

CLAIMS

What is claimed is:

1. A method of identifying a marker useful for identifying a condensate of interest associated with a disease, the method comprising:
 - (a) assessing a plurality of candidate markers, or precursors thereof, for a level of association with one or more disease-associated factors of the disease;
 - (b) assessing at least one of the plurality of candidate markers for a condensate affinity factor; and
 - (c) identifying the marker from the plurality of candidate markers based on the marker having a desired level of association with the one or more disease-associated factors of the disease and having a desired condensate affinity factor.
2. The method of claim 1, further comprising identifying the one or more disease-associated factors of the disease.
3. The method of claim 1 or 2, wherein each of the one or more disease-associated factors of the disease is selected from the group consisting of a genetic variant, post translational modification variant, exogenous genetic material, presence of an endogenous compound, presence of an exogenous compound, a physical process, and environmental stimulus.
4. The method of any one of claims 1-3, further comprising identifying a gene or a non-coding variant associated with the one or more disease-associated factors of the disease.
5. The method of any one of claims 1-4, wherein the level of association of each candidate marker with the one or more disease-associated factors of the disease is based on a disease-causal factor score.
6. The method of claim 5, wherein the disease-causal factor score reflects the strength of association of each candidate marker with the one or more disease-associated factors of the disease.

7. The method of claim 5 or 6, further comprising assigning each candidate marker with the disease-causal factor score.
8. The method of any one of claims 1-7, wherein the condensate affinity factor is based on a condensate-association score.
9. The method of claim 8, wherein the condensate-association score reflects the strength of association of the candidate marker, or a portion thereof, with any condensate, a specific condensate, and/or a macromolecule associated with a condensate.
10. The method of claim 8 or 9, further comprising assigning the candidate marker with the condensate-association score.
11. The method of any one of claims 8-10, wherein the condensate-association score is a composite score of a condensate function score and a condensate affinity score.
12. The method of claim 11, wherein the condensate function score is determined based on one or more factors of whether a genetic variation of the candidate marker or a portion thereof or the gene or the non-coding variant associated with the one or more disease-associated factors of the disease:
 - i) is within an intrinsically disordered region (IDR);
 - ii) is subject to a post-translational modification;
 - iii) affects splicing of the candidate marker or the gene associated with the one or more disease-associated factors of the disease;
 - iv) affects a chromatin state close to the gene or the non-coding variant associated with the one or more disease-associated factors of the disease; and
 - v) affects expression of the gene associated with the one or more disease-associated factors of the disease.
13. The method of claim 12, where the one or more factors each has a weight contributing to the condensate function score.

14. The method of claim 11, wherein the condensate affinity score is determined based on, in the candidate marker or a portion thereof or the gene associated with the one or more disease-associated factors of the disease, one or more factors of:
 - i) the presence, absence, amount, and/or degree of an IDR;
 - ii) the presence, absence, amount, and/or degree of a condensate-favoring motif; and
 - iii) the presence, absence, amount, and/or valency of an interacting domain.
15. The method of claim 14, where the one or more factors each has a weight contributing to the condensate affinity score.
16. The method of any one of claims 4-15, further comprising identifying a gene expression product based on the identified gene, wherein the gene expression product, or a portion thereof, is used to populate the plurality of candidate markers, or precursors thereof.
17. The method of any one of claims 1-16, wherein identifying the marker from the plurality of candidate markers comprises a cumulative score based on the desired level of association with the one or more disease-associated factors of the disease and the desired condensate affinity factor.
18. The method of any one of claims 1-17, wherein the marker is a biological marker.
19. The method of any one of claims 1-18, wherein the marker is identified *in silico*.
20. The method of any one of claims 1-19, further comprising verifying the marker as useful for identifying a condensate of interest associated with the disease.
21. A method of identifying a condensate of interest associated with a disease, the method comprising:
 - (a) comparing a first condensate phenotype from a first cell model with a second condensate phenotype from a second cell model,
wherein the first condensate phenotype and the second condensate phenotype are obtained using a marker identified using the method of any one of claims 1-20,

- wherein a difference between the first cell model and the second cell model is attributable to one or more disease-associated factors of the disease, and wherein at least one of the one or more disease-associated factors is introduced to either the first cell model or the second cell model; and
- (b) identifying the condensate of interest as associated with the disease based on a difference between the first condensate phenotype and the second condensate phenotype for the condensate of interest.
22. A method of identifying a compound that modulates a condensate phenotype, the method comprising:
- (a) admixing the compound and a composition comprising a cell model; and
- (b) obtaining a resulting condensate phenotype of the composition using a marker identified using the method of any one of claims 1-20,
- wherein a difference between the resulting condensate phenotype and a reference condensate phenotype identifies the compound as modulating the condensate phenotype.
23. A method of identifying a compound useful for treating a disease, the method comprising:
- (a) admixing the compound and a composition comprising a first cell model;
- (b) obtaining a resulting condensate phenotype of the composition using a marker identified using the method of any one of claims 1-20,
- wherein the compound is identified as useful for treating the disease when the resulting condensate phenotype has a desired modulation of a phenotypic identifier associated with one or more disease-associated factors of the disease.