OncoOmics approaches to reveal essential genes in breast cancer: a panoramic view from pathogenesis to precision medicine

Andrés López-Cortés,[1,2,*] César Paz-y-Miño,[1] Santiago Guerrero,[1] Alejandro Cabrera-Andrade,[2,3,4] Stephen J. Barigye,[5] Cristian R. Munteanu,[2,6] Humberto González-Díaz,[7,8] Alejandro Pazos,[2,6] Yunierkis Pérez-Castillo,[4,9] and Eduardo Tejera[4,10,*]

[1] Centro de Investigación Genética y Genómica, Facultad de Ciencias de la Salud Eugenio Espejo, Universidad UTE, Mariscal Sucre Avenue, Quito 170129, Ecuador

[2] RNASA-IMEDIR, Computer Science Faculty, University of Coruna, Coruna 15071, Spain

[3] Carrera de Enfermería, Facultad de Ciencias de la Salud, Universidad de Las Américas, Avenue de los Granados, Quito 170125, Ecuador

[4] Grupo de Bio-Quimioinformática, Universidad de Las Américas, Avenue de los Granados, Quito 170125, Ecuador

[5] Department of Chemistry, McGill University, 801 Sherbrooke Street West, Montreal, QC H3A 0B8, Canada

[6] INIBIC, Institute of Biomedical Research, CHUAC, UDC, Coruna 15006, Spain

[7] Department of Organic Chemistry II, University of the Basque Country UPV/EHU, Leioa 48940, Biscay, Spain

[8] IKERBASQUE, Basque Foundation for Science, Bilbao 48011, Biscay, Spain

[9] Escuela de Ciencias Físicas y Matemáticas, Universidad de Las Américas, Avenue de los Granados, Quito 170125, Ecuador

[10] Facultad de Ingeniería y Ciencias Agropecuarias, Universidad de Las Américas, Avenue de los Granados, Quito 170125, Ecuador

*** Authors to whom correspondence should be addressed**

Andrés López-Cortés, MSc.

Centro de Investigación Genética y Genómica, Facultad de Ciencias de la Salud Eugenio Espejo, Universidad UTE, Mariscal Sucre Avenue, Quito 170129, Ecuador. E-mail: aalc84@gmail.com

Eduardo Tejera, PhD.

Facultad de Ciencias de la Salud, Universidad de Las Américas, Avenue de los Granados, Quito 170125, Ecuador. E-mail: eduardo.tejera@udla.edu.ec

SUMMARY

Breast cancer (BC) is a heterogeneous disease where each OncoOmics approach needs to be fully understood as a part of a complex network. Therefore, the main objective of this study was to analyze genetic alterations, signaling pathways, protein-protein interaction networks, protein expression, dependency maps and enrichment maps in 230 previously prioritized genes by the Consensus Strategy, the Pan-Cancer Atlas, the Pharmacogenomics Knowledgebase and the Cancer Genome Interpreter, in order to reveal essential genes to accelerate the development of precision medicine in BC. The OncoOmics essential genes were rationally filtered to 144, 48 (33%) of which were hallmarks of cancer and 20 (14%) were significant in at least three OncoOmics approaches: RAC1, AKT1 CCND1, PIK3CA, ERBB2, CDH1, MAPK14, TP53, MAPK1, SRC, RAC3, PLCG1, GRB2, MED1, TOP2A, GATA3, BCL2, CTNNB1, EGFR and CDK2. According to the Open Targets Platform, there are 111 drugs that are currently being analyzed in 3151 clinical trials in 39 genes. Lastly, there are more than 800 clinical annotations associated with 94 genes in BC pharmacogenomics.

INTRODUCTION

Breast cancer (BC) is a heterogeneous disease characterized by an intricate interplay between different biological aspects such as ethnicity, genomic alterations, gene expression deregulation, hormone disruption, signaling pathway alterations and environmental determinants[1,2]. Over the last years, prevention, treatment and survival strategies have evolved favorably; however, there are BC profiles that remain incurable[3]. Nowadays, BC is the leading cause of cancer-related death among women (626,679; 15% cases) and the most commonly diagnosed cancer (2,088,849; 24% cases) worldwide[4].

The development of large-scale DNA sequencing, gene expression, proteomics, large-scale RNA interference (RNAi) screens and large-scale CRISPR-Cas9 screens has allowed us to better understand the molecular landscape of oncogenesis. Significant progress has been made in discovering gene coding regions[5], cancer driver genes[6,7], cancer driver mutations[8,9], germline variants[10], driver fusion genes[11,12], alternatively spliced transcripts[13], expression-based stratification[14], molecular subtyping[15], biomarkers[16], druggable enzymes[17], cancer dependencies[18–21], and drug sensitivity and resistance[22].

Scientific advances made to date mark the era called the “end of the beginning” of cancer omics. In other words, each approach that was previously mentioned needs to be fully understood as a part of a complex network, analyzing the mechanistic interplay of signaling pathways, protein-protein interaction (PPi) networks, enrichment maps, gene ontology (GO), deep learning, molecular dependencies and genomic alterations per intrinsic molecular subtype: basal-like (estrogen receptor (ER)$^-$, progesterone receptor (PR)$^-$, human epidermal growth factor receptor 2 (Her2)$^-$, cytokeratin 5/6$^+$ and/or EGFR$^+$); Her2-enriched (ER$^-$, PR$^-$, Her2$^+$); luminal A (ER$^+$ and/or PR$^+$, Her2$^-$, low Ki67); luminal B with Her2$^-$ (ER$^+$ and/or PR$^+$, Her2$^-$, low Ki67); luminal B with Her2$^+$ (ER$^+$ and/or PR$^+$, Her2$^-$, any Ki67); and normal like[23–29]. We will herein analyze previously prioritized genes/biomarkers by the Consensus Strategy (CS)[28], the Pan-

Cancer Atlas (PCA)[3,12,30–36], the Pharmacogenomics Knowledgebase (PharmGKB)[37] and the Cancer Genome Interpreter (CGI)[38].

In our previous studies, López-Cortés *et al*. and Tejera *et al*., developed a Consensus Strategy that was proved to be highly efficient in the recognition of gene-disease association[28,39]. The main objective was to apply several bioinformatics methods to explore BC pathogenic genes. The CS identified both well-known pathogenic genes and prioritized genes that will be further explored through the OncoOmics approaches. On the other hand, The Cancer Genome Atlas (TCGA) has concluded the most sweeping cross-cancer analysis yet undertaken, namely the PCA project[31]. PCA reveals how genetic alterations, such as putative mutations, fusion genes, mRNA expression, copy number variants (CNVs) and protein expression collaborate in BC progression, providing insights to prioritize the development of new treatments and immunotherapies[3,12,30–36]. The CGI flags genomic biomarkers of drug response with different levels of clinical relevance[38]. Lastly, PharmGKB is a comprehensive resource that curates and spreads knowledge of the impact of clinical annotations on BC drug response[37,40]. PharmGKB collects the precise guidelines for the application of pharmacogenomics in clinical practice published by the European Society for Medical Oncology (ESMO), the National Comprehensive Cancer Network (NCCN), the Royal Dutch Association for the Advancement of Pharmacy (DPWG), the Canadian Pharmacogenomics Network for Drug Safety (CPNDS) and the Clinical Pharmacogenetics Implementation Consortium (CPIC)[41–44]. Hence, the aim of this study was to implement OncoOmics approaches to analyze genetic alterations, signaling pathways, PPi networks, protein expression, BC dependencies and enrichment maps in order to reveal essential genes/biomarkers to accelerate the development of precision medicine in BC.

RESULTS

OncoPrint of genetic alterations according to the Pan-Cancer Atlas. PCA has reported the clinical data of 1084 individuals with BC and it can be visualized in the Genomic Data Commons of the National Cancer Institute and in the cBioPortal[45,46]. In regard to molecular subtypes and tumor stages, 46% were lumina A, 18% luminal B, 7% Her2-enriched, 16% basal-like and 3% normal-like, whereas 17% were stage T1, 58% stage T2, 23% stage T3 and 2% stage T4 (Table S1).

Figure 1A shows the average frequency of genetic alterations per gene set. The average frequency of the PCA gene set was 1.3, followed by CS gene set (1.2), PharmGKB/CGI gene set (1.1), BC driver genes (0.8) and non-cancer genes (0.4) (Table S2). Significant p-values ($p < 0.001$) were found among all gene sets. Therefore, the fact that gene sets of interest (CS, PCA and PharmGKB/CGI) presented an average frequency of genetic alterations greater than the non-cancer gene set and the BC driver gene set indicates that we are analyzing potential essential genes in BC. Figure 1B shows the percentage of genetic alterations per type. The most common genetic alterations were mRNA upregulation (55.8%), CNV amplification (17.1%) and missense mutations (8.4%). Figure 1C shows the ratio of genetic alterations in the 230 genes per sample and molecular subtype. Basal-like had the highest ratio (n = 33), followed by Her2-enriched (29), luminal B (24), normal-like (17) and luminal A (15). The ratio of all BC samples was 19.6. Figure 1D shows the ratio of genetic alterations in the 230 genes per sample and tumor stage. Stage T2 had the highest ratio (23), followed by T3 (22), T1 (17) and T4 (8). Figures 1E and 1F show the percentage of genetic alterations per subtype and tumor stage, respectively. mRNA upregulation and CNV amplification were the most common alterations in all molecular subtypes and tumor stages.

Figure 2 shows the ranking of genes with the greatest number of genetic alterations per molecular subtype and tumor stage. Regarding molecular subtypes, *PIK3CA* was the most altered gene in luminal A, *CCND1* in luminal B, TP53 in basal-like and normal-like, and *ERBB2* in Her2-enriched, with significant p-values < 0.001 (Figure 2A). On the other hand, the

most altered genes per tumor stage were *PIK3CA* in stage T1, TP53 in stages T2 and T3, and *ERBB2* in stage T4, with significant p-value < 0.001 (Figure 2B). Figures 2C, 2E, 2G, 2I and 2K show the top mutated genes, CNV amplified genes, CNV deep deleted genes, mRNA upregulated genes and mRNA downregulated genes per molecular subtype, respectively (Tables S3-S7). On the other hand, Figures 2D, 2F, 2H, 2J and 2L show the top mutated genes, CNV amplified genes, CNV deep deleted genes, mRNA upregulated genes and mRNA downregulated genes per tumor stage, respectively (Tables S8-S13).

Regarding the first OncoOmics approach, Figure 3A shows an OncoPrint of 73 genes with a number of genetic alterations greater than the average (> 86). For this analysis driver mutations were taken into account, discarding passenger mutations (Figure S1 and Table S14). Figure 3B shows a circos plot of interactions between molecular subtypes and genetic alterations of the 73 most altered genes. mRNA downregulated plus CNV deep deleted genes and mRNA upregulated plus CNV amplified genes were more related with basal-like, whereas fusion genes, and driver mutations were more related with Her2-enriched. Finally, Figure 3C shows a circos plot of interactions between tumor stages and genetic alterations of the 73 most altered genes. Fusion genes, mRNA downregulated plus CNV deep deleted genes, and mRNA upregulated plus CNV amplified genes were more related with stage T4, whereas driver mutations were more related with stage T3.

Pathway enrichment analysis. The pathway enrichment analysis was performed using David Bioinformatics Resource to obtain integrated information from the Kyoto Encyclopedia of Genes and Genomes (KEGG)[47–50]. The enrichment analysis of signaling pathways was carried on in the 230 genes, obtaining more than 50 terms with a false discovery rate (FDR) < 0.01 (Table S15). Subsequently, genetic alterations of genes that make up each signaling pathway were analyzed according to the molecular subtype and tumor stage. Figure 4A shows a circos plot correlating molecular subtypes with signaling pathways (Table S16). NF-kappa ß, NOD-like receptor, adipocytokine, GnRH, RIG-like receptor, TNF, TGFß, FOXO, glucagon, MAPK,

prolactin, cAMP, PI3K-AKT, neurotrophin, VEGF, notch, p53, sphingolipid and Wnt signaling pathways were more altered in basal-like; estrogen, HIF1, toll-like receptor, ras, insulin, T-cell receptor, rap1, ERBB, AMPK, chemokine, B-cell receptor, mTOR, Fc-epsilon RI, Jak-STAT, phosphatidylinositol and thyroid hormone signaling pathways were more altered in Her2-enriched; and Hippo signaling pathway in normal-like. On the other hand, Figure 4B shows the ranking of the most altered signaling pathways per molecular subtype. Jak-STAT signaling pathway was more altered in luminal A; Wnt signaling pathway in luminal B; p53 signaling pathway in basal-like; ERBB signaling pathway in Her2-enriched; and Hippo signaling pathway in normal-like (Table S17).

Figure 4C shows a circos plot correlating tumor stages with signaling pathways according to the frequency of genetic alterations (Table S16). NOD-like receptor, adipocytokine, GnRH, TNF, estrogen, prolactin, FOXO, glucagon, ras, MAPK, T-cell receptor, cAMP, rap1, PI3K-AKT, B-cell receptor, VEGF, mTOR, Fc epsilon RI, NOTCH, p53, sphingolipid and Wnt signaling pathways were more altered in stage T2; NF-kappa ß, Hippo and phosphatidylinositol signaling pathways were more altered in stage T3; and RIG-like receptor, HIF1, TGFß, toll-like receptor, insulin, AMPK, ERBB, chemokine, neurotrophin, mTOR, jak-STAT and thyroid hormone signaling pathways were more altered in stage T4. On the other hand, Figure 4D shows the ranking of the most altered signaling pathways per tumor stage. Wnt signaling pathway was more altered in stages T1, T2 and T3; and thyroid hormone signaling pathway was more altered in stage T4 (Table S18).

Protein-protein interaction network. Regarding the second OncoOmics approach, the PPi network was performed to better understand BC behavior using the String Database and Cytoscape[51,52]. With the indicated cutoff of 0.9, the final interaction network had 258 nodes conformed by 198 (86%) genes from the CS, PCA and PharmGKB/CGI gene sets, and enriched with 60 previously known BC driver genes. Regarding the OncoPrint genes, 65 (89%) nodes integrated this network (Figure 5A). On the other hand, out of the 258 genes that make up our

String PPi network, 16 (6%) genes and 18 edges were part of the OncoPPi BC network[53,54]. The degree centrality made it possible to establish a significant correlation (Spearman $p < 0.05$) between our String PPi network and the OncoPPi BC network (Figure 5B).

Considering the degree centrality and the consensus score of our previous study[28], there was enrichment among sub-networks (Figures 5A and 5B). The average of degree centrality of the 258 nodes network was 48.8; out of the 198 nodes network was 52.7; out of the 65 nodes network was 61.7; and out of the OncoPPi BC network was 124.4. Meanwhile, the average of consensus score of the 258 nodes network was 0.803, out of the 198 nodes network was 0.812, out of the 65 nodes network was 0.833, and out of the OncoPPi BC network was 0.885. Additionally, the second OncoOmics approach was made up of genes with the highest degree centrality (> 52.7) such as *TP53*, *AKT1*, *SRC*, *CREBBP*, *EP300*, *JUN*, *CTNNB1*, *PIK3CA*, *RAC1* and *EGFR*, genes with the highest consensus score such as *TP53*, *ESR1*, *CCND1*, *BRCA2*, *BRCA1*, *ERBB2*, *CHEK2*, *AR*, *MYC* and *PTEN*, and genes with both of them such as *TP53*, *ESR1*, *CCND1*, *ERBB2*, *PTEN*, *CDKN1B*, *ATM*, *AKT1*, *STAT3*, *CDH1* and *EGFR* (Table S19).

Protein expression analysis. The third OncoOmics approach was related to the expression analysis of the 230 proteins. Figure 6A shows 43 proteins with significant high expression (Z-scores ≥ 2) and low expression (Z-scores ≤ -2) analyzed with the reverse-phase protein array (RPPA) and mass spectrometry, according to TCGA. The top ten proteins with the highest expression levels in a cohort of 994 individuals were *ERBB2*, *SERPINE2*, *CDH2*, *CCND1*, *EGFR*, *ERCC1*, *IRS1*, *NOTCH1*, *ERBB3* and *INPP4B*, and the ones with the lowest expression levels were *CDH1*, *ATM*, *JAK2*, *MAPK1*, *AKT1*, *AKT3*, *MAPK14*, *ABL1*, *CTNNB1* and *IRF1* (Table S20). On the other hand, the Human Protein Atlas (HPA) presented a map of the human tissue proteome based on tissue microarray-based immunohistochemistry. HPA has analyzed 202 (88%) of the 230 proteins of our study, classifying the protein expression in high, medium, low and non-detected. As a result, *RAC1*, *GJB2*, *MED1*, *PIK3CA*, *PIK3R3*, *FGFR2*, *HCFC2*,

MAP2K4, *NQO2* and *RAC3* were proteins with high and medium expression in normal tissue, and low and non-detected expression in BC tissue, acting as tumor suppressor genes. Meanwhile, *CDK2*, *CYP2D6*, *NCOR1*, *RRM1*, *FOXA1* and *TOP2A* were proteins with high and medium expressions in BC tissue, and low and non-detected expressions in normal tissue, acting as oncogenes (Figure 6B and Table S21). Lastly, according to the HPA, Figure 6C shows the overall survival analysis of *RAD51*, *PERP* and *MORC4* as BC biomarkers with unfavorable prognosis and $p < 0.001$ (Table S22)[55,56]. All these altered proteins made up the third OncoOmics approach.

Breast cancer dependency map. The fourth OncoOmics approach consisted in identifying genes that are essential for cancer cell proliferation and survival performing systematic loss-of-function screens in a large number of well-annotated cancer cell lines and BC cell lines representing the tumor heterogeneity[18–21]. Figure 7A shows the distribution of dependency scores of 227/230 genes through DEMETER2, an analytical framework for analyzing genome-scale RNAi loss-of-function screens in 73 BC cell lines (Table S23). Our results showed 563 dependencies with at least one score ≤ -1 in 57 (25%) essential genes. The top 10 genes with the greatest number of significant dependency scores in BC cell lines were *RPL5* (68; 93%), *SF3B1* (67; 92%), *RPA1* (61; 84%), *RRM1* (53; 73%), *BUB1B* (26; 36%), *RPA3* (25; 34%), *RAD51* (23; 32%), *PPP2R1A* (21; 29%), *CHD4* (19; 26%) and *POLE* (13, 18%). At the same time, Figure 7A shows the distribution of dependency scores of 217/230 genes through CERES, an analytical framework for analyzing genome-scale CRISPR-Cas9 loss-of-function screens in 28 BC cell lines (Table S24). Our results showed 310 dependencies with at least one score ≤ -1 in 34 (16%) essential genes. The top 10 genes with the greatest number of significant dependency score in BC cell lines were *RPA1* (27; 96%), *RRM1* (27; 96%), *TOP2A* (26; 93%), *BUB1B* (24; 86%) *CTCF* (24; 86%), *POLE* (23; 82%), *SF3B1* (19; 68%), *RPL5* (17; 61%), *CCND1* (13; 46%) and *SOD2* (13; 46%). Figure 7B shows the distribution of dependency scores of DEMETER2 and CERES per molecular subtype. The genome-scale RNAi loss-of-function screens detected 165 (29%) dependencies in 19 Her2-enriched cell lines (ratio = 8.7), 110 (20%)

in 13 luminal A cell lines (8.5), 57 (10%) in 7 luminal B cell lines (8.1), and 231 (41%) in 34 basal-like cell lines (6.8), whereas the genome-scale CRISPR-Cas9 loss-of-function screens detected 85 (27%) dependencies in 7 luminal A cell lines (ratio = 12.1), 176 (15%) in 16 basal-like cell lines (11), and 49 (16%) in 5 Her2-enriched cell lines (9.8). Figure 7C shows violin plots of dependencies per molecular subtype. DEMETER2 has detected a greatest number of significant dependencies in basal-like, followed by Her2-enriched, luminal A and luminal B, whereas CERES has detected a greatest number of significant dependencies in basal-like, followed by luminal A and Her2-enriched. Figure 7D shows a Venn diagram of 66 essential genes with at least one significant dependency in different molecular subtypes, where 22 were strongly selective genes, 26 were common essential genes, and 5 were both of them in all cancer cell lines (Figure 7E).

OncoOmics approaches to reveal essential genes in BC. Figure 8A shows a Venn diagram integrated by the OncoOmics essential genes, the most relevant genes of the CS, PCA and PharmGKB/CGI gene sets per approach. *RAC1*, *AKT1*, *CCND1*, *PIK3CA* and *ERBB2* were relevant genes in all OncoOmics approaches; *CDH1*, *MAPK14*, *TP53*, *MAPK1*, *SRC* and *RAC3* were relevant genes in the OncoPrint, networking and protein expression analyses; *PLCG1* and *GJB2* were relevant genes in the OncoPrint, networking and DepMap analyses; *MED1*, *TOP2A* and *GATA3* were relevant genes in the DepMap, OncoPrint and protein expression analyses; *BCL2*, *CTNNB1*, *EGFR* and *CDK2* were relevant in the DepMap, networking and protein expression analyses; *EP300* and *CREBBP* were relevant in the networking and the OncoPrint analyses; *PTEN*, *MRE11*, *CDKN2A*, *WWNTR1*, *ABL1*, *BRCA2*, *NF2*, *AKT3*, *ARDID1A* and *RB1* were relevant in the OncoPrint and protein expression analyses; *RPA1*, *TOP3A*, *FGFR1*, *SF3B1*, *ATR*, *KRAS*, *PDPK1*, *RELA*, *SMARCE1*, *SPOP*, *CCNK* and *MDM4* were relevant in the DepMap and OncoPrint analyses; *CDKN1B*, *LCK* and *NOTCH1* were relevant in the networking and protein expression analyses, *CDK4* and *ESR1* were relevant in the DepMap and networking analyses; and *RAD51*, *IRS1*, *FGFR2*, *JAK2*, *RRM1*, *PIK3R3*, *FOXA1* and *ERBB3* were relevant in the DepMap and protein expression analyses (Table S25).

Out of the 144 OncoOmics essential genes, 21% were oncogenes, 24% were tumor suppressor genes, 50% were tier 1, according to the Cancer Gene Census (COSMIC)[60], and 59% were driver genes in other types of cancer, according to The Network of Cancer Genes[61] (Figure 8B). On the other hand, *FGF4*, *INPP4B*, *WWNTR1*, *MAPK8*, *PIGB*, *RRM1*, *CASP8*, *FCGR2A*, *SMARCB1*, *SF3B1* and *CTCF* were cancer immunotherapy genes[62]; *LCK*, *MAP3K1*, *EGFR*, *SRC*, *FGFR1*, *MAP2K4*, *ABL1*, *ERBB3*, *FGFR2* and *ERBB2* were kinome genes[63]; *CDKN1B*, *BLM*, *BUB1B* and *BARD1* were cell cycle genes[64]; *XRCC1*, *RAD51*, *ERCC1*, *NBN*, *ERCC2*, *MLH1*, *BRCA2*, *PMS2*, *RPA1* and *PALB2* were DNA repair genes[65]; lastly, *YAP1*, *CDKN2A*, *GNL3*, *ZC3H13*, *JUN*, *LARP7*, *KMT2C*, *HMGB1*, *GSTP1* and *GRB2* were RNA-binding proteins (RBPs) (Figure 8C and Table S26)[66].

Figure 8D shows a circos plot of the 48 (33%) OncoOmics essential genes that are hallmarks of cancer. The top 10 genes with the greatest number of interactions with the hallmarks of cancer were *TP53*, *CTNNB1*, *PTEN*, *KRAS*, *AKT1*, *RAC1*, *EGFR*, *ABL1*, *RB1* and *NOTCH1*. Suppression of growth was promoted by *AKT1*, *CTNNB1*, *PTEN*, *RB1* and *TP53*; escaping immune response to cancer was promoted by *CTNNB1*, *EGFR* and *RAC1*, and suppressed by *ABLI*, *PTEN* and *TP53*; cell replicative immortality was promoted by *CTNNB1*, *KRAS* and *NOTCH1*, suppressed by *PTEN*, and promoted/suppressed by *TP53*; tumor promoting inflammation was promoted by *KRAS* and suppressed by *TP53*; metastasis was promoted by *ABL1*, *CTNNB1*, *EGFR*, *KRAS*, *RAC1* and *RB1*, suppressed by *PTEN* and *TP53*, and promoted/suppressed by *AKT1*; angiogenesis was promoted by *ABL1*, *CTNNB1*, *EGFR*, *KRAS*, *NOTCH1* and *RAC1*, suppressed by *TP53* and promoted/suppressed by *AKT1*; genome instability was promoted by *ABL1* and *RB1*, and suppressed by *AKT1*, *CTNNB1*, *PTEN*, *RAC1* and *TP53*; escaping programmed cell death was promoted by *AKT1*, *CTNNB1*, *EGFR*, *NOTCH1* and *RAC1*, suppressed by *PTEN*, and promoted/suppressed by *KRAS*, *RB1* and *TP53*; change of cellular energetics was promoted by *ABL1*, *AKT1*, *CTNNB1*, *EGFR*, *KRAS*,

335 *NOTCH1*, *PTEN*, *RB1* and *TP53*; finally, proliferative signaling was promoted by *ABL1*, *AKT1*,
336 *CTNNB1*, *EGFR*, *KRAS*, *NOTCH* and *RAC1* (Table S27).
337
338 **Enrichment map of the OncoOmics essential genes in BC.** Figure 8E shows the enrichment
339 map of the 144 OncoOmics essential genes in BC. g:Profiler searches for a collection of gene
340 sets representing pathways, networks, GO terms and disease phenotypes[67]. The most significant
341 GO: biological process with a FDR < 0.001 was positive regulation of macromolecule
342 metabolic process (Table S28); the most significant GO: molecular function was
343 phosphatidylinositol 3-kinase activity (Table S29); the most significant Reactome pathway was
344 generic transcriptor pathway (Table S30)[68]; additionally, the most significant disease, according
345 the Human Phenotype Ontology, was breast carcinoma (Table S31)[69]. Subsequently, g:Profiler
346 annotations were analyzed with the EnrichmentMap software and visualized using Cytoscape,
347 in order to generate network interactions of the most relevant GO: biological processes (Figure
348 S2) and Reactome pathways (Figure 9) related to immune system, tyrosine kinase, cell cycle
349 and DNA repair pathways[52,67].
350
351 **Precision medicine.** Figure 10 shows the current status of clinical trials for BC, according to
352 the Open Targets Platform[70]. There are 111 drugs that are being analyzed in 3151 clinical trials
353 in 39/230 genes. The top 10 genes with the highest number of clinical trials in process or
354 completed were *TUBB1*, *ERBB2*, *ESR1*, *TOP2A*, *EGFR*, *ESR2*, *VEGFA*, *CDK4*, *POLE* and
355 *RRM1*. The greatest number of clinical trials was in phase 2. Small molecules were the most
356 analyzed type of drug, followed by antibodies and proteins. Lastly, the target classes with the
357 greatest number of clinical trials were tyrosine kinases, structural proteins and nuclear hormone
358 receptors (Table S32).
359
360 Regarding precise guidelines for the application of BC pharmacogenomics in clinical practice,
361 PharmGKB details 154 clinical annotations in 70/230 (30%) genes (Table S33)[41–44]; the CGI

details 76 clinical annotations in 26/230 (11%) genes (Table S34)[71]; and PCA details 648 clinical annotations in 14/230 (6%) genes (Table S35)[72].

Additionally, Figure S3 shows a drug-gene interaction matrix conformed by 109 clinical annotations in phase 4, according to the OTP; 9 clinical annotations in levels 1A, 2A and 2B, according to PharmGKB; 9 clinical annotations approved by the US Food and Drug Administration (FDA), according to CGI; and 648 clinical annotations, according to PCA.

DISCUSSION

In this study we proposed a compendium of OncoOmics approaches that analyze genetic alterations, protein expression, signaling pathways, PPi networks, enrichment maps, gene ontology and dependency maps in three gene sets. The first gene set was taken from our previous study where we developed a Consensus Strategy that was proved to be highly efficient in the recognition of BC pathogenic genes[28]. The second gene set was taken from several studies of PCA, which provides a panoramic view of the oncogenic processes that contributes to BC progression[3,12,30–36]. The third gene set was taken from the CGI and PharmGKB. On the one hand, the CGI flags genomic biomarkers of drug response with different levels of clinical relevance[38]. On the other hand, PharmGKB collects clinical annotations applied in BC patients and taken from the NCCN, ESMO, CPNDS, DPWG and CPIC guidelines[41–44]. Finally, the compendium of these 230 potential essential genes in BC was analyzed through four different OncoOmics approaches.

The first OncoOmics approach consisted in the analysis of genetic alterations using the PCA data[45,46]. The frequency of genetic alterations in the CS (average = 1.2), PCA (1.3) and PharmGKB/CGI (1.1) gene sets were higher than the non-cancer gene set (0.4) and the previously known BC driver genes (0.8). This means that these 230 genes had a greater number

of genetic alterations and might be strongly associated with BC (Figure 1A). The most common genetic alterations in a cohort of 994 individuals were mRNA upregulation, CNV amplification and missense mutations. Molecular subtypes with the greatest number of genetic alterations were basal-like, Her2-enriched, luminal B, normal-like and luminal A, whereas tumor stages with the greatest number of genetic alterations were T2, T3, T1 and T4 (Figures 1B-F). Genes with the greatest number of genetic alterations per subtype were *PIK3CA* in luminal A, *CCND1* in luminal B, *TP53* in basal-like and normal-like, and *ERBB2* in Her2-enriched (Figure 2A), whereas *PIK3CA* was the most altered gene in stage T1, *TP53* in stages T2 and T3, and *ERBB2* in stage T4 (Figure 2B).

After a thorough analysis of genetic alterations in the 230 genes, the first OncoOmics approach was generated by an OncoPrint conformed by the top 73 genes with the greatest number of genetic alterations and with a frequency of alterations greater than the average (> 86) (Figure 3A). The top 10 most altered genes were *PIK3CA*, *TP53*, *MDM4*, *CCND1*, *NBN*, *MED1*, *CREBBP*, *PALB2*, *ERBB2* and *SPOP*[3,12,30–36].

Subsequently, the enrichment analysis of signaling pathways was carried on taking into account all genetic alterations in the 230 genes using David Bioinformatics Resource and KEGG[47,50]. The signaling pathways with the greatest number of genetic alterations per intrinsic molecular subtype were Jak-STAT in luminal A, Wnt in luminal B, p53 in basal-like, ERBB in Her2-enriched and Hippo in normal-like (Figure 4B); and per tumor stage were Wnt in stages T1, T2 and T3, and thyroid hormone in stage T4 (Figure 4D).

Regarding the previously mentioned signaling pathways, Jak-STAT is involved in the control of processes, such as stem cell maintenance, hematopoiesis and inflammatory response. However, the mechanism underlying inappropriate Jak-STAT pathway activation is not well-known in BC[73]. The Wnt signaling pathway actively functions in embryonic development and helps in homeostasis in mature tissues by regulating cell survival, migration, proliferation and polarity[74].

418 The p53 tumor suppressor is the most frequently mutated gene in human cancer[75], and acting as
419 a transcription factor, the p53 signaling pathway plays a critical role in growth-inhibition,
420 apoptosis, cell migration and angiogenesis[76]. The ERBB signaling pathway members form cell-
421 surface receptors with extracellular domains yielding ligand-binding specificity[77]. Downstream
422 signaling proceeds via tyrosine phosphorylation mediating signal transduction events that
423 control cell survival, migration and proliferation. However, aberrant ERBB activation can
424 increase transcriptional expression[78]. The Hippo pathway plays important roles in immune
425 response, stem cell function and tumor suppression. However, alterations in this pathway are
426 involved in the BC tumorigenesis and metastasis[79]. Lastly, the thyroid hormone signaling
427 pathway is an important regulator of growth and metabolism. Nevertheless, deregulation of the
428 T3 hormone levels could promote abnormal responsiveness of mammary epithelial cells
429 developing BC[80].
430
431 The second OncoOmics approach consisted in the PPi network analysis and its validation with
432 the OncoPPi BC network. According to Li *et al*. and Ivanov *et al*.[54,81], PPi with therapeutic
433 significance can be revealed by the integration of cancer genes into networks. PPi regulates
434 essential oncogenic signals to cell proliferation and survival, and thus, represents potential
435 targets for drug development and drug discovery. Regarding our networking analysis, the final
436 interaction network consisted in 258 nodes with an average of degree centrality of 48.8 and an
437 average of consensus scoring of 0.803[28]; the sub-network integrated by 198 of 230 nodes had
438 52.7 of degree centrality and 0.812 of consensus scoring; finally, the sub-network integrated by
439 65 of 73 genes with the greatest number of genetic alterations had 61.7 of degree centrality and
440 0.833 of consensus scoring. Hence, a sub-network of genes with greatest number of genetic
441 alterations presented a greater degree centrality and consensus scoring, suggesting that there is
442 strong correlation between these genes and BC. Additionally, the oncogenomics validation
443 showed a significant correlation between our String PPi network (Figure 5A) and the OncoPPi
444 BC network (Figure 5B), identifying 16 nodes strongly associated with BC[28]. The second
445 OncoOmics approach was made up with the top 40 genes with the highest degree centrality and

446 consensus scoring, such as *TP53*, *ESR1*, *CCND1*, *ERBB2*, *PTEN*, *CDKN1B*, *ATM*, *AKT1*,
447 *STAT3*, *CDH1* and *EGFR*.
448
449 The third OncoOmics approach was related to the BC proteome. More than 500 proteins have
450 been identified as strongly involved in oncogenesis. Loss of expression, overexpression or
451 expression of dysfunctional proteins contribute to uncontrolled tumor growth, causing
452 chromosomal rearrangements, gene amplification and ungoverned methylation[59]. Regarding our
453 230 proteins, 43 showed significant high and low expression ($p < 0.001$), according to TCGA.
454 The top ten proteins with the highest expression levels were *ERBB2*, *SERPINE2*, *CDH2*,
455 *CCND1*, *EGFR*, *ERCC1*, *IRS1*, *NOTCH1*, *ERBB3* and *INPP4B*, whereas the top ten proteins
456 with the lowest expression levels were *CDH1*, *ATM*, *JAK2*, *MAPK1*, *AKT1*, *AKT3*, *MAPK14*,
457 *ABL1*, *CTNNB1* and *IRF1*. On the other hand, the HPA has analyzed 202 of 230 proteins, where
458 *FOXA1*, *TOP2A*, *CDK2*, *CYP2D6*, *NCOR1* and *RRM1* were involved in oncogenic processes,
459 and *RAC1*, *GJB2*, *MED1*, *PIK3CA*, *PIK3R3*, *FGFR2*, *HCFC2*, *MAP2K4*, *NQO2* and *RAC3*
460 were involved in tumor suppression processes. Lastly, genes with unfavorable prognosis in BC
461 were *RAD51*, *PERP* and *MORC4* (Figure 6)[55,56]. The compendium of all these 60 proteins with
462 significant high and low expression made up the third OncoOmics approach.
463
464 The fourth OncoOmics approach was related to the BC dependency map. According to
465 Tsherniak *et al*., the mutations that trigger the growth of cancer cells also confer specific
466 vulnerabilities that normal cells lack, and these dependencies are compelling therapeutic
467 targets[82]. The cancer dependency map identifies essential genes in proliferation and survival of
468 well-annotated cell lines through systematic loss-of-function screens[18–21]. On the one hand,
469 DETEMER2 analyzed the genome-scale RNAi loss-of-function screens. The top 10 genes with
470 the greatest number of significant dependency scores in BC cell lines were *RPL5*, *SF3B1*,
471 *RPA1*, *RRM1*, *BUB1B*, *RPA3*, *RAD51*, *PPP2R1A*, *CHD4* and *POLE*. On the other hand,
472 CERES analyzed the genome-scale CRISPR-Cas9 loss-of-function screens. The top 10 genes
473 with the greatest number of significant dependencies in BC cell lines were *RPA1*, *RRM1*,

TOP2A, *BUB1B*, *CTCF*, *POLE*, *SF3B1*, *RPL5*, *CCND1* and *SOD2* (Figure 7A). Additionally, the fourth OncoOmics approach was made up of genes with significant dependencies in BC cell lines and all cancer cell lines. *PLCG1*, *CDK4*, *KRAS*, *SPOP*, *CTNNB1*, *EGFR*, *AKT1*, *JAK2*, *MDM4*, *FGFR1*, *IRS1*, *BCL2*, *RELA*, *GATA3*, *PIK3CA*, *PIK3RE*, *PIK3CB*, *FOXA1*, *ERBB3*, *FGFR2*, *ESR1* and *ERBB2* were strongly selective genes, whereas *CDH4*, *TOP2A*, *GNL3*, *RBBP8*, *TOP3A*, *SMARCB1*, *UROD*, *RPL5*, *RAD51*, *PDPK1*, *CCNK*, *SF3B1*, *CDC42*, *ERCC2*, *BUB1B*, *CTCF*, *MAX*, *CCND1*, *BARD1*, *RAC1*, *RPA3*, *SMARCE1*, *PPP2R1A*, *POLE*, *RPA1* and *GRB2* were common essential genes, and *SOD2*, *CDK2*, *ATR*, *RRM1* and *MED1* were both (Figure 7E).

