Molecular and Clinical Analyses of Gene Fusions Identify Therapeutic Targets in Paired Primary and Metastatic Breast Cancer from the AURORA Program (BIG 14-01)

Chiara Biagioni[1], Danai Fimereli[2], Alexandre Irrthum[3], Angel Guerrero-Zotano[4,5], Martine Piccart[3,6], David Cameron[7], Carmela Caballero[3], Alice Raimbault[3], Diogo Martins-Branco[6], Dario Romagnoli[1,8], Elisa Agostinetto[6], David Venet[2], Arnau Llinas-Bertran[9], Jose A. Seoane[9], Gabriele Zoppoli[10,11,12], Ana Vivancos[9], Andrea Vingiani[13,14], Mafalda Oliveira[15], Giuseppe Curigliano[14,16], Thayane Crestani[3], Christos Sotiriou[2], Einav Nili Gal-Yam[17], Florentine S. Hilbers[18], Philippe Aftimos[6], and Matteo Benelli[1,8]

ABSTRACT

Purpose: Gene fusions are molecular rearrangements with oncogenic potential. However, their role in disease progression and therapy resistance remains largely unexplored, particularly in breast cancer.

Experimental Design: In this study, we utilized multiomics data from the AURORA program to characterize gene fusions in metastatic breast cancer. We analyzed RNA sequencing data from 325 primary tumors and 350 metastatic lesions across 476 patients to develop a high-confidence catalog of gene fusions.

Results: In a subset of 398 matched samples from 199 patients, we observed a higher burden of gene fusions in metastatic tumors compared with their corresponding primary tumors. Metastatic-specific (acquired) gene fusions were characterized by few recurrent rearrangements and were generally absent in patients with *de novo* metastatic disease. We observed a colocalization of gene fusions with subtype-specific copy-number gains. Various scores indicative of genomic instability were found to be associated with the gene fusion burden. Fusions involving genes located within the same topologically associating domain (TAD) were common in HER2-positive tumors and frequently acquired in metastatic triple-negative breast cancer. The presence of gene fusions, particularly those that were acquired or involved in genes within the same TAD, was associated with poor prognosis in estrogen receptor–positive/HER2-negative tumors. Patients with acquired *ESR1* fusions exhibited a more aggressive disease course, characterized by shorter treatment response and poorer clinical outcomes.

Conclusions: The AURORA gene fusion catalog may serve as a valuable resource for identifying targetable genetic alterations, thereby supporting the development of more effective therapeutic strategies for breast cancer.

Introduction

Each year, metastatic breast cancer accounts for approximately 600,000 deaths worldwide (1). Understanding the oncogenic drivers and therapeutic vulnerabilities of metastatic breast cancer is crucial for developing treatment strategies to prevent metastases and improve the outcomes for patients with advanced disease. The Breast International Group is conducting the AURORA (Aiming to Understand the Molecular Aberrations in Metastatic Breast Cancer) program (NCT02102165). This initiative aims to enhance the understanding of metastatic breast cancer through multiomics profiling of paired primary tumors and metastatic samples, along with cell-free DNA from plasma, complemented with curated clinical data (2–4). Gene fusions, as a class of molecular alterations, have emerged as significant drivers of oncogenesis in various cancers, yet few studies have examined their role in breast cancer progression and therapeutic resistance. Gene fusions can disrupt critical cellular pathways, promoting cancer cell proliferation and invasion, and

[1]Department of Oncology, Hospital of Prato, Azienda USL Toscana Centro, Prato, Italy. [2]Breast Cancer Translational Research Laboratory, Institut Jules Bordet, Hôpital Universitaire de Bruxelles (H.U.B), Université Libre de Bruxelles (ULB), Brussels, Belgium. [3]Breast International Group (BIG), Brussels, Belgium. [4]Instituto Valenciano de Oncologia, Valencia, Spain. [5]Grupo Español de Investigación en Cáncer de Mama (GEICAM), Madrid, Spain. [6]Institut Jules Bordet, Hôpital Universitaire de Bruxelles (H.U.B), Université Libre de Bruxelles (ULB), Brussels, Belgium. [7]Institute for Genomics and Cancer, University of Edinburgh, Edinburgh, United Kingdom. [8]Department of Experimental and Clinical Biomedical Sciences, University of Florence, Florence, Italy. [9]Vall d'Hebron Institute of Oncology, Barcelona, Spain. [10]Gruppo Oncologico Italiano di Ricerca Clinica (GOIRC), Parma, Italy. [11]Department of Internal Medicine, University of Genova, Genova, Italy. [12]Ospedale Policlinico San Martino IRCCS, Genova, Italy. [13]Department of Advanced Diagnostics, Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy. [14]Istituto Europeo di Oncologia (IEO), IRCCS, Milan, Italy. [15]Vall d'Hebron University Hospital, Vall d'Hebron Institute of Oncology, Barcelona, Spain. [16]Department of Oncology and Hemato-Oncology, University of Milano, Milano, Italy. [17]Breast Oncology Institute, Sheba Medical Center, Ramat Gan, Israel. [18]Department of Molecular Pathology, The Netherlands Cancer Institute, Amsterdam, the Netherlands.

P. Aftimos and M. Benelli contributed equally to this article.

Corresponding Authors: Philippe Aftimos, Institut Jules Bordet, Rue Meylemeersch 90, Brussels 1070, Belgium. E-mail: philippe.aftimos@hubruxelles.be; and Matteo Benelli, Department of Experimental and Clinical Biomedical Sciences, University of Florence, V.le Morgagni 50, Florence 50134, Italy. E-mail: matteo.benelli@unifi.it

Clin Cancer Res 2026;32:947–62

doi: 10.1158/1078-0432.CCR-25-2707

Translational Relevance

This study provides one of the most comprehensive analyses of gene fusions in metastatic breast cancer to date, using multiomics data from the AURORA program. We demonstrate that gene fusions are more prevalent in metastatic lesions and are associated with genomic instability, subtype-specific alterations, and poor prognosis, particularly in estrogen receptor–positive/HER2-negative disease. Gene fusions within the same topologically associated domain, particularly those acquired in metastases, were more frequent in triple-negative breast cancer. The identification of clinically relevant and potentially targetable fusions, including acquired *ESR1* rearrangements, highlights the importance of fusion profiling in guiding therapeutic strategies and improving patient outcomes in advanced breast cancer.

have gained attention as potential therapeutic targets and biomarkers (5, 6). Advances in molecular diagnostics, such as the shift in clinical practice from hotspot-only next-generation sequencing testing to comprehensive DNA hybrid capture–based molecular assays and DNA-/RNA-targeted panels, have expanded the array of targets in the precision medicine field (7). Indeed, gene fusions have proven to be highly actionable in specific tumor types such as with FGFR fusions in bladder (8) and biliary tract cancers (9) but also as tissue-agnostic therapies such as for NTRK and RET fusions (10–12). In the context of metastatic breast cancer, *ESR1* fusions have been found to induce endocrine therapy resistance and contribute to metastatic progression (13–16). However, the full spectrum of gene fusions in metastatic breast cancer and their roles in tumor progression, molecular interplay, and treatment resistance are still largely unexplored.

In this article, we leveraged the unique characteristics of the AURORA dataset, which contains a large collection of RNA sequencing (RNA-seq) data from paired primary and metastatic breast cancer samples, collected prospectively at metastatic relapse before treatment or after a maximum of one line of therapy for metastatic disease. We compiled a comprehensive and high-confidence catalog of gene fusions in metastatic breast cancer and investigated their evolution during the disease progression. By integrating gene fusions with other molecular alterations from the AURORA dataset, we explored their associations with mutations, copy-number alterations (CNA), and gene expression signatures. We also examined the link between gene fusions and genomic instability in metastatic breast cancer and finally investigated their potential prognostic role using clinical outcome data from the program.

Materials and Methods

The AURORA molecular screening program

The objectives, study design, and eligibility criteria of AURORA have been previously described (2, 3). The AURORA study (NCT02102165) is a multicenter study to investigate the molecular basis of metastatic breast cancer. The study has enrolled about 1,150 patients with metastatic breast cancer who have been diagnosed with metastatic disease or who have received one line of therapy. Samples are collected from the primary tumor, metastatic biopsy, whole blood, plasma, and serum. Target gene sequencing (TGS) of 411 cancer genes was performed on DNA extracted from the primary tumor, the formalin-fixed, paraffin-embedded (FFPE) metastatic sample, whole blood, and baseline plasma sample (not included in this study). RNA-seq and CNA analyses by the OncoScan FFPE Assay Kit (Affymetrix) were also performed on the available tumor samples. The study is conducted in collaboration with Institut Jules Bordet (IJB) and Frontier Science (FS). IJB's Clinical Trials Support Unit holds the clinical database for AURORA, in which clinical data, including information on previous and subsequent treatments, as well as survival follow-up for up to 10 years, are collected.

RNA-seq

RNA was extracted from FFPE samples using the RNeasy FFPE kit (Qiagen) and eluded in RNase-free water. RNA concentration was determined using Qubit Fluorometric Quantitation (Thermo Fisher Scientific; RRID: SCR_020553) and its integrity (DV200) was assessed using the Agilent Bioanalyzer. Samples with an RNA DV200 <30% were discarded. Sequencing libraries were prepared from 100 ng of RNA from each FFPE sample using the transcriptome capture library Illumina TruSeq RNA Access library following the manufacturer's instructions and sequenced on Illumina HiSeq 2500 (RRID: SCR_016383) in 2 × 100 bp paired-end mode.

Samples were sequenced at a median of 119 million reads (IQR, 90 million–148 million; minimum of 32 million). The median duplication rate as estimated by Picard tools (https://broadinstitute.github.io/picard/) was 84% (IQR, 74.9%–91.7%) and the median insert size was 1,547.7 (IQR, 1,383–1,687.8).

Gene fusion identification

Reads were analyzed with the nf-core/rnafusion (17) version 1.0.2, including STAR-Fusion (18) version 1.5.0 (RRID: SCR_025853), FusionCatcher (https://doi.org/10.1101/011650) version 1.00 (RRID: SCR_000060), EricScript (19) version 0.5.5, pizzly (https://doi.org/10.1101/166322) version 0.37.3, and squid (20) version 1.5, and the nf-core/rnafusion v 1.2 to include Arriba (21) version 1.0 (RRID: SCR_025854), for a total of six fusion callers. For each sample, the pipeline provides the fusions called by consensus of at least two tools using a scoring formula based on the number of tool consensus and databases hits [FusionGDB (22), Mitelman (23), and Cosmic (24)], as well as executing common quality control steps browsable through the MultiQC (25) version 1.7 report (RRID: SCR_014982). The MultiQC report and running logs were reviewed and no major bias in sequences was discovered. For each sample, a list of fusions called by at least two tools or present in at least one database was *in silico* validated by FusionInspector (26) v2.1.0 (using GENCODE version 31, RRID: SCR_014966, annotation and examine_coding_effect). For paired samples, we considered a patient-specific list, including the list of detected gene fusions in primary and metastatic samples, and then validated by FusionInspector (paired-sample validation). Only validated (single-sample or paired-sample validation) fusions were kept for downstream analysis.

Relevant data related to the fusions, including protein effect, fusion fragments per million total reads (FFPM), fusion-coding DNA sequence, fusion protein sequence, junction read count, spanning fragment count, and breakpoints, were retrieved by FusionInspector output. In case of more than one validation per fusion, we selected the one with the highest FFPM and according to the fusion model (in-frame > frameshift > other).

A single-sample validation rate was used to select the best filtering strategy among the possible filters. The validation rate increased with the growing consensus among the number of tools used, reaching >90% when fusions were identified by at least four tools in agreement. The mean number of fusions called per sample with these criteria was also in line with previous results by The Cancer Genome Atlas (TCGA) consortium (6), thus confirming its validity. The use of six different tools enabled the identification of putative gene fusions with different features as highlighted by the clustering reported in Supplementary Fig. S1. The code used to generate the AURORA gene fusion catalog is available at https://github.com/bighq/aurora-gene-fusion-catalog.

Gene fusion filtering

Validated fusions called by at least four tools were selected. Fusions present in normal tissues [TCGA normal samples (27), GTEx tissues (28), and noncancer cell study] were removed as previously done (29). Paralog and same gene fusions were also discarded (list downloaded from https://github.com/STAR-Fusion/STAR-Fusion_benchmarking_data/tree/master/resources). Fusions involving genes classified by oncoKB (30) as actionable with any level of confidence in any tumor types and present in the list of fusions detected by at least three tools and validated were added. Reciprocal fusions were also added. The final list was completed with annotations using BiomaRt (RRID: SCR_019214; ref. 31). Protein kinase domains were retrieved using annoFuse (32) and domains were considered present if at least partially retained. Detailed description on the number of gene fusion calls and filtering is reported in the flow diagram of Supplementary Fig. S2.

Gene expression analysis

Reads were processed with the nf-core/rnaseq version 3.8.1, aligning against GRCh38 using STAR version 2.7.10a (33). Gene expression levels were estimated by transcript per million using RSEM (34) version 1.3.1 (RRID: SCR_000262). The PAM50 molecular subtype (35) was estimated by Genefu. The Hallmark signatures from MSigDB (36) version 7.5.1 (RRID: SCR_016863) were estimated by GSVA (37) version 1.40.1 (RRID: SCR_021058), using Gaussian kernel. Functional enrichment analysis of genes involved in fusions was done using clusterProfiler (38) version 4.6.0 (RRID: SCR_016884).

