be accessed through the WWW via a JSP-based web interface. Hormones, receptors and tissues can be queried for their names, and detailed information on all identified

components is available through individual characteristics pages. Each hormone's individual entry page displays its source and target tissues (donor and acceptor cells) as well as its receptors, along with some molecular data. Similarly, each receptor entry exhibits the tissues in which the receptor is expressed, and the hormones it interacts with. Finally, each tissue entry lists the hormone receptors that are found in, as well as the hormones that are synthesized and secreted by the tissue. It is also indicated whether the corresponding tissue exerts gender-specific properties; additional information based on the CYTOMER[®] ontology is available through the corresponding link on the tissue detail page.

Instead of searching for a name of a hormone, one can also browse the hierarchical hormone classification featured by EndoNet (available at the 'Search' page).

Each query result can be used as starting point for an extended query. Different items of interest can be selected and added to a common result set. Since several search results can be combined, it is possible to create sets with multiple search parameters.

Table 1. Contents of EndoNet

Components	Number of entries (17/10/2005)
Molecules	
Hormones	109
Receptors	117
Cellular sources	
Cells/tissues	112
Relations	
Hormone–receptor	149
Donor cell–hormone	184
Receptor–acceptor cell	292
Information sources	
References	264

At any step of this incremental retrieval process, the sets of hormones, receptors and tissues obtained so far can be used as starting points for reconstructing a network by a depth-first graph traversal algorithm. The maximum number of steps can be selected separately for the upstream and downstream part of the reconstruction process. Subsequently, the graph will be displayed using a Graphviz-based visualization method [(22), <http://www.graphviz.org>]. In the resulting image, hormones and receptors are represented as nodes grouped together into subgraphs representing the tissues (cells/organs) they are secreted from or expressed in, respectively (Figure 2). Autocrine loops (donor and acceptor cell being identical) are treated specially for visualization. On demand, the hormones' transport media can be included in the visualization of intercellular edges, enabling the user to choose between different complexities of output. Intracellular edges, which connect a receptor to a hormone, represent the influence of a receptor's activation on the secretion of a hormone from the same cell and are displayed differently, depending on whether this influence is triggering or inhibitory in nature.

