

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
20 September 2001 (20.09.2001)

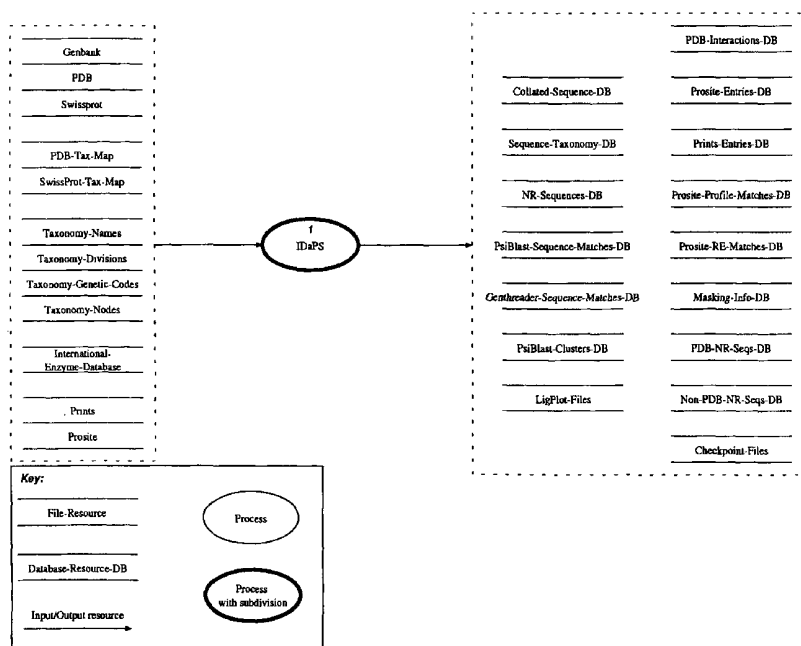
PCT

(10) International Publication Number
WO 01/069507 A3

- (51) International Patent Classification⁷: **G06F 19/00** (74) Agents: **MERCER, Christopher, Paul** et al.; Carpmals & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).
- (21) International Application Number: PCT/GB01/01105
- (22) International Filing Date: 14 March 2001 (14.03.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
0006153.1 14 March 2000 (14.03.2000) GB
- (71) Applicant (for all designated States except US): **INPHARMATICA LIMITED** [GB/GB]; 60 Charlotte Street, London W1T 2NU (GB).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **SWINDELLS, Mark** [GB/GB]; 12 High Street, Easton-On-The-Hill, Stamford, Lincs PE9 3LR (GB). **THORNTON, Janet** [GB/GB]; 63 Hillfield Road, Hemel Hempstead, Herts HP2 4AB (GB). **JONES, David** [GB/GB]; 5 Twineham Park, Woodside Park, London N12 7ER (GB).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— with international search report
- (88) Date of publication of the international search report:
12 September 2002

[Continued on next page]

(54) Title: PROTEOMICS DATABASE



(57) Abstract: The invention concerns methods and systems for predicting the function of proteins. In particular, the invention relates to databases in which details of sequence homologies, biological functions and structures that are shared between proteins of differing sequence have been compiled. The invention also relates to methods, systems and computer software that allows the prediction of protein function and structure and, optionally, the ligand binding properties of the proteins within such a database.



WO 01/069507 A3



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 01/01105

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F19/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ECKMAN B A ET AL: "The Merck Gene Index browser: an extensible data integration system for gene finding, gene characterization and EST data mining" BIOINFORMATICS, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 14, no. 1, February 1998 (1998-02), pages 2-13, XP002132418 ISSN: 1367-4803 abstract	1,47
Y	page 3, section "Data integration" page 5, section "Algorithm" --- -/--	2-46, 48-57



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- * & * document member of the same patent family

Date of the actual completion of the international search

27 May 2002

Date of mailing of the international search report

25/06/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Fillooy García, E

INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 01/01105

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RITTER O ET AL: "PROTOTYPE IMPLEMENTATION OF THE INTEGRATED GENOMIC DATABASE" COMPUTERS AND BIOMEDICAL RESEARCH, ACADEMIC PRESS, LONDON, GB, vol. 27, 1 April 1994 (1994-04-01), pages 97-115, XP002039573 ISSN: 0010-4809	1,47
Y	abstract page 103, paragraph 3 page 113, paragraph 5 ---	2-46, 48-57
X	STOECKERT C J ET AL: "EpoDB: a prototype database for the analysis of genes expressed during vertebrate erythropoiesis" NUCLEIC ACIDS RESEARCH, 'Online! vol. 27, no. 1, 1999, pages 200-203, XP002200128 Retrieved from the Internet: <URL:http://www.cbil.upenn.edu/EpoDB/release/papers/gkc044_gml.pdf> 'retrieved on 2002-05-22! the whole document ---	1,47
X,P	DAVIDSON S B ET AL: "K2/Kleisli and GUS: Experiments in Integrated Access to Genomic Data Sources" IBM SYSTEMS JOURNAL, 'Online! vol. 40, no. 2, 2001, pages 512-531, XP002200129 Retrieved from the Internet: <URL:http://www.research.ibm.com/journal/sj/402/davidson.pdf> 'retrieved on 2002-05-21! abstract page 519, left-hand column, paragraph 5 ---	1,47
A	DAVIDSON S B ET AL: "BIOKLEISLI: A DIGITAL LIBRARY FOR BIOMEDICAL RESEARCHERS" INTERNATIONAL JOURNAL ON DIGITAL LIBRARIES, HEIDELBERG, DE, vol. 1, no. 1, April 1997 (1997-04), pages 36-53, XP000904387 ISSN: 1432-5012 the whole document --- -/--	1-57

INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 01/01105

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BUNEMAN P ET AL: "A Data Transformation System for Biological Data Sources" PROCEEDINGS OF THE TWENTY-FIRST INTERNATIONAL CONFERENCE ON VERY LARGE DATABASES, 'Online! September 1995 (1995-09), pages 158-169, XP002200130 Zurich Retrieved from the Internet: <URL:http://citeseer.nj.nec.com/buneman95data.html> 'retrieved on 2002-05-22! abstract; figure 2 page 162, right-hand column, paragraph 1 - paragraph 2	1-57
A	DAVIDSON S B ET AL: "Challenges in Integrating Biological Data Sources" JOURNAL OF COMPUTATIONAL BIOLOGY, 'Online! vol. 2, no. 4, 12 July 1995 (1995-07-12), pages 557-572, XP002200131 Retrieved from the Internet: <URL:http://citeseer.nj.nec.com/196076.html> 'retrieved on 2002-05-22! abstract section 3	1-57
A	CHUNG S Y ET AL: "Kleisli: a new tool for data integration in biology" TRENDS IN BIOTECHNOLOGY, ELSEVIER, AMSTERDAM, NL, vol. 17, no. 9, 1 September 1999 (1999-09-01), pages 351-355, XP004179984 ISSN: 0167-7799 abstract	1-57