CNA and genomic instability scores

Copy-number aberration profiling using the Affymetrix OncoScan FFPE array was performed as previously described (4). Copy numbers (CN) were extracted from Affymetrix SNP arrays when available. ASCAT (39) version 3 (RRID: SCR_014695) was used with default parameters. When both primary and metastasis were available, they were co-segmented. The resulting fits were checked manually, and samples for which no copy-number aberrations were found, because of lack of aberrations or lack of purity, were discarded. Samples for which the fits reported were not optimal were reoptimized. Cases with ASCAT-provided different ploidies for the primary and metastatic samples of the same patients were manually assessed. If there was no clear evidence for that change of ploidy, a fit with similar ploidies was chosen instead. Homologous Recombination Deficiency (HRD) score was computed by scarHRD (40) version 0.1.0. Chromosomal instability signatures (CIN; CX signatures) by the CINSignatureQuantification R package version 1.1.1 (using scaledAct3 value; ref. 41) were calculated. Ploidy was determined through the calculation of the weighted median of the total CN, with the alteration size serving as the weighting parameter. Gain, neutral, and loss regions were identified as regions with total CN values exceeding, equal to, or below the ploidy level. In the genomic profile presented in **Fig. 3A**, amplifications were designated for regions with three or more copies above the ploidy, whereas deletions were attributed to regions with one copy below the ploidy.

Somatic mutations

Mutations (i.e., single-nucleotide variants and insertion/deletions) were identified by combining protein-affecting somatic calls identified in single-sample mode from OncoDNA (https://oncodna.com) and by the somatic caller PipeIT2 (42) using matching blood samples as controls from a TGS panel, including 411 cancer genes (4). PipeIT2 was also used to identify variants acquired in metastatic samples using corresponding primary samples as controls. After the exclusion of variants due to technical biases, variants were filtered based on the following filtering criteria: tumor alternative allele count (FAO) > 10, tumor read depth (FDP) > 100, strand bias (STB) < 0.90, control alternative allele percentage (computed as AO/DP) < 1%, and variant quality call (QUAL) > 100. For this study, we considered only protein-affecting mutations [based on SnpEFF annotation (43); RRID: SCR_005191] in *TP53*, *PIK3CA*, and *ESR1*. As *TP53* and *PIK3CA* mutations are mostly clonal and shared between primary and metastatic samples, we classify a tumor as mutated in these genes if a mutation is present in metastatic samples. If metastatic samples are not available, we determined the mutational status from primary samples.

Statistical analysis

Statistical analyses were performed with the R software environment for statistical computing and graphics (RRID: SCR_001905; R core team, 2022 https://www.R-project.org/).

Survival analyses used the Cox proportional hazards method, with the log-rank test. Overall survival (OS) was left truncated between the time of diagnosis of the metastatic disease and the time of inclusion in AURORA. Time to relapse (TTR) was computed from the diagnosis of primary breast cancer to diagnosis of metastatic disease (resectable locoregional recurrences were not counted as events). The median duration of follow-up was estimated with the Kaplan–Meier reverse method on the time interval between enrollment and death or between enrollment and database freeze for patients still alive on September 30, 2024. A multivariate Cox regression model was fitted to evaluate the independent effect of each covariate on OS and TTR. Relevant covariates were selected according to clinical relevance and differently when considering primary or metastatic samples. These included Ki67 measured on metastatic samples (continuous), type of metastatic disease (visceral vs. nonvisceral), metastatic burden (<3 sites vs. ≥3 sites), IHC subtypes [estrogen receptor–positive (ER+)/HER2− vs. ER+/HER2+ vs. ER−/HER2+ vs. triple-negative breast cancer (TNBC)], and presence of mutated TP53 in metastatic samples (True vs. False) for OS and primary tumor size (continuous), node status (positive vs. negative), Ki67 measured on primary (≥20 vs. <20), IHC subtypes, and presence of mutated TP53 in primary samples for TTR.

Mann–Whitney (for two groups) or Kruskal–Wallis (for three groups or more) tests were used to assess the significance of the relationship between continuous and categorical variables. All correlations are Spearman correlations. Significances between two categorical variables were assessed using the Fisher exact test.

Ethics declarations

Written informed consent was obtained from all patients registered in the AURORA program. The study was conducted in accordance with the Declaration of Helsinki and was approved by the ethics committees of all participating hospitals.

Results

The AURORA dataset

Details on the AURORA program have already been published elsewhere (3). The present analysis focuses on patients from the AURORA program for whom RNA-seq data were available at the time of the analysis. This cohort includes 476 patients, with 325 primary RNA-seq results and 350 metastatic RNA-seq results, including biopsies from a wide range of metastatic sites (**Fig. 1A**). The IHC subtype for each patient was based on that of primary sample. Supplementary Table S1 presents a summary of the clinical characteristics of these patients. Additionally, the availability of the different types of molecular data utilized in this study is illustrated in Supplementary Fig. S3.

The AURORA gene fusion catalog

To generate a high-confidence catalog of gene fusions characterizing primary and metastatic samples collected as part of the AURORA program, we used a pipeline that included the following steps: (i) gene fusion detection by six state-of-the-art tools, (ii) stringent filtering criteria to exclude fusions commonly observed in normal tissue or known as artifacts, and (iii) *in silico* validation of passing filter fusions (see "Materials and Methods" for more details). A total of 2,007 high-confidence gene fusions were identified across 675 samples, corresponding to 816 and 1,191 gene fusions in primary and metastatic samples, respectively (Supplementary Table S2). No gene fusions were detected in 127 primary and 98 metastatic samples, corresponding to 190 patients (among the 199 patients with paired samples, 93 patients had no fusions in either sample). Tumor purity, available from matched OncoScan data (see "Materials and Methods"), was not associated with the likelihood of fusion detection in either primary or metastatic samples (Supplementary Fig. S4). The average number of gene fusions per primary sample was 2.5 (range, 0–30), in line with recent TCGA findings in the TCGA-BRCA dataset (6). A significantly higher number of gene fusions were observed in metastatic versus primary samples ($P < 0.001$; **Fig. 1B**). When examining IHC subtypes, this trend was significant in ER+/HER2− samples ($P < 0.001$) and present, although not statistically significant, in ER−/HER2+ (**Fig. 1C**). There was a trend for numerically higher number of fusions across different metastatic biopsy sites compared with their corresponding primary samples, which was statistically significant for chest wall ($P = 0.01$) and liver ($P = 0.02$) biopsies (Supplementary Fig. S5).

Nineteen fusions, representing 1% of the total, were detected in two or more patients and therefore considered recurrent in our dataset (**Fig. 1D**). Top recurrent fusions included TTC6–MIPOL1 ($n = 9$), THSD4–LRRC49 ($n = 8$), ESR1–CCDC170 ($n = 4$), EIF3H–TRPS1 ($n = 3$), LRBA–SH3D19 ($n = 3$), and PCD4–RBM20 ($n = 3$). IHC subtype–specific recurrent fusions were observed. ESR1–CCDC170 and LRBA–SH3D19 were detected only in ER+/HER2− samples. In particular, ESR1–CCDC170 was detected only in metastatic samples. Other recurrent fusions identified exclusively in metastases included EIF3H–TRPS1, CDK12–MIEN1, MAP3K1–GPBP1, PHF6–HPRT1, PPM1D–BCAS3, and UR1–GPI. Seven of the 19 recurrent fusions were previously reported in public databases and 18 of 19 were intrachromosomal. Most of these fusions showed recurrent transcript breakpoints involving 5′ and/or 3′ genes. A higher frequency of alteration was observed for 5′ and 3′ fusion partner genes (**Fig. 1E**, fused in three or more patients). *TTC6* and *MIPOL1* were the most frequently fused 5′ and 3′ genes, respectively, occurring in both ER+ primary and metastatic samples (irrespective of HER2 status). Genes in the same amplicon as or near *CCND1* (*SHANK2*, *CTTN*, *CPT1A*, *UVRAG*, and *NARS2*) were commonly involved in fusions as both 5′ and 3′ partners, primarily in ER+ cases across primary and metastatic samples. *ESR1* was predominantly a 5′ partner, specifically in ER+/HER2− metastases. In HER2+ cancers, genes within the *ERBB2* cytoband, such as *CDK12* and *STARD3*, were frequently fused. Of the 20 most frequently fused genes, 16 also had recurrent breakpoints. Half of these genes (*TTC6*, *SHANK2*, *ESR1*, *ETV6*, *KDM2A*, *CDK12*, *STARD3*, *VPS13B*, and *MIPOL1*) exhibited at least three recurrent 5′ or 3′ breakpoints. With regard to fusions proposed for tissue-agnostic therapeutics (7), *FGFR1* fusions were detected in five patients (1%). A total of 149 (9%) fusions involved a gene coding for a kinase, aligning with TCGA findings (11%), 77 of which retained an intact or partial kinase domain (Supplementary Fig. S6).

Paired sample analysis for characterizing gene fusions in metastatic breast cancer

Next, we compared paired samples (i.e., primary and metastatic samples from the same patient) to study the evolution of gene fusions with disease progression. Exploiting 398 RNA-seq samples from 199 patients, we first observed a high correlation between the number of fusions detected in primary versus corresponding metastatic samples (Spearman rho = 0.6, $P < 1e{-}20$). A significantly higher number of fusions were detected in metastatic compared with primary samples (paired $P < 1e{-}4$; **Fig. 2A**). We also found a statistically significant higher number of fusions in samples of chest wall (paired $P = 0.03$), liver ($P = 0.001$), and skin ($P = 0.04$) metastases compared with primary tumors (**Fig. 2B**). When considering IHC subtypes, these results were confirmed for liver and skin metastases only in ER+/HER2− cases, which represented the most numerous group (Supplementary Fig. S7).

Gene fusions were categorized into three classes: shared, (i.e., present in both primary and metastatic samples), lost (private to the primary tumors), and acquired (detected only in the metastasis). Of all fusions detected in paired samples ($n = 1{,}216$), 43%, 19%, and 38% were identified as shared, lost, and acquired, respectively, showing comparable distribution across IHC subtypes (**Fig. 2C**). Lost fusions were associated with lower purity in metastatic compared with matched primary tumors (Spearman rho = 0.33, $P = 0.01$), suggesting that tumor purity may influence the classification of these events, whereas no association was observed for acquired fusions (Supplementary Fig. S8). A median of one acquired gene fusion per patient was observed (range, 0–16), with comparable distribution across IHC subtypes (**Fig. 2D**). For patients with *de novo* metastatic disease, a higher fraction of shared and a lower fraction of acquired fusions were observed compared with those with a metastatic relapse ($P < 1e{-}6$; **Fig. 2E**). When considering biopsies collected before the first-line therapy for metastatic disease, a significantly higher number of acquired gene fusions in patients with relapsing disease compared with *de novo* metastatic disease were observed ($P = 0.01$), with no difference for metastatic samples collected after the first-line treatment (**Fig. 2F**). Leveraging ER+/HER2− tumors, the most prevalent subtype, we observed no association with the adjuvant regimen and in particular

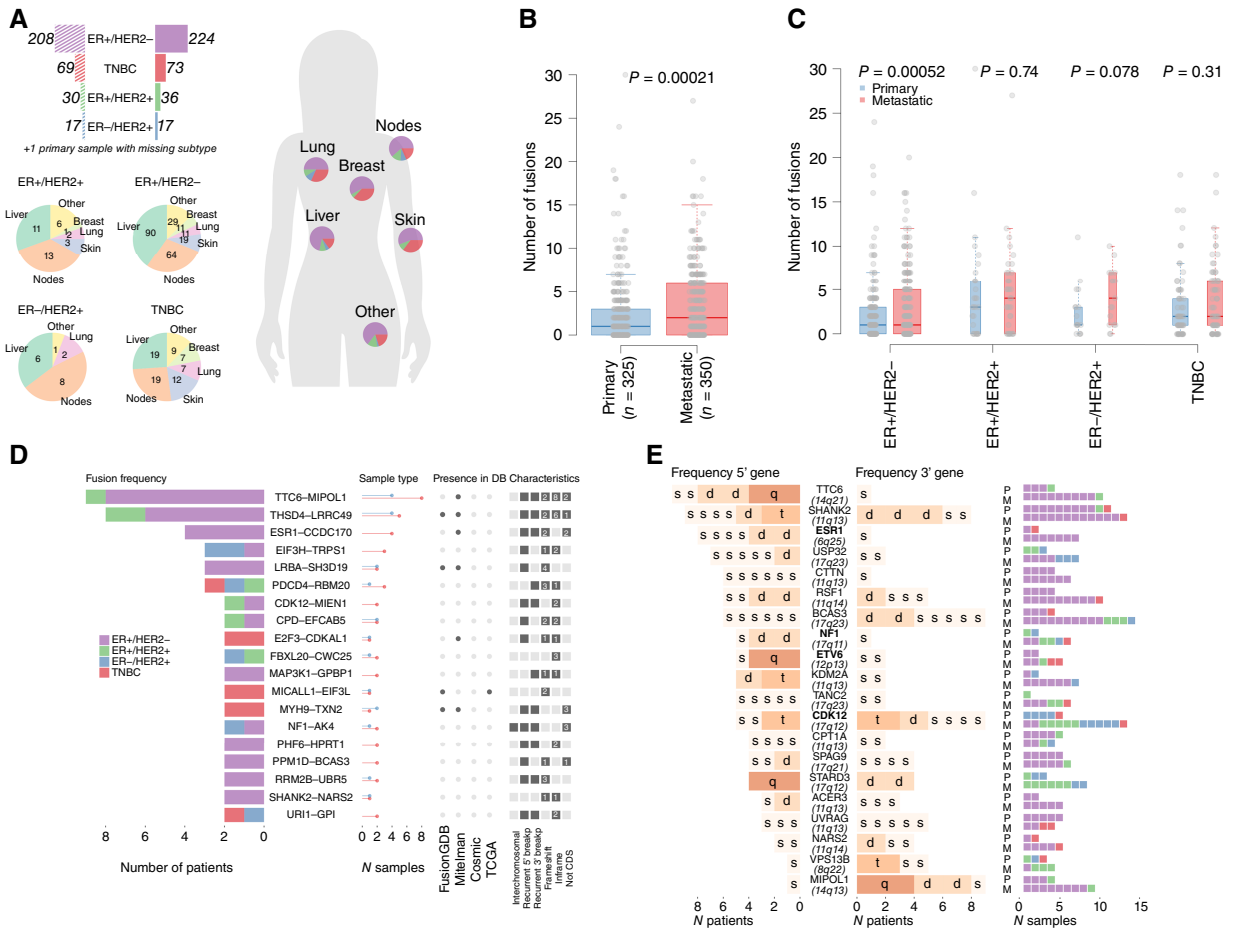

Figure 1.