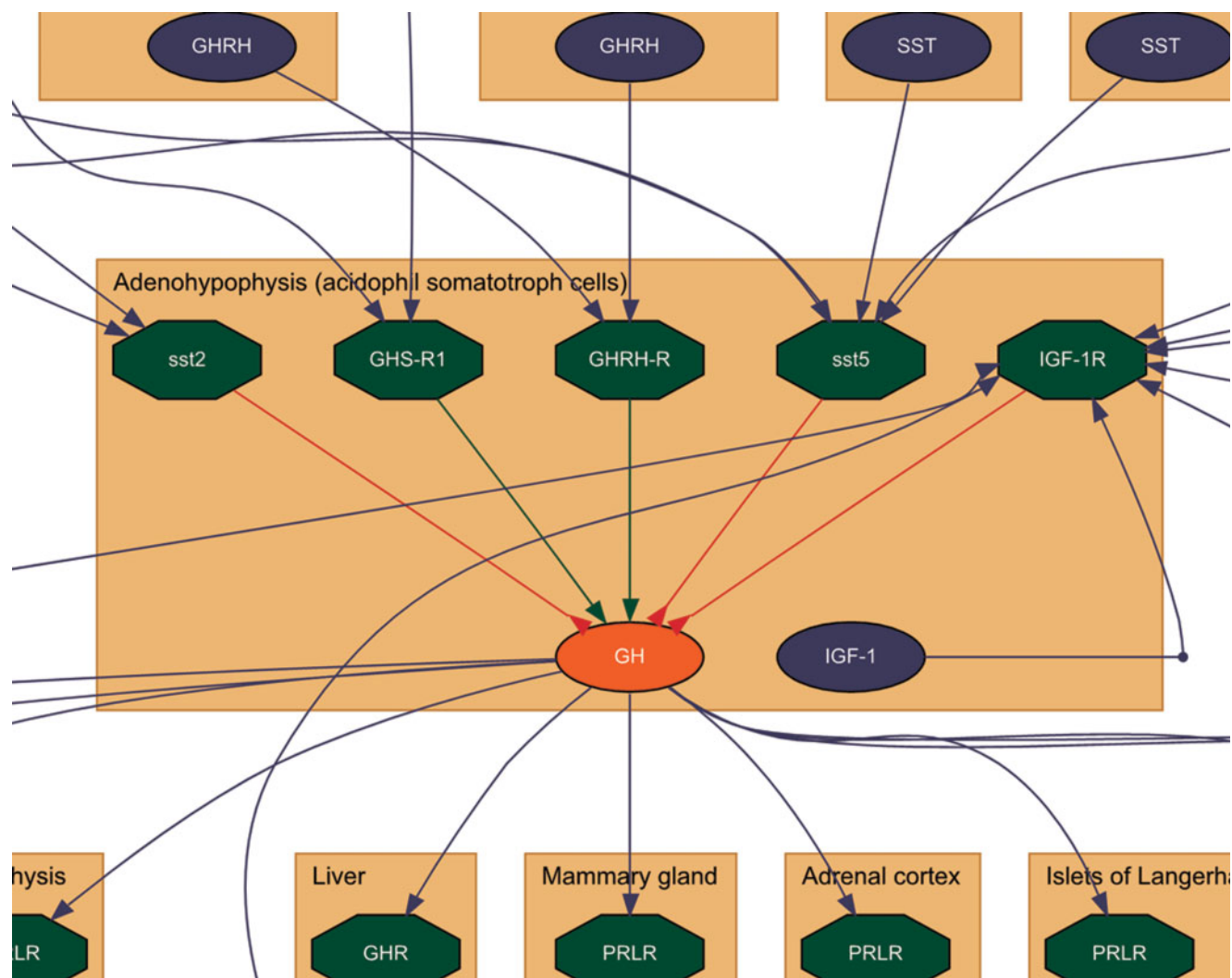


Figure 2. Screenshot with an example of network visualization, only partially drawn for clarity. The blue ovals represent hormones and the green hexagons hormone receptors. The same symbols in red indicate the queried entities. The blue, green and red edges display hormone binding to a receptor, stimulation of synthesis/secretion of another hormone or inhibition of the latter effect, respectively. The brown boxes represent a defined tissue (organ, cell).

The graph is displayed as a clickable image map, linking each entity to its detailed characteristics page, thus making the database entries accessible from the graphical overview of a network, too. The graph is available in two different formats: PNG and SVG. While virtually every browser supports PNG (a 'pixel' or 'bitmap' format), only a few of them provide a zoom functionality for bitmap pictures. Scalable vector graphics, (SVG) provides more functionality (including perfect image quality throughout all zoom factors), but until now most browsers do not support SVG natively and require an SVG plugin (www.adobe.com/svg/viewer/install/main.html) for displaying this vector-based format.

Some well-established endocrine pathways are offered as showcases in the form of predefined result sets. For instance, sets representing the hypothalamic–hypophyseal axis with a focus on either thyroid hormones, adrenal hormones, growth hormones or prolactin are provided. These predefined sets can be used for obtaining a quick overview or as starting points for more complex queries.

FUTURE DEVELOPMENTS

Among the important improvements of EndoNet which we are currently working on is the possibility to represent intercellular communication at different levels of the hierarchical organization of organs, tissues and cells in the organism, as well as to distinguish between such communications in male and female organisms. These options will be introduced by a tighter integration with the CYTOMER-based ontology (19,20).

