

(19)



SUOMI - FINLAND
(FI)

PATENTTI- JA REKISTERIHALLITUS
PATENT- OCH REGISTERSTYRELSEN
FINNISH PATENT AND REGISTRATION OFFICE

- (10) **FI/EP2951754 T3**
- (12) **EUROOPPAPATENTIN KÄÄNNÖS
ÖVERSÄTTNING AV EUROPEISKT PATENT
TRANSLATION OF EUROPEAN PATENT SPECIFICATION**
- (45) Käännöksen kuulutuspäivä - Kungörelsedag av översättning - **16.05.2024**
Translation available to the public
- (97) Eurooppapatentin myöntämispäivä - Meddelandedatum för **20.03.2024**
det europeiska patentet - Date of grant of European patent
- (51) Kansainvälinen patenttiluokitus - Internationell patentklassificering -
International patent classification
G16B 35/00 (2019 . 01)
- (96) Eurooppapatenttihakemus - Europeisk patentansökan - **EP14746677.5**
European patent application
- (22) Tekemispäivä - Ingivningsdag - Filing date **29.01.2014**
- (97) Patenttihakemuksen julkiseksitilopäivä - Patentansökans **09.12.2015**
publiceringsdag - Patent application available to the public
- (86) Kansainvälinen hakemus - Internationell **29.01.2014 PCT/US2014013666**
ansökan - International application
- (30) Etuoikeus - Prioritet - Priority
31.01.2013 US US201361759276 P 15.03.2013 US US201361799377 P

(73) Haltija - Innehavare - Holder

1 • Codexis, Inc., 200 Penobscot Drive , Redwood City, CA 94063 , (US)

(72) Keksijä - Uppfinnare - Inventor

1 • COPE, Gregory Allan, 200 Penobscot Drive , Redwood City, California 94603 , (US)

(74) Asiamies - Ombud - Agent

Boco IP Oy Ab, Kansakoulukatu 3 , 00100 Helsinki , (FI)

(54) Keksinnön nimitys - Uppfinningens benämning - Title of the invention

**MENETELMÄT, JÄRJESTELMÄT JA OHJELMISTO BIOMOLEKYYLIEN TUNNISTAMISEKSI VUOROVAIKUTTAVIEN KOMPONENTTIEN AVULLA
METHODS, SYSTEMS, AND SOFTWARE FOR IDENTIFYING BIO-MOLECULES WITH INTERACTING COMPONENTS**

MENETELMÄT, JÄRJESTELMÄT JA OHJELMISTO BIOMOLEKYYLIEN TUNNISTAMISEKSI VUOROVAIKUTTAVIEN KOMPONENTTIEN AVULLA

Patenttivaatimukset

5

1. Tietokoneella toteutettu menetelmä sellaisten biologisten molekyylien tunnistamiseksi, joilla on parannettu haluttu aktiivisuus, joka menetelmä käsittää sen, että:

10 (a) vastaanotetaan useita biologisia molekyyliä koskevaa sekvenssi- ja aktiivisuusdataa;

(b) valmistetaan sekvenssi- ja aktiivisuusdatasta perusmalli, jolloin perusmalli ennustaa aktiivisuutta sekvenssin alayksiköiden läsnäolon tai poissaolon funktiona ja perusmalli ei sisällä vuorovaikutustermejä määritellystä vuorovaikutustermien joukosta;

15 (b1) asetetaan perusmalliin nykyinen sekvenssimalli ja paras sekvenssimalli;

(c1) luodaan uusi sekvenssimalli lisäämällä nykyiseen sekvenssimalliin vuorovaikutustermi, jota ei ole vielä lisätty määritellystä vuorovaikutustermien joukosta;

20 (c2) arvioidaan uuden sekvenssimallin ennustetehto ja jos uuden sekvenssimallin ennustetehto on suurempi kuin parhaan sekvenssimallin ennustetehto, asetetaan uusi sekvenssimalli parhaaksi sekvenssimalliksi;

(c3) jos määritellyssä vuorovaikutustermien joukossa on vuorovaikutustermi, jota ei ole lisätty nykyiseen sekvenssimalliin, toistetaan (c1)–(c3);

25 (d) jos uusi malli asetettiin parhaaksi sekvenssimalliksi vaiheessa (c2), asetetaan nykyinen sekvenssimalli parhaaksi sekvenssimalliksi ja toistetaan (c1)–(d);

(e) asetetaan paras sekvenssimalli lopulliseksi malliksi; ja

(f) käytetään lopullista mallia ohjaamaan suunnattua evoluutiota ja valitsemaan yksi tai useampia sekvenssejä, joilla on parannettu haluttu aktiivisuus,

tunnettu siitä, että vuorovaikutustermi valitaan satunnaisesti vaiheessa (c1).

