IMMEDIATE COMMUNICATION

Precision medicine for psychotic disorders: objective assessment, risk prediction, and pharmacogenomics

M. D. Hill[1,2,8], S. S. Gill[1,8], H. Le-Niculescu[1,3,8], O. MacKie[1,2,8], R. Bhagar[1], K. Roseberry[1], O. K. Murray[1], H. D. Dainton[1,5], S. K. Wolf[1,6], A. Shekhar[1,7], S. M. Kurian[4] and A. B. Niculescu[1,2,3]✉

Psychosis occurs inside the brain, but may have external manifestations (peripheral molecular biomarkers, behaviors) that can be objectively and quantitatively measured. Blood biomarkers that track core psychotic manifestations such as hallucinations and delusions could provide a window into the biology of psychosis, as well as help with diagnosis and treatment. We endeavored to identify objective blood gene expression biomarkers for hallucinations and delusions, using a stepwise discovery, prioritization, validation, and testing in independent cohorts design. We were successful in identifying biomarkers that were predictive of high hallucinations and of high delusions states, and of future psychiatric hospitalizations related to them, more so when personalized by gender and diagnosis. Top biomarkers for hallucinations that survived discovery, prioritization, validation and testing include PPP3CB, DLG1, ENPP2, ZEB2, and RTN4. Top biomarkers for delusions include AUTS2, MACROD2, NR4A2, PDE4D, PDP1, and RORA. The top biological pathways uncovered by our work are glutamatergic synapse for hallucinations, as well as Rap1 signaling for delusions. Some of the biomarkers are targets of existing drugs, of potential utility in pharmacogenomics approaches (matching patients to medications, monitoring response to treatment). The top biomarkers gene expression signatures through bioinformatic analyses suggested a prioritization of existing medications such as clozapine and risperidone, as well as of lithium, fluoxetine, valproate, and the nutraceuticals omega-3 fatty acids and magnesium. Finally, we provide an example of how a personalized laboratory report for doctors would look. Overall, our work provides advances for the improved diagnosis and treatment for schizophrenia and other psychotic disorders.

Molecular Psychiatry; https://doi.org/10.1038/s41380-024-02433-8

INTRODUCTION

> "A hallucination is a fact, not an error; what is erroneous is a judgment based upon it."
> - Bertrand Russell

Schizophrenia and related psychotic disorders affect over 3 million Americans, usually have their onset in young adulthood, and are a major cause of diminished quality of life and disability. Due to lack of objective tests, they are often sub-optimally diagnosed and treated, leading to a downward socio-economic spiral and shortened lifespan. Psychotic disorders have as key pathognomonic symptoms hallucinations (perceptual abnormalities) and delusions (conceptual abnormalities). Psychiatric patients may have an increased vulnerability to transient or permanent psychosis, regardless of their primary diagnosis. As such, they are an enriched population in which to try to identify blood biomarkers for hallucinations and for delusions that are generalizable and trans-diagnostic. Such biomarkers would eliminate subjectivity from assessments, provide some indication of risk, and help guide treatments [1]. At the level of population health, if used early as part of routine primary care, they can identify prodromal risk and lead to preventive approaches.

First, we used a powerful longitudinal within-subject design in individuals with psychiatric disorders to discover blood gene expression changes between self-reported no hallucinations and high hallucinations states, and between no delusions and high delusions states. Second, we prioritized this list of candidate biomarkers with a Bayesian-like Convergent Functional Genomics (CFG) approach, comprehensively integrating previous human and animal model evidence in the field of schizophrenia, from our previous work and that of others. Third, we validated our top biomarkers from discovery and prioritization in independent cohorts of psychiatric subjects with high scores on psychosis rating scales. We prioritized a list of 98 candidate biomarkers for hallucinations and 70 for delusions that had the most evidence from the first three steps. Fourth, we tested if these candidate biomarkers are able to predict hallucinations and delusions severity state (i.e., analytical validity), and future clinical worsening (hospitalizations with hallucinations and delusions as the primary cause) (i.e., clinical validity), in additional independent cohorts of psychiatric subjects. We tested the biomarkers in all subjects in the

[1]Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, USA. [2]Indianapolis VA Medical Center, Indianapolis, IN, USA. [3]Stark Neuroscience Research Institute, Indiana University School of Medicine, Indianapolis, IN, USA. [4]Scripps Health, La Jolla, CA, USA. [5]Present address: Department of Neurology, Medical University of South Carolina, Charleston, SC, USA. [6]Present address: Department of Neurology, Ohio State University Medical Center, Columbus, OH, USA. [7]Present address: Office of the Dean, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [8]These authors contributed equally: M. D. Hill, S. S. Gill, H. Le-Niculescu, O. MacKie. ✉email: anicules@iupui.edu

Received: 8 August 2023 Revised: 16 December 2023 Accepted: 15 January 2024
Published online: 08 February 2024

test cohorts, as well as in a more personalized fashion by gender and psychiatric diagnosis, showing increased accuracy with the personalized approach. Fifth, we analyzed the biological pathways and networks they are involved in, as well as which of our top biomarkers have evidence for involvement in other psychiatric and related disorders. Sixth, we identified which of our biomarkers are targets of existing drugs and thus can be used for pharmacogenomic matching of patient to treatment and measuring of response to treatment.

Our work provides a two-step solution to psychotic disorders. The first step is objective awareness that a patient has something. The second step is helping find the right treatment. Current medication treatments for psychosis (e.g., antipsychotics) do not work well in everybody (e.g., low response/remission rates, trial-and-error prescription, problematic side effects, etc.). Matching the right individuals to the right medications using their biomarker profile is a key actionable outcome of our work.

MATERIALS AND METHODS

Cohorts

Our study utilized 3 independent cohorts: 1. discovery cohort (psychiatric disorders subjects with changes from visit to visit in the hallucinations item, and in the delusions item, of the Positive and Negative Syndrome Scale-PANSS); 2. Validation cohort (psychiatric disorders subjects with clinically severe PANSS positive symptom scale scores); and 3. Testing cohort (an independent psychiatric disorders subjects cohort for predicting *state* hallucinations and delusions, and for predicting *trait* hallucinations and delusions -future hospitalization with hallucinations and with delusions as the primary reason) (Fig. 1A, F).

The psychiatric subjects are part of a larger longitudinal cohort of adults that we are continuously collecting [2–4]. Subjects were recruited from the patient population at the Indianapolis VA Medical Center. All subjects understood and signed informed consent forms detailing the research goals, procedures, caveats, and safeguards, per IRB approved protocol. Subjects completed extensive structured neuropsychological testing at each testing visit, 3–6 months apart or whenever a new psychiatric hospitalization occurred. At each testing visit, they received a series of rating scales, including the PANSS scale. The P3 item from the PANSS was used to quantify Hallucinations severity, and the P1 item was used to quantify Delusions severity (Fig. S1). These items are on a scale of 1 to 7, with a score of 1 being no hallucinations and delusions, and a score of 4 and above being high hallucinations and delusions. As such, they generate temporal, quantitative, and targeted data.

At each testing visit we collected whole blood (10 ml) in two RNA-stabilizing PAXgene tubes, labeled with an anonymized ID number, and stored at −80 °C in a locked freezer until the time of future processing. Whole-blood RNA was extracted for gene expression studies from the PAXgene tubes, as detailed below.

For this study, our within-subject discovery cohort for hallucinations consisted of 25 subjects (20 males, 5 females) with a total of 65 visits, and for delusions consisted of 31 subjects (27 males, 4 females) with a total of 95 visits (Table S1). Each of the subjects had at least one diametric change in hallucinations and delusions *state* for consecutive visits, from a no hallucinations or delusions state (P3 or P1 scores of 1) to a high hallucinations or delusions state (P3 or P1 scores of ≥4), or vice versa, from one visit to another (Figs. 1B, G and S1A, B).

Our independent validation cohorts of clinically severe patients, in which the top candidate biomarker findings were validated for being even more changed in expression compared to in the discovery cohort, consisted of 36 subjects (27 male and 9 female) with a total of 52 visits for hallucinations, and 43 subjects (36 male and 7 female) with a total of 62 visits for delusions. These subjects were selected for clinically severe psychosis (PANSS Positive scores ≥21) concordant with high hallucinations or delusions scores (P3 or P1 scores of ≥4).

For testing the biomarkers, we used independent testing cohorts.

For *state* predictions for hallucinations, we predicted high hallucinations scores (P3 ≥ 4) in 196 subjects (162 male and 34 female) with a total of 513 visits. For *state* predictions for delusions, we predicted high delusions scores (P1 ≥ 4) in 120 subjects (109 male and 11 female) with a total of 315 visits (Fig. 1A, F, and Table S1).

For *trait* predictions of future hospitalizations with hallucinations or delusions as a contributory reason (Fig. 1A, F and Table S1), we used a subset of the independent test cohort for which we had longitudinal follow-up with electronic medical records. The subjects' subsequent number of hospitalizations with hallucinations or delusions was tabulated from electronic medical records.

Medications. The subjects in our study were all diagnosed with various psychiatric disorders (Table S1) and had various medical co-morbidities. Their medications were listed in their electronic medical records and documented by us at the time of each testing visit. Medications can have a strong influence on gene expression. However, there was no consistent pattern of any particular type of medication. Our subjects were on a wide variety of different medications, psychiatric and non-psychiatric. Furthermore, the independent validation and testing cohort's gene expression data was Z-scored by gender and by diagnosis before being combined, to normalize for any such effects. Some subjects may be non-compliant with their treatment and may have changes in medications or drugs of abuse not reflected in their medical records. Our goal is to find biomarkers that track hallucinations or delusions, regardless if the reason for it is internal biology or it is driven by external medications or drugs. In fact, one would expect some of these biomarkers to be targets of medications, as we show in this paper. Furthermore, the prioritization step that occurs after the discovery step is based on a field-wide convergence with literature that includes genetic data and animal model data, that are unrelated to medication effects. Overall, the discovery, validation, and replication by testing in independent cohorts of the biomarkers, with our design, occurs despite the subjects having different genders, diagnoses, being on various different medications, and other variables.

Blood gene expression experiments

RNA extraction. Whole blood (2.5 ml) was collected into each PaxGene tube by routine venipuncture. PaxGene tubes contain proprietary reagents for the stabilization of RNA. Total RNA was extracted and processed as previously described [2–4].

Microarrays. Microarray work was carried out using previously described methodology [2–5].

Of note, all genomic data was normalized (RMA for technical variability, then z-scoring for biological variability), by gender and psychiatric diagnosis, before being combined and analyzed.

Biomarkers

Step 1: Discovery. We have used the subject's score from the PANSS Scale Items P3 hallucinations and P1 delusions, assessed at the time of blood collection (Fig. S1). We analyzed gene expression differences between visits with no hallucinations or delusions (defined as a score ≤1) and visits with high hallucinations or delusions (defined as a score ≥4), using a powerful within-subject design, then an across-subjects summation (Fig. 1B, G).

We analyzed the data in two ways: an Absent-Present (AP) approach, and a differential expression (DE) approach, as in previous work by us on suicide biomarkers. The AP approach may capture turning on and off of genes, and the DE approach may capture gradual changes in expression. Analyses were performed as previously described. In brief, we imported all Affymetrix microarray data as CEL. files into Partek Genomic Suites 6.6 software package (Partek Incorporated, St Louis, MI, USA). Using only the perfect match values, we ran a robust multi-array analysis (RMA) by gender and diagnosis, background corrected with quantile normalization and a median polish probeset summarization of all chips, to obtain the normalized expression levels of all probesets for each chip. Then, to establish a list of differentially expressed probesets we conducted a within-subject analysis, using a fold change in expression of at least 1.2 between consecutive high and no hallucinations or delusions visits within each subject. Probesets that have a 1.2-fold change are then assigned either a 1 (increased in high hallucinations or delusions) or a −1 (decreased in high hallucinations or delusions) in each comparison. Fold changes between 1.1 and 1.2 are given 0.5, and fold changes less than 1.1 are given 0. These values were then summed for each probeset across all the comparisons and subjects, yielding a range of raw scores. The probesets above the 33.3% of raw scores were carried forward in analyses (Fig. 1), and received an internal score of 2 points; those above 50% 4 points, and those above 80% 6 points. We have developed in our labs R scripts to automate and conduct all these large dataset analyses in bulk, checked against human manual scoring.

