



- (51) **International Patent Classification:**
G06F 19/22 (2011.01) *G06N 3/12* (2006.01)
- (21) **International Application Number:**
PCT/EP2013/061300
- (22) **International Filing Date:**
31 May 2013 (31.05.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/654,295 1 June 2012 (01.06.2012) US
- (71) **Applicant:** EUROPEAN MOLECULAR BIOLOGY LABORATORY [DE/DE]; Meyerhofstr. 1, 69117 Heidelberg (DE).
- (72) **Inventors:** GOLDMAN, Nick; European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge Cambridgeshire CB10 1SD (GB). BIRNEY, John; European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge Cambridgeshire CB10 1SD (GB).
- (74) **Agent:** 24IP LAW GROUP SONNENBERG FORTMANN; Patent- und Rechtsanwälte, Dr. Robert Harrison, Herzogspitalstraße 10a, 80331 München (DE).

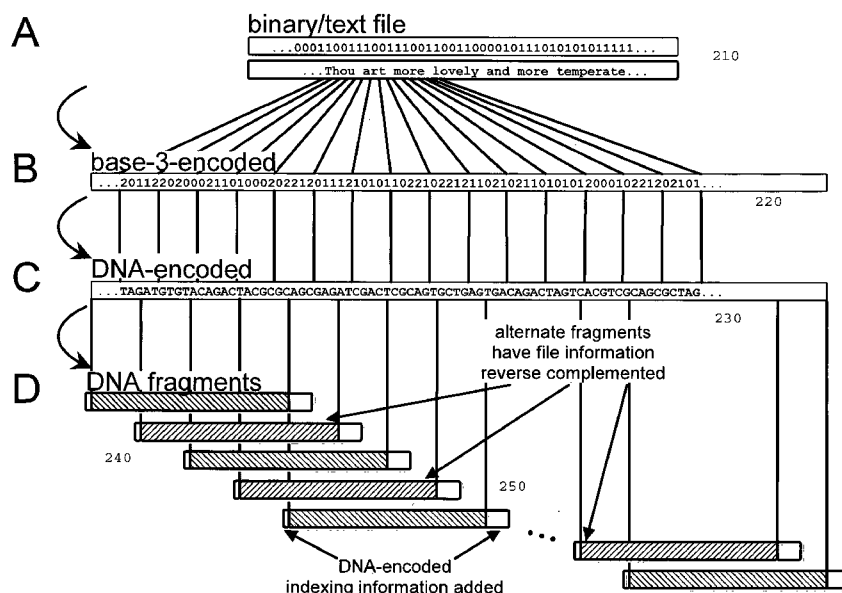
(81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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.

(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

[Continued on next page]

(54) **Title:** HIGH-CAPACITY STORAGE OF DIGITAL INFORMATION IN DNA**Figure 2**

(57) **Abstract:** A method for storage of an item of information (210) is disclosed. The method comprises encoding bytes (720) in the item of information (210), and representing using a schema the encoded bytes by a DNA nucleotide to produce a DNA sequence (230). The DNA sequence (230) is broken into a plurality of overlapping DNA segments (240) and indexing information (250) added to the plurality of DNA segments. Finally, the plurality of DNA segments (240) is synthesized (790) and stored (795).



(88) Date of publication of the international search report:
23 January 2014