



(51) International Patent Classification:

G16B 20/00 (2019.01) G16H 50/20 (2018.01)
G06N 20/20 (2019.01) G16B 5/00 (2019.01)

(21) International Application Number:

PCT/US2023/068787

(22) International Filing Date:

21 June 2023 (21.06.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/357,803 01 July 2022 (01.07.2022) US

(71) Applicant: **THE BOARD OF REGENTS OF THE UNIVERSITY OF TEXAS SYSTEM** [US/US]; 210 West 7th Street, Austin, Texas 78701 (US).

(72) Inventors: **XU, Darui**; 5323 Harry Hines Boulevard, Dallas, Texas 75390 (US). **BU, Chun Hui**; 5323 Harry Hines Boulevard, Dallas, Texas 75390 (US). **LYON, Stephen**

Arthur; 5323 Harry Hines Boulevard, Dallas, Texas 75390 (US). **XIE, Yang**; 5323 Harry Hines Boulevard, Dallas, Texas 75390 (US). **WANG, Tao**; 5323 Harry Hines Boulevard, Dallas, Texas 75390 (US). **ZHAN, Xiaowei**; 5323 Harry Hines Boulevard, Dallas, Texas 75390 (US). **BEUTLER, Bruce**; 5323 Harry Hines Boulevard, Dalls, Texas 75039 (US).

(74) Agent: **CLEARY, Zachary D.**; Polsinelli PC, 2950 N Harwood St #2100, Dallas, Texas 75201 (US).

(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,

(54) Title: SYSTEMS AND METHODS TO IDENTIFY MUTATION AND PHENOTYPE ASSOCIATION

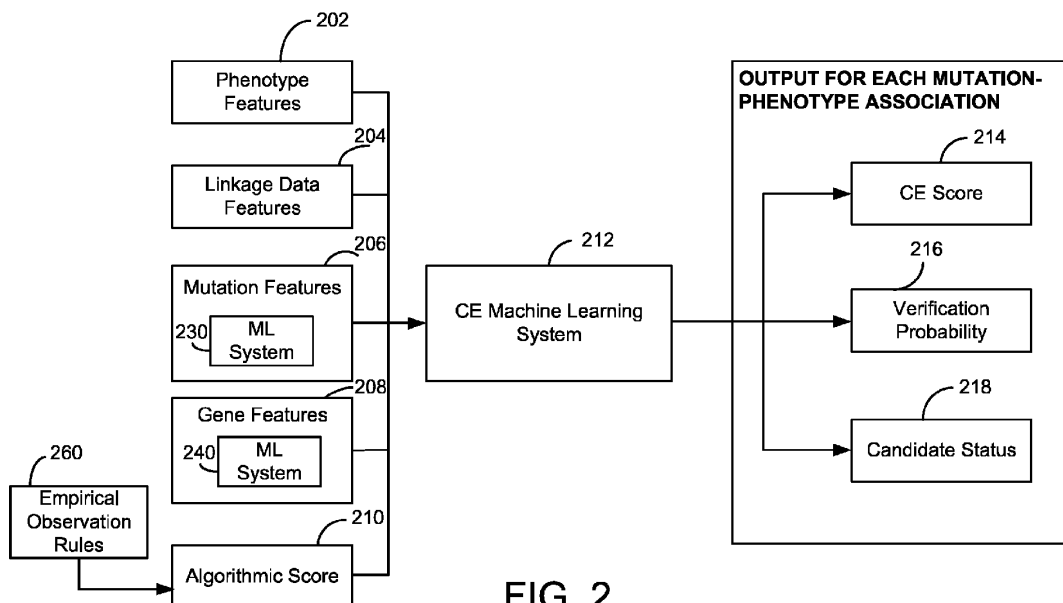


FIG. 2

(57) Abstract: Aspects of the present inventive concept generally relate to systems and methods for mutation processing, and more specifically, for identifying associations between phenotypes and mutations. One example method generally includes receiving one or more input features including phenotype data and mutation data, generating, via a machine learning model, a candidate explorer (CE) score indicating a probability of association between a phenotype and a mutation based on the one or more input features, and outputting an indication of the association between the phenotype and the mutation based on the CE score.

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).

Published:

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