Subsequently, the compendium of the most relevant genes per OncoOmics approach reveals the 144 OncoOmics essential genes in BC (Figure 8A). *RAC1*, *AKT1*, *CCND1*, *PIK3CA* and *ERBB2* were relevant genes in all OncoOmics approaches; *CDH1*, *MAPK14*, *TP53*, *MAPK1*, *SRC* and *RAC3* were relevant genes in the OncoPrint, networking and protein expression analyses; *PLCG1* and *GJB2* were relevant genes in the OncoPrint, networking and DepMap analyses; *MED1*, *TOP2A* and *GATA3* were relevant genes in the DepMap, OncoPrint and protein expression analyses; and *BCL2*, *CTNNB1*, *EGFR* and *CDK2* were relevant in the DepMap, networking and protein expression analyses. Lastly, the top 10 genes with the greatest number of interactions with the hallmarks of cancer were *TP53*, *CTNNB1*, *PTEN*, *KRAS*, *AKT1*, *RAC1*, *EGFR*, *ABL1*, *RB1* and *NOTCH1* (Figure 8D).

According to Reimand *et al*., g:Profiler lets us know the enrichment map of the 144 OncoOmics essential genes in BC[83]. The most significant GO: biological process was the positive regulation of macromolecule metabolic process, the GO: molecular function was phosphatidylinositol 3-kinase activity, the Reactome pathway was generic transcriptor pathway, and the most significant Human Phenotype Ontology term was breast carcinoma[69]. Subsequently, the most relevant network interactions of the GO: biological process and the Reactome pathways were

related to immune system, tyrosine kinase, cell cycle and DNA repair terms (Figures 9 and S2)[52,67].

There is currently great enthusiasm about immunotherapeutic strategies to treat BC. The first approval of an immune checkpoint blockade agent for treatment of BC came in March 2019 when the anti-PD-L1 antibody atezolizumab was approved to be used in combination with nab-paclitaxel for patients with triple-negative BC[84]. 17 OncoOmics essential genes were associated with immunotherapy[62]. Kinases have been recognized as highly tractable targets for BC treatment due to their druggability and critical roles they play in regulating cellular migration, differentiation, growth and survival[85]. 17 OncoOmics essential genes in BC were kinome genes[63]. The cell cycle comprises a series of tightly controlled events that drive cell division and the DNA replication[86]. 12 OncoOmics essential genes in BC were involved in cell cycle[64]. DNA repair constitutes several signaling pathways working in concert to eliminate DNA lesions and maintain genome stability. Defective components in DNA repair machinery are an underlying cause for the development of BC[87]. 19 OncoOmics essential genes in BC were involved in the DNA repair system[65]. RBPs are key players in post-transcriptional events[88,89]. Three recent reports using high-throughput bioinformatics profiling of thousands of tumors now reveal a consistent pattern of alterations in RBPs expression levels across different cancer types[90–92]. Lastly, 11 OncoOmics essential genes were RBPs (Figure 8C)[66].

Precision medicine provides BC patients with the most appropriate diagnostics and targeted therapies based on the omics profile and other predictive and prognostic tests. Additionally, it is relevant to know the composition of their breast tissue, tumor microenvironment, comorbid conditions and lifestyle[93].

The OTP is an available resource for the integration of genetics, omics and chemical data to aid systematic drug target identification and prioritization[70]. Currently, there are 111 drugs that are being analyzed in 3151 clinical trials in 39 of the 230 genes. Most of clinical trials are in phase

529 2; most of the analyzed drugs are small molecules; and most of target classes belong to tyrosine
530 kinases. Finally, the top ten genes with the greatest number of clinical trials in process or
531 completed are *TUBB1*, *ERBB2*, *ESR1*, *TOP2A*, *EGFR*, *ESR2*, *VEGFA*, *CDK4*, *POLE* and
532 *RRM1*[70] (Figure 10).
533
534 PharmGKB collects the precise guidelines for the application of pharmacogenomics in clinical
535 practice[41–44]. This database details 154 clinical annotations associated with 70 genes in BC. The
536 CGI is a platform that annotates clinical evidence and tumor variants that constitute state-of-art
537 biomarkers of drug response. The CGI details 76 clinical annotations associated with 26 genes
538 in BC[71]. According to TCGA, PCA details 648 clinical annotations associated with 14 genes in
539 BC[72]. Lastly, the drug-gene interaction matrix is a compendium of the most relevant clinical
540 annotations made up of 32 genes and 51 drugs in order to facilitate the treatment of patients
541 with BC (Figure S3).
542
543 In conclusion, since BC is a complex and heterogeneous disease, the study of different
544 OncoOmics approaches is an effective way to reveal essential genes to better understand the
545 molecular landscape of processes behind oncogenesis, and to develop better therapeutic
546 treatments focused on pharmacogenomics and precision medicine.
547
548
549 **METHODS**
550
551 **OncoPrint of genetic alterations according to the Pan-Cancer Atlas.** PCA has reported the
552 clinical data of 1084 individuals with BC and it can be visualized in the Genomic Data
553 Commons of the National Cancer Institute (https://gdc.cancer.gov/) and in the cBioPortal
554 (http://www.cbioportal.org/)[45,46]. The clinical annotations were age, pTNM classification, tumor
555 type, tumor stage and race/ethnicity.
556

Additionally, PCA has reported genetic alterations (mRNA upregulation, mRNA downregulation, CNV amplification, CVN deep deletion, missense mutation, truncating mutation, inframe mutation and fusion gene) in 994 individuals. Putative mutations were analyzed through exome sequencing, CNVs through the Genomic Identification of Significant Targets in Cancer (GISTIC 2.0)[94,95], and mRNA expression through RNA Seq V2. We analyzed five gene sets in order to compare the average frequency of genetic alterations among them. The first gene set (n = 177) was integrated by the non-cancer genes[96]. We calculated the OncoScore of non-cancer genes, taking out all genes from our study. The second gene set (n = 119) was the BC driver genes, according to The Network of Cancer Genes[61]. The third gene set (n = 84) was taken from our previous study where we developed a Consensus Strategy of prioritized genes related to BC pathogenesis[28]. The fourth gene set (n = 85) was made up of genes associated with BC development, according to several PCA studies[30,31,57]. The fifth gene set (n = 91) consisted of BC biomarkers and druggable enzymes taken from PharmGKB and the CGI (Table S2)[37,38,40]. Finally, the significant differentiation of the average frequency of genetic alterations among gene sets was analyzed (p-value < 0.001).

The OncoOmics approaches were performed in 230 genes conformed by the CS, PCA and PharmGKB/CGI gene sets. Firstly, we calculated the percentage and ratio of genetic alterations per intrinsic molecular subtype and tumor stage, and we established a ranking of genes with the greatest number of different genetic alterations. Subsequently, we performed an OncoPrint of genes with more genetic alterations than the average. The final list of genes made up the first OncoOmics approach.

Pathway enrichment analysis. The enrichment analysis of signaling pathways was performed using David Bioinformatics Resource to obtain integrated information from KEGG[47–50]. It was carried on in the 230 genes, taking into account terms with a significant FDR < 0.01. After that, genetic alterations that comprise each signaling pathway were analyzed, taking into account the molecular subtype and tumor stage of individuals from PCA. Circos plots and violin plots were

585 designed to visualize all data. Lastly, in order to compare the ratio of genetic alterations among
586 subtypes and tumor stages, normalization was carried out dividing the number of genetic
587 alterations by the number of individuals per subtype and tumor stage. Regarding molecular
588 subtypes, 499 individuals were luminal A, 197 were luminal B, 171 were basal-like, 78 were
589 Her2-enriched and 36 were normal-like, and regarding tumor stage, 255 were stage T1, 586
590 were stage T2, 113 were stage T3 and 103 were stage T4.
591
592 **Protein-protein interaction network.** The PPi network with a highest confidence cutoff of 0.9
593 and zero node addition was created using the String Database, which takes into account
594 predicted and known interactions[51]. The confidence scoring is the approximate probability that a
595 predicted link exists between two enzymes in the same metabolic map, whereas the degree
596 centrality of a node means the number of edges the node has to other nodes in a network. The
597 centrality indexes calculation and network visualization were analyzed through the Cytoscape
598 software[52]. Genes with the highest degree centrality, consensus score and sub-networks were
599 differentiated by colors in the PPi network. On the other hand, OncoPPi
600 (http://oncoppi.emory.edu/) reports the development of a cancer-focused PPi network,
601 identifying more than 260 high-confidence cancer-associated PPi[53,54]. In addition, the OncoPPi
602 BC network consisted of 16 genes and 18 PPi experimentally analyzed in BC cell lines[53,54]. The
603 correlation of the degree centrality by means of Spearman p-value test between our String PPi
604 network and the OncoPPi BC network allowed for the validation of all the high-confidence BC-
605 focused PPi analyzed in cell lines[28]. Lastly, genes with the highest degree centrality and
606 consensus scoring made up the second OncoOmics approach.
607
608 **Protein expression analysis.** TCGA has reported the protein expression data of 994 individuals
609 with BC through RPPA and mass spectrometry by the Clinical Proteomic Tumor Analysis
610 Consortium (CPTAC), and it can be visualized in the cBioPortal[45,46]. We analyzed the protein
611 expression of 230 genes (CS, PCA and PharmGKB/CGI gene sets) where Z-scores $\geq$ 2 mean a
612 significant high protein expression and Z-scores $\leq$ -2 mean a significant low protein expression.

On the other hand, the Human Protein Atlas (https://www.proteinatlas.org/) explains the diverse molecular signatures of proteomes in the human tissues based on an integrated omics approach that involves quantitative transcriptomics and tissue microarray-based immunohistochemistry[56,58,59]. We compared the protein gene levels (high, medium, low and non-detected) of our 230 genes between normal and BC tissues. Finally, we analyzed the overall survival curve of our 230 genes and reveled all biomarkers with significant unfavorable prognostic ($p < 0.001$)[55,56]. All genes with the altered protein expression made up the third OncoOmics approach.

Breast cancer dependency map. The DepMap project (https://depmap.org/portal/) is a collaboration between the Broad Institute and the Welcome Sanger Institute. Multiple genetic or epigenetic changes provide cancer cells with specific vulnerabilities that normal cells lack. Even though the landscape of genetic alterations has been extensively studied to date, we have limited understanding of the biological impact of these alterations in the development of specific tumor vulnerabilities, which triggers a limited use of precision medicine in the clinical practice worldwide. Therefore, the main goal of DepMap is to create a comprehensive preclinical reference map connecting tumor features with tumor dependencies to accelerate the development of precision treatments[18–21].

In order to identify essential genes for BC cell proliferation and survival, DepMap performed systematic loss-of-function screens in a large number of well-annotated BC cell lines representing the tumor heterogeneity and their molecular subtypes. The DEMETER2 algorithm was applied to analyze genome-scale RNAi loss-of-function screens in 73 BC cell lines and 711 cancer cell lines, whereas the CERES algorithm was applied to analyze genome-scale CRISPR-Cas9 loss-of-function screens in 28 BC cell lines and 558 cancer cell lines[19,21]. In addition to existing cell lines, the Cancer Cell Line Encyclopedia (CCLE) project will greatly expand the collection of characterized cell lines to improve precision treatments[97].

Regarding dependency scores, a lower score means that a gene is more likely to be dependent in a specific cancer cell line. A score of 0 means that a gene is not essential, whereas a score of -1 corresponds to the median of all common essential genes. A strongly selective gene means that its dependency is at least 100 times more likely to have been sampled from a skewed distribution than a normal distribution. Lastly, a common essential gene is when in a pan-cancer screen its gene ranks in the top most depleting genes in at least 90% of cell lines[18]. All genes with a dependency score ≤ -1 made up the fourth OncoOmics approach.

Enrichment map of the OncoOmics essential genes in BC. The pathway enrichment analysis gives scientists curated interpretation of gene lists generated from genome-scale experiments[67]. The OncoOmics essential genes in BC were analyzed by using g:Profiler (https://biit.cs.ut.ee/gprofiler/) in order to obtain significant annotations (FDR < 0.001) related to GO terms, pathways, networks and disease phenotypes. Subsequently, g:Profiler annotations were analyzed with the EnrichmentMap software in order to generate network interactions of the most relevant GO: biological processes and Reactome pathways, and these networks were visualized using Cytoscape[52,67].

Precision medicine. We analyzed drug-gene interactions for BC using four selective databases: 1) OTP[70], 2) PharmGKB[37,40], 3) CGI[38], and 4) PCA[98]. The Open Targets Platform (https://www.targetvalidation.org) is comprehensive and robust data integration for access to and visualization of potential drug targets associated with BC. Additionally, this platform shows all drugs in clinical trials associated with BC genes, detailing its phase, status, type and target class[70]. PharmGKB (https://www.pharmgkb.org/) collects complete guidelines for application of pharmacogenomics in clinical practice, according to several consortiums worldwide[41–44]. The CGI (https://www.cancergenomeinterpreter.org/home) flags genomic biomarkers of drug response with different levels of clinical relevance[38]. Finally, PCA reveals genetic alterations, druggable enzymes and clinical annotations in a cohort of 994 individuals[3,12,30–36]. The clinical

annotations of these four databases were analyzed in order to create a drug-gene interaction matrix.

