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.

Distinct transcriptomic programs and proximal AKT activation associated with AKT1 E17K and PIK3CA hotspot mutations in HR-positive/HER2-negative breast cancer Running title: Molecular differences between AKT1- and PIK3CA-mutant breast cancer

Qin Zhang
Cangzhou Central Hospital
Hong Liu
liuhong_submit@tjmuch.com

Tianjin Medical University

Research Article

Keywords: AKT1 E17K, PIK3CA, breast cancer, HR-positive/HER2-negative, RPPA, PI3K-AKT signaling

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11331642/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

PIK3CA and AKT1 alterations are commonly grouped within the therapeutically relevant PI3K-AKT pathway, but whether they represent equivalent breast-cancer states remains uncertain. We compared molecular features associated with AKT1 E17K and recurrent PIK3CA hotspots.

Methods

METABRIC was the discovery cohort and TCGA-BRCA the independent validation cohort. Prespecified hotspots were AKT1 E17K and PIK3CA E542K, E545K, H1047R, or H1047L. Primary transcriptomic analyses were restricted to HR-positive/HER2-negative tumors and adjusted for age and PAM50 subtype. Differential expression, Hallmark/Reactome enrichment, genomic context, and TCGA reverse-phase protein array data were analyzed.

Results

AKT1 hotspots occurred in 79 of 2,433 METABRIC tumors (3.2%) and 19 of 816 TCGA tumors (2.3%); PIK3CA hotspots occurred in 733 (30.1%) and 210 (25.7%), respectively. The alterations were largely mutually exclusive. In adjusted HR-positive/HER2-negative analyses, 341 genes differed between AKT1-mutant and PIK3CA-mutant-only METABRIC tumors; 280 showed concordant TCGA directions and 53 achieved same-direction nominal replication. Nine pathways achieved FDR-controlled TCGA replication, including higher G2M, glycolysis, chromosome-segregation, and ER-Golgi transport and lower translation-elongation and nonsense-mediated-decay programs in AKT1-mutant tumors. TCGA RPPA showed a higher AKT activation score (0.98 SD; FDR = 0.0025), AKT pS473 (1.00 SD; FDR = 0.0107), and pT308 (0.90 SD; FDR = 0.0166), without differences in pan-AKT, aggregate AKT-substrate, or mTOR-effector scores.

Conclusions

AKT1 E17K and PIK3CA hotspots are associated with overlapping but non-equivalent molecular states. AKT1-mutant tumors show distinct transcriptional programs and preferential proximal AKT phosphorylation without uniform downstream mTOR activation. Confirmation in treatment-annotated and experimental cohorts is required.

Introduction

The PI3K-AKT pathway coordinates cell growth, survival, metabolism, and therapeutic response and is among the most frequently altered signaling systems in breast cancer. PIK3CA encodes the catalytic

p110-alpha subunit of class I phosphatidylinositol 3-kinase, whereas AKT1 encodes a central serine-threonine kinase activated downstream of PI3K. Recurrent PIK3CA mutations cluster in the helical and kinase domains, particularly at E542, E545, and H1047, while the transforming AKT1 E17K mutation alters the pleckstrin homology domain and promotes membrane localization [1–3].

These alterations have increasing clinical relevance in hormone receptor-positive (HR-positive), HER2-negative advanced breast cancer. PI3K-alpha inhibition improves outcomes in appropriately selected PIK3CA-mutant disease, and the AKT inhibitor capivasertib has demonstrated benefit in tumors with qualifying PIK3CA, AKT1, or PTEN alterations [15, 16]. This shared therapeutic framework can imply that activating PIK3CA and AKT1 mutations are interchangeable markers of pathway activation. However, the mutations occur at different positions in the signaling hierarchy and may produce different amplitudes, feedback responses, and downstream transcriptional states.

Large genomic studies have established the frequency, hotspot distribution, clinicopathologic associations, and near-mutual exclusivity of PIK3CA and AKT1 alterations [1, 4, 5]. A recent clinicogenomic analysis of more than 50,000 breast tumors further defined the diversity and clinical relevance of PIK3CA, AKT1, and PTEN variants [1]. Nevertheless, direct cross-cohort comparisons of the transcriptomic programs and protein signaling states associated with AKT1 E17K and recurrent PIK3CA hotspots remain limited. This distinction is important because genomic grouping alone may obscure mutation-specific biology.

We therefore performed a retrospective multi-omic analysis using METABRIC as the primary discovery cohort and TCGA-BRCA as an independent validation cohort. We first characterized hotspot frequencies, mutual exclusivity, and clinicopathologic distributions. We then compared AKT1-mutant and PIK3CA-mutant-only tumors within the clinically relevant HR-positive/HER2-negative population using adjusted transcriptomic models, pathway-level replication, genomic-context analyses, and TCGA reverse-phase protein array (RPPA) data. We hypothesized that the two mutation groups would be associated with distinguishable downstream molecular programs despite converging on the PI3K-AKT pathway.

Methods

Study design and data sources

This retrospective study used publicly available METABRIC and TCGA-BRCA data downloaded in cBioPortal study format [4-7]. METABRIC served as the primary discovery cohort because of its larger number of AKT1-mutant tumors and long-term clinical follow-up. TCGA-BRCA served as an independent cohort for transcriptomic, pathway, genomic, and RPPA analyses. Raw downloaded files were treated as read-only inputs, and all derived tables were generated through versioned R scripts.

The initial clinical datasets contained 2,509 METABRIC samples and 818 TCGA samples representing 817 patients. Expression matrices contained 1,980 METABRIC and 817 TCGA samples. Analyses were restricted to primary tumors with evaluable mutation status. Sample identifiers, patient identifiers,

missing values, and clinical categories were harmonized before analysis. Duplicate identifiers or unexpected sample counts caused the pipeline to stop. Detailed sample flow, exclusions, and mutation-evaluable denominators are provided in Table S1.

Mutation definitions and analysis groups

The prespecified AKT1 hotspot was p.E17K. Prespecified PIK3CA hotspots were p.E542K, p.E545K, p.H1047R, and p.H1047L. Variants were harmonized at the protein-change level. Four initial categories were defined: AKT1 hotspot only, PIK3CA hotspot only, co-mutated, and pair-wild-type. For the three-group clinicopathologic and molecular analyses, tumors carrying AKT1 E17K were assigned to the AKT1-mutant group, including the single AKT1/PIK3CA co-mutant TCGA tumor; tumors with a PIK3CA hotspot but no AKT1 hotspot were assigned to the PIK3CA-mutant-only group; and tumors without either hotspot were assigned to the pair-wild-type group.

METABRIC pair-wild-type classification required documented panel coverage of both AKT1 and PIK3CA. TCGA pair-wild-type classification was based on the exome-profiled cohort. Hotspot frequencies were reported with Wilson 95% confidence intervals. Mutual exclusivity was assessed using two-sided Fisher exact tests with odds ratios and 95% confidence intervals. Cohorts were analyzed separately and were not pooled. Complete hotspot counts, group denominators, and mutual-exclusivity statistics are reported in Table S1.

Clinical variable harmonization and Table 1

Estrogen receptor (ER), progesterone receptor (PR), HER2, HR/HER2 category, PAM50 subtype, histologic type, grade, stage, age, menopausal status, tumor size, and nodal variables were harmonized where available. METABRIC PAM50-like categories were reconstructed from the supplied CLAUDIN_SUBTYPE field. For TCGA, PAM50 annotations were reconstructed from available study case lists; samples not covered by those annotations were labeled unknown. HER2 status in TCGA prioritized FISH results and used immunohistochemistry when FISH was unavailable.