The next step will be to expand the contents of EndoNet towards the details of the processes occurring in the transport medium, usually the blood. Since not all hormones are transported as free molecules and some hydrophobic hormones (e.g. steroids and thyroids) need to be bound to specific carrier proteins, proper description of such transporters and their interaction with the corresponding hormones will be required. It is planned to involve quantitative data about the regular or pathological levels of the hormones, the overall kinetics of each hormone in the blood (monotonous decay, oscillating concentrations, increase in response to certain stimuli, etc.), its turnover and metabolic products, etc. That will allow utilization of EndoNet contents for diagnostic purposes. In future, the EndoNet data model will be extended in order to incorporate external stimuli (e.g. light) and physiological states (stress, age, etc.) in a formalized manner, allowing to determine whether or not and in which quantity a hormone will be secreted under the given circumstances. EndoNet will then link the physiological effects of hormones with the intracellular molecular processes leading to its synthesis and secretion in the donor cells, and to the effects on its acceptor cells. At many places of these intracellular and intercellular networks, genetically determined aberrations may cause specific, sometimes pathological phenotypes. Thus, EndoNet will enable to bridge the gap between known genotypes and their molecular and clinical phenotypes in this area of medical research and its applications.

DISCUSSION

At its present state, EndoNet provides a high coverage of molecules that are conventionally considered as hormones

as well as other molecules that are involved in intercellular communication, such as growth factors, lymphokines and chemokines and their known receptors. The aim of the database is to provide a useful resource for studying the principal features of hormonal networks in a comprehensive way, as it was done more exemplarily in the past for these kinds of networks (23), but was done globally for many other intracellular network types, such as metabolic, protein interaction and transcription networks [reviewed in (24,25)]. EndoNet database certainly is not yet complete but will grow rapidly.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the help of Michael Tillberg (BIOBASE Corp., Beverly, MA) with establishing the links to HumanPSD. This work was funded in part by 'Forschungs- und Berufungspool des Niedersächsischen Ministeriums für Wissenschaft und Kultur' (Kap. 06 08 TG 74) and by the Federal Ministry of Education, Research and Technology (BMBF) as part of the National Genome Research Network (NGFN, grant NGFN2—Biomed—SMP—Bioinformatik, FKZ: 01GR0-480). Funding to pay the Open Access publication charges for this article was provided by the University of Göttingen/Medical School, Department of Bioinformatics.

Conflict of interest statement. None declared.

REFERENCES

1. Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. and Hattori, M. (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res.*, **32**, D277–D280.
2. Lemer, C., Antezana, E., Couche, F., Fays, F., Santolaria, X., Janky, R., Deville, Y., Richelle, J. and Wodak, S.J. (2004) The aMAZE LightBench: a web interface to a relational database of cellular processes. *Nucleic Acids Res.*, **32**, D443–D448.
3. Bader, G.D., Betel, D. and Hogue, C.W. (2003) BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res.*, **31**, 248–250.
4. Salwinski, L., Miller, C.S., Smith, A.J., Pettit, F.K., Bowie, J.U. and Eisenberg, D. (2004) The Database of Interacting Proteins: 2004 update. *Nucleic Acids Res.*, **32**, D449–D451.
5. Takai-Igarashi, T., Nadaoka, Y. and Kaminuma, T. (1998) A database for cell signaling networks. *J. Comput. Biol.*, **5**, 747–754.
6. Demir, E., Babur, O., Dogrusoz, U., Gursoy, A., Nisanci, G., Cetin-Atalay, R. and Ozturk, M. (2002) PATIKA: an integrated visual environment for collaborative construction and analysis of cellular pathways. *Bioinformatics*, **18**, 996–1003.
7. Krull, M., Voss, N., Choi, C., Pistor, S., Potapov, A. and Wingender, E. (2003) TRANSPATH[®]: an integrated database on signal transduction and a tool for array analysis. *Nucleic Acids Res.*, **31**, 97–100.
8. Choi, C., Crass, T., Kel, A., Kel-Margoulis, O., Krull, M., Pistor, S., Potapov, A., Voss, N. and Wingender, E. (2004) Consistent re-modeling of signaling pathways and its implementation in the TRANSPATH database. *Genome Inform. Ser. Workshop Genome Inform.*, **15**, 244–254.
9. Dupont, A., Tokarski, C., Dekeyser, O., Guihot, A.L., Amouyel, P., Rolando, C. and Pinet, F. (2004) Two-dimensional maps and databases of the human macrophage proteome and secretome. *Proteomics*, **4**, 1761–1778.
10. Klee, E.W., Carlson, D.F., Fahrenkrug, S.C., Ekker, S.C. and Ellis, L.B. (2004) Identifying secretomes in people, pufferfish and pigs. *Nucleic Acids Res.*, **32**, 1414–1421.
11. Dupont, A., Corseaux, D., Dekeyser, O., Drobecq, H., Guihot, A.L., Susen, S., Vincentelli, A., Amouyel, P., Jude, B. and Pinet, F. (2005) The proteome and secretome of human arterial smooth muscle cells. *Proteomics*, **5**, 585–596.

