

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
28. Februar 2002 (28.02.2002)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2002/017127 A3

(51) Internationale Patentklassifikation⁷: **G06F 19/00**,
17/10

(21) Internationales Aktenzeichen: PCT/EP2001/009710

(22) Internationales Anmeldedatum:
22. August 2001 (22.08.2001)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
100 41 325.0 23. August 2000 (23.08.2000) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von
US): **GENPROFILE AG** [DE/DE]; Robert-Rössle-Strasse
10, 13125 Berlin (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): **TERHALLE, Werner**
[DE/DE]; c/o GenProfile AG, Robert-Rössle-Strasse 10,
13125 Berlin (DE).

(74) Anwalt: **V. BEZOLD & SOZIEN**; Akademiestrasse 7,
80799 München (DE).

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,

CU, CZ, DK, DM, DZ, EC, 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, PH, PL, PT, RO, RU, SD, SE, SG, SI,
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU,
ZA, ZW.

(84) Bestimmungsstaaten (regional): ARIPO-Patent (GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW),
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,
TM), europäisches 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, GQ, GW,
ML, MR, NE, SN, TD, TG).

Veröffentlicht:

- mit internationalem Recherchenbericht
- vor Ablauf der für Änderungen der Ansprüche geltenden
Frist; Veröffentlichung wird wiederholt, falls Änderungen
eintreffen

(88) Veröffentlichungsdatum des internationalen

Recherchenberichts: 15. Januar 2004

*Zur Erklärung der Zweibuchstaben-Codes und der anderen Ab-
kürzungen wird auf die Erklärungen ("Guidance Notes on Co-
des and Abbreviations") am Anfang jeder regulären Ausgabe der
PCT-Gazette verwiesen.*

(54) Title: METHOD AND DEVICE FOR THE CORRELATION ANALYSIS OF DATA SERIES

(54) Bezeichnung: VERFAHREN UND VORRICHTUNG ZUR KORRELATIONSANALYSE VON DATENFOLGEN

(57) Abstract: A method for processing data series, comprising a number of data points in a predetermined sequence of positions is disclosed, comprising the following steps: determination of correlation values for all doublets, triplets, or n-tuplets of positions in a set of data series on the basis of a pre-determined correlation scale, determination of position weightings from the correlation values for each position of the data series, determination of groups of positions correlated together in the data series, the position weightings of which are not zero and are different from a given threshold value and preparation of derived data series, formed from data in the correlated positions. A correlation device for the conversion method is also disclosed.

(57) Zusammenfassung: Es wird ein Verfahren zur Bearbeitung von Datenfolgen, die jeweils eine Anzahl von Daten in einer vorbestimmten Reihenfolge von Positionen umfassen, beschrieben, das die Schritte aufweist: Ermittlung von Korrelationswerten für alle Paare, Tripel oder n-Tupel von Positionen in einem Satz von Datenfolgen auf der Grundlage eines vorbestimmten Korrelationsmasses, Ermittlung von Positionsgewichtungen aus den Korrelationswerten für jede Position der Datenfolgen, Erfassung von Gruppen zueinander korrelierter Positionen in den Datenfolgen, deren Positionsgewichtungen ungleich Null sind und von einem vorbestimmten Schwellwert abweichen, und Bereitstellung von abgeleiteten Datenfolgen, die durch Daten an den korrelierten Positionen gebildet werden. Es wird auch eine Korrelatorvorrichtung zur Umsetzung des Verfahrens beschrieben.



WO 2002/017127 A3

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 01/09710

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G06F19/00 G06F17/10

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, PAJ, INSPEC, IBM-TDB, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	RIVALS E ET AL: "DETECTION OF SIGNIFICANT PATTERNS BY COMPRESSION ALGORITHMS: THE CASE OF APPROXIMATE TANDEM REPEATS IN DNA SEQUENCES" CABIOS COMPUTER APPLICATIONS IN THE BIOSCIENCES, IRL PRESS, OXFORD, GB, vol. 13, no. 2, 1997, pages 131-136, XP008023983 ISSN: 0266-7061 abstract page 131, left-hand column, paragraph 1 -page 132, right-hand column, paragraph 3 page 133, left-hand column, paragraph 7 page 134, right-hand column, paragraph 6 page 135, left-hand column, paragraph 5 -right-hand column, paragraph 1 --- -/--	1-18



