

(51) International Patent Classification:
G16B 50/00 (2019.01)(21) International Application Number:
PCT/EP2023/061192(22) International Filing Date:
27 April 2023 (27.04.2023)

(25) Filing Language: English

(26) Publication Language: English

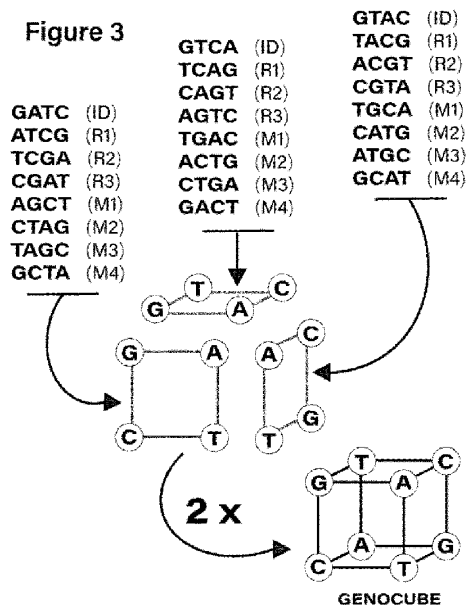
(30) Priority Data:
22170387.9 27 April 2022 (27.04.2022) EP(71) Applicant: INSTITUT PASTEUR DE DAKAR
[SN/SN]; 32, Avenue Pasteur, Dakar, BP 220 (SN).

(72) Inventor; and

(71) Applicant: BERTHET, François Xavier [FR/ES]; Ronda
Mitre, 126, 08021 Barcelona (ES).(72) Inventor: SALL, Amadou Alpha; INSTITUT PASTEUR
DE DAKAR, 32, Avenue Pasteur, Dakar, BP 220 (SN).(74) Agent: HOFFMANN EITL S.L.U.; Paseo de la Castel-
lana 140, 3ª Planta, Edificio Lima, 28046 Madrid (ES).(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, CV, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG,
KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA,
NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO,
RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,
TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, CV,
GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST,
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, ME, 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).

(54) Title: ENCODING GENETIC SEQUENCING INFORMATION AND USES THEREOF

Figure 3

(57) Abstract: A computer-implemented method of representing a
selected sequence of RNA, and/or DNA nucleotides and/or nucleotide
analogues, comprising the steps of (a) defining a cube using nu-
cleotide bases, wherein each of the eight vertices of the cube is as-
signed to a nitrogen-containing base from the subset [A (adenine), C
(cytosine), G (guanine), T (thymine)], wherein nitrogen-containing
bases T (thymine) and U (uracil) are herein understood as equivalent
and thus interchangeable, and wherein each of the vertices of the cube
is assigned to a nucleotide base from the subset such that for each ver-
tex, the assigned base is directly connected to every other base type of
the subset through an edge, (b) assigning the first nucleotide base of
the selected sequence of nucleotides to a vertex of the cube to which
the nitrogen-containing base type of the nucleotide has been assigned
and sequentially assigning each subsequent nucleotide of the selected
sequence of nucleotides to a corresponding vertex in the cube or the
tetrahedron, such that the assigned vertex is either directly connec-
ted to the vertex of the previous nucleotide base through an edge of
the cube; wherein in case the nucleotide base is equal to the previous
nucleotide base, the nucleotide base is assigned to the same vertex
of the cube as the vertex to which the previous nucleotide base has
been assigned to; and wherein the selected sequence of nucleotides
and/or nucleotide analogues comprises nitrogen-containing bases se-
lected from A (adenine), C (cytosine), G (guanine), T (thymine), and/
or U (uracil).

[Continued on next page]

Published:

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