716 **References**
717
718 1. Espinal-Enríquez, J., Fresno, C., Anda-Jáuregui, G. & Hernández-Lemus, E.
719 RNA-Seq based genome-wide analysis reveals loss of inter-chromosomal
720 regulation in breast cancer. *Sci. Rep.* **7**, 1760 (2017).
721 2. Guerrero, S. *et al.* Analysis of Racial / Ethnic Representation in Select Basic and
722 Applied Cancer Research Studies. 1–8 (2018). doi:10.1038/s41598-018-32264-x
723 3. Ding, L. *et al.* Perspective on Oncogenic Processes at the End of the Beginning
724 of Cancer Genomics. *Cell* **173**, 305-320.e10 (2018).
725 4. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence
726 and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.*
727 (2018). doi:10.3322/caac.21492
728 5. Sjöblom, T. *et al.* The consensus coding sequences of human breast and
729 colorectal cancers. *Science (80-.).* (2006). doi:10.1126/science.1133427
730 6. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer
731 types. *Nature* (2013). doi:10.1038/nature12634
732 7. Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across
733 21 tumour types. *Nature* (2014). doi:10.1038/nature12912
734 8. Porta-Pardo, E. *et al.* Comparison of algorithms for the detection of cancer
735 drivers at subgene resolution. *Nat. Methods* (2017). doi:10.1038/nmeth.4364
736 9. Tamborero, D. *et al.* Comprehensive identification of mutational cancer driver
737 genes across 12 tumor types. *Sci. Rep.* **3**, 2650 (2013).
738 10. Lu, C. *et al.* Patterns and functional implications of rare germline variants across
739 12 cancer types. *Nat. Commun.* (2015). doi:10.1038/ncomms10086
740 11. Klijn, C. *et al.* A comprehensive transcriptional portrait of human cancer cell
741 lines. *Nat. Biotechnol.* (2015). doi:10.1038/nbt.3080
742 12. Gao, Q. *et al.* Driver Fusions and Their Implications in the Development and
743 Treatment of Human Cancers. *Cell Rep.* (2018).
744 doi:10.1016/j.celrep.2018.03.050
745 13. Oltean, S. & Bates, D. O. Hallmarks of alternative splicing in cancer. *Oncogene*
746 (2014). doi:10.1038/onc.2013.533
747 14. Stricker, T. P. *et al.* Robust stratification of breast cancer subtypes using
748 differential patterns of transcript isoform expression. *PLoS Genet.* (2017).
749 doi:10.1371/journal.pgen.1006589
750 15. Lawrence, R. T. *et al.* The Proteomic Landscape of Triple-Negative Breast
751 Cancer. *Cell Rep.* (2015). doi:10.1016/j.celrep.2015.03.050
752 16. Sogawa, K. *et al.* Identification of a novel serum biomarker for pancreatic cancer,
753 C4b-binding protein α-chain (C4BPA) by quantitative proteomic analysis using
754 tandem mass tags. *Br. J. Cancer* (2016). doi:10.1038/bjc.2016.295
755 17. Rubio-Perez, C. *et al.* In Silico Prescription of Anticancer Drugs to Cohorts of 28
756 Tumor Types Reveals Targeting Opportunities. *Cancer Cell* **27**, 382–396 (2015).
757 18. Tsherniak, A. *et al.* Defining a Cancer Dependency Map. *Cell* (2017).
758 doi:10.1016/j.cell.2017.06.010
759 19. Meyers, R. M. *et al.* Computational correction of copy number effect improves
760 specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.*
761 (2017). doi:10.1038/ng.3984
762 20. Stransky, N. *et al.* Pharmacogenomic agreement between two cancer cell line
763 data sets. *Nature* (2015). doi:10.1038/nature15736
764 21. McFarland, J. M. *et al.* Improved estimation of cancer dependencies from large-
765 scale RNAi screens using model-based normalization and data integration. *Nat.*

Commun. (2018). doi:10.1038/s41467-018-06916-5

22. Shah, P. *et al.* Integrated Proteomic and Glycoproteomic Analyses of Prostate Cancer Cells Reveal Glycoprotein Alteration in Protein Abundance and Glycosylation. *Mol. Cell. Proteomics* (2015). doi:10.1074/mcp.M115.047928
23. Bernard, P. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).
24. Kumar, R., Sharma, A. & Tiwari, R. K. Application of microarray in breast cancer: An overview. *J. Pharm. Bioallied Sci.* **4**, 21–6 (2012).
25. Banerji, S. *et al.* Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* **486**, 405–409 (2012).
26. López-Cortés, A. *et al.* Breast cancer risk associated with gene expression and genotype polymorphisms of the folate-metabolizing MTHFR gene: a case-control study in a high altitude Ecuadorian mestizo population. *Tumor Biol.* **36**, 6451–6461 (2015).
27. Prat, A., Ellis, M. J. & Perou, C. M. Practical implications of gene-expression-based assays for breast oncologists. *Nature Reviews Clinical Oncology* **9**, 48–57 (2012).
28. López-Cortés, A. *et al.* Gene prioritization, communality analysis, networking and metabolic integrated pathway to better understand breast cancer pathogenesis. *Sci. Rep.* **8**, 16679 (2018).
29. López-cortés, A. *et al.* Mutational Analysis of Oncogenic AKT1 Gene Associated with Breast Cancer Risk in the High Altitude Ecuadorian Mestizo Population. **2018**, (2018).
30. Huang, K. lin *et al.* Pathogenic Germline Variants in 10,389 Adult Cancers. *Cell* 355–370 (2018). doi:10.1016/j.cell.2018.03.039
31. Bailey, M. H. *et al.* Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **173**, 371-385.e18 (2018).
32. Thorsson, V. *et al.* The Immune Landscape of Cancer. *Immunity* 1–19 (2018). doi:10.1016/j.immuni.2018.03.023
33. Liu, J. *et al.* An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* (2018). doi:10.1016/j.cell.2018.02.052
34. Sanchez-Vega, F. *et al.* Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **173**, 321-337.e10 (2018).
35. Hoadley, K. A. *et al.* Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**, 291-304.e6 (2018).
36. Ellrott, K. *et al.* Scalable Open Science Approach for Mutation Calling of Tumor Exomes Using Multiple Genomic Pipelines. *Cell Syst.* (2018). doi:10.1016/j.cels.2018.03.002
37. Thorn, C. F., Klein, T. E. & Altman, R. B. PharmGKB: The pharmacogenomics knowledge base. *Methods Mol. Biol.* **1015**, 311–320 (2013).
38. Tamborero, D. *et al.* Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. 1–8 (2018). doi:10.1186/s13073-018-0531-8
39. Tejera, E. *et al.* Consensus strategy in genes prioritization and combined bioinformatics analysis for preeclampsia pathogenesis. *BMC Med. Genomics* **10**, 50 (2017).
40. Barbarino, J. M., Whirl-Carrillo, M., Altman, R. B. & Klein, T. E. PharmGKB: A worldwide resource for pharmacogenomic information. *Wiley Interdisciplinary*

Reviews: Systems Biology and Medicine (2018). doi:10.1002/wsbm.1417

41. Ross, C. J. D. *et al.* The Canadian Pharmacogenomics Network for Drug Safety: a model for safety pharmacology. *Thyroid* **20**, 681–7 (2010).
42. Saito, Y. *et al.* CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. *Clinical Pharmacology and Therapeutics* **99**, 36–37 (2016).
43. Swen, J. J. *et al.* Pharmacogenetics: From bench to byte an update of guidelines. *Clin. Pharmacol. Ther.* **89**, 662–673 (2011).
44. European Society for Medicinal Oncology. Breast Cancer: A guide for patients. *European Society for Medical Oncology* (2018).
45. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* (2012). doi:10.1158/2159-8290.CD-12-0095
46. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* (2013). doi:10.1126/scisignal.2004088
47. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
48. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
49. Antonov, A. V, Schmidt, E. E., Dietmann, S., Krestyaninova, M. & Hermjakob, H. R spider: a network-based analysis of gene lists by combining signaling and metabolic pathways from Reactome and KEGG databases. *Nucleic Acids Res.* **38**, W78-83 (2010).
50. Ogata, H. *et al.* KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* **27**, 29–34 (1999).
51. Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447-52 (2015).
52. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–504 (2003).
53. Ivanov, A. A. *et al.* The OncoPPi Portal: an integrative resource to explore and prioritize protein–protein interactions for cancer target discovery. *Bioinformatics* 1–9 (2017). doi:10.1093/bioinformatics/btx743
54. Li, Z. *et al.* The OncoPPi network of cancer-focused protein-protein interactions to inform biological insights and therapeutic strategies. *Nat. Commun.* **8**, (2017).
55. Uhlen, M. *et al.* A pathology atlas of the human cancer transcriptome. *Science (80-.).* (2017). doi:10.1126/science.aan2507
56. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science (80-.).* (2015). doi:10.1126/science.1260419
57. Berger, A. C. *et al.* A Comprehensive Pan-Cancer Molecular Study of Gynecologic and Breast Cancers. *Cancer Cell* 1–16 (2018). doi:10.1016/j.ccell.2018.03.014
58. Uhlen, M. *et al.* Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* (2010). doi:10.1038/nbt1210-1248
59. Uhlén, M. *et al.* A Human Protein Atlas for Normal and Cancer Tissues Based on Antibody Proteomics. *Mol. Cell. Proteomics* (2005). doi:10.1074/mcp.M500279-MCP200
60. Sondka, Z. *et al.* The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nat. Rev. Cancer* 1 (2018).

doi:10.1038/s41568-018-0060-1

61. Repana, D. *et al.* The Network of Cancer Genes (NCG): a comprehensive catalogue of known and candidate cancer genes from cancer sequencing screens. *Genome Biol.* (2019). doi:10.1186/s13059-018-1612-0

62. Patel, S. J. *et al.* Identification of essential genes for cancer immunotherapy. *Nature* (2017). doi:10.1038/nature23477

63. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* (2002). doi:10.1126/science.1075762

64. Bar-Joseph, Z. *et al.* Genome-wide transcriptional analysis of the human cell cycle identifies genes differentially regulated in normal and cancer cells. *Proc. Natl. Acad. Sci.* (2008). doi:10.1073/pnas.0704723105

65. Chae, Y. K. *et al.* Genomic landscape of DNA repair genes in cancer. *Oncotarget* (2016). doi:10.18632/oncotarget.8196

66. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nature Reviews Molecular Cell Biology* (2018). doi:10.1038/nrm.2017.130

67. Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat. Protoc.* (2019). doi:10.1038/s41596-018-0103-9

68. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* (2018). doi:10.1093/nar/gkx1132

69. Posey, J. E. *et al.* Resolution of Disease Phenotypes Resulting from Multilocus Genomic Variation. *N. Engl. J. Med.* (2016). doi:10.1056/nejmoa1516767

70. Carvalho-Silva, D. *et al.* Open Targets Platform: New developments and updates two years on. *Nucleic Acids Res.* (2019). doi:10.1093/nar/gky1133

71. Tamborero, D. *et al.* Cancer Genome Interpreter Annotates The Biological And Clinical Relevance Of Tumor Alterations. *bioRxiv* (2017). doi:10.1101/140475

72. The Cancer Genome Atlas Research Network. The Cancer Genome Atlas Pan-Cancer Analysis project. *Nat. Genet.* **45**, 1113–8 (2013).

73. Thomas, S. J., Snowden, J. A., Zeidler, M. P. & Danson, S. J. The role of JAK/STAT signalling in the pathogenesis, prognosis and treatment of solid tumours. *Br. J. Cancer* (2015). doi:10.1038/bjc.2015.233

74. Kazi, M., Trivedi, T., Kobawala, T. & Ghosh, N. The Potential of Wnt Signaling Pathway in Cancer: A Focus on Breast Cancer. *Cancer Transl. Med.* (2016). doi:10.4103/2395-3977.181437

75. Murphy, M. E. *et al.* A functionally significant SNP in TP53 and breast cancer risk in African-American women. *npj Breast Cancer* **3**, 5 (2017).

76. Xie, B. *et al.* Benzyl Isothiocyanate potentiates p53 signaling and antitumor effects against breast cancer through activation of p53-LKB1 and p73-LKB1 axes. *Sci. Rep.* **7**, (2017).

77. Paz-y-Miño, C. *et al.* Incidence of the L858R and G719S mutations of the epidermal growth factor receptor oncogene in an Ecuadorian population with lung cancer. *Cancer Genet. Cytogenet.* **196**, (2010).

78. Ali, R. & Wendt, M. K. The paradoxical functions of EGFR during breast cancer progression. *Signal Transduct. Target. Ther.* **2**, 16042 (2017).

79. Wu, L. & Yang, X. Targeting the Hippo Pathway for Breast Cancer Therapy. *Cancers (Basel).* (2018). doi:10.3390/cancers10110422

80. Uzair, I. D., Conte Grand, J., Flamini, M. I. & Sanchez, A. M. Molecular Actions of Thyroid Hormone on Breast Cancer Cell Migration and Invasion via

Cortactin/N-WASP. *Front. Endocrinol. (Lausanne).* (2019). doi:10.3389/fendo.2019.00139

81. Ivanov, A. A. *et al.* The OncoPPi Portal: An integrative resource to explore and prioritize protein-protein interactions for cancer target discovery. *Bioinformatics* (2018). doi:10.1093/bioinformatics/btx743
82. Tsherniak, A. *et al.* Defining a Cancer Dependency Map. *Cell* **170**, 564-576.e16 (2017).
83. Reimand, J. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat. Protoc.* (2019). doi:10.1038/s41596-018-0103-9
84. Adams, S. *et al.* Current Landscape of Immunotherapy in Breast Cancer. *JAMA Oncol.* 1–10 (2019). doi:10.1001/jamaoncol.2018.7147
85. Miller, S. M., Goulet, D. R. & Johnson, G. L. Targeting the Breast Cancer Kinome. *J. Cell. Physiol.* (2017). doi:10.1002/jcp.25427
86. Caldon, C. E., Daly, R. J., Sutherland, R. L. & Musgrove, E. A. Cell cycle control in breast cancer cells. *Journal of Cellular Biochemistry* (2006). doi:10.1002/jcb.20690
87. Majidinia, M. & Yousefi, B. DNA repair and damage pathways in breast cancer development and therapy. *DNA Repair* (2017). doi:10.1016/j.dnarep.2017.03.009
88. Pereira, B., Billaud, M. & Almeida, R. RNA-Binding Proteins in Cancer: Old Players and New Actors. *Trends in Cancer* (2017). doi:10.1016/j.trecan.2017.05.003
89. Wurth, L. *et al.* UNR/CSDE1 Drives a Post-transcriptional Program to Promote Melanoma Invasion and Metastasis. *Cancer Cell* **30**, 694–707 (2016).
90. Kechavarzi, B. & Janga, S. C. Dissecting the expression landscape of RNA-binding proteins in human cancers. *Genome Biol.* (2014). doi:10.1186/gb-2014-15-1-r14
91. Wang, J., Liu, Q. & Shyr, Y. Dysregulated transcription across diverse cancer types reveals the importance of RNA-binding protein in carcinogenesis. *BMC Genomics* (2015). doi:10.1186/1471-2164-16-S7-S5
92. Sebestyén, E. *et al.* Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. *Genome Res.* (2016). doi:10.1101/gr.199935.115
93. Harris, E. E. R. Precision Medicine for Breast Cancer: The Paths to Truly Individualized Diagnosis and Treatment. *Int. J. Breast Cancer* (2018). doi:10.1155/2018/4809183
94. Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* (2010). doi:10.1038/nature08822
95. Mermel, C. H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* (2011). doi:10.1186/gb-2011-12-4-r41
96. Rocco, P. *et al.* OncoScore: A novel, Internet-based tool to assess the oncogenic potential of genes. *Sci. Rep.* (2017). doi:10.1038/srep46290
97. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–7 (2012).
98. Ding, L. *et al.* Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics. *Cell* (2018). doi:10.1016/j.cell.2018.03.033

Acknowledgments

This work was supported by Universidad UTE (Quito, Ecuador), Universidad de Las Américas (Quito, Ecuador), University of Coruna (Coruña, Spain), University of the Basque Country (Bilbao, Spain), and McGill University (Montreal, Canada). Additionally, this work was supported by "Collaborative Project in Genomic Data Integration (CICLOGEN) " PI17/01826 funded by the Carlos III Health Institute from the Spanish National plan for Scientific and Technical Research and Innovation 2013-2016 and the European Regional Development Funds (FEDER).

Author Contributions

ALC and ET conceived the subject and the conceptualization of the study. ALC wrote the manuscript. ET, SJB, CRM, HGD and CPyM supervised the project. ALC and CPyM did founding acquisition. ALC, SG and ACA did data curation and supplementary data. ET, SG, ACA, SJB, CRM, HGD, AP, YPC and CPyM gave conceptual advice and valuable scientific input. Finally, all authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Data availability statement

All data generated or analysed during this study are included in this published article (and its Supplementary Information files).

994 **Figure legends**

995

996 **Figure 1. Genetic alterations of the breast cancer cohort according to PCA.** (A) Frequency
997 of genetic alterations per gene set (non-cancer genes, BC driver genes according to the Network
998 of Cancer Genes, Consensus Strategy, BC genes according to PCA, BC biomarkers according to
999 the PharmGKB and CGI). (B) Percentage of genetic alterations per type. (C) Ratio of genetic
1000 alterations per intrinsic molecular subtype. (D) Ratio of genetic alterations per tumor stage. (E)
1001 Percentage of genetic alterations per type and per molecular subtype. (F) Percentage of genetic
1002 alterations per type and per tumor stage.

1003

1004 **Figure 2. Ranking of genes with the highest number of genetic alterations per molecular**
1005 **subtype and tumor stage.** (A) Frequency of genetic alterations (punctual mutations, copy
1006 number variants and mRNA expression) per molecular subtype. (B) Frequency of genetic
1007 alterations per tumor stage. (C) Frequency of punctual mutations per molecular subtype. (D)
1008 Frequency of punctual mutations per tumor stage. (E) Frequency of CNV amplifications per
1009 molecular subtype. (F) Frequency of CNV amplifications per tumor stage. (G) Frequency of
1010 CNV deep deletions per molecular subtype. (H) Frequency of CNV deep deletions per tumor
1011 stage. (I) Frequency of mRNA upregulation per molecular subtype. (J) Frequency of mRNA
1012 upregulation per tumor stage. (K) Frequency of mRNA downregulation per molecular subtype.
1013 (L) Frequency of mRNA downregulation per tumor stage.

1014

1015 **Figure 3. OncoPrint of genetic alterations according to the Pan-Cancer Atlas.** (A)
1016 OncoPrint of genes with more genetic alterations than the average (>86) per molecular subtype.
1017 (B) Circos plot between molecular subtypes and the highest number of genetic alterations
1018 (fusion genes, mRNA downregulation plus CNV deep deletion, mRNA upregulation plus CNV
1019 amplification and driver mutations). (C) Circos plot between tumor stages and the highest
1020 number of genetic alterations.

