

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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
06 February 2020 (06.02.2020)



(10) International Publication Number
WO 2020/024917 A1

- (51) **International Patent Classification:**
G16B 30/00 (2019.01)
- (21) **International Application Number:**
PCT/CN2019/098258
- (22) **International Filing Date:**
30 July 2019 (30.07.2019)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
PCT/CN2018/097745
30 July 2018 (30.07.2018) CN
- (71) **Applicant:** NANJINGJINSIRUI SCIENCE & TECHNOLOGY BIOLOGY CORP. [CN/CN]; 28 Yongxi Road, Jiangning Science Park, Nanjing, Jiangsu 211100 (CN).
- (72) **Inventor:** FAN, Long; 28 Yongxi Road, Jiangning Science Park, Nanjing, Jiangsu 211100 (CN).
- (74) **Agent:** CHENG & PENG INTELLECTUAL PROPERTY LAW OFFICE; 704, Block B, Xinyu Commercial Building, 90 Guangqumen Inner Street, Dongcheng District, Beijing 100062 (CN).

- (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, SA, 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, 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, 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))
- with sequence listing part of description (Rule 5.2(a))

(54) **Title:** CODON OPTIMIZATION

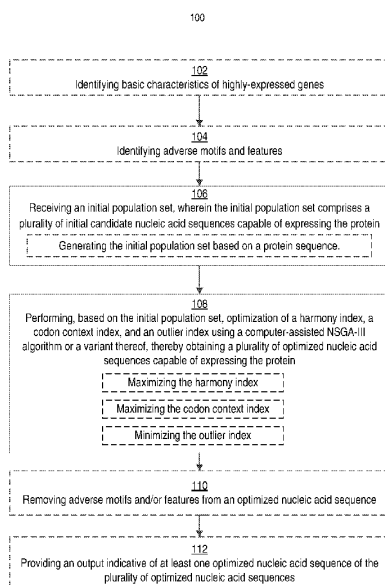


FIG. 1

(57) **Abstract:** An exemplary computer-implemented method for optimizing a nucleic acid sequence for expression of a protein in a host, comprises: a) receiving an initial population set, wherein the initial population set comprises a plurality of initial candidate nucleic acid sequences capable of expressing the protein (106); and b) performing, based on the initial population set, optimization of a harmony index, a codon context index, and an outlier index using a computer-assisted NSGA-III algorithm or a variant thereof, thereby obtaining a plurality of optimized nucleic acid sequences capable of expressing the protein (108).

WO 2020/024917 A1