PATENT
Attorney Docket No.: 117123-5001-US

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: Lee Chae Application No.: 15/282,052 Filed: September 30, 2016 For: DISCOVERY SYSTEMS FOR IDENTIFYING ENTITIES THAT HAVE A TARGET PROPERTY	Confirmation No.: 9183 Examiner: Lut Wong Art Unit: 2129 CERTIFICATE OF ELECTRONIC TRANSMISSION I hereby certify that this correspondence, including listed enclosures is being electronically transmitted in Portable Document Form (PDF) through EFS-Web via Hyper Text Transfer Protocol to the United States Patent and Trademark Office's Patent Electronic Business Center on: Dated: June 13, 2017 Signed: / Peter F. Roidmaier / Peter F. Roidmaier

Mail Stop Office Action
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

RESPONSE TO NON-FINAL OFFICE ACTION

Commissioner:

In response to the Office Action mailed May 3, 2017, please amend the above-identified application as follows:

Amendment to the Claims begin on page 2 of this paper.

Remarks/Arguments begin on page 12 of this paper.

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings of claims in the application:

Listing of Claims:

1. (Currently Amended) A discovery system for inferentially screening a test entity to determine whether it exhibits a target property without directly measuring the test entity for the target property, the discovery system comprising:

at least one processor and memory addressable by the at least one processor, the memory storing at least one program for execution by the at least one processor, the at least one program comprising instructions for:

A) obtaining a training set that comprises a plurality of reference entities and, for each respective reference entity, (i) a respective measurement of each first feature in a respective subset of first features in an N-dimensional feature space and (ii) a respective measurement of each second feature in a respective subset of an M-dimensional feature space, wherein

N is a positive integer of two or greater,

M is a positive integer,

the training set collectively provides at least one measurement for each first feature in the N-dimensional feature space,

the training set collectively provides at least one measurement for each second feature in the M-dimensional feature space,

at least one second feature in the M-dimensional feature space is a metric for the target property,

the N-dimensional feature space does not include any of the second features in the M-dimensional space,

the M-dimensional feature space does not include any of the first features in the N-dimensional space, ~~and~~

the test entity comprises a protein, a fragment thereof, or a mixture of the protein with one or more other proteins**[[;]],**

the obtaining (A) associates the test entity with a data structure comprising one or more extraction parameters used to extract the test entity from the test member, and

the one or more extraction parameters comprises an extraction parameter in the group consisting of (i) an elution pH or time for the test entity, (ii) a buffer type used to extract the test entity from the test member, (iii) a specific pH or pH range used to extract the test entity from the test member, (iv) a specific ionic strength or an ionic strength range used to extract the test entity from the test member, and (v) a specific temperature or temperature range used to extract the test entity from the test member;

B) identifying two or more first features, or one or more combinations thereof, in the N-dimensional feature space using a feature selection method and the training set, thereby selecting a set of first features $\{p_1, \ldots, p_{N-K}\}$ from the N-dimensional feature space, wherein N-K is a positive integer less than N;

C) training a model using measurements for the set of first features $\{p_1, \ldots, p_{N-K}\}$ across the training set, thereby obtaining a trained model;

D) obtaining measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity;

E) inputting the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity into the trained model thereby obtaining a trained model output value for the test entity; and

F) comparing the trained model output value of the test entity to one or more trained model output values computed using measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of one or more reference entities that exhibits the target property thereby determining whether the test entity exhibits the target property.

2. (Withdrawn) The discovery system of claim 1, wherein the trained model is a linear regression model of the form:

$$f(X) = \beta_0 + \sum_{j=1}^{t} X_j \beta_j$$

wherein *t* is a positive integer,

f(*X*) are the measurements for a second feature in the M-dimensional feature space across the training set,

$\beta_0, \beta_1, \ldots, \beta_t$ are parameters that are determined by the training C), and

each X_j in $\{X_1, \ldots, X_t\}$ is a first feature p_i in the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the training set, a transformation of the first feature p_i, a basis expansion of the first feature p_i, an interaction between two or more first features in the set of first features $\{p_1, \ldots, p_{N-K}\}$, or a principal component derived from one or more first features in the set of first features $\{p_1, \ldots, p_{N-K}\}$.

3. (Withdrawn) The discovery system of claim 2, wherein at least one X_j in $\{X_1, \ldots, X_t\}$ represents an interaction between two or more features in the set of first features $\{p_1, \ldots, p_{N-K}\}$.

4. (Withdrawn) The discovery system of claim 2, wherein $\{X_1, \ldots, X_t\}$ is determined by the identifying B) or training C) from the N-dimensional feature space using a subset selection or shrinkage method.

5. (Withdrawn) The discovery system of claim 1, wherein the trained model is a nonlinear regression model.

6. (Withdrawn) The discovery system of claim 1, wherein

the trained model is a clustering applied to the measurements for the set of first features $\{p_1, \ldots, p_{N-K}\}$ across the training set without use of respective measurements of each second feature in the M-dimensional feature space, and

the inputting E) comprises clustering the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity together with the measurements for the set of first features $\{p_1, \ldots, p_{N-K}\}$ across the training set, and

the comparing F) comprises determining whether the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity co-clusters with the set of first features $\{p_1, \ldots, p_{N-K}\}$ of one or more reference entities in the training set that exhibit the target property.

7. (Withdrawn) The discovery system of claim 6, wherein the clustering comprises unsupervised clustering.