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

7 November 2003

Date of mailing of the international search report

25/11/2003

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

Swarén, P.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 01/09710

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NELSON M R: "DATA COMPRESSION WITH THE BURROWS WHEELER TRANSFORM" DR. DOBB'S JOURNAL, M&T PUBL., REDWOOD CITY, CA, US, vol. 21, no. 9, September 1996 (1996-09), pages 46,48-50, XP000972675 ISSN: 1044-789X	15,17
A	page 46, left-hand column, paragraph 2 page 46, middle column, paragraph 3 -right-hand column, paragraph 4 page 49, left-hand column, paragraph 5 -middle column, paragraph 3 figures 1-6 ---	2,8,9, 14,16
X	US 5 769 074 A (ZHANG ZHEN ET AL) 23 June 1998 (1998-06-23) abstract	15
A	column 1, paragraph 2 -column 2, paragraph 1 column 5, paragraphs 2,3 column 6, paragraphs 2,3 column 7, paragraph 4 -column 9, paragraph 1 column 11, paragraph 4 -column 13, paragraph 4 column 19, paragraph 6 -column 20, paragraph 1 column 21, paragraphs 2,3 column 28, paragraph 4 figures 3,4,10,12,13,17,21,25,26 claims 1,2,16-18 ---	6,7, 10-13
A	D'HAESELEER ET AL: "Gene Expression Data Analysis and Modeling" PROCEEDINGS OF THE PACIFIC SYMPOSIUM ON BIOCOMPUTING, HAWAII, USA, 4 January 1999 (1999-01-04), pages 1-34, XP002203022 page 2, paragraph 4 -page 3, paragraph 1 page 12, paragraph 5 -page 14, paragraph 2 page 25, paragraph 1 -page 29, paragraph 2 figures 6,7,18 --- -/--	1-18

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 01/09710

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LANCTOT J K ET AL: "Estimating DNA sequence entropy" PROCEEDINGS OF THE ELEVENTH ANNUAL ACM-SIAM SYMPOSIUM ON DISCRETE ALGORITHMS, PROCEEDINGS OF SODA 2000. ELEVENTH ANNUAL ACM-SIAM SYMPOSIUM ON DISCRETE ALGORITHMS, SAN FRANCISCO, CA, USA, 9-11 JAN. 2000, pages 409-418, XP002260648 2000, New York, NY, USA & Philadelphia, PA, USA, ACM & SIAM, USA ISBN: 0-89871-453-2 the whole document -----	

INTERNATIONAL SEARCH REPORT

information on patent family members

International Application No

PCT/EP 01/09710

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 5769074	A	23-06-1998	AT 197511 T	11-11-2000
			AU 717449 B2	23-03-2000
			AU 6713196 A	26-02-1997
			CA 2227543 A1	13-02-1997
			CN 1194045 A	23-09-1998
			DE 69610926 D1	14-12-2000
			DE 69610926 T2	21-06-2001
			DK 842475 T3	27-11-2000
			EP 0842475 A1	20-05-1998
			ES 2152548 T3	01-02-2001
			GR 3035370 T3	31-05-2001
			JP 11504739 T	27-04-1999
			NZ 315428 A	28-02-2000
			PT 842475 T	30-04-2001
			WO 9705553 A1	13-02-1997
			US 6248063 B1	19-06-2001
			AU 1837495 A	06-05-1996
			WO 9612187 A1	25-04-1996
			US 6306087 B1	23-10-2001

INTERNATIONALER RECHERCHENBERICHT

Internationaler Aktenzeichen

PCT/EP 01/09710

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 7 G06F19/00 G06F17/10

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
IPK 7 G06F

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal, WPI Data, PAJ, INSPEC, IBM-TDB, BIOSIS

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie ^a	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	RIVALS E ET AL: "DETECTION OF SIGNIFICANT PATTERNS BY COMPRESSION ALGORITHMS: THE CASE OF APPROXIMATE TANDEM REPEATS IN DNA SEQUENCES" CABIOS COMPUTER APPLICATIONS IN THE BIOSCIENCES, IRL PRESS, OXFORD, GB, Bd. 13, Nr. 2, 1997, Seiten 131-136, XP008023983 ISSN: 0266-7061 Zusammenfassung Seite 131, linke Spalte, Absatz 1 -Seite 132, rechte Spalte, Absatz 3 Seite 133, linke Spalte, Absatz 7 Seite 134, rechte Spalte, Absatz 6 Seite 135, linke Spalte, Absatz 5 -rechte Spalte, Absatz 1 --- -/--	1-18

☒ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

☒ Siehe Anhang Patentfamilie

^a Besondere Kategorien von angegebenen Veröffentlichungen :

"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

7. November 2003

Absenddatum des internationalen Recherchenberichts

25/11/2003

Name und Postanschrift der Internationalen Recherchenbehörde
Europäisches Patentamt, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Swarén, P.

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	NELSON M R: "DATA COMPRESSION WITH THE BURROWS WHEELER TRANSFORM" DR. DOBB'S JOURNAL, M&T PUBL., REDWOOD CITY, CA, US, Bd. 21, Nr. 9, September 1996 (1996-09), Seiten 46,48-50, XP000972675 ISSN: 1044-789X	15,17
A	Seite 46, linke Spalte, Absatz 2 Seite 46, mittlere Spalte, Absatz 3 -rechte Spalte, Absatz 4 Seite 49, linke Spalte, Absatz 5 -mittlere Spalte, Absatz 3 Abbildungen 1-6 -----	2,8,9, 14,16
X	US 5 769 074 A (ZHANG ZHEN ET AL) 23. Juni 1998 (1998-06-23)	15
A	Zusammenfassung Spalte 1, Absatz 2 -Spalte 2, Absatz 1 Spalte 5, Absätze 2,3 Spalte 6, Absätze 2,3 Spalte 7, Absatz 4 -Spalte 9, Absatz 1 Spalte 11, Absatz 4 -Spalte 13, Absatz 4 Spalte 19, Absatz 6 -Spalte 20, Absatz 1 Spalte 21, Absätze 2,3 Spalte 28, Absatz 4 Abbildungen 3,4,10,12,13,17,21,25,26 Ansprüche 1,2,16-18 ----	6,7, 10-13
A	D'HAESELEER ET AL: "Gene Expression Data Analysis and Modeling" PROCEEDINGS OF THE PACIFIC SYMPOSIUM ON BIOCOMPUTING, HAWAII, USA, 4. Januar 1999 (1999-01-04), Seiten 1-34, XP002203022 Seite 2, Absatz 4 -Seite 3, Absatz 1 Seite 12, Absatz 5 -Seite 14, Absatz 2 Seite 25, Absatz 1 -Seite 29, Absatz 2 Abbildungen 6,7,18 ----	1-18
A	LANCTOT J K ET AL: "Estimating DNA sequence entropy" PROCEEDINGS OF THE ELEVENTH ANNUAL ACM-SIAM SYMPOSIUM ON DISCRETE ALGORITHMS, PROCEEDINGS OF SODA 2000. ELEVENTH ANNUAL ACM-SIAM SYMPOSIUM ON DISCRETE ALGORITHMS, SAN FRANCISCO, CA, USA, 9-11 JAN. 2000, Seiten 409-418, XP002260648 2000, New York, NY, USA & Philadelphia, PA, USA, ACM & SIAM, USA ISBN: 0-89871-453-2 das ganze Dokument -----	

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationaler Kassenzeichen

PCT/EP 01/09710

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
US 5769074	A	23-06-1998	AT 197511 T 11-11-2000
		AU 717449 B2 23-03-2000	
		AU 6713196 A 26-02-1997	
		CA 2227543 A1 13-02-1997	
		CN 1194045 A 23-09-1998	
		DE 69610926 D1 14-12-2000	
		DE 69610926 T2 21-06-2001	
		DK 842475 T3 27-11-2000	
		EP 0842475 A1 20-05-1998	
		ES 2152548 T3 01-02-2001	
		GR 3035370 T3 31-05-2001	
		JP 11504739 T 27-04-1999	
		NZ 315428 A 28-02-2000	
		PT 842475 T 30-04-2001	
		WO 9705553 A1 13-02-1997	
		US 6248063 B1 19-06-2001	
		AU 1837495 A 06-05-1996	
		WO 9612187 A1 25-04-1996	
		US 6306087 B1 23-10-2001	