The AURORA gene fusion catalog. **A,** Schematic representation of the distribution of metastatic sites among the patients included in the study, categorized by IHC subtypes (one patient not reported because of missing subtype). **B,** Box plots displaying the distribution of the number of gene fusions detected in both primary (in red) and metastatic (in blue) samples from the AURORA dataset. **C,** Box plots representing the distribution of the number of gene fusions detected in primary (in red) and metastatic (in blue) samples, stratified by various IHC subtypes. *P* values were estimated by Wilcoxon test. **D,** From left to right: bar plots indicating the number of patients with recurrent gene fusions (present in at least two patients), with sub-bars corresponding to different IHC subtypes. Dot chart presenting the number of primary (in blue) and metastatic (in red) samples associated with the recurrent gene fusions. Table displaying the presence of recurrent gene fusions in publicly available datasets (FusionGDB, Mitelman, COSMIC, and TCGA). Detailed features of recurrent gene fusions, including intra- or interchromosomal rearrangement, recurrent 5′ and 3′ breakpoints, and the number of samples showing different fusion products (such as inframe, frameshift, and noncoding). **E,** Recurrent fused 5′ (left) and 3′ (right) genes annotated with cytoband. The bar plots show the frequency of genes by patient annotated with breakpoint recurrence (s, single occurrence of breakpoint; d, duplet; t, triplet; q, quadruplet). Genes involved in breast cancer are reported in bold. The waffle chart represents the number of samples with fusion involving the corresponding gene, by sample type and primary subtype.

with the type of chemotherapy received (Supplementary Fig. S9). Among acquired fusions, we identified breast cancer–relevant and potentially actionable genes based on OncoKB (version November 2022; ref. 30), including *ESR1*, *BRIP1*, *PPM1D*, *RAD51C*, *STAT5B*, and *NF1* in ER+/HER2−, *ETV6* in ER+/HER2+, *CDK12* in ER−/HER2+, and *BRAF*, *FANCL*, *NF1*, and *PTEN* in TNBC (**Fig. 2G**). We did not identify acquired gene fusions with current drug approvals in specific disease settings such as *FGFR2* or in a tissue-agnostic fashion like NTRK and RET. Functional enrichment analysis of genes involved in acquired fusions showed enrichment of mitotic spindle assembly (q-value = 0.003; Supplementary Fig. S10). Among recurrently (in at least two patients) acquired fused genes, relevant breast cancer 5′ genes (*TTC6*, *ESR1*, *BCAS3*, and *NF1*) and 3′ genes (*MIPOL1*, *BCAS3*, *CCDC170*, and *CSMD3*) were identified.

A total of five recurrent acquired fusions were observed, all intrachromosomal, including TTC6–MIPOL1, THSD4–LRRC49, ESR1–CCDC170, EIF3H–TRPS1, and URI1–GPI, with the last two fusions not previously reported in public databases. TTC6–MIPOL1 and ESR1–CCDC170 were specifically detected in ER+/HER2− patients (Supplementary Fig. S11). Focusing on the 78 fusions involving kinase domain, disrupted domains seemed to be more frequent in lost (82%, $n = 9$) than acquired (44%, $n = 14$) and shared (50%, $n = 18$) fusions ($P = 0.11$, Fisher exact test; Supplementary Fig. S12).

Focus on *ESR1* gene fusions

Next, we examined gene fusions involving *ESR1*. In metastatic samples, *ESR1* fusions were detected in seven ER+/HER2− tumors;

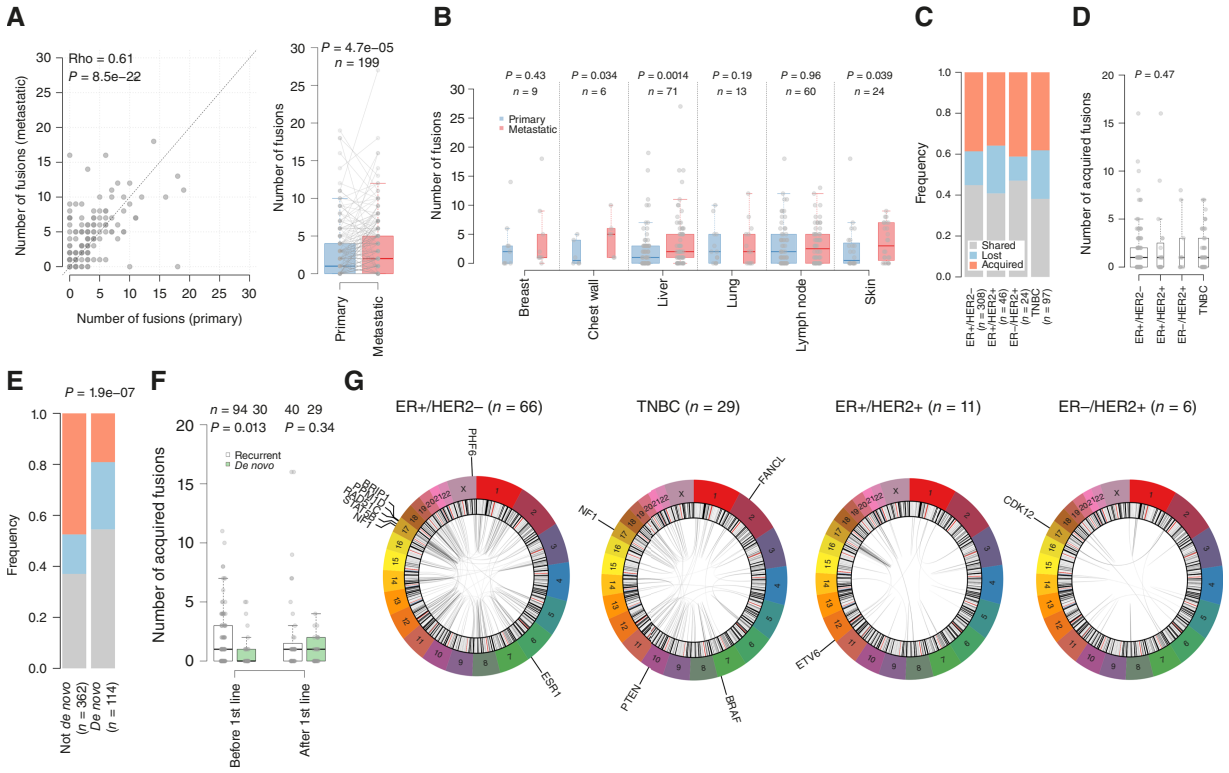

Figure 2.

Gene fusions characterizing metastatic breast cancer. **A,** Left, scatter plot reporting the number of fusions detected in metastatic (*y*-axis) vs. corresponding primary (*x*-axis) samples, with correlation estimated by the Spearman correlation test. Right, box plots displaying the distribution of the number of fusions detected in paired primary (blue) and metastatic (red) samples. *P* values were estimated by paired Wilcoxon test. **B,** Box plots displaying the distribution of the number of fusions detected in paired primary (blue) and metastatic (red) samples across the different metastatic sites. *P* values were estimated by paired Wilcoxon test. **C,** Bar plots reporting the fraction of shared (gray), lost (blue), and acquired (orange) gene fusions detected in paired samples. **D,** Box plots showing the distribution of the number of acquired fusions across the IHC subtypes. *P* value was estimated by Kruskal–Wallis test. **E,** Bar plots of the fraction of shared, lost, and acquired fusion in samples from patients with recurrent and *de novo* metastatic disease. *P* value was estimated by Fisher's exact test. **F,** Box plots displaying the number of acquired fusions for samples from patients with recurrent and *de novo* metastatic disease collected before and after the first line. *P* values were estimated by paired Wilcoxon test. **G,** Circle plots reporting acquired gene fusions detected in ER+/HER2−, ER+/HER2+, ER−/HER2+, and TNBC tumors. Actionable genes based on OncoKB are shown.

paired primary–metastatic samples were available for four, all showing fusion acquisition in metastasis. The recurrent ESR1–CCDC170 fusion was detected in four metastatic samples; three were acquired in metastasis, whereas for one, the primary sample was not available. In two of four cases of ESR1–CCDC170, the first two untranslated exons of *ESR1* and exon 3 to 11 of CCDC170 were involved, resulting in a truncated form of *CCDC170*, found previously associated with endocrine resistance (15, 44). In the other two cases, the first six exons of *ESR1* were retained, affecting the ER ligand–binding domain and potentially leading to a constitutionally active protein (45, 46). Consistently, *ESR1* gene expression was higher in the cases with retained domain (Supplementary Fig. S13). While looking at the PAM50 subtypes, ER+/HER2− harboring *ESR1* fusions were classified as luminal B (LumB; $n = 6$) or Her2 enriched, ($n = 1$), suggesting a more aggressive biology.

Two *ESR1* fusions were found exclusively in the primary tumors. AKAP7–ESR1 was identified in a TNBC primary tumor classified as basal by PAM50 in both the primary and metastatic samples. ESR1–IYD was detected in a patient with ER+/HER2− primary tumor who exhibited ER loss at relapse; this case showed PAM50 subtype switching from LumB in the primary tumor to HER2 enriched in the metastasis.

Most cases (seven of nine, 78%) developed visceral metastases (mostly in the liver, lung, brain, pleura, pericardium, and ovary). In two cases of *ESR1* fusions detected in the metastatic samples, breast cancer was *de novo* metastatic. In both cases, samples were collected after first-line treatment with chemotherapy, followed by aromatase inhibitor. In the relapsed cases, all patients had received endocrine therapy ± chemotherapy in the early setting. Details of *ESR1* fusions per single patient and main baseline characteristics and clinical outcomes are reported in Supplementary Tables S3 and S4.

Gene fusions and the other omics generated in AURORA

We investigated the association between gene fusions and the other molecular data generated within the AURORA program, including CNAs, mutations, and gene expression signatures estimated from the same RNA-seq data. Overall, we observed a marked overlap between the genomic locations of gene fusions and subtype-specific CN amplifications (**Fig. 3A**). Specifically, in ER+/HER2− tumors, gene fusions frequently involved cytoband 11q13 (containing *CCND1*), whereas in HER2+ tumors, fusions frequently

affected cytoband 17q12 (*ERBB2*). No clear association was found for TNBC, likely because of the lack of recurrent, specific CN gain events. These associations were confirmed by testing the correlation between cytoband-wise frequencies of gene fusions and CN gains. We observed a stronger correlation between fusion frequency and CN gain when comparing IHC subtypes with concordant ER/HER2 status (except for TNBC; Supplementary Fig. S14). We then checked gene fusion alterations occurring more frequently in metastatic compared with primary samples (P value < 0.2 from an FDR-unadjusted proportion test between primary and metastatic samples). In ER+/HER2−, we observed prevalent alteration of 17q23 (containing *BRIP1*), 17q24, 10q23 (containing *PTEN*), 19p13, 14q21, 14q13 (TTC6–MIPOL1 fusions), 11p15, 6q16, and 6q21. The cytoband 8q24 (*MYC*) was more frequently involved in gene fusions in metastatic versus primary TNBC. These results, except for the cytobands 17q23 and 17q24, were consistent with the frequency of acquired versus shared gene fusions in paired samples (Supplementary Table S5; Supplementary Fig. S15). Fusions involving the *ESR1* cytoband (6q25) were predominantly observed as acquired events (n = 4, 3.2%) than shared (n = 1, 0.8%) in ER+/HER2− tumors. In TNBC, the most frequently fused cytoband was 12p13, containing *ETV6*. Using matched CNA data, we verified that fused genes (both 5′ and 3′) tended to occur more frequently in CN gain and less in CN loss events than expected by chance (P < 1e−10; **Fig. 3B** for metastatic samples). Consistent results were observed across IHC subtypes in both primary and metastatic samples (Supplementary Fig. S16).