8. (Withdrawn) The discovery system of claim 1, wherein

the model is a *k*-nearest neighbors classifier,

the inputting E) and the comparing F) comprises obtaining the trained model output value as the outcome of the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity against the *k* nearest neighbors of the test entity in the training set using the trained *k*-nearest neighbors classifier, and

the *k* nearest neighbors of the test entity includes one or more reference entities that exhibit the target property.

9. (Withdrawn) The discovery system of claim 1, wherein the model is a support vector machine.

10. (Withdrawn) The discovery system of claim 1, wherein

the respective measurement of each first feature in a respective subset of first features in the N-dimensional feature space for each corresponding reference entity in the training set is taken when the corresponding reference entity is in the form of an emulsion or a liquid, and

the set of first features $\{p_1, \ldots, p_{N-K}\}$ comprises protein concentration, hydrophobicity, fat content, color, or phospholipid concentration of the corresponding reference entity.

11. (Withdrawn) The discovery system of claim 1, wherein

the respective measurement of each first feature in a respective subset of first features in the N-dimensional feature space for each corresponding reference entity in the training set is taken when the corresponding reference entity is in the form of an emulsion or a liquid, and

the set of first features $\{p_1, \ldots, p_{N-K}\}$ comprises an amount of inter- or intra-molecular bonds within the corresponding reference entity.

12. (Withdrawn) The discovery system of claim 1, wherein the training C) further comprises training the model using measurements of each corresponding reference entity in the training set for a single second feature, wherein

the single second feature is selected from the group consisting of dye penetration, viscosity, gelation, texture, angled layering, layer strength, flow consistency, and gelling speed, or

the single second feature is hardness, fracturability, cohesiveness, springiness, chewiness, or adhesiveness as determined by a texture profile analysis assay.

13. (Withdrawn) The discovery system of claim 1, wherein

N is 10 or more, and

N-K is 5 or less.

14. (Withdrawn) The discovery system claim 1, wherein the respective measurement of each first feature in the N-dimensional feature space for a single reference entity in the plurality of reference entities is obtained from a molecular assay set comprising three or more different molecular assays.

15. (Withdrawn) The discovery system of claim 1, wherein the respective measurement of each second feature in a respective subset of the M-dimensional feature space for a single reference entity in the plurality of reference entities is obtained from a functional assay set comprising three or more different functional assays of the single reference entity.

16. (Withdrawn) The discovery system of claim 1, wherein the feature selection method comprises regularization across the training set using the N-dimensional feature space and a single second feature in the M-dimensional feature space.

17. (Withdrawn) The discovery system of claim 1, wherein the feature selection method comprises application of a decision tree to the training set using the N-dimensional feature space and all or a portion of the M-dimensional feature space.

18. (Withdrawn) The discovery system of claim 1, wherein the feature selection method comprises application of a Gaussian process regression to the training set using the N-dimensional feature space and a single second feature in the M-dimensional feature space.

19. (Withdrawn) The discovery system of claim 1, wherein

the feature selection method comprises application of principal component analysis to the training set thereby identifying a plurality of principal components wherein the plurality of principal components collectively represent the set of first features $\{p_1, \ldots, p_{N-K}\}$ from the M-dimensional feature space across the training set, and

the training of the model using measurements for the set of first features $\{p_1, \ldots, p_{N-K}\}$ across the training set C) comprises training the model using the plurality of principal components samples for each reference entity in the plurality of reference entities and measurements for one or more second features in each reference sample in the training set.

20. (Withdrawn) The discovery system of claim 1, wherein

a plurality of first features in the N-dimensional feature space is obtained from a molecular assay of each reference entity in the training set,

the feature selection method comprises:

(i) application of a kernel function to the respective measurement of each measured first feature in the plurality of first features in the N-dimensional feature space for each reference entity in the plurality of reference entities thereby deriving a kernel matrix, and

(ii) applying principal component analysis to the kernel matrix thereby identifying a plurality of principal components wherein the plurality of principal components collectively represent the set of first features $\{p_1, \ldots, p_{N-K}\}$ from the N-dimensional feature space; and

the training of the model using measurements for the set of first features $\{p_1, \ldots, p_{N-K}\}$ across the training set comprises training the model using the plurality of principal components samples for each reference entity in the plurality of reference entities.

21. (Withdrawn) The discovery system of claim 1, wherein

a first plurality of first features in the N-dimensional feature space is obtained from a first molecular assay of each reference entity in the training set,

a second plurality of first features in the N-dimensional feature space is obtained from a second molecular assay of each reference entity in the training set,

the feature selection method comprises:

(i) applying a first kernel function to the respective measurement of each measured first feature in the first plurality of first features in the N-dimensional feature space for each reference entity in the plurality of reference entities, thereby deriving a first kernel matrix,

(ii) applying a second kernel function to the respective measurement of each measured first feature in the second plurality of first features in the N-dimensional feature space for each reference entity in the plurality of reference entities, thereby deriving a second kernel matrix, and

(iii) applying principal component analysis to the first kernel matrix and the second kernel matrix thereby identifying a plurality of principal components wherein the plurality of principal components collectively represent the set of first features $\{p_1, \ldots, p_{N-K}\}$ from the N-dimensional feature space; and

the training the model using measurements for the set of first features $\{p_1, \ldots, p_{N-K}\}$ across the training set comprises training the model using the plurality of principal components samples for each reference entity in the plurality of reference entities.

22. (Withdrawn) The discovery system of claim 21, wherein the model is a support vector machine.

23. (Original) The discovery system of claim 1, wherein the test entity originates from a test member of the Fungi, Protista, Archaea, Bacteria, or Plant Kingdom.

