

## ORIGINAL ARTICLE OPEN ACCESS

# Identification of Schizophrenia-Risk Regulatory Variant rs1399178 in the Non-coding Region With Its Impact on NRF1 Binding

Lei Ji<sup>1,2,3</sup> | Bojian Huang<sup>1,2,3</sup> | Decheng Ren<sup>1,2,3</sup> | Xiaoxi Wei<sup>1,2,3</sup> | Liangjie Liu<sup>1,2,3</sup> | Yan Bi<sup>1,2,3</sup> | Zhiqiang Li<sup>1,2,3,4</sup> | Fan Yuan<sup>1,2,3</sup> | Ke Han<sup>1,2,3</sup> | Keyi Li<sup>1,2,3</sup> | Fengping Yang<sup>1,2,3</sup> | Xingwang Li<sup>1,2,3</sup> | Tao Yu<sup>1,2,3</sup> | Yi Shi<sup>1,2,3</sup> | Lin He<sup>1,2,3</sup> | Qing Lu<sup>1,5,6,7</sup> | Guang He<sup>1,2,3</sup>

<sup>1</sup>Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, Shanghai Jiao Tong University, Shanghai, China | <sup>2</sup>The Collaborative Innovation Center for Brain Science, and Brain Science and Technology Research Center, Shanghai Jiao Tong University, Shanghai, China | <sup>3</sup>Shanghai Key Laboratory of Psychotic Disorders, Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China | <sup>4</sup>Affiliated Hospital of Qingdao University and Biomedical Sciences Institute of Qingdao University (Qingdao Branch of SJTU Bio-X Institutes), Qingdao University, Qingdao, China | <sup>5</sup>Department of Otorhinolaryngology-Head and Neck Surgery, Chongqing General Hospital, Chongqing, China | <sup>6</sup>Ear Institute, Shanghai Jiao Tong University School of Medicine, Shanghai, China | <sup>7</sup>Shanghai Key Laboratory of Translational Medicine on Ear and Nose Diseases, Shanghai, China

**Correspondence:** Qing Lu ([luqing67@sjtu.edu.cn](mailto:luqing67@sjtu.edu.cn)) | Guang He ([heguang@sjtu.edu.cn](mailto:heguang@sjtu.edu.cn))

**Received:** 19 September 2024 | **Revised:** 28 November 2024 | **Accepted:** 31 January 2025

**Funding:** This project is supported by the National Key Research and Development Program (2024YFC2707002, 2022YFE0125300), Innovation Program of Shanghai Municipal Education Commission (2023ZKZD16), National Natural Science Foundation of China (82071262, 32300464, 82301682), Shanghai Municipal Science and Technology Major Project (2017SHZDZX01, 20JC1418600), Natural Science Foundation of Shanghai (23ZR1432700), China Postdoctoral Science Foundation (2023M732266), the Shanghai Leading Academic Discipline Project (B205), Key Technology Breakthrough Program of Ningbo Sci-Tech Innovation YONGJIANG 2035 (2024Z221), Municipal Public Welfare Project (2022S035) and Shanghai Jiao Tong University STAR Grant (YG2023ZD26, YG2022ZD024, YG2022QN111, YG2023LC14, YG2024QNA59).

**Keywords:** functional genomics | *NRF1* | rs1399178 | schizophrenia | structural analysis | TF-binding regulation

## ABSTRACT

**Aims:** The challenges in identifying functional variants from genome-wide association studies (GWAS) and unraveling regulatory mechanisms in schizophrenia research persist, particularly in intronic regions. A non-coding regulatory variant, rs1399178, associated with schizophrenia risk, is identified and its impact on NRF1 binding is investigated.

**Methods:** This study focuses on schizophrenia GWAS risk loci, using functional genomics, expression analyses and structural analysis to identify 736 schizophrenia risk single-nucleotide polymorphisms (SNPs) that disrupt transcription factor (TF) binding.

**Results:** Among these SNPs, rs1399178 stands out as a bifunctional intergenic SNP that can switch between acting as a promoter and an enhancer, potentially influencing *MLH1* and *LRRFIP2* expression via expression quantitative trait loci analysis. Importantly, mutation of the G allele of rs1399178 to A significantly diminishes its binding affinity to nuclear respiratory factor 1 (NRF1). Structural analysis provides further insight into this alteration in the protein–nucleic acid complex formation.

**Conclusion:** Based on our data, a model is proposed in which rs1399178 confers schizophrenia risk by modifying *NRF1* binding profiles, thereby regulating the abundance of target genes through promoter-enhancer switching. This study provides novel insights into the regulatory mechanisms of schizophrenia risk variants, highlighting the intricate nature of genetic interactions and potential therapeutic targets.

Lei Ji and Bojian Huang contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *CNS Neuroscience & Therapeutics* published by John Wiley & Sons Ltd.

## 1 | Introduction

Schizophrenia, considered one of the most severe mental disorders, is a chronic psychiatric condition with a diverse genetic and neurobiological foundation that impacts early brain development [1]. Schizophrenia is usually characterized as a combination of psychotic symptoms, including delusions, hallucinations, disordered thought, and reduced emotional expression [2, 3]. According to global estimates from the World Health Organization (WHO), ~24 million individuals (0.32% of the global population) are affected by schizophrenia, equating to approximately one in every 300 people. The global lifetime prevalence of schizophrenia is ~1% [4]. Although its prevalence is comparatively lower than other psychiatric disorders, such as major depression or bipolar disorder, schizophrenia is notable for having one of the highest global rates of disability. In fact, schizophrenia leads to significant economic costs and places a substantial mental burden on patients, their families, and society as a whole, collectively contributing to a diminished quality of life [5, 6].

Despite significant progress in schizophrenia research over the past decade, a comprehensive understanding of its pathogenesis remains elusive. Genomic studies have unveiled the intricate genetic architecture of schizophrenia, with recent emphasis on identifying rare and common risk alleles to deepen insights into its etiology. Notably, over 200 risk loci have been identified across diverse populations [7–13]. Lee et al.'s research emphasizes the significance of single-nucleotide polymorphisms (SNPs), contributing to ~23% of the heritability of schizophrenia. This underscores the critical role of common variants in the disorder [14]. However, challenges persist in identifying functional or causal variants at reported loci and elucidating their regulatory mechanisms, mainly owing to the complexities of linkage disequilibrium (LD) and gene regulation [15, 16]. Individual risk alleles may govern multiple genes through intricate chromatin interactions, resulting in effects on genes located distally in the linear genome [17, 18]. On the other hand, the majority of risk loci identified through genome-wide association studies (GWAS) are located in non-coding regions [13, 19]. Bacanu et al. found that variations associated with schizophrenia exhibit significant enrichment in regulatory regions, suggesting that schizophrenia may exert its influence by disrupting regulatory function and altering gene expression [20, 21]. Luo et al. used functional genomics approaches to identify 132 schizophrenia risk variants influencing transcription factor (TF) binding [22]. Considering the presumed neurobiological origin of schizophrenia, research efforts focused on identifying risk variants regulating human brain tissue or neuronal cells [23–25]. Despite the existence of up to 1600 TFs in the human genome, Luo et al. deliberately included a limited number of TFs (34) in their study [22, 26]. However, this restriction poses challenges to discerning risk SNP loci that disrupt the binding of other TFs [22]. The proliferation of sequencing data in recent years, along with the progress of additional databases, has significantly expanded the available information during this five-year period. Leveraging this expanded and novel data pool is poised to enhance our understanding of the regulatory mechanisms underlying schizophrenia.

Here, we systematically identify functional variants at schizophrenia risk loci derived from GWAS, and explore how these variants exert their biological effects and regulatory mechanisms in the context of schizophrenia [11]. Leveraging functional genomics and expression quantitative trait loci (eQTL) analyses, coupled with a series of functional validations, we annotated SNPs from GWAS, pinpointing schizophrenia regulatory loci. Based on the results of the reporter gene assay, we selected the rs1399178 locus for functional genomics validation, which indicated this variant disrupts the binding of NRF1. Structural evidence elucidating the molecular mechanisms behind this recognition alteration further underscored the significance of our results. Our findings highlight the intricate genetic regulatory structure of schizophrenia risk variants, providing novel insights into target genes.

## 2 | Methods

### 2.1 | Functional Genomics Pipelines Used to Identify Risk SNPs Affecting TF Binding

#### 2.1.1 | GWAS Used in This Study

GWAS schizophrenia risk SNPs identified by Li et al. were used in this study [11]. Li et al. first carried out a GWAS approach in the Chinese population (containing 7699 cases and 18,327 controls). A total of 124 genome-wide significant ( $p < 5 \times 10^{-8}$ ) SNPs were identified based on the combined results of the trans-ancestry meta-analysis and replication samples.

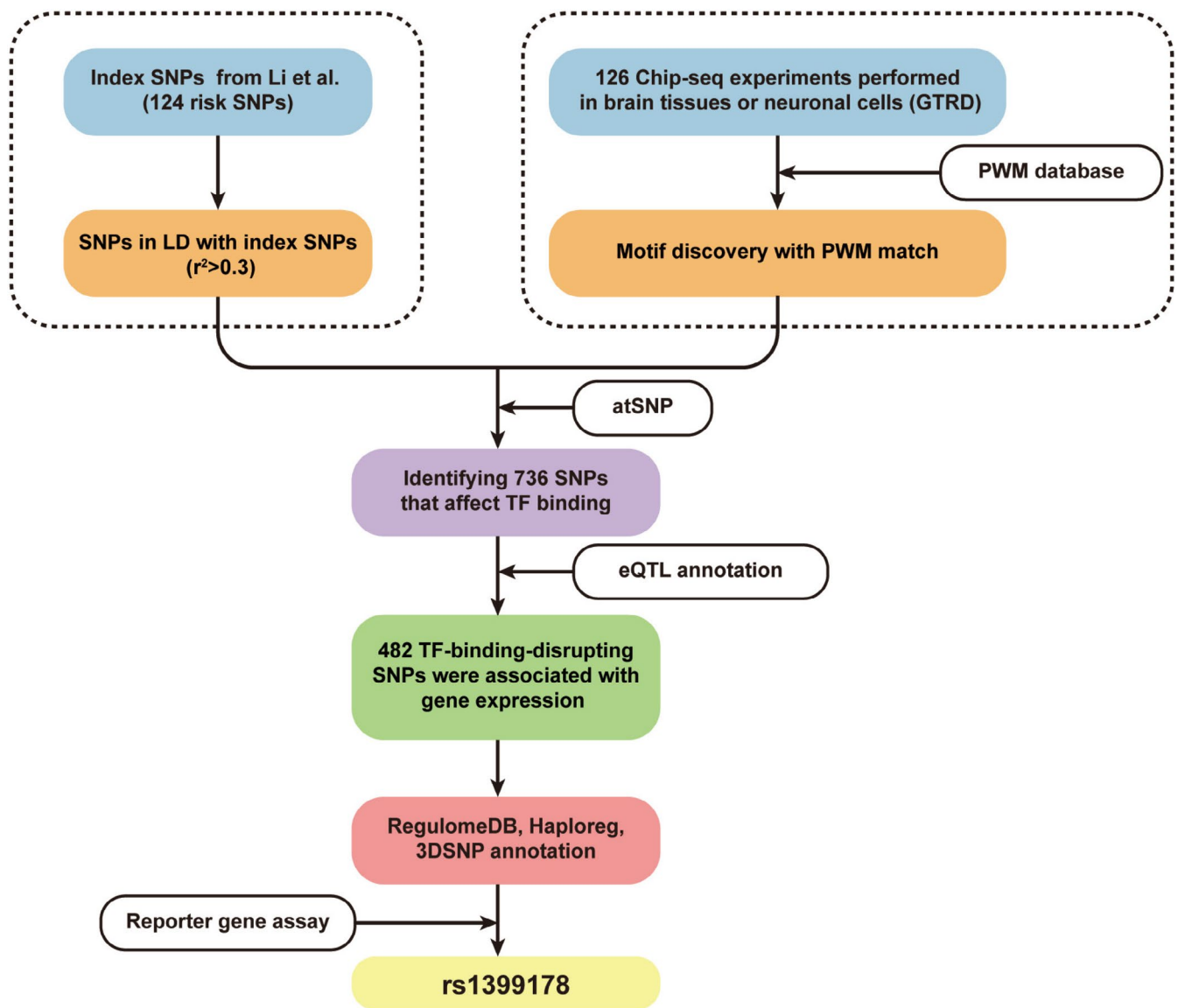
#### 2.1.2 | Extraction of SNPs in LD With the Index SNPs

For each risk locus in the East-Asian individuals from the 1000 Genomes Project (Phase 3\_Version 5) [27], we extracted SNPs that were in LD with the 124 index SNPs using the “LDlinkR” R package. All downstream analyses were performed using R v4.1.3. We selected the widely used  $r^2$  threshold ( $r^2 > 0.3$ ) following previous work [22, 28]. In total, 21,300 SNPs (including the index SNPs and SNPs in LD with the index SNPs) were obtained.

#### 2.1.3 | Functional Genomics Pipelines Used to Identify Risk SNPs That Disrupt TF Binding

The functional genomics flow chart for identifying TF-disrupting SNPs is shown in Figure 1. Briefly, we first utilized 126 chromatin immunoprecipitation (ChIP)-seq data from the Gene Transcription Regulation Database (GTRD, version 21.12), including only ChIP-seq experiments performed in human brain tissues or neuronal cells (including neuronal cell lines and iPSCs-derived neural progenitor cells) [29]. The detailed information about the TFs and ChIP-seq experiments is found in [Supporting Information](#).

To obtain DNA-binding motifs (PWMs) enriched in the genomic sequences surrounding ChIP-seq signals, the top 500 ChIP-seq peaks ( $\pm 20$  bp, ranked by peak height) were used for



**FIGURE 1** | Flowchart of functional genomics analysis. ChIP-seq assays performed using neuronal cell lines and brain tissues were used for peak calling and motif discovery. The identified motifs were then compared with the PWMs from the respective database (see Section 2), and the matched motifs were used for downstream analyses. SNPs in linkage disequilibrium (LD,  $r^2 \geq 0.3$ ) with the index SNPs (identified by SCZ GWAS, a total of 124 index SNPs) were extracted. A total of 23,000 SNPs from 124 SCZ risk loci were mapped to the identified motifs to investigate their association with the disruption of transcription factor binding. In total, 736 TF binding-disrupting SNPs were identified.

motif discovery with runStreme function of the “memes” R package (with parameters “nmotifs=5, minw=6, maxw=20”) [30]. Taking a shuffled set of input sequences as background, peaks with a false discovery rate (FDR) > 5% were excluded. We then examined the binding specificity of the included TFs by comparing the motifs from ChIP-seq with the 6091 human motif sequences available from the “MotifDb” R package (created from public sources containing HOCOMOCO, JASPAR, SwissRegulon and HOMER). The matched PWMs were utilized to investigate whether the potential schizophrenia SNPs were located in the binding motif of these TFs.

To quantify the in silico effect of SNPs on TF binding, we used the “atSNP” R package to calculate P values for each SNP-transcription factor pair (with parameters “half.window.size=20”) [31]. For this analysis, we required a significant

binding likelihood ( $p < 0.001$ ) for at least one of the alleles (reflecting potential TF occupancy of at least one allele), and a significant difference in binding likelihood between the two alleles ( $p$  value\_rank < 0.05 after Bonferroni correction). In total, we identified 736 SNPs that disrupted the binding of TFs (Table S1).

## 2.2 | eQTL Analysis

The eQTL summary data utilized in this analysis were sourced from the eQTL Catalogue (<https://www.ebi.ac.uk/eql/>), a repository housing uniformly standardized eQTL summary data compiled from various published sources [32]. Given that schizophrenia primarily manifests as a brain condition, only eQTL datasets derived from human brain tissues were considered. Specifically, datasets from the Common Mind Consortium

( $N=586$ ), the Genotype-Tissue Expression v8 (encompassing 13 brain regions, with  $N$  ranging from 114 to 209), the Religious Orders Study and Memory and Aging Project ( $N=560$ ), and the Braineac2 database (specifically, putamen with  $N=102$  and substantia nigra with  $N=65$ ) were used. Only SNPs exhibiting  $p$ -values below 0.001 after FDR correction were considered for subsequent analysis.

### 2.3 | Functional Annotation of TF Binding-Disrupting SNPs Using RegulomeDB, Haploreg, and 3DSNP

We used three widely recognized online functional annotation tools to enrich for the most functional SNPs at each risk loci, including RegulomeDB, HaploReg v4.1, and 3DSNP v2.0 [33–35]. All TF binding-disrupting SNPs identified in this study have been scored to evaluate the functional significance of SNPs using RegulomeDB, HaploReg, and 3DSNP, one by one. We also investigated whether the variation in rs1399178 within topologically associating domain (TAD) structures could be associated with gene regulation [17]. Briefly, we collected in situ Hi-C data for the GM12878 (lymphoblastoid cells), IMR90 (fetal lung fibroblasts), and hESC (human embryonic stem cells) cell lines from the 4DN data portal (<https://data.4dnucleome.org/>), and analyzed the TAD distribution with resolutions of 50 kb [36, 37]. Detailed information is provided in [Supporting Information](#).

### 2.4 | Association Analysis and Meta-Analysis

Schizophrenia cases used in this study have been described in previous studies [11]. The association between rs1399178 and schizophrenia, and the meta-analysis, were assessed with PLINK [38].

### 2.5 | Dual Luciferase Reporter Gene Assays

We incorporated these target DNA sequences into the regulatory region, i.e., the promoter or enhancer, of the reporter gene vector. Briefly, the DNA fragments (~300–800 bp) containing the different alleles of target SNPs were cloned into pGL3 promoter (RRID: Addgene\_212939) or pGL4.11[luc2P] vectors. HEK293T and SH-SY5Y cells were transfected with the constructed pGL3 promoter or pGL4.11[luc2P] vectors. The pRL-TK Renilla vector served as the internal control.

### 2.6 | Construction of Expression Plasmids, Protein Expression, and Protein Purification

The full-length *NRF1* gene was PCR amplified from a human cDNA library (reference sequence NCBI: NP\_001035199.1, 503 aa). Shorter fragments of *NRF1* were generated using standard PCR-based methods and confirmed by DNA sequencing. Recombinant proteins were expressed in *Escherichia coli* BL21 (DE3) cells in LB medium at 16°C and purified using a nickel-NTA agarose (for Trx-His6-/GB1-His6-/His6-tagged proteins), followed by size-exclusion chromatography with a column buffer containing 50 mM Tris pH8.0, 100 mM NaCl, 1 mM EDTA,

and 1 mM DTT. When needed, the tags were cleaved using HRV 3C protease and separated by another step of size-exclusion chromatography. The protein purities were evaluated by SDS-PAGE.

### 2.7 | Chromatin Immunoprecipitation-qPCR

ChIP-qPCR was conducted to assess the binding of TFs NRF1 to the genomic sequence encompassing rs1399178 in SH-SY5Y cells. The SimpleChIP Enzymatic Chromatin IP Kit (CST, Cat. No: #9003) was used in accordance with manufacturer's instructions. Three immunoprecipitation (IP) groups were established using IgG (CST, Cat. No: 2729, RRID: AB\_1031062), H3 (CST, Cat. No: 4620, RRID: AB\_1904005), and NRF1 (Abcam, Cat. No: ab175932, RRID: AB\_2629496) antibodies. Cross-linked chromatin fragments (8 µg) were used for the NRF1 antibody as the experimental group, IgG served as the negative control, and H3 as the positive control. DNA purification preceded subsequent qPCR analysis, for which primers were designed to span 100 bp up- and down-stream from rs1399178.

### 2.8 | Electrophoretic Mobility Shift Assay

Nuclear extracts of SH-SY5Y were obtained using the nuclear protein extraction kit (Beyotime, Cat. No: P0028) and quantified using the BCA protein assay kit (Pierce, Cat. No: 23227). Oligonucleotides harboring the G and A alleles of rs1399178 (37 bp) were labeled at their 3' ends with biotin using the EMSA probe biotin labeling kit (Beyotime, Cat. No: GS008). Single-strand oligonucleotides were annealed to form double-stranded complementary sequences. The EMSA procedure was conducted using the Chemiluminescent EMSA Kit (Thermo Scientific, Cat. No: 20148) in accordance with manufacturer's instructions.

### 2.9 | Isothermal Titration Calorimetry Assay

Isothermal titration calorimetry (ITC) assay measurements were performed on a Microcal VP-ITC calorimeter at 25°C. The proteins used for ITC measurements were dissolved in an assay buffer composed of 50 mM Tris pH8.0, 100 mM NaCl, 1 mM EDTA, and 1 mM DTT. High concentration of DNA (500 µM for RS(G) and RS(A)) was individually loaded into the syringe and titrated into the cell containing low concentration of corresponding interactors (50 µM for DD-DBD). For each titration point, a 10 µL aliquot of a protein sample in the syringe was injected into the interacting protein in the cell at a time interval of 2–3 min. The ITC titration data were analyzed using the Origin7.0 software and fitted with the two-site binding model.

### 2.10 | Statistical Analysis

Statistical analyses were performed using R software (version 4.3.2). The details of the R packages, functions, and parameter settings used for the analysis are provided in [Supporting Information](#). The Shapiro–Wilk test was applied to assess data normality. Since all data conformed to a normal distribution, Student's  $t$ -test was used to evaluate the significance of differences in luciferase activity between cells transfected with



vectors containing different alleles of the target SNP. Statistical significance was defined as  $p < 0.05$ . Quantitative data are presented as mean  $\pm$  standard deviation (SD). The exact sample size (n) and the statistical test used are provided in the legend of each figure.

### 3 | Results

#### 3.1 | Functional Genomics Identified 736 Schizophrenia Risk SNPs Disrupting TF Binding

To identify functional SNPs in the risk loci reported by Li et al., we used a series of functional genomics approaches, as illustrated in Figure 1. Briefly, by integrating ChIP-seq experimental data and PWM data from public databases, we identified 736 SNPs that affect TF binding whose P-values were lower than 0.05 after Bonferroni correction (Table S1). The disruptions involve changes that either impair or enhance the binding affinity between SNPs and motifs, leading to a classification as “TF binding-disrupting SNPs” rather than damaging. Notably, the majority are located in intronic or intergenic regions (717/736), suggesting potential regulatory roles in transcription (Figure 2A). Among the 736 TF binding-disrupting SNPs, 111 SNPs interfere with JUN binding (13%), 98 with ZNF770 binding (12.13%), 68 with FOXK1 (7.69%), and 37 with *NRF1* (7.69%) (Figure 2B). A high number of these SNPs exhibit two or more allelic variations, and the binding affinities with TFs vary based on the specific alleles, with examples represented by rs11014185 and rs10208097. In addition, we observed several SNPs that concurrently disrupted binding to two or more TFs. As shown in Figure 2C, 30 SNPs disrupted the binding of BATF::JUN and JUN, 17 SNPs disrupted the binding of ZNF460 and ZNF770, and 7 SNPs disrupted the binding of KLF12 and SPI. Specifically, rs2271319 affected the binding of RAD21 and SMC3. These findings highlight the significance of SNP genetic variation in regulating crucial and intricate molecular interactions within the gene expression regulatory network.

#### 3.2 | Functional Annotation of TF Binding-Disrupting SNPs

To identify the most functional SNPs of these TF binding-disrupting SNPs and elucidate their regulatory mechanisms, we systematically annotated these SNPs by using several bioinformatic databases for subsequent screening. The selection criteria of top functional SNPs for reporter gene assays were as follows: first, the SNP must be significantly associated with gene expression in human brains ( $P_{FDR} < 0.001$ ); second, the SNP should have a RegulomeDB score of 1. After this, the remaining SNPs were scored based on functional information derived from the HaploReg and 3DSNP databases, specifically considering (i) promoter histone marks, (ii) enhancer histone marks, (iii) DNase protein binding, and (iv) motif changes. The resulting sites were then sorted. Detailed information is found in [Supporting Information](#).

Among the initial set of 736 TF binding-disrupting SNPs, we observed that 464 SNPs exhibited a significant association with gene expression in human brains in at least one brain eQTL

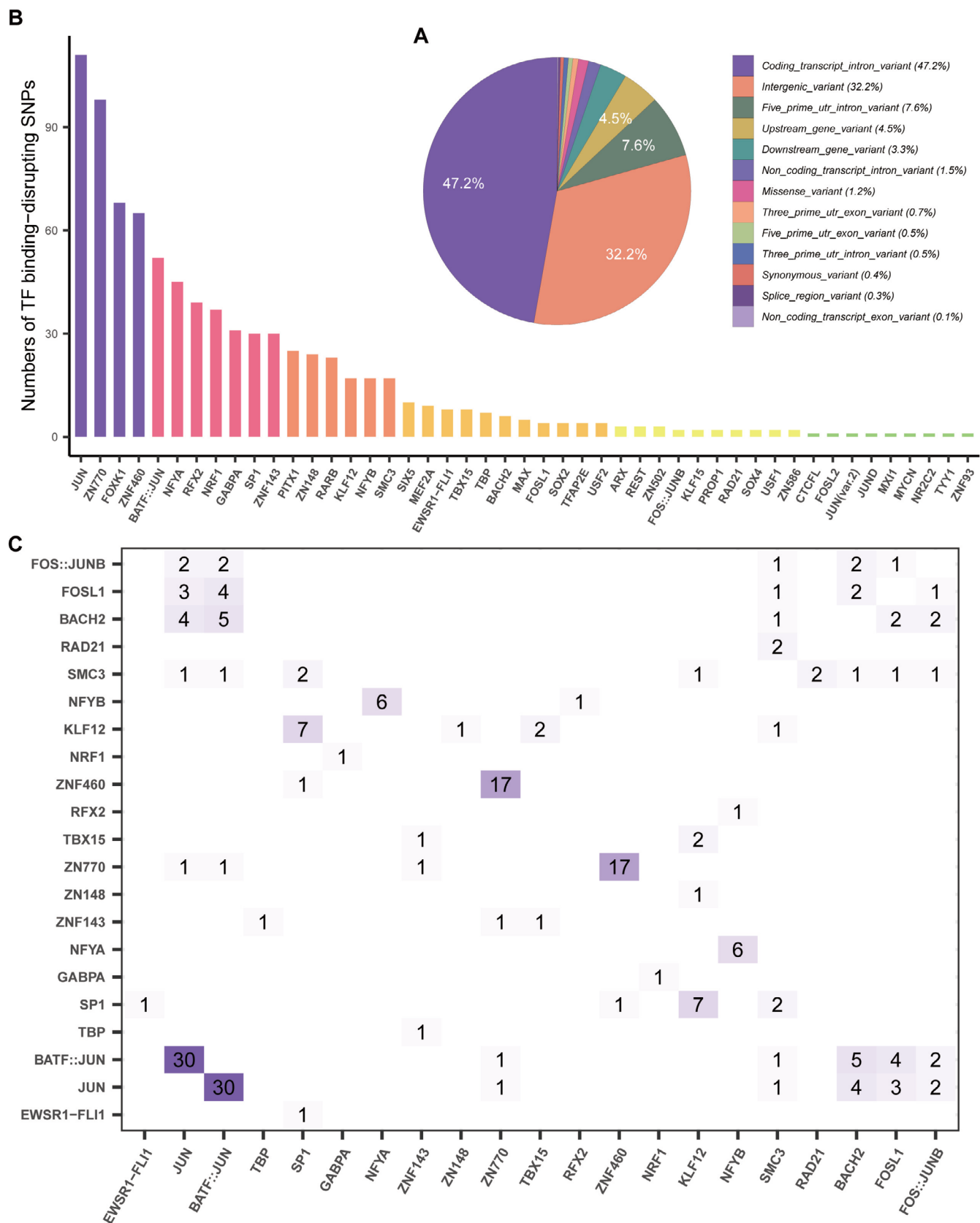
dataset. Notably, 323 identified regulatory SNPs were associated with the expression of the same gene in at least two independent brain eQTL datasets, implying these genes are regulated by these TF binding-disrupting SNPs. Out of the 464 SNPs with significant eQTL analysis, 335 SNPs have a RegulomeDB score of one. Integrating the information retrieved from HaploReg and 3DSNP, we systematically scored and ranked the remaining SNPs. The top three ranked loci were rs12146541 (G>C, which reduces binding to SMC3 after transition), rs2234053 (G>C, which enhances binding to ZNF148 after transition), and rs223333 (C>A, which strengthens binding to RFX2 after transition). Based on the ranking, we randomly selected several of the top 20 SNPs for subsequent reporter gene assays, including rs1399178, rs223333, rs3776130, rs4788211, rs655293, rs7204852, rs12146541, and rs2016875.

#### 3.3 | Functional Validation of Identified SNPs: rs1399178 Mediates Promoter-Enhancer Switch

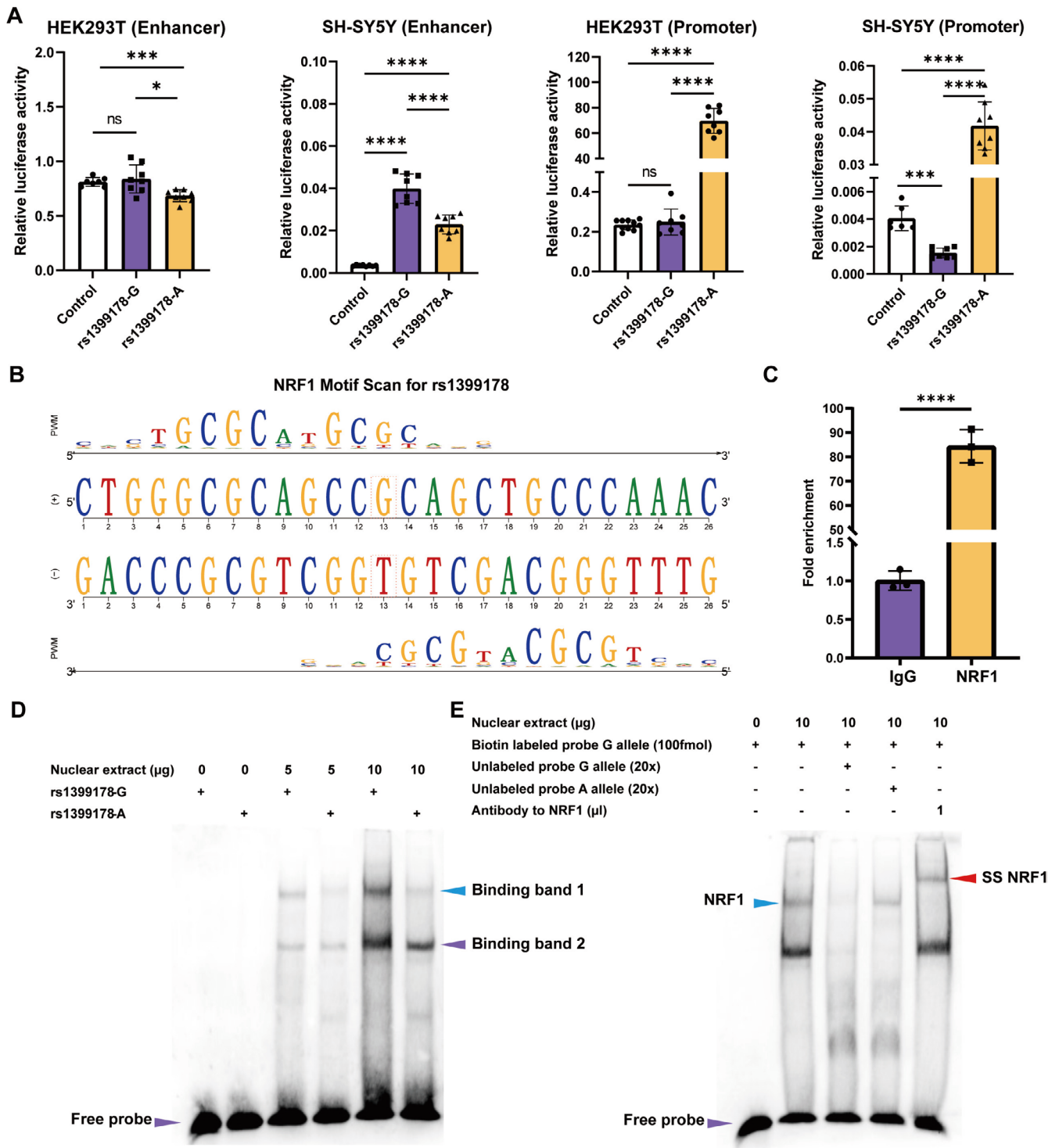
We conducted dual-luciferase reporter gene assays to validate the impact of the identified functional SNPs in HEK293T cells (Figure S1). With the exception of rs7204852, five SNPs exhibited significant changes in enhancer luciferase activity, indicating significant enhancer regulatory effects. Notably, minor alleles of rs1399178, rs223333, rs3776130, and rs655293 displayed significantly reduced luciferase activity compared to wild-type variants, whereas the opposite was observed in rs4788211.