We then studied the association between the number of gene fusions per sample and the three most frequently mutated genes in metastatic breast cancer (*TP53*, *PIK3CA*, and *ESR1*) across breast cancer subtypes. No evident associations were found between mutational status of *TP53* and *PIK3CA* and gene fusions detected in primary and metastatic samples or when considering paired samples (Supplementary Figs. S17 and S18). In metastatic samples from patients with relapsing ER+/HER2− disease, *ESR1* mutations tended to co-occur with *ESR1* fusions (P = 0.11, Supplementary Fig. S19), with statistically significant association when considering fusions involving the cytoband of *ESR1* (P = 0.02; **Fig. 3C**). In patients with both paired mutation and RNA-seq data, no concurrent acquired *ESR1* mutations and *ESR1* fusions were detected in our dataset.

We then explored the association with gene expression signatures. Using the PAM50 molecular subtype (35) estimated in primary samples, we observed a higher gene fusion burden in metastatic versus primary samples of the LumB subtype (P = 0.043). Higher burden was also observed for normal (P = 0.005) and borderline significant for luminal A (LumA; P = 0.053; **Fig. 3D**). We and others (4, 48, 49) previously reported a significant fraction of PAM50 subtype switching events between primary and metastatic samples, with the majority involving LumA/B tumors switching to Her2 enriched. Among these cases, a higher number of acquired fusions were identified in LumA/B tumors that switched to Her2 enriched compared with those that remaining LumA/B in metastases (P = 0.017; **Fig. 3E**).

We then interrogated gene expression signatures associated with gene fusion burden in primary and metastatic samples across the different breast cancer subtypes (**Fig. 3F**). Correlation analyses revealed subtype-independent association between the number of fusions and hallmarks associated with proliferation and DNA damage. These included statistically significant associations for the terms MYC target, E2F target, DNA repair, and UV response down in the most numerous ER+/HER2− and TNBC subtypes (Supplementary Table S6). Subtype-specific associations were also noted. These included metabolic terms, such glycolysis in metastatic ER+/HER2−, epithelial–mesenchymal transition in primary TNBC, negative association for ER signaling in TNBC subtypes, and immune-related terms in metastatic ER+/HER2+.

Gene fusions and genomic instability

We conducted specific analyses examining the association between gene fusions and genomic instability. We first observed higher gene fusion burden in HRD score–high (>41) versus –low samples, regardless of the sample type (P < 5e−3 for primary and P < 1e−4 for metastatic samples; **Fig. 4A**; Supplementary Fig. S20). This trend was confirmed in metastatic ER+/HER2− (P = 0.001) and primary and metastatic TNBC (P = 0.001) tumors (**Fig. 4B**; Supplementary Fig. S21). The change of HRD score (rho = 0.3, P = 0.03) and mutational load (rho = 0.24, P = 0.0046) in metastatic versus primary samples was positively associated with the number of acquired fusions (**Fig. 4C**; Supplementary Fig. S22), driven by ER+/HER2− samples (rho = 0.41, P = 0.015 and rho = 0.26, P = 0.016 for HRD score and mutational load, respectively; **Fig. 4C**; Supplementary Fig. S23). We therefore considered a recently published catalog of CIN signatures (CX signatures; ref. 41). The top positively associated signature was CX13, originally linked to extrachromosomal DNA (ecDNA) circularization and amplifications (41). We observed positive association for signatures related to replication stress (CX13 and CX8), involving short-to-medium–sized events and impaired homologous recombination (CX3 and CX7; long sized), whereas negative association was detected for CX1, CX14, and CX6, all related to whole-arm/chromosome alterations (**Fig. 4D**; Supplementary Figs. S24 and S25). Notably, these associations were also observed as statistically significant in the TCGA-BRCA dataset (Supplementary Fig. S26). We therefore exploited our gene fusion catalog to investigate the relationship between the size of underlying genomic alterations (genomic distance between 5′ and 3′ breakpoint) and genomic instability. A shorter genomic distance between 5′ and 3′ breakpoints of gene fusion partners was observed for TNBC, indeed typically characterized by higher levels of genomic instability and DNA repair defects (4, 50–53), compared with ER+/HER2− patients considering both primary (P < 1e−2) and metastatic (P < 1e−3) samples. This finding was validated in the TCGA-BRCA dataset (P < 1e−10; **Fig. 4E**). Consistent with this finding, intrachromosomal fusions involving genes within the same topologically associating domains (TAD; ref. 54) were more frequent in ER−/HER2+ and TNBC compared with ER+/HER2−, both in primary (46% in ER−/HER2+ and 33% for TNBC vs. 26% in ER+/HER2−) and metastatic (36% and 33% vs. 25%) samples. This different distribution of intrachromosomal fusions involving genes within the same TAD was also validated in the TCGA-BRCA dataset (Supplementary Fig. S27, P < 1e−10). Interestingly, acquired intrachromosomal fusions in TNBC were enriched for those involving genes within the same TAD compared with other subtypes (48%, P = 0.029 and P = 0.014 vs. ER+/HER2− and ER+/HER2+, respectively). In ER−/HER2+ samples, 46% of intrachromosomal fusions in primary samples involved genes within the same TAD and were frequently located (five of 12 fusions, 42%) in the TAD including *ERBB2* (**Fig. 4F**). Of note, considering metastatic samples collected before any systemic therapy for metastatic

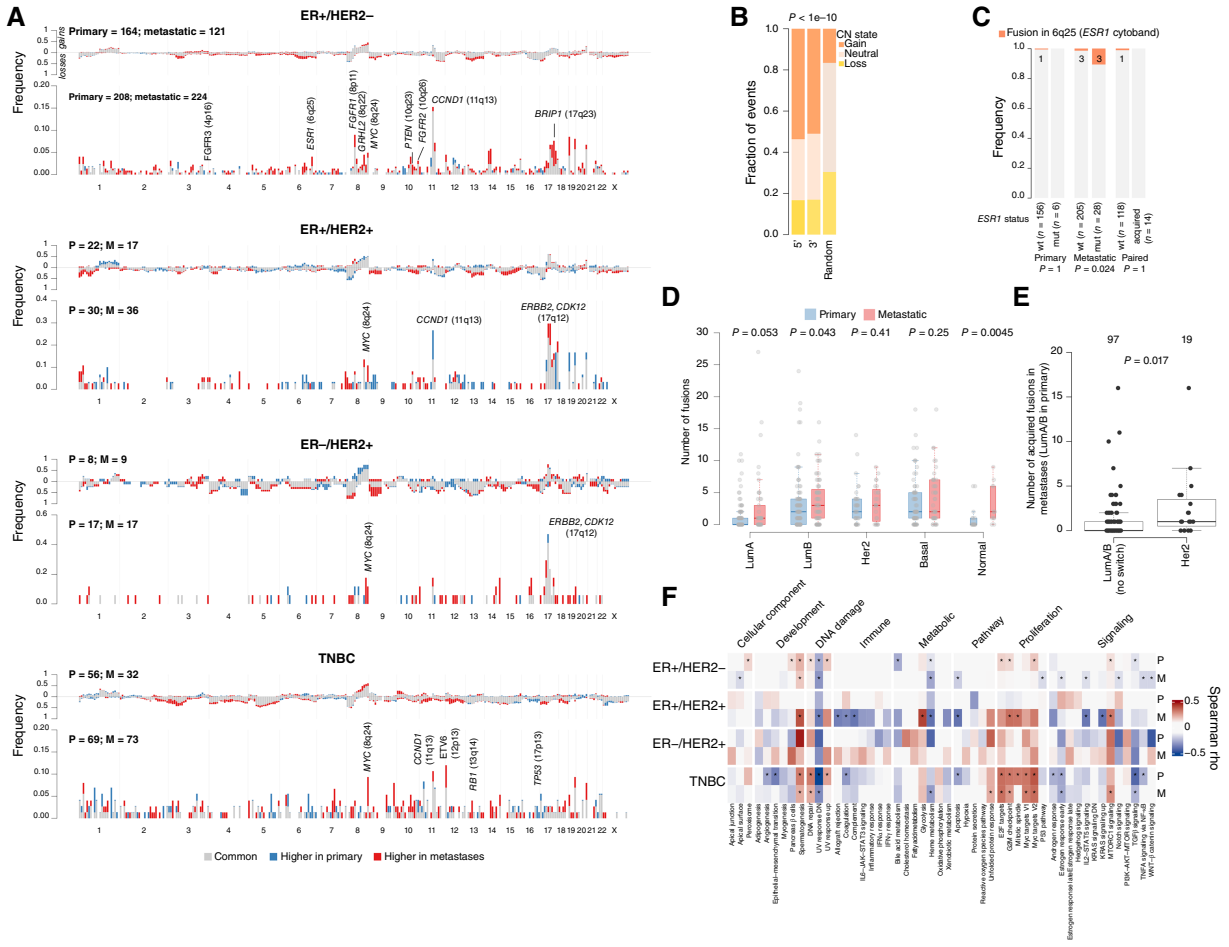

Figure 3.

Multiomics analysis of gene fusions. **A,** Frequency plot showing cytoband-wise alteration frequency of CNA (top) and gene fusions (bottom) for the different IHC subtypes. A selection of breast cancer–relevant genes involved in fusions with corresponding cytobands is reported. Blue and red sub-bars refer to cytobands showing higher alteration frequency in primary and metastatic samples, respectively. **B,** Bar plots of the fraction of 5′, 3′, and a random selection ($n = 10{,}000$ events, genes x samples) of genes involved in fusions and detected as CN loss (yellow), neutral (pink), or gain (orange) in metastatic samples. *P* value was estimated by Fisher's exact test. **C,** Bar plots showing the frequency of fusions involving genes mapping to the cytoband 6q25 (containing *ESR1*), by the mutational status of *ESR1* in primary and metastatic samples, and acquired mutations in metastases, considering paired samples. **D,** Box plots showing the number of fusions in primary (blue) and metastatic (red) samples across PAM50 subtypes estimated in primary samples. *P* values were estimated by Wilcoxon test. **E,** Box plots showing the number of acquired fusions in LumA/B primary samples remaining LumA/B and for those classified as Her2 enriched in metastatic samples. *P* value was estimated by Wilcoxon test. **F,** Heatmap showing the Spearman rho between the number of gene fusions in primary (P) and metastatic (M) samples and the levels of the hallmarks' signatures grouped by process categories as defined by Liberzon and colleagues (47). Significant associations identified by Spearman correlation test are marked with asterisks.

disease, we observed that these fusions involving genes within the same TAD were enriched in TNBC samples having received previous adjuvant treatment compared with patients with *de novo* disease ($P = 0.032$; Supplementary Fig. S28).

Impact of gene fusions on clinical outcomes

Next, we evaluated the impact of gene fusions on the outcome of patients with metastatic breast cancer ($n = 350$). To do that, we studied the association between gene fusions and OS and TTR. The median duration of follow-up from study entry was 73 months (IQR, 58–83 months). A total of 341 patients had died, including 214 patients with ER+/HER2−, 24 ER+/HER2+, 18 ER−/HER2+, and 84 TNBC. When considering the overall study population, the presence of gene fusions in primary tumor samples (not of patients with *de novo* metastatic breast cancer) was associated with a shorter TTR [31 vs. 43 months; HR, 1.33; 95% confidence interval (CI), 1.02–1.74], although not reaching statistical significance when controlling for covariates. Patients with gene fusions in metastatic biopsies ($n = 252$) also showed a statistically significantly shorter OS (25 vs. 41 months; HR, 1.58; 95% CI, 1.19–2.11) compared with those without ($n = 98$). This association was mostly driven by the patients with ER+/HER2− subtype, which remained independently associated with shorter OS (HR, 1.5; 95% CI, 1.05–2.2) in multivariate analysis controlling for relevant covariates, including proliferation (Ki67 in metastases) and *TP53* mutational status as a proxy for genomic instability. Considering 199 patients with paired primary and metastatic samples, we observed that the presence of acquired gene fusions ($n = 113$) was significantly associated with a shorter OS (21 vs.