Gene Symbol for the probesets were identified using NetAffyx (Affymetrix) for Affymetrix HG-U133 Plus 2.0 GeneChips, followed by

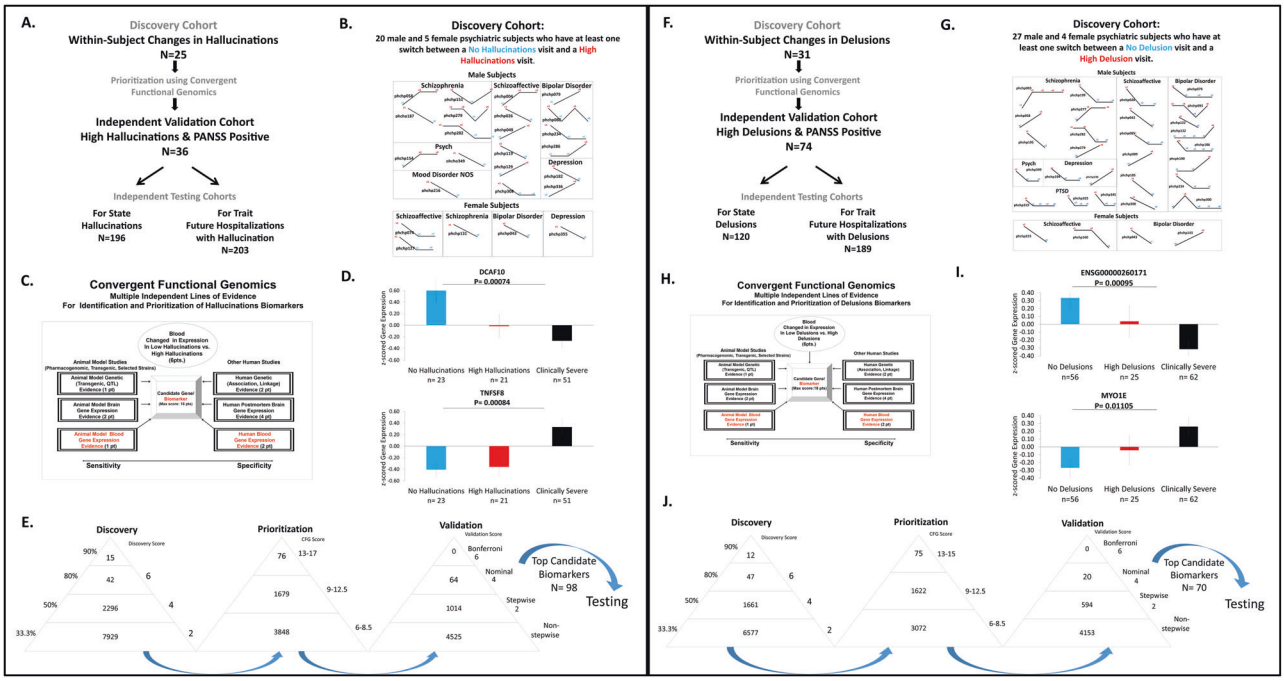


Fig. 1 Steps 1-3: Discovery, Prioritization and Validation of Biomarkers for Hallucinations and Delusions. A, **F** Cohorts used in study, depicting flow of discovery, prioritization, and validation of biomarkers from each step. **B**, **G** Discovery cohort longitudinal within-subject analysis. Phchp### is study ID for each subject. V# denotes visit number. **C**, **H** Prioritization using Convergent Functional Genomics (CFG). **D**, **I** Validation -biomarkers are assessed for stepwise change from discovery subjects with no symptoms, high symptoms to the validation subjects with clinically severe symptoms, using ANOVA. The histograms depict a top increased and a top decreased biomarker (**E**, **J**). Number of probesets and scoring at each of the Steps. Step 1 -Discovery probesets are identified based on their score for tracking symptoms and ranked 33% (2 pt), 50% (4 pt) and 80% (6 pt). Step 2- Prioritization with CFG for prior evidence of involvement in psychotic disorders. Maximum of 12 pt. Genes scoring at least 6 pt out of a maximum possible of 18 pt after Discovery and Prioritization are carried forward to the validation step. Step 3- Validation in an independent cohort of psychiatric patients with clinically severe hallucination (P3 or P1 ≥ 4, PANSS ≥ 21) We selected the top CFE score ≥ 14 ($n = 98$ for hallucinations, $n = 70$ for delusions) for further testing and characterization.

GeneCards to confirm the primary gene symbol. In addition, for those probesets that were not assigned a gene symbol by NetAffyx, we used GeneAnnot (https://genecards.weizmann.ac.il/geneannot/index.shtml), or if need be UCSC (https://genome.ucsc.edu), to obtain gene symbol for these uncharacterized probesets, followed by GeneCard. Genes were then scored using our manually curated CFG databases as described below (Fig. 1C, H).

Step 2: Prioritization using Convergent Functional Genomics (CFG)

Databases: We have established in our laboratory (Laboratory of Neurophenomics, www.neurophenomics.info) manually curated databases of the human gene expression/protein expression studies (postmortem brain, peripheral tissue/fluids: CSF, blood and cell cultures), human genetic studies (association, copy number variations), and animal model gene expression and genetic studies, published to date on psychiatric disorders. Only findings deemed significant in the primary publication, by the study authors, using their particular experimental design and thresholds, are included in our databases. Our databases include only primary literature data and do not include review papers or other secondary data integration analyses to avoid redundancy and circularity. We also favored unbiased discovery studies over candidate genes hypothesis-driven studies. These large and constantly updated databases have been used in our CFG cross validation and prioritization platform (Fig. 1C, H). Data from 1565 papers on psychotic disorders were present in the databases at the time of the CFG analyses (human genetic studies-787, human brain studies-362, human peripheral tissue/fluids- 313, non-human genetic studies-139, non-human brain/studies-135, non-human peripheral tissue/fluids- 12). We have developed in our lab a computerized CFG Wizard to automate and score in bulk large lists of genes by integrating evidence from these large databases, checked against manual scoring. Analyses were performed as previously described.

Step 3: Validation analyses. We examined which of the top candidate genes (score of 6 or above after the first two steps), were stepwise changed in expression from the no psychosis discovery group to the high psychosis discovery group to the clinically severe validation cohort (Fig. 1D, I). A total score of 6 or above after the first two steps permits the inclusion of potentially novel genes with maximal score of 6 from Discovery but no external evidence from Prioritization.

The AP derived and DE derived lists of genes were combined, and the gene expression data corresponding to them was used for the validation analysis. The groups were assembled out of Affymetrix.cel data that was RMA normalized by gender and diagnosis. We transferred the log transformed expression data to an Excel sheet, and non-log transformed the data by taking 2 to the power of the transformed expression value. We then Z-scored the values by gender and diagnosis. We then imported the Excel sheets with the Z-scored by gender and diagnosis expression data into Partek, and statistical analyses were performed using a one-way ANOVA for the stepwise changed probesets, and also did a stringent Bonferroni correction for all the probesets tested in ANOVA (Fig. 1E, J).

Top candidate biomarkers (after the first 3 steps)

Adding the scores from the first three steps into an overall convergent functional evidence (CFE) score (Fig. 1E, J), we ended up with a list of 98 top candidate biomarkers (98 probesets in 74 genes) for hallucinations and a list of 70 top candidate biomarkers for delusions (70 probesets in 64 genes). These top candidate biomarkers were carried forward into additional analyses for biological understanding and for clinical utility testing (Table 1).

Biological understanding

Pathway analyses. IPA (Ingenuity Pathway Analysis, version 24390178, Qiagen), DAVID Functional Annotation Analysis (National Institute of Allergy and Infectious Diseases) version 6.8 (August 2016), and Kyoto Encyclopedia of Genes and Genomes (KEGG) (through DAVID) were used to analyze the biological roles, including top canonical pathways and diseases (Table 2). We performed the pathway analyses for the 98

Table 1. Top biomarkers: convergent functional evidence (CFE).

A. Top 10 Hallucinations Biomarkers

Gene Symbol/Name	Probesets	Step 1 Discovery (Direction of Change in High Hallucinations) Method/Score/%	Step 2 Prioritization Convergent Functional Genomics (CFG) Evidence for Involvement in Psychosis Score	Step 3 Validation ANOVA p-value/ Score	Step 4 Significant Prediction of High Hallucinations State ROC AUC/ *p* value	Step 4 Significant Predictions of First Year Hosp for Hallucinations ROC AUC/ *p* value	Step 4 Significant Predictions of Future Hosp for Hallucinations OR/OR *p* value	Other Psychiatric and Related Disorders Evidence	Drugs that Modulate the Biomarker in Opposite Direction to High Hallucinations	CFE Polyevidence Score for Involvement in Hallucinations (Based on Steps 1–4)
PPP3CB Protein Phosphatase 3 Catalytic Subunit Beta	215586_at	(I) AP/4 53.06%	10	3.51E-02/4 Nominal	**ALL** **L:**(29/317) 0.6/4.51E-02 **Gender** **Males** **L:**(29/264) 0.6/4.68E-02 **M-PSYCHOSIS** **C:**(35/170) 0.62/1.30E-02 **L:**(19/101) 0.64/2.58E-02 **M-SZA** **C:**(21/96) 0.69/4.41E-03 **L:**(11/57) 0.67/4.10E-02 **M-BP** **L:**(5/107) 0.8/1.24E-02	**ALL** **C:**(22/531) 0.63/2.01E-02 **Gender** **Females** **C:**(4/77) 0.89/4.44E-03 **Gender-Dx** **F-BP** **C:**(2/27) 0.9/3.20E-02 **M-MDD** **C:**(3/84) 0.81/3.63E-02	**ALL C:**(109/525) 1.21/5.60E-03 **Gender Females** **C:**(9/79) 1.99/1.73E-02 **Gender** **Males C:**(100/446) 1.16/3.40E-02 **Gender/Dx** **M-BP C:**(23/142) 1.29/4.69E-02 **M-MDD C:**(9/80) 1.61/3.71E-02 **M-PTSD C:**(10/42) 1.8/1.47E-02	Addictions Aging Alcohol ASD BP Dementia Depression Pain PTSD Stress Suicide	Clozapine Valproate	30
DLG1 Discs Large MAGUK Scaffold Protein 1	233869_x_at	(I) AP/4 55.1%	11	3.79E-02/4 Nominal	**ALL** **L:**(29/317) 0.61/2.43E-02 **Gender** **Males** **L:**(29/264) 0.61/2.67E-02 **Gender/Dx** **M-SZA** **C:**(21/96) 0.64/2.64E-02 **L:**(11/57) 0.68/3.44E-02 **M-BP** **L:**(5/107) 0.8/1.11E-02 **M-PSYCHOSIS** **L:**(19/101) 0.64/3.34E-02	**Gender** **Females** **C:**(4/77) 0.84/1.09E-02 **Males** **L:**(10/279) 0.68/3.01E-02 **Gender-Dx** **M-MDD** **L:**(2/50) 0.92/2.38E-02 **M-PTSD** **C:**(4/42) 0.77/3.95E-02	**ALL** **C:**(109/525) 1.18/7.89E-03 **L:**(61/321) 1.46/6.24E-03 **Gender** **Males C:**(100/446) 1.15/2.86E-02 **L:**(58/274) 1.42/1.26E-02 **Gender/Dx** **M-MDD C:**(9/80) 1.75/1.17E-02 **M-PTSD C:**(10/42) 2.05/5.32E-04 **L:**(7/27) 3.94/2.00E-03	Addictions Alcohol Anxiety Pain Psychosis PTSD Stress Suicide	Clozapine	29
ENPP2 Ectonucleotide Pyrophosphatase/ Phosphodiesterase 2	210839_s_at	(D) AP/6 94%	8	4.13E-01/0 Not Stepwise	**ALL** **C:**(62/513) 0.57/3.98E-02 **L:**(29/317) 0.63/9.00E-03 **Gender** **Males** **L:**(29/264) 0.63/1.10E-02 **Gender/Dx** **M-PSYCHOSIS** **C:**(35/170) 0.61/2.64E-02 **L:**(19/101) 0.68/6.47E-03 **M-SZA** **C:**(21/96) 0.67/1.00E-02	**ALL** **C:**(22/531) 0.64/1.35E-02 **Gender** **Males** **C:**(18/454) 0.65/1.40E-02 **L:**(10/279) 0.66/3.85E-02 **Gender-Dx** **M-PSYCHOSIS** **C:**(8/156) 0.73/1.56E-02 **M-SZ** **C:**(4/72) 0.76/3.84E-02 **L:**(3/43) 0.83/2.83E-02	**ALL** **L:**(61/321) 1.43/3.40E-02 **Gender** **Males L:**(58/274) 1.5/2.06E-02 **Gender/Dx** **M-PSYCHOSIS** **C:**(58/161) 1.46/2.90E-02 **L:**(34/96) 1.63/3.37E-02 **M-SZ L:**(13/43) 3.24/1.91E-02	Addictions Aging Alcohol Anxiety BP Dementia Depression Pain Suicide Stress	American Ginseng Estrogen Fluoxetine Kososan Menaquinone-4 Valproate	26

Table 1. continued

A. Top 10 Hallucinations Biomarkers

Gene Symbol/Name	Probesets	Step 1 Discovery (Direction of Change in High Hallucinations) Method/Score/%	Step 2 Prioritization Convergent Functional Genomics (CFG) Evidence for Involvement in Psychosis Score	Step 3 Validation ANOVA p-value/ Score	Step 4 Significant Prediction of High Hallucinations State ROC AUC/ *p* value	Step 4 Significant Predictions of First Year Hosp for Hallucinations ROC AUC/ *p* value	Step 4 Significant Predictions of Future Hosp for Hallucinations OR/OR *p* value	Other Psychiatric and Related Disorders Evidence	Drugs that Modulate the Biomarker in Opposite Direction to High Hallucinations	CFE Polyevidence Score for Involvement in Hallucinations (Based on Steps 1–4)
RTN4 Reticulon 4	243031_at	(I) DE/4 64.8%	11	3.79E-01/2 Stepwise	**ALL** **L:**(29/317) 0.64/6.96E-03 **Gender** **Males** **L:**(29/264) 0.64/6.84E-03 **Gender/Dx** **M-PTSD** **C:**(5/45) 0.76/3.02E-02 **M-BP** **L:**(5/107) 0.88/1.91E-03 **M-PSYCHOSIS** **L:**(19/101) 0.66/1.66E-02 **M-SZ** **L:**(8/44) 0.79/5.77E-03	**ALL** **L:**(11/326) 0.65/4.70E-02 **Gender** **Males** **L:**(10/281) 0.69/2.31E-02 **Gender-Dx** **M-MDD** **L:**(2/50) 0.86/4.16E-02 **M-PTSD** **C:**(4/42) 0.89/5.07E-03 **L:**(3/27) 0.83/3.20E-02	**Gender/Dx** **M-PTSD C:**(10/42) 2.83/1.12E-04 **L:**(7/27) 3.79/5.92E-03	Addictions Alcohol BP Dementia Depression Neurological Pain PTSD Stress Suicide	Omega-3 fatty acids Valproate	26
ZEB2 Zinc Finger E-Box Binding Homeobox 2	239296_at	(I) AP/2 40.8%	10	1.13E-01/2 Stepwise	**ALL** **C:**(62/513) 0.58/1.77E-02 **Gender** **Males** **C:**(58/426) 0.58/2.30E-02 **Gender/Dx** **M-PSYCHOSIS** **C:**(35/170) 0.64/6.71E-03 **M-SZA** **C:**(21/96) 0.68/6.30E-03	**ALL** **C:**(22/531) 0.62/2.73E-02 **Gender** **Females** **C:**(4/77) 0.81/1.84E-02 **Gender-Dx** M-PTSD **C:**(4/42) 0.82/1.79E-02 **L:**(3/27) 0.89/1.54E-02	**ALL C:**(109/525) 1.2/3.82E-03 **Gender Females** **C:**(9/79) 1.51/3.81E-02 **Males C:**(100/446) 1.17/1.74E-02 **Gender/Dx** **F-BP** **C:**(2/27) 2.05/3.25E-02 **M-PTSD C:**(10/42) 2.53/6.44E-04 **L:**(7/27) 5.75/2.96E-03	Addictions Aging Alcohol Anxiety Dementia Depression Psychosis PTSD Sleep Stress Suicide	Celastrol Clozapine Omega-3 fatty acids	26
ZNF24 Zinc Finger Protein 24	203247_s_at	(I) DE/4 50%	9	4.73E-01/2 Stepwise	**ALL** **L:**(29/317) 0.63/1.29E-02 **Gender Males** **L:**(29/264) 0.63/1.20E-02 **Gender/Dx** **M-SZA** **C:**(21/96) 0.64/2.33E-02 **L:**(8/44) 0.78/6.86E-03 **M-BP** **L:**(5/107) 0.83/6.57E-03 **M-PSYCHOSIS** **L:**(19/101) 0.67/9.26E-03	**Gender** **Females** **C:**(4/77) 0.78/2.84E-02 **Gender-Dx** **F-BP** **C:**(2/27) 0.86/4.78E-02 **M-PTSD** **L:**(3/27) 0.81/4.48E-02	**ALL C:**(109/525) 1.2/1.43E-02 **L:**(61/321) 1.33/3.41E-02 **Gender Females** **C:**(9/79) 1.64/4.65E-02 **Males C:**(100/446) 1.17/4.20E-02 **Gender/Dx** **M-BP C:**(23/142) 1.47/1.17E-02 **M-PTSD C:**(10/42) 3.12/2.99E-04 **L:**(7/27) 4.04/1.34E-02	Alcohol BP Dementia Pain Stress Suicide	Celastrol Clozapine Valproate	25

