Differential impact of mutations in cell cycle checkpoint kinases on breast cancer presentation and prognosis.

Seker S[1,2], Punturi NB[1,2], Carell S[1,2], Mazumder AM[1,2], Fusco N[3,4], Bainbridge MN[5*], Haricharan S[1,2]*.

Affiliations:
[1] Aging and Cancer Immunology, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA
[2] NCI-designated Cancer Center, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA
[3] Department of Oncology and Hemato-oncology, University of Milan, Italy
[4] Division of Pathology, IRCCS European Institute of Oncology (IEO), Italy
[5] Rady Children's Institute for Genomic Medicine, San Diego, CA

*Correspondence to be addressed to Matthew Bainbridge (mbainbridge@rcigm.edu) and to Svasti Haricharan (sharicharan@sbpdiscovery.org)

Acknowledgements: Funding (to SH): Komen CCR18548157, NCI K22 CA229613.

The authors have no conflicts of interest to declare.

Abstract

Cell cycle checkpoint kinases ATM/Chk2 and ATR/Chk1 are important tumor suppressors that are activated by DNA damage and other stressors to arrest replication. Germline variants in *ATM* and *CHEK2* associate with increased incidence of estrogen receptor positive (ER+)/HER2- breast cancer and some evidence suggests that somatic inactivation of ATM/Chk2 induces therapy resistance in ER+/HER2- breast cancer patients. However, the relative contributions of germline and somatic inactivation of each of the four cell cycle checkpoint kinases to ER+/HER2- and ER-/PR-/HER2- (or triple negative breast cancer; TNBC) breast cancer incidence and outcomes has not yet been systematically assessed. Here, we comprehensively analyze associations between mutations in each cell cycle checkpoint kinase gene with age at diagnosis, tumor characteristics, and patient survival outcomes using six breast cancer datasets. Overall, we find that of all the cell cycle checkpoint kinases, *CHEK2* (p=0.001) is the only gene that is preferentially mutated in ER+/HER2- breast cancer samples relative to TNBCs. *CHEK2* mutations also associate with younger age at diagnosis of metastatic ER+/HER2- breast cancer (median age 41.5 years compared to 50 years for *CHEK2* wildtype patients, p=0.03) that is highly PR positive (10% increase in PR positivity in metastatic vs primary *CHEK2* mutant tumors, as opposed to a 25% decrease in PR positivity in *CHEK2* wildtype tumors, p=0.03). These associations appear to be driven by germline variants, rather than somatic mutations. Whether germline or somatic, deleterious (truncating, frameshift or splice site) mutations in *CHEK2* associate with rapid disease progression in response to standard first line endocrine therapy when compared to tumors with wildtype *CHEK2* in two independent datasets (HR=6.15, p=0.01). Conversely, neither *ATM* nor *ATR* mutations are enriched in ER+/HER2- breast cancers relative to TNBC. However, co-incident mutation of *ATR* and *TP53* is 2-fold enriched in this subset (p=0.002), and further enriched in metastatic (3.3%) vs primary (1%) ER+/HER2- breast cancers (p=0.002). This co-incident mutation of *ATR* and *TP53* associates with decreased PR positivity in primary ER+/HER2- breast cancer diagnoses (p<0.0001) and significantly worse metastasis-free survival than mutations in either *ATR* or *TP53* alone in two independent datasets (HR=2.01, p=0.007). This association is not specific to ER+/HER2- breast cancer but is observable in ER- breast cancer patients as well. These results provide the first systematic analysis of the contribution of germline and somatic mutations in cell cycle checkpoint kinases to tumor characteristics affecting patient prognosis and treatment response. They demonstrate a unique role for CHK2 in ER signaling and response to endocrine therapy, and a more universal role for ATR in combination with TP53 in metastatic progression across breast cancer subtypes. These findings provide rationale for streamlined testing of checkpoint kinase mutations as prognostic markers of metastatic potential and/or predictive markers of responsiveness to endocrine treatment in breast cancer patients.