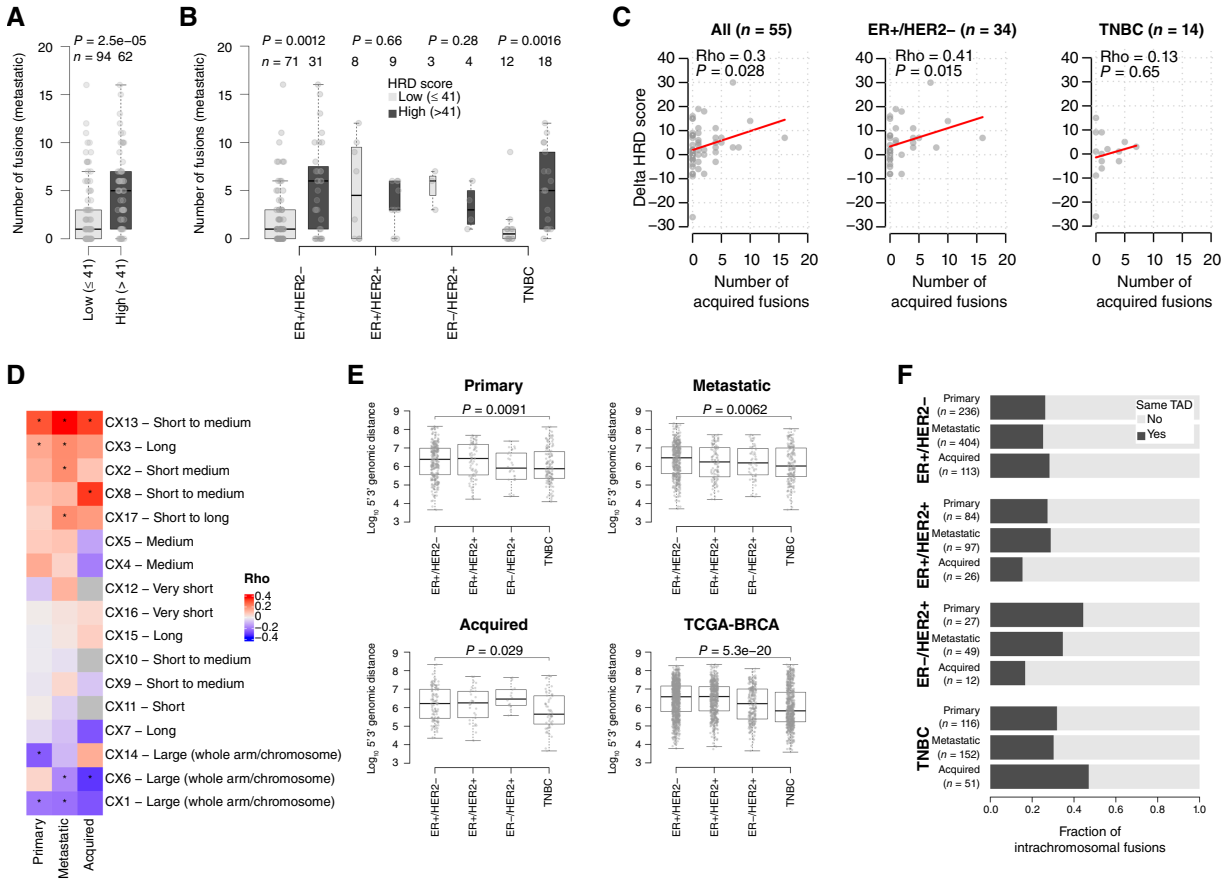

Figure 4.

Gene fusions and genomic instability. **A,** Box plots displaying the distribution of the number of fusions detected in metastatic samples based on HRD score classification (high vs. low), and **B,** by subtype. *P* values were estimated by Wilcoxon test. **C,** Scatter plot of change in HRD score in paired metastatic vs. primary samples (*y*-axis) vs. the number of acquired fusions (*x*-axis) for all, ER+/HER2−, and TNBC samples. Correlation was estimated by Spearman correlation test. **D,** Heatmap showing the Spearman correlation coefficient between the number of fusions and the different CINs (CX) defined by Drews and colleagues (41) for fusions detected in primary and metastatic samples and for acquired fusions. Significant associations identified by Spearman correlation test are marked with asterisks. **E,** Box plots displaying the $\log_{10}$ of genomic distance between 5′ and 3′ genes across the different subtypes in primary and metastatic samples, for acquired fusions, and in the TCGA-BRCA dataset. *P* values were estimated by Wilcoxon test. **F,** Bar plots showing the frequency of intrachromosomal fusions involving genes mapping to the same TAD (dark gray) vs. not across IHC subtypes for fusions detected in primary, metastatic, and acquired fusions.

42 months; HR, 1.89; 95% CI, 1.35–2.64) compared with those without ($n = 86$). Consistent with previous analyses, the association was significant exclusive for the ER+/HER2− subtype, although not reaching statistical significance when controlling for covariates (**Fig. 5A**). We then considered fusions involving genes within the same TAD and studied their association with OS and TTR. No association was found between the presence of these fusions and TTR. In ER+/HER2−, detection of same TAD fusions in metastases ($n = 59$) was statistically significantly associated with worse OS (21 vs. 46 months; HR, 1.94; 95% CI, 1.27–2.97) compared with the absence of fusions ($n = 69$), remaining independently associated when controlling for covariates. Similar results were obtained for acquired fusions involving genes in the same TAD (21 vs. 45 months; HR, 2.43; 95% CI, 1.38–4.28; **Fig. 5B**).

Discussion

In this study, we conducted an extensive characterization of gene fusions in metastatic breast cancer leveraging the rich dataset obtained through the AURORA molecular screening program. The program includes the largest collection to date of RNA-seq data from paired primary tumors and metastases, along with mutational and CNA data from a cohort of 476 patients with prospectively collected clinical data. Our program, along with the AURORA US project (48), is aimed at advancing knowledge of metastatic disease and guiding the development of new treatment strategies.

Using multiple gene fusion callers and strict filtering criteria, we identified a low number of gene fusions per patient, consistent with previous studies (6, 55). Differently from most existing studies on primary tumors, our study provides a characterization of gene fusions associated with metastatic disease in breast cancer, offering valuable insights to better understand the role of gene fusions in disease progression. Although few recurrent gene fusions were identified, some of them exhibited consistent transcriptomic breakpoints, suggesting potential utility in developing assay tests specifically targeting these recurrent events. Focusing on the

A

		n	*n* ref	HR (95% CI)	*P* value	*P* value (multivar)
TTR, primary samples						
Overall	*Detected **	*123*	*104*	*1.33 (1.02–1.74)*	*0.032*	*0.32 $*
ER+/HER2–	Detected *	75	78	1.18 (0.86–1.62)	0.32	0.43 $
TNBC	Detected *	36	14	1.15 (0.61–2.17)	0.66	0.44 $
OS, metastatic samples						
Overall	*Detected **	*252*	*98*	*1.58 (1.19–2.11)*	*0.0018*	*0.12 $$*
ER+/HER2–	Detected *	155	69	1.63 (1.14–2.32)	0.0074	**0.029** $$
TNBC	Detected *	56	17	1.05 (0.58–1.90)	0.87	0.54 $$
OS, metastatic (paired)						
Overall	*Acquired ***	*113*	*86*	*1.89 (1.35–2.64)*	*0.00021*	*0.16 $$*
ER+/HER2–	Acquired **	67	57	2.09 (1.37–3.19)	0.00067	0.084 $$
TNBC	Acquired **	29	16	1.29 (0.65–2.57)	0.46	0.77 $$

* vs. Not detected
** vs. Not acquired
$ primary: size, nodal status, Ki67, and TP53 status +/– subtype
$$ metastatic: Ki67, visceral metastasis, metastatic burden, and TP53 status +/– subtype

B

		n	*n* ref	HR (95% CI)	*P* value	*P* value (multivar)
TTR, primary samples						
Overall	*Not same TAD **	*84*	*104*	*1.29 (0.97–1.73)*	*0.08*	*0.44 $*
	*Same TAD **	*39*	*104*	*1.43 (0.98–2.08)*	*0.062*	*0.27 $*
ER+/HER2–	Not same TAD *	55	78	1.18 (0.84–1.67)	0.35	0.52 $
	Same TAD *	20	78	1.17 (0.70–1.93)	0.55	0.43 $
TNBC	Not same TAD *	23	14	1.25 (0.64–2.45)	0.51	0.29 $
	Same TAD *	13	14	1.00 (0.46–2.19)	1	0.92 $
OS, metastatic samples						
Overall	*Not same TAD **	*147*	*98*	*1.44 (1.06–1.97)*	*0.021*	*0.37 $$*
	*Same TAD **	*105*	*98*	*1.83 (1.31–2.56)*	*0.00037*	**0.032** $$
ER+/HER2–	Not same TAD *	96	69	1.48 (1.01–2.17)	0.046	0.15 $$
	Same TAD *	59	69	1.94 (1.27–2.97)	0.0022	**0.0054** $$
TNBC	Not same TAD *	31	17	0.99 (0.52–1.89)	0.98	0.5 $$
	Same TAD *	25	17	1.13 (0.58–2.22)	0.72	0.72 $$
OS, metastatic (paired)						
Overall	*Not same TAD ***	*72*	*86*	*1.81 (1.25–2.61)*	*0.0017*	*0.16 $$*
	*Same TAD ***	*41*	*86*	*2.04 (1.33–3.15)*	*0.0012*	*0.38 $$*
ER+/HER2–	Not same TAD **	43	57	1.95 (1.22–3.10)	0.0049	0.26 $$
	Same TAD **	24	57	2.43 (1.38–4.28)	0.002	**0.029** $$
TNBC	Not same TAD **	16	16	1.44 (0.65–3.20)	0.37	0.88 $$
	Same TAD **	13	16	1.16 (0.52–2.60)	0.72	0.39 $$

* vs. Not detected
** vs. Not acquired
$ primary: size, nodal status, Ki67, and TP53 status +/– subtype
$$ metastatic: Ki67, visceral metastasis, metastatic burden, and TP53 status +/– subtype

Figure 5.

Impact of gene fusions on clinical outcomes. **A,** Univariate and multivariate analysis of TTR by the presence or absence of gene fusions detection in primary samples of not *de novo* metastatic patients, left-truncated OS by the presence or absence of gene fusions detection in metastatic samples, and left-truncated OS by the presence or absence of acquired gene fusions detection in metastasis of patients with paired samples. **B,** Univariate and multivariate analysis of TTR by the absence of gene fusions, presence of gene fusions not in the same TAD, and in the same TAD in primary samples of not *de novo* metastatic; left-truncated OS by the absence of gene fusions, presence of gene fusions not in the same TAD, and in the same TAD in metastatic samples; and left-truncated OS by the absence of gene fusions, presence of acquired gene fusions not in same TAD, and presence of acquired gene fusions in the same TAD in metastasis of patients with paired samples. All analyses were performed in all populations (overall) and by IHC subtypes with adequate sample size. Cox regression hazards ratio estimate with 95% CI is reported. Multivariate analysis of survival outcomes and relevant clinicopathologic factors, reported in the table footnotes, was performed with the Cox proportional hazards model.

partners of fusions, breast cancer–relevant genes, including *ESR1*, *CDK12*, *NF1*, and *BRIP1*, were found recurrently fused. Data exist on the oncogenic role of *ESR1* fusions in mediating endocrine resistance (14, 56, 57). There is, however, to the best of our knowledge no clinical or nonclinical data on the role of gene fusions involving *CDK12*, *BRIP1*, or *NF1* in the development or progression of breast cancer despite evidence on the pathogenicity of point mutations and their vulnerability to specific therapies with variable levels of evidence such as *CDK12* and *BRIP1* with PARP inhibitors (58) and *NF1* with MEK inhibitors (59, 60).

We exploited the unique feature of the AURORA dataset that includes paired primary tumors and metastases to characterize fusions specific to metastatic samples and to classify gene fusion events into shared (present in primary tumors and metastases) and acquired (private to metastases) fusions. Our data showed that gene fusion burden can increase during the recurrence or progression of the disease as demonstrated by the higher number of gene fusion events detected in metastases compared with paired primary tumors. As gene fusions are mostly generated through genomic rearrangements, this is consistent with the typical increase of CNA burden during disease evolution (53, 61, 62). Interestingly, treatment-naïve biopsies from patients with *de novo* metastatic disease showed significantly lower number of acquired fusions compared with those collected from patients with recurrent disease, in line with our previous findings showing less acquired mutations and more similar gene expression profiles in *de novo* compared with recurrent disease (4). This finding suggests that acquired fusion events may be induced by the pressure of neoadjuvant therapies as a means of drug resistance as previously suggested for mutations private to metastases in different tumor types (63). Technically relevant, this further demonstrated the reliability of the computational approach we used for gene fusion identification and of the catalog of gene fusions provided in this study. Acquired gene fusions were characterized by few recurrent but actionable (as defined in OncoKB) rearrangements, such as fusions involving *ESR1* (only in ER+ tumors), *NF1*, *CDK12*, and *CSMD3*. With regard to fusions that might be targetable with tissue-agnostic therapeutics (7), such as those involving *FGFR1*, *FGFR2*, *BRAF*, *NTRK1*, *RET*, *ROS1*, *ALK*, *ERBB2*, and *MET*, *FGFR1* ($n = 5$ patients, 1%), *BRAF* ($n = 2$, 0.4%), and *ERBB2* ($n = 2$, 0.4%) were identified in our cohort. Of note, although *ESR1* fusions had been previously described as a non-targetable mechanism of endocrine resistance, recent nonclinical experiments have suggested RET inhibition for the treatment of ESR1–TAF–driven endocrine-resistant breast cancer (64, 65). Proteasome inhibition could also be a therapeutic strategy (66). A randomized phase II study investigating proteasome inhibitor bortezomib was completed but did not select for tumors harboring *ESR1* fusions (67). Designing clinical trials to test novel therapeutics poses a recruitment challenge when the prevalence of alterations is low. Successful single–cancer type examples include bladder cancer in which *FGFR1* to *FGFR4* gene rearrangements are found in 2.8% of cases (in addition to 14% short variants) and 8.6% of biliary tract cancers (68), leading to registrational phase II and phase III clinical trials and drug approvals in these settings (9, 69). The other regulatory approvals for therapies targeting gene fusions came after the results of multi-histology basket trials, investigating the efficacy of a drug in different types of cancers characterized by the same biomarker. Therapies targeting NTRK fusions found in <1% of cancers and RET fusions in ~1% are successful examples of drug development in niches. These clinical data would have been extremely difficult to generate in tumor-specific clinical trials. This is particularly relevant in the breast cancer field as studies designed to investigate targeted therapies against targetable gene fusions in patients with breast cancer would not be feasible. Indeed, a large molecular profiling study on kinase fusions in metastatic breast cancer identified kinase fusions in only 27 of 4,854 patients (0.6%; ref. 70). The expanding use of comprehensive genomic profiling will allow the identification of patients with metastatic breast cancer with gene fusions involving actionable genes that can be enrolled in ongoing basket studies (7) in a tissue-agnostic fashion.