24. (Currently Amended) The discovery system of claim 1, wherein

the test entity is extracted from a plant ~~and the at least one program further comprises instructions for associating one or more data structures with the test entity~~, and

the one or more data structures identify the test entity, **[[**an**]]** the extraction parameter for the test entity, and a characteristic of the plant.

25. (Cancelled)

26. (Currently Amended) The discovery system of claim **[[**24**]]** 1, wherein the one or more data structures comprises at least three extraction parameters used to extract the test entity from the ~~plant~~ test member selected from the group consisting of : (i) an elution pH or time for the test entity, (ii) a buffer type used to extract the test entity from the test member ~~plant~~, (iii) a specific pH or pH range used to extract the test entity from the test member ~~plant~~, (iv) a specific ionic strength or an ionic strength range used to extract the test entity from the test member ~~plant~~, or (v) a specific temperature or temperature range used to extract the test entity from the test member ~~plant~~.

27. (Original) The discovery system of claim 24, wherein the characteristic of the plant is a plant taxonomy feature.

28. (Currently Amended) **[[**The**]]** A discovery system ~~of claim 1~~ for inferentially screening a test entity to determine whether it exhibits a target property without directly measuring the test entity for the target property, the discovery system comprising:

at least one processor and memory addressable by the at least one processor, the memory storing at least one program for execution by the at least one processor, the at least one program comprising instructions for:

A) obtaining a training set that comprises a plurality of reference entities and, for each respective reference entity, (i) a respective measurement of each first feature in a respective subset of first features in an N-dimensional feature space and (ii) a respective measurement of each second feature in a respective subset of an M-dimensional feature space, wherein

N is a positive integer of two or greater,

M is a positive integer,

the training set collectively provides at least one measurement for each first feature in the N-dimensional feature space,

the training set collectively provides at least one measurement for each second feature in the M-dimensional feature space,

at least one second feature in the M-dimensional feature space is a metric for the target property,

the N-dimensional feature space does not include any of the second features in the M-dimensional space,

the M-dimensional feature space does not include any of the first features in the N-dimensional space, and

the test entity comprises a mixture of two or more proteins from a single plant species;

B) identifying two or more first features, or one or more combinations thereof, in the N-dimensional feature space using a feature selection method and the training set, thereby selecting a set of first features $\{p_1, \ldots, p_{N-K}\}$ from the N-dimensional feature space, wherein N-K is a positive integer less than N;

C) training a model using measurements for the set of first features $\{p_1, \ldots, p_{N-K}\}$ across the training set, thereby obtaining a trained model;

D) obtaining measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity;

E) inputting the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity into the trained model thereby obtaining a trained model output value for the test entity; and

F) comparing the trained model output value of the test entity to one or more trained model output values computed using measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of one or more reference entities that exhibits the target property thereby determining whether the test entity exhibits the target property.

29. (Original) The discovery system of claim 1, the at least one program further comprising instructions for repeating the obtaining D), inputting E), and comparing F) for each test entity in a plurality of test entities, wherein

each respective test entity in the plurality of test entities comprises a different protein, a different fragment thereof, or a mixture of the different protein with one or more other proteins.

30. (Original) The discovery system of claim 29, wherein the plurality of test entities comprises more than 50 different test entities each from a single plant species.

REMARKS/ARGUMENTS

In the May 3, 2017 Office Action, the Examiner

- withdrew claims 2-22 from consideration;
- rejected claims 1, 23-24 and 29 under 35 U.S.C. § 102(a)(2) as being anticipated by United States Patent Publication No. US2009/0269772A1 to Califano *et al.* (hereinafter, "Califano");
- rejected claims 25-26 under 35 U.S.C. § 103 as being unpatentable over Califano in view of United States Patent Publication No. US2007/0275397 to Wehrman *et al.* (hereinafter, "Wehrman");
- rejected claim 27 under 35 U.S.C. § 103 as being unpatentable over Califano in view of Thomas G. McCloud, 2010, in "High Throughput Extraction of Plant, Marine and Fungal Specimens for Preservation of Biologically Active Molecules" (hereinafter, "Thomas"); and
- rejected claims 28 and 30 under 35 U.S.C. § 103 as being unpatentable over Califano in view of G. Colebatch, 2001, in "Functional genomics: tools of the trade," hereinafter ("Colebatch").

Status of the Claims

Claims 1-30 are pending, with claims 2-22 withdrawn from consideration. With this amendment, claim 25 is cancelled without prejudice. Accordingly, upon entry of the present claim amendments claims 1-24 and 26-30 will be pending.

With this Amendment, claim 1 is amended to specify associating the test entity is associated with a data structure comprising one or more extraction parameters, where the one or more extraction parameters comprises an extraction parameter in the group consisting of (i) an elution pH or time for the test entity, (ii) a buffer type used to extract the test entity from the test member, (iii) a specific pH or pH range used to extract the test entity from the test member, (iv) a specific ionic strength or an ionic strength range used to extract the test entity from the test member, and (v) a specific temperature or temperature range used to extract the test entity from

the test member. Example support for this claim amendment is found in claim 25 as originally filed and paragraph [0039] of the specification.

With this Amendment, the limitations of original claim 1 were incorporated into claim 28 and claim 28 was therefore made independent.

With this Amendment, claims 24 and 26 were amended to correct for antecedent basis in view of the amendments to claim 1.