1021

Figure 4. Pathway enrichment analysis per molecular subtype and tumor stage. (A) Circos plot between molecular subtypes and the most altered genetic pathways. (B) Violin plots showing the frequency of the most altered signaling pathways per molecular subtype. (C) Circos plot between tumor stages and the most altered genetic pathways. (D) Violin plots showing the frequency of the most altered signaling pathways per tumor stage.

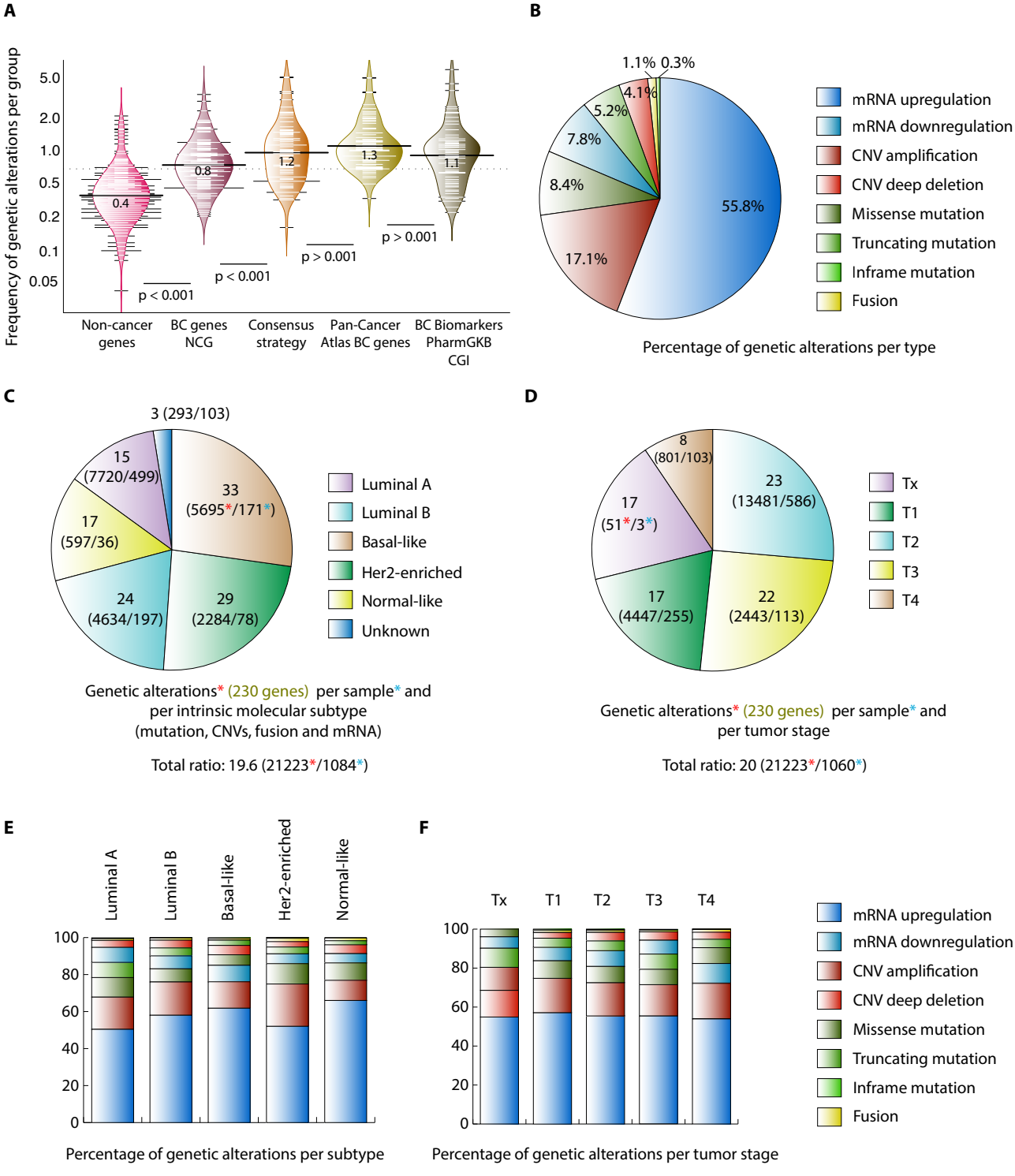
Figure 5. Breast cancer integrated network. (A) Network composed of BC driver genes and genes of our study (PCA gene set, consensus strategy gene set and PharmGKB gene set. (B) Significant correlation ($p < 0.05$) of degree centrality and consensus score between the OncoPPi BC network and or BC integrated network.

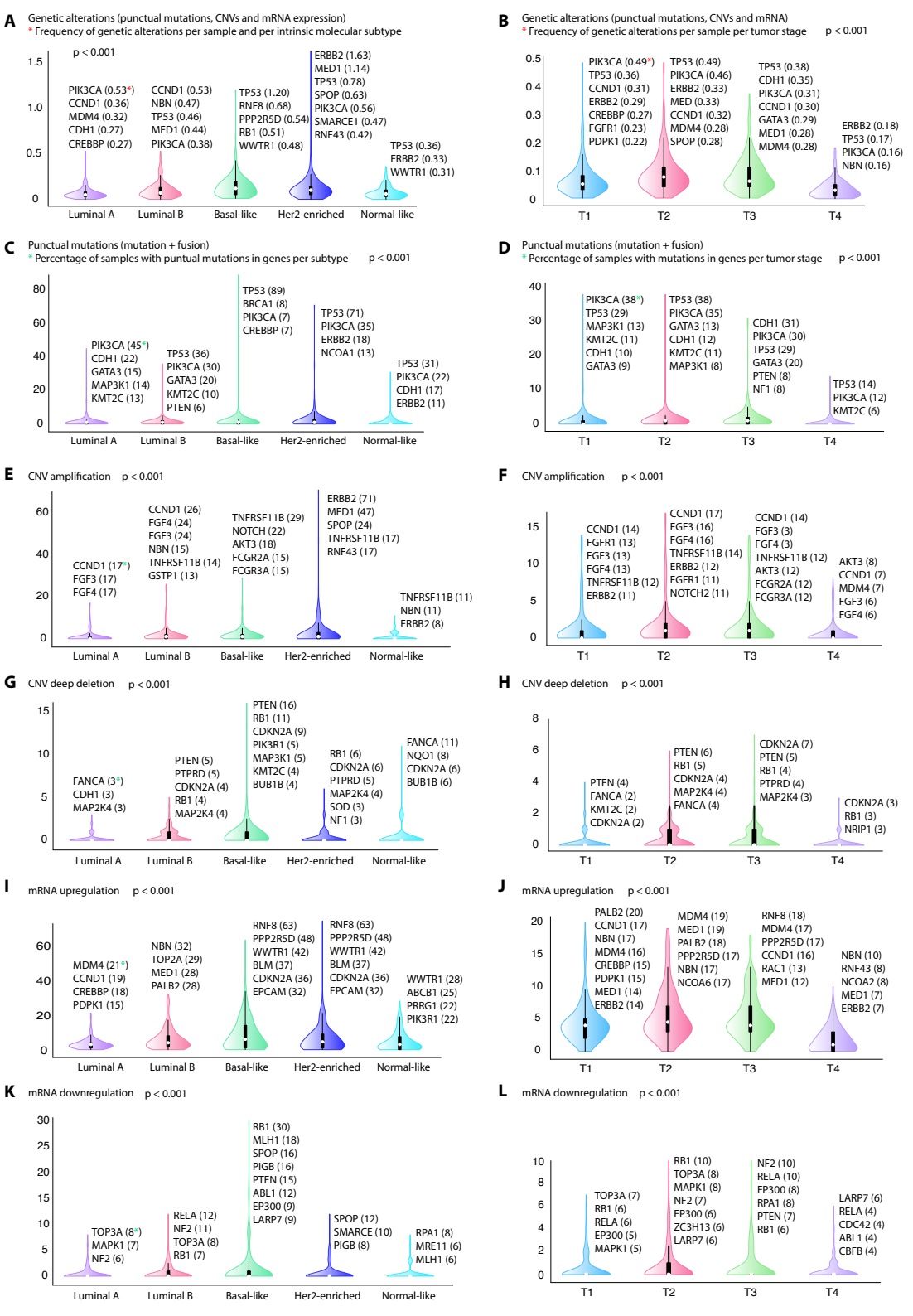
Figure 6. Analysis of protein expression. (A) Ranking of genes with the highest number of protein alterations (high and low expression with Z-score ≥ 2) according to The Cancer Genome Atlas. (B) Comparison of protein expression levels between BC tissue and normal tissue according to The Human Protein Atlas. (C) Overall survival of genes with prognosis unfavorable ($p < 0.001$) in BC according to The Human Protein Atlas.

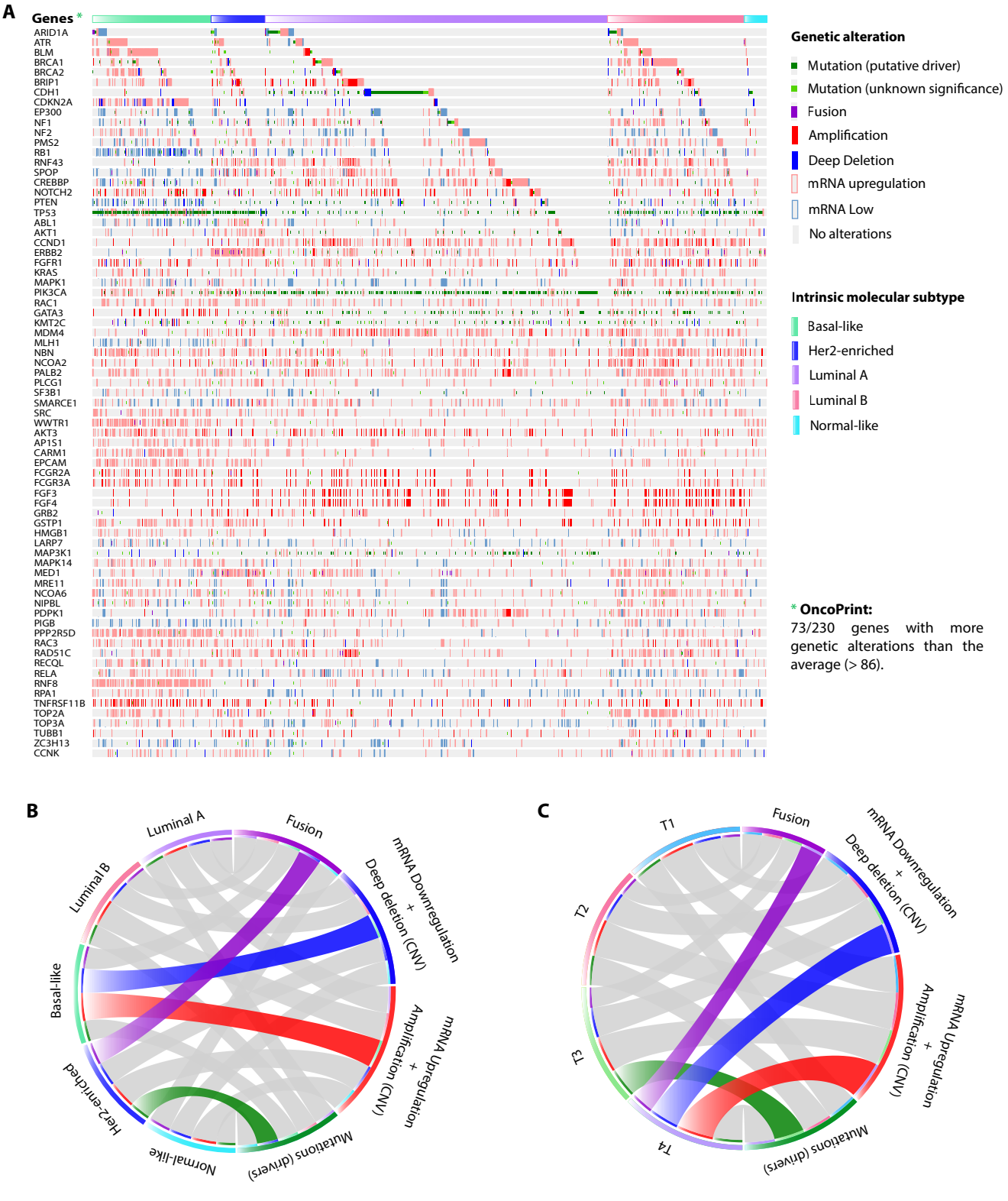
Figure 7. Analysis of dependencies in BC cell lines. (A) Dependency score of BC gene sets using RNAi DIMETER2 and CRISPR-Cas9 CERES algorithms in BC cell lines. (B) Dependency score of BC gene sets per molecular subtypes. (C) Violin plots of dependencies per molecular subtypes. All significant dependencies < -1 are in black. (D) Venn diagram of genes with at least one dependency < -1 in cell lines belonging to each molecular subtype. (E) Venn diagram of strongly selective and common essential genes in all cancer cell lines.

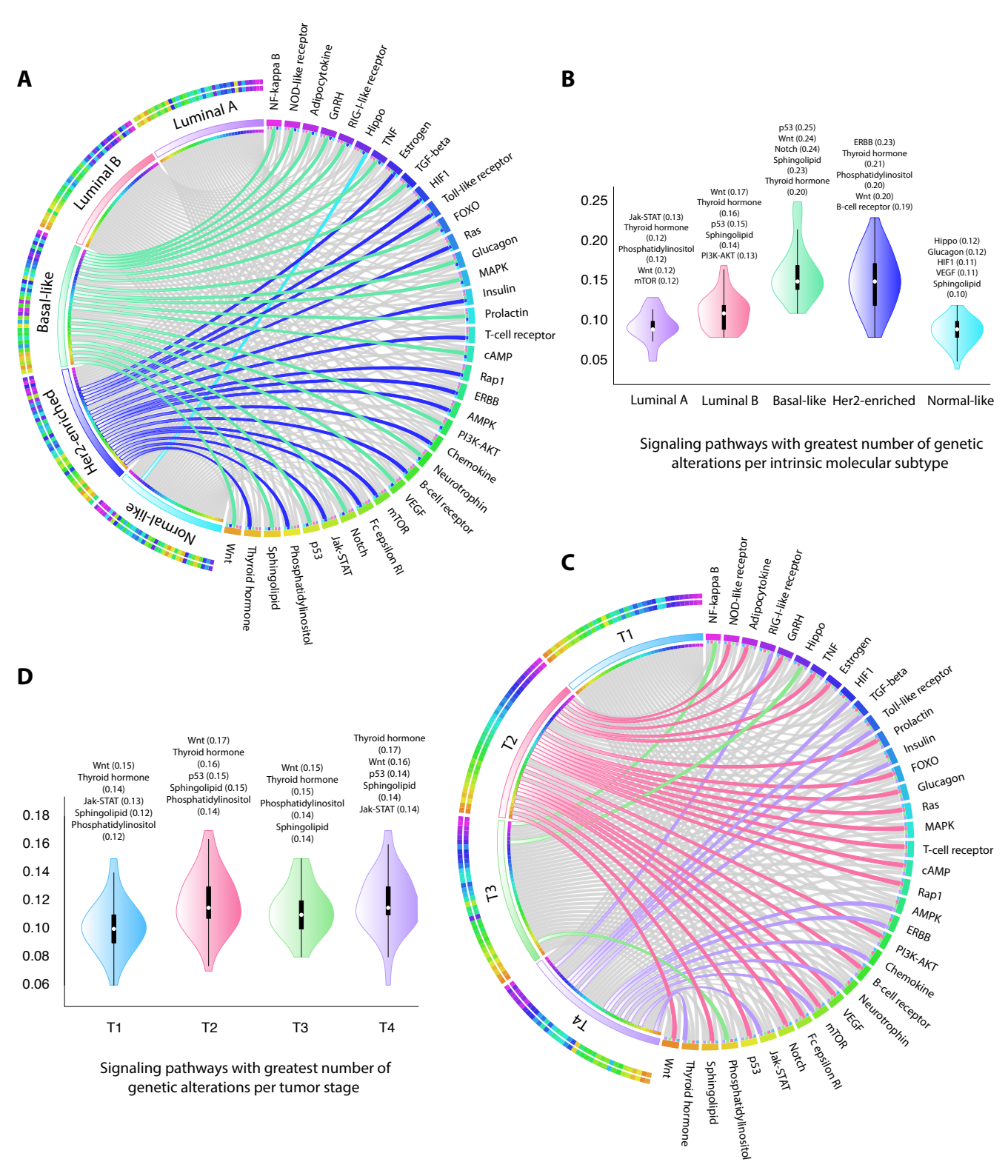
Figure 8. The OncoOmics essential genes of breast cancer. (A) Venn diagram of the most relevant genes per genomics approach (PCA genetic alterations, networking, protein expression and DepMap). (B) Percentage of oncogenes, tumor suppressor genes, tier 1 genes, BC driver genes and driver genes in other cancer types. (C) Venn diagram of the most relevant genes