The multiomics profiling of primary and metastatic samples included in the AURORA study enabled the integrative analysis of gene fusions with CNA and mutations, with the goal of evaluating the interplay between fusion rearrangements and other molecular alterations. We observed evident colocalization of gene fusions and subtype-specific copy-number gains and amplifications, such as fusions involving 11q13 (*CCND1*) for ER+ and 17q12 (*ERBB2*) for HER2+. This finding is consistent with previous studies conducted in breast cancer tumors (55) and cell lines (71), for which in the latter, the authors proposed classifying these events as passenger aberrations. Some hallmark gene expression signatures seemed to be enriched in specific breast cancer subtypes or differ between primary and metastatic samples. These data may suggest a potential driver role for fusions although further studies are needed to confirm these findings. Notably, our data show that *ESR1* fusions predominantly occurred in cases with no CN gain (four of five, three neutral CN and one deletion, considering samples with both RNA-seq and CNA data), suggesting mechanisms independent of gene amplification underlying the generation of these fusions. In TNBC, we observed frequent fusions involving genes mapping to 12p13. This cytoband includes *ETV6*, a member of the family of ETS transcription factors like the commonly fused protein ERG (72, 73). *ETV6* has been found fused with several partners, and in particular with NTRK1-2-3, in various types of both hematologic and solid tumors (74), including breast cancer (75). Consistently, the TCGA consortium reported detection of *ETV6* fusions prevalently in TNBC (three of five cases, one of which with no available IHC subtype; ref. 6). A trend toward co-occurrence between mutations and fusions of *ESR1* was observed, consistent with previous studies (14). However, upon analyzing paired samples, we did not observe any cases of acquired *ESR1* fusion in tumors with acquired *ESR1* mutations. Although novel endocrine therapies are currently being investigated in endocrine resistance mediated by *ESR1* mutations, positive phase III data have been reported with selective estrogen receptor degraders (SERD) and PROTACs, and oral SERDs elacestrant and imlunestrant have already received regulatory approval in this setting. Data are still however lacking on the efficacy of this class of drugs in the presence of *ESR1* fusions.

The evident association of gene fusion burden with gene expression proliferation signatures, in addition to the aforementioned integrative analyses, suggested a link between gene fusions and genomic instability. Conducting specific analyses to evaluate this relationship, we indeed confirmed strong positive association between gene fusion burden and HRD score and that, in paired ER+/HER2− samples, the increase in genomic instability by HRD score and mutational load was positively correlated with the number of acquired fusions. By using a recent catalog of 17 CN signatures specifically designed to characterize different types of chromosomal instability (41), we observed positive association between gene fusion burden and signatures associated with replication stress and impaired homologous recombination and negative with those involved in chromosome missegregation via defective mitosis. Interestingly, a signature originally associated with ecDNA circularization and amplifications (CX13) resulted as having the highest correlation with gene fusion burden. Also, gene fusion burden tended to positively correlate with signatures characterized by short-to-medium size alterations. Consistently, among subtypes, the typically most genomically instable TNBC showed intrachromosomal fusions characterized by shorter genomic distance between 5′ and 3′ genes. In primary HER2+, fusions were characterized by genes located in the same TAD of *ERBB2*.

Of note, in TNBC, intrachromosomal acquired fusions tended to involve genes mapping to the same TAD. The higher fraction of these fusions in metastatic samples from recurrent compared with *de novo* patients enrolled before the start of the first-line suggests that these events are associated with resistance to adjuvant chemotherapy, similar to what has been reported for ecDNA (76). Collectively, our data confirm a role for gene fusions in depicting genomic instability in breast cancer, with fusions within the same TAD exhibiting similar characteristics to ecDNA alterations (77, 78) and representing, *per se*, a proxy of genomic instability. Elucidating how gene fusion formation is mechanistically linked to HRD or other forms of genomic instability could provide important insights into tumor biology and guide strategies to target these alterations and their underlying causes. The detection of fusions in tumor samples of ER+/HER2− breast cancer was found to be associated with a poor prognosis in our study. Indeed, in general, patients with fusions detected in the primary tumor tended to have a shorter TTR, and those with fusions detected, particularly those with acquired gene fusions in the metastasis, had a shorter OS. Notably, in our study, this association retained prognostic value after controlling for proliferation and *TP53* mutational status as a proxy for genomic instability. This finding could allow the selection of patients for treatment escalation in the neoadjuvant setting and the identification of ideal candidates for novel therapies in the metastatic setting. This is particularly relevant for patients with tumors harboring *ESR1* fusions. Indeed, such patients with ER+/HER2− metastatic breast cancer in our cohort had a short TTR on adjuvant endocrine therapy and a progression-free survival on first-line CDK4/6 inhibitor (CDK4/6i) below the median expected from the pivotal phase III trials. The patients also had an aggressive biology (LumB or HER2 enriched) and a dismal course of disease as demonstrated by the presence of visceral metastases, consistent with previous reports (16). Recent advances in the escalation of neoadjuvant therapy in ER+/HER2− breast cancer include the approvals of the CDK4/6is abemaciclib and ribociclib in combination with adjuvant endocrine therapy for patients with ER+/HER2− early breast cancer selected following high-risk clinical features (79). Nonclinical data suggest that CDK4/6is are active in the setting of *ESR1* fusions (13) that can be therefore used to rationalize patient selection. Novel endocrine therapies such as SERDs, PROTACs, and CERANs are in clinical development, specifically in tumors harboring *ESR1* mutations (80). Elacestrant and imlunestrant, two oral SERDs, are already approved in ER+/HER2− metastatic breast cancer with *ESR1* mutations following the results of the EMERALD and EMBER 3 phase III trials (81, 82). Phase III studies in the adjuvant setting are ongoing. *ESR1* fusions could be additional biomarkers to be studied in such studies. Their predictive and prognostic impact can additionally be investigated in correlative studies in the pivotal phase III trials that tested novel endocrine agents. Indeed, *ESR1* fusions can be found in up to 10% of ER+/HER2− metastatic breast cancers with no detected *ESR1* mutations (57). These studies included a substantial number of patients with no detected *ESR1* mutations: 249 in EMERALD, 405 in the monotherapy arms of EMBER 3 (82), and 354 in VERITAC-2 (83). This is, however, more challenging for fusions involving genes other than *ESR1* that have a much lower prevalence in breast cancer. Although our analyses demonstrate that gene fusions influence clinical outcomes in HR+/HER2− metastatic breast cancer, it is unlikely that the prognostic impact of fusions involving driver genes can be adequately assessed within the recent pivotal clinical trials that supported new drug approvals in this setting. The sample sizes of these studies—ranging from 325 patients in INAVO-120, which evaluated the PI3Kα inhibitor inavolisib (84), to 866 patients in DESTINY-Breast06, which evaluated the anti-HER2 antibody–drug conjugate trastuzumab deruxtecan (85)—would include only a very limited number of patients harboring recurrent pathogenic fusions. This rarity poses challenges for future prospective studies, which would require comprehensive RNA-seq and dedicated bioinformatics pipelines for fusion detection during screening and patient stratification.

Limitations

The limitations of our study must be recognized. First, multiomics data, in particular CNA, were not available for all RNA-seq samples, limiting a more extensive evaluation of the associations tested in this study. Second, primary tumor samples were collected at the time of the initial breast cancer diagnosis, whereas metastatic biopsies were obtained specifically for the AURORA program. These differences may have affected sample quality and, consequently, influenced fusion detection; nonetheless, we have taken measures to address this limitation through careful analyses, including the *in silico* validation of gene fusions in paired samples; we also verified that tumor purity was not associated with the likelihood of fusion detection in either primary or metastatic samples. We observed a weak but statistically significant association between lost fusions and lower tumor purity in metastatic compared with matched primary samples, which may limit the reliability of classifying these events. Third, the sample size of our study could have affected the ability to detect some statistically significant associations. Despite these limitations, our study provides valuable insights into the role of gene fusions in metastatic breast cancer, how they evolve with disease progression, how they can contribute to genomic instability, and their role as a prognostic biomarker independent from other clinical factors especially in ER+/HER2− tumors.

We also acknowledge the limitations of our findings about prognostication. AURORA has prospectively collected treatment information, including treatment lines and outcomes on therapy. The very small number of recurrent fusions, including fusions in actionable genes, did not allow us to study the predictive and/or prognostic role of individual gene fusions beyond what was described in the article with regard to the identification of gene fusions, acquired fusions, and their association with prognosis. Furthermore, results of gene fusions were not reported to the treating physicians as RNA-seq was performed in batches at a subsequent time. Thus, drug–genomic alteration matching in AURORA did not include any matching based on the identification of a specific gene fusion (4). AURORA recruited patients irrespective of the subtype of breast cancer, and treatment choices were at the discretion of the treating physician: neoadjuvant chemotherapy, endocrine ± targeted therapy in all settings, anti-HER2 therapy, immune checkpoint inhibitors, and PARP inhibitors. The prognostic findings in our study are therefore applicable to the patient population of AURORA, a patient population seen in the clinic. They would require further validation in less heterogeneous groups of patients recruited in the same setting and treated with the same therapies such as in randomized controlled clinical trials. Investigating the prognostic potential of the detection of gene fusions and acquired gene fusions in correlative translational studies within the pivotal trials mentioned above could lead, if validated, to the

integration of our findings as stratification factors in future randomized clinical trials.

Data Availability

Instructions to access the processed data other than gene fusions for reproducibility purposes are available at the webpage https://aurora.bigagainstbreastcancer.org and can be obtained upon signature of an appropriate data transfer agreement subject to applicable laws. Instructions to access processed or raw article data to perform original research are also available on the webpage and investigators can contact aurora.researchproposals@bigagainstbc.org for inquiries. Access to data for research will be granted upon review of a project proposal and endorsement by the study Steering Committee and after entering into an appropriate data access agreement subject to applicable laws.

Authors' Disclosures

A. Irrthum reports grants from Breast Cancer Research Foundation, Fondation Cancer Luxembourg, Pfizer, Fondation contre le cancer Belgique, and National Lottery Belgium and other support from Think Pink Belgium, NIF foundation, Candriam, King Baudouin Foundation, Sogerim, Cognizant Foundation, Eurofins Foundation, and Fondation Futur 21 during the conduct of the study, as well as other support from AstraZeneca, Roche/Genentech, Novartis, Pfizer, GlaxoSmithKline, and Sanofi/Aventis outside the submitted work. A. Guerrero-Zotano reports other support from Novartis and grants and personal fees from Roche, AstraZeneca, Eli Lilly and Company, Menarini, Daiichi Sankyo, and Pierre Fabre during the conduct of the study. M. Piccart reports other support from Oncolytics, grants and personal fees from AstraZeneca, Gilead Sciences, Eli Lilly and Company, Menarini, MSD, Novartis, Pfizer, and Roche-Genentech, personal fees from Mersana and Summit Therapeutics, and grants from Servier outside the submitted work. D. Cameron reports other support from AstraZeneca, Roche/Genentech, Pfizer, Menarini/Stemline, Novartis, Daiichi Sankyo, and Eli Lilly and Company outside the submitted work. C. Caballero reports grants from Breast Cancer Research Foundation, Fondation Cancer Luxembourg, Pfizer, Think Pink Belgium, NIF Foundation, and Eurofins Foundation and other support from Candriam, King Baudoin Foundation, Sogerim, Cognizant Foundation, and Fondation Futur 21 during the conduct of the study, as well as other support from AstraZeneca, Roche/Genentech, Novartis, Pfizer, GlaxoSmithKline, and Sanofi/Aventis outside the submitted work. A. Raimbault reports grants from Breast Cancer Research Foundation, Fondation Cancer Luxembourg, Pfizer, Fondation contre le cancer Belgique, National Lottery Belgium, and Eurofins Foundation and other support from Think Pink Belgium, NIF foundation, Candriam, King Baudouin Foundation, Sogerim, Cognizant Foundation, and Fondation Futur 21 during the conduct of the study, as well as other support from AstraZeneca, Roche/Genentech, Novartis, Pfizer, GlaxoSmithKline, and Sanofi/Aventis outside the submitted work. D. Martins-Branco reports employment with the European Society for Medical Oncology (ESMO) and other support (board member) from Associação de Investigação e Cuidados de Suporte em Oncologia, American Society of Clinical Oncology, Associação Portuguesa de Cuidados Paliativos, Multinational Association of Supportive Care in Cancer, and Sociedade Portuguesa de Oncologia. E. Agostinetto reports grants from Gilead Sciences, personal fees and other support from AstraZeneca, Eli Lilly and Company, Abscint, and Gilead Sciences, personal fees from Bayer, and other support from Daiichi Sankyo and Menarini outside the submitted work. G. Zoppoli reports other support from Immunomica Ltd. during the conduct of the study, as well as personal fees from Menarini Stemline outside the submitted work. A. Vivancos reports personal fees from Bayer, Bristol Myers Squibb, Guardant Health, Incyte, and Reveal Genomics and grants and personal fees from Roche during the conduct of the study, as well as a patent for PCT/IB20 15/052245 licensed to Novartis. A. Vingiani reports personal fees from Roche, Illumina, and Eli Lilly and Company outside the submitted work. M. Oliveira reports grants from BIG during the conduct of the study, as well as grants, personal fees, and nonfinancial support from AstraZeneca, Gilead Sciences, Hoffmann-La Roche, and Pfizer and personal fees from Curio Science, Daiichi Sankyo, Eisai, i-One, Eli Lilly and Company, Medscape, MSD, ProteinQure, and Relay Therapeutics outside the submitted work. G. Curigliano reports personal fees from Roche, Daiichi Sankyo, Eli Lilly and Company, Novartis, AstraZeneca, Menarini, Gilead Sciences, Pfizer, and Exact Sciences outside the submitted work. T. Crestani reports grants from Breast Cancer Research Foundation, Fondation Cancer Luxembourg, Pfizer, Fondation contre le Cancer, National Lottery Belgium, and Eurofins Foundation and other support from Think Pink Belgium, NIF Foundation, Candriam, King Baudouin Foundation, Sogerim, Cognizant Foundation, and Fondation Futur 21 during the conduct of the study, as well as grants from Roche, AstraZeneca, Novartis, Sanofi, Pfizer, GlaxoSmithKline, and Genentech outside the submitted work. E. Nili Gal-Yam reports other support from Roche, Eli Lilly and Company, Pfizer, Novartis, AstraZeneca, MSD, and Honoraria outside the submitted work. P. Aftimos reports personal fees from Menarini, Eli Lilly and Company, Roche, Gilead Sciences, Novartis, Daiichi Sankyo, AstraZeneca, and Olema, grants from Roche, and nonfinancial support from Daiichi Sankyo, Gilead Sciences, Pfizer, MSD, Amgen, and Roche outside the submitted work. M. Benelli reports grants from BIG during the conduct of the study. No disclosures were reported by the other authors.