Table 1. continued

A. Top 10 Hallucinations Biomarkers

Gene Symbol/Name	Probesets	Step 1 Discovery (Direction of Change in High Hallucinations) Method/Score/%	Step 2 Prioritization Convergent Functional Genomics (CFG) Evidence for Involvement in Psychosis Score	Step 3 Validation ANOVA p-value/ Score	Step 4 Significant Prediction of High Hallucinations State ROC AUC/ *p* value	Step 4 Significant Predictions of First Year Hosp for Hallucinations ROC AUC/ *p* value	Step 4 Significant Predictions of Future Hosp for Hallucinations OR/OR *p* value	Other Psychiatric and Related Disorders Evidence	Drugs that Modulate the Biomarker in Opposite Direction to High Hallucinations	CFE Polyevidence Score for Involvement in Hallucinations (Based on Steps 1–4)
FNBP1 Formin Binding Protein 1	244286_at	(I) DE/4 50%	8	2.84E-01/2 Stepwise	**ALL** **L:**(29/317) 0.59/4.93E-02 **Gender** **Males** **L:**(29/264) 0.6/4.31E-02 **M-BP** **L:**(5/107) 0.83/6.85E-03 **M-PSYCHOSIS** **C:**(35/170) 0.6/3.96E-02 **L:**(19/101) 0.63/4.04E-02 **M-SZ** **L:**(8/44) 0.69/4.42E-02	**Gender** **Females** **C:**(4/77) 0.8/2.17E-02 **Gender-Dx** **M-PTSD** **C:**(4/42) 0.76/4.73E-02 **M-SZ** **L:**(3/43) 0.79/4.76E-02	**ALL C:**(109/525) 1.18/1.33E-02 **Gender** **Males C:**(100/446) 1.17/2.36E-02 **Gender/Dx** **M-PTSD C:**(10/42) 3.41/5.55E-05 **L:**(7/27) 4.44/3.25E-03	Aging Alcohol ASD BP Dementia Depression Pain PTSD Suicide Stress	Olanzapine Valproate	24
DST Dystonin	215016_x_at	(I) AP/4 51%	10	6.44E-02/0 Not Stepwise	**ALL** **C:**(62/513) 0.57/3.93E-02 **L:**(29/317) 0.65/4.61E-03 **Gender** **Males** **C:**(58/426) 0.57/4.80E-02 **L:**(29/264) 0.65/3.98E-03 **Gender/Dx** **M-PSYCHOSIS** **C:**(35/170) 0.61/1.99E-02 **L:**(19/101) 0.78/6.13E-05 **M-SZ** **L:**(8/44) 0.84/1.43E-03 **M-SZA** **C:**(21/96) 0.63/3.10E-02 **L:**(11/57) 0.75/5.42E-03	**Gender-Dx** **M-MDD** **L:**(2/50) 0.96/1.47E-02	**ALL** **C:**(109/525) 1.26/2.64E-03 **L:**(61/321) 1.46/1.91E-03 **Gender** **Males C:**(100/446) 1.28/1.96E-03 **L:**(58/274) 1.46/1.82E-03 **Gender/Dx** **M-MDD C:**(9/80) 2.4/1.86E-05 **L:**(6/48) 6.96/1.83E-04 **M-SZ L:**(13/43) 1.42/4.97E-02	Alcohol Anxiety ASD Dementia Depression Memory Pain PTSD Stress Suicide	Clozapine Risperidone	23
FAT4 FAT Atypical Cadherin 4	219427_at	(D) AP/6 80%	8	8.09E-01/0 Not Stepwise	**Gender/Dx** **M-SZ** **L:**(8/44) 0.69/4.71E-02	**ALL** **C:**(22/531) 0.61/4.64E-02 **Gender-Dx** **F-BP** **C:**(2/27) 0.86/4.78E-02	**ALL** **C:**(109/525) 1.27/9.51E-03 **L:**(61/321) 1.98/8.70E-05 **Gender** **Males C:**(100/446) 1.29/9.62E-03 **L:**(58/274) 1.89/4.27E-04 **Gender/Dx** **M-MDD L:**(6/48) 2.53/4.73E-02 **M-PSYCHOSIS** **L:**(34/96) 1.63/2.86E-02 **M-PTSD C:**(10/42) 1.83/4.84E-02 **L:**(7/27) 3.54/2.96E-02 **M-SZA L:**(21/53) 1.65/4.74E-02	Aging Alcohol Stress Suicide	Citalopram Jujuboside A Valproate	23

Table 1. continued

A. Top 10 Hallucinations Biomarkers

Gene Symbol/Name	Probesets	Step 1 Discovery (Direction of Change in High Hallucinations) Method/Score/%	Step 2 Prioritization Convergent Functional Genomics (CFG) Evidence for Involvement in Psychosis Score	Step 3 Validation ANOVA p-value/ Score	Step 4 Significant Prediction of High Hallucinations State ROC AUC/ *p* value	Step 4 Significant Predictions of First Year Hosp for Hallucinations ROC AUC/ *p* value	Step 4 Significant Predictions of Future Hosp for Hallucinations OR/OR *p* value	Other Psychiatric and Related Disorders Evidence	Drugs that Modulate the Biomarker in Opposite Direction to High Hallucinations	CFE Polyevidence Score for Involvement in Hallucinations (Based on Steps 1–4)
PDE4B Phosphodiesterase 4B	222326_at	(I) DE/4 66.7%	11	9.43E-01/0 Not Stepwise	**ALL** **L:**(29/317) 0.63/1.04E-02 **Gender Males** **L:**(29/264) 0.63/1.03E-02 **Gender/Dx** **M-BP** **L:**(5/107) 0.89/1.66E-03 **M-PSYCHOSIS** **L:**(19/101) 0.67/1.17E-02 **M-SZA** **C:**(21/96) 0.62/4.92E-02		**ALL** **L:**(61/323) 1.31/3.64E-02 **Gender** **Males L:**(58/276) 1.31/3.46E-02 **Gender/Dx** **M-PSYCHOSIS** **L:**(34/96) 1.51/1.28E-02 **M-PTSD C:**(10/42) 2.73/2.91E-04 **L:**(7/27) 5.8/4.64E-03	Addictions Alcohol ASD BP Depression Pain PTSD Suicide Stress	Clozapine Fluoxetine Lithium Valproate	23

B. Top 10 Delusions Biomarkers.

Gene Symbol/Name	Probesets	Step 1 Discovery (Direction of Change in High Delusions) Method/ Score/ % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence for Involvement in Psychosis Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Best Significant Prediction of High Delusions State ROC AUC/ p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 4 Best Significant Predictions of First Year Hosp for Delusions ROC AUC/ p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 4 Best Significant Predictions of All Future Hosp for Delusions OR/OR p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 5 Other Psychiatric and Related Disorders Evidence	Step 6 Drugs that Modulate the Biomarker in Opposite Direction to High Delusions	CFE Polyevidence Score for Involvement in Delusions (Based on Steps 1-4)
AUTS2 Activator Of Transcription And Developmental Regulator AUTS2	242721_at	(I) DE/4 52.8%	10	4.96E-01 /0 Not Stepwise	**ALL** **C:**(36/315) 0.62/1.14E-02 **Gender** **Males** **C:**(32/285) 0.62/1.28E-02 **Gender/Dx** **M-PSYCHOSIS** **C:**(29/138) 0.67/2.74E-03 **M-SZA** **C:**(16/76) 0.71/4.85E-03	**ALL** **L:**(7/306) 0.74/1.38E-02 **Gender Females** **L:**(2/54) 0.86/4.51E-02	**Gender Females** **C:**(12/90) 1.47/1.43E-02 **Gender/Dx** **F-MDD** **C:**(2/31) 1.71/1.86E-02 **M-PSYCHOSIS** **L:**(23/74) 1.49/2.63E-02 **M-SZ** **L:**(12/36) 2.91/9.33E-03 **F-SZA** **C:**(6/9) 4.45/2.21E-02	Addictions Aging Alcohol Anxiety ASD BP Dementia Depression Memory PTSD	Citalopram Lithium Psychotherapy Valproate	24
MACROD2 Mono-ADP Ribosylhydrolase 2	242468_at	(I) AP/6 100%	8	1.08E-01 /0 Not Stepwise	**Gender** **Females** **C:**(4/30) 0.77/4.38E-02	**Gender/Dx** **M-BP** **C:**(2/146) 0.87/3.72E-02	**ALL** **C:**(84/488) 1.2/2.76E-02 **Gender** **Males** **C:**(72/398) 1.25/1.07E-02 **Gender/Dx** **M-BP** **C:**(24/135) 1.32/4.62E-02 **M-SZA** **C:**(22/66) 1.32/4.97E-02	Alcohol BP Depression Pain PTSD Stress Suicide	Clozapine	21

Table 1. continued

B. Top 10 Delusions Biomarkers.

Gene Symbol/Name	Probesets	Step 1 Discovery (Direction of Change in High Delusions) Method/ Score/ % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence for Involvement in Psychosis Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Best Significant Prediction of High Delusions State ROC AUC/ p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 4 Best Significant Predictions of First Year Hosp for Delusions ROC AUC/ p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 4 Best Significant Predictions of All Future Hosp for Delusions OR/OR p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 5 Other Psychiatric and Related Disorders Evidence	Step 6 Drugs that Modulate the Biomarker in Opposite Direction to High Delusions	CFE Polyevidence Score for Involvement in Delusions (Based on Steps 1-4)
NR4A2 Nuclear Receptor Subfamily 4 Group A Member 2	204621_s_at	(I) DE/2 33.3%	11	9.55E-01 /2 Stepwise	NA	**ALL** **C:**(15/500) 0.68/7.77E-03 **L:**(7/306) 0.78/6.36E-03 **Gender Females** **C:**(4/88) 0.81/1.86E-02 **L:**(2/54) 0.94/1.76E-02 **Gender/Dx** **M-BP** **C:**(2/146) 0.98/1.05E-02 **L:**(1/96) 1/4.33E-02 **F-PSYCHOSIS** **C:**(4/11) 0.96/7.01E-03 **L:**(2/6) 1/3.20E-02 **F-SZA** **C:**(3/9) 0.94/1.94E-02 **L:**(2/6) 1/3.20E-02	**Gender Females** **C:**(12/90) 1.74/1.29E-02 **Gender/Dx** **F-PSYCHOSIS** **C:**(7/11) 3.61/1.24E-02 **L:**(4/6) 7.9/4.92E-02 **F-SZA** **C:**(6/9) 3.42/2.36E-02 **L:**(4/6) 7.9/4.92E-02	Addictions Aging Alcohol Anxiety ASD BP Dementia Depression Pain Sleep Stress	Gamma frequency Ibuprofen Modafinil Psychotherapy Vilazodone	21
PDE4D Phosphodiesterase 4D	236610_at	(D) AP/4 60.3%	10	4.18E-01 /0 Not Stepwise	**ALL** **L:**(20/194) 0.62/4.08E-02	**Gender/Dx** **M-PSYCH** **C:**(3/7) 1/1.69E-02 **M-PSYCHOSIS** **L:**(2/74) 0.94/1.79E-02 **M-SZ** **L:**(2/36) 0.96/1.61E-02	**Gender Females** **L:**(7/56) 3.06/3.74E-02	Aging Alcohol Anxiety Dementia Depression PTSD Sleep Stress Suicide	Clozapine Olanzapine TCA Valproate	21
PDP1 Pyruvate DehyrogeNAse Phosphatase Catalytic Subunit 1	218273_s_at	(I) DE/4 55.6%	8	9.72E-01 /2 Stepwise	**ALL** **L:**(20/194) 0.67/5.28E-03 **Gender** **Females** **L:**(2/18) 0.97/1.75E-02 **Males** **L:**(18/176) 0.64/2.72E-02 **Gender/Dx** **F-MDD** **L:**(2/18) 0.97/1.75E-02 **M-PSYCHOSIS** **L:**(17/80) 0.68/1.22E-02 **M-SZ** **L:**(8/36) 0.73/2.62E-02	**Gender Males** **C:**(11/412) 0.67/2.50E-02 **Gender/Dx** **M-BP** **C:**(2/146) 1/7.67E-03	**Gender/Dx** **M-SZA** **C:**(22/66) 1.49/2.02E-02	Alcohol Dementia Depression PTSD Stress	Fluoxetine Gamma frequency Lithium Omega-3 fatty acids	21