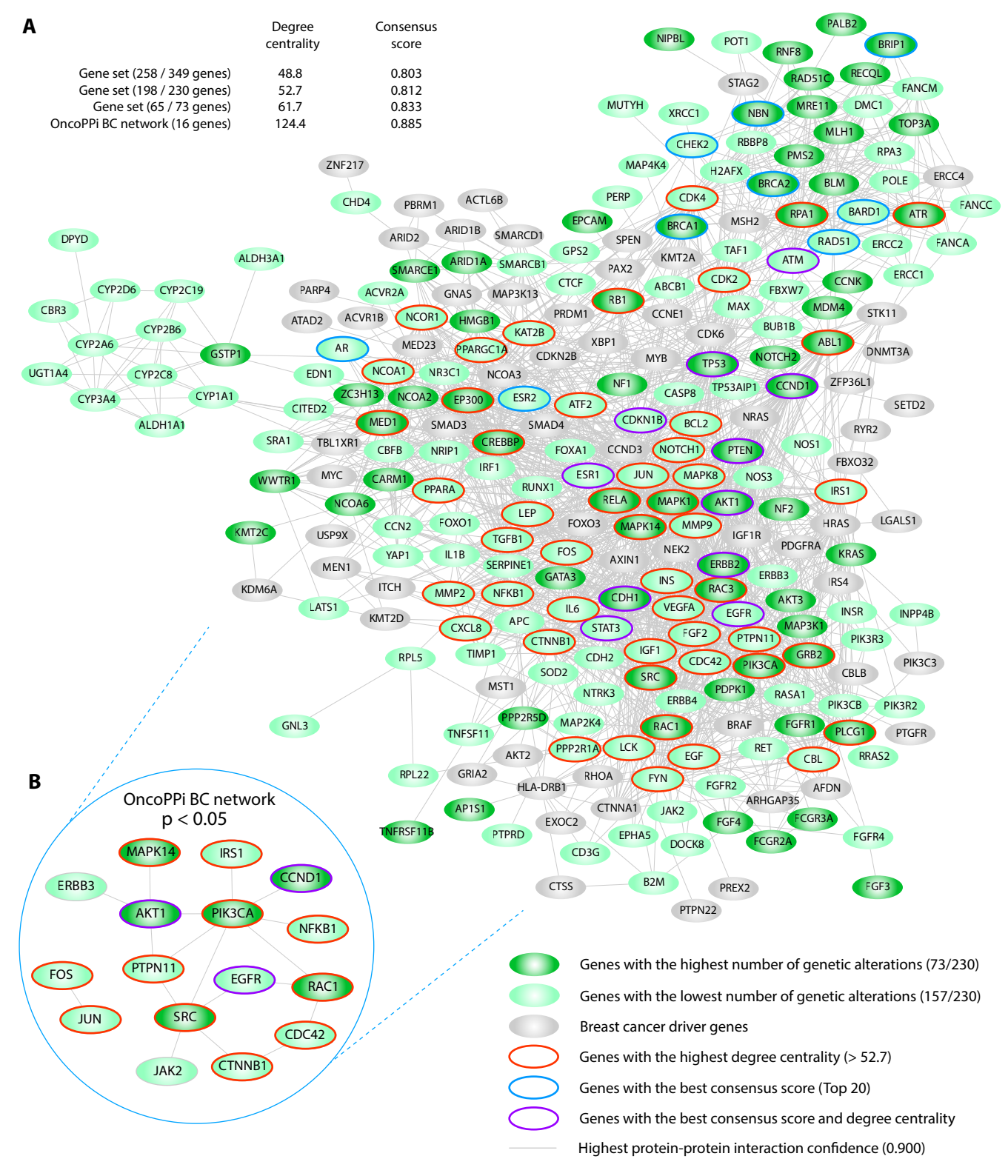
1050 related with cancer immunotherapy, kinome, cell cycle, DNA repair and RNA-binding proteins.
1051 (D) Circos plot of the hallmarks of cancer genes. (E) Most significant g:Profiler features of the
1052 most relevant genes according to the gene ontology biological processes, Reactome pathways,
1053 wikipathways and the human phenotype ontology.
1054
1055 **Figure 9. Pathway enrichment analysis of the most relevant genes using g:Profiler and**
1056 **EnrichmentMap.** Most significant Reactome pathways related to immune system, tyrosine
1057 kinases, cell cycle, DNA repair and genetic transcription.
1058
1059 **Figure 10. A panoramic view of clinical trial features in breast cancer.**
1060
1061
1062
1063

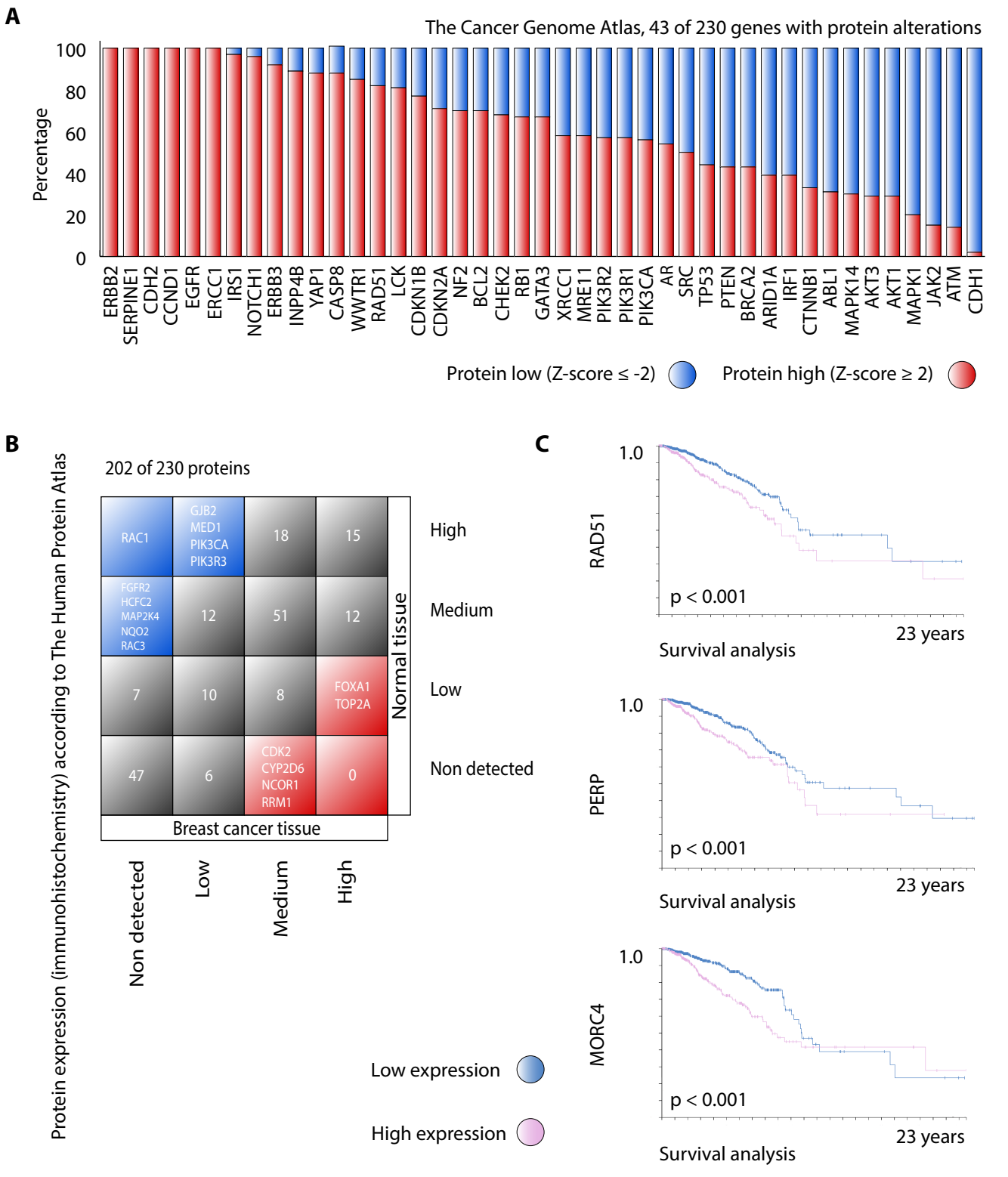

A
Frequency of genetic alterations per group
5.0
2.0
1.0
0.5
0.2
0.1
0.05
0.4
0.8
1.2
1.3
1.1
p < 0.001
p < 0.001
p > 0.001
p > 0.001
Non-cancer genes
BC genes NCG
Consensus strategy
Pan-Cancer Atlas BC genes
BC Biomarkers PharmGKB CGI
B
1.1%
0.3%
4.1%
5.2%
7.8%
8.4%
55.8%
17.1%
mRNA upregulation
mRNA downregulation
CNV amplification
CNV deep deletion
Missense mutation
Truncating mutation
Inframe mutation
Fusion
Percentage of genetic alterations per type
C
3 (293/103)
15 (7720/499)
33 (5695*/171*)
17 (597/36)
24 (4634/197)
29 (2284/78)
Luminal A
Luminal B
Basal-like
Her2-enriched
Normal-like
Unknown
Genetic alterations* (230 genes) per sample* and per intrinsic molecular subtype (mutation, CNVs, fusion and mRNA)
Total ratio: 19.6 (21223*/1084*)
D
8 (801/103)
23 (13481/586)
17 (51*/3*)
17 (4447/255)
22 (2443/113)
Tx
T1
T2
T3
T4
Genetic alterations* (230 genes) per sample* and per tumor stage
Total ratio: 20 (21223*/1060*)
E
Luminal A
Luminal B
Basal-like
Her2-enriched
Normal-like
100
80
60
40
20
0
Percentage of genetic alterations per subtype
F
Tx
T1
T2
T3
T4
100
80
60
40
20
0
Percentage of genetic alterations per tumor stage
mRNA upregulation
mRNA downregulation
CNV amplification
CNV deep deletion
Missense mutation
Truncating mutation
Inframe mutation
Fusion

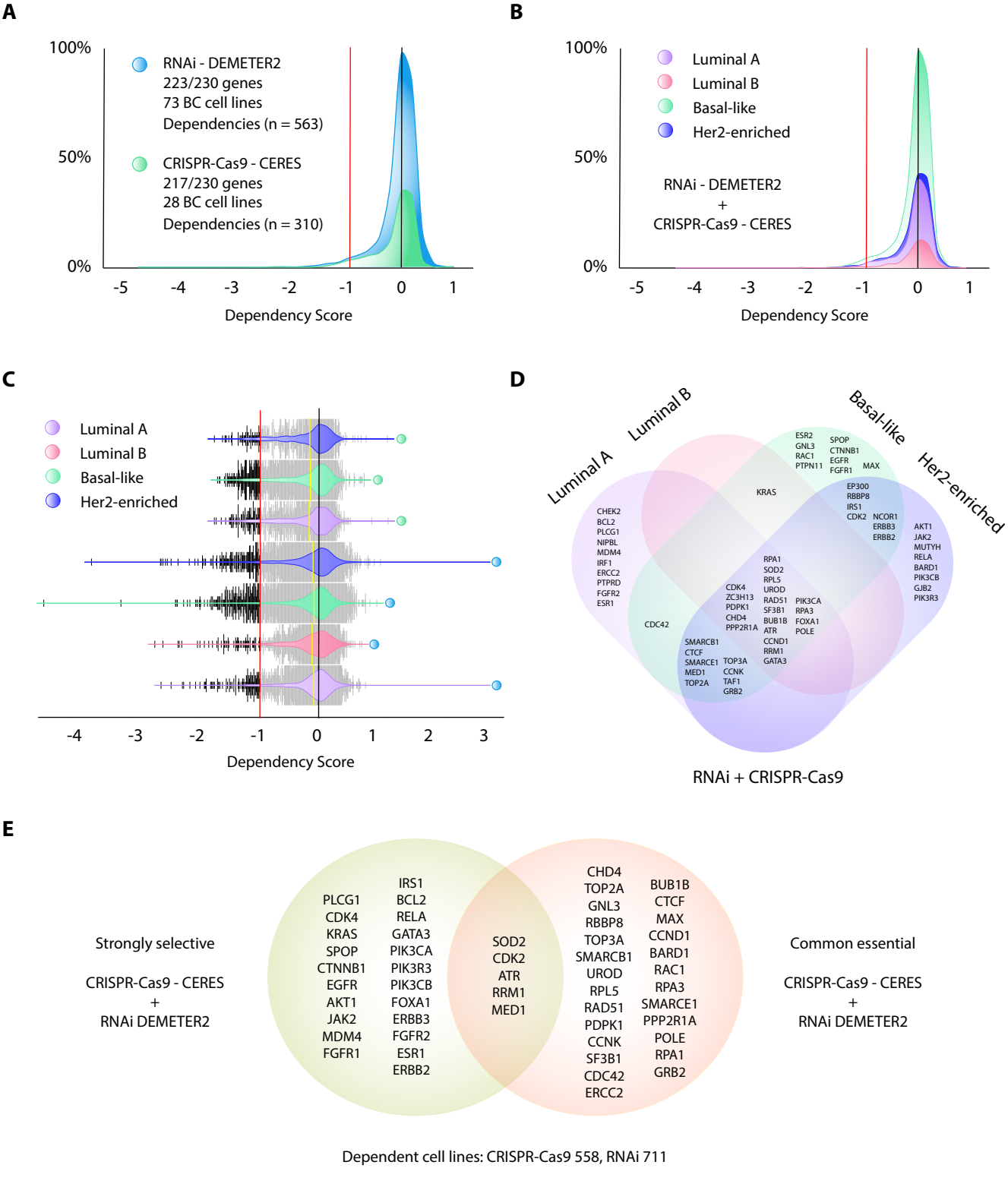

A
Genetic alterations (punctual mutations, CNVs and mRNA expression)
* Frequency of genetic alterations per sample and per intrinsic molecular subtype
p < 0.001
PIK3CA (0.53*)
CCND1 (0.36)
MDM4 (0.32)
CDH1 (0.27)
CREBBP (0.27)
CCND1 (0.53)
NBN (0.47)
TP53 (0.46)
MED1 (0.44)
PIK3CA (0.38)
TP53 (1.20)
RNF8 (0.68)
PPP2R5D (0.54)
RB1 (0.51)
WWTR1 (0.48)
ERBB2 (1.63)
MED1 (1.14)
TP53 (0.78)
SPOP (0.63)
PIK3CA (0.56)
SMARCE1 (0.47)
RNF43 (0.42)
TP53 (0.36)
ERBB2 (0.33)
WWTR1 (0.31)
Luminal A
Luminal B
Basal-like
Her2-enriched
Normal-like
B
Genetic alterations (punctual mutations, CNVs and mRNA)
* Frequency of genetic alterations per sample per tumor stage
p < 0.001
PIK3CA (0.49*)
TP53 (0.36)
CCND1 (0.31)
ERBB2 (0.29)
CREBBP (0.27)
FGFR1 (0.23)
PDPK1 (0.22)
TP53 (0.49)
PIK3CA (0.46)
ERBB2 (0.33)
MED (0.33)
CCND1 (0.32)
MDM4 (0.28)
SPOP (0.28)
TP53 (0.38)
CDH1 (0.35)
PIK3CA (0.31)
CCND1 (0.30)
GATA3 (0.29)
MED1 (0.28)
MDM4 (0.28)
ERBB2 (0.18)
TP53 (0.17)
PIK3CA (0.16)
NBN (0.16)
T1
T2
T3
T4
C
Punctual mutations (mutation + fusion)
* Percentage of samples with puntual mutations in genes per subtype
p < 0.001
PIK3CA (45*)
CDH1 (22)
GATA3 (15)
MAP3K1 (14)
KMT2C (13)
TP53 (36)
PIK3CA (30)
GATA3 (20)
KMT2C (10)
PTEN (6)
TP53 (89)
BRCA1 (8)
PIK3CA (7)
CREBBP (7)
TP53 (71)
PIK3CA (35)
ERBB2 (18)
NCOA1 (13)
TP53 (31)
PIK3CA (22)
CDH1 (17)
ERBB2 (11)
Luminal A
Luminal B
Basal-like
Her2-enriched
Normal-like
D
Punctual mutations (mutation + fusion)
* Percentage of samples with mutations in genes per tumor stage
p < 0.001
PIK3CA (38*)
TP53 (29)
MAP3K1 (13)
KMT2C (11)
CDH1 (10)
GATA3 (9)
TP53 (38)
PIK3CA (35)
GATA3 (13)
CDH1 (12)
KMT2C (11)
MAP3K1 (8)
CDH1 (31)
PIK3CA (30)
TP53 (29)
GATA3 (20)
PTEN (8)
NF1 (8)
TP53 (14)
PIK3CA (12)
KMT2C (6)
T1
T2
T3
T4
E
CNV amplification
p < 0.001
CCND1 (17*)
FGF3 (17)
FGF4 (17)
CCND1 (26)
FGF4 (24)
FGF3 (24)
NBN (15)
TNFRSF11B (14)
GSTP1 (13)
TNFRSF11B (29)
NOTCH (22)
AKT3 (18)
FCGR2A (15)
FCGR3A (15)
ERBB2 (71)
MED1 (47)
SPOP (24)
TNFRSF11B (17)
RNF43 (17)
TNFRSF11B (11)
NBN (11)
ERBB2 (8)
Luminal A
Luminal B
Basal-like
Her2-enriched
Normal-like
F
CNV amplification
p < 0.001
CCND1 (14)
FGFR1 (13)
FGF3 (13)
FGF4 (13)
TNFRSF11B (12)
ERBB2 (11)
CCND1 (17)
FGF3 (16)
FGF4 (16)
TNFRSF11B (14)
ERBB2 (12)
FGFR1 (11)
NOTCH2 (11)
CCND1 (14)
FGF3 (3)
FGF4 (3)
TNFRSF11B (12)
AKT3 (12)
FCGR2A (12)
FCGR3A (12)
AKT3 (8)
CCND1 (7)
MDM4 (7)
FGF3 (6)
FGF4 (6)
T1
T2
T3
T4
G
CNV deep deletion
p < 0.001
FANCA (3*)
CDH1 (3)
MAP2K4 (3)
PTEN (5)
PTPRD (5)
CDKN2A (4)
RB1 (4)
MAP2K4 (4)
PTEN (16)
RB1 (11)
CDKN2A (9)
PIK3R1 (5)
MAP3K1 (5)
KMT2C (4)
BUB1B (4)
RB1 (6)
CDKN2A (6)
PTPRD (5)
MAP2K4 (4)
SOD (3)
NF1 (3)
FANCA (11)
NQO1 (8)
CDKN2A (6)
BUB1B (6)
Luminal A
Luminal B
Basal-like
Her2-enriched
Normal-like
H
CNV deep deletion
p < 0.001
PTEN (4)
FANCA (2)
KMT2C (2)
CDKN2A (2)
PTEN (6)
RB1 (5)
CDKN2A (4)
MAP2K4 (4)
FANCA (4)
CDKN2A (7)
PTEN (5)
RB1 (4)
PTPRD (4)
MAP2K4 (3)
CDKN2A (3)
RB1 (3)
NRIP1 (3)
T1
T2
T3
T4
I
mRNA upregulation
p < 0.001
MDM4 (21*)
CCND1 (19)
CREBBP (18)
PDPK1 (15)
NBN (32)
TOP2A (29)
MED1 (28)
PALB2 (28)
RNF8 (63)
PPP2R5D (48)
WWTR1 (42)
BLM (37)
CDKN2A (36)
EPCAM (32)
RNF8 (63)
PPP2R5D (48)
WWTR1 (42)
BLM (37)
CDKN2A (36)
EPCAM (32)
WWTR1 (28)
ABCB1 (25)
PRRG1 (22)
PIK3R1 (22)
Luminal A
Luminal B
Basal-like
Her2-enriched
Normal-like
J
mRNA upregulation
p < 0.001
PALB2 (20)
CCND1 (17)
NBN (17)
MDM4 (16)
CREBBP (15)
PDPK1 (15)
MED1 (14)
ERBB2 (14)
MDM4 (19)
MED1 (19)
PALB2 (18)
PPP2R5D (17)
NBN (17)
NCOA6 (17)
RNF8 (18)
MDM4 (17)
PPP2R5D (17)
CCND1 (16)
RAC1 (13)
MED1 (12)
NBN (10)
RNF43 (8)
NCOA2 (8)
MED1 (7)
ERBB2 (7)
T1
T2
T3
T4
K
mRNA downregulation
p < 0.001
TOP3A (8*)
MAPK1 (7)
NF2 (6)
RELA (12)
NF2 (11)
TOP3A (8)
RB1 (7)
RB1 (30)
MLH1 (18)
SPOP (16)
PIGB (16)
PTEN (15)
ABL1 (12)
EP300 (9)
LARP7 (9)
SPOP (12)
SMARCE (10)
PIGB (8)
RPA1 (8)
MRE11 (6)
MLH1 (6)
Luminal A
Luminal B
Basal-like
Her2-enriched
Normal-like
L
mRNA downregulation
p < 0.001
TOP3A (7)
RB1 (6)
RELA (6)
EP300 (5)
MAPK1 (5)
RB1 (10)
TOP3A (8)
MAPK1 (8)
NF2 (7)
EP300 (6)
ZC3H13 (6)
LARP7 (6)
NF2 (10)
RELA (10)
EP300 (8)
RPA1 (8)
PTEN (7)
RB1 (6)
LARP7 (6)
RELA (4)
CDC42 (4)
ABL1 (4)
CBFB (4)
T1
T2
T3
T4

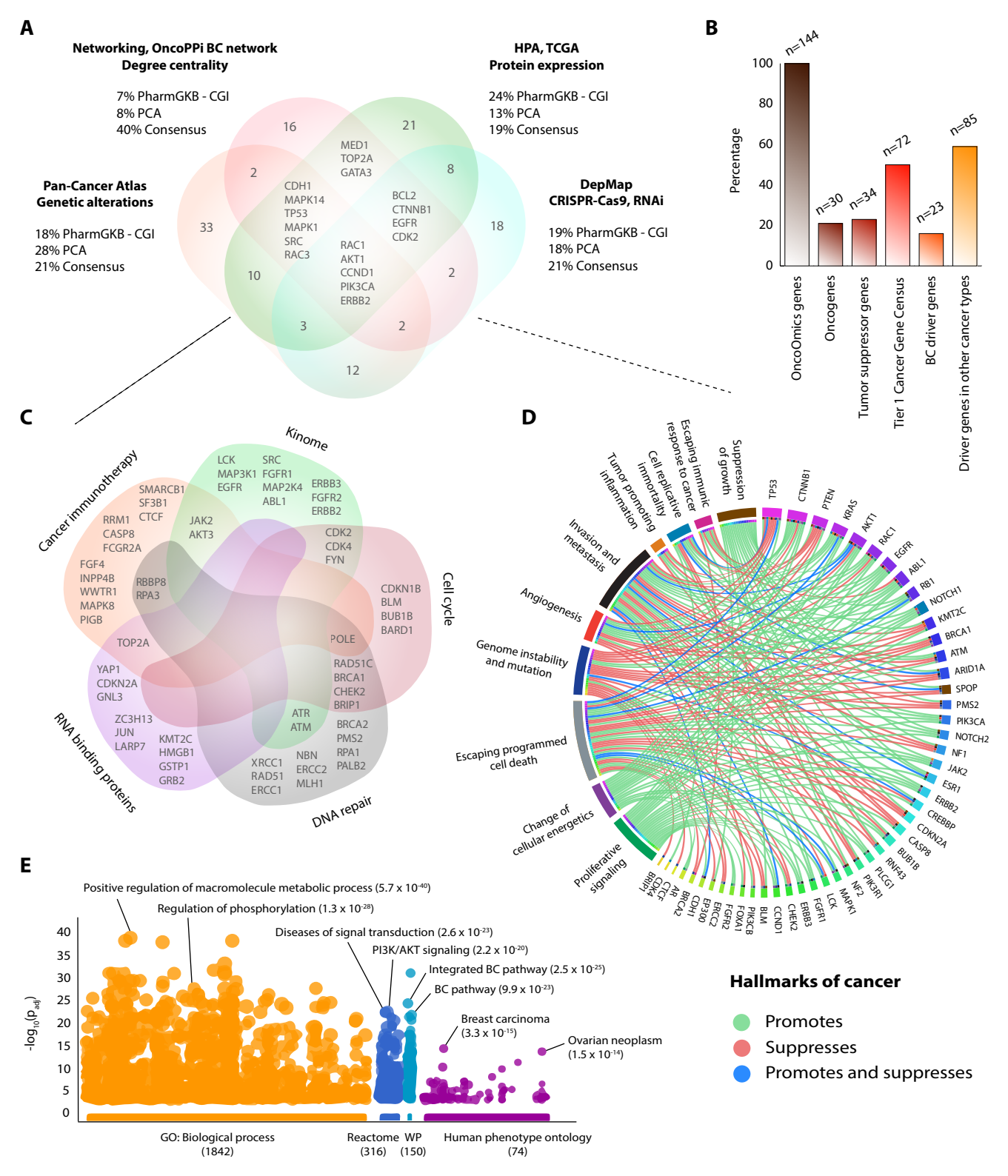

A
Genes *
ARID1A
ATR
BLM
BRCA1
BRCA2
BRIP1
CDH1
CDKN2A
EP300
NF1
NF2
PMS2
RB1
RNF43
SPOP
CREBBP
NOTCH2
PTEN
TP53
ABL1
AKT1
CCND1
ERBB2
FGFR1
KRAS
MAPK1
PIK3CA
RAC1
GATA3
KMT2C
MDM4
MLH1
NBN
NCOA2
PALB2
PLCG1
SF3B1
SMARCE1
SRC
WWTR1
AKT3
AP1S1
CARM1
EPCAM
FCGR2A
FCGR3A
FGF3
FGF4
GRB2
GSTP1
HMGB1
LARP7
MAP3K1
MAPK14
MED1
MRE11
NCOA6
NIPBL
PDPK1
PIGB
PPP2R5D
RAC3
RAD51C
RECQL
RELA
RNF8
RPA1
TNFRSF11B
TOP2A
TOP3A
TUBB1
ZC3H13
CCNK
Genetic alteration
Mutation (putative driver)
Mutation (unknown significance)
Fusion
Amplification
Deep Deletion
mRNA upregulation
mRNA Low
No alterations
Intrinsic molecular subtype
Basal-like
Her2-enriched
Luminal A
Luminal B
Normal-like
* OncoPrint:
73/230 genes with more genetic alterations than the average (> 86).
B
Luminal A
Fusion
mRNA Downregulation + Deep deletion (CNV)
mRNA Upregulation + Amplification (CNV)
Mutations (drivers)
Normal-like
Her2-enriched
Basal-like
Luminal B
C
T1
Fusion
mRNA Downregulation + Deep deletion (CNV)
mRNA Upregulation + Amplification (CNV)
Mutations (drivers)
T4
T3
T2

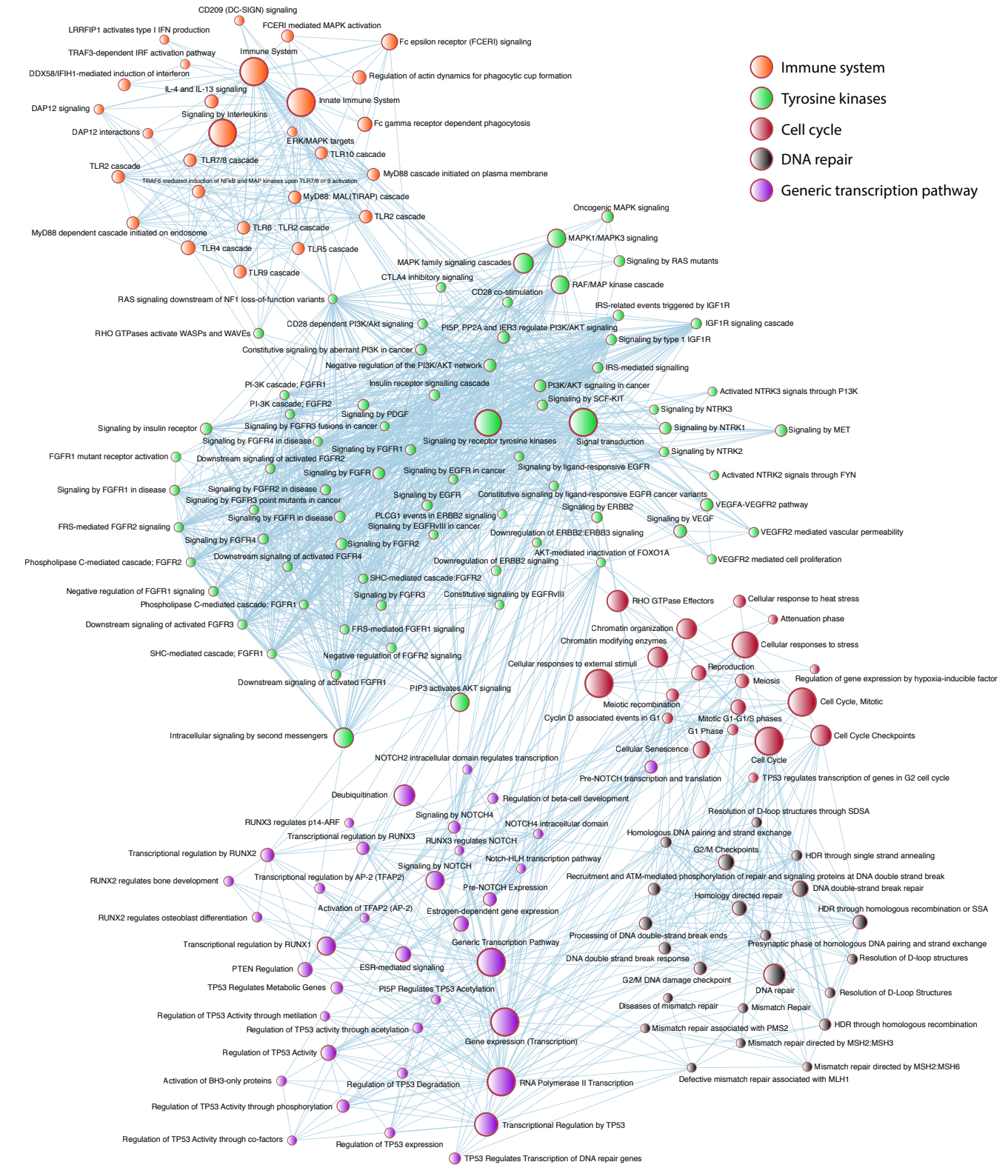

A
Luminal A
Luminal B
Basal-like
Her2-enriched
Normal-like
NF-kappa B
NOD-like receptor
Adipocytokine
GnRH
RIG-I-like receptor
Hippo
TNF
Estrogen
TGF-beta
HIF1
Toll-like receptor
FOXO
Ras
Glucagon
MAPK
Insulin
Prolactin
T-cell receptor
cAMP
Rap1
ERBB
AMPK
PI3K-AKT
Chemokine
Neurotrophin
B-cell receptor
VEGF
mTOR
Fc epsilon RI
Notch
Jak-STAT
p53
Phosphatidylinositol
Sphingolipid
Thyroid hormone
Wnt
B
0.25
0.20
0.15
0.10
0.05
Jak-STAT (0.13)
Thyroid hormone (0.12)
Phosphatidylinositol (0.12)
Wnt (0.12)
mTOR (0.12)
Wnt (0.17)
Thyroid hormone (0.16)
p53 (0.15)
Sphingolipid (0.14)
PI3K-AKT (0.13)
p53 (0.25)
Wnt (0.24)
Notch (0.24)
Sphingolipid (0.23)
Thyroid hormone (0.20)
ERBB (0.23)
Thyroid hormone (0.21)
Phosphatidylinositol (0.20)
Wnt (0.20)
B-cell receptor (0.19)
Hippo (0.12)
Glucagon (0.12)
HIF1 (0.11)
VEGF (0.11)
Sphingolipid (0.10)
Luminal A
Luminal B
Basal-like
Her2-enriched
Normal-like
Signaling pathways with greatest number of genetic alterations per intrinsic molecular subtype
C
T1
T2
T3
T4
NF-kappa B
NOD-like receptor
Adipocytokine
RIG-I-like receptor
GnRH
Hippo
TNF
Estrogen
HIF1
TGF-beta
Toll-like receptor
Prolactin
Insulin
FOXO
Glucagon
Ras
MAPK
T-cell receptor
cAMP
Rap1
AMPK
ERBB
PI3K-AKT
Chemokine
B-cell receptor
Neurotrophin
VEGF
mTOR
Fc epsilon RI
Notch
Jak-STAT
p53
Phosphatidylinositol
Sphingolipid
Thyroid hormone
Wnt
D
0.18
0.16
0.14
0.12
0.10
0.08
0.06
Wnt (0.15)
Thyroid hormone (0.14)
Jak-STAT (0.13)
Sphingolipid (0.12)
Phosphatidylinositol (0.12)
Wnt (0.17)
Thyroid hormone (0.16)
p53 (0.15)
Sphingolipid (0.15)
Phosphatidylinositol (0.14)
Wnt (0.15)
Thyroid hormone (0.15)
Phosphatidylinositol (0.14)
Sphingolipid (0.14)
Thyroid hormone (0.17)
Wnt (0.16)
p53 (0.14)
Sphingolipid (0.14)
Jak-STAT (0.14)
T1
T2
T3
T4
Signaling pathways with greatest number of genetic alterations per tumor stage

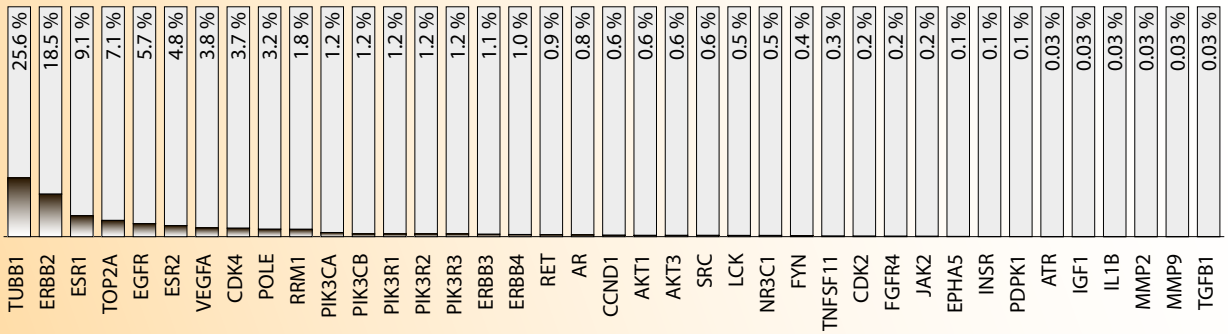

A
Degree centrality
Consensus score
Gene set (258 / 349 genes) 48.8 0.803
Gene set (198 / 230 genes) 52.7 0.812
Gene set (65 / 73 genes) 61.7 0.833
OncoPPi BC network (16 genes) 124.4 0.885
B
OncoPPi BC network
p < 0.05
MAPK14
IRS1
CCND1
ERBB3
AKT1
PIK3CA
NFKB1
FOS
PTPN11
EGFR
RAC1
JUN
SRC
CDC42
JAK2
CTNNB1
Genes with the highest number of genetic alterations (73/230)
Genes with the lowest number of genetic alterations (157/230)
Breast cancer driver genes
Genes with the highest degree centrality (> 52.7)
Genes with the best consensus score (Top 20)
Genes with the best consensus score and degree centrality
Highest protein-protein interaction confidence (0.900)

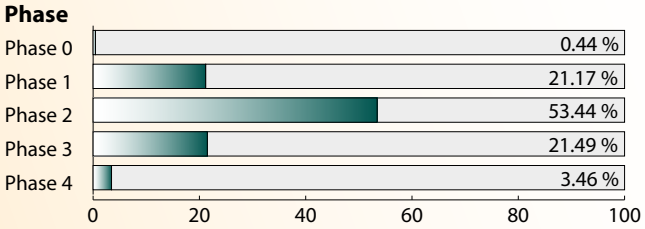

A
The Cancer Genome Atlas, 43 of 230 genes with protein alterations
Percentage
100
80
60
40
20
0
ERBB2
SERPINE1
CDH2
CCND1
EGFR
ERCC1
IRS1
NOTCH1
ERBB3
INPP4B
YAP1
CASP8
WWTR1
RAD51
LCK
CDKN1B
CDKN2A
NF2
BCL2
CHEK2
RB1
GATA3
XRCC1
MRE11
PIK3R2
PIK3R1
PIK3CA
AR
SRC
TP53
PTEN
BRCA2
ARID1A
IRF1
CTNNB1
ABL1
MAPK14
AKT3
AKT1
MAPK1
JAK2
ATM
CDH1
Protein low (Z-score ≤ -2)
Protein high (Z-score ≥ 2)
B
202 of 230 proteins
RAC1
GJB2
MED1
PIK3CA
PIK3R3
18
15
FGFR2
HCFC2
MAP2K4
NQO2
RAC3
12
51
12
7
10
8
FOXA1
TOP2A
47
6
CDK2
CYP2D6
NCOR1
RRM1
0
High
Medium
Low
Non detected
Normal tissue
Breast cancer tissue
Non detected
Low
Medium
High
Protein expression (immunohistochemistry) according to The Human Protein Atlas
Low expression
High expression
C
RAD51
1.0
p < 0.001
23 years
Survival analysis
PERP
1.0
p < 0.001
23 years
Survival analysis
MORC4
1.0
p < 0.001
23 years
Survival analysis


A
100%
50%
0%
RNAi - DEMETER2
223/230 genes
73 BC cell lines
Dependencies (n = 563)
CRISPR-Cas9 - CERES
217/230 genes
28 BC cell lines
Dependencies (n = 310)
-5 -4 -3 -2 -1 0 1
Dependency Score
B
100%
50%
0%
Luminal A
Luminal B
Basal-like
Her2-enriched
RNAi - DEMETER2
+
CRISPR-Cas9 - CERES
-5 -4 -3 -2 -1 0 1
Dependency Score
C
Luminal A
Luminal B
Basal-like
Her2-enriched
-4 -3 -2 -1 0 1 2 3
Dependency Score
D
Luminal A
Luminal B
Basal-like
Her2-enriched
CHEK2
BCL2
PLCG1
NIPBL
MDM4
IRF1
ERCC2
PTPRD
FGFR2
ESR1
CDC42
ESR2
GNL3
RAC1
PTPN11
SPOP
CTNNB1
EGFR
FGFR1
MAX
KRAS
EP300
RBBP8
IRS1
CDK2
NCOR1
ERBB3
ERBB2
AKT1
JAK2
MUTYH
RELA
BARD1
PIK3CB
GJB2
PIK3R3
CDK4
ZC3H13
PDPK1
CHD4
PPP2R1A
RPA1
SOD2
RPL5
UROD
RAD51
SF3B1
BUB1B
ATR
CCND1
RRM1
GATA3
PIK3CA
RPA3
FOXA1
POLE
SMARCB1
CTCF
SMARCE1
MED1
TOP2A
TOP3A
CCNK
TAF1
GRB2
RNAi + CRISPR-Cas9
E
Strongly selective
CRISPR-Cas9 - CERES
+
RNAi DEMETER2
PLCG1
CDK4
KRAS
SPOP
CTNNB1
EGFR
AKT1
JAK2
MDM4
FGFR1
IRS1
BCL2
RELA
GATA3
PIK3CA
PIK3R3
PIK3CB
FOXA1
ERBB3
FGFR2
ESR1
ERBB2
SOD2
CDK2
ATR
RRM1
MED1
CHD4
TOP2A
GNL3
RBBP8
TOP3A
SMARCB1
UROD
RPL5
RAD51
PDPK1
CCNK
SF3B1
CDC42
ERCC2
BUB1B
CTCF
MAX
CCND1
BARD1
RAC1
RPA3
SMARCE1
PPP2R1A
POLE
RPA1
GRB2
Common essential
CRISPR-Cas9 - CERES
+
RNAi DEMETER2
Dependent cell lines: CRISPR-Cas9 558, RNAi 711

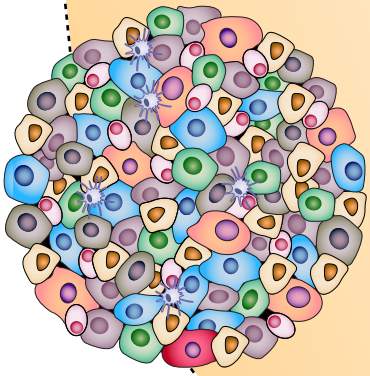

A
Networking, OncoPPi BC network
Degree centrality
7% PharmGKB - CGI
8% PCA
40% Consensus
HPA, TCGA
Protein expression
24% PharmGKB - CGI
13% PCA
19% Consensus
Pan-Cancer Atlas
Genetic alterations
18% PharmGKB - CGI
28% PCA
21% Consensus
DepMap
CRISPR-Cas9, RNAi
19% PharmGKB - CGI
18% PCA
21% Consensus
16
21
2
8
33
18
10
2
3
2
12
MED1
TOP2A
GATA3
CDH1
MAPK14
TP53
MAPK1
SRC
RAC3
BCL2
CTNNB1
EGFR
CDK2
RAC1
AKT1
CCND1
PIK3CA
ERBB2
B
Percentage
n=144
n=30
n=34
n=72
n=23
n=85
OncoOmics genes
Oncogenes
Tumor suppressor genes
Tier 1 Cancer Gene Census
BC driver genes
Driver genes in other cancer types
C
Kinome
Cancer immunotherapy
Cell cycle
RNA binding proteins
DNA repair
LCK
MAP3K1
EGFR
SRC
FGFR1
MAP2K4
ABL1
ERBB3
FGFR2
ERBB2
SMARCB1
SF3B1
CTCF
RRM1
CASP8
FCGR2A
JAK2
AKT3
CDK2
CDK4
FYN
FGF4
INPP4B
WWTR1
MAPK8
PIGB
RBBP8
RPA3
CDKN1B
BLM
BUB1B
BARD1
TOP2A
POLE
RAD51C
BRCA1
CHEK2
BRIP1
YAP1
CDKN2A
GNL3
ZC3H13
JUN
LARP7
KMT2C
HMGB1
GSTP1
GRB2
ATR
ATM
BRCA2
PMS2
RPA1
PALB2
XRCC1
RAD51
ERCC1
NBN
ERCC2
MLH1
D
Suppression of growth
Escaping immune response to cancer
Cell replicative immortality
Tumor promoting inflammation
Invasion and metastasis
Angiogenesis
Genome instability and mutation
Escaping programmed cell death
Change of cellular energetics
Proliferative signaling
TP53
CTNNB1
PTEN
KRAS
AKT1
RAC1
EGFR
ABL1
RB1
NOTCH1
KMT2C
BRCA1
ATM
ARID1A
SPOP
PMS2
PIK3CA
NOTCH2
NF1
JAK2
ESR1
ERBB2
CREBBP
CDKN2A
CASP8
BUB1B
RNF43
PLCG1
PIK3R1
NF2
MAPK1
LCK
FGFR1
ERBB3
CHEK2
CCND1
BLM
PIK3CB
FOXA1
FGFR2
ERCC2
EP300
CDH1
BRCA2
ATR
CCNE1
CDK12
BRIP1
E
Positive regulation of macromolecule metabolic process (5.7 x 10-40)
Regulation of phosphorylation (1.3 x 10-28)
Diseases of signal transduction (2.6 x 10-23)
PI3K/AKT signaling (2.2 x 10-20)
Integrated BC pathway (2.5 x 10-25)
BC pathway (9.9 x 10-23)
Breast carcinoma (3.3 x 10-15)
Ovarian neoplasm (1.5 x 10-14)
-log10(padj)
GO: Biological process (1842)
Reactome (316)
WP (150)
Human phenotype ontology (74)
Hallmarks of cancer
Promotes
Suppresses
Promotes and suppresses

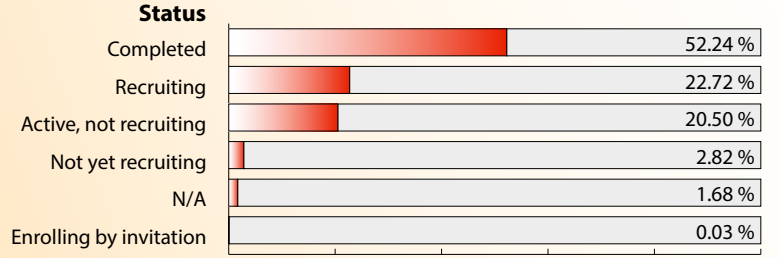

Immune system
Tyrosine kinases
Cell cycle
DNA repair
Generic transcription pathway
CD209 (DC-SIGN) signaling
LRRFIP1 activates type I IFN production
FCERI mediated MAPK activation
Fc epsilon receptor (FCERI) signaling
TRAF3-dependent IRF activation pathway
Immune System
DDX58/IFIH1-mediated induction of interferon
Regulation of actin dynamics for phagocytic cup formation
IL-4 and IL-13 signaling
Innate Immune System
DAP12 signaling
Signaling by Interleukins
Fc gamma receptor dependent phagocytosis
DAP12 interactions
ERK/MAPK targets
TLR10 cascade
TLR2 cascade
TLR7/8 cascade
MyD88 cascade initiated on plasma membrane
TRAF6 mediated induction of NFkB and MAP kinases upon TLR7/8 or 9 activation
MyD88: MAL(TIRAP) cascade
TLR2 cascade
MyD88 dependent cascade initiated on endosome
TLR6 : TLR2 cascade
TLR4 cascade
TLR5 cascade
TLR9 cascade
Oncogenic MAPK signaling
MAPK1/MAPK3 signaling
MAPK family signaling cascades
Signaling by RAS mutants
CTLA4 inhibitory signaling
RAF/MAP kinase cascade
CD28 co-stimulation
RAS signaling downstream of NF1 loss-of-function variants
IRS-related events triggered by IGF1R
CD28 dependent PI3K/Akt signaling
PI5P, PP2A and IER3 regulate PI3K/AKT signaling
IGF1R signaling cascade
RHO GTPases activate WASPs and WAVEs
Signaling by type 1 IGF1R
Constitutive signaling by aberrant PI3K in cancer
Negative regulation of the PI3K/AKT network
IRS-mediated signalling
PI-3K cascade; FGFR1
Insulin receptor signalling cascade
PI3K/AKT signaling in cancer
Activated NTRK3 signals through P13K
PI-3K cascade; FGFR2
Signaling by SCF-KIT
Signaling by NTRK3
Signaling by PDGF
Signaling by FGFR3 fusions in cancer
Signaling by NTRK1
Signaling by MET
Signaling by insulin receptor
Signaling by receptor tyrosine kinases
Signal transduction
Signaling by FGFR4 in disease
Signaling by FGFR1
Signaling by NTRK2
FGFR1 mutant receptor activation
Downstream signaling of activated FGFR2
Signaling by ligand-responsive EGFR
Signaling by FGFR
Signaling by EGFR in cancer
Activated NTRK2 signals through FYN
Signaling by FGFR1 in disease
Signaling by FGFR2 in disease
Constitutive signaling by ligand-responsive EGFR cancer variants
Signaling by FGFR3 point mutants in cancer
Signaling by EGFR
Signaling by ERBB2
VEGFA-VEGFR2 pathway
Signaling by FGFR in disease
PLCG1 events in ERBB2 signaling
Signaling by VEGF
FRS-mediated FGFR2 signaling
Signaling by EGFRvIII in cancer
Downregulation of ERBB2:ERBB3 signaling
VEGFR2 mediated vascular permeability
Signaling by FGFR4
Signaling by FGFR2
AKT-mediated inactivation of FOXO1A
Downstream signaling of activated FGFR4
Downregulation of ERBB2 signaling
VEGFR2 mediated cell proliferation
Phospholipase C-mediated cascade; FGFR2
SHC-mediated cascade;FGFR2
Negative regulation of FGFR1 signaling
Signaling by FGFR3
Constitutive signaling by EGFRvIII
RHO GTPase Effectors
Cellular response to heat stress
Phospholipase C-mediated cascade: FGFR1
Attenuation phase
Downstream signaling of activated FGFR3
FRS-mediated FGFR1 signaling
Chromatin organization
Chromatin modifying enzymes
Cellular responses to stress
SHC-mediated cascade; FGFR1
Negative regulation of FGFR2 signaling
Cellular responses to external stimuli
Reproduction
Regulation of gene expression by hypoxia-inducible factor
Downstream signaling of activated FGFR1
PIP3 activates AKT signaling
Meiosis
Meiotic recombination
Cell Cycle, Mitotic
Cyclin D associated events in G1
Mitotic G1-G1/S phases
Intracellular signaling by second messengers
G1 Phase
Cell Cycle Checkpoints
Cellular Senescence
Cell Cycle
NOTCH2 intracellular domain regulates transcription
Pre-NOTCH transcription and translation
TP53 regulates transcription of genes in G2 cell cycle
Deubiquitination
Regulation of beta-cell development
Resolution of D-loop structures through SDSA
RUNX3 regulates p14-ARF
Signaling by NOTCH4
NOTCH4 intracellular domain
Homologous DNA pairing and strand exchange
Transcriptional regulation by RUNX3
RUNX3 regulates NOTCH
G2/M Checkpoints
Transcriptional regulation by RUNX2
Notch-HLH transcription pathway
HDR through single strand annealing
Signaling by NOTCH
Recruitment and ATM-mediated phosphorylation of repair and signaling proteins at DNA double strand break
RUNX2 regulates bone development
Transcriptional regulation by AP-2 (TFAP2)
Pre-NOTCH Expression
DNA double-strand break repair
Homology directed repair
Activation of TFAP2 (AP-2)
Estrogen-dependent gene expression
HDR through homologous recombination or SSA
RUNX2 regulates osteoblast differentiation
Processing of DNA double-strand break ends
Generic Transcription Pathway
Presynaptic phase of homologous DNA pairing and strand exchange
Transcriptional regulation by RUNX1
DNA double strand break response
Resolution of D-loop structures
PTEN Regulation
ESR-mediated signaling
G2/M DNA damage checkpoint
DNA repair
TP53 Regulates Metabolic Genes
PI5P Regulates TP53 Acetylation
Resolution of D-Loop Structures
Diseases of mismatch repair
Regulation of TP53 Activity through metilation
Mismatch Repair
Regulation of TP53 activity through acetylation
Mismatch repair associated with PMS2
HDR through homologous recombination
Gene expression (Transcription)
Mismatch repair directed by MSH2:MSH3
Regulation of TP53 Activity
Mismatch repair directed by MSH2:MSH6
Activation of BH3-only proteins
Regulation of TP53 Degradation
RNA Polymerase II Transcription
Defective mismatch repair associated with MLH1
Regulation of TP53 Activity through phosphorylation
Transcriptional Regulation by TP53
Regulation of TP53 Activity through co-factors
Regulation of TP53 expression
TP53 Regulates Transcription of DNA repair genes

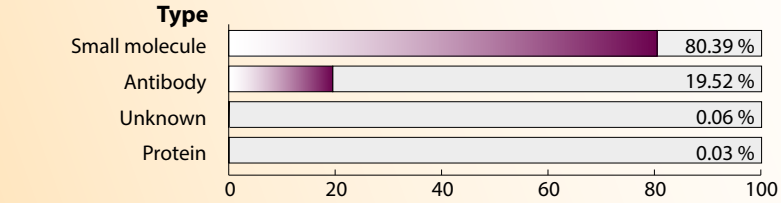




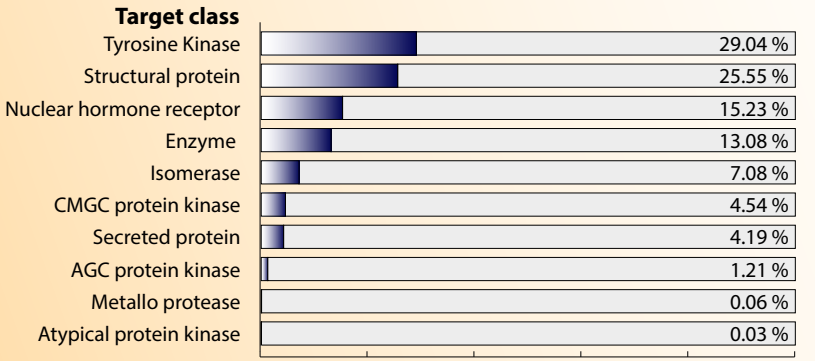
Breast tumor heterogeneity







Clinical trials
n = 3151



Precision Medicine
Open Targets