EXAMINER INTERVIEW SUMMARY

An Examiner Interview was held on Wednesday May 31, 2017, with Examiner Lut Wong and attorneys for Applicant, Wayne Szeto and Brett Lovejoy, in attendance. During the Interview, attorneys for Applicant provided a detailed explanation on the shortcomings of Califano as set forth in the present response to Office Action. During the Interview, Examiner Wong stated that incorporation of claim 25 into claim 1 would likely overcome Califano, subject to further review. In this response, Applicant has incorporated the extraction parameters recited in claim 25 as originally filed into claim 1 and have cancelled claim 25. During the interview, the Examiner stated that he would reconsider rejoining withdrawn claims 2-22 upon determination that claim 1 was allowable.

THE 35 U.S.C. § 102 REJECTION OF CLAIMS 1, 23-24 AND 29 SHOULD BE WITHDRAWN

In the Office Action, the Examiner rejected claims 1, 23-24 and 29 under 35 U.S.C. § 102(a)(2) as being anticipated by Califano. Applicant respectfully traverses the rejection for the reasons discussed below.

A. CALIFANO DOES NOT DISCLOSE A DISCOVERY SYSTEM FOR INFERENTIALLY SCREENING A TEST ENTITY TO DETERMINE WHETHER IT EXHIBITS A TARGET PROPERTY WITHOUT DIRECTLY MEASURING THE TEST ENTITY FOR THE TARGET PROPERTY

Applicant's claim 1 requires:

> A discovery system for inferentially **screening a test entity to determine whether it exhibits a target property without directly measuring the test entity for the target property**

Califano does not disclose this limitation. The features of all compounds in Califano are directly measured. In the Office Action, the Examiner relies on paragraph [0003] in the background section of Califano to satisfy this claim limitation. See the paragraph bridging pages 2-3 of the Office Action. However, paragraph [0003] of Califano, reproduced below, is merely describing conventional high throughput screening:

> [0003] Despite what appears to be a plethora of new drugs making their way to the clinic, there is a rapidly emerging crisis in traditional drug development for malignant diseases. The crisis is triggered by a paucity of new or lead drugs in the pipeline of most pharmaceutical companies. Large pharmaceutical firms have the means to generate many new potential lead compounds. Applications for increasingly smaller percentage of drugs are submitted to the United States Food and Drug Administration (FDA) for approval over time because many of these drugs have not been developed in a manner that respects the underlying systems biology perspective. It is also becoming increasingly clear that high-throughput screening approaches have exhausted the opportunities to focus strictly on single drug target candidates. As a result, pharmaceutical and biotech companies are being trapped between the demand for new blockbuster drugs that work on every patient and the dramatically smaller niches of diseases that are traceable to a common molecular mechanism.
>
> (Califano paragraph [0003])

Paragraph [0003] of Califano does not disclose screening a test entity to determine whether it exhibits a target property **without directly measuring the test entity for the target property** as required by Applicant's claim 1. On this basis alone, Califano fails as a 35 U.S.C. 102 reference.

B. CALIFANO DOES NOT DISCLOSE SCREENING A TEST ENTITY WHERE THE TEST ENTITY COMPRISES A PROTEIN, A FRAGMENT THEREOF, OR A MIXTURE OF THE PROTEIN WITH ONE OR MORE OTHER PROTEINS

Applicant's claim 1 requires:

> A discovery system for inferentially screening **a test entity** to determine whether it exhibits **a target property without directly measuring the test entity for the target property**
>
> [...]
>
> **the test entity comprises a protein, a fragment thereof, or a mixture of the protein with one or more other proteins**;
>
> [...]
>
> **D) obtaining measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity;**

Califano does not disclose these limitations. In the Office Action, the following is asserted:

CLAIM TERM	LOCATION IN CALIFANO	LOCATION IN OFFICE ACTION
Test entity	Paragraph [0014]	Page 4
First Feature	Fig. 1-48-1 to 1-48-N	Page 3
Target Property	Fig. 2A-210	Page 4
Measurement Values (for the first set of features of the test entity)	Fig. 2A-206 on obtaining MAP 52	Page 4

For the reasons discussed below, Applicant contends that the Office Action fails to disclose how measurement values for the set of first features (Califano Fig. 1-48-1 to 1-48-N) are made of the proteins of Califano.

First, there is no disclosure in Califano that measurement values of first features (1-48-1 to 1-48-N) are for a test entity that is a protein. Paragraph [0039] of Califano makes it clear that features 48 are only derived from cell based assays in which individual compounds from one or more compound libraries are exposed cell lines:

> [0038] cell based activity screen assay data 46 from cell based assays in which individual compounds from one or more of the compound libraries are exposed cell lines thereby resulting in assay result data 48;
>
> (Califano [0038])

Thus, to arrive at assay result data 48, it is necessary to expose a compound to a cell. This point is confirmed by paragraph [0045] of Califano which asserts that all screen assay data 46 (that results in assay result data 48 per Califano paragraph [0038]) is from single compound exposure:

> [0045] As illustrated in Fig. 1, computer 10 comprises compound libraries 44, cell based activity screen data 46 (single compound exposure)
>
> (Califano [0045])

Califano does not disclose obtaining data 48 by exposing a protein, a fragment thereof, or a mixture of the protein with one or more other proteins to a cell. All paragraph [0014] of Califano, relied upon by the Examiner for teaching a test entity is "a protein, a fragment thereof, or a mixture of the protein with one or more other proteins," states regarding proteins is a need to concurrently identify (a) proteins in synergistic pathways whose inhibition would produce the desired end-point phenotype, and (b) compounds able to target these proteins. That is, paragraph [0014] of Califano does not teach or suggest exposing a protein, a fragment thereof, or a mixture of the protein with one or more other proteins to a cell in order to obtain first features 48-1 through 48-N:

[0014] Recent advances in systems biology have shown that synergistic pathways and corresponding targets can be efficiently and systematically mapped in specific cellular contexts. This is achieved though perturbation studies using libraries of small chemical compounds. Similarly, it has been shown that perturbation studies using chemical compound libraries can also help identify the specific pathways and even targets affected by an individual compound (e.g.: assigning an "address" to a compound). **One aspect combines these two approaches to concurrently identify (a) proteins in synergistic pathways whose inhibition would produce the desired end-point phenotype, and (b) compounds able to target these proteins.** A second aspect involves using perturbation based on these compounds to directly identify compounds that can implement the desired end-point phenotype. Given a specific end-point phenotype, the systems and methods disclosed herein may reduce the number of potential synergistic com-pounds from $> 10^{10}$ to a few thousand that can be efficiently screened in experimental assays under a multitude of concen-trations, delays, and other experimental conditions. Furthermore, since the target biology can be further investigated using available databases mapping tissue specific expression, a handful of candidate combinations can be selected such that they maximize availability in the diseased tissue while mini-mizing availability in other healthy tissues. In some embodi-ments, the inventive strategy is complemented by a traditional high-throughput screening assay approach in which individual compounds that show some potential towards the desired endpoint phenotype are identified, and which may be further combined with compounds emerging from the bioin-formatics screening. The novel combination of bioinformatics with a standardized high-throughput screening strategy allows for the search a significantly bigger space of potential drug combinations that are likely to have a higher probability of success. The novel platform described herein for the development of combinatorial therapies against diseases, such as cancer, allows for the rapid develop of multiple promising drug combinations and also allows for the generation of revenue from services provided to pharmaceutical and biotechnology companies.

As such, Califano does not disclose screening a test entity where the test entity comprises a protein, a fragment thereof, or a mixture of the protein with one or more other proteins as required by Applicant's claims.

C. CALIFANO DOES NOT DISCLOSE OBTAINING MEASUREMENT VALUES FOR THE SET OF FIRST FEATURES $\{P_1, \ldots, P_{N-K}\}$ OF THE TEST ENTITY

Applicant's claim 1 requires:

A discovery system for inferentially screening **a test entity**

[...]

the test entity comprises a protein, a fragment thereof, or a mixture of the protein with one or more other proteins;

B) identifying two or more first features, or one or more combinations thereof, in the N-dimensional feature space using a feature selection method and the training set, thereby selecting a set of first features $\{p_1, \ldots, p_{N-K}\}$ from the N-dimensional feature space, wherein N-K is a positive integer less than N;

[...]

D) obtaining measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity;

Califano does not disclose obtaining measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity. Even if, *arguendo,* Califano disclosed a test entity that was a protein, which it does not, Califano does not disclose obtaining measurement values for the **set** of first features $\{p_1, \ldots, p_{N-K}\}$ (where the set necessary comprises two or more first features, or one or more combinations thereof) of the test entity. The Examiner relies upon element 206 of Fig. 2A of Califano pertaining to obtaining MAP 52 for disclosure of this feature (Office Action at Page 4). However, as explained in paragraph [0090] of Califano, which accompanies element 206 of Fig. 2A of Califano, MAP 52 is obtained through the exposure of compounds to cells:

[0090] Step 206. In step 206, gene expression profiling is performed with each compound from a reserve library of compounds, such as drugs that have been approved by the FDA regardless of the performance of such drugs in step 202 and regardless of whether such compounds were in fact tested in step 202. In some embodiments, all or a portion of the compounds in the reserve library of compounds are tested in step 202. In some embodiments, none of the compounds in the reserve library of compounds are tested in step 202. Such compounds are referred to herein as validated compounds because such compounds have been approved by a regulatory agency. This does not mean, nor is there any requirement, that such compounds have demonstrated activity against the condition or disease of interest in this screening method. **For each respective compound in the reserve library of compounds, the respective compound is exposed to one or more cell lines and then cellular constituent abundance values for a plurality of cellular constituents in the one or more cell lines is**

> **measured using microarray profiles.** In some embodiments, the reserve library of compounds initially contains compounds approved by the United States Food and Drug Administration (and/or some other governing authority that has the power to approve the use of drugs in a country) and is then extended to include additional compounds of known activity. Over time, these compounds are profiled to identify the specific pathways and targets they uniquely affect. In some embodiments, each of the compounds in the reserve library is exposed to two or more cell lines, three or more cell lines, five or more cell lines, or ten or more cell lines resulting in two or more MAPs 52, three or more MAPs 52, five or more MAPs 52, or ten or more MAPs 52.

Even if the Califano cellular constituent was somehow equated to Applicant's test entity, and the Califano cellular constituent abundance was equated to Applicant's measurement value for a first feature, Califano still fails to anticipate Applicant's claims because Applicant's claims require obtaining a set of features $\{p_1, \ldots, p_{N-K}\}$, where the set of features comprises "two or more first features, or one or more combinations thereof" whereas only a single measurement is made for any given protein in Califano – its abundance. Moreover, Califano makes it clear that the test entity is a compound, as described at paragraphs [0055] through [0059] of Califano, which is not the same thing as a cellular constituent in Califano.