Table 1. continued

B. Top 10 Delusions Biomarkers.

Gene Symbol/Name	Probesets	Step 1 Discovery (Direction of Change in High Delusions) Method/ Score/ % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence for Involvement in Psychosis Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Best Significant Prediction of High Delusions State ROC AUC/ p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 4 Best Significant Predictions of First Year Hosp for Delusions ROC AUC/ p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 4 Best Significant Predictions of All Future Hosp for Delusions OR/OR p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 5 Other Psychiatric and Related Disorders Evidence	Step 6 Drugs that Modulate the Biomarker in Opposite Direction to High Delusions	CFE Polyevidence Score for Involvement in Delusions (Based on Steps 1-4)
RORA RAR Related Orphan Receptor A	240951_at	(I) DE/4 52.8%	11	1.08E-01 /0 Not Stepwise	**Gender/Dx** **M-SZ** **C:**(13/62) 0.65/4.93E-02	**Gender/Dx** **M-BP** **C:**(2/146) 0.9/2.64E-02 **L:**(1/96) 1/4.33E-02	**ALL** **C:**(84/488) 1.25/1.30E-02 **Gender** **Males** **C:**(72/398) 1.33/3.43E-03 **Gender/Dx** **M-BP** **C:**(24/135) 2.08/1.08E-05 **L:**(15/88) 2.26/3.10E-03 **M-SZA** **L:**(11/38) 2.06/1.28E-02	Addictions Alcohol Anxiety BP Dementia Depression Memory Pain Stress	Clozapine Fluoxetine Lithium Risperidone Vortioxetine	21
CHD9 Chromodomain Helicase DNA Binding Protein 9	220586_at	(I) AP/6 98.7%	8	1.21E-02 /0 Not Stepwise		**Gender Females** **L:**(2/54) 0.91/2.44E-02 **Gender/Dx** **M-BP** **C:**(2/146) 0.92/2.17E-02 **L:**(1/96) 0.99/4.67E-02 **F-PSYCHOSIS** **L:**(2/6) 1/3.20E-02 **F-SZA** **C:**(3/9) 0.89/3.54E-02**L:**(2/6) 1/3.20E-02	**ALL** **C:**(84/488) 1.19/3.03E-02 **Gender/Dx** **F-MDD** **C:**(2/31) 2.41/1.85E-02 **M-PSYCHOSIS** **C:**(43/128) 1.26/4.18E-02 **M-SZA** **C:**(22/66) 1.4/2.19E-02 **L:**(11/38) 2.99/5.69E-04	Alcohol Anxiety BP Dementia Memory Pain Psychosis PTSD Suicide	Clozapine Valproate	20
FOXP1 Forkhead Box P1	240666_at	(D) DE/4 51.4%	10	1.92E-01 /0 Not Stepwise	**ALL** **C:**(36/315) 0.6/2.25E-02 **L:**(20/194) 0.62/3.46E-02 **Gender** **Females** **C:**(4/30) 0.8/2.93E-02 **Gender/Dx** **M-PSYCHOSIS** **C:**(29/138) 0.63/1.86E-02 **M-SZ** **C:**(13/62) 0.7/1.57E-02 **L:**(8/36) 0.71/3.39E-02	**Gender/Dx** **M-SZ** **L:**(2/36) 0.87/4.21E-02	**Gender/Dx** **F-PSYCHOSIS** **C:**(7/11) 2.53/2.52E-02	Addictions Aging Alcohol BP Depression Memory PTSD Sleep Stress Suicide	Clozapine Lithium Omega-3 fatty acids Valproate	20
GNAS GNAS Complex Locus	242975_s_at	(I) DE/4 58.3%	10	2.14E-01 /0 Not Stepwise		**ALL** **C:**(15/500) 0.65/2.30E-02 **Gender Males** **C:**(11/412) 0.67/2.80E-02 **Gender/Dx** **M-BP** **C:**(2/146) 0.97/1.10E-02 **L:**(1/96) 1/4.33E-02	**Gender** **Males** **L:**(41/242) 1.29/4.96E-02 **Gender/Dx** **M-BP** **C:**(24/135) 1.35/4.75E-02 **L:**(15/88) 1.81/6.50E-03 **M-SZA** **C:**(22/66) 1.49/1.32E-02	Addictions Aging Alcohol Anxiety BP Depression Memory PTSD Stress Suicide	Clozapine Risperidone Valproate Vortioxetine	20

Table 1. continued

B. Top 10 Delusions Biomarkers.

Gene Symbol/Name	Probesets	Step 1 Discovery (Direction of Change in High Delusions) Method/ Score/ % Up to 6pts	Step 2 External Convergent Functional Genomics (CFG) Evidence for Involvement in Psychosis Up to 12pts	Step 3 Validation ANOVA p-value/ Score Up to 6 pts	Step 4 Best Significant Prediction of High Delusions State ROC AUC/ p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 4 Best Significant Predictions of First Year Hosp for Delusions ROC AUC/ p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 4 Best Significant Predictions of All Future Hosp for Delusions OR/OR p-value 4 pts All 2pts Gender 1 pt Gender /Dx	Step 5 Other Psychiatric and Related Disorders Evidence	Step 6 Drugs that Modulate the Biomarker in Opposite Direction to High Delusions	CFE Polyevidence Score for Involvement in Delusions (Based on Steps 1-4)
ZBTB20 Zinc Finger And BTB Domain Containing 20	239955_at	(D) AP/2 35.3%	10	6.22E-01 /2 Stepwise	**ALL** **C:**(36/315) 0.63/6.23E-03 **L:**(20/194) 0.65/1.57E-02 **Gender** **Females** **C:**(4/30) 0.81/2.55E-02 **Males** **C:**(32/285) 0.61/2.55E-02 **L:**(18/176) 0.65/1.79E-02 **Gender/Dx** **F-MDD** **C:**(3/28) 0.8/4.73E-02 **M-PSYCHOSIS** **C:**(29/138) 0.64/1.17E-02 **L:**(17/80) 0.65/3.37E-02 **M-SZ** **C:**(13/62) 0.74/3.58E-03 **L:**(8/36) 0.79/7.44E-03		**Gender Females** **C:**(12/90) 2.25/2.25E-02 **Gender/Dx** **F-PSYCHOSIS** **C:**(7/11) 6.67/2.59E-02 **F-SZA** **C:**(6/9) 5.31/4.76E-02	Aging Anxiety BP Depression Memory Stress Suicide	Fluoxetine Acetaminophen Risperidone Valproate	20

After Step 4 Testing in independent cohorts for state and trait predictive ability. For Step 4 Predictions, **C**: -cross-sectional (using levels from one visit), **L**: -longitudinal (using levels and slopes from multiple visits). In ALL, by Gender, and personalized by Gender and Diagnosis (Gender/Dx).M-Males, F-Females. MDD-depression, BP-bipolar, SZ-schizophrenia, SZA-schizoaffective, PSYCHOSIS- schizophrenia and schizoaffective combined, PTSD-post-traumatic stress disorder. For Step 4 predictions scoring, 4 pts if significant in **ALL**, 2pts **Gender**, 1 pt **Gender /Dx**. Capped at 4.

biomarkers (84 unique genes) for hallucinations and 70 biomarkers (64 unique genes) for delusions that were the top candidate biomarkers after the discovery, prioritization, and validation.

Networks. For network analyses we performed STRING Interaction network (https://string-db.org) by inputting the genes into the search window, and performed Multiple Proteins Homo sapiens analysis. (Fig. S2A, B).

CFG beyond Psychosis: evidence for involvement in other psychiatric and related disorders. We also used a CFG approach to examine evidence from other psychiatric and related disorders, as exemplified for the list of top biomarkers after Step 4 testing (Table S3A, B). This was not used to prioritize genes, but rather to understand the molecular basis of co-morbidities. We also calculated a genomic co-morbidities % based on number of genes on our list that matched to different other disorders (Table 2E).

Testing for clinical utility in independent cohorts

We tested in independent cohorts of psychiatric patients the ability of each of the top candidate biomarkers (hallucinations $n = 98$, delusions $n = 70$) to assess state severity hallucinations (measured by PANSS, P3), delusions (measured by PANSS, P1), and predict trait risk (future hospitalizations with hallucinations, future hospitalizations with delusions). We conducted our analyses across all patients, as well as personalized by gender and diagnosis. We then predict with the biomarkers from the list in independent cohorts state (high hallucinations PANSS, P3 ≥ 4, high delusions PANSS, P1 ≥ 4), and trait (future hospitalizations with hallucinations and future hospitalizations with delusions) in the first year of follow-up, and in all future years of follow-up.

The test cohorts for predicting hallucinations and delusions (state), and the test cohorts for predicting future hospitalizations with hallucinations and delusions (trait), were assembled out of data that was RMA normalized by gender and diagnosis. The cohorts were completely independent from the discovery and validation cohorts, there was no subject overlap with them. Individual markers used for predictions were Z scored by gender and diagnosis, to be able to combine different biomarkers into panels and to avoid potential artefacts due to different ranges of expression in different gender and diagnoses. Predictions were performed using R-studio. For cross-sectional analyses, we used biomarker expression levels, z-scored by gender and diagnosis. For longitudinal analyses, we combined four measures: biomarker expression levels, slope (defined as ratio of levels at current testing visit vs. previous visit, divided by time between visits), maximum levels (at any of the current or past visits), and maximum slope (between any adjacent current or past visits). For decreased biomarkers, we used the minimum rather than the maximum for level calculations.

Predicting state- Hallucinations/Delusions. Receiver-operating characteristic (ROC) analyses between marker levels and hallucinations\delusions state were performed by assigning subjects visits with a hallucinations PANSS, P3 score ≥4 in the high hallucinations category, and subjects with a delusions PANSS, P1 score of ≥4 into the high delusions category. We used the pROC package of R (Xavier Robin et al. BMC Bioinformatics 2011). (Table 1 and Fig. 2). Additionally, a one-tailed *t* test was performed between high hallucinations/high delusions group vs. the rest, and Pearson R (one-tail) was calculated between hallucinations/delusions scores and biomarker levels.

Predicting trait- future psychiatric hospitalization with hallucinations/ delusions as a symptom/reason for admission. We conducted analyses for predicting future psychiatric hospitalizations with hallucinations \delusions as a symptom/reason for admission in the first year following each testing visit, in subjects that had at least one year of follow-up in the VA system, in which we have access to complete electronic medical records. ROC analyses between biomarkers measures (cross-sectional, longitudinal) at a specific testing visit and future hospitalizations were performed as described above, based on assigning if subjects had been admitted to the hospital with hallucinations\delusions or not. Additionally, a one tailed t-test with unequal variance was performed between groups of subject visits with and without future hospitalization with hallucinations \delusions. Pearson R (one-tail) correlation was performed between hospitalization frequency (number of hospitalizations with hallucinations \delusions divided by duration of follow-up) and marker levels. A Cox regression was performed using the time in days from the testing visit date to first hospitalization date in the case of patients who had been hospitalized, or 365 days for those who did not. The odds ratio was calculated such that a value greater than 1 always indicates increased risk for hospitalization, regardless if the biomarker is increased or decreased in expression.

We also conducted Cox regression and Pearson R analyses for all future hospitalizations with hallucinations\delusions, including those occurring beyond one year of follow-up, in the years following testing (hallucinations: on average 7.34 years per subject, range.07–15.24 years, delusions: on average 7.34 years per subject, range 0.07–15.24 years), as these calculations, unlike the ROC and *t* test, account for the actual length of follow-up, which varied from subject to subject. The ROC and *t* test might in fact, if used, under-represent the power of the markers to predict, as the more severe psychiatric patients are more likely to move geographically and/or be lost to follow-up. The Cox regression was performed using the time in days from visit date to first hospitalization date in the case of patients who had hospitalizations with hallucinations\delusions, or from visit date to last note date in the electronic medical records for those who did not.

Therapeutics

Pharmacogenomics. We analyzed which of the top biomarkers for delusions and for hallucinations after Steps 1–4 are known to be changed in expression by existing drugs in a direction opposite to the one in disease, using our CFG databases (Tables 3 and S4B).

Report generation

We present examples of how a report to doctors might look, using the above insights. We used a panel of the top 10 biomarkers for each to generate a score for delusions severity and hallucinations severity. For both hallucinations and delusions, out of a dataset of 794 subject visits, we chose a case study of a patient with a high past severity score (CFI SZ).

First, we removed the subject from the dataset, and divided the remaining dataset into two populations: those who had a high hallucinations or delusions score (P3 or P1 ≥ 4) and those who had no hallucinations or delusions (P3 or P1 = 1). For the two groups, we calculated the average Z-scored expression values for each biomarker in the panel. Biomarkers whose average Z-scored expression values were non-stepwise were not used moving forward with the report analyses. We then compared the biomarkers for the subject of interest to these reference levels. If a biomarker was higher than the average of the high group it got a 1, if it was below the average of the no group it got a 0, and if it was in between, it got a 0.5 for increased biomarkers. For decreased biomarkers, if it was lower than the average of the high group it got a 1, if it was higher than the average of the no group it got a 0, and if it was in between in got a 0.5. The hallucinations/delusions state risk score is the average score of all stepwise biomarkers multiplied by 100, generating 4 risk categories: high (red), intermediate high (orange), intermediate low (yellow), and low (green) for delusion/hallucination severity. The chronic hallucinations/delusions risk score was calculated the same way using future hospitalization biomarkers. These percentile scores of the patient are provided in the report (Fig. 3).

Second, for each stepwise biomarker in the panel, we also have a list of existing psychiatric medications that modulate the expression of the biomarker for hallucinations/delusions. Each medication got a score determined by the score of each biomarker that is modulated by it. A medication can modulate more than one biomarker. We then calculated an average score for each medication based on its effects on all the biomarkers in the panel, and multiplied that by 100, resulting in a score of 0 to 100 for each medication. Thus, psychiatric medications are matched to the patient and ranked in order of impact on the panel.

RESULTS

In Step 1 Discovery, we identified candidate blood gene expression biomarkers that: 1. change in expression in blood between no and high hallucinations, or delusions states, 2. track the hallucinations or delusions state across visits in a subject, and 3. track the hallucinations or delusions states in multiple subjects. We used quantitative measures for hallucinations state (item P3 in PANSS) and delusions state (item P1 in PANSS). At a phenotypic level, these items quantify hallucinations and delusions state at a particular moment in time, based on the rater interview of the

Table 2. Biological analyses: A. Hallucinations Pathways. B. Hallucinations Diseases C. Delusions Pathways. D. Delusions Diseases E. Genomic Co-Morbidity.

A. Hallucinations Pathways	**DAVID GO Functional Annotation Biological Processes**					**KEGG Pathways**				**Ingenuity Pathways (Fold change)**		
	#	**Term**	**Count**	**%**	***P* Value**	**Term**	**Count**	**%**	***P* Value**	**Top Canonical Pathways**	***P* Value**	**Overlap**
Hallucinations Top Candidate Biomarkers ($n = 98$ probesets, 84 genes)	1	Nervous System Development	34	40.5	1.1E-9	**Glutamatergic synapse**	8	9.5	8.90E-06	**nNOS Signaling in Neurons**	3.21E-10	14.9% (7/47)
	2	Cellular Component Morphogenesis	19	22.6	1.0E-6	Calcium signaling pathway	10	11.9	2.40E-05	Gαq Signaling	1.01E-08	5.3% (9/170)
	3	Phosphate-Containing Compound Metabolic Process	33	39.3	1.2E-6	Inflammatory mediator regulation of TRP channels	7	8.3	4.20E-05	Cardiac Hypertrophy Signaling (Enhanced)	6.28E-08	5.3% (7/131)
	4	Cell Morphogenesis	18	21.4	1.2E-6	Dopaminergic synapse	7	8.3	2.20E-04	fMLP Signaling in Neutrophils	4.48E-07	5.3%
	5	Neurgenesis	23	27.4	1.9E-6	Circadian entrainment	6	7.1	4.30E-04	Role of NFAT in Cardiac Hypertrophy	1.40E-06	3.6% (8/224)

B. Hallucinations Diseases	**David Genetic Association Disease**					**Ingenuity Pathways Disease**		
	#	**Term**	**Count**	**%**	***P* Value**	**Diseases and Disorders**	***P* Value**	**# Molecules**
Hallucinations Top Candidate Biomarkers ($n = 98$ probesets, 84 genes)	1	Schizophrenia	19	22.6	2.80E-09	Heriditary Disorder	8.42-05 - 1.61E-12	48
	2	Schizophrenia	10	11.9	2.30E-04	Neurological Disease	7.39E-05 – 1.61E-12	75
	3	Tobacco Use Disorder	33	39.3	2.30E-04	Organismal Injury and Abnormalities	8.82E-05 – 1.61E-12	84
	4	Myocardial Infarction	11	13.1	7.40E-04	Cancer	8.82E-05 – 2.87E-12	82
	5	Schizophrenia and Bipolar disorder	3	3.6	6.10E-03	Gastrointestinal Disease	8.68E-0.05 – 2.87E-12	80

C. Delusions Pathways	**DAVID GO Functional Annotation Biological Processes**					**KEGG Pathways**				**Ingenuity Pathways (Fold change)**		
	#	**Term**	**Count**	**%**	***P* Value**	**Term**	**Count**	**%**	***P* Value**	**Top Canonical Pathways**	***P* Value**	**Overlap**
Delusions Top Candidate Biomarkers ($n = 70$ probesets, 64 genes)	1	Branching Involved in Salivary Gland Morphogenesis	4	6.2	7.7E-6	**Rap1 Signaling Pathway**	9	14.1	9.6E-6	**Sperm Motility**	4.87E-07	3.1% 8/254
	2	Positive Regulation of Cardiac Muscle Cell Proliferation	4	6.2	7.0E-5	Adherens Junction	5	7.8	4.8E-4	Insulin Secretion Signaling Pathway	7.30E-07	3.0% 8/268
	3	Positive Regulation of Transcription from RNA Polymerase II Promotor	13	20.3	2.1E-4	Aldosterone Synthesis and Secretion	5	7.8	7.9E-4	CREB Signaling in Neurons	5.58E-06	1.7% 10/602
	4	Multicellular Organism Growth	5	7.8	2.2E-4	Insulin Secretion	5	7.8	9.5E-4	CSCR4 Signaling	6.80E-06	3.6% 6/167

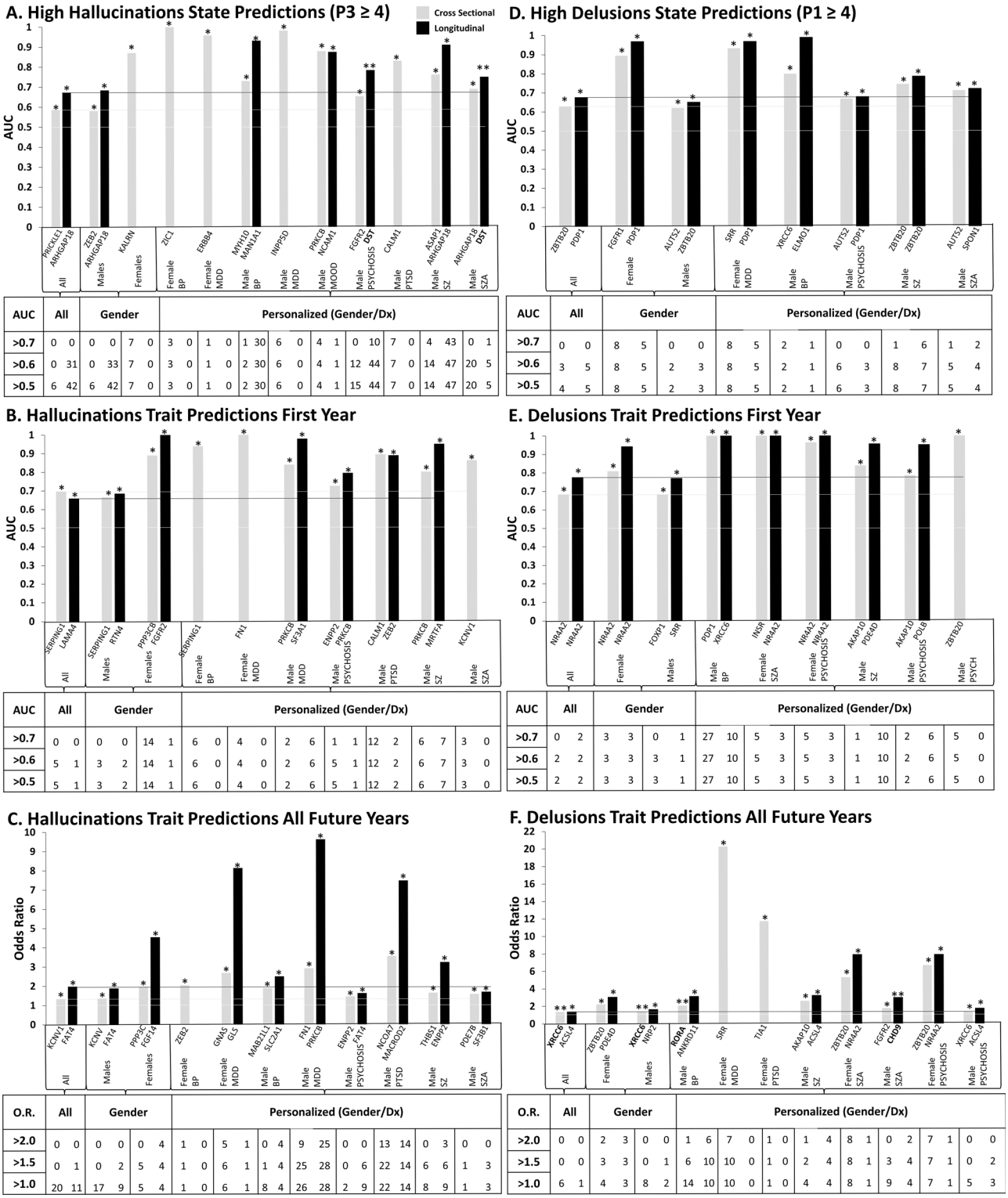
Table 2. continued

C. Delusions Pathways	DAVID GO Functional Annotation Biological Processes					KEGG Pathways				Ingenuity Pathways (Fold change)		
	#	Term	Count	%	*P* Value	Term	Count	%	*P* Value	Top Canonical Pathways	*P* Value	Overlap
	5	Positive Regulation of MAPK Cascade	5	7.8	2.3E-4	Aldosterone-Regulated Sodium Reabsorption	4	6.2	1.1E−3	G-Protein Coupled Receptor Signaling	1.10E-05	2.5% 7/276

D. Delusions Diseases	David Genetic Association Disease					Ingenuity Pathways Disease		
	#	Term	Count	%	*P* Value	Diseases and Disorders	*P* Value	# Molecules
Delusions Top Candidate Biomarkers ($n = 70$ probesets, 64 genes)	1	Schizophrenia	27	42.2	7.7E-17	Neurological Disease	2.25E-05 - 1.92E-19	58
	2	Tobacco Use Disorder	34	53.1	7.0E-8	Organismal Injury and Abnormalities	2.30E-05 - 1.92E-19	64
	3	Depression	8	12.5	6.4E-6	Psychological Disorders	1.84E-05 - 1.31E-16	38
	4	Type 2 Diabetes| Edema| Rosiglitazone	24	37.5	4.6E-5	Developmental Disorder	1.84E-05 - 3.72E-16	41
	5	Several Psychiatric Disorders	9	14.1	5.8E-5	Hereditary Disorder	2.25E-05 - 1.54E-14	35

E. Genomic co-morbidity

Hallucinations Candidate Biomarkers ($n = 98$ probesets in 84 genes)		Delusions Candidate Biomarkers ($n = 70$ probesets in 64 genes)	
Comorbidity	**Percentile**	**Comorbidity**	**Percentile**
Alcohol	94%	Alcohol	96.9%
Depression	86.9%	Depression	85.9%
Stress	81%	Dementia	79.7%
Suicide	76.2%	Stress	78.1%
Dementia	72.6%	Suicide	73.4%
Pain	58.3%	Bipolar	56.3%
Bipolar	51.2%	Pain	53.1%
Aging	45.2%	Anxiety	48.4%
PTSD	40.5%	Aging	34.4%
Anxiety	35.7%	ASD	31.3%
ASD	23.8%	PTSD	29.7%
Sleep	9.5%	Sleep	18.8%
ADHD	2.4%	Panic	6.3%
Panic	2.4%	ADHD	3.1%



AUC	All		Gender				Personalized (Gender/Dx)											
>0.7	0	0	0	0	7	0	3	0	1	0	1	30	6	0	4	1	0	10	7	0	4	43	0	1
>0.6	0	31	0	33	7	0	3	0	1	0	2	30	6	0	4	1	12	44	7	0	14	47	20	5
>0.5	6	42	6	42	7	0	3	0	1	0	2	30	6	0	4	1	15	44	7	0	14	47	20	5

AUC	All		Gender				Personalized (Gender/Dx)											
>0.7	0	0	8	5	0	0	8	5	2	1	0	0	1	6	1	2		
>0.6	3	5	8	5	2	3	8	5	2	1	6	3	8	7	5	4		
>0.5	4	5	8	5	2	3	8	5	2	1	6	3	8	7	5	4		

AUC	All		Gender				Personalized (Gender/Dx)											
>0.7	0	0	0	0	14	1	6	0	4	0	2	6	1	1	12	2	6	7	3	0		
>0.6	5	1	3	2	14	1	6	0	4	0	2	6	5	1	12	2	6	7	3	0		
>0.5	5	1	3	2	14	1	6	0	4	0	2	6	5	1	12	2	6	7	3	0		

AUC	All		Gender				Personalized (Gender/Dx)											
>0.7	0	2	3	3	0	1	27	10	5	3	5	3	1	10	2	6	5	0
>0.6	2	2	3	3	3	1	27	10	5	3	5	3	1	10	2	6	5	0
>0.5	2	2	3	3	3	1	27	10	5	3	5	3	1	10	2	6	5	0

O.R.	All		Gender				Personalized (Gender/Dx)											
>2.0	0	0	0	0	0	4	1	0	5	1	0	4	9	25	0	0	13	14	0	3	0	0
>1.5	0	1	0	2	5	4	1	0	6	1	1	4	25	28	0	6	22	14	6	6	1	3
>1.0	20	11	17	9	5	4	1	0	6	1	8	4	26	28	2	9	22	14	8	9	1	3

O.R.	All		Gender				Personalized (Gender/Dx)											
>2.0	0	0	2	3	0	0	1	6	7	0	1	0	1	4	8	1	0	2	7	1	0	0
>1.5	0	0	3	3	0	1	6	10	10	0	1	0	2	4	8	1	3	4	7	1	0	2
>1.0	6	1	4	3	8	2	14	10	10	0	1	0	4	4	8	1	9	4	7	1	5	3

Fig. 2 Best single biomarkers predictors for state and trait. A–C Hallucinations (**D**, **E**, **F**). Delusions. From top candidate biomarkers after Steps 1–3 (Discovery, Prioritization, Validation) ($n = 98$ for hallucinations, $n = 70$ for delusions). Bar graph shows nominally significant predictive biomarkers in each group. Table underneath the figures displays the actual number of biomarkers for each group whose ROC AUC p values (**A**, **B**, **D**, **E**) and Cox Odds Ratio p values (**C**, **F**) are at least nominally significant. Some gender and diagnosis group are missing from the graph as they did not have any significant biomarkers. Cross-sectional is based on levels at one visit. Longitudinal is based on levels at multiple visits (integrates levels at most recent visit, maximum levels, slope into most recent visit, and maximum slope). Dividing lines represent the cutoffs for a test performing at chance levels (white), and at the same level as the best biomarkers for all subjects in cross-sectional (gray) and longitudinal (black) based predictions. Biomarkers perform better than chance. Biomarkers performed better when personalized by gender and diagnosis. *nominally significant with $p < 0.05$ ** survived Bonferroni correction for the number of candidate biomarkers tested.

Table 3. Therapeutics: A, B. Top 10 Biomarkers after Step 4 from Table 1 ($n = 10$ each).

A. Hallucinations Top 10 Biomarkers		***B. Delusions Top 10 Biomarkers***	
Treatment	**Percentile**	**Treatment**	**Percentile**
Clozapine	50%	Lithium	50%
Valproate	30%	Clozapine	30%
Celastrol	20%	Fluoxetine	30%
Fluoxetine	20%	Risperidone	30%
Lithium	20%	*Gamma frequency*	*20%*
Magnesium	20%	*Magnesium*	*20%*
Omega-3 fatty acids	20%	*Omega-3 fatty acids*	*20%*
Ginseng	10%	*Therapy*	*20%*
Citalopram	10%	Valproate	20%
Desipramine	10%	Vortioxetine	20%
Estradiol	10%	Citalopram	10%
Genistein	10%	Ibuprofen	10%
Indomethacin	10%	Olanzapine	10%
Jujuboside A	10%	Acetaminophen	10%
Kososan	10%	Vilazodone	10%
Menaquinone-4	10%		
Olanzapine	10%		
C. Hallucinations Top Candidate Biomarkers (n = 98)		***D. Delusions Top Candidate Biomarkers (n = 70)***	
Treatment	**Percentile**	**Treatment**	**Percentile**
Clozapine	32.1%	Clozapine	37.5%
Omega-3 fatty acids	28.6%	Lithium	34.4%
Lithium	26.2%	Fluoxetine	23.4%
Valproate	17.9%	*Omega-3 fatty acids*	20.3%
Magnesium	16.7%	*Magnesium*	18.8%
Fluoxetine	14.3%	Risperidone	14.1%
Citalopram	7.1%	Valproate	14.1%
Risperidone	7.1%	Vortioxetine	12.5%
Escitalopram	6%	Escitalopram	7.8%
Indomethacin	6%	Carbamazepine	6.3%
Minocycline	6%	Citalopram	6.3%
Vortioxetine	6%	*Gamma frequency*	6.3%
Desipramine	4.8%	Ibuprofen	6.3%
Prednisolone	4.8%	Olanzapine	6.3%
Gamma frequency	3.6%	Desipramine	4.7%
Ginseng	3.6%	*Ginseng*	4.7%
Haloperidol	3.6%	Nortriptyline	4.7%
Nortriptyline	3.6%	Dexamethasone	3.1%

Table 3. continued

C. Hallucinations Top Candidate Biomarkers (n = 98)		*D. Delusions Top Candidate Biomarkers (n = 70)*	
Treatment	**Percentile**	**Treatment**	**Percentile**
Olanzapine	3.6%	Diazepam	3.1%
Sertraline	3.6%	Haloperidol	3.1%
Aripiprazole	2.4%	Imipramine	3.1%
Carbamazepine	2.4%	Indomethacin	3.1%
Celastrol	2.4%	Modafinil	3.1%
Dexamethasone	2.4%	Prednisolone	3.1%
Guava	2.4%	Sertraline	3.1%
Kososan	2.4%	*Therapy*	3.1%
Mianserin	2.4%	Trazodone	3.1%
Vitamin D3	2.4%	Vilazodone	3.1%
		Ziprasidone	3.1%

See also Table S4. C, D. Top Candidate Biomarkers after Step 3. Italic- nutraceuticals/non-pharmacological treatments.

subjects (Fig. S1). They have moderate correlations with a self-reported new VAS scale for psychosis we have developed, the Simplified Psychosis Scale -SPS4 (Fig. S2) (hallucinations $R = 0.441$, $p < 0.001$, delusions $R = 0.154$, $p = 0.02$).

For the discovery step, we used a powerful within –subject and then across-subject design in a longitudinally followed cohort of subjects (for hallucinations, $n = 25$ subjects, with 65 visits; for delusions, $n = 31$ subjects with 95 visits) who displayed at least a change in the hallucinations or delusions measure (from 1 to 4 and above, and vice-versa) between at least two consecutive testing visits, to identify differentially expressed genes that track hallucinations and delusions state. Using our 33% of maximum raw score threshold (internal score of 2 pt) [3, 4], for hallucinations we identified 10,282 unique probesets from Affymetrix Absent/Present (AP) analyses and Differential Expression (DE) analyses (Fig. 1). For delusions, we identified 8302 unique probesets. These were carried forward to the prioritization step. These represents approximately an over 5 fold enrichment of the 54,625 probesets on the Affymetrix array.

In Step 2 Prioritization, we used a CFG approach to prioritize the candidate biomarkers identified in the discovery step (33% cutoff, internal score of ≥2 pt.) by using prior published literature evidence (genetic, gene expression and proteomic), from human and animal model studies, for involvement in schizophrenia and other psychotic disorders (Fig. 1 and Table S2). For hallucinations, there were 5603 probesets that had a total score (combined discovery score and prioritization CFG score) of 6 and above. For delusions, there were 4769 probesets. These were carried forward to the validation step. These represent approximately a 10-fold enrichment of the probesets on the Affymetrix array.

In Step 3 Validation, we validated the prioritized candidate biomarkers for change in clinically severe psychosis, in age-matched cohorts ($n = 36$ subjects with 52 visits for hallucinations, and $n = 43$ subjects with 62 visits for delusions), We assessed which biomarkers were stepwise changed in expression from no hallucinations or delusions in discovery cohort, to high hallucinations or delusions in discovery cohort, to clinically severe hallucinations or delusions in validation cohort (Fig. 1). For hallucinations, of the 5603 probesets after the prioritization step, 1078 probesets were stepwise changed. Of these, 64 probesets were nominally significant. For delusions, of the 4769 probesets after the prioritization step, 614 probesets were stepwise, and 20 were nominally significant.

Adding the scores from the first three steps into an overall CFE score (Fig. 1), we ended up with a list of 98 top candidate biomarkers for hallucinations, and 70 top candidate biomarkers for delusions, that had a CFE3 score ≥14, better than 50% of the maximum possible score of 24 after the first three steps, which we decided to use as an empirical cutoff. This represents approximately an over 500-fold enrichment of the probesets on the Affymetrix array. These top candidate biomarkers were carried forward into analyses for understanding biological underpinnings. They were also tested in Step 4 for clinical utility/predictive ability in additional independent cohorts (Fig. 2 and Table 1).

Biological understanding

Biological pathways. We carried out biological pathway analyses using the list of top candidate biomarkers for hallucinations ($n = 84$ genes, 98 probesets), which suggest that glutamatergic synapses and nNOS signaling are involved. For delusions ($n = 64$ genes, 70 probesets), the top pathways were Rap1 signaling and sperm motility (Table 2). Schizophrenia and tobacco use disorders were top diseases identified by the pathway analyses using DAVID, pointing out to a molecular underpinning for this well-known clinical co-morbidity.

Networks and interactions. We carried out a STRING analysis (Fig. S2) of the top candidate biomarkers that revealed groups of

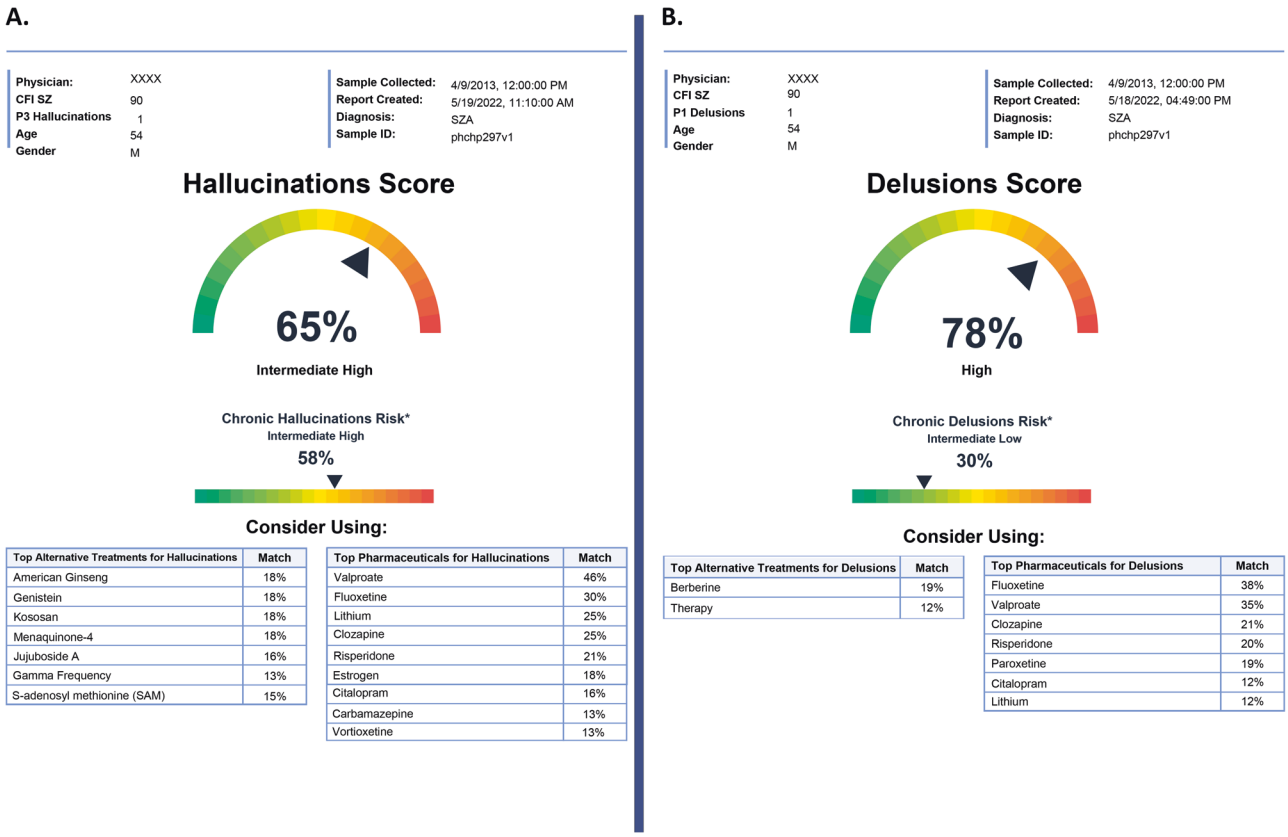


Fig. 3 Example of prototype report for physicians. A Hallucinations. **B** Delusions. Using a panel of the best predictive biomarkers for state and trait in all and by gender. The raw expression values of the biomarkers in the 794 microarrays gene expression database were Z-scored by gender. For state score, the Z-scored expression value of each increased biomarker was compared to the average value for the biomarker in the high hallucinations group (P3, hallucinations ≥4), and the average value of the no hallucinations group (P3, hallucination = 1) resulting in scores of 1 or 0 respectively, and 0.5 if it was in between. The reverse was done for decreased biomarkers. For trait chronic risk score, we calculated the average expression value for a biomarker in the first-year hospitalizations group, and in the not hospitalized in the first-year group, and for all future hospitalizations, and no future hospitalizations. The digitized biomarkers were then added into a polygenic risk score. A similar approach was done for delusions. The digitized biomarkers were also used for matching with existing psychiatric medications and alternative treatments (nutraceuticals and others). We used our large datasets and literature databases to match biomarkers to medications that had effects on gene expression opposite to their expression in high hallucinations or delusions. Each medication matched to a biomarker got the biomarker score of 1, 0.5 or 0. The scores for the medications were added, normalized for the number of biomarkers that were 1 or 0.5 in that patient, resulting in a percentile match. Subject 297v1 is a 54 year old African-American male with a diagnosis of schizoaffective disorder. He had a high CFI-SZ score of 90, indicating he had been severely ill in the past. At the time of testing, his PANSS P3 Hallucinations items was 1, and his P1 delusions item was 1, indicating no symptoms by clinical assessment. His psychiatric medications were divalproex 1250 mg at bedtime and fluphenazine 7.5 mg at bedtime. His nutraceuticals were cyanocobalamin 500 mcg daily, vitamin E 400 units daily, and a multivitamin. The reports indicate that he was a good match for valproate, a medication he was on, and while externally he was rated low on PANSS items, he still had significant internal levels of biomarkers of disease severity.

interacting proteins. For hallucinations, PPP3CB is at the overlap of a network centered on ERBB4, containing DLG1, ENPP2, and DISC1, and one centered on FN1, that contains ZEB2 and RTN4. A third network is centered on ATP5C1 (which encodes a subunit of mitochondrial ATP synthase). For delusions, AUTS2 is at the overlap of a network centered on TCF4, that contains NR4A2, and a network centered on GNAS, that contains PDE4D. A third network is centered on NR3C2 (the mineralocorticoid receptor). These networks may have biological significance and could be targeted therapeutically.

Testing for clinical utility

In Step 4 Testing, we examined in independent cohorts from the ones used for discovery or validation whether the top candidate biomarkers after the first three steps can assess high hallucinations or delusions states, as well as predict of future psychiatric hospitalizations due to hallucinations or delusions (Figs. 1 and 2, and Table S1), using electronic medical records follow-up data of our study subjects (up to 15.24 years from initial visit at the time of the analyses). The gene expression data in the test cohorts was normalized (Z-scored) across genders and various psychiatric diagnoses, before those different demographic groups were combined. This reduces bias from larger demographic groups, and permits them to be combined. We used as predictors biomarker levels information cross-sectionally, as well as expanded longitudinal information about biomarker levels at multiple visits. We tested the biomarkers in all subjects in the independent test cohort, as well as in a more personalized fashion by gender and psychiatric diagnosis (Fig. 2).

Convergent Functional Evidence (CFE)

For the top candidate biomarkers ($n = 98$ for hallucinations, $n = 70$ for delusions), we computed into a CFE score all the evidence from discovery (up to 6 points), CFG prioritization (up to 12 points), validation (up to 6 points), and testing (predicting state high hallucinations or delusions, first year hospitalization with hallucinations or delusions, all future hospitalizations with hallucinations or delusions- up to 4 points each if it significantly predicts in all

subjects, 2 points if in gender, 1 points if in gender/diagnosis). The total score can be up to 36 points: 24 from our own new data, and 12 from literature data used for CFG. We weigh our new data more than the literature data, as it is functionally related to psychosis in 3 independent cohorts (discovery, validation, testing). The goal is to highlight, based on the totality of our data and of the evidence in the field to date, biomarkers that have all around evidence: track hallucinations or delusions, have convergent evidence for involvement in psychotic disorders, and predict hallucinations or delusions state, and future clinical events (Table 1).

Top 10 biomarkers for hallucinations. The top 10 blood biomarkers with the strongest overall CFE for tracking and predicting hallucinations, after all four steps (Table 1) were, in order of CFE4 score: PPP3CB (Protein Phosphatase 3 Catalytic Subunit Beta), DLG1 (Discs Large MAGUK Scaffold Protein 1), ENPP2 (Ectonucleotide Pyrophosphatase/Phosphodiesterase 2), RTN4 (Reticulon 4), ZEB2 (Zinc Finger E-Box Binding Homeobox 2), ZNF24 (Zinc Finger Protein 24), FNBP1 (Formin Binding Protein 1), DST (Dystonin), FAT4 (FAT Atypical Cadherin 4), PDE4B And (Phosphodiesterase 4B).