30

2. Tietokoneella toteutettu menetelmä sellaisten biologisten molekyylien tunnistamiseksi, joilla on parannettu haluttu aktiivisuus, joka menetelmä käsittää sen, että:

- (a) vastaanotetaan useita biologisia molekyyliä koskevaa sekvenssi- ja aktiivisuusdataa;
- (b) valmistetaan sekvenssi- ja aktiivisuusdatasta perusmalli, jolloin perusmalli ennustaa aktiivisuutta sekvenssin alayksiköiden läsnäolon tai poissaolon funktiona ja perusmalli sisältää kaikki vuorovaikutustermit määritellystä vuorovaikutustermien joukosta;
- 5 (b1) asetetaan perusmalliin nykyinen sekvenssimalli ja paras sekvenssimalli;
- (c1) luodaan uusi sekvenssimalli vähentämällä nykyisestä sekvenssimallista vuorovaikutustermi, jota ei ole vielä vähennetty määritellystä vuorovaikutustermien
- 10 joukosta;
- (c2) arvioidaan uuden sekvenssimallin ennustetehto ja jos uuden sekvenssimallin ennustetehto on suurempi kuin parhaan sekvenssimallin ennustetehto, asetetaan uusi sekvenssimalli parhaaksi sekvenssimalliksi;
- (c3) jos määritellyssä vuorovaikutustermien joukossa on vuorovaikutustermi, jota ei
- 15 ole vähennetty nykyisestä sekvenssimallista, toistetaan (c1)–(c3);
- (d) jos uusi malli asetettiin parhaaksi sekvenssimalliksi vaiheessa (c2), asetetaan nykyinen sekvenssimalli parhaaksi sekvenssimalliksi ja toistetaan (c1)–(d);
- (e) asetetaan paras sekvenssimalli lopulliseksi malliksi; ja
- (f) käytetään lopullista mallia ohjaamaan suunnattua evoluutiota ja valitsemaan yksi
- 20 tai useampia sekvenssejä, joilla on parannettu haluttu aktiivisuus,
- tunnettu** siitä, että vuorovaikutustermi valitaan satunnaisesti vaiheessa (c1).
3. Patenttivaatimuksen 1 tai patenttivaatimuksen 2 menetelmä, jossa uuden sekvenssimallin luominen vaiheessa (c1) käsittää sen, että käytetään uuden sekvenssimallin
- 25 parametreja koskevaa aiempaa tietoa uuden sekvenssimallin parametrien myöhempien todennäköisyysjakaumien määrittämiseen.
4. Patenttivaatimuksen 3 menetelmä, jossa perusmallin valmistaminen ja/tai uuden sekvenssimallin luominen käsittää sen, että käytetään Gibbs-näytteistämistä mallin
- 30 sovittamiseen sekvenssi- ja aktiivisuusdataan.

5. Patenttivaatimuksen 1 tai patenttivaatimuksen 2 menetelmä, jossa uuden sekvenssimallin ennusteteho vaiheessa (c2) mitataan Akaiken informaatiokriteerin tai Bayesin informaatiokriteerin avulla.
- 5 6. Patenttivaatimuksen 1 tai patenttivaatimuksen 2 menetelmä, jossa sekvenssi on koko genomi, koko kromosomi, kromosomisegmentti, geenisekvenssien kokoelma vuorovaikuttavia geenejä varten, geeni tai proteiini.
7. Jommankumman patenttivaatimuksen 1 tai patenttivaatimuksen 2 menetelmä, jossa
10 alayksiköt ovat kromosomeja, kromosomisegmenttejä, haplotyypppejä, geenejä, nukleotideja, kodoneja, mutaatioita, aminohappoja tai tähteitä.
8. Patenttivaatimuksen 1 tai patenttivaatimuksen 2 menetelmä, jossa useat biologiset molekyylit muodostavat proteiinivarianttikirjaston opetusdatan.
- 15 9. Tietokoneohjelmatuote, joka käsittää yhden tai useampia tietokoneella luettavia pysyväisluonteisia tallennusvälineitä, joille on tallennettu tietokoneella suoritettavia käskyjä, jotka tietokonejärjestelmän yhden tai useampien prosessorien suorittamina saavat tietokonejärjestelmän toteuttamaan jonkin patenttivaatimuksista 1–8 menetelmän.
- 20 10. Tietokonejärjestelmä, joka käsittää:
- yhden tai useampia prosessoreita;
järjestelmämuistin; ja
- 25 yhden tai useampia tietokoneella luettavia tallennusvälineitä, joille on tallennettu tietokoneella suoritettavia käskyjä, jotka yhden tai useampien prosessorien suorittamina saavat tietokonejärjestelmän toteuttamaan jonkin patenttivaatimuksista 1–8 menetelmän.