Authors' Contributions

C. Biagioni: Formal analysis, investigation, methodology, writing–original draft, writing–review and editing. **D. Fimereli:** Formal analysis, investigation, writing–original draft, writing–review and editing. **A. Irrthum:** Data curation, investigation, writing–original draft. **A. Guerrero-Zotano:** Validation, writing–original draft, writing–review and editing. **M. Piccart:** Conceptualization, resources, supervision, funding acquisition, writing–original draft, writing–review and editing. **D. Cameron:** Resources, supervision, validation, writing–original draft, writing–review and editing. **C. Caballero:** Data curation, validation, writing–original draft, project administration, writing–review and editing. **A. Raimbault:** Resources, data curation, validation, project administration. **D. Martins-Branco:** Data curation, formal analysis, investigation, writing–original draft, writing–review and editing. **D. Romagnoli:** Data curation, formal analysis, methodology, writing–original draft. **E. Agostinetto:** Data curation, formal analysis, validation, investigation, writing–original draft, writing–review and editing. **D. Venet:** Data curation, formal analysis, validation, investigation, methodology, writing–review and editing. **A. Llinas-Bertran:** Data curation, formal analysis, methodology, writing–review and editing. **J.A. Seoane:** Data curation, formal analysis, methodology, writing–review and editing. **G. Zoppoli:** Data curation, formal analysis, validation, writing–original draft, writing–review and editing. **A. Vivancos:** Validation, writing–review and editing. **A. Vingiani:** Data curation, validation, writing–review and editing. **M. Oliveira:** Data curation, supervision, validation, writing–original draft, writing–review and editing. **G. Curigliano:** Validation, writing–original draft, writing–review and editing. **T. Crestani:** Data curation, validation, project administration, writing–review and editing. **C. Sotiriou:** Resources, validation, writing–original draft, writing–review and editing. **E. Nili Gal-Yam:** Validation, writing–original draft, writing–review and editing. **F.S. Hilbers:** Data curation, validation, writing–original draft, project administration, writing–review and editing. **P. Aftimos:** Conceptualization, supervision, validation, investigation, writing–original draft, writing–review and editing. **M. Benelli:** Conceptualization, resources, data curation, formal analysis, supervision, investigation, methodology, writing–original draft, writing–review and editing.

Acknowledgments

The authors thank all the patients who participated in the program and their families. Current or former study Steering Committee members include P. Aftimos, M. Oliveira, M. Piccart, J. Albanell, M. Balic, P. Bedard, M. Benelli, K. Benn, J. Bliss, C. Caballero, F. Cardoso, D. Cameron, E. Ciruelos, M. Colleoni, G. Curigliano, N. Davidson, E. De Azambuja, A. Di Leo, C. Duhem, N. Harbeck, C. Herremans, F.S. Hilbers, A. Irrthum, O. Johannsson, S. Knox, B. Linderholm, S. Loi, S. Loibl, S. Marreaud, J.-L. Martinez, E. Nil Gal-Yam, D. Rea, J. Reis-Filho, E. Scheepers, C. Sotiriou, G. Viale, D. Venet, A. Vingiani, L. Yates, G. Zoppoli, H. Azim, R. Bartsch, P. Campbell, E. Carrasco, P. Dinh, V. Golfinopoulos, T. Goulioti, C. Herremans, M. Maetens, C. Tryfonidis, D. Zardavas, C. Fallato, F. Daly, S. Fielding, T. Crestani, and A. Guerrero Zotano. Local principal investigators participating in the program (who are neither authors nor part of other committees) included R. Bartsch, C. Singer, and R. Greil from Austria; J.L. Canon, L. Dirix, F. Duhoux, G. Jerusalem, K. Papadimitriou, D. Taylor, and H. Wildiers from Belgium; C. Cirkel, S. Kümmel, S. Seiler, and C. Jackisch from Germany; F. Artioli, A. Ballestrero, A. Bernardo, L. Biganzoli, G. Bisagni, E. Cretella, F. Giovanardi, L. Moscetti, A. Musolino, E. Seles, K. Cagossi, A. Mercanti, L. D. Locati, and M. Colleoni from Italy; B. Bermejo, L. Calvo Martinez, L. Garcia Estevez, J.A. Garcia Saenz, M. Gonzalez Cao, S. Gonzalez-Santiago, E. Martinez de Dueñas, S. Morales, V. Iranzo, M. Ruiz Borrego, S. Servitja Tormo, and B. Rojas from Spain; M. Ekholm and H. Lindman from Sweden; S. Aebi, C. Leo, M. Rabaglio, and M. Schwitter from Switzerland; and J. Abraham, A. Armstrong, U. Barthakur, J. Braybrooke, S. Chan, D. Davies, J. Macaskill, I. MacPherson, A. Thomson, and A. Tutt from United Kingdom. The participating groups

included Austrian Breast & Colorectal Cancer Study Group; Centre Hospitalier de Luxembourg; European Organisation for Research and Treatment of Cancer Breast Cancer Cooperative Group; German Breast Group; Grupo Espanol de Investigacion en Cancer de Mama; Gruppo Oncologico Italiano di Ricerca Clinica; Icelandic Breast Cancer Group; International Breast Cancer Study Group; National Cancer Research Institute Breast Clinical Studies Group; Swedish Association of Breast Oncologists; Spanish Breast Cancer Study Group; and Westdeutsche Studien Gruppe. The members of the Data Analysis Committee included M. Benelli, A. Irrthum, G. Zoppoli, C. Biagioni, M. Paoli, D. Romagnoli, S. Vosberg, M. Dadiani, L. Ferrando, S. Sommariva, A. Llinas-Bertran, J.A. Seoane, D. Fimereli, D. Venet, F. Santaniello, E. Agostinetto, D. Martins-Branco, and F. Santaniello. The members of the study Data Monitoring Committee included M. Castiglione, S. Litiére, and F. Orsi. Current or former members of the Molecular Advisory Board included Aftimos and colleagues, P. Aftimos, F.S. Hilbers, A. Irrthum, E. Nil Gal-Yam, M. Oliveira, A. Vingiani, M. Robson, T. Severson, and F. Santaniello. Breast International Group Headquarters members included P. Boussis, B. Herpoel, A. Karabogdan, B. Jacques, S. Massey, C. Saravia, C. Schurmans, O. Spagnolo, C. Straehle, V. Van der Veeken, A. Raimbault, A. Arahmani, I. Diaz Mesa, L. Lamcaj, J. Leemans, L. Beaupère, K. Kwok, K. Van Impe, T. Goulioti, A. Irrthum, C. Caballero, and T. Crestani. Frontier Science Scotland members included I. Bradbury. Institut Jules Bordet's Clinical Trials Support Unit members included E. Agostinetto, B. Andre, M. Bajji, S. Guillaume, M.-P. Gauthier, M. Krieguer, and E. De Azambuja. Other collaborators were Bioinformatics Unit Hospital of Prato, Breast Cancer Translational Research Laboratory, Clinigen Clinical Supplies Management, Eurofins, Europa Donna, Integrated BioBank of Luxembourg, and OncoDNA. The authors thank patients and their families. AURORA would not be possible without the generous grants from the Breast Cancer Research Foundation; Fondation Cancer (Luxembourg); Pfizer grant for non-drug research; Fondation contre le Cancer (Belgium); National Lottery (Belgium); NIF Foundation; Barrie & Dena Webb; Candriam; the Fund Friends of BIG (King Baudouin Foundation); Martine Piccart; the Hotimsky family; Sogerim; Think Pink Belgium (SMART Fund); Cognizant Foundation; Eurofins Foundation; Fondation Futur 21; and many individual donors.

Note

Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Received July 21, 2025; revised October 27, 2025; accepted December 16, 2025; posted first December 22, 2025.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71:209–49.
2. Zardavas D, Maetens M, Irrthum A, Goulioti T, Engelen K, Fumagalli D, et al. The AURORA initiative for metastatic breast cancer. Br J Cancer 2014;111:1881–7.
3. Caballero C, Irrthum A, Goulioti T, Cameron D, Norton L, Piccart M. International research to address the challenges of metastatic breast cancer: the AURORA Program (BIG 14-01). NPJ Breast Cancer 2023;9:42.
4. Aftimos P, Oliveira M, Irrthum A, Fumagalli D, Sotiriou C, Gal-Yam EN, et al. Genomic and transcriptomic analyses of breast cancer primaries and matched metastases in AURORA, the Breast International Group (BIG) molecular screening initiative. Cancer Discov 2021;11:2796–811.
5. Kumar-Sinha C, Kalyana-Sundaram S, Chinnaiyan AM. Landscape of gene fusions in epithelial cancers: seq and ye shall find. Genome Med 2015;7:129.
6. Gao Q, Liang W-W, Foltz SM, Mutharasu G, Jayasinghe RG, Cao S, et al. Driver fusions and their implications in the development and treatment of human cancers. Cell Rep 2018;23:227–38.e3.
7. Hilbers FS, Aftimos P. Expanding the landscape of actionable genomic alterations in metastatic breast cancer: comprehensive genomic profiling for all? Ann Oncol 2020;31:967–9.
8. Loriot Y, Necchi A, Park SH, Garcia-Donas J, Huddart R, Burgess E, et al. Erdafitinib in locally advanced or metastatic urothelial carcinoma. N Engl J Med 2019;381:338–48.
9. Abou-Alfa GK, Sahai V, Hollebecque A, Vaccaro G, Melisi D, Al-Rajabi R, et al. Pemigatinib for previously treated, locally advanced or metastatic cholangiocarcinoma: a multicentre, open-label, phase 2 study. Lancet Oncol 2020;21:671–84.
10. Subbiah V, Cassier PA, Siena S, Garralda E, Paz-Ares L, Garrido P, et al. Pan-cancer efficacy of pralsetinib in patients with RET fusion-positive solid tumors from the phase 1/2 ARROW trial. Nat Med 2022;28:1640–5.
11. Drilon A, Laetsch TW, Kummar S, DuBois SG, Lassen UN, Demetri GD, et al. Efficacy of larotrectinib in TRK fusion-positive cancers in adults and children. N Engl J Med 2018;378:731–9.
12. Doebele RC, Drilon A, Paz-Ares L, Siena S, Shaw AT, Farago AF, et al. Entrectinib in patients with advanced or metastatic NTRK fusion-positive solid tumours: integrated analysis of three phase 1–2 trials. Lancet Oncol 2020;21:271–82.
13. Lei JT, Shao J, Zhang J, Iglesia M, Chan DW, Cao J, et al. Functional annotation of ESR1 gene fusions in estrogen receptor-positive breast cancer. Cell Rep 2018;24:1434–44.e7.
14. Hartmaier RJ, Trabucco SE, Priedigkeit N, Chung JH, Parachoniak CA, Vanden Borre P, et al. Recurrent hyperactive ESR1 fusion proteins in endocrine therapy-resistant breast cancer. Ann Oncol 2018;29:872–80.
15. Veeraraghavan J, Tan Y, Cao X-X, Kim JA, Wang X, Chamness GC, et al. Recurrent ESR1-CCDC170 rearrangements in an aggressive subset of oestrogen receptor-positive breast cancers. Nat Commun 2014;5:4577.
16. Priedigkeit N, Lebrón-Torres A, Liao J, Alberge J-B, Morganti S, Weiss J, et al. Abstract GS03-09: characterization and proposed therapeutic exploitation of fusion RNAs in metastatic breast cancers. Cancer Res 2024;84(Suppl 9):GS03-09.
17. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, et al. The nf-core framework for community-curated bioinformatics pipelines. Nat Biotechnol 2020;38:276–8.
18. Haas BJ, Dobin A, Li B, Stransky N, Pochet N, Regev A. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. Genome Biol 2019;20:213.
19. Benelli M, Pescucci C, Marseglia G, Severgnini M, Torricelli F, Magi A. Discovering chimeric transcripts in paired-end RNA-seq data by using EricScript. Bioinformatics 2012;28:3232–9.
20. Ma C, Shao M, Kingsford C. SQUID: transcriptomic structural variation detection from RNA-seq. Genome Biol 2018;19:52.
21. Uhrig S, Ellermann J, Walther T, Burkhardt P, Fröhlich M, Hutter B, et al. Accurate and efficient detection of gene fusions from RNA sequencing data. Genome Res 2021;31:448–60.
22. Kim P, Zhou X. FusionGDB: fusion gene annotation DataBase. Nucleic Acids Res 2019;47:D994–1004.
23. Wang J, Zheng J, Lee EE, Aguilar B, Phan J, Abdilleh K, et al. A cloud-based resource for genome coordinate-based exploration and large-scale analysis of chromosome aberrations and gene fusions in cancer. Genes Chromosomes Cancer 2023;62:441–8.
24. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the catalogue of somatic mutations in cancer. Nucleic Acids Res 2019;47:D941–7.
25. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 2016;32:3047–8.
26. Haas BJ, Dobin A, Ghandi M, Van Arsdale A, Tickle T, Robinson JT, et al. Targeted in silico characterization of fusion transcripts in tumor and normal tissues via FusionInspector. Cell Rep Methods 2023;3:100467.
27. Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, et al. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. Cell 2018;173:291–304.e6.
28. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580–5.
29. Babiceanu M, Qin F, Xie Z, Jia Y, Lopez K, Janus N, et al. Recurrent chimeric fusion RNAs in non-cancer tissues and cells. Nucleic Acids Res 2016;44:2859–72.
30. Chakravarty D, Gao J, Phillips S, Kundra R, Zhang H, Wang J, et al. OncoKB: a precision oncology knowledge base. JCO Precis Oncol 2017;2017:PO.17.00011.

31. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc 2009;4:1184–91.
32. Gaonkar KS, Marini F, Rathi KS, Jain P, Zhu Y, Chimicles NA, et al. annoFuse: an R Package to annotate, prioritize, and interactively explore putative oncogenic RNA fusions. BMC Bioinformatics 2020;21:577.
33. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.
34. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 2011;12:323.
35. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009;27:1160–7.
36. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.
37. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013;14:7.
38. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation (Camb) 2021;2:100141.
39. Van Loo P, Nordgard SH, Lingjærde OC, Russnes HG, Rye IH, Sun W, et al. Allele-specific copy number analysis of tumors. Proc Natl Acad Sci U S A 2010;107:16910–5.
40. Sztupinszki Z, Diossy M, Krzystanek M, Reiniger L, Csabai I, Favero F, et al. Migrating the SNP array-based homologous recombination deficiency measures to next generation sequencing data of breast cancer. NPJ Breast Cancer 2018;4:16.
41. Drews RM, Hernando B, Tarabichi M, Haase K, Lesluyes T, Smith PS, et al. A pan-cancer compendium of chromosomal instability. Nature 2022;606:976–83.
42. Schnidrig D, Garofoli A, Benjak A, Rätsch G, Rubin MA; SOCIBP consortium; et al. PipeIT2: a tumor-only somatic variant calling workflow for molecular diagnostic Ion Torrent sequencing data. Genomics 2023;115:110587.
43. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012;6:80–92.
44. Giltnane JM, Hutchinson KE, Stricker TP, Formisano L, Young CD, Estrada MV, et al. Genomic profiling of ER$^+$ breast cancers after short-term estrogen suppression reveals alterations associated with endocrine resistance. Sci Transl Med 2017;9:eaai7993.
45. Li S, Shen D, Shao J, Crowder R, Liu W, Prat A, et al. Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. Cell Rep 2013;4:1116–30.
46. Gou X, Anurag M, Lei JT, Kim B-J, Singh P, Seker S, et al. Transcriptional reprogramming differentiates active from inactive ESR1 fusions in endocrine therapy-refractory metastatic breast cancer. Cancer Res 2021;81:6259–72.
47. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015;1:417–25.
48. Garcia-Recio S, Hinoue T, Wheeler GL, Kelly BJ, Garrido-Castro AC, Pascual T, et al. Multiomics in primary and metastatic breast tumors from the AURORA US network finds microenvironment and epigenetic drivers of metastasis. Nat Cancer 2023;4:128–47.
49. Garcia-Recio S, Thennavan A, East MP, Parker JS, Cejalvo JM, Garay JP, et al. FGFR4 regulates tumor subtype differentiation in luminal breast cancer and metastatic disease. J Clin Invest 2020;130:4871–87.
50. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534:47–54.
51. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature 2012;486:400–4.
52. Shah SP, Roth A, Goya R, Oloumi A, Ha G, Zhao Y, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature 2012;486:395–9.
53. Yates LR, Knappskog S, Wedge D, Farmery JHR, Gonzalez S, Martincorena I, et al. Genomic evolution of breast cancer metastasis and relapse. Cancer Cell 2017;32:169–84.e7.
54. Gillani R, Seong BKA, Crowdis J, Conway JR, Dharia NV, Alimohamed S, et al. Gene fusions create partner and collateral dependencies essential to cancer cell survival. Cancer Res 2021;81:3971–84.
55. Fimereli D, Fumagalli D, Brown D, Gacquer D, Rothé F, Salgado R, et al. Genomic hotspots but few recurrent fusion genes in breast cancer. Genes Chromosomes Cancer 2018;57:331–8.
56. Lei JT, Gou X, Ellis MJ. ESR1 fusions drive endocrine therapy resistance and metastasis in breast cancer. Mol Cell Oncol 2018;5:e1526005.
57. Brett JO, Ritterhouse LL, Newman ET, Irwin KE, Dawson M, Ryan LY, et al. Clinical implications and treatment strategies for ESR1 fusions in hormone receptor-positive metastatic breast cancer: a case series. Oncologist 2023;28:172–9.
58. de Bono J, Mateo J, Fizazi K, Saad F, Shore N, Sandhu S, et al. Olaparib for metastatic castration-resistant prostate cancer. N Engl J Med 2020;382:2091–102.
59. Nissan MH, Pratilas CA, Jones AM, Ramirez R, Won H, Liu C, et al. Loss of NF1 in cutaneous melanoma is associated with RAS activation and MEK dependence. Cancer Res 2014;74:2340–50.
60. Ameratunga M, McArthur G, Gan H, Cher L. Prolonged disease control with MEK inhibitor in neurofibromatosis type I-associated glioblastoma. J Clin Pharm Ther 2016;41:357–9.
61. Hieronymus H, Murali R, Tin A, Yadav K, Abida W, Moller H, et al. Tumor copy number alteration burden is a pan-cancer prognostic factor associated with recurrence and death. eLife 2018;7:e37294.
62. Bakhoum SF, Ngo B, Laughney AM, Cavallo J-A, Murphy CJ, Ly P, et al. Chromosomal instability drives metastasis through a cytosolic DNA response. Nature 2018;553:467–72.
63. Hu Z, Li Z, Ma Z, Curtis C. Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. Nat Genet 2020;52:701–8.
64. Wu J, Subbiah V. RETooling the RET inhibitor pralsetinib for ESR1 fusion-positive breast cancer and beyond. Cancer Res 2023;83:3159–61.
65. Gou X, Kim B-J, Anurag M, Lei JT, Young MN, Holt MV, et al. Kinome reprogramming is a targetable vulnerability in ESR1 fusion-driven breast cancer. Cancer Res 2023;83:3237–51.
66. Gates LA, Gu G, Chen Y, Rohira AD, Lei JT, Hamilton RA, et al. Proteomic profiling identifies key coactivators utilized by mutant ERα proteins as potential new therapeutic targets. Oncogene 2018;37:4581–98.
67. Adelson K, Ramaswamy B, Sparano JA, Christos PJ, Wright JJ, Raptis G, et al. Randomized phase II trial of fulvestrant alone or in combination with bortezomib in hormone receptor-positive metastatic breast cancer resistant to aromatase inhibitors: a New York Cancer Consortium trial. NPJ Breast Cancer 2016;2:16037.
68. Murugesan K, Necchi A, Burn TC, Gjoerup O, Greenstein R, Krook M, et al. Pan-tumor landscape of fibroblast growth factor receptor 1–4 genomic alterations. ESMO Open 2022;7:100641.
69. Loriot Y, Matsubara N, Park SH, Huddart RA, Burgess EF, Houede N, et al. Erdafitinib or chemotherapy in advanced or metastatic urothelial carcinoma. N Engl J Med 2023;389:1961–71.
70. Ross DS, Liu B, Schram AM, Razavi P, Lagana SM, Zhang Y, et al. Enrichment of kinase fusions in ESR1 wild-type, metastatic breast cancer revealed by a systematic analysis of 4854 patients. Ann Oncol 2020;31:991–1000.
71. Kalyana-Sundaram S, Shankar S, Deroo S, Iyer MK, Palanisamy N, Chinnaiyan AM, et al. Gene fusions associated with recurrent amplicons represent a class of passenger aberrations in breast cancer. Neoplasia 2012;14:702–8.
72. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun X-W, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644–8.
73. Lorenzin F, Demichelis F. Past, current, and future strategies to target ERG fusion-positive prostate cancer. Cancers (Basel) 2022;14:1118.
74. Biswas A, Rajesh Y, Mitra P, Mandal M. ETV6 gene aberrations in non-haematological malignancies: a review highlighting ETV6 associated fusion genes in solid tumors. Biochim Biophys Acta Rev Cancer 2020;1874:188389.
75. Lee S, Hu Y, Loo SK, Tan Y, Bhargava R, Lewis MT, et al. Landscape analysis of adjacent gene rearrangements reveals BCL2L14-ETV6 gene fusions in more aggressive triple-negative breast cancer. Proc Natl Acad Sci U S A 2020;117:9912–21.
76. Bailey C, Pich O, Thol K, Watkins TBK, Luebeck J, Rowan A, et al. Origins and impact of extrachromosomal DNA. Nature 2024;635:193–200.
77. Hung KL, Mischel PS, Chang HY. Gene regulation on extrachromosomal DNA. Nat Struct Mol Biol 2022;29:736–44.

78. Bailey C, Shoura MJ, Mischel PS, Swanton C. Extrachromosomal DNA-relieving heredity constraints, accelerating tumour evolution. Ann Oncol 2020; 31:884–93.
79. Johnston SRD, Toi M, O'Shaughnessy J, Rastogi P, Campone M, Neven P, et al. Abemaciclib plus endocrine therapy for hormone receptor-positive, HER2-negative, node-positive, high-risk early breast cancer (monarchE): results from a preplanned interim analysis of a randomised, open-label, phase 3 trial. Lancet Oncol 2023;24:77–90.
80. Corti C, De Angelis C, Bianchini G, Malorni L, Giuliano M, Hamilton E, et al. Novel endocrine therapies: what is next in estrogen receptor positive, HER2 negative breast cancer? Cancer Treat Rev 2023;117:102569.
81. Bidard F-C, Kaklamani VG, Neven P, Streich G, Montero AJ, Forget F, et al. Elacestrant (oral selective estrogen receptor degrader) versus standard endocrine therapy for estrogen receptor-positive, human epidermal growth factor receptor 2-negative advanced breast cancer: results from the randomized phase III EMERALD trial. J Clin Oncol 2022;40:3246–56.
82. Jhaveri KL, Neven P, Casalnuovo ML, Kim S-B, Tokunaga E, Aftimos P, et al. Imlunestrant with or without abemaciclib in advanced breast cancer. N Engl J Med 2025;392:1189–202.
83. Campone M, De Laurentiis M, Jhaveri K, Hu X, Ladoire S, Patsouris A, et al. Vepdegestrant, a PROTAC estrogen receptor degrader, in advanced breast cancer. N Engl J Med 2025;393:556–68.
84. Jhaveri KL, Im S-A, Saura C, Loibl S, Kalinsky K, Schmid P, et al. Overall survival with inavolisib in PIK_3CA-mutated advanced breast cancer. N Engl J Med 2025;393:151–61.
85. Bardia A, Hu X, Dent R, Yonemori K, Barrios CH, O'Shaughnessy JA, et al. Trastuzumab deruxtecan after endocrine therapy in metastatic breast cancer. N Engl J Med 2024;391:2110–22.