PPP3CB, the overall top biomarker for hallucinations in this study, is a calcium-dependent, calmodulin-stimulated protein phosphatase which plays an essential role in the transduction of intracellular Ca(2+)-mediated signals [6, 7], and dephosphorylates DARPP32 [6]. Abnormalities in calcium signaling may be a central abnormality in schizophrenia [8]. PPP3CB, increased in expression in blood in high hallucinations in our work, has previous, convergent evidence for involvement in schizophrenia. It is increased in expression in dorsolateral PFC (left hemisphere, Broadman area 46) [9]. There is also previous evidence of increased in expression in human blood [10], as well as human genetic evidence [11, 12]. PPP3CB in our studies modestly predicts severe hallucinations state in all patients in the independent testing cohort (AUC 60%, $p = 0.045$), with results being somewhat better in men with schizoaffective disorders (AUC 69%, $p = 0.0044$). It also predicts future hospitalizations with hallucinations in all, in the first year (AUC 64%, $p = 0.015$), and in future years (OR 1.2, $p = 0.0056$).

Among the best individual biomarkers was RTN4 (reticulon 4). It had an AUC of 64% ($p = 0.007$) for predicting high hallucinations state in all. RTN4 also had an AUC of 76% ($p = 0.03$) in males with PTSD, and an odds ratio of 2.83 ($p = 0.0003$) for predicting future hospitalizations with psychosis in males with PTSD, suggestive of a stress component. RTN4 is a potent neurite outgrowth inhibitor, consistent with decreased connectivity in psychosis.

Top 10 biomarkers for delusions. The top 10 blood biomarkers with the strongest overall CFE for tracking and predicting hallucinations, after all four steps (Table 1) were, in order of CFE4 score: AUTS2 (Activator Of Transcription And Developmental Regulator AUTS2), MACROD2 (Mono-ADP Ribosylhydrolase 2), NR4A2 (Nuclear Receptor Subfamily 4 Group A Member 2),

PDE4D (Phosphodiesterase 4D), PDP1 (Pyruvate Dehydrogenase Phosphatase Catalytic Subunit 1), RORA (RAR Related Orphan Receptor A), CHD9 (Chromodomain Helicase DNA Binding Protein 9), FOXP1 (Forkhead Box P1), GNAS (GNAS Complex Locus), and ZBTB20 (Zinc Finger And BTB Domain Containing 20).

AUTS2, the overall top biomarkers for delusions in this study, is a component of a Polycomb group (PcG) multiprotein PRC1-like complex, a complex class required to maintain the transcriptionally repressive state of many genes, including Hox genes, This gene has been implicated in neurodevelopment and as a candidate gene for numerous neurological disorders, including ASDs, intellectual disability, and developmental delay. AUTS2, increased in expression in blood in high delusions in our work, has previous, convergent evidence for involvement in schizophrenia. It is hypomethylated in the brain in schizophrenia [13], as well as increased in expression in fibroblasts from schizophrenia patients [14]. AUTS2 also has previous independent evidence of genetic association with schizophrenia [15–17]. It plays a role in axon and dendrite elongation and in neuronal migration during embryonic brain development, consistent with schizophrenia being a neurodevelopmental disorder. Of note, AUTS2 is also increased in expression by cannabis [18], in the same direction as high delusions, and its expression may be decreased by lithium [19] and valproate [20].

Other top biomarkers for delusions include TCF4 (which was genome-wide significant in previous GWAS studies and serves as a de facto positive control), and DISC1, a top candidate biomarker as well, decreased in expression in high delusions states (Supplementary Information-Additional Data).

Therapeutics

Pharmacogenomics. Overall, based on number of the top 10 biomarkers modulated in expression in opposite direction to hallucinations or delusions, clozapine (50% match for hallucinations, 30% match for delusions) had the best evidence for efficacy in psychotic disorders (Table 3A, B), followed closely by lithium (20% for hallucinations, and 50% for delusions). If we look at the longer list of candidate biomarkers ($n = 98$ for hallucinations, and $n = 70$ for delusions), then clozapine is the best overall match (32.1% for hallucinations, 37.5% for delusions) (Table 3C, D), followed by lithium (26.2% for hallucinations, 34.4% for delusions). Clozapine is known to be a broad-spectrum, gold standard antipsychotic [21]. The evidence pointing at lithium is consistent with some previous older clinical studies in the field [22, 23], and interestingly, may be used for augmenting clozapine treatment [24]. In general, mood stabilizers such as lithium and valproate (also a top match, at 17.9%) are useful adjunctive agents in schizophrenia [25]. Another alternative treatment that was a top match was omega-3 fatty acids, at 28.6%. This may be a widely deployable preventive treatment, with minimal side-effects, including in women who are or may become pregnant.

A number of individual top biomarkers are known to be modulated by medications in current clinical use for treating schizophrenia such as by clozapine (PPP3CB, DLG1, ZEB2, ZNF24, DST, PDE4B for hallucinations, MACROD2, PDE4D, RORA, CHD9, FOXP1, GNAS for delusions), olanzapine (FNBP1 for hallucinations, PDE4D for delusions), risperidone (RORA, GNAS, ZBTB20 for delusions), as well as the nutraceutical omega-3 fatty acids(RTN4, ZEB2, DST for hallucinations, PDP1, FOXP1 for delusions). This is of potential utility in pharmacogenomics approaches matching schizophrenia patients to the right medications, and monitoring response to treatment.

DISCUSSION

We describe a novel and comprehensive effort to discover and validate blood biomarkers of relevance to psychosis, including testing them in independent cohorts to evaluate predictive ability and clinical utility. These biomarkers also open a window into understanding the biology of schizophrenia and related disorders, as well as indicate new and more precise therapeutic approaches.

Current clinical practice and the need for biomarkers

Assessing a persons' internal subjective sensory perceptions and thoughts, along with more objective external ratings of actions and behaviors, are used in clinical practice to assess psychosis and diagnose clinical psychotic disorders, such as schizophrenia and schizoaffective disorder. Such an approach is insufficient, and lagging those used in other medical specialties. Moreover, there is a delay between illness onset and proper specialty diagnosis and treatment, which can lead to disease progression and chronicity [26]. Blood biomarkers related to psychosis, if used as part of routine primary care annual exams, would provide a critical

objective measurement to inform clinical assessments and treatment decisions.

Advantages of biomarkers

Blood biomarkers offer real-world clinical practice advantages. As the brain cannot be readily biopsied in live individuals, and CSF is less easily accessible than blood, we have endeavored over the years to identify blood biomarkers for neuropsychiatric disorders. A whole –blood approach facilitates field deployment of sample collection. The assessment of gene expression changes focuses our approach on immune cells. The ability to identify peripheral gene expression changes that reflect brain activities is likely due to the fact that the brain and immune system have developmental commonalities, marked by shared reactivity and ensuing gene expression patterns. There is also a bi-directional interaction between the brain and immune system. Not all changes in expression in peripheral cells are reflective of or germane to brain activity. By carefully tracking a phenotype with our within-subject design in the discovery step, and then using CFGs prioritization, we are able to extract the peripheral changes that do track and are relevant to the brain activity studied, in this case hallucinations and delusions.

Subsequent validation and testing in independent cohorts narrow the list to the best markers. In the end, we do not expect to recapitulate in the blood all that happens in the brain. We just want to have good accessible peripheral biomarkers- "liquid biopsies", as they are called in cancer.

Comprehensiveness

In this current work, we carried out extensive blood gene expression studies in male and female subjects with major psychiatric disorders, an enriched population in terms of co-morbidity with psychotic disorders. In fact, besides their primary clinical diagnosis, overall, 13% of the subjects in our study that had diagnoses other than schizophrenia and related disorders, had a co-morbid clinical psychotic disorders diagnosis, the highest percentage (18%) being those with bipolar disorder as their primary diagnosis, followed by depression at 12% (Table S1C). The potential molecular-level co-morbidity between other psychiatric disorders and psychosis is underlined by the fact that medications for mood disorders are also used to treat schizophrenia and schizoaffective disorders. Our primary goal was to discover and validate biomarkers for psychosis, that are transdiagnostic. Secondarily, we aimed to understand their universality vs. their specificity by gender, and even by psychiatric diagnosis.

Our studies were arranged in a stepwise fashion. First, we endeavored to discover blood gene expression biomarkers for psychosis using a longitudinal design, looking at differential expression of genes in the blood of male and female subjects with major psychiatric disorders (bipolar disorder, major depressive disorder, schizophrenia/schizoaffective, and post-traumatic stress disorder (PTSD)), high risk populations prone to psychosis, which constitute and enriched pool in which to look for biomarkers. We compared no psychosis states to high psychosis states using a powerful within-subject design [2–4, 27], to generate a list of differentially expressed genes. Second, we used a comprehensive CFG approach with the whole body of knowledge in the field to prioritize from the list of differentially expressed genes/biomarkers of relevance to psychosis. CFG integrates multiple independent lines of evidence- genetic, gene expression, and protein data, from brain and periphery, from human and animal model studies, as a Bayesian strategy for identifying and prioritizing findings, reducing the false-positives and false-negatives inherent in each individual approach. Third, we examined if the expression levels of the top biomarkers identified by us as tracking psychosis state are changed even more strongly in blood samples from an independent cohort of subjects with clinically severe psychosis, to validate these biomarkers. Fourth, the biomarkers thus discovered, prioritized, and validated were tested in corresponding independent cohorts of psychiatric subjects. Fifth, we used the biomarkers to match to existing psychiatric medications, as well as to identify and potentially repurpose drugs for psychotic disorders treatment using bioinformatics analyses. The series of studies was a systematic and comprehensive approach to move the field forward towards precision medicine.

Power

We used a systematic discovery, prioritization, validation, and testing approach, as we have done over the years for other disorders [5, 28–32]. For discovery, we used a hard to accomplish but powerful within-subject design, with an N of 25 subjects with 65 visits for hallucinations, and 31 subjects with 95 visits for delusions. A within-subject design factors out genetic variability, as well as some medications, lifestyle, and demographic effects on gene expression, permitting identification of relevant signal with Ns as small as 1 [27]. Another benefit of a within-subject design may be accuracy/consistency of psychiatric symptoms ("phene expression"), as it is the same person reporting different states. This is similar in rationale to the signal detection benefits it provides in gene expression.

Based on our work of over two decades in genetics and gene expression, along with the results of others in the field, we estimate that using a quantitative phenotype is up to 1 order of magnitude more powerful than using a categorical diagnosis. The within-subject longitudinal design, by factoring out all genetic and some environmental variability, is up to 3 orders of magnitude more powerful than an inter-subject case-control cross-sectional design. Moreover, gene expression, by integrating the effects of many SNPs and environment, is up to 3 orders of magnitude more powerful than a genetic study. Combined, our approach may be up to 6 orders of magnitude more powerful than a GWAS study, even prior to the CFG literature-based prioritization step, which encompasses all the independent work in the field prior to our studies, which may add up to 1 order of magnitude as well. In addition, the Validation and the Testing steps add additional 1 order of magnitude power each. As such, our approach might be up to 10 orders of magnitude more powered to detect signal than most current genetic study designs as used in GWAS.

Reproducibility

We reproduced and expanded our earlier findings in an animal model (Le-Niculescu et al.) [33] of PP3CB, RTN4, ZEB2, NR4A2, and RORA as top genes involved in psychosis. We also reproduced top blood biomarkers from an early pilot study of ours- FN1 for hallucinations, IQCH and KLK2 for delusions (Le-Niculescu, Kurian et al.) [34].

Additionally, there is reproducibility with findings generated by other independent large scale studies that came out after our analyses were completed, and were thus not included in our CFG approach (see Supplementary Information-Additional Data). A number of their top findings were present in our candidate gene expression biomarkers for hallucinations and delusions lists that had survived our initial whole-genome, unbiased, within-subject Discovery step, before any CFG literature prioritization: 48 for hallucinations and 42 for delusions out of their 120 top genes in a recent large GWAS of schizophrenia [35], and 29 each for hallucinations and delusions out of 63 top genes from another recent GWAS of schizophrenia and vitamin D levels [36]. This independent reproducibility of findings between our studies and these other genetic studies, which are done in independent cohorts from ours, with independent methodologies, is reassuring, and provides strong convergent evidence for the validity and relevance of our approach and of their approaches. Our work also provides functional evidence for some of their top genetic hits.

Pathophysiology

A number of top candidate biomarkers identified by us overlap with genes implicated in autism spectrum disorders (ASD) (Table 1A, B). A top biological pathway for hallucinations biomarkers is glutamatergic signaling. Glutamatergic signaling is excitatory, and under its influence the brain may be over-responding to sensory information, or filling in gaps in information [37]. A top pathway for delusions biomarkers is Rap1 signaling pathway. Rap1 is involved in the formation of synapses, and changes in Rap1 activity may lead to changes in the way that synapses are formed, which can contribute to the development of delusions [38].

The majority of top blood biomarkers we have identified have prior evidence in human or animal model brain data from schizophrenia studies, which indicates their relevance to the pathophysiology of psychotic disorders (Table S2). The co-directionality of blood changes in our work and brain changes reported in the literature needs to be interpreted with caution, as it may depend on brain region.

The top candidate biomarkers also had prior evidence of involvement in other psychiatric and related disorders (Tables 2 and S3), providing a molecular basis for co-morbidity, and the possible precursor effects of some these disorders on psychosis, and conversely, the precursor role of psychosis in some of them. In particular, over 90% of them have an overlap with genes involved in alcoholism and in depression, followed to a lesser extent by stress, dementia, suicidality, pain, bipolar disorder, consistent with psychosis being a common and often under-treated and under-appreciated factor in most mental health disorders. In particular, 7 out of the top 10 biomarkers for delusions have evidence in suicidality (Table 1B), which points to potential delusional aspects of suicidality, and the possibility of using antipsychotics to resolve suicidality states.

Phenomenology

In addition to using the standard PANSS scale, we developed two scales to assess trait, respectively state, psychosis.

The trait Convergent Functional Information for Schizophrenia (CFI-SZ) scale (Fig. S1C, E) is a 10-items checklist of past disease severity and social impairment. The CFI-SZ correlates moderately with the PANSS Total ($R = 0.307$, $p = 4.00E-15$).

The state Simplified Psychosis Scale (SPS-4) consists of 4 self-report visual analog scale (VAS) items (Fig. S1D). The overall SPS-4 score correlates moderately with the PANSS Total ($p = 0.339$, $p = 3.04E-3$). The Hallucinations item in it correlates moderately with the PANSS P3 Hallucinations item ($R = 0.441$, $p = 8.11E-12$), whereas the Delusions item in SPS-4 correlates less well with the PANSS P1 Delusions item ($R = 0.154$, $p = 0.02$) (Fig. S1E), reflecting the fact that delusions may be more difficult for an external rater to evaluate, as is the case with the PANSS.

Biomarkers vs. Scales

In general, the best predictive biomarkers were better than the standard or new rating scales at predicting state and trait hallucinations or delusions. This may reflect the fact that these are difficult phenotypes to assess by clinicians, and reinforces the need for using objective blood biomarkers to assess psychotic disorders (Table S5).

Diagnostics

For the biomarkers identified by us, combining all the available evidence from this current work into a CFE score, brings to the fore biomarkers that have clinical utility for objective assessment and risk prediction for psychotic disorders (Table 1). These biomarkers should be tested individually as well as tested as polygenic panels of biomarkers in future clinical studies and practical clinical applications in the field. They may permit to distinguish, upon an initial clinical presentation of psychosis, whether the person is in fact severely psychotic and at chronic risk (Fig. 3). The integration of phenomic data, such as repeated measures of SPS-4 (perhaps via a phone app in a daily fashion), can further substantiate and elucidate the diagnosis of a psychotic disorder, distinguishing between an intermittent type such as transient psychosis, and continuous type such as schizophrenia.

In general, our predictive results with biomarkers were stronger in females than in males, by an order of 10–20% points on AUCs. While some of it may be biological, in terms of immune system reactivity and brain-blood interplay being perhaps higher in women, it is also possible that men are not as accurate as women in terms of reporting psychosis symptoms (affecting our results on state predictions), and do not seek help as much (affecting our results on future hospitalizations predictions). If so, this under-reporting makes the use of objective biomarker tests in men even more necessary.

In regard to how our biomarker discoveries might be applied in clinical laboratory settings, we suggest that panels of top biomarkers for hallucinations and delusions be used (Fig. 3). In practice, every new patient tested would be normalized against the database of similar patients already tested, and compared to them for ranking and risk prediction purposes, regardless if a platform like microarrays, RNA sequencing, or a more targeted one like PCR is used in the end clinically. As databases get larger, normative population levels can and should be established, similar to any other laboratory measures. Moreover, longitudinal monitoring of changes in biomarkers within an individual, measuring most recent slope of change, maximum levels attained, and maximum slope of change attained in the past, may be even more informative than simple cross-sectional comparisons of levels within an individual with normative populational levels, as we have shown in our studies. For future point of care approaches, research and development should focus on top individual biomarkers, including at a protein level. One might look at a combination of the best universal biomarkers (that are predictive in all), for reliability, and of the best personalized biomarkers (that are predictive by gender, and even diagnosis), for higher accuracy.

Treatment

Biomarkers may also be useful for matching patients to medications and measuring response to treatment (pharmacogenomics) (Fig. 3, Tables 3 and S4), as well as new drug discovery clinical trials, and drug repositioning (Table 3). From the pharmacogenomics analyses, lithium was a top hit, second only to clozapine, the gold-standard antipsychotic. Other interesting novel candidates were omega-3 fatty acids, fluoxetine, valproate, and magnesium. All these drugs and nutraceuticals are relatively safe if used appropriately, and have been used in clinical practice for other indications for decades, which facilitates the direct translation to clinical practice of our findings.

CONCLUSIONS

Overall, this work is a major step forward towards better understanding, diagnosing, and treating psychotic disorders. We hope that our trait biomarkers for future risk may be useful in preventive approaches, before the full-blown disorder manifests itself (or re-occurs). Prevention could be accomplished with biological interventions (i.e., early targeted use of medications or nutraceuticals), social measures to help with integration in society, and psychological support. Given the fact that the annual prevalence of diagnosed schizophrenia in the US is over 1 in 200 people [39], that the prevalence seems to be increasing worldwide [40], that psychotic disorders can severely affect quality of life and lead to shortened lifespan, and that not all patients respond to current treatments, the need for and importance of efforts such as ours cannot be overstated.

DATA AVAILABILITY

Additional data are available in Supplementary Information and upon request to the corresponding author.

REFERENCES

1. Le-Niculescu H, Niculescu AB. Precision medicine in psychiatry: biomarkers to the forefront. Neuropsychopharmacology. 2022;47:422–3.
2. Le-Niculescu H, Levey DF, Ayalew M, Palmer L, Gavrin LM, Jain N, et al. Discovery and validation of blood biomarkers for suicidality. Mol Psychiatry. 2013;18:1249–64.
3. Niculescu AB, Levey DF, Phalen PL, Le-Niculescu H, Dainton HD, Jain N, et al. Understanding and predicting suicidality using a combined genomic and clinical risk assessment approach. Mol psychiatry. 2015;20:1266–85.
4. Levey DF, Niculescu EM, Le-Niculescu H, Dainton HL, Phalen PL, Ladd TB, et al. Towards understanding and predicting suicidality in women: biomarkers and clinical risk assessment. Mol Psychiatry. 2016;21:768–85.
5. Niculescu AB, Le-Niculescu H, Levey DF, Phalen PL, Dainton HL, Roseberry K, et al. Precision medicine for suicidality: from universality to subtypes and personalization. Mol Psychiatry. 2017;22:1250–73.
6. Kilka S, Erdmann F, Migdoll A, Fischer G, Weiwad M. The proline-rich N-terminal sequence of calcineurin Abeta determines substrate binding. Biochemistry. 2009;48:1900–10.
7. Li SJ, Wang J, Ma L, Lu C, Wang J, Wu JW, et al. Cooperative autoinhibition and multi-level activation mechanisms of calcineurin. Cell Res. 2016;26:336–49.
8. Parnell E, Culotta L, Forrest MP, Jalloul HA, Eckman BL, Loizzo DD, et al. Excitatory dysfunction drives network and calcium handling deficits in 16p11.2 duplication schizophrenia induced pluripotent stem cell-derived neurons. Biol Psychiatry. 2023;94:153–63.
9. Hakak Y, Walker JR, Li C, Wong WH, Davis KL, Buxbaum JD, et al. Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. Proc Natl Acad Sci USA. 2001;98:4746–51.
10. Leirer DJ, Iyegbe CO, Di Forti M, Patel H, Carra E, Fraietta S, et al. Differential gene expression analysis in blood of first episode psychosis patients. Schizophr Res. 2019;209:88–97.
11. Liu CM, Fann CS, Chen CY, Liu YL, Oyang YJ, Yang WC, et al. ANXA7, PPP3CB, DNAJC9, and ZMYND17 genes at chromosome 10q22 associated with the subgroup of schizophrenia with deficits in attention and executive function. Biol Psychiatry. 2011;70:51–8.
12. Forero DA, Herteleer L, De Zutter S, Norrback KF, Nilsson LG, Adolfsson R, et al. A network of synaptic genes associated with schizophrenia and bipolar disorder. Schizophr Res. 2016;172:68–74.
13. Jaffe AE, Gao Y, Deep-Soboslay A, Tao R, Hyde TM, Weinberger DR, et al. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. Nat Neurosci. 2016;19:40–7.
14. Brennand KJ, Simone A, Jou J, Gelboin-Burkhart C, Tran N, Sangar S, et al. Modelling schizophrenia using human induced pluripotent stem cells. Nature. 2011;473:221–5.
15. Mozhui K, Wang X, Chen J, Mulligan MK, Li Z, Ingles J, et al. Genetic regulation of Nrxn1 [corrected] expression: an integrative cross-species analysis of schizophrenia candidate genes. Transl Psychiatry. 2011;1:e25.
16. McCarthy SE, Gillis J, Kramer M, Lihm J, Yoon S, Berstein Y, et al. De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. Mol Psychiatry. 2014;19:652–8.
17. Ozsoy F, Karakus NB, Yigit S, Kulu M. Effect of AUTS2 gene rs6943555 variant in male patients with schizophrenia in a Turkish population. Gene. 2020;756:144913.
18. Oksenberg N, Ahituv N. The role of AUTS2 in neurodevelopment and human evolution. Trends Genet. 2013;29:600–8.
19. Pisanu C, Merkouri Papadima E, Melis C, Congiu D, Loizedda A, Orru N et al. Whole genome expression analyses of miRNAs and mRNAs suggest the involvement of miR-320a and miR-155-3p and their targeted genes in lithium response in bipolar disorder. Int J Mol Sci. 2019;20:6040.
20. Schulpen SH, Pennings JL, Piersma AH. Gene expression regulation and pathway analysis after valproic acid and carbamazepine exposure in a human embryonic stem cell-based neurodevelopmental toxicity assay. Toxicol Sci. 2015;146:311–20.
21. Stroup TS, Gerhard T, Crystal S, Huang C, Olfson M. Comparative effectiveness of clozapine and standard antipsychotic treatment in adults with schizophrenia. Am J Psychiatry. 2016;173:166–73.
22. Alexander PE, van Kammen DP, Bunney WE Jr. Antipsychotic effects of lithium in schizophrenia. Am J Psychiatry. 1979;136:283–7.
23. Zemlan FP, Hirschowitz J, Sautter FJ, Garver DL. Impact of lithium therapy on core psychotic symptoms of schizophrenia. Br J Psychiatry. 1984;144:64–9.
24. Moldavsky M, Stein D, Benatov R, Sirota P, Elizur A, Matzner Y, et al. Combined clozapine-lithium treatment for schizophrenia and schizoaffective disorder. Eur Psychiatry. 1998;13:104–6.
25. Puranen A, Koponen M, Lahteenvuo M, Tanskanen A, Tiihonen J, Taipale H. Real-world effectiveness of mood stabilizer use in schizophrenia. Acta Psychiatr Scand. 2023;147:257–66.
26. Lieberman JA, Fenton WS. Delayed detection of psychosis: causes, consequences, and effect on public health. Am J Psychiatry. 2000;157:1727–30.
27. Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, Miriami E, et al. Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell. 2012;148:1293–307.
28. Niculescu AB, Le-Niculescu H, Levey DF, Roseberry K, Soe KC, Rogers J, et al. Towards precision medicine for pain: diagnostic biomarkers and repurposed drugs. Mol Psychiatry. 2019;24:501–22.
29. Le-Niculescu H, Roseberry K, Levey DF, Rogers J, Kosary K, Prabha S et al. Towards precision medicine for stress disorders: diagnostic biomarkers and targeted drugs. Mol Psychiatry. 2020;25:918–38.
30. Niculescu AB, Le-Niculescu H, Roseberry K, Wang S, Hart J, Kaur A et al. Blood biomarkers for memory: toward early detection of risk for Alzheimer disease, pharmacogenomics, and repurposed drugs. Mol Psychiatry. 2020;25:1651–72.
31. Le-Niculescu H, Roseberry K, Gill SS, Levey DF, Phalen PL, Mullen J, et al. Precision medicine for mood disorders: objective assessment, risk prediction, pharmacogenomics, and repurposed drugs. Mol Psychiatry. 2021;26:2776–804.
32. Roseberry K, Le-Niculescu H, Levey DF, Bhagar R, Soe K, Rogers J, et al. Towards precision medicine for anxiety disorders: objective assessment, risk prediction, pharmacogenomics, and repurposed drugs. Mol Psychiatry. 2023;28:2894–912.
33. Le-Niculescu H, Balaraman Y, Patel S, Tan J, Sidhu K, Jerome RE, et al. Towards understanding the schizophrenia code: an expanded convergent functional genomics approach. Am J Med Genet B Neuropsychiatr Genet. 2007;144B:129–58.
34. Le-Niculescu H, Kurian SM, Yehyawi N, Dike C, Patel SD, Edenberg HJ, et al. Identifying blood biomarkers for mood disorders using convergent functional genomics. Mol Psychiatry. 2009;14:156–74.
35. Trubetskoy V, Pardinas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature. 2022;604:502–8.
36. Jaholkowski P, Hindley GFL, Shadrin AA, Tesfaye M, Bahrami S, Nerhus M et al. Genome-wide Association Analysis of Schizophrenia and Vitamin D Levels Shows Shared Genetic Architecture and Identifies Novel Risk Loci. Schizophr Bull. 2023;49:1654–64.
37. Hjelmervik H, Craven AR, Johnsen E, Kompus K, Bless JJ, Sinkeviciute I, et al. Negative valence of hallucinatory voices as predictor of cortical glutamatergic metabolite levels in schizophrenia patients. Brain Behav. 2022;12:e2446.
38. Zhao XF, Kohen R, Parent R, Duan Y, Fisher GL, Korn MJ, et al. PlexinA2 forward signaling through Rap1 GTPases regulates dentate gyrus development and schizophrenia-like behaviors. Cell Rep. 2018;22:456–70.
39. Wu EQ, Shi L, Birnbaum H, Hudson T, Kessler R. Annual prevalence of diagnosed schizophrenia in the USA: a claims data analysis approach. Psychol Med. 2006;36:1535–40.
40. He H, Liu Q, Li N, Guo L, Gao F, Bai L, et al. Trends in the incidence and DALYs of schizophrenia at the global, regional and national levels: results from the Global Burden of Disease Study 2017. Epidemiol Psychiatr Sci. 2020;29:e91.

ACKNOWLEDGEMENTS

We would like to acknowledge our gratitude for the work and results of the many other groups, cited in our paper, who have conducted and published studies (clinical, genetic, and biological) in schizophrenia and related disorders. Combining their work with ours makes a convergent approach possible. We also would particularly like to thank the subjects in these studies and their families. Without their contribution, such work to advance the understanding of psychotic disorders would not be possible. This work was supported by NIH grants (R01MH117431) and a VA Merit Award (2I01CX000139) to ABN.

AUTHOR CONTRIBUTIONS

ABN designed the study and wrote the manuscript. MH, SSG, HLN, OM, RB, KR, OKM, HDD, SKW analyzed the data. MH and OM organized, conducted, and scored testing in psychiatric subjects. SSG assisted with sample report generation. AS assisted with data interpretation. SMK conducted microarray experiments. All authors discussed the results and commented on the manuscript.

COMPETING INTERESTS

ABN is listed as inventor on a patent application filed by Indiana University. ABN and AS are co-founders, SMK is a consultant, and SSG is a part-time employee of MindX Sciences.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41380-024-02433-8.

Correspondence and requests for materials should be addressed to A. B. Niculescu.