



(51) **International Patent Classification:**
G06F 17/30 (2006.01) *G06F 19/00* (2006.01)
G06F 17/27 (2006.01)

(74) **Agent:** WATKIN, Timothy, Lawrence, Harvey; Marks & Clerk Singapore LLP, Tanjong Pagar, P.O. Box 636, Singapore 910816 (SG).

(21) **International Application Number:**
PCT/SG2010/000118

(81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(22) **International Filing Date:**
26 March 2010 (26.03.2010)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
200902159-3 27 March 2009 (27.03.2009) SG

(71) **Applicant** (*for all designated States except US*): **AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH** [SG/SG]; 1 Fusionopolis Way, #20-10 Connexis, Singapore 138 632 (SG).

(72) **Inventors; and**

(75) **Inventors/Applicants** (*for US only*): **KANAGASABAI, Rajaraman** [SG/SG]; c/o PKM, Institute for Infocomm Research, 1 Fusionopolis Way, #21-01 Connexis, South Tower Singapore 138 632 (SG). **BAKER, Christopher James Oliver** [GB/SG]; c/o PKM, Institute for Infocomm Research, 1 Fusionopolis Way, #21-01 Connexis, South Tower Singapore 138 632 (SG).

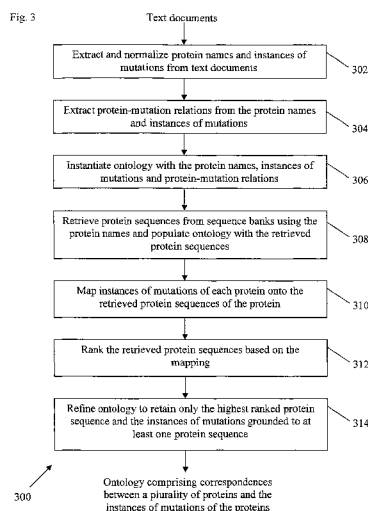
(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

[Continued on next page]

(54) **Title:** A METHOD OF OBTAINING A CORRESPONDENCE BETWEEN A PROTEIN AND A SET OF INSTANCES OF MUTATIONS OF THE PROTEIN



(57) **Abstract:** A method of obtaining a correspondence between a protein and a set of instances $i = 1, \dots, k$ of mutations of the protein is disclosed. The method includes: a) matching a plurality of protein sequences in a sequence bank to at least one expression formed using the set of instances of mutations wherein each protein sequence comprises a plurality of amino acid residues of the protein's constituent peptides and wherein the at least one expression includes wild type residues of a subset of the instances of mutations in the order of their positions within the protein, and differences in said positions of the successive wild type residues in the expression; b) ranking the protein sequences according to the similarity of the protein sequences to the set of instances of mutations, the similarity of the protein sequences to the set of instances of mutations being determined by the matching in step (a); and c) retaining the protein sequence with the highest similarity ranking, said correspondence being a relationship between the retained protein sequence and a subset of the instances of mutations corresponding to the retained protein sequence.



— *before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of*

amendments (Rule 48.2(h))