D. CALIFANO DOES NOT DISCLOSE THE N-DIMENSIONAL FEATURE SPACE DOES NOT INCLUDE ANY OF THE SECOND FEATURES IN THE M-DIMENSIONAL SPACE

Applicant's claim 1 requires:

> the N-dimensional feature space does not include any of the second features in the M-dimensional space,

Califano does not disclose this limitation. In the Office Action, the following is asserted:

CLAIM TERM	LOCATION IN CALIFANO	LOCATION IN OFFICE ACTION
N-dimensional space	Fig. 1 48-1 to 48-N	Page 3
M-dimensional space	Fig. 1 66-1 to 66-N	Page 3

However, the compound combinations in Califano 66 are in fact combinations of the single compounds in Califano 48. See, for example the following paragraphs of Califano:

> **[0016]** *Step 214.* In step 214, the compounds that have been tested are filtered to form a filtered set of compound combinations.
>
> **[0126]** In general, step 214 serves to identify each of the compounds suitable for further analysis. Combinations of compounds (*e.g.* combinations of two compounds, combinations of three compounds, combinations of four compounds) are of interest in some embodiments. Because combinations will be selected, in some embodiments the filtering imposed in this step does not impose the requirement that a respective compound have observed efficacy in step 202.
>
> **[0128]** *Step 216.* Among all the possible compound combinations from the filtered list of step 214, a top number of the most synergistic combinations (*e.g.* 10,000 combinations) are screened again against the phenotype of interest as well as background cell types in combination form using, for example, the experimental assay used in step 202, to assess their synergistic behavior.

Paragraph [0016] of Califano addresses the single compounds in Califano 48. Paragraph [0126] of Califano addresses identifying combinations of the single compounds in Califano 48 to generate combinations of compounds. Paragraph [0128] of Califano addresses the compound combinations in Califano 66. Thus, Califano fails to disclose an N-dimensional feature space that does not include any of the second features in the M-dimensional space.

E. CALIFANO DOES NOT DISCLOSE IDENTIFYING TWO OR MORE FIRST FEATURES, OR ONE OR MORE COMBINATIONS THEREOF, IN THE N-DIMENSIONAL FEATURE SPACE USING A FEATURE SELECTION METHOD AND THE TRAINING SET, THEREBY SELECTING A SET OF FIRST FEATURES {P1, …, PN-K} FROM THE N-DIMENSIONAL FEATURE SPACE

Applicant's claims require:

> B) identifying two or more first features, or one or more combinations thereof, in the N-dimensional feature space using a feature selection method and the training

set, thereby selecting a set of first features $\{p_1, \ldots, p_{N-K}\}$ from the N-dimensional feature space, wherein N-K is a positive integer less than N;

The Office Action relies upon paragraph [0003] of Califano for the candidate selection of Applicant's claim element B) (Office Action at page 4). However, such reliance is misplaced because paragraph [0003] of Califano describes selection of a compound, not selection of a compound features as required by Applicant's claims. The Office Action further relies upon elements 58-M-1 to 58-M-N of Figures 1 of Califano for the candidate selection of Applicant's claim element B) (Office Action at page 4). However, this reliance is again misplaced. Elements 58-1-N through 58-M-N of Califano constitute the cellular constituent profiles of cells upon exposure of respective compounds. There is no disclosure in Califano of selecting some subset of first features from the N-dimensional feature space which page 3 of the Office Action states is found in Fig. 1 48-1 to 48-N of Califano. In other words, there is no disclosure in Califano, even when paragraph [0003] of Califano is taken into consideration, with limiting the features of the N-dimensional feature space (Fig. 1 48-1 to 48-N of Califano) using elements 58-1-N through 58-M-N of Califano. As such, Califano fails to disclose identifying two or more first features, or one or more combinations thereof, in the N-dimensional feature space using a feature selection method and the training set, thereby selecting a set of first features $\{p_1, \ldots, p_{N-K}\}$ from the N-dimensional feature space, wherein N-K is a positive integer less than N, as required by Applicant's claims.

F. CALIFANO DOES NOT DISCLOSE TRAINING A MODEL USING MEASUREMENTS FOR THE SET OF FIRST FEATURES $\{P_1, \ldots, P_{N-K}\}$ ACROSS THE TRAINING SET, THEREBY OBTAINING A TRAINED MODEL

Applicant's claim 1 requires:

C) training a model using measurements for the set of first features $\{p_1, \ldots, p_{N-K}\}$ across the training set, thereby obtaining a trained model;

For this claim limitation, the Examiner relies upon step 204 of Fig. 2A of Califano which pertains to obtaining a molecular abundance map for each active compound of step 202 of Figure

2A of Califano. However, step 204 is silent on training a model. For instance, paragraph [0078] of Califano states:

> [0078] Step 204. Molecular abundance maps (MAPs) 52 of active compounds from step 202 are obtained in step 204. For each respective compound tested, one or more cell lines are treated with the respective compound and then the abundance values of cellular constituents in the one or more cell lines are obtained using high throughput techniques such as gene expression profile microarrays. In some embodiments where a compound is exposed to cells at multiple concentrations, the smallest concentration to achieve a differential end-point phenotype in malignant cells versus normal cells is used in step 204. n some embodiments, where a compound is exposed to cells at multiple concentrations, the concentration used in step 204 is determined on a case by case basis upon review of data from step 202.

As such, the cited portions of Califano fail to disclose training a model using measurements for the set of first features $\{p_1, \ldots, p_{N-K}\}$ across the training set, thereby obtaining a trained model.