Table 1 was constructed separately for each cohort across the AKT1-mutant, PIK3CA-mutant-only, and pair-wild-type groups. Continuous variables were summarized using medians and interquartile ranges and compared with Kruskal-Wallis tests. Categorical variables were summarized as counts and percentages and compared using Pearson chi-square or Fisher exact tests, with Monte Carlo estimation for sparse multi-category tables. Percentages were calculated among non-missing observations. The main table presents the prespecified core clinicopathologic characteristics, whereas complete cohort-specific distributions and missingness are reported in Table S2.

Expression preprocessing and quality control

Downloaded normalized Illumina microarray expression values were used for METABRIC. Duplicate gene symbols were collapsed using the limma average-replicate procedure, resulting in 20,374 genes. TCGA RSEM values were filtered to retain genes with raw expression greater than 1 in at least 10% of samples and were transformed as log2(RSEM+1), resulting in 18,089 genes. Principal-component analysis used the 1,000 most variable genes within each cohort and analysis population. Expression-based outlier checks used a median absolute deviation cutoff of 5; no manual sample exclusion was performed solely on visual inspection. Cohort- and population-specific principal-component plots and adjusted-model sample counts are shown in Figures S5-S6, with preprocessing and sample-audit data in Table S3.

Differential-expression analysis

The primary transcriptomic population comprised HR-positive/HER2-negative tumors. A Luminal A/B analysis was retained as a secondary population. Differential expression was modeled separately in METABRIC and TCGA using limma with empirical Bayes moderation, robust estimation, and mean-variance trend modeling [8]. The adjusted design included mutation group, standardized age, and PAM50 category (Luminal A, Luminal B, or other), with Luminal A as the reference. Samples missing age or PAM50 were excluded from this adjusted analysis.

Three contrasts were estimated: AKT1-mutant versus PIK3CA-mutant-only, AKT1-mutant versus pair-wild-type, and PIK3CA-mutant-only versus pair-wild-type. The AKT1-mutant versus PIK3CA-mutant-only contrast was prespecified as the core comparison. P values were adjusted using the Benjamini-Hochberg method [17]. METABRIC genes with FDR<0.05 were considered discovery findings. Replication in TCGA required the same effect direction and nominal P<0.05; transcriptome-wide TCGA FDR was also reported and was not replaced by candidate-level adjustment. Expression-analysis summaries and all 53 same-direction nominally replicated genes are supplied in Tables S3-S4; complete matrices can be regenerated using Supplementary Code.

Pathway enrichment and cross-cohort replication

Pathway analysis used Hallmark and Reactome gene sets from locally versioned GMT files [10,11]. Competitive enrichment was performed using cameraPR applied to the limma moderated t statistic, accounting for an inter-gene correlation of 0.01 [9]. Gene sets with 10-500 mapped genes were tested. Enrichment was calculated separately for each cohort, population, and contrast, and FDR was controlled within each analysis.

A pathway was considered discovered when METABRIC FDR was below 0.05. Independent FDR-controlled replication required concordant direction and TCGA FDR below 0.05. Standardized pathway scores were additionally calculated by standardizing expression for each member gene, averaging mapped genes per sample, and standardizing the final score within cohort. Adjusted score models used

age and PAM50 as covariates. Pathway-analysis summaries and all nine independently replicated pathways are provided in Table S5.

Exploratory candidate genes were required to meet three prespecified conditions: METABRIC FDR<0.05, same-direction nominal replication in TCGA, and membership in the leading edge of an independently replicated pathway. Candidate-gene FDR values shown in the focused validation panel were adjusted across the three selected candidates and were not interpreted as transcriptome-wide TCGA significance. Pathway-score and candidate-gene effect estimates and sample-level values are provided in Table S6.

Genomic-context analysis

Secondary genomic analyses were restricted to HR-positive/HER2-negative AKT1-mutant and PIK3CA-mutant-only tumors. To maintain comparable targeted coverage, co-mutation testing used the 171 non-index genes from the METABRIC 173-gene panel. Non-synonymous alterations were evaluated for genes with at least five events in the corresponding cohort. Fisher exact tests were performed separately by cohort, followed by Benjamini-Hochberg correction.

Tumor mutational burden (TMB) was obtained from the cohort clinical files. Copy-number burden was calculated from discrete copy-number matrices as the fraction of genes with any copy-number alteration, high-level alteration, amplification, or deep deletion. Individual high-level copy-number events were tested when pooled prevalence was at least 3%. Expression and pathway robustness models further adjusted the primary age- and PAM50-adjusted estimates for log-transformed TMB, copy-number burden, and the number of FDR-significant co-mutations. Key co-mutation, genomic-burden, and robustness results are supplied in Table S7.

RPPA signaling analysis

TCGA RPPA z scores were analyzed in HR-positive/HER2-negative tumors with available protein data. A prespecified 21-antibody PI3K-AKT-mTOR panel included total and phosphorylated proteins. Thirteen phosphoproteins were assigned to three topology-based domains: AKT activation (AKT pS473 and AKT pT308), AKT substrates (PRAS40 pT246, GSK3 phosphorylation, and TSC2 pT1462), and mTOR effectors (mTOR pS2448, p70S6K pT389, S6 phosphorylation, and 4E-BP1 phosphorylation). Domain scores were calculated as the mean of standardized component proteins and then standardized across samples.

Protein and domain-score comparisons used limma models adjusted for standardized age and PAM50 category. To preserve the limited AKT1-mutant sample size, missing PAM50 was retained as an explicit unknown category. FDR was controlled within phosphoprotein, total-protein, or domain-score families for each contrast. Because the total AKT antibody detects AKT1, AKT2, and AKT3, total-AKT results were interpreted as pan-AKT abundance rather than AKT1-specific protein abundance. The RPPA panel

definition, sample manifest, sample-level values, group summaries, and adjusted effects are provided in Table S10.

Immune and survival analyses

Immune analyses were secondary and used three complementary expression-based methods: MCP-counter, ESTIMATE, and ssGSEA [12-14]. Scores were standardized within cohort and compared using age- and PAM50-adjusted models. Cross-cohort replication required concordant direction and nominal significance in TCGA. Method manifests, sample-level scores, adjusted effects, and cross-cohort summaries are provided in Table S9.

Survival analyses included overall survival, disease-specific survival, and relapse-free survival in METABRIC and overall and disease-free survival in TCGA. Kaplan-Meier and log-rank analyses were supplemented by cohort-specific Cox models. METABRIC adjusted models included age, PAM50, grade, tumor size, and nodal burden; TCGA models included age, PAM50, and stage when estimable. Proportional-hazards assumptions were tested. Analyses with no events in a comparison group were reported as non-estimable rather than assigned artificial estimates. Endpoint audits and primary adjusted Cox estimates are provided in Table S8.

Statistical analysis

All tests were two-sided. Unless otherwise specified, statistical significance was defined as FDR<0.05. Effect estimates, 95% confidence intervals, nominal P values, and FDR values were retained in machine-readable output tables. Analyses were performed in R version 4.4.3 with a fixed random seed. The complete pipeline used relative paths and could be rerun from the project root without modifying raw data. Supplementary figures and readable statistical tables are provided in a single Supplementary Material PDF. All R scripts, configuration files, locally versioned gene-set definitions, validation scripts, package versions, and a single master runner are provided separately as Supplementary Code.