Background

ATM/Chk2 and ATR/Chk1 are cell cycle checkpoints activated by DNA damage as well as many other stressors[1,2]. Although interactions between these proteins in regulating cell cycle progression are complex and often redundant a simplified representation is that ATM/Chk2 and ATR/Chk1 inhibit the cell cycle at G1/S and G2/M phases, respectively, to allow the cell time to repair damaged DNA before proceeding with the cell cycle[1,3,4]. In cases where prolonged cell cycle arrest is not sufficient to facilitate DNA repair, these checkpoint kinases trigger cell death through both p53-dependent and p53-independent mechanisms[5–8].

Breast cancer is one of the most commonly diagnosed cancers globally[9]. Estrogen receptor (ER) status of breast cancer stratifies diagnoses as ER+ and ER-[10]. ER positivity predicts response to endocrine therapies that inhibit ER signaling and are highly effective initially[11,12]. However, ~20% of ER+ breast cancer patients are intrinsically resistant to endocrine therapy and ~40% of ER+ breast cancer patients acquire resistance to endocrine therapy over time[11,13,14]. Discovering predictive biomarkers for endocrine therapy resistance that can be assayed at diagnosis is a significant clinical need. Triple negative breast cancer (TNBC) which is characterized by a lack of significant expression of ER, its downstream effector progesterone receptor (PR), and receptor tyrosine kinase growth factor HER2 is the most aggressive subtype of breast cancer, has few available targeted therapeutics, and associates with poor patient outcome[10]. Understanding molecular contributions to the evolution of treatment-resistant ER+ and TNBC can refine current diagnostics and present new therapeutic opportunities.

The importance of cell cycle regulators for the evolution of breast cancer is well established[15,16]. For instance, somatic inactivation of ATM/CHK2 signaling, through loss of upstream DNA repair pathways, induces resistance to endocrine therapy in ER+/HER2- breast cancer cells and patient tumors[17–19]. A role for germline variants in *ATM/CHEK2* in ER+/HER2- breast cancer initiation and outcomes has also been identified by large and multiple independent epidemiological studies[20–23]. Mutations in *CHEK2* and *ATM* associate specifically with increased incidence of ER+/HER2- breast cancer that associates with worse survival outcomes[21,22,24]. Other studies have independently demonstrated a significant association between high levels of ATM and responsiveness to endocrine therapy[25–27]. However, while many epidemiological studies have investigated the impact of germline *ATM/CHEK2* variants on breast cancer incidence, there are few studies investigating the impact of somatic mutations in these genes on breast cancer incidence or prognosis. The few studies that have been published have conflicting or inconclusive results[24,28,29]. Previous studies of the association of *ATR/CHEK1* mutations, either germline or somatic, with breast cancer incidence or prognosis are similarly inconclusive. Germline mutations in *ATR*, for instance, are observed in familial breast cancer, but whether they contribute substantially to tumor incidence or a specific tumor subtype remains uncertain[30,31]. There is some evidence pointing to *ATR/CHEK1* as therapeutic targets in TNBCs[32–34].

Overall, cell cycle checkpoint kinases are important guardians of the genome and well-established tumor suppressors[1]. Inactivation of any of these genes has the potential to alter tumor evolution by directing the route of cell cycle dysregulation, a fundamental feature of most tumors. Understanding the relative contribution of each of these cell cycle checkpoint kinases to the incidence of different breast cancer subtypes is critical for identifying their potential as prognostic/predictive biomarkers as well as therapeutic targets[24,35]. Here, we undertake a systematic evaluation of the relative contribution of germline and somatic mutations in cell cycle

checkpoint kinase genes to ER+/HER2- and TNBC incidence, tumor characteristics, and patient outcomes as described below.