G. CALIFANO DOES NOT DISCLOSE INPUTTING THE SET OF FIRST FEATURES $\{P_1, \ldots, P_{N-K}\}$ OF THE TEST ENTITY INTO THE TRAINED MODEL THEREBY OBTAINING A TRAINED MODEL OUTPUT VALUE FOR THE TEST ENTITY

Applicant's claim 1 requires:

> E) inputting the set of first features $\{p_1, \ldots, p_{N-K}\}$ of the test entity into the trained model thereby obtaining a trained model output value for the test entity

For this claim limitation, the Examiner relies upon step 206 of Fig. 2A of Califano which pertains to obtaining a molecular abundance map for each compound in a reserve library of compounds, such as drugs approved by the United States Food and Drug Administration, regardless of the performance of such drugs. Paragraph [0090] of Califano, reproduced in relevant part below states:

> [0090] Step 206. In step 206, gene expression profiling is performed with each compound from a reserve library of compounds, such as drugs that have been approved by the FDA regardless of the performance of such drugs in step 202 and regardless of whether such compounds were in fact tested in step 202. In some

> embodiments, all or a portion of the compounds in the reserve library of compounds are tested in step 202. In some embodiments, none of the compounds in the reserve library of compounds are tested in step 202. Such compounds are referred to herein as validated compounds because such compounds have been approved by a regulatory agency. This does not mean, nor is there any requirement, that such compounds have demonstrated activity against the condition or disease of interest in this screening method. For each respective compound in the reserve library of compounds, the respective compound is exposed to one or more cell lines and then cellular constituent abundance values for a plurality of cellular constituents in the one or more cell lines is measured using microarray profiles.

First, the Califano reserve library of compounds cannot be equated to Applicant's test compound because the Califano reserve library of compound's functions are already known and Califano is not directed to trying to inferentially screening them to determine whether they exhibit target properties without directly measuring them for the target property as required by the preamble of Applicant's claim 1. Second, step 206 of Figure 2A of Califano makes no reference to a trained model, which accordingly to the Office Action, was trained at step 204 of Figure 2 of Califano. As such, the cited portions of Califano fail to disclose inputting the set of first features {p1, …, pN-K} of the test entity into a trained model thereby obtaining a trained model output value for the test entity.

H. CALIFANO DOES NOT DISCLOSE COMPARING THE TRAINED MODEL OUTPUT VALUE OF THE TEST ENTITY TO ONE OR MORE TRAINED MODEL OUTPUT VALUES COMPUTED USING MEASUREMENT VALUES FOR THE SET OF FIRST FEATURES $\{P_1, \ldots, P_{N-K}\}$ OF ONE OR MORE REFERENCE ENTITIES THAT EXHIBITS THE TARGET PROPERTY THEREBY DETERMINING WHETHER THE TEST ENTITY EXHIBITS THE TARGET PROPERTY

Applicant's claim 1 requires:

> comparing the trained model output value of the test entity to one or more trained model output values computed using measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of one or more reference entities that exhibits the target property thereby determining whether the test entity exhibits the target property

For this claim limitation, the Examiner relies upon element 210 of Califano Fig. 2A:

> Integrate protein-DNA interactions (e.g., from ARACNc) and transcription factor modulatory interactions (e.g., from MINDy) and optionally protein-protein interactions (e.g., from curated databases obtained by data mining) into a mixed-interaction network using a Bayesian evidence integration framework.

Applicant respectfully submits that element 210 of Fig. 2A of Califano fails to disclose **comparing** the trained model output value **of the test entity** to one or more trained model output values computed using measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of one or more reference entities that exhibits the target property thereby determining whether the test entity exhibits the target property.

At best, step 210 of Fig. 2A of Califano relates to **building** a model using ARACNe and/or MINDy analysis of MAPs. Step 210 does not disclose **comparing** the trained model output value **of the test entity** to one or more trained model output values computed using measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of one or more reference entities that exhibits the target property thereby determining whether the test entity exhibits the target property. Indeed, the reference compounds of Califano do not even need to possess a target property. See, for example, paragraph [0092] of Califano, reproduced below in relevant part:

> In some embodiments, none of the compounds in the reserve library of compounds are tested in step 202. Such compounds are referred to herein as validated compounds because such compounds have been approved by a regulatory agency. **This does not mean, nor is there any requirement, that such compounds have demonstrated activity against the condition or disease of interest in this screening method.**

As such, the cited portions of Califano fail to disclose comparing the trained model output value of the test entity to one or more trained model output values computed using measurement values for the set of first features $\{p_1, \ldots, p_{N-K}\}$ of one or more reference entities that exhibits the target property thereby determining whether the test entity exhibits the target property.

Claims 24 and 29 depend from claim 1 and thus are patentable over Califano for at the least the same reasons discussed above for claim 1.

Claim 23 is patentable over Califano for additional reasons. Claim 23 is patentable over Califano for the additional reason that specifies:

> the test entity originates from a test member of the Fungi, Protista, Archaea, Bacteria, or Plant Kingdom,

Califano does not disclose a test entity that is from a member of the Fungi, Protista, Archaea, Bacteria, or Plant Kingdom. In rejecting this claim limitation, page 5 of the Office Action relies upon paragraph [0214] of Califano:

> [0214] Furthermore, the present invention is not limited to application to humans but may be used in other mammals, plants, yeast, or any other biological organisms. In such instances, transcription factors for such organisms would be used in preferred embodiments

All that paragraph [0214] of Califano discloses is that one may wish to identify a test composition **for** other mammals, plants, yeast, or any other biological organisms. Paragraph [0214] of Califano does not say that such a test compound would be from a member of the Fungi, Protista, Archaea, Bacteria, or Plant Kingdom. Indeed, the Califano test compound would arise from the compound libraries 44 of Figure 1 of Califano as disclosed in paragraphs [0055] through [0059] of Califano. Califano is silent on the original of such compounds.

Claim 24 is patentable over Califano for additional reasons. Claim 24 is patentable of Califano for the additional reason that Califano does not disclose the test entity is extracted from a plant and the at least one program further comprises instructions for associating one or more data structures with the test entity, and the one or more data structures identify the test entity, an extraction parameter for the test entity, and a characteristic of the plant. For this, page 5 of the Office Action relies upon paragraph [0049] of Califano regarding an isolated constituent and paragraph [0214] of Califano for teachings on a plant. As discussed above, paragraph [0214] of Califano does not disclose or suggest the use of a test entity that is from a plant. All that paragraph [0214] of Califano discloses is that one may wish to identify a test composition **for** a

plant. Paragraph [0214] of Califano does not say that such a test compound would be **from** a plant. Paragraph [0049] of Califano does not remedy the deficiencies in paragraph [0214] of Califano because paragraph [0049] of Califano is directed to describing the properties of a cellular constituent. At a minimum, the Califano cellular constituent cannot be construed as a test entity because a set of features (comprising two or more features) are not measured for cellular constituents in Califano. Moreover, Califano is testing compound libraries (element 40 of Califano Fig. 1) for a target property, not cellular constituents. See, for example, the first sentences of the Califano abstract "Systems, methods, and apparatus for searching for a combination of **compounds** of therapeutic interest are provided." There is no teaching in Califano that such compounds are cellular constituents. Such compounds are described at paragraphs [0055] through [059] of Califano. As such, paragraph [0049] of Califano fails to disclose a test entity that is extracted from a plant.

Accordingly, Applicant respectfully requests that the 35 U.S.C. § 102 rejection of claims 1, 23-24, and 29 be withdrawn.

THE 35 U.S.C. § 103 REJECTION OF CLAIMS 25-26 SHOULD BE WITHDRAWN

In the Office Action, the Examiner rejected claims 25-26 under 35 U.S.C. § 103 as being unpatentable over Califano in view of Wehrman. Applicant respectfully traverses the rejection for the reasons discussed below. Applicant notes that the rejection is moot as it applies to claim 25 in view of the cancellation of this claim. Claim 26 depends from claim 1 and thus is patentable over Califano for the reasons discussed above. Wehrman, which is relied upon in the Office Action for its teachings on extraction parameters, fails to remedy the deficiencies in Califano. Accordingly, Applicant respectfully requests that the 35 U.S.C. § 103 rejection of claims 25-26 be withdrawn.

THE 35 U.S.C. § 103 REJECTION OF CLAIM 27 SHOULD BE WITHDRAWN

In the Office Action, the Examiner rejected claim 27 under 35 U.S.C. § 103 as being unpatentable over Califano in view of Thomas. Applicant respectfully traverses the rejection for the reasons discussed below. Claim 27 depends from claim 1 and thus is patentable over

Califano for the reasons discussed above. Thomas, which is relied upon in the Office Action for its teachings on high throughput extraction of plants and uses of plant taxonomy for high throughput screening, fails to remedy the deficiencies in Califano. Accordingly, Applicant respectfully requests that the 35 U.S.C. § 103 rejection of claim 27 be withdrawn.

THE 35 U.S.C. § 103 REJECTION OF CLAIMS 28 AND 30 SHOULD BE WITHDRAWN

In the Office Action, the Examiner rejected claims 28 and 30 under 35 U.S.C. § 103 as being unpatentable over Califano in view of Colebatch. Applicant respectfully traverses the rejection for the reasons discussed below. Claims 28 and 30 depend from claim 1 and thus are patentable over Califano for the reasons discussed above. Colebatch, which is relied upon in the Office Action for its teachings on high throughput screening of plant proteins, fails to remedy the deficiencies in Califano. Accordingly, Applicant respectfully requests that the 35 U.S.C. § 103 rejection of claims 28 and 30 be withdrawn.

CONCLUSION

All the stated grounds of objection and rejection have been properly traversed, accommodated, or rendered moot. Applicant therefore respectfully requests that the Examiner reconsider all presently outstanding objections and rejections and that they be withdrawn.

Applicant believes that a full and complete reply has been made to the outstanding Office Action, and as such, the present application is in condition for allowance. If the Examiner believes, for any reason, that personal communication will expedite prosecution of this application, the Examiner is invited to telephone the undersigned at the number provided below.

Except for issue fees payable under 37 C.F.R. § 1.18, the Commissioner is hereby authorized by this paper to charge any necessary fees during the entire pendency of this application including fees due under 37 C.F.R. §§ 1.16 and 1.17, which may be required, including any required extension of time fees, or credit any overpayment to Deposit Account 50-0310 (Order No.: 117123-5001-US).

Prompt and favorable consideration of this Amendment and Response is respectfully requested.

Respectfully submitted,
MORGAN, LEWIS & BOCKIUS LLP

Dated: June 13, 2017 By: / Brett Lovejoy /
Brett A. Lovejoy, Reg. No. 42,813

Customer No.: 43850
Morgan, Lewis & Bockius LLP
One Market, Spear Street Tower
San Francisco, CA 94105
Telephone: (415) 442-1000
Facsimile: (415) 442-1001