Results

AKT1 and PIK3CA hotspot mutations define largely mutually exclusive breast cancer subsets

After mutation-evaluation and panel-coverage requirements, 2,433 METABRIC tumors and 816 TCGA tumors were available for hotspot analysis (Figure 1A; Table S1). In METABRIC, AKT1 E17K was present in 79 tumors (3.2%; 95% CI, 2.6%-4.0%) and a prespecified PIK3CA hotspot in 733 tumors (30.1%; 95% CI, 28.3%-32.0%). In TCGA, the corresponding frequencies were 19 of 816 (2.3%; 95% CI, 1.5%-3.6%) and 210 of 816 (25.7%; 95% CI, 22.9%-28.8%) (Figure 1B; Table S1).

No AKT1/PIK3CA co-mutant tumor was observed in METABRIC, and one was observed in TCGA. Fisher exact testing supported mutual exclusivity in the full cohorts (METABRIC P=4.98x10^-13; TCGA odds ratio=0.16, P=0.036), the HR-positive/HER2-negative subsets (METABRIC P=1.80x10^-11; TCGA odds ratio=0.14, P=0.028), and the Luminal A/B subsets (METABRIC P=2.12x10^-9; TCGA odds ratio=0.13, P=0.027) (Figure 1C-D).

Clinicopathologic distributions differed across the three mutation groups (Table 1, Table S2, and Figure 2). ER positivity, PR positivity, and PAM50 subtype composition differed significantly in both cohorts (overall P<0.001). AKT1-mutant and PIK3CA-mutant-only tumors were more frequently hormone receptor positive and Luminal A than pair-wild-type tumors. For example, Luminal A tumors comprised 50.8%, 50.2%, and 28.2% of evaluable AKT1-mutant, PIK3CA-mutant-only, and pair-wild-type METABRIC tumors, respectively, and 76.5%, 68.8%, and 44.9% in TCGA. Stage distributions were not significantly different. Age patterns were not consistent between cohorts and were not interpreted as mutation-specific.

AKT1-mutant and PIK3CA-mutant-only tumors display distinct transcriptional profiles

The adjusted HR-positive/HER2-negative expression analyses included 52 AKT1-mutant, 504 PIK3CA-mutant-only, and 807 pair-wild-type METABRIC tumors and 14, 102, and 203 TCGA tumors, respectively (Figure S6; Table S3). Principal-component analysis did not show complete separation of mutation groups, indicating substantial within-group heterogeneity and supporting adjusted feature-level rather than cluster-based comparisons (Figure 3A-B and Figure S5).

For the core AKT1-mutant versus PIK3CA-mutant-only contrast, 341 genes achieved FDR<0.05 in METABRIC. Of these, 280 (82.1%) showed the same direction of effect in TCGA and 53 achieved same-direction nominal replication. No individual gene passed transcriptome-wide FDR correction in TCGA. Across all 16,620 genes measured in both cohorts, effect sizes were positively correlated (Spearman rho=0.226, P=6.71x10^-191), although the modest correlation emphasized substantial platform and sampling heterogeneity (Figure 3C-D; Tables S3-S4).

Replicated genes included components or regulators relevant to PI3K-AKT signaling and cellular metabolism. AKT1 expression itself was higher in AKT1-mutant tumors in both cohorts (METABRIC log2 fold change=0.45, FDR=9.39x10^-6; TCGA log2 fold change=0.32, P=0.025). FOXO4, MYC, and PDK4 were lower in AKT1-mutant tumors, whereas PIK3R3 and several metabolic or secretory-pathway genes were higher. These individual-gene findings were treated as supportive because TCGA replication was nominal rather than transcriptome-wide FDR controlled.

Cross-cohort pathway analysis identifies reproducible biological programs

Pathway-level analysis produced stronger cross-cohort replication than individual-gene testing. For the AKT1-mutant versus PIK3CA-mutant-only contrast, 13 Hallmark and 31 Reactome pathways were significant in METABRIC. Two Hallmark and seven Reactome pathways showed concordant TCGA replication at FDR<0.05 (Figure S1; Table S5).

The replicated programs distinguished several biological themes. AKT1-mutant tumors showed higher G2M checkpoint and glycolysis enrichment and higher Reactome programs related to resolution of sister-chromatid cohesion and COPI-mediated or ER-Golgi transport. They showed lower eukaryotic translation elongation and nonsense-mediated mRNA decay programs. Focused standardized-score analyses supported the same directions in both cohorts, although score-level confidence intervals were wider in TCGA because of the small AKT1-mutant group (Figure 4A-C). In METABRIC, adjusted differences for AKT1-mutant versus PIK3CA-mutant-only tumors were 0.26 SD for G2M checkpoint (FDR=0.043), 0.52 SD for glycolysis (FDR<0.001), -0.27 SD for translation elongation (FDR=0.063), and -0.25 SD for nonsense-mediated decay (FDR=0.073).

GALE, CHPF, and NUSAP1 met the prespecified exploratory-candidate rule by combining METABRIC FDR significance, same-direction nominal TCGA replication, and membership in the leading edge of a replicated pathway. In models standardized to permit cross-platform comparison, all three showed positive AKT1-mutant versus PIK3CA-mutant-only effects in both cohorts (Figure 4D; Table S6). GALE and CHPF were linked to the glycolysis leading edge, whereas NUSAP1 was linked to the G2M-checkpoint leading edge. These genes were considered pathway-anchored candidates rather than independently validated biomarkers.

Genomic context does not fully explain the expression differences

Among the 171 non-index genes assessed in both cohorts, GATA3 was the only differential co-mutation passing FDR correction. In TCGA, GATA3 mutations were detected in 7 of 16 AKT1-mutant tumors (43.8%) and 9 of 138 PIK3CA-mutant-only tumors (6.5%; odds ratio=10.82; FDR=0.0042). The corresponding enrichment was directionally present but not significant in METABRIC (29.6% versus 13.7%), and therefore the GATA3 association was considered exploratory rather than independently replicated (Figure 5A-B; Table S7).

TMB, the fraction of the genome affected by any copy-number alteration, and the fraction affected by high-level copy-number events did not differ significantly between AKT1-mutant and PIK3CA-mutant-only tumors in either cohort (all primary FDR values>0.38). No individual amplification or deep deletion achieved FDR<0.05 (Figure 5C-D). Adding TMB, copy-number burden, and significant co-mutations to the clinical models did not materially attenuate the principal candidate-gene and pathway-score effects in METABRIC (Figure 5E), suggesting that the expression findings were not explained solely by broad genomic burden.

RPPA demonstrates preferential proximal AKT activation in AKT1-mutant tumors

TCGA RPPA data were available for 12 AKT1-mutant, 112 PIK3CA-mutant-only, and 241 pair-wild-type HR-positive/HER2-negative tumors (Table S10). The phosphoprotein landscape showed markedly higher AKT pS473 and pT308 in AKT1-mutant tumors (Figure 6A). In age- and PAM50-adjusted models, the aggregate AKT activation score was 0.98 SD higher in AKT1-mutant than PIK3CA-mutant-only tumors (95% CI, 0.41-1.56; FDR=0.0025) (Figure 6B).

At the individual-protein level, AKT pS473 was 1.00 SD higher (95% CI, 0.41-1.59; FDR=0.0107) and AKT pT308 was 0.90 SD higher (95% CI, 0.32-1.49; FDR=0.0166) in AKT1-mutant tumors (Figure 6C-D). Pan-AKT protein abundance did not differ between AKT1-mutant and PIK3CA-mutant-only tumors (difference, -0.08 SD; FDR=0.887). Aggregate AKT-substrate and mTOR-effector scores were also not different (differences, -0.16 and -0.02 SD; FDR=0.866 and 0.953, respectively). Thus, the RPPA data supported preferential proximal AKT phosphorylation rather than uniform activation across all measured downstream effectors.

Secondary immune and survival analyses provide limited additional evidence

In METABRIC, AKT1-mutant tumors had lower ESTIMATE stromal, immune, and combined scores than PIK3CA-mutant-only tumors, and the IL6-JAK-STAT3 inflammatory program was lower after multiple-testing correction (Figure S3; Table S9). The descriptive immune landscape similarly showed lower median lymphoid and myeloid MCP-counter scores in METABRIC AKT1-mutant tumors and lower adjusted ESTIMATE immune and stromal scores (Figure S4A-C). However, individual tumors were heterogeneous, and the corresponding group-level patterns were absent or directionally inconsistent in TCGA (Figure S4A-C). No MCP-counter feature met the prespecified cross-cohort replication criterion (Figure S3; Table S9). The immune findings were therefore interpreted as cohort-specific and hypothesis generating.

No survival contrast achieved FDR<0.05 (Figure S2; Table S8). In adjusted METABRIC HR-positive/HER2-negative models, AKT1-mutant versus PIK3CA-mutant-only hazard ratios were 0.80 for overall survival (95% CI, 0.54-1.21), 1.00 for disease-specific survival (95% CI, 0.59-1.70), and 0.83 for relapse-free survival (95% CI, 0.51-1.37). TCGA comparisons were not estimable in the primary population because no overall-survival or disease-free-survival events occurred among AKT1-mutant tumors. These results did not support an independent prognostic interpretation.

Discussion

In this two-cohort study, AKT1 E17K and recurrent PIK3CA hotspot mutations identified largely mutually exclusive, predominantly hormone receptor-positive breast cancer subsets but were associated with distinguishable downstream molecular states. The principal findings were threefold. First, adjusted transcriptomic comparisons showed reproducible directional differences between AKT1-mutant and PIK3CA-mutant-only tumors. Second, pathway-level analysis identified replicated differences in cell-cycle, glycolytic, chromosome-segregation, secretory-transport, translation, and RNA-surveillance programs. Third, TCGA RPPA data showed preferential phosphorylation of AKT at S473 and T308 in AKT1-mutant tumors without corresponding increases in pan-AKT abundance or uniform downstream mTOR-effector activation.

The results extend prior clinicogenomic descriptions of the PI3K-AKT pathway. Large studies have shown that PIK3CA is altered more frequently than AKT1 and that AKT1 mutations are strongly concentrated at E17K [1]. Our frequency and mutual-exclusivity results were consistent with that literature. The contribution of the present analysis is the direct integration of adjusted transcriptomic discovery, independent pathway replication, and protein phosphorylation within a clinically relevant HR-positive/HER2-negative population. This design asks a narrower question than broad alteration profiling: whether two canonical activating mutations commonly grouped within the same therapeutic pathway are associated with equivalent molecular phenotypes.

The RPPA findings provide a plausible signaling correlate for the transcriptomic differences. AKT1 E17K promotes membrane localization and constitutive signaling competence [2], whereas PIK3CA hotspots act upstream by increasing PI3K catalytic activity. Higher S473 and T308 phosphorylation in AKT1-mutant tumors is therefore biologically coherent. Importantly, the absence of broad differences in measured AKT substrates and mTOR effectors argues against a simple model of uniformly greater pathway output. Feedback regulation, phosphatase activity, substrate availability, parallel signaling, and tissue context may uncouple proximal AKT phosphorylation from individual downstream nodes. The data support non-equivalence, but they do not establish that one mutation produces globally stronger PI3K-AKT-mTOR signaling.

The replicated pathway findings suggest that mutation-specific effects extend beyond canonical pathway annotations. Higher glycolysis and G2M-related activity in AKT1-mutant tumors was supported by pathway-level replication and by the exploratory candidates GALE, CHPF, and NUSAP1. Increased ER-Golgi transport and chromosome-segregation programs, together with reduced translation-elongation and nonsense-mediated-decay programs, point to broader differences in biosynthetic and RNA-processing states. These observations are hypothesis generating. In particular, the candidate genes were selected from the analyzed data and require confirmation in external datasets or experimental models before being considered biomarkers or therapeutic targets.

The genomic-context analysis provided limited evidence that broad genomic differences explained the transcriptomic findings. TMB and copy-number burden were similar between groups, and adjustment for genomic context did not substantially alter the principal METABRIC expression estimates. GATA3 co-

mutation was enriched in TCGA AKT1-mutant tumors but was not statistically replicated in METABRIC. Given the small TCGA AKT1-mutant group, this finding should not be interpreted as a stable defining feature without additional confirmation.

The clinical implications are not yet established. Current treatment selection can group PIK3CA, AKT1, and PTEN alterations as biomarkers for AKT inhibition, whereas PI3K-alpha inhibitors are specifically used for PIK3CA-mutant disease [15, 16]. Our results do not compare treatment response and should not be used to infer differential sensitivity to PI3K, AKT, mTOR, endocrine, or CDK4/6-directed therapies. They instead suggest that pathway-level eligibility does not imply molecular interchangeability. Treatment-annotated cohorts will be required to determine whether the observed signaling and transcriptional differences modify therapeutic response.

This study has several limitations. AKT1 E17K was uncommon, leaving only 14 adjusted-expression and 12 RPPA samples in TCGA. METABRIC and TCGA used different expression platforms, which reduces direct single-gene comparability but also provides a degree of cross-platform validation. The analysis was retrospective and based on bulk tumors, so cellular composition and intratumoral heterogeneity could not be fully resolved. The RPPA pan-AKT antibody measured AKT1, AKT2, and AKT3 rather than AKT1 alone. PAM50 status was unavailable for a subset of TCGA samples, and the RPPA model retained missing values as an unknown category. The single TCGA co-mutant tumor was assigned to the AKT1-mutant group according to the prespecified rule. Finally, no experimental validation or treatment-response cohort was available, and the immune and survival findings did not provide stable independent support.

Conclusions

AKT1 E17K and PIK3CA hotspot mutations are associated with overlapping but non-equivalent molecular phenotypes in HR-positive/HER2-negative breast cancer. AKT1-mutant tumors showed distinct cross-cohort transcriptional programs and preferential proximal AKT phosphorylation, without evidence of uniform downstream mTOR-effector activation. Mutation-specific biological interpretation should therefore accompany broad PI3K-AKT pathway classification. Prospective treatment-annotated studies and experimental models are needed to determine the therapeutic significance of these differences.

Declarations

Ethics approval and consent to participate

This study analyzed de-identified, publicly available data from previously published METABRIC and TCGA studies. No new patient recruitment or intervention was performed. Ethical approval and informed-consent procedures were reported by the original data-generating studies.

Consent for publication

Not applicable.

Data availability

The METABRIC and TCGA-BRCA data analyzed in this study are publicly available through cBioPortal. Versioned R scripts, local configuration files, and machine-readable analysis outputs are contained in the accompanying project repository.

Code availability

The complete analysis code is provided as Supplementary Code, including all R scripts, configuration files, fixed gene-set resources, validation checks, package-version records, and the master runner R/run_all_manuscript_analysis.R. Raw downloaded data are read-only and are not overwritten by the pipeline.

Competing interests

[The authors declare no competing interests / insert author-specific disclosures.]

Funding

This research received no external funding.

Author contributions

QZ：dataset downloading, code analysis, figure plotting, manuscript drafting.

HL:revision and proofreadingAcknowledgments

Acknowledgments

Not applicable.

References

1. Tao JJ, Sisoudiya SD, Tukachinsky H, Schrock A, Sivakumar S, Sokol ES, et al. Clinicogenomic landscape and function of PIK3CA, AKT1, and PTEN mutations in breast cancer. JCO Precis Oncol. 2026;10(4):e2500609. 10.1200/PO-25-00609.
2. Carpten JD, Faber AL, Horn C, Donoho GP, Briggs SL, Robbins CM, et al. A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. Nature. 2007;448:439–44. 10.1038/nature05933.
3. Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, et al. High frequency of mutations of the PIK3CA gene in human cancers. Science. 2004;304:554. 10.1126/science.1096502.
4. CCurtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature. 2012;486:346–52. 10.1038/nature10983.
5. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490:61–70. 10.1038/nature11412.
6. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer Genomics Portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2:401–4. 10.1158/2159-8290.CD-12-0095.
7. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6:pl1. 10.1126/scisignal.2004088.
8. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43:e47. 10.1093/nar/gkv007.
9. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. Nucleic Acids Res. 2012;40:e133. 10.1093/nar/gks461.
10. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database hallmark gene set collection. Cell Syst. 2015;1:417–25. 10.1016/j.cels.2015.12.004.
11. Gillespie M, Jassal B, Stephan R, Milacic M, Rothfels K, Senff-Ribeiro A, et al. The Reactome pathway knowledgebase 2022. Nucleic Acids Res. 2022;50:D687–92. 10.1093/nar/gkab1028.
12. Hanzelmann S, Castelo R, Guinney J. BMC Bioinformatics. 2013;14:7. 10.1186/1471-2105-14-7. GSVA: gene set variation analysis for microarray and RNA-seq data.
13. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17:218. 10.1186/s13059-016-1070-5.
14. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612. 10.1038/ncomms3612.
15. Turner NC, Oliveira M, Howell SJ, Dalenc F, Cortes J, Gomez Moreno HL, et al. Capivasertib in hormone receptor-positive advanced breast cancer. N Engl J Med. 2023;388:2058–70.

10.1056/NEJMoa2214131.

16. AAndré F, Ciruelos E, Rubovszky G, Campone M, Loibl S, Rugo HS, et al. Alpelisib for PIK3CA-mutated, hormone receptor-positive advanced breast cancer. N Engl J Med. 2019;380:1929–40. 10.1056/NEJMoa1813904.
17. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B. 1995;57:289–300. 10.1111/j.2517-6161.1995.tb02031.x.

Table

Table 1 is available in the Supplementary Files section.

Figures

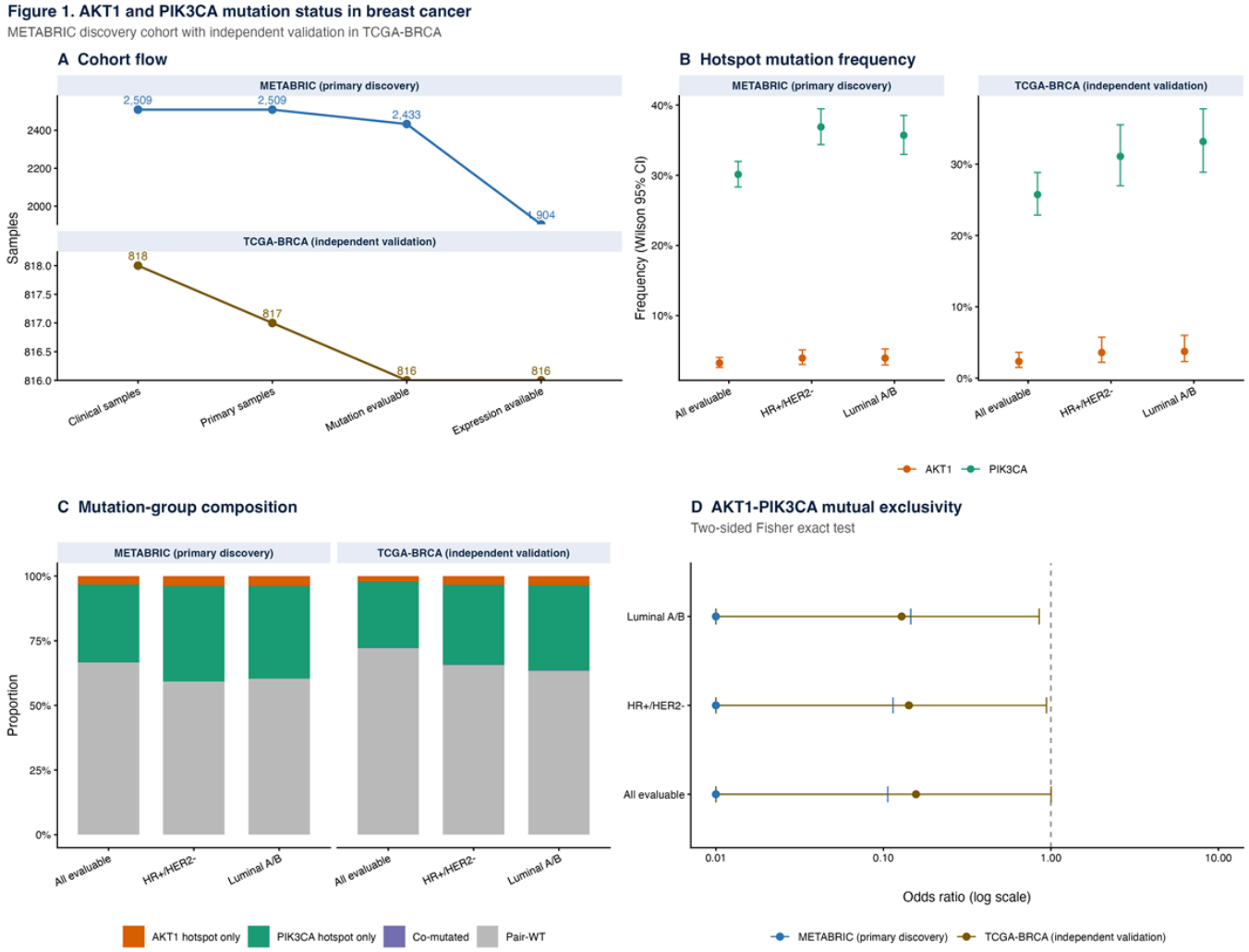


Figure 1

AKT1 and PIK3CA hotspot-mutation status in breast cancer.

(A) Cohort flow from clinical samples to primary, mutation-evaluable, and expression-available samples. (B) AKT1 and PIK3CA hotspot frequencies with Wilson 95% confidence intervals in all mutation-evaluable, HR-positive/HER2-negative, and Luminal A/B populations. (C) Proportions of AKT1-hotspot-only, PIK3CA-hotspot-only, co-mutated, and pair-wild-type tumors. (D) Odds ratios and 95% confidence intervals from two-sided Fisher exact tests of AKT1-PIK3CA mutual exclusivity. METABRIC was the discovery cohort and TCGA-BRCA the independent validation cohort.