We selected four SNPs to evaluate promoter activity, specifically rs1399178, rs2016875, rs655293, and rs12146541, but only rs1399178 and rs12146541 demonstrated altered luciferase activity. Remarkably, a substantial increase in promoter activity was observed following variation of these two SNPs ( $p < 0.001$ ). Collectively, these findings provide evidence for the regulatory function of TF binding-disrupting SNPs. To delve further into the regulatory mechanisms of the functional SNPs, particular attention was given to SNP rs1399178, which exhibited significant changes in luciferase activity in both enhancer and promoter contexts. At the same time, association analysis showed that rs1399178 was significantly associated with schizophrenia in the European population ( $p = 1.27 \times 10^{-3}$ ; 33,640 schizophrenia cases and 43,456 controls) and trans-ancestry meta-analysis ( $p = 2.92 \times 10^{-3}$ ; 48,223 schizophrenia cases and 60,717 controls) (Table S2).

To further confirm the regulatory effects of rs1399178, we conducted reporter gene assays in both HEK293T and SH-SY5Y cell lines (Figure 3A). Rs1399178 is located in an intergenic region of chromosome 3p22.2, positioned between the LRRFIP2 and GOLGA4. Reporter gene assays revealed that the enhancer luciferase activity of the fragments containing the G allele of rs1399178 was higher than controls in SH-SY5Y cell lines. The opposite was observed in promoter activity. Interestingly, no significant differences were observed between the G allele and controls in HEK293T, suggesting a cell-type specific regulatory effect of rs1399178 (i.e., rs1399178 has a regulatory effect in neuroblastoma cell lines but not in HEK293T). Of note, the G allele exhibited significantly higher enhancer activity than the A allele in both tested cell lines. Intriguingly, the A allele conferred significant higher promoter activity compared to the G allele.



**FIGURE 2** | Overview of the TF binding-disrupting SNPs. The distribution of TF binding-disrupting SNPs in the human genome. (A) The number of SNPs that affect the binding affinity of each TF. (B) Distribution of the binding-disrupting SNPs in the human genome. (C) Heatmap showing the number of SNPs that affect binding of two or more TFs.



**FIGURE 3** | Validation of the regulatory effects of rs1399178. (A) The genomic region containing rs1399178 has enhancer and promoter activities in SH-SY5Y but not in HEK293T. Compared with the A allele, the G allele exhibits higher enhancer activity and lower promoter activity in HEK293T and SH-SY5Y ( $n = 8$  for each group). (B) rs1399178 is located in the binding motif of NRF1. (C) Significant enrichment of NRF1 on genomic sequence containing rs1399178 ( $n = 3$  for each group). (D, E) SNP rs1399178 affects NRF1-binding affinity. Competitive experiments showing that the G allele had stronger NRF1-binding affinity compared with the A allele. Super-shift assay showing that rs1399178 could bind to NRF1 (the purple arrow indicates the free probe, the blue arrow marks the band representing NRF1 binding, and the red arrow denotes the super-shift band observed in the lane with added NRF1 antibody). \* $p < 0.05$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ .

These results indicate that the genomic sequence containing rs1399178 exhibits both enhancer and promoter activities, with the regulatory effects of rs1399178 on enhancer activity differing from effects on promoter.

Based on the Hi-C results, we observed that rs1399178 is positioned near the inner boundary of a TAD (Figure S2) that includes interacting genes, such as LRRFIP2, MLH1, and GOLGA4. This suggests that rs1399178 may not only impact

the expression of nearby genes but could also influence distant genes within this loop via enhancer–promoter interactions. The findings from the eQTL analysis further support this observation, indicating that the presence of the risk allele is associated with decreased LRRFIP2 and increased MLH1 transcript abundance (Figure S3).

### 3.4 | Identification of NRF1 as the Binder for the G Allele of rs1399178

Motif analysis showed that rs1399178 disrupts the binding of the nuclear respiratory factor 1 (*NRF1*), which displays a heightened preference for binding to the DNA sequence containing the G allele of rs1399178 (Figure 3B). ChIP-qPCR analysis also revealed a significant enrichment of NRF1 binding on the sequence containing rs1399178 compared to the control in the SH-SH5Y cells (Figure 3C, Figure S4). We performed EMSA to further elucidate the effect of different rs1399178 alleles on the binding affinity of NRF1, which revealed a heightened binding capacity of the G allele compared to the A allele, as evidenced by a more pronounced band (band 1 in Figure 3D). Additionally, the results of the competitive experiment indicated significant disparities in binding affinities between the two alleles. Specifically, the TF in binding band 1 demonstrated a preference for the G allele. Moreover, the super-shift assays identified NRF1 as the TF in binding band 1, which was detected by electrophoresis upon the introduction of an anti-NRF1 antibody (indicated by the SS NRF1 arrowhead on the right side of Figure 3E). These observations suggest a preferential binding of NRF1 to the genomic sequence harboring the G allele of rs1399178 in vitro.

### 3.5 | Structural Insights Into NRF1 Recognition of DNA

To gain further insight into the role of rs1399178 in interfering NRF1–DNA recognition, we headed to generate a 3D structure model of full-length NRF1 by AlphaFold2. Sequence alignment revealed that NRF1 is highly conserved among species and belongs to the MADS-box family (Figures S5 and S6). AlphaFold2 prediction suggests that full-length NRF1 forms a dimer and each monomer comprises three domains: a Dimerization domain (DD, residues 49–177), a DNA-binding domain (DBD, residues 201–288), and a Transactivation domain (TAD, residues 301–476) (Figure 4A, Figure S7). Traditionally, MADS-box proteins bind to CArG-box sequences (CC-“adenine-rich”-GG), though lineage-specific amino acid variations in the MADS-box influence DNA-binding specificity [39, 40]. The NRF1-responsive element is characterized by a GC-rich palindrome with the core sequence YGCGCAYGCGCR (Y can be C or T, R can be A or G) [41]. To test the impact of rs1399178, we synthesized two DNA fragments: RS(WT) (5′-GGCGCAGCCGAGC-3′) and RS(G-A) (5′-GGCGCAGCCACAGC-3′) and used ITC to measure the binding affinity of the DD-DBD complex to these sequences. The G to A mutation at rs1399178 resulted in a tenfold decrease in binding strength, in line with observed biological effects (Figure 4B). We analyzed the complex structure of NRF1 (DD-DBD) with a consensus DNA sequence ATGCGCATGCGCAT (PDB code:8k4l). The DD-DBD dimer-dsDNA structure suggests that DD mediate dimerization of the protein, whereas DBD

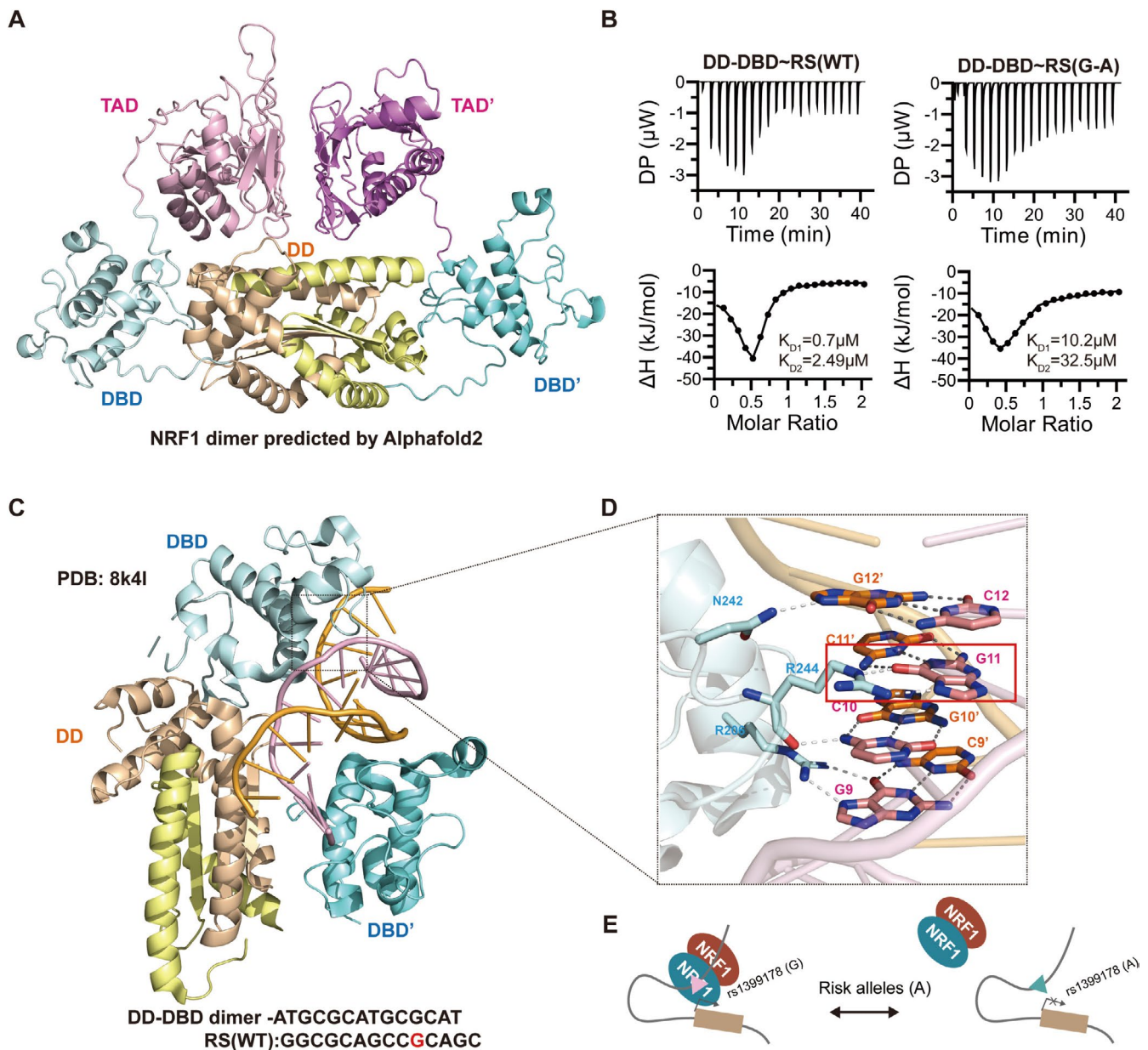
recognizes short TGCGC DNA sequence [42]. Interface analysis showed that R206 and G9, R244 and C10, G11, N242 and G12′ form hydrogen bonds for the specific DNA recognition (Figure 4C,D). The mutation from G11 to A in rs1399178 disrupted the hydrogen bond formation between R244 and G11, thereby affecting the stability of the complex.

## 4 | Discussion

Following the successful initial GWAS screening of schizophrenia-related alleles, which identified a single locus around the ZNF804A gene, numerous similar studies have been reported [7]. Each study builds upon the iterative findings of the earlier works, both confirming and expanding upon these discoveries [8, 10–13, 43, 44]. Presently, over 200 risk loci have been documented, which contributed to the substantial progress in understanding the genetics of schizophrenia, surpassing insights gained in any other field within biological psychiatry [45]. However, formidable challenges remain in interpreting causal variants from these reported risk loci and elucidating their mechanisms of action in the post-GWAS era. Here, to investigate potential functional regulatory mechanisms of schizophrenia risk loci, we selected 124 risk loci from a GWAS study conducted by Li et al. in 2017 [11]. We leverage several comprehensive datasets and used a multifaceted functional genomics approach to explore the regulatory consequences of these SNPs. Our analysis identified 736 TF binding-disrupting SNPs that affect TF binding, notably concentrated in intronic and intergenic regions. eQTL analysis suggests that they may influence the expression of multiple genes (Figure S8). Reporter gene assay of randomly selected TF binding-disrupting SNPs confirmed that the majority of these mutations exhibit regulatory functions. We note that changes of the same SNP to different genotypes may lead to varying binding affinities for the same TF. Furthermore, many SNPs disrupted the binding of two or more TFs. Together, these findings demonstrate the complexity of identifying causal variants and interpreting their regulatory mechanisms based on GWAS-derived SNP loci, urging some caution in the interpretation. Moreover, our results underscore the significance of integrative approaches combining genomics, functional assays, and structural analysis to unravel the intricate nature of psychiatric disorders, such as schizophrenia.

Regulatory elements are categorized as either promoters or enhancers based on their genomic location and can be associated with histone markers and regulatory functions. Here, we studied in detail a bifunctional regulatory element, rs1399178, which serves as both promoter and enhancer. Reporter gene assays conducted in SH-SY5Y and HEK293T cell lines demonstrated the bifunctional regulatory effect of rs1399178. The G allele consistently displayed significantly higher enhancer activity and lower promoter activity than the A allele in both cell lines. This observation aligns with the hypothesis that the genomic sequence harboring rs1399178 may mediate a switch between enhancer and promoter activities through the TF. Similarly, Li et al. demonstrated that high-risk schizophrenia SNPs rs1801311 exhibit cell line-specific regulatory activity [46]. Furthermore, leveraging 3D genome data, we ascertained the presence of rs1399178 in an active gene-regulating TAD, which justifies its classification as a bifunctional element. Our results





**FIGURE 4** | NRF1 comprises a MADS-box like domain via dimerization. (A) Cartoon diagram depicting the full-length NRF1 structure. (B) ITC-based measurements comparing DD-DBD binding to RS(WT) and RS(G-A). (C) Cartoon diagram illustrating the binding of DD-DBD to DNA (PDB ID: 8k4l). (D) Close-up view of the interactions between DBD and DNA. (E) A schematic diagram illustrating the mechanism by which SNP rs1399178 affects NRF1 binding and downstream gene expression. The G allele of rs1399178 enhances NRF1-binding affinity, promoting its interaction with the DNA-binding site. This interaction regulates the transcriptional activity of target genes, potentially influencing downstream biological processes.

also suggest the involvement of this variant in long-range regulatory activities, possibly acting as an enhancer to modulate the expression of other genes. Within the local TAD, factors such as local chromatin architecture and the abundance/occupancy of different TFs may influence dual gene regulation. Based on the eQTL analysis, we found that the risk allele A of rs1399178 is associated with lower expression of *LRRFIP2* and higher expression of *MLH1*. *MLH1* encodes a mismatch repair protein that is essential for maintaining genomic stability. Dysregulation of *MLH1* can lead to genomic instability and abnormal neurodevelopment, thereby increasing the risk of schizophrenia [47]. Using a Bayesian framework to integrate heterogeneous genomic data, *MLH1* has been identified as a top-ranked risk

gene for schizophrenia. Additionally, experimental studies have validated its association with the disorder for the first time. Furthermore, *MLH1* may have genetic overlap with other neuropsychiatric disorders, such as bipolar disorder and depression, highlighting its central role in mental illness. Notably, *LRRFIP2* has also been implicated in other brain and mental disorders, such as Alzheimer's disease [48, 49]. The biological mechanisms involving *LRRFIP2* are evident across multiple pathological contexts, suggesting that this gene may influence schizophrenia risk through diverse regulatory pathways. The mechanisms through which rs1399178 exerts its effects remain unclear, specifically whether it induces changes in loop structures by altering TF-binding properties. To determine whether

this functional SNP was associated with schizophrenia across multiple populations, we conducted a case–control association study. In both European populations and in the trans-ancestry analyses spanning diverse ethnic groups, rs1399178 exhibits a significant association with schizophrenia. Interestingly, this effect is not observed in the Chinese population, which underscores the importance of stratifying schizophrenia populations for more nuanced research into the genetic basis of this condition [13, 45]. Notably, rs1399178 has a different minor allele frequency (MAF) between Europeans (MAF, 0.431, minor allele: A) and Chinese (MAF, 0.069) populations. Further research is required to evaluate the impact of genetic variation on hierarchical genome structures [50].

Intriguingly, the disruption of NRF1 binding by rs1399178 was clearly observed, with a heightened preference for binding to the G allele. Lower binding affinity of NRF1 was observed at the risk alleles through EMSA and ITC. Recently, Liu elucidated the molecular mechanism by which NRF1 recognizes specific DNA sequences in dimeric form [42]. Our sequences exhibit striking similarity to the consensus binding sequence GCGCATGCGC of NRF1 as enumerated in the article. The results demonstrate that substitution of G11 in the conserved binding sequence GCGCATGCGC with either T or A results in a respective decrease in binding affinity by 4- and 6-fold, consistent with our findings. The diminished binding affinity significantly compromises NRF1's transcriptional activity on downstream genes. NRF1 has been reported to regulate ~471 downstream genes, including mitochondria- and neurodevelopment-related loci [51, 52]. The *in vitro* binding assay demonstrated that the A allele of rs1399178 significantly weakens the binding strength. The structural insights into NRF1's DNA recognition and dimer formation provide a basis for understanding its role in transcriptional regulation. Through the validation of rs1399178, we have further substantiated the rationale of our analytical framework, establishing that TF binding-disrupting SNPs have a significant influence on the binding of TFs.

