(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 April 2004 (01.04.2004)

PCT

(10) International Publication Number
WO 2004/027540 A3

(51) International Patent Classification[7]: G01N 33/48

(21) International Application Number:
PCT/US2002/030656

(22) International Filing Date:
23 September 2002 (23.09.2002)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK** [US/US]; 116th Street and Broadway, New York, NY 10027 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **PANG, Phillip, S.** [US/US]; 100 Haven Avenue, Apt. 14B, New York, NY 10032 (US). **JANKOWSKY, Eckhard** [DE/US]; 333 Escuela Avenue, Apt. 333, Mountain View, CA 94040 (US). **PYLE, Anna, Marie** [US/US]; 560 Riverside Drive, Apt. 18G, New York, NY 10027 (US).

(74) Agents: **TANG, Henry** et al.; Baker Botts L.L.P., 30 Rockefeller Plaza, New York, NY 10112-4498 (US).

(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, 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, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
24 March 2005

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.

WO 2004/027540 A3

(54) Title: A COMPUTATIONAL METHOD FOR PREDICTING INTRAMOLECULAR AND INTERMOLECULAR BIOPOLYMER INTERACTIONS

(57) Abstract: Disclosed is a computational method for predicting intramolecular and intermolecular biopolymer interactions which provides an improved way of determining structure and function information, including intramolecular and intermolecular interactions using an*ab initio*-type approach, i.e. using only sequence information. The method is a widely applicable sequence-mining tool capable of detecting both intramolecular and intermolecular interactions for all biopolymers, including, but not limited to, DNA, RNA and protein. It possesses an adaptive screening process that allows for high accuracy. It can be an entirely rule-free, unbiased methodology, and thus can detect novel interactions for all biopolymers. Due to the incorporation of a misalignment process, it can be used iteratively and is capable of refining its own predictions and detecting and managing errors. Therefore, the disclosed method also provides a technique for more accurately determining (and refining) sequence alignments.