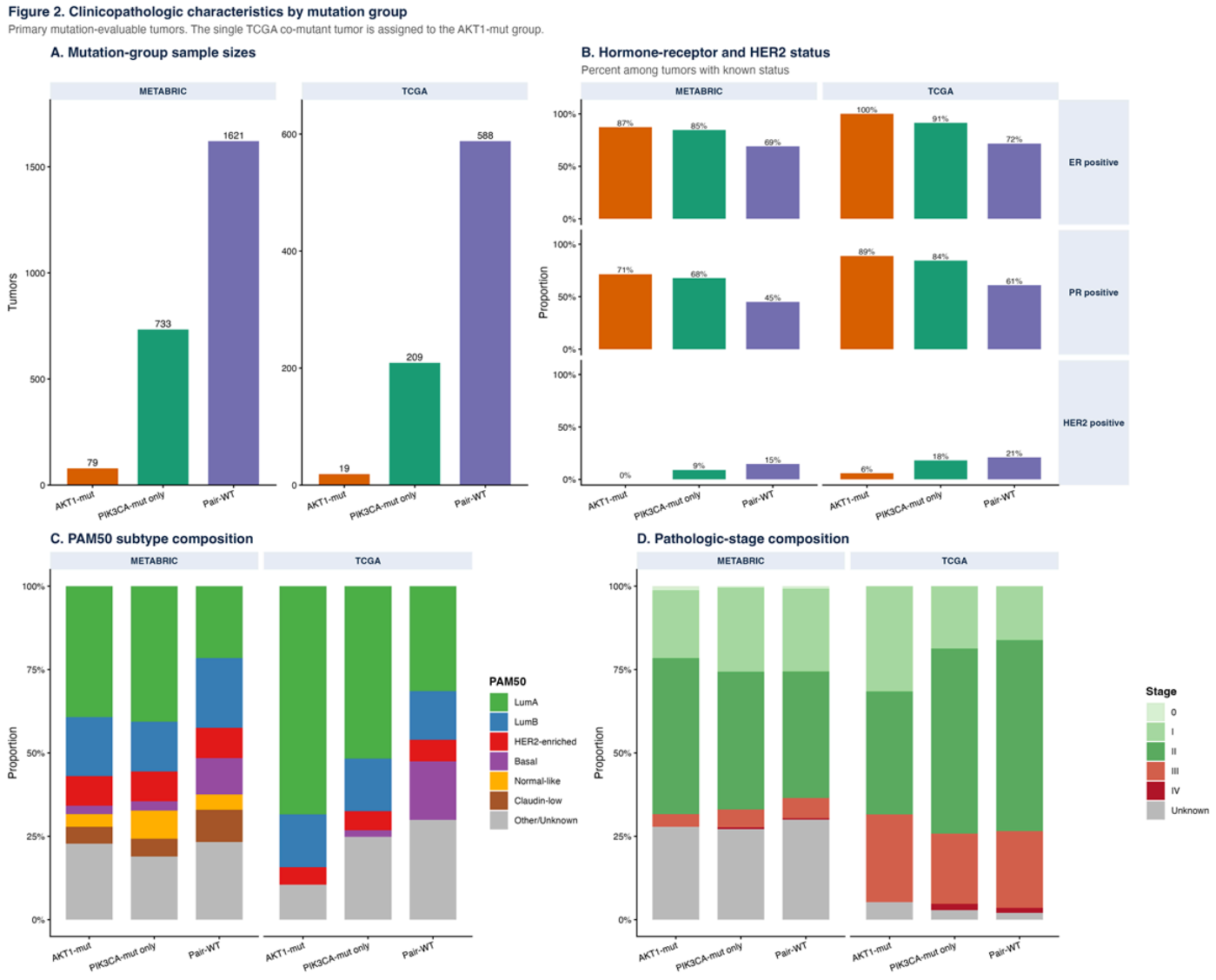


Figure 2

Clinicopathologic characteristics by mutation group.

(A) Numbers of primary mutation-evaluable AKT1-mutant, PIK3CA-mutant-only, and pair-wild-type tumors in each cohort. (B) ER-positive, PR-positive, and HER2-positive proportions among tumors with known status. (C) PAM50 subtype composition. (D) Pathologic-stage composition. The single TCGA co-

mutant tumor was assigned to the AKT1-mutant group. Full distributions, missingness, and overall statistical comparisons are provided in Table 1.

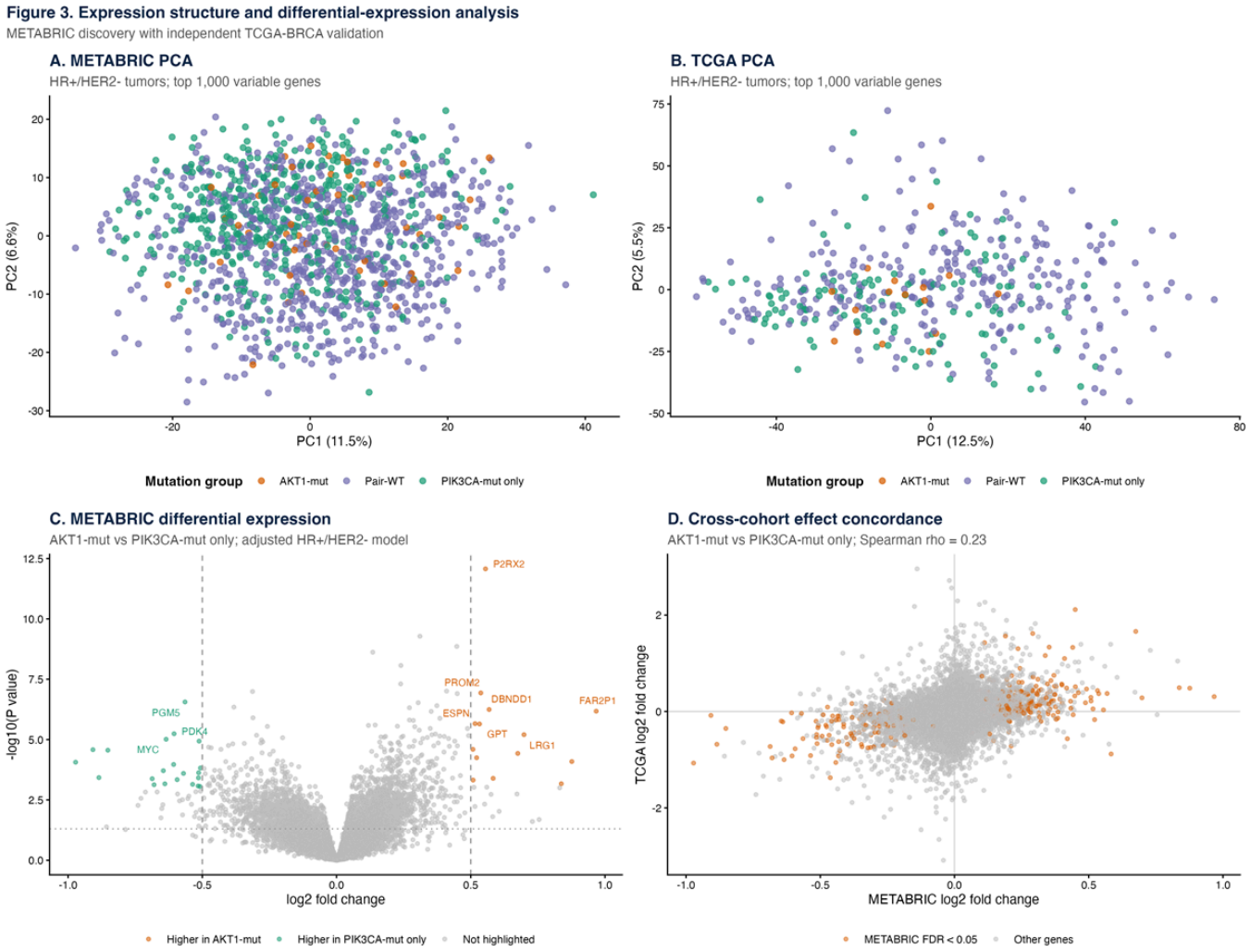


Figure 3

Expression structure and differential-expression analysis.