Materials and Methods

Datasets Six datasets with mutational data from primary and metastatic breast cancer patients were combined for all initial analyses (see **Figure S1**). Analyses regarding patient outcomes were conducted in two independent datasets: MSKCC and METABRIC. Details of each dataset are described below.

MSKCC, Cancer Cell 2019 [14]: clinical and mutational data (*ESR1, TP53, ATM, ATR, CHEK2* and *CHEK1*) collected from 1756 patients with hormone receptor positive (i.e. ER and/or PR+) HR+/HER2- (n=1365), HR-/HER2+(n=58), TNBC (n=168) and HR+/HER2+(n=165) primary and metastatic breast cancers.
TCGA, Nature 2012 [36]: Downloaded from cBioPortal[37] for clinical and mutational (*ESR1*,*TP53*, *ATM*, *ATR*, *CHEK2* and *CHEK1*) analysis in December 2019.The data set is composed of clinical and mutational data collected from 825 patients with HR+/HER2- (n=486), HR-/HER2+ (n=31), TNBC (n=123) and HR+/HER2+(n=79) and undetermined (n=106) primary breast cancers.
METABRIC Nature 2012 [38] *& Nature Communication 2016* [39]*:* Downloaded from cBioPortal for clinical and mutational (*ESR1, TP53, ATR* and *CHEK2*) analysis in December 2019.The data set is composed of clinical and mutational data collected from 2509 patients with HR+/HER2- (n=1398), HR-/HER2+ (n=139), TNBC (n=320) and HR+/HER2+ (n=79108) and undetermined (n=544) primary breast cancers.
Broad, Nature 2012 [40]: Downloaded from cBioPortal for clinical and mutational (*ESR1, TP53, ATM, ATR, CHEK2* and *CHEK1*) analysis in March 2020.The data set is composed of clinical and mutational data collected from 825 patients with HR+/HER2- (n=37), HR-/HER2+ (n=6), TNBC (n=320) and HR+/HER2+ (n=108) and undetermined (n=544) primary breast cancers.
MBCP, Provisional, February 2020 [41]: Downloaded from cBioPortal for clinical and mutational (*ESR1, TP53, ATM, ATR, CHEK2* and *CHEK1*) analysis in March 2020.The data set is composed of clinical and mutational data collected from 825 patients with HR+/HER2- (n=50), HR-/HER2+ (n=8), TNBC (n=8) and HR+/HER2+ (n=21) and undetermined (n=93) metastatic breast cancers.
British Columbia, Nature 2012 [42]*:* Downloaded from cBioPortal for clinical and mutational (*ESR1, TP53, ATM, ATR, CHEK2* and *CHEK1*) analysis in March 2020.The data set is composed of clinical and mutational data collected from 825 patients with HR+/HER2- (n=9), TNBC (n=90) and undetermined (n=8) primary breast cancers.

Mutational analysis All non-synonymous mutations in *ATM, CHEK2, ATR* and *CHEK1* genes were included irrespective of category (i.e., missense, nonsense, frameshift, etc) or predicted pathogenicity. Non-synonymous mutations in *TP53* and *ESR1* genes served as controls. Mutational data were compared in three main categories: Receptor status (ER+ vs TN), sample site (primary, Pri vs metastatic, Met) and mutation type (germline vs somatic). Mutational frequency was calculated based on total number of mutations identified in a category and the total patient size of the same category. Samples without mutations in any of the indicated genes are labeled as wildtype (WT) or Rest. Truncating mutations preceding the kinase domain and all non-synonymous mutations in the kinase domain were considered kinase-affecting for analyses in **Figure 3**. Predicted frequency of kinase-affecting mutations was calculated for each gene as follows:

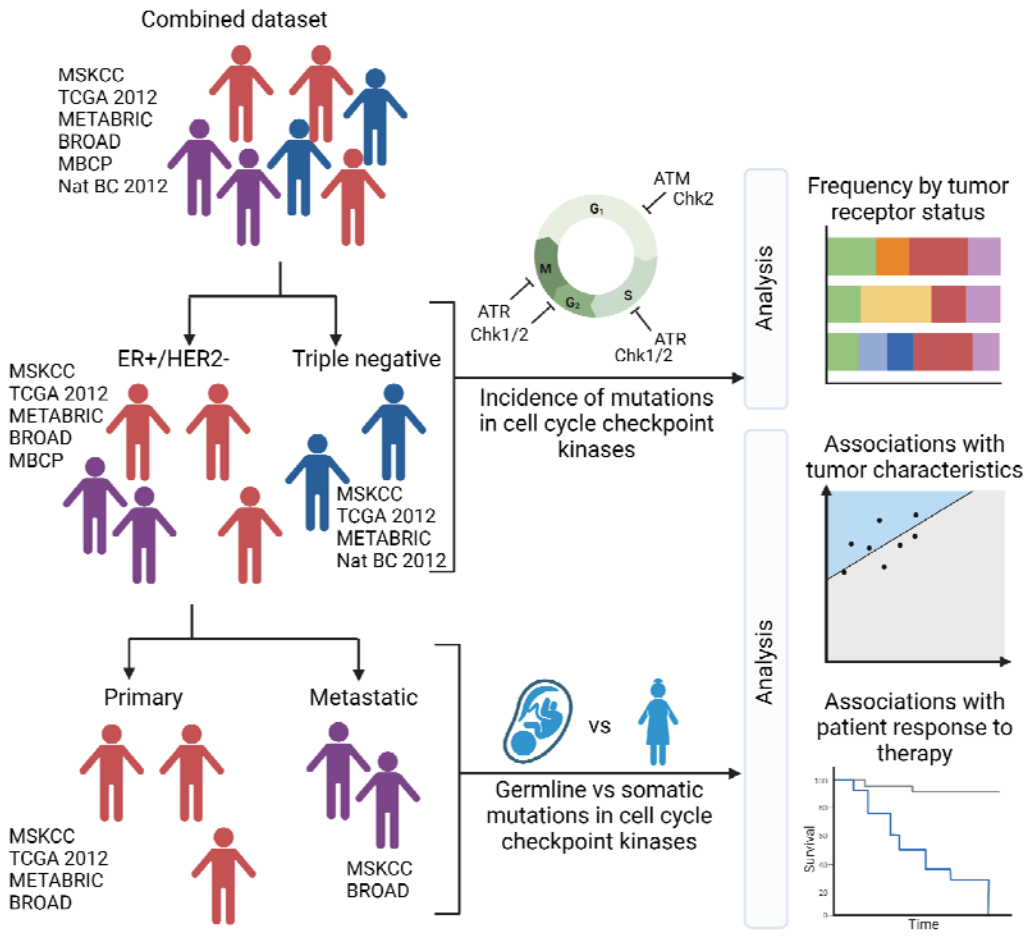

Figure 1: Outline of study. Data collected from: MSKCC, Cancer Cell 2019**,** TCGA Nature 2012, METABRIC, Nature 2012 & Nat Commun 2016, Broad, Nature 2012, MBCP, Provisional, February 2020, and British Columbia, Nature 2012. Patient tumors stratified by molecular subtype of disease and analyzed for frequency of incidence of mutations in cell cycle checkpoint kinase genes, *ATM*, *ATR*, *CHEK1* and *CHEK2*. Hormone receptor positive (ER+) patient tumors were further stratified based on primary vs metastatic disease. Associations of germline vs somatic mutations in each gene with age at diagnosis, tumor characteristics of PR status and tumor stage, and response to therapy were analyzed. Supporting data with breakup of datasets presented in **Supplementary Figure 1**.

sum((prob(truncating mutations preceding/including the kinase domain)), (prob(missense mutation in the kinase domain)). Similarly, the predicted frequency of mutations not affecting the kinase domain was calculated as sum((prob(truncