Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Neuro-Genetic Mechanisms Connecting Brain Structure to Breast Cancer

Jing Li
the Sixth Medical Center of PLA General Hospital https://orcid.org/0009-0005-5928-1147
Jun Chen
Shaanxi University of Chinese Medicine
Xue Li
the Sixth Medical Center of PLA General Hospital
Mingyue Hao
the Sixth Medical Center of PLA General Hospital
Fei Ji
the Sixth Medical Center of PLA General Hospital
Wei Ma
langmawei@bjmu.edu.cn
the Sixth Medical Center of PLA General Hospital

Article

Keywords: Breast cancer, Brain imaging-derived phenotypes, Mendelian randomization, Population-specific genetics

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8676570/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Background:

Breast cancer incidence has continued to rise over recent decades. Beyond established genetic and hormonal factors, neuropsychiatric conditions including depression, anxiety, and sleep disturbances, have emerged as independent risk factors. Brain structure, as a core neuroendocrine substrate, may share genetic determinants with breast cancer risk; however, this relationship remains insufficiently characterized.

Objective:

This study investigated the potential genetic correlations between brain imaging-derived phenotypes (IDPs) and breast cancer using genome-wide association study (GWAS) summary statistics via Mendelian randomization (MR) analysis. We hypothesized that shared genetic architecture between neuropsychiatric traits and breast cancer may reveal previously unrecognized biological pathways.

Methods:

We conducted MR analysis to systematically evaluate potential causal associations between 4,013 IDPs and 14 breast cancer (BC) traits, encompassing overall BC and molecular subtypes. Genetic instruments for IDPs were obtained from the UK Biobank and ENIGMA consortium. Summary-level BC data were sourced from IEU OpenGWAS, FinnGen, and the GWAS Catalog, covering European and African ancestry populations. A combined multi-ancestry linkage disequilibrium (LD) reference panel was constructed to optimize instrument clumping. The inverse-variance weighted (IVW) method served as the primary analysis, supported by sensitivity tests and colocalization to assess robustness and identify shared causal loci.

Results:

Temporal lobe structure demonstrated the strongest causal associations with estrogen receptor-positive (ER+) breast cancer. Within this region, the fusiform gyrus showed the most consistent negative causal effects across European ancestry datasets: right fusiform volume (IDP: GCST90002807, *P*-value = 2.23×10^{-7}), right fusiform surface area (IDP: GCST90003180, *P*-value = 1.11×10^{-6}), and left fusiform surface area (IDP: GCST90003079, *P*-value = 5.23×10^{-6}). Cross-ancestry analyses highlighted substantial population heterogeneity, as associations identified in European cohorts were not replicated in African ancestry datasets. Colocalization further localized shared genetic signals to chromosomes 17 and 12, implicating genes such as *ARL17B* and *NSFP1*, which are enriched in pathways related to intracellular vesicular transport and membrane dynamics.

Conclusion:

Our findings reveal a causal relationship between brain structure and breast cancer risk, identifying the fusiform gyrus as a key region associated with ER+ BC. The study uncovers a shared genetic basis underlying this brain-breast axis, centered on genes involved in vesicular transport and membrane regulation. These insights suggest that neural processes may contribute to cancer susceptibility and provide new directions for biomarker development and mechanistic research.

Introduction

Breast cancer (BC) remains the most common malignancy among women worldwide, and its global burden continues to rise[1]. According to the WHO, incidence rates are increasing by approximately 1–5% per year in nearly half of monitored countries[2], and global projections estimate that by 2050 the number of new cases will increase by 38%, reaching nearly 3.2 million annually[3]. In China, this trend is characterized by a “dual increase”, with an annual incidence of 423,000 new cases and a distinct shift toward earlier onset, as 28.7% of patients are diagnosed before the age of 50[4].

In addition to classical risk factors such as BC mutations[5] and prolonged reproductive hormone exposure[6], chronic psychological stress has increasingly been recognized as a significant contributor to breast cancer risk[7, 8]. Consistent with epidemiological findings, our previous work demonstrated that depression can actively promote breast cancer progression[9]. Mechanistically, persistent activation of the hypothalamic-pituitary-adrenal (HPA) axis disrupts glucocorticoid receptor signaling, thereby reshaping the tumor immune microenvironment[10] and facilitating angiogenesis[11, 12]. Despite these observations, the biological mechanisms linking brain structure, mental health, and breast carcinogenesis remain poorly understood.

Imaging-derived phenotypes (IDPs), as robust neurobiological markers, exhibit causal relationships with multiple systemic diseases. For example, patients recovering from acute COVID-19 display accelerated hippocampal atrophy and reduced functional connectivity within the default mode network[13, 14]. The severity of atrophy correlates strongly with serum neurofilament light chain (NfL) levels, suggesting axonal injury as a central pathological driver of structural degeneration[15, 16]. In multiple sclerosis (MS), the expansion rate of whole-brain white matter lesion volume is influenced by the *HLA-DRB1*15:01* allele, whereas progressive decline in corpus callosum fractional anisotropy (FA) predicts 5-year disability progression and colocalizes spatially with microglial activation markers[17, 18]. Collectively, these findings establish IDPs as anatomical anchors for investigating neuro-immune axis mechanisms. However, while IDPs have been extensively studied in neuropsychiatric and immune-mediated disorders, their potential role in breast cancer remains largely unexplored, representing a critical gap in current knowledge.

In this study, we apply MR analysis to examine the potential causal relationships between a broad spectrum of brain structural phenotypes and breast cancer susceptibility. Specifically, we aim to address the following questions: (i) Do specific imaging-derived phenotypes exert causal effects on the risk of different molecular subtypes of breast cancer? (ii) Which brain regions and structural features show the

most robust and subtype-specific associations? (iii) What shared genetic mechanisms and functional pathways underlie these observed brain-cancer relationships?

Methods

1. Study Design

As shown in Fig. 1, we conducted MR analysis to systematically screen brain IDPs for potential neuroimaging biomarkers of BC and to characterize alterations in the IDP profile associated with BC development. MR relies on three core assumptions: (i) genetic variants are robustly associated with the exposure; (ii) these variants are independent of potential confounders; and (iii) the variants influence the outcome exclusively through their effects on the exposure. All original genome-wide association study (GWAS) received ethical approval from their respective institutional review boards, and written informed consent was obtained from each participant prior to data collection.

2. Data Sources

(1) Brain Imaging-Derived Phenotypes (IDPs)

Genetic data for brain structural traits were obtained from two major sources.

UK Biobank IDPs:

The first dataset was derived from the 2021 publication “An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank”, which reported GWAS summary statistics for 3,935 IDPs in 22,138 individuals. These datasets are available in the GWAS Catalog under accession numbers GCST90002426 to GCST90006360 and are detailed on a dedicated project website.

ENIGMA Consortium:

The second dataset was obtained from the Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) consortium (https://enigma.ini.usc.edu/research/- download-enigma-gwas-results/). We applied for all eight available data freezes and received access to the 2015, 2018, 2020, and 2022 datasets. The 2018 and 2022 releases lacked sufficient variant-level information for MR and were excluded. The 2015 dataset provided GWAS data for 8 brain region volumes, including intracranial volume and seven subcortical structures (n = 30,717). The 2020 dataset contributed 70 GWAS datasets covering total brain surface area, average cortical thickness, and the area and thickness of 34 functionally specialized cortical regions. In total, ENIGMA contributed 78 brain-related GWASs, resulting in 4,013 IDPs included in the final analysis.

(2) Breast Cancer Data

Breast cancer summary statistics were obtained from three repositories.

The FinnGen R12 summary statistics were accessed through a successful data application. Ieu-a-1127 and ieu-a-1128 were obtained from IEU OpenGWAS (https://opengwas.io/). Additional datasets were retrieved from the GWAS Catalog, including summary statistics from two major publications (PMID: 32424353 ref pb53 dataset and 38741014 ref pb14 dataset). Altogether, 14 BC GWAS datasets were incorporated, covering overall BC and major molecular subtypes: ER-positive, ER-negative, triple-negative BC (TNBC), HER2-enriched, Luminal A, and Luminal B. All datasets included more than 20,000 participants, except for one smaller cohort (GCST90296722, n = 19122). With the exception of a single dataset based on an African American or Afro-Caribbean population, all studies were conducted in populations of European ancestry. A detailed list of datasets is provided in Supplementary Table 1.

3. MR Analysis

The instrumental variables (IVs) were selected according to the following criteria: (1) SNPs were first identified across the genome based on their genome-wide significant association with IDPs (P-value < 5×10^{-8}); (2) Only variants with a minor allele frequency (MAF) greater than 0.01 were retained; (3) To avoid linkage disequilibrium (LD), we applied LD clumping using a threshold of $R^2 < 0.001$ and a clumping window of 10,000 kb, ensuring that the remaining SNPs were independently associated with the IDPs; (4) If a selected SNP was not available in the outcome summary statistics, we searched for a proxy SNP in strong LD ($R^2 > 0.8$) to substitute for the missing variant; (5) For each SNP, we calculated the F statistic to evaluate instrument strength and eliminate potential weak instruments. The F statistic was computed using the formula $F = R^2 \times (N - 2)/(1 - R^2)$, where R^2 denotes the proportion of variance in the exposure explained by the SNP. SNPs with $F > 10$ were considered sufficiently strong instruments. We applied the inverse-variance weighted (IVW) approach to estimate the genetically predicted effect of the exposure on the outcome, expressed as the coefficient β. Heterogeneity within the IVW analysis was assessed using Cochran's Q statistic. When significant heterogeneity was detected (P-value < 0.05), a random-effects IVW model was adopted; otherwise, a fixed-effects model was used. Substantial heterogeneity may indicate that the core assumptions of MR are not fully satisfied. To minimize the loss of informative SNPs during LD clumping, a common issue when relying on a single-population linkage disequilibrium (LD) reference panel, we constructed a unified multi-ancestry reference panel. Specifically, we merged the Phase 3 1000 Genomes LD reference panels from five major ancestral groups. These panels were downloaded from:

http://fileserve.mrcieu.ac.uk/ld/1kg.v3.tgz

Because LD clumping retains only SNPs present in the reference panel, expanding the SNP coverage increases the number of eligible variants and consequently improves the stability and sensitivity of MR estimations. The composition of the merged LD reference panel was as follows:

- EUR: 503 individuals; 8,550,156 SNPs

- AFR: 661 individuals; 14,806,113 SNPs
- AMR: 347 individuals; 9,768,133 SNPs
- EAS: 504 individuals; 7,550,541 SNPs
- SAS: 489 individuals; 8,954,568 SNPs
- Merged 5-ancestry union panel: 2,054 individuals; 19,218,135 SNPs

This substantial increase in variant density markedly reduced SNP loss during clumping. All LD clumping procedures were performed using Plink software (https://www.cog-genomics.org/plink/, version v1.9.0-b.7.7 b4-bit).

4. Specificity Calculation of MR Results

Specificity was evaluated for each of the 14 BC outcome datasets. For a given outcome, we first identified the union of significant IDPs associated with the remaining 13 outcomes. Specificity was then defined as 1 minus the proportion of significant IDPs for that outcome that overlapped with this union set, i.e., 1 - (overlapping IDPs/total significant IDPs). As shown in Fig. 2 (with outcomes ranked by the number of significant IDPs), the ER-positive breast cancer dataset from the pb14 study exhibited the highest specificity, likely reflecting ancestral population differences. In contrast, the other ER-positive datasets (including Luminal A) showed relatively low specificity, which may be due to the high genetic similarity among these European-ancestry datasets.

5. Similarity and Significance of MR Results

Of the 4,013 brain IDPs analyzed, 2,219 exhibited a significant causal association with at least one breast cancer outcome. The similarity between any two outcomes was quantified using the Jaccard index, defined as the ratio of the intersection to the union of their significant IDP sets. To evaluate whether the observed similarity exceeded what would be expected by chance, we conducted a resampling-based significance test. For each outcome pair, 10,000 random samples of IDPs were drawn from the total pool of 2,219 to generate the null distribution of overlap sizes. The false discovery rate (*FDR*) was calculated as the proportion of simulated overlaps that were greater than or smaller than the observed overlap. An $FDR < 0.1$ was interpreted as significant similarity (“over”) or significant dissimilarity/independence (“under”).

6. Colocalization Analysis

Genetic colocalization analysis was conducted using the coloc R package. A posterior probability indicating a shared causal variant ($PP.H4 > 0.8$) was taken as strong evidence of colocalization. For MR associations of interest, we first identified genomic regions containing SNPs with P-value $< 1\times10^{-5}$ in

both the exposure (brain IDP) and outcome (breast cancer) datasets. Colocalization analyses were then performed within ± 100 kb of the transcription start sites (TSS) of genes overlapping these regions.

7. Enrichment Analysis

For the significant causal pairs prioritized for colocalization, we performed gene-level functional enrichment analysis. Genes overlapping the genomic regions defined by shared SNPs ($P < 1\times10^{-5}$ in both exposure and outcome) were identified and submitted to the Metascape (https://metascape.org/gp/index.html) platform for enrichment analysis.

8. Statistical analysis

MR analyses were performed using the TwoSampleMR R package (version 0.6.15). The IVW method is employed to assess the causal relationship between IDPs and BC. Colocalization analyses were conducted with the coloc R package (version 5.2.3). All statistical analyses and visualizations were carried out in R version 4.4.1, which is freely available software (https://www.r-project.org/).

Results

1. Brain IDPs collectively exhibit significant causal associations with breast cancer risk

A total of 4,013 brain IDPs were compiled from the GWAS Catalog and the ENIGMA consortium, and MR analyses were conducted across multiple BC GWAS datasets (Supplementary Tables 2 and 3). Overall, brain IDPs demonstrated widespread and statistically significant causal associations with breast cancer risk.

Across the BC datasets derived from three independent sources, numerous IDPs surpassed the significance threshold (*P*-value < 0.05). In the FinnGen dataset (European ancestry), 113 IDPs exhibited causal associations with overall BC. When stratified by subtype, 101 IDPs showed causal relationships with ER-positive BC, whereas 83 IDPs were associated with ER-negative BC.

For the European ancestry pb53 dataset (GWAS Catalog; PMID: 32424353), 169 IDPs demonstrated causal associations with overall BC (GCST010098). Further subtype stratification revealed that the Luminal A subtype (GCST90454345) exhibited the highest number of causal associations (n = 187), followed by the Luminal B subtype (GCST90454346) with 172 associations. In the ieu-a-1127 dataset, representing a meta-analysis of European ancestry populations for ER-positive BC, 158 IDPs showed significant causal effects.

In contrast, the pb14 dataset (GCST90296719; PMID: 38741014), based on an African American or Afro-Caribbean population, demonstrated causal associations with only 47 IDPs, substantially fewer than the

numbers observed in European ancestry datasets. This divergence likely reflects ancestry-specific genetic architectures that influence both brain structural variation and BC susceptibility.

2. ER-positive breast cancer exhibits the strongest causal associations with brain IDPs, with the fusiform gyrus as a key region

We analyzed overall BC and its molecular subtypes as parallel phenotypes (14 datasets in total) to determine whether genetic effects were broad-spectrum or subtype-specific. Both broad-spectrum and subtype-specific risk loci were identified. Cross-subtype comparison in European ancestry populations revealed that most loci in ER-negative and TNBC datasets showed only nominal significance (*P*-value = 0.01–0.05). Applying a stricter threshold (*P*-value < 1×10^{-4}), the ER-positive subtype displayed the most significant and abundant associations, reflecting a more detectable genetic architecture.

In the pb53 dataset (GCST90454345, Luminal A), this subtype had the highest number of significant causal associations across all datasets at the p < 1×10^{-4} level. The top five IDPs were:

a. Average thickness of the right superior temporal sulcus (GCST90003749, *P*-value = 3.44×10^{-8})
b. Volume of the right fusiform gyrus (GCST90002807, *P*-value = 2.23×10^{-7})
c. Surface area of the right fusiform gyrus (GCST90003180, *P*-value = 1.11×10^{-6})
d. Cortical thickness of the left anterior transverse collateral sulcus (GCST90003652, *P*-value = 1.71×10^{-6})
e. Surface area of the left fusiform gyrus (GCST90003079, *P*-value = 5.23×10^{-6})

In the ieu-a-1127 dataset (predominantly ER-positive), the three most significant IDPs (*P*-value < 1.00×10^{-5}) were:

a. Surface area of the left fusiform gyrus (GCST90003079, *P*-value = 4.00×10^{-7}))
b. Surface area of the left H-shaped orbital sulcus (GCST90003360, *P*-value = 1.59×10^{-5}))
c. Volume of the right fusiform gyrus (GCST90002807, *P*-value = 6.78×10^{-5})

In the FinnGen ER-positive dataset, the most significant IDPs included the average thickness of the lateral orbitofrontal cortex (*P*-value = 4.95×10^{-5}), surface area of the right middle temporal gyrus (GCST90003408, *P*-value = 0.00155), and volume of the right lateral ventral diencephalon (GCST90002644, *P*-value = 0.00243). A portion of Broca's area (B45) in the right hemisphere (GCST90003535, *P*-value = 8.08×10^{-6}) showed a Luminal B-specific causal association.

Collectively, these data indicate that the temporal lobe is strongly associated with ER-positive breast cancer, with the bilateral fusiform gyrus consistently demonstrating prominent causal relationships across datasets and subtypes (Brain schematic: http://atlas.brainnetome.org/, as illustrated in Fig. 3.

3. Cortical Area, Volume, and Thickness Exhibit Causal Effects on Breast Cancer in European Populations

Classical IDP metrics, including cortical area and cortical volume, showed broad and substantial causal effects on breast cancer risk across multiple European datasets. Cortical thickness also demonstrated significant causal associations, and in some datasets, evidence suggested potential bidirectional effects. Notably, in ER-positive European datasets, the fusiform gyrus consistently displayed cross-dataset associations, particularly through area- and volume-based IDPs. In addition to these traditional structural parameters, cortical thickness emerged as an independent and meaningful contributor to breast cancer susceptibility.

In the FinnGen ER+ dataset, the strongest association was a negative correlation with lateral orbitofrontal cortex thickness (IDP: Mean_lateralorbitofrontal_thickavg_ 20200522_noGC, *P*-value = 4.95×10^{-5}). The “noGC” annotation indicates that this metric specifically reflects cortical grey-matter thickness rather than gray-white contrast, tissue composition, or signal intensity.

In the pb53 Luminal A dataset, the top association was a negative correlation with right superior temporal sulcus thickness (IDP: GCST90003749, *P*-value = 3.44×10^{-8}). The third-strongest association was a positive correlation with anterior transverse collateral sulcus thickness (IDP: GCST90003652, *P*-value = 1.71×10^{-6}).

For the Luminal B subtype in the pb53 dataset, the strongest causal effect was a negative association with BA45 thickness in Broca’s area (IDP: GCST90003535, *P*-value = 8.08×10^{-6}). This same cortical region also demonstrated negative associations with overall breast cancer (*P*-value = 0.000143) and with the Luminal A subtype (*P*-value = 0.0121), highlighting its potential cross-subtype relevance.

4. Emerging Evidence for Causal Roles of the Thalamus and Diencephalon in Breast Cancer

Our analyses identified significant causal associations between breast cancer and IDPs related to the diencephalon and thalamic nuclei across multiple datasets, including FinnGen (overall breast cancer), pb53 (Luminal B subtype), pb14 (overall breast cancer), and ieu-a-1127.

In the FinnGen dataset, diencephalon volume showed particularly strong associations. For overall breast cancer, right diencephalon volume (IDP: GCST90002644, *P*-value = 4.93×10^{-5}) and left diencephalon volume (IDP: GCST90002627, *P*-value = 6.64×10^{-5}) were the second- and third-ranked associations, respectively. The right diencephalon volume (GCST90002644) also exhibited significant subtype-specific effects, ranking third in ER-positive (*P*-value = 0.00243) and second in ER-negative breast cancer (*P*-value = 0.00106).

In the pb53 dataset (European ancestry), for the Luminal B subtype, left amygdala volume (IDP: GCST90002625, *P*-value = 0.000471) displayed a significant inverse causal relationship with breast cancer. No significant associations were observed for other subcortical structures, including the caudate nucleus, putamen, pallidum, hippocampus, nucleus accumbens, or lentiform nucleus.

5. Population-Specific Genetic Associations between Brain Structure and ER-positive Breast Cancer

To examine cross-ancestry heterogeneity in breast cancer genetics, we analyzed four independent datasets, one African ancestry dataset (pb14) and three European ancestry datasets. The pb14 dataset identified 17 significant brain IDP associations with overall BC. The strongest association in pb14 was the mean fractional anisotropy of the right posterior limb of the internal capsule (GCST90003897, *P*-value = 5.19×10^{-5}). This IDP was also the second-ranked association for ER-positive breast cancer in pb14, whereas pontine volume (GCST90002765, *P*-value = 0.000434) showed the strongest ER-positive association. No significant IDPs were detected for ER-negative or TNBC subtypes in this cohort. Cross-dataset network analysis demonstrated markedly distinct genetic architectures between ancestral groups (Fig. 4). The African ancestry ER-positive dataset (pb14_ERplus) exhibited no overlap (similarity = 0) with any European ER-positive datasets, suggesting ancestry-specific genetic determinants.

In contrast, European datasets showed strong internal similarity. Notably, resting-state fMRI amplitudes of ICA100 nodes 21 and 27 showed inverse causal associations with overall breast cancer in pb14, associations absent in European datasets. These results highlight the importance of incorporating diverse ancestral populations to comprehensively understand disease mechanisms.

6. Genetic Colocalization Analysis Reveals Shared Genetic Architecture Between Breast Cancer and Brain Imaging-Derived Phenotypes

We performed systematic genetic colocalization analyses to clarify the biological mechanisms linking BC risk to brain IDPs. No significant pleiotropic loci were found in the fg_ERneg, pb53_luminalB, pb53_TNBC, pb53_her2, pb14_ERneg, pb14_ERplus, pb_14_TNBC, or ieu_1128_ERneg datasets. In contrast, robust and consistent colocalization signals were detected in the pb53_luminal A and ieu1127 datasets, primarily on chromosomes 17, 11, 12, and 5 (showed in Supplementary Table 3 and Fig. 1).

Chromosome 17 harbored the most prominent and stable pleiotropic hotspot (chr17:45,577,102–46,496,919), as illustrated in Fig. 5 and Supplementary Table 4. Multiple IDPs colocalized within this region. For example, the right fusiform gyrus volume (GCST90002807) showed strong evidence for shared causal variants with genes such as *NSFP1* (*PP.H4* = 0.836) and ENSG00000294196 (*PP.H4* = 0.835). The left fusiform gyrus area (GCST90003709) replicated this core gene set, while the thickness of the S-collat-transv-ant region (GCST90003652) identified four high-confidence genes (*PP.H4* > 0.8)

ENSG00000267198, *ARL17B*, *NSFP1*, and *ENSG00000294196*. Notably, the right superior temporal sulcus (GCST90003749) also demonstrated significant colocalization for ENSG00000267198 and *ARL17B* within a more confined interval (chr17:45,577,102–46,287,283). In analyses of the ieu1128 ER-positive dataset with the left fusiform gyrus area, this hotspot extended to chr17:43,460,181–44,865,603, containing 22 genes with strong colocalization support, including *ARL17A*, *ARL17B*, *NSFP1*, and several long non-coding RNAs. These results highlight pronounced ancestry- and population-specific association patterns.

On chromosome 12 (chr12:28,102,864–28,586,694), colocalization signals were consistently detected for both the left (GCST90003709) and right (GCST90003180) fusiform gyrus areas, as illustrated in Fig. 6. Two genes, *RPL29P27* (a ribosomal pseudogene) and *CCDC91* (coiled-coil domain–containing protein), showed high-confidence colocalization (*PP.H4* > 0.8). *CCDC91* also demonstrated near-significant support (*PP.H4* = 0.774) in the right fusiform analysis, accompanied by several suggestive signals (e.g., *ENSG00000247934*, *IFT57P1*) repeatedly observed across datasets.

Chromosome 11 displayed two distinct patterns, showed in Fig. 6. In the FinnGen ER-positive dataset, colocalization signals within chr11:69,635,537 − 69,639,167 implicated genes such as *CCND1*, *LTO1*, *FGF19*, and *DNAJB6P5* (*PP.H4* = 0.4–0.6). Independently, the right superior temporal sulcus (GCST90003749) showed high-confidence colocalization (*PP.H4* > 0.8) with *FAM76B*, *CEP57*, and *MTMR2L* in chr11:95,758,521 − 95,939,265. On chromosome 5, the lateral orbitofrontal thickness phenotype localized primarily to a single SNP (rs11739135), involving genes including *SLC22A4*, *MIR3936*, and *SLC22A5*, which showed moderate evidence of shared genetic origin (*PP.H4* = 0.4–0.6) .

Based on enrichment analysis (Supplementary Tables 4–5 and Supplementary Fig. 1), the identified loci include key proteins in intracellular vesicular transport. Specifically, *NSFP* facilitates vesicle fusion with target membranes (e.g., presynaptic membranes), while *ARL17B*—an ADP-ribosylation factor-like GTPase—resides in a genomic duplication region with high homology to ARL17A and proximity to the *MAPT* locus. *ARL* proteins typically regulate intracellular membrane trafficking. *CCDC91* participates in retrograde transport from the Golgi to the endoplasmic reticulum. Additionally, non-coding elements such as *ENSG00000294196, ENSG00000267198*, and *RN7SL656P et.al*, suggest that membrane dynamics, assembly, and trafficking are critically involved in this process.

Discussion

This study systematically evaluated the causal links between brain IDPs and BC risk across molecular subtypes using large-scale MR. By integrating 4,013 brain IDPs with 14 BC GWAS datasets from European and African ancestry populations, we reveal population-specific and subtype-specific associations, providing novel insight into the neuro-breast cancer axis.

Our analysis highlights pronounced heterogeneity across ancestries. European datasets formed a highly consistent network of associations, whereas the African ancestry dataset showed virtually no overlap, with white matter microstructure, rather than grey matter, predominating. This pattern is consistent with

well-established population differences in breast cancer epidemiology, subtype distribution, and genetic risk architecture [19, 20]. For example, variants involved in estrogen-metabolism pathways show ancestry-specific effects in Asian women [21]. Women of African ancestry also have a markedly higher incidence of aggressive triple-negative breast cancer, and many risk SNPs discovered in European cohorts display reduced or absent predictive power in African populations [22]. Recent GWAS in African ancestry cohorts continue to identify unique risk loci, reinforcing the presence of ancestry-specific biological mechanisms underlying breast cancer susceptibility [23, 24].

ER-positive (ER+) breast cancer exhibited the most significant genetic links to brain structure, whereas TNBC and HER2-enriched subtypes showed weaker associations. This pattern likely reflects the distinct biology of ER+ tumors: estrogen signaling drives peripheral tumorigenesis [25, 26] and its receptors are widely expressed in the brain, where they regulate neuronal survival, synaptic plasticity, and cognitive function [27, 28].

Consequently, genetic variants affecting estrogen pathways may underlie the observed connections between ER+ breast cancer and specific brain regions [29]. Among these, the temporal lobe emerged as a central hub. Sub-regions including the fusiform gyrus, essential for high-level visual processing and facial recognition [30]; the superior temporal sulcus, critical for social cognition and language [31]; and the middle temporal gyrus, involved in semantic memory [30], collectively form a network supporting perception, memory, and social behavior. Notably, the fusiform gyrus consistently showed the strongest and most stable association with breast cancer risk in European ancestry datasets across multiple sources and brain atlases, suggesting a robust biological link beyond statistical correlation. Traditionally recognized for visual processing [34, 35], it also acts as a multimodal integration hub, converging sensory information [36] and interacting with emotion and memory circuits [37], with connections to limbic structures and the hypothalamus that influence autonomic and endocrine regulation [38]. Its high heritability and experience-dependent plasticity [39] suggest that these neurostructural features may modulate long-term physiological homeostasis and disease susceptibility, including cancer, via the neuro-endocrine-immune axis, linking brain structure to stress response and tumor risk [40].

Associations between the diencephalon and ER-positive (ER+) breast cancer were observed in European ancestry datasets. FinnGen data revealed a significant positive causal relationship between ventral diencephalon volume and overall breast cancer, which was also supported, though ranked slightly lower, by the Luminal B subtype in the GWAS Catalog (pb58) and the ER+ subtype in the ieu-1127 dataset. The ventral diencephalon, as defined in the subcortical segmentation, includes the hypothalamus, a key neuroendocrine regulator. The hypothalamus controls pulsatile gonadotropin-releasing hormone (GnRH) secretion, which governs anterior pituitary release of luteinizing hormone (LH) and follicle-stimulating hormone (FSH), thereby regulating ovarian production of estradiol and progesterone—primary drivers of Luminal breast cancer [41]. It also integrates energy balance, stress, and immune signals [42]. Structural variations in this region may therefore reflect individual differences in neuroendocrine set points,

influencing lifelong hormonal and inflammatory environments and modulating susceptibility to hormone-sensitive breast cancer.

IDPs provide quantitative measures of brain structure and function from macroscopic to microscopic levels [43]. Unlike cortical surface area and volume, cortical thickness showed bidirectional effects on breast cancer risk: increased thickness of the left superior frontal gyrus (SFG) was positively associated, whereas greater thickness of the right BA45 was negatively associated. These patterns may reflect differential influences of distinct neural circuits on peripheral physiology. Cortical thickness changes across development and aging, with childhood/adolescent thickening reflecting synaptogenesis and myelination, early adult thinning reflecting synaptic pruning, and later-life thinning indicating aging or neurodegeneration [40]. Greater SFG thickness, a key node of the executive control network, may relate to cognitive or emotional traits such as rumination or anxiety [44, 45], fostering a subtle pro-inflammatory state that promotes hormone-sensitive breast cancer. Thinning in right BA45, involved in social cognition, may reflect reduced stress-buffering capacity, enhancing sympathetic-adrenal-medullary activation and inflammatory responses [46].

A key finding of this study is the pronounced population heterogeneity in neuro–breast cancer associations. While grey matter structural alterations dominated European ancestry datasets, white matter microstructure emerged as the primary feature in the African ancestry cohort. Specifically, reduced fractional anisotropy (FA) in the right posterior limb of the internal capsule was significantly associated with breast cancer risk, alongside changes in pontine volume, suggesting potential disruptions in sensorimotor integration, autonomic control, and immune regulation. The posterior limb of the internal capsule contains major ascending and descending fibers connecting the cortex with the brainstem and spinal cord, and its integrity is essential for autonomic function. Disrupted white matter may therefore perturb autonomic output, altering peripheral immune homeostasis and fostering a tumor-promoting microenvironment [47]. Consistent with this, reduced FA is an early marker of Alzheimer's disease, reflecting impaired axonal transport, myelin loss, and Wallerian degeneration [48, 49], and is also observed in depression, where chronic stress and HPA axis overactivation impair oligodendrocyte function, causing dysmyelination [50]. These population-specific differences likely arise from a combination of genetic background, socio-environmental factors, and neurodevelopmental influences, highlighting the necessity of including diverse cohorts in future studies.

To investigate genetic mechanisms linking brain structure and breast cancer, we performed systematic colocalization analyses. The most significant pleiotropic signals were enriched on chromosomes 17, 11, and 12, which harbor key functional genes: TP53 and BRCA1 as tumor suppressors and MAPT in neurology on chr17 [51, 52, 19, 53]; CCND1 regulating the cell cycle on chr11 [54]; and TBX3 involved in development and cancer on chr12 [55], indicating localization to functionally enriched genomic regions.

Specific genes included NSF and ARL17B on chr17, CCDC91 on chr12, and CEP57 on chr11, collectively highlighting a coordinated "cellular logistics system" for intracellular trafficking and structural

maintenance [56]. KEGG enrichment confirmed involvement in “Endocytosis” and “Protein processing in endoplasmic reticulum.”

Notably, many hotspots contained long non-coding RNAs (lncRNAs), e.g., ENSG00000294196, consistent with their established role in complex trait susceptibility [57–59], analogous to BACE1-AS in Alzheimer’s [60] and XIST in cancer [61]. These lncRNAs may regulate protein-coding networks via ceRNA or chromatin modification.

IDPs were derived from multiple parcellation schemes (DK, DKT, Destrieux); despite anatomical variations, large-scale GWAS show genetic effects on brain structure are macroscopic, yielding consistent signals across parcellations [62]. For example, fusiform gyrus associations were robustly detected under both DK and DKT, confirming reproducibility.

In summary, our study identifies population-specific genetic links between brain structure and breast cancer, particularly ER+ subtypes. These findings provide insight into the brain-body axis in cancer and offer a foundation for novel biomarkers and preventive strategies.

Declarations

Ethics approval and consent to participate

This study utilized publicly available summary data from genome-wide association studies. All original studies obtained ethical approval from their respective institutional review boards and informed consent from all participants.

Consent for publication

Not applicable. (This study does not contain any individual person's data in any form.)

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

This work was supported by the National Natural Science Foundation of China (Grant Nos. 82474637; Nos. 52076213), the Beijing Traditional Chinese Medicine Science and Technology Development Fund Project (Grant Nos. BJZYYB-2023-45),and the Innovation Cultivating Foundation of The Sixth Medical Center of Chinese PLA General Hospital (Grant Nos: CXPY202007).

Author Contribution

Wei Ma and Jing Li conceived and designed the study. Jing Li provided the data used for the analyses. Wei Ma performed the data analysis and interpretation. Jing Li and Jun Chen processed the data and drafted the manuscript. Mingyue Hao contributed to part of the data analysis and visualization. Fei Ji and Xue Li were responsible for revising the manuscript and preparing additional figures.

Acknowledgement

We want to acknowledge the participants and investigators of the FinnGen study.We also thank the participants and researchers of the UK Biobank, the ENIGMA consortium, the IEU OpenGWAS project, and the GWAS Catalog for making their summary statistics publicly available.

Data Availability

All data used in this study were obtained from publicly available sources.

Data Availability Statement

All data used in this study were obtained from publicly available sources.

References

1. Sung, H. et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **71** (3), 209–249 (2021).
2. Kim, J. et al. Global patterns and trends in breast cancer incidence and mortality across 185 countries. *Nat. Med.* **31** (4), 1154–1162 (2025).
3. Sun, K. et al. Global patterns and trends in cancer-related premature death and their impact on life expectancy across 185 countries: a population-based analysis. *Mil Med. Res.* **12** (1), 56 (2025).
4. Sun, K. et al. Trends in Incidence Rates, Mortality Rates, and Age-Period-Cohort Effects of Female Breast Cancer - China, 2003–2017. *China CDC Wkly.* **5** (15), 340–346 (2023).
5. Kuchenbaecker, K. B. et al. Risks of Breast, Ovarian, and Contralateral Breast Cancer for BRCA1 and BRCA2 Mutation Carriers. *JAMA* **317** (23), 2402–2416 (2017).
6. Nunes, G. G. C. et al. Genomic Variants and Worldwide Epidemiology of Breast Cancer: A Genome-Wide Association Studies Correlation Analysis. *Genes (Basel)* **15**(2). (2024).
7. Iacoviello, L., Bonaccio, M., de Gaetano, G. & Donati, M. B. Epidemiology of breast cancer, a paradigm of the common soil hypothesis. *Semin Cancer Biol.* **72**, 4–10 (2021).
8. Hessami, A. et al. Depression, anxiety, and stress in breast cancer patients: prevalence, associated risk factors, and clinical correlates. *Sci. Rep.* **15** (1), 31084 (2025).

9. Chen, J., Li, J., Qiao, H., Hu, R. & Li, C. Disruption of IDO signaling pathway alleviates chronic unpredictable mild stress-induced depression-like behaviors and tumor progression in mice with breast cancer. *Cytokine* **162**, 156115 (2023).
10. Reiche, E. M. V., Nunes, S. O. V. & Morimoto, H. K. Stress, depression, the immune system, and cancer. *Lancet Oncol.* **5** (10), 617–625 (2004).
11. Qin, J. et al. Adrenergic receptor β2 activation by stress promotes breast cancer progression through macrophages M2 polarization in tumor microenvironment. *BMB Rep.* **48** (5), 295–300 (2015).
12. Lamkin, D. M. et al. α2-Adrenergic blockade mimics the enhancing effect of chronic stress on breast cancer progression. *Psychoneuroendocrinology* **51**, 262–270 (2015).
13. Chen, H. et al. Functional Connectivity Alterations Associated with COVID-19-Related Sleep Problems: A Longitudinal Resting-State fMRI Study. *Nat. Sci. Sleep.* **17**, 97–113 (2025).
14. Sarvandani, M. N. et al. COVID-19 and the Brain: A Psychological and Resting-state Functional Magnetic Resonance Imagin (fMRI) Study of the Whole-brain Functional Connectivity. *Basic. Clin. Neurosci.* **14** (6), 753–771 (2023).
15. Zheng, Y. et al. α-Synuclein Promotes the Expression and Release of Neurofilament Light Chain from Oligodendrocytes in a Murine Model: Implication for Multiple System Atrophy. *Ann. Neurol.* (2025).
16. Giannini, G. et al. Early onset sleep disorders predict severity, progression and death in multiple system atrophy. *J. Neurol.* **272** (3), 239 (2025).
17. Zhang, J. et al. Clinically Relevant Patterns of Co-Fluctuating Structure and Function in Multiple Sclerosis. *Eur. J. Neurol.* **32** (10), e70367 (2025).
18. Ammitzbøll, C. et al. NfL and GFAP in serum are associated with microstructural brain damage in progressive multiple sclerosis. *Mult Scler. Relat. Disord*. **77**, 104854 (2023).
19. Winer, E. et al. Clinical cancer advances 2008: major research advances in cancer treatment, prevention, and screening–a report from the American Society of Clinical Oncology. *J. Clin. Oncol.* **27** (5), 812–826 (2009).
20. Sung, H. et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **71** (3), 209–249 (2021).
21. Zhang, H. et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat. Genet.* **52** (6), 572–581 (2020).
22. Zavala, V. A. et al. Cancer health disparities in racial/ethnic minorities in the United States. *Br. J. Cancer*. **124** (2), 315–332 (2021).
23. Roberts, E. et al. Breast cancer polygenic risk scores derived in White European populations are not calibrated for women of Ashkenazi Jewish descent. *Genet. Med.* **25** (10), 100963 (2023).
24. Borde, J. et al. Performance of Breast Cancer Polygenic Risk Scores in 760 Female CHEK2 Germline Mutation Carriers. *J. Natl. Cancer Inst.* **113** (7), 893–899 (2021).

25. Luo, A. & Zhang, X. ERRF is essential for Estrogen-Estrogen Receptor alpha signaling pathway in ER positive breast cancer cells. *Biochem. Biophys. Res. Commun.* **474** (2), 400–405 (2016).
26. Guo, P. et al. Estrogen-induced interaction between KLF5 and estrogen receptor (ER) suppresses the function of ER in ER-positive breast cancer cells. *Int. J. Cancer*. **126** (1), 81–89 (2010).
27. Hill, J. W. & Elias, C. F. Neuroanatomical Framework of the Metabolic Control of Reproduction. *Physiol. Rev.* **98** (4), 2349–2380 (2018).
28. Li, Z. et al. EstroGene database reveals diverse temporal, context-dependent and directional estrogen receptor regulomes in breast cancer. bioRxiv. (2023).
29. Stacey, S. N. et al. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat. Genet.* **40** (6), 703–706 (2008).
30. Rossion, B., Jacques, C. & Jonas, J. The anterior fusiform gyrus: The ghost in the cortical face machine. *Neurosci. Biobehav Rev.* **158**, 105535 (2024).
31. Zhang, R. et al. Reduced grey matter volume in adolescents with conduct disorder: a region-of-interest analysis using multivariate generalized linear modeling. *Discov Ment Health*. **3** (1), 25 (2023).
32. Moran, C. et al. Brain atrophy in type 2 diabetes: regional distribution and influence on cognition. *Diabetes Care*. **36** (12), 4036–4042 (2013).
33. Kamarck, T. W. et al. Exaggerated blood pressure responses during mental stress are associated with enhanced carotid atherosclerosis in middle-aged Finnish men: findings from the Kuopio Ischemic Heart Disease Study. *Circulation* **96** (11), 3842–3848 (1997).
34. Kanwisher, N., McDermott, J. & Chun, M. M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17** (11), 4302–4311 (1997).
35. Grill-Spector, K., Weiner, K. S., Kay, K. & Gomez, J. The Functional Neuroanatomy of Human Face Perception. *Annu. Rev. Vis. Sci.* **3**, 167–196 (2017).
36. Beauchamp, M. S. See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Curr. Opin. Neurobiol.* **15** (2), 145–153 (2005).
37. Vuilleumier, P. & Pourtois, G. Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia* **45** (1), 174–194 (2007).
38. Critchley, H. D. & Harrison, N. A. Visceral influences on brain and behavior. *Neuron* **77** (4), 624–638 (2013).
39. McCandliss, B. D., Cohen, L. & Dehaene, S. The visual word form area: expertise for reading in the fusiform gyrus. *Trends Cogn. Sci.* **7** (7), 293–299 (2003).
40. Dhabhar, F. S. Effects of stress on immune function: the good, the bad, and the beautiful. *Immunol. Res.* **58** (2–3), 193–210 (2014).
41. Boehm, U. et al. Expert consensus document: European Consensus Statement on congenital hypogonadotropic hypogonadism–pathogenesis, diagnosis and treatment. *Nat. Rev. Endocrinol.* **11** (9), 547–564 (2015).

42. Morton, G. J. & Schwartz, M. W. Leptin and the central nervous system control of glucose metabolism. *Physiol. Rev.* **91** (2), 389–411 (2011).
43. Miller, K. L. et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19** (11), 1523–1536 (2016).
44. Seeley, W. W. et al. Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* **27** (9), 2349–2356 (2007).
45. Dixon, M. L. et al. Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks. *Proc. Natl. Acad. Sci. U S A*. **115** (7), E1598–E1607 (2018).
46. Hamzei, F. et al. The Dual-Loop Model and the Human Mirror Neuron System: an Exploratory Combined fMRI and DTI Study of the Inferior Frontal Gyrus. *Cereb. Cortex*. **26** (5), 2215–2224 (2016).
47. MariPaz, L., Gabriel, T. I. & de Gonzalo, S. M. D. Immune checkpoint-based biomarkers for therapeutic response in patients with multiple sclerosis. *Front. Immunol.* **16**, 1694021 (2025).
48. Oishi, K. et al. Critical role of the right uncinate fasciculus in emotional empathy. *Ann. Neurol.* **77** (1), 68–74 (2015).
49. Phan, T. X. et al. Increased Cortical Thickness in Alzheimer's Disease. *Ann. Neurol.* **95** (5), 929–940 (2024).
50. Liao, Y. et al. Is depression a disconnection syndrome? Meta-analysis of diffusion tensor imaging studies in patients with MDD. *J. Psychiatry Neurosci.* **38** (1), 49–56 (2013).
51. Vasconcelos, D. S. et al. Numerical aberrations of chromosome 17 and TP53 in brain metastases derived from breast cancer. *Genet. Mol. Res.* **12** (3), 2594–2600 (2013).
52. Orsetti, B. et al. Genomic and expression profiling of chromosome 17 in breast cancer reveals complex patterns of alterations and novel candidate genes. *Cancer Res.* **64** (18), 6453–6460 (2004).
53. Chia, S. et al. Single-nucleus transcriptomics reveals a distinct microglial state and increased MSR1-mediated phagocytosis as common features across dementia subtypes. *Genome Med.* **17** (1), 92 (2025).
54. Li, B. & Nabavi, S. A multimodal graph neural network framework for cancer molecular subtype classification. *BMC Bioinform.* **25** (1), 27 (2024).
55. Andersen, R. E. et al. Chromosomal structural rearrangements implicate long non-coding RNAs in rare germline disorders. *Hum. Genet.* **143** (7), 921–938 (2024).
56. Bonsergent, E. et al. Quantitative characterization of extracellular vesicle uptake and content delivery within mammalian cells. *Nat. Commun.* **12** (1), 1864 (2021).
57. He, F. et al. Integrative Analysis of Somatic Mutations in Non-coding Regions Altering RNA Secondary Structures in Cancer Genomes. *Sci. Rep.* **9** (1), 8205 (2019).
58. Vangoor, V. R., Gomes-Duarte, A. & Pasterkamp, R. J. Long non-coding RNAs in motor neuron development and disease. *J. Neurochem*. **156** (6), 777–801 (2021).
59. Chen, K. W. & Chen, J. A. Functional Roles of Long Non-coding RNAs in Motor Neuron Development and Disease. *J. Biomed. Sci.* **27** (1), 38 (2020).

60. Zhou, Y. et al. LncRNA BACE1-AS Promotes Autophagy-Mediated Neuronal Damage Through The miR-214-3p/ATG5 Signalling Axis In Alzheimer's Disease. *Neuroscience* **455**, 52–64 (2021).
61. Cook, B., Nayar, S., Filson, S. & Yap, T. The incidence of male breast cancer in Klinefelter Syndrome and its proposed mechanisms. *Breast* **78**, 103827 (2024).
62. Morita, T., Asada, M. & Naito, E. Contribution of Neuroimaging Studies to Understanding Development of Human Cognitive Brain Functions. *Front. Hum. Neurosci.* **10**, 464 (2016).

Figures

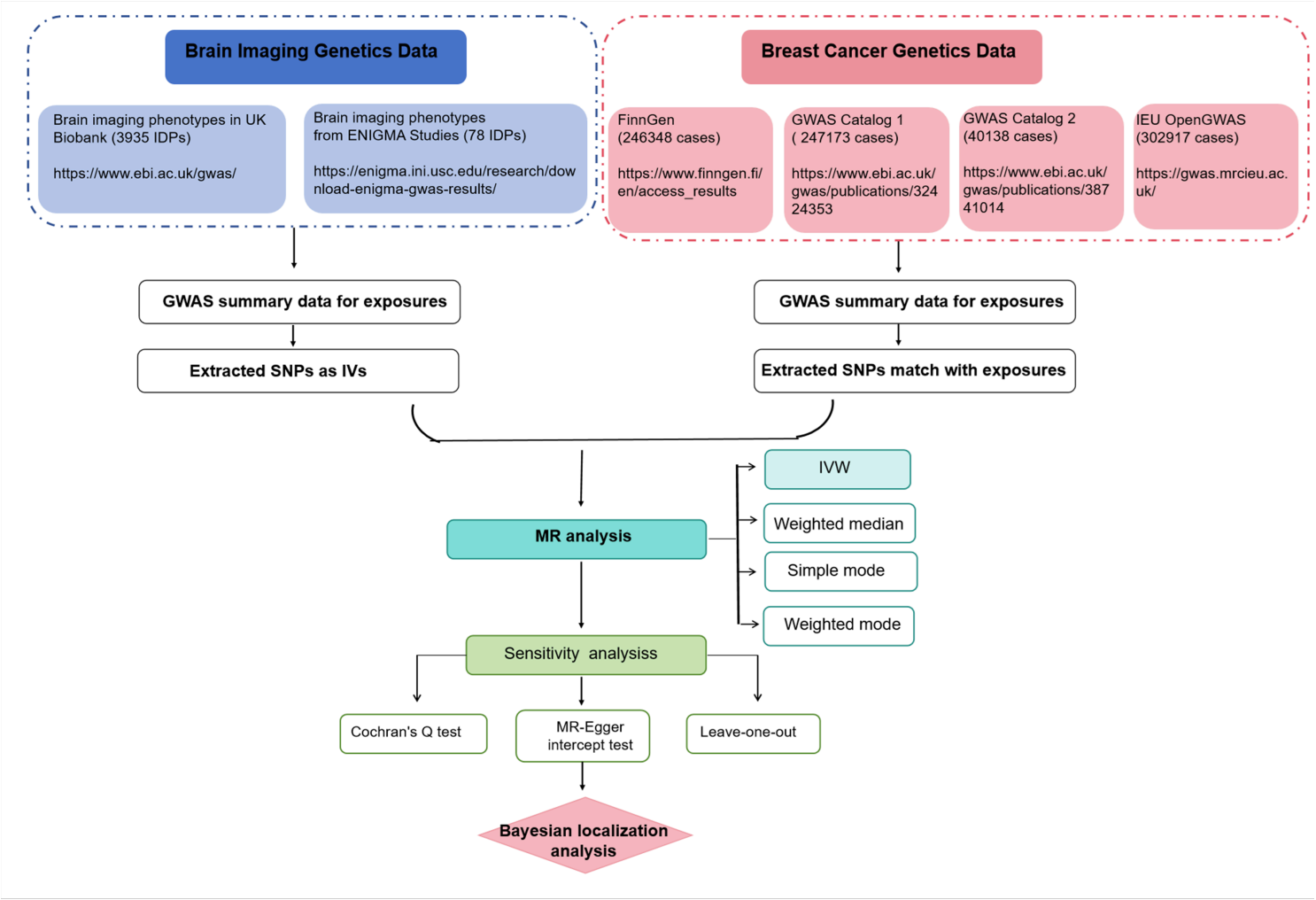


Figure 1

Study design overview.
Schematic workflow illustrating the systematic screening of brain imaging-derived phenotypes (IDPs) as potential biomarkers for breast cancer (BC), and the characterization of brain IDP profile changes in response to BC development. The figure outlines data sources, analysis pipeline, and methodological steps.

A

Source	IDPs $_{All}$	IDPs $_{p<0.05}$
FinnGen All	726	113
FinnGen ER+	733	101
FinnGen ER-	735	83
Pb53 All	759	169
Pb53 luminal A	753	187
Pb53 luminal B	756	172
pb53 TNBC	756	101
pb53 HER2rich	756	55
pb14 BC All	719	47
pb14 ER+	719	45
pb14 ER-	720	46
pb14 TNBC	719	70
ieu1127 ER+	730	158
ieu1127 ER-	731	88

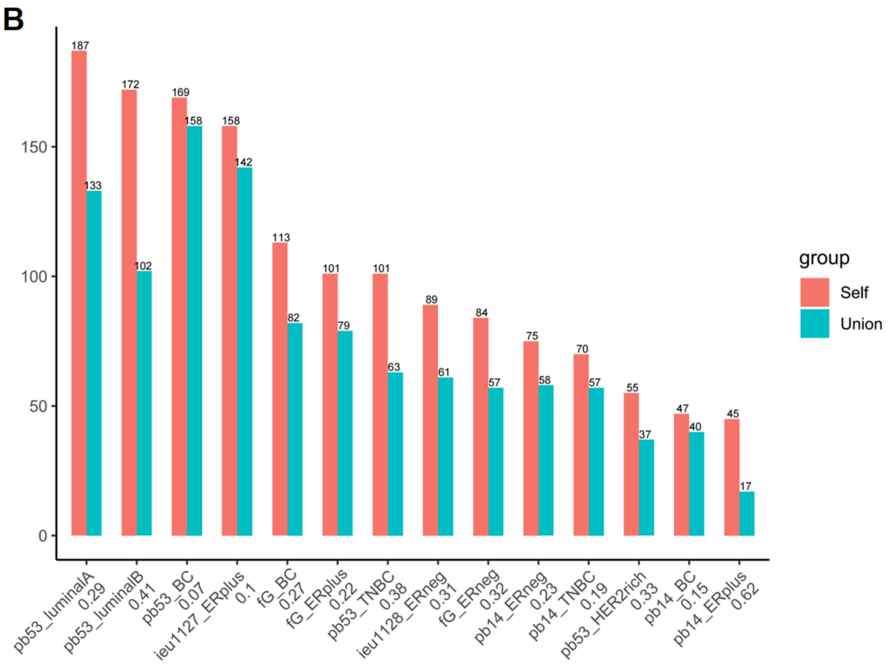


Figure 2

Systematic screening of causal associations between brain IDPs and breast cancer subtypes.
(A) Tabular summary of MR results across 14 BC subtypes. Each row represents a subtype, showing the total number of tested brain IDPs and the number with significant causal associations (P < 0.05). Subtypes are color-coded by data source to distinguish cohorts.
(B) Specificity analysis of significant IDPs for each BC subtype (red). Specificity was calculated as 1 minus the proportion of a subtype's significant IDPs overlapping with the union of significant IDPs from the other 13 subtypes (blue). Subtypes are ordered from left to right by the absolute number of significant IDPs.

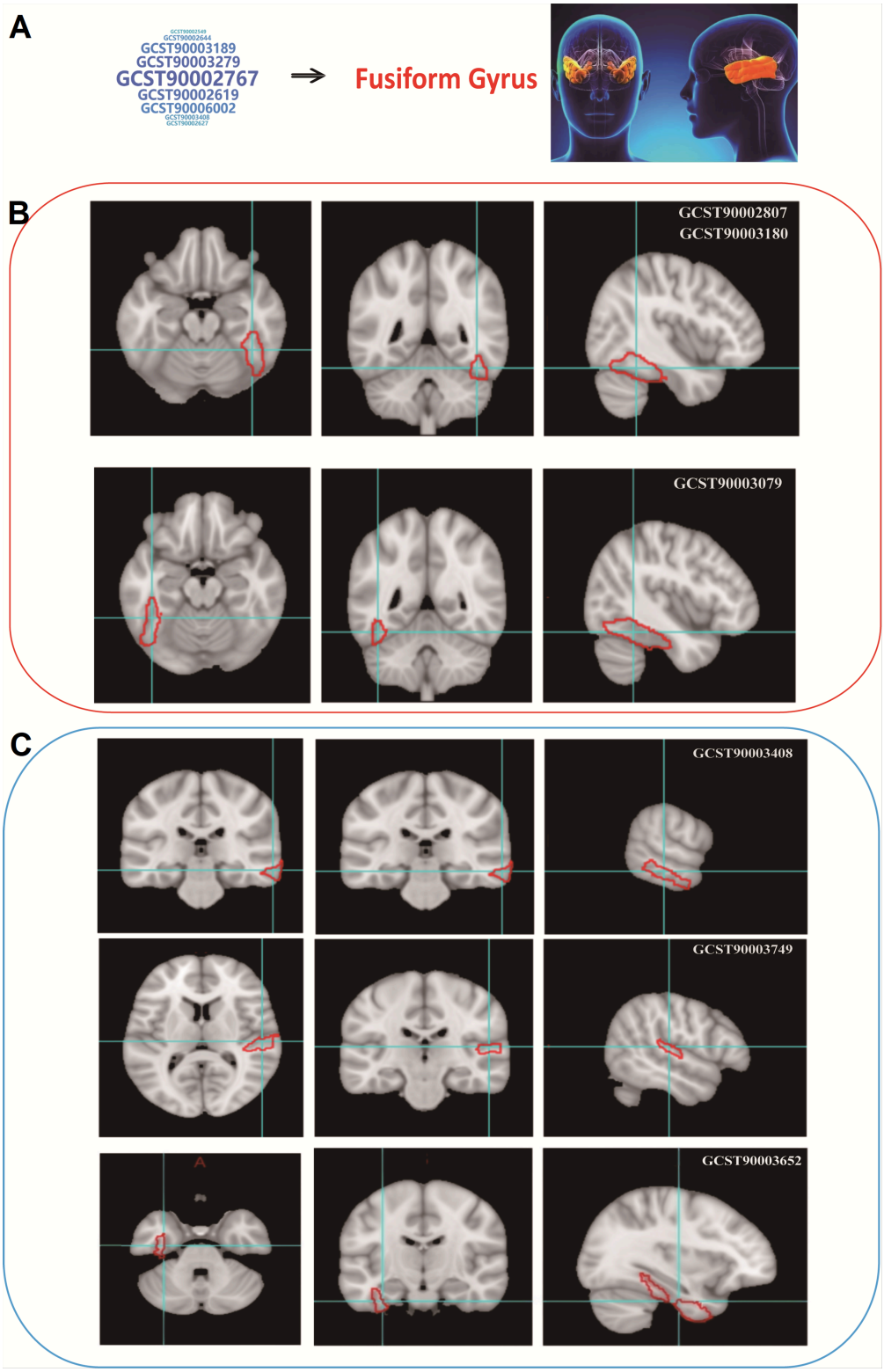


Figure 3

The temporal lobe, particularly the fusiform gyrus, as a core hub for ER-positive breast cancer.
(A) Word cloud showing the most frequently occurring IDPs across all significant associations. The size of each term corresponds to its frequency, with GCST90002767 (fusiform gyrus) being most prominent, highlighting its central role.
(B) Fine-grained mapping localizes the strongest causal signal to the fusiform gyrus in both

hemispheres.
(C) Additional temporal lobe regions contribute to this hub, reinforcing its centrality in the network. Brain renderings (B, C) are based on the Brainnetome Atlas (http://atlas.brainnetome.org/)

fg_BC	fg_ERneg	fg_ERplus	pb53_BC	pb53_luminalA	pb53_luminalB	pb53_TNBC	pb53_her-2	pb 14 BC	pb14_ERneg	pb14_ERplus	pb14_TNBC	ieu1127_ERplus	pval.ieu1128_ERneg
Mean_lateralorbitofrontal_thickavg_20200522_noGC	GCST90002768	Mean_lateralorbitofrontal_thickavg_20200522_noGC	GCST90003180	GCST90003749	GCST90003535	GCST90003065	Mean_inferiorparietal_thickavg_20200522_noGC	GCST90003897	GCST90004471	GCST90002765	GCST90002724	GCST90003079	Mean_middletemporal_thickavg_20200522_noGC
GCST90002644	GCST90002644	GCST90003408	GCST90003079	GCST90002807	GCST90003322	GCST90004229	GCST90003098	GCST90004615	GCST90004609	GCST90003897	GCST90005893	GCST90003360	Mean_fusiform_thickavg_20200522_noGC
GCST90002627	GCST90002602	GCST90002644	GCST90003556	GCST90003180	GCST90003263	GCST90003448	Mean_middletemporal_thickavg_20200522_noGC	GCST90004609	GCST90002845	GCST90002768	GCST90003121	GCST90002807	GCST90002872
GCST90002767	GCST90003523	GCST90002767	GCST90002872	GCST90003652	GCST90003479	GCST90003542	GCST90003571	GCST90003278	GCST90002893	GCST90002450	GCST90002549	Mean_inferiortemporal_thickavg_20200522_noGC	GCST90003714
GCST90003408	GCST90002627	GCST90002627	GCST90002807	GCST90003079	GCST90003218	GCST90003320	Mean_inferiortemporal_thickavg_20200522_noGC	GCST90002765	GCST90003121	GCST90002859	GCST90005814	GCST90002725	GCST90003263
GCST90002549	GCST90002562	GCST90006345	GCST90003472	Mean_inferiortemporal_thickavg_20200522_noGC	GCST90002625	GCST90004465	GCST90004514	GCST90002638	GCST90004615	GCST90004544	GCST90003972	GCST90003503	GCST90003103
GCST90002619	GCST90003145	GCST90003952	GCST90003293	GCST90004045	GCST90003180	GCST90004332	GCST90004551	GCST90002768	GCST90004633	GCST90004075	GCST90006342	GCST90003594	GCST90003065
GCST90003279	GCST90003169	GCST90002619	GCST90003726	GCST90003950	GCST90002716	GCST90003935	GCST90003899	GCST90004397	GCST90003922	Mean_parsopercularis_surfavg_20190429_noGC	GCST90004615	GCST90003589	GCST90003322
GCST90003189	GCST90003084	GCST90003189	GCST90003535	GCST90003927	GCST90003227	GCST90003257	GCST90004056	GCST90002859	GCST90003434	GCST90003129	GCST90003878	GCST90003556	GCST90002707
GCST90006002	GCST90003077	GCST90003538	GCST90003475	GCST90002443	Mean_fusiform_surfavg_20190429_noGC	GCST90003097	GCST90003972	GCST90002723	GCST90004581	GCST90002602	Mean_supramarginal_surfavg_20190429_noGC	GCST90004115	GCST90002633

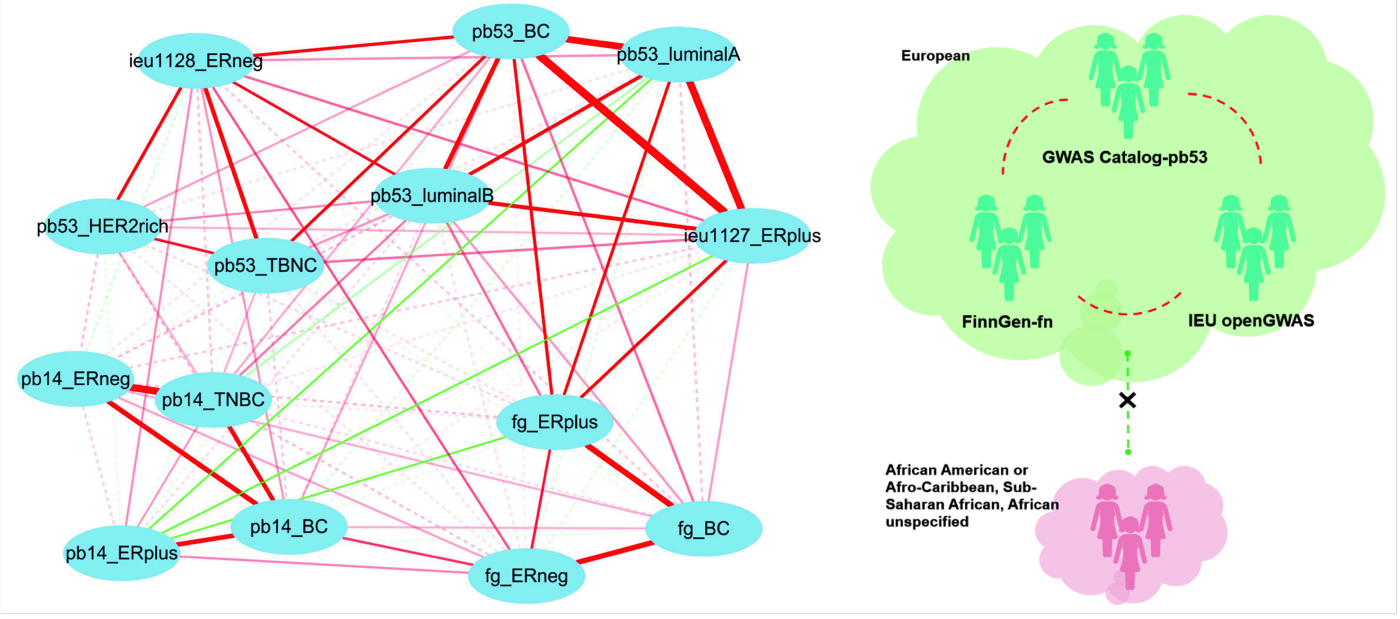


Figure 4

Population-specific landscapes of causal brain IDPs across breast cancer subtypes.

(A) Top 10 most significant IDPs for each of the 14 BC subtypes. IDPs unique to, or originating from, the African ancestry cohort (pb14) are highlighted in pink, showing the distinct neuroimaging genetic profile of this population.

(B) Network analysis depicting similarity in causal IDP signatures. Node color indicates population/data source. Solid red edges represent statistically significant similarity (FDR < 0.1), with thickness and color intensity proportional to correlation strength. Dashed green edges indicate significant dissimilarity, with thickness proportional to divergence. European ER+ datasets form a tightly connected network, whereas the African ancestry cohort shows no correlation (similarity = 0) with European cohorts.

(C) Conceptual illustration of divergent genetic architectures. The green cloud represents overlapping IDPs among European GWAS datasets; the pink cloud represents African ancestry BC data, showing no significant overlap, indicative of distinct genetic mechanisms.

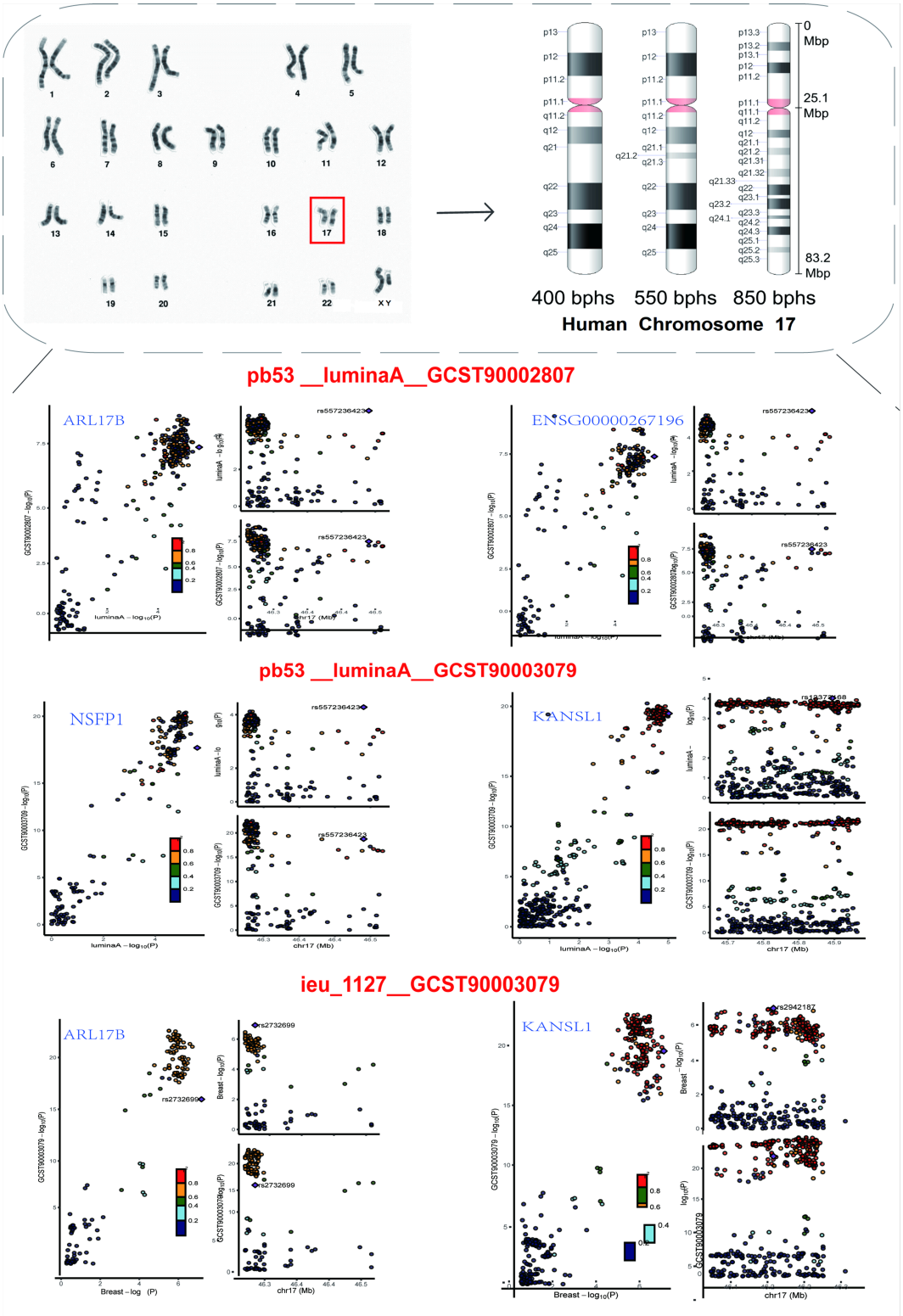


Figure 5

Genetic convergence on chromosome 17 underpins the brain-ER+ BC axis.

Illustration of a pleiotropic hotspot on chromosome 17 acting as shared genetic ground for variations in brain structure and susceptibility to ER+ breast cancer. Scatter plot shows SNP effects on brain IDPs and ER+ BC, with the cloud of red and green points forming a diagonal trend, indicating proportional genetic

effects likely driven by the same causal variants. The most significant SNPs cluster at the upper end of the trend, pinpointing the location of shared genetic influence.

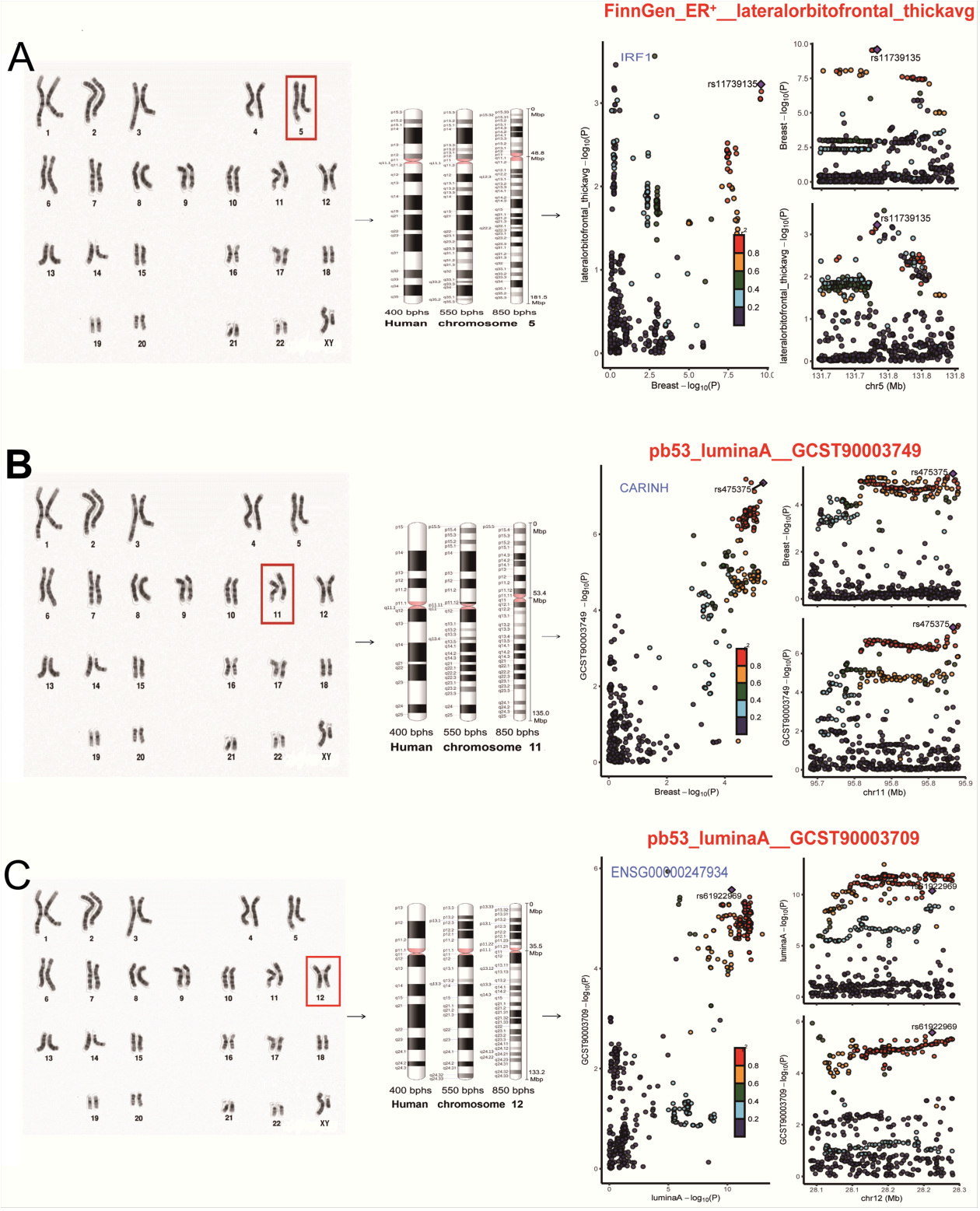


Figure 6

Polygenic network connecting brain structure to ER+ BC risk.

Colocalization scatter plots highlight significant shared genetic signals on (A) chr5, (B) chr11, and (C) chr12. Red and green points indicate SNPs with evidence of a shared genetic basis, with the top-ranking SNPs representing the primary drivers of pleiotropy linking brain structure to ER+ BC susceptibility.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarytables.xlsx
- Supplementaryfigure1.tif
- Supplementarymaterialslegend.docx