(A-B) Principal-component analysis of the 1,000 most variable genes in HR-positive/HER2-negative METABRIC and TCGA tumors. (C) METABRIC volcano plot for the age- and PAM50-adjusted AKT1-mutant versus PIK3CA-mutant-only comparison. Dashed vertical lines indicate absolute log2 fold change of 0.5 and the horizontal line nominal P=0.05. (D) Cross-cohort comparison of adjusted log2 fold changes for 16,620 common genes; genes with METABRIC FDR<0.05 are highlighted.

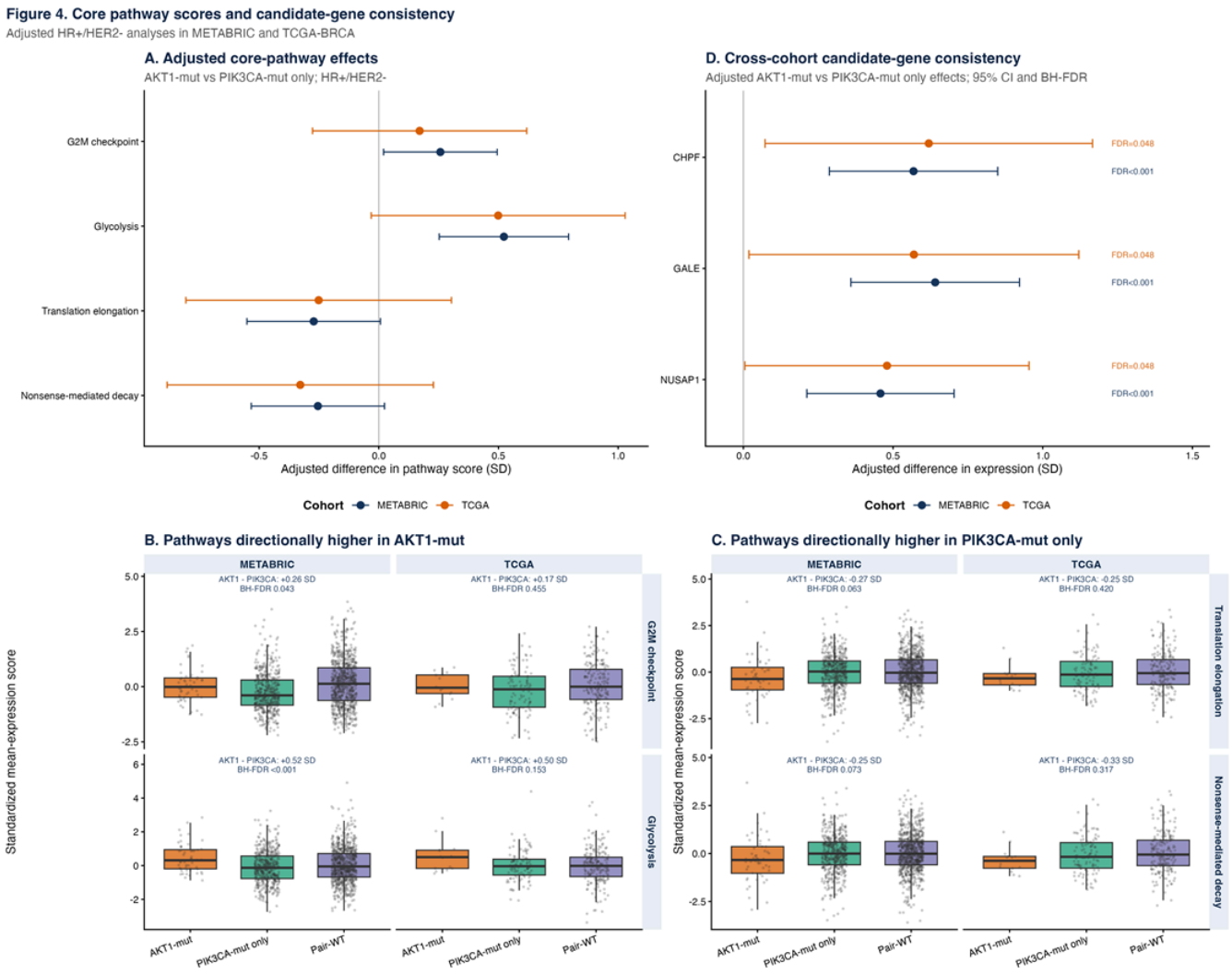


Figure 4

Core pathway scores and candidate-gene consistency.

(A) Adjusted AKT1-mutant versus PIK3CA-mutant-only effects and 95% confidence intervals for selected replicated biological programs. (B) Standardized G2M-checkpoint and glycolysis scores, which were directionally higher in AKT1-mutant tumors. (C) Standardized translation-elongation and nonsense-mediated-decay scores, which were directionally lower in AKT1-mutant tumors. (D) Adjusted standardized expression differences for the exploratory pathway-anchored candidates CHPF, GALE, and NUSAP1. Models adjusted for age and PAM50 subtype. Candidate-panel FDR values are corrected across the three displayed genes and do not represent transcriptome-wide TCGA significance.

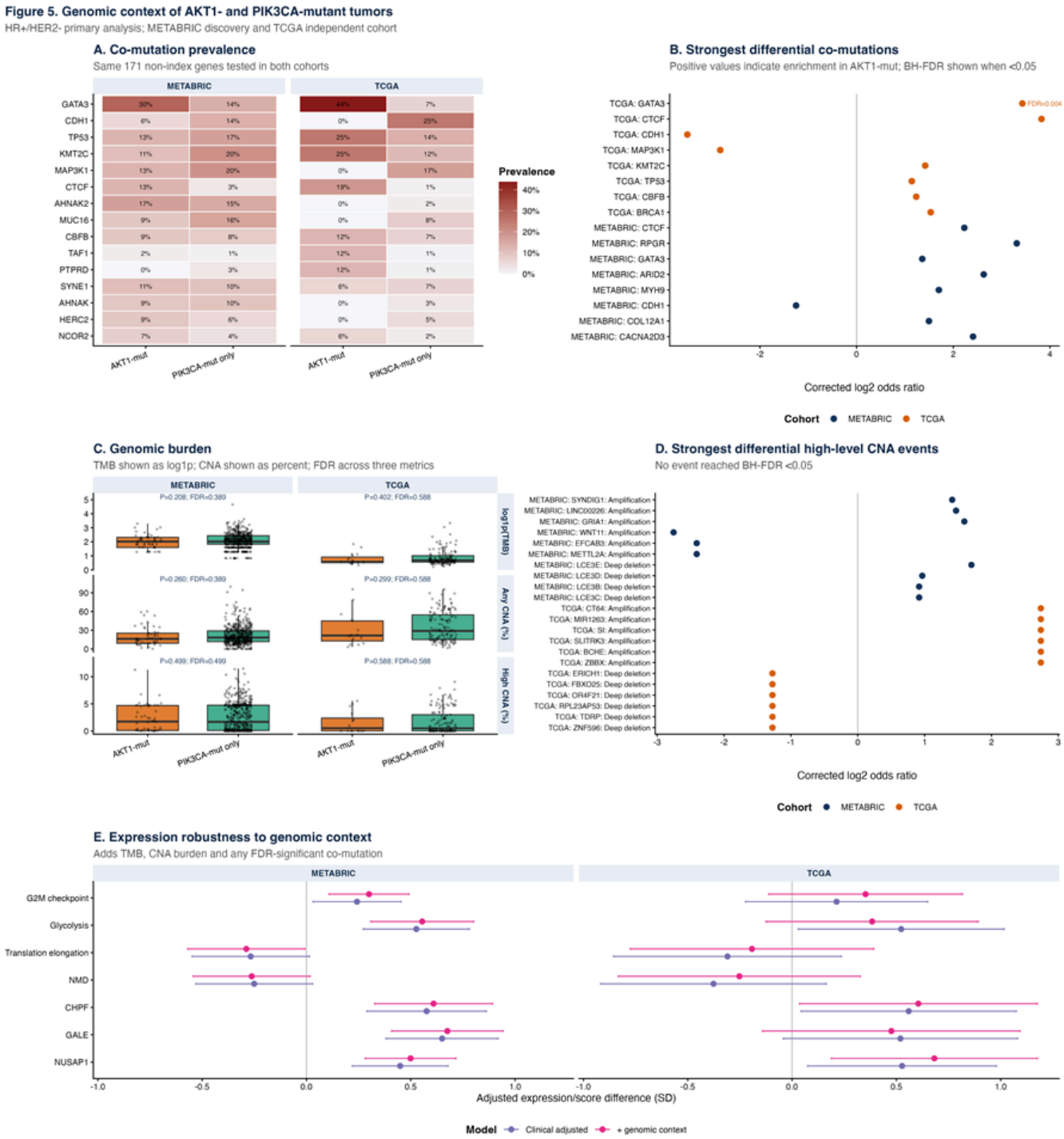


Figure 5

Genomic context of AKT1- and PIK3CA-mutant tumors.

(A) Co-mutation prevalence among non-index genes tested in both cohorts. (B) Differential co-mutation effects; positive corrected log2 odds ratios indicate enrichment in AKT1-mutant tumors. GATA3 was significant in TCGA but not independently replicated in METABRIC. (C) TMB and copy-number burden distributions. (D) Strongest differential high-level copy-number events; none achieved FDR<0.05. (E)

Comparison of clinical and clinical-plus-genomic adjusted effects for selected pathways and candidate genes.

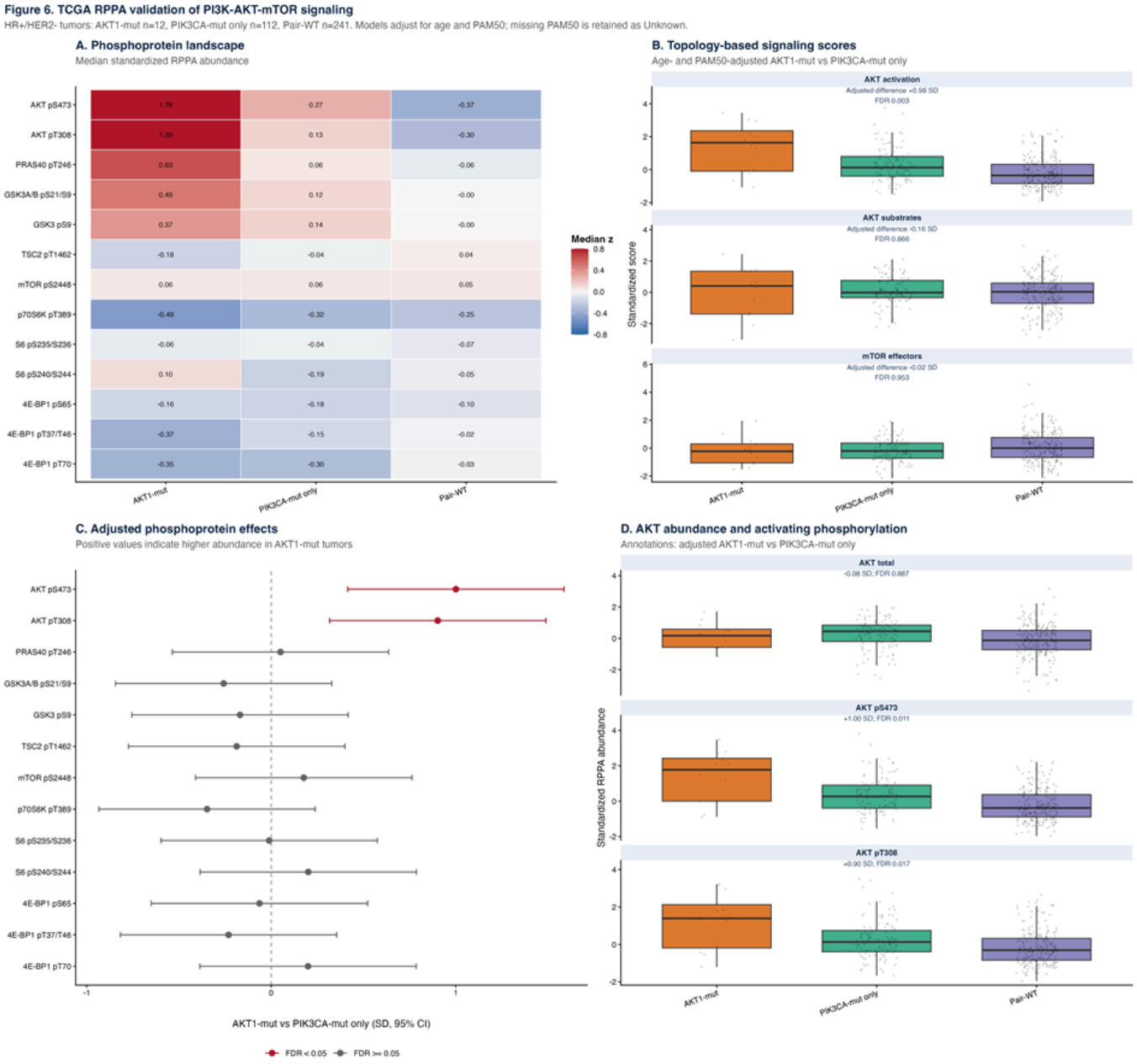


Figure 6

TCGA RPPA validation of PI3K-AKT-mTOR signaling.

(A) Median standardized phosphoprotein abundance across the three mutation groups. (B) Topology-based AKT activation, AKT-substrate, and mTOR-effector scores. (C) Adjusted AKT1-mutant versus PIK3CA-mutant-only phosphoprotein effects with 95% confidence intervals. (D) Distributions of pan-AKT abundance and AKT pS473 and pT308. Models adjusted for age and PAM50 category, retaining missing

PAM50 as unknown. RPPA sample sizes were 12 AKT1-mutant, 112 PIK3CA-mutant-only, and 241 pair-wild-type tumors.

Supplementary Files

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

- SupplementaryMaterial.docx
- Table1.docx