12. Chen, Y., Zhang, Y., Yin, Y., Gao, G., Li, S., Jiang, Y., Gu, X. and Luo, J. (2005) SPD—a web-based secreted protein database. *Nucleic Acids Res.*, **33**, D169–D173.
13. Kronenberg, H., Melmed, S., Larsen, P.R. and Polonsky, K. (2003) Principles of endocrinology. In Larsen, P.R., Kronenberg, H.M., Melmed, S. and Polonsky, K.S. (eds), *Williams Textbook of Endocrinology*, Saunders, Elsevier Science, pp. 1–9.
14. Harris, M.A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis, S., Marshall, B., Mungall, C. *et al.* (2004) The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.*, **32**, D258–D261.
15. Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D. and Darnell, J. (2000) *Molecular Cell Biology*. Media Connected, W.H. Freeman and Company, NY.
16. Nussey, S.S. and Whitehead, S.A. (2001) *Endocrinology: An Integrated Approach*. BIOS Scientific Publishers Ltd., Oxford.
17. Schomburg, I., Chang, A., Ebeling, C., Gremse, M., Heldt, C., Huhn, G. and Schomburg, D. (2004) BRENDA, the enzyme database: updates and major new developments. *Nucleic Acids Res.*, **32**, D431–D433.
18. Hodges, P.E., Carrico, P.M., Hogan, J.D., O'Neill, K.E., Owen, J.J., Mangan, M., Davis, B.P., Brooks, J.E. and Garrels, J.I. (2002) Annotating the human proteome: the Human Proteome Survey Database (HumanPSD™) and an in-depth target database for G protein-coupled receptors (GPCR-PD™) from Incyte Genomics. *Nucleic Acids Res.*, **30**, 137–141.
19. Heinemeyer, T., Chen, X., Karas, H., Kel, A.E., Kel, O.V., Liebich, I., Meinhardt, T., Reuter, I., Schacherer, F. and Wingender, E. (1999) Expanding the TRANSFAC database towards an expert system of regulatory molecular mechanisms. *Nucleic Acids Res.*, **27**, 318–322.
20. Michael, H., Chen, X., Fricke, E., Haubrock, M., Ricanek, R. and Wingender, E. (2005) Deriving an ontology for human gene expression sources from the CYTOMER® database on human organs and cell types. *In Silico Biol.*, **5**, 61–66.
21. Martini, L. (2004) *Encyclopedia of Endocrine Diseases*, Vol. 1–4, Elsevier Inc., Academic Press, Amsterdam.
22. Gansner, E.R. and North, S.C. (2000) An open graph visualization and its applications to software engineering. *Softw. Pract. Exper.*, **30**, 1203–1233.
23. Heyland, A., Hodin, J. and Reitzel, A.M. (2005) Hormone signaling in evolution and development: a non-model system approach. *Bioessays*, **27**, 64–75.
24. Barabási, A.-L. and Oltvai, Z.N. (2004) Network biology: understanding the cell's functional organization. *Nature Rev. Genet.*, **5**, 101–113.
25. Dorogovtsev, S.N. and Mendes, J.F.F. (2002) Evolution of networks. *Adv. Phys.*, **51**, 1079–1187.