Our study possesses several notable strengths. First, given the highly tissue-specific nature of genetic regulatory elements, this investigation exclusively incorporates the most recent and comprehensive ChIP-Seq data from brain tissues or neuronal cell lines. This stringent criterion ensures the sole examination of risk variants located within actively regulated regions of the brain that are characterized by corresponding TF binding. Second, we systematically characterize the regulatory mechanisms of reported schizophrenia risk loci using a relatively high-throughput approach. By integrating functional genomics and experimental validation, we verified several previously identified TF-associated SNPs. Third, our transformation of SNPs into specific target genes through eQTL analysis provides novel insights for future research into the mechanisms of schizophrenia and drug development. Fourth, the use of multiple molecular approaches revealed that rs1399178 is a bifunctional regulatory element located within an active TAD. This suggests that the locus may regulate various genes within the loop through distinct regulatory mechanisms, contributing to a deeper understanding of the mechanisms underlying schizophrenia. Finally, we provided evidence for the molecular mechanism of the binding between rs1399178 and NRF1, further

validating the characteristic of NRF1 in recognizing specific DNA sequences and offering new insights into investigating schizophrenia susceptibility loci.

Our study also has several limitations. First, our analysis relied solely on SH-SY5Y and HEK293T cell lines for reporter gene assays. Given the observed cell-type-specific effects, emphasizing the importance of considering tissue-specific regulatory mechanisms, it is imperative to incorporate additional cell lines relevant to schizophrenia for a more comprehensive investigation in the future. Second, this study exclusively focused on single nucleotide variants, neglecting other types of genetic mutations, such as copy number variations, chromatin structural loci, and rare mutations, all of which are implicated in schizophrenia. Further research is thus required to understand the genetic mechanisms of schizophrenia in relation to these types of variants. Third, we did not experimentally validate the downstream target genes affected by rs1399178.

## 5 | Conclusion

We identified 736 SNPs that impact the binding of 47 TFs and validated the regulatory effects of several of these SNPs through reporter gene experiments. We elucidated the bifunctional regulatory element of rs1399178, an inter-genetic schizophrenia risk variant that mediates promoter-enhancer switch. We demonstrated that the A allele of rs1399178 attenuates the binding affinity of NRF1, due to specific amino acids of NRF1's MADS-box domain. Our findings contribute to a more nuanced understanding of the regulatory mechanisms associated with rs1399178 in the context of schizophrenia. The combination of genomics, functional assays, structural analysis, and in-depth exploration of NRF1's recognition mechanisms provide a comprehensive framework for future research into the genetic mechanisms of schizophrenia and potential therapeutic interventions.

---

## Author Contributions

G.H. and Q.L. conceived the concept; L.J., B.H., G.H., and Q.L. designed experiments; G.H. led the project with assistance from L.J., B.H., D.R., X.W., L.H., and Q.L.; L.J. performed experiments with assistance from B.H., D.R., L.L., Y.B., F.Y., K.H., K.L., X.L., and T.Y.; L.J., B.H., L.L., Z.L., Y.S., and Q.L. analyzed data; L.J., B.H., Q.L., and G.H. wrote the manuscript with input from all coauthors.

## Acknowledgments

We extend our sincere appreciation to the authors and laboratories who generously shared their public data, which was instrumental in the completion of this research. Their efforts in collecting and making the data publicly available have significantly contributed to the scientific community. The invaluable contributions of Miss JunYang Wang in the experimental details were crucial to the success of this study. This project is supported by the National Key Research and Development Program (2024YFC2707002, 2022YFE0125300), Innovation Program of Shanghai Municipal Education Commission (2023ZKZD16), National Natural Science Foundation of China (82071262, 32300464, 82301682), Shanghai Municipal Science and Technology Major Project (2017SHZDZX01, 20JC1418600), Natural Science Foundation of Shanghai (23ZR1432700), China Postdoctoral Science Foundation (2023M732266), the Shanghai Leading Academic Discipline Project (B205), Key Technology

## Conflicts of Interest

The authors declare no conflicts of interest.

## Data Availability Statement

The data that support the findings of this study are available in GTRD, eQTL Catalogue, RegulomeDB, HaploReg, 3DSNP v2.0 and 4DN data portal. These data were derived from the following resources available in the public domain: GTRD, <http://gtrd.biouml.org/>; eQTL Catalogue, <https://www.ebi.ac.uk/eqtl/>; RegulomeDB, <https://www.regulomedb.org/>; HaploReg, <http://www.broadinstitute.org/mammals/haploreg/haploreg.php>; 3DSNP v2.0, <https://omic.tech/3dsnpv2/>; 4DN data portal, <https://data.4dnucleome.org/>. Other data can be available from the corresponding author upon reasonable request.

## References

1. R. S. Kahn, I. E. Sommer, R. M. Murray, et al., "Schizophrenia," *Nature Reviews. Disease Primers* 1 (2015): 15067.
2. R. Tandon, "Schizophrenia and Other Psychotic Disorders in Diagnostic and Statistical Manual of Mental Disorders (DSM)-5: Clinical Implications of Revisions From DSM-IV," *Indian Journal of Psychological Medicine* 36, no. 3 (2014): 223–225.
3. K. T. Mueser and S. R. McGurk, "Schizophrenia," *Lancet* 363, no. 9426 (2004): 2063–2072.
4. S. Saha, D. Chant, J. Welham, and J. McGrath, "A Systematic Review of the Prevalence of Schizophrenia," *PLoS Medicine* 2, no. 5 (2005): e141.
5. M. Knapp, R. Mangalore, and J. Simon, "The Global Costs of Schizophrenia," *Schizophrenia Bulletin* 30, no. 2 (2004): 279–293.
6. T. M. Laursen, M. Nordentoft, and P. B. Mortensen, "Excess Early Mortality in Schizophrenia," *Annual Review of Clinical Psychology* 10 (2014): 425–448.
7. M. C. O'Donovan, N. Craddock, N. Norton, et al., "Identification of Loci Associated With Schizophrenia by Genome-Wide Association and Follow-Up," *Nature Genetics* 40, no. 9 (2008): 1053–1055.
8. J. Shi, D. F. Levinson, J. Duan, et al., "Common Variants on Chromosome 6p22.1 Are Associated With Schizophrenia," *Nature* 460, no. 7256 (2009): 753–757.
9. W. H. Yue, H. F. Wang, L. D. Sun, et al., "Genome-Wide Association Study Identifies a Susceptibility Locus for Schizophrenia in Han Chinese at 11p11.2," *Nature Genetics* 43, no. 12 (2011): 1228–1231.
10. Schizophrenia Working Group of the Psychiatric Genomics C, "Biological Insights From 108 Schizophrenia-Associated Genetic Loci," *Nature* 511, no. 7510 (2014): 421–427.
11. Z. Li, J. Chen, H. Yu, et al., "Genome-Wide Association Analysis Identifies 30 New Susceptibility Loci for Schizophrenia," *Nature Genetics* 49, no. 11 (2017): 1576–1583.
12. A. F. Pardinas, P. Holmans, A. J. Pocklington, et al., "Common Schizophrenia Alleles Are Enriched in Mutation-Intolerant Genes and in Regions Under Strong Background Selection," *Nature Genetics* 50, no. 3 (2018): 381–389.
13. V. Trubetskoy, A. F. Pardinas, T. Qi, et al., "Mapping Genomic Loci Implicates Genes and Synaptic Biology in Schizophrenia," *Nature* 604, no. 7906 (2022): 502–508.
14. S. H. Lee, T. R. DeCandia, S. Ripke, et al., "Estimating the Proportion of Variation in Susceptibility to Schizophrenia Captured by Common SNPs," *Nature Genetics* 44, no. 3 (2012): 247–250.
15. J. K. Pritchard and M. Przeworski, "Linkage Disequilibrium in Humans: Models and Data," *American Journal of Human Genetics* 69, no. 1 (2001): 1–14.
16. J. van de Leemput, J. L. Hess, S. J. Glatt, and M. T. Tsuang, "Genetics of Schizophrenia: Historical Insights and Prevailing Evidence," *Advances in Genetics* 96 (2016): 99–141.
17. H. Won, L. de la Torre-Ubieta, J. L. Stein, et al., "Chromosome Conformation Elucidates Regulatory Relationships in Developing Human Brain," *Nature* 538, no. 7626 (2016): 523–527.
18. W. Mah and H. Won, "The Three-Dimensional Landscape of the Genome in Human Brain Tissue Unveils Regulatory Mechanisms Leading to Schizophrenia Risk," *Schizophrenia Research* 217 (2020): 17–25.
19. S. L. Edwards, J. Beesley, J. D. French, and A. M. Dunning, "Beyond GWASs: Illuminating the Dark Road From Association to Function," *American Journal of Human Genetics* 93, no. 5 (2013): 779–797.
20. S. A. Bacanu, J. Chen, J. Sun, et al., "Functional SNPs Are Enriched for Schizophrenia Association Signals," *Molecular Psychiatry* 19, no. 3 (2014): 276–277.
21. A. O. Degtyareva, E. V. Antontseva, and T. I. Merkulova, "Regulatory SNPs: Altered Transcription Factor Binding Sites Implicated in Complex Traits and Diseases," *International Journal of Molecular Sciences* 22, no. 12 (2021): 6454, <https://doi.org/10.3390/ijms22126454>.
22. Y. Huo, S. Li, J. Liu, X. Li, and X. J. Luo, "Functional Genomics Reveal Gene Regulatory Mechanisms Underlying Schizophrenia Risk," *Nature Communications* 10, no. 1 (2019): 670.
23. A. Senturk, S. Pfennig, A. Weiss, K. Burk, and A. Acker-Palmer, "Ephrin Bs Are Essential Components of the Reelin Pathway to Regulate Neuronal Migration," *Nature* 472, no. 7343 (2011): 356–360.
24. S. H. Fatemi and T. D. Folsom, "The Neurodevelopmental Hypothesis of Schizophrenia, Revisited," *Schizophrenia Bulletin* 35, no. 3 (2009): 528–548.
25. J. L. Rapoport, J. N. Giedd, and N. Gogtay, "Neurodevelopmental Model of Schizophrenia: Update 2012," *Molecular Psychiatry* 17, no. 12 (2012): 1228–1238.
26. S. A. Lambert, A. Jolma, L. F. Campitelli, et al., "The Human Transcription Factors," *Cell* 172, no. 4 (2018): 650–665.
27. M. J. Machiela and S. J. Chanock, "LDlink: A Web-Based Application for Exploring Population-Specific Haplotype Structure and Linking Correlated Alleles of Possible Functional Variants," *Bioinformatics* 31, no. 21 (2015): 3555–3557.
28. S. Li, Y. Li, X. Li, et al., "Regulatory Mechanisms of Major Depressive Disorder Risk Variants," *Molecular Psychiatry* 25, no. 9 (2020): 1926–1945.
29. S. Kolmykov, I. Yevshin, M. Kulyashov, et al., "GTRD: An Integrated View of Transcription Regulation," *Nucleic Acids Research* 49, no. D1 (2021): D104–D111.
30. T. L. Bailey, "STREME: Accurate and Versatile Sequence Motif Discovery," *Bioinformatics* 37, no. 18 (2021): 2834–2840.
31. C. Zuo, S. Shin, and S. Keles, "atSNP: Transcription Factor Binding Affinity Testing for Regulatory SNP Detection," *Bioinformatics* 31, no. 20 (2015): 3353–3355.
32. N. Kerimov, J. D. Hayhurst, K. Peikova, et al., "A Compendium of Uniformly Processed Human Gene Expression and Splicing Quantitative Trait Loci," *Nature Genetics* 53, no. 9 (2021): 1290–1299.
33. A. P. Boyle, E. L. Hong, M. Hariharan, et al., "Annotation of Functional Variation in Personal Genomes Using RegulomeDB," *Genome Research* 22, no. 9 (2012): 1790–1797.
34. L. D. Ward and M. Kellis, "HaploReg: A Resource for Exploring Chromatin States, Conservation, and Regulatory Motif Alterations Within Sets of Genetically Linked Variants," *Nucleic Acids Research* 40 (2012): D930–D934.



35. C. Quan, J. Ping, H. Lu, G. Zhou, and Y. Lu, "3DSNP 2.0: Update and Expansion of the Noncoding Genomic Variant Annotation Database," *Nucleic Acids Research* 50, no. D1 (2022): D950–D955.
36. J. Dekker, A. S. Belmont, M. Guttman, et al., "The 4D Nucleome Project," *Nature* 549, no. 7671 (2017): 219–226.
37. S. B. Reiff, A. J. Schroeder, K. Kirli, et al., "The 4D Nucleome Data Portal as a Resource for Searching and Visualizing Curated Nucleomics Data," *Nature Communications* 13, no. 1 (2022): 2365.
38. S. Purcell, B. Neale, K. Todd-Brown, et al., "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses," *American Journal of Human Genetics* 81, no. 3 (2007): 559–575.
39. M. R. Abdullah-Zawawi, N. F. Ahmad-Nizammuddin, N. Govender, S. Harun, N. Mohd-Assaad, and Z. A. Mohamed-Hussein, "Comparative Genome-Wide Analysis of WRKY, MADS-Box and MYB Transcription Factor Families in Arabidopsis and Rice," *Scientific Reports* 11, no. 1 (2021): 19678.
40. X. Lai, R. Vega-Leon, V. Hugouvieux, et al., "The Intervening Domain Is Required for DNA-Binding and Functional Identity of Plant MADS Transcription Factors," *Nature Communications* 12, no. 1 (2021): 4760.
41. N. Gleyzer, K. Vercauteren, and R. C. Scarpulla, "Control of Mitochondrial Transcription Specificity Factors (TFB1M and TFB2M) by Nuclear Respiratory Factors (NRF-1 and NRF-2) and PGC-1 Family Co-activators," *Molecular and Cellular Biology* 25, no. 4 (2005): 1354–1366.
42. K. Liu, W. Li, Y. Xiao, M. Lei, M. Zhang, and J. Min, "Molecular Mechanism of Specific DNA Sequence Recognition by NRF1," *Nucleic Acids Research* 52, no. 2 (2024): 953–966.
43. International Schizophrenia C, S. M. Purcell, N. R. Wray, et al., "Common Polygenic Variation Contributes to Risk of Schizophrenia and Bipolar Disorder," *Nature* 460, no. 7256 (2009): 748–752.
44. Schizophrenia Psychiatric Genome-Wide Association Study C, "Genome-Wide Association Study Identifies Five New Schizophrenia Loci," *Nature Genetics* 43, no. 10 (2011): 969–976.
45. M. J. Owen, S. E. Legge, E. Rees, J. T. R. Walters, and M. C. O'Donovan, "Genomic Findings in Schizophrenia and Their Implications," *Molecular Psychiatry* 28, no. 9 (2023): 3638–3647, <https://doi.org/10.1038/s41380-023-02293-8>.
46. Y. Li, C. Ma, W. Li, et al., "A Missense Variant in NDUFA6 Confers Schizophrenia Risk by Affecting YY1 Binding and NAGA Expression," *Molecular Psychiatry* 26, no. 11 (2021): 6896–6911.
47. D. He, L. Li, H. S. Zhang, et al., "Accurate Identification of Genes Associated With Brain Disorders by Integrating Heterogeneous Genomic Data Into a Bayesian Framework," *eBioMedicine* 107 (2024): 105286.
48. M. A. A. DeMichele-Sweet, E. A. Weamer, L. Klei, et al., "Genetic Risk for Schizophrenia and Psychosis in Alzheimer Disease," *Molecular Psychiatry* 23, no. 4 (2018): 963–972.
49. X. Peng, J. S. Bader, and D. Avramopoulos, "Schizophrenia Risk Alleles Often Affect the Expression of Many Genes and Each Gene May Have a Different Effect on the Risk: A Mediation Analysis," *American Journal of Medical Genetics* 186, no. 4 (2021): 251–258.
50. M. E. Tabangin, J. G. Woo, and L. J. Martin, "The Effect of Minor Allele Frequency on the Likelihood of Obtaining False Positives," *BMC Proceedings* 3 (2009): S41.
51. I. U. Haussmann, K. White, and M. Soller, "Erect Wing Regulates Synaptic Growth in Drosophila by Integration of Multiple Signaling Pathways," *Genome Biology* 9, no. 4 (2008): R73.
52. J. Fu, J. Zhang, F. Jin, J. Patchefsky, K. H. Braunewell, and A. J. Klein-Szanto, "Promoter Regulation of the Visinin-Like Subfamily of Neuronal Calcium Sensor Proteins by Nuclear Respiratory Factor-1," *Journal of Biological Chemistry* 284, no. 40 (2009): 27577–27586.

## Supporting Information

Additional supporting information can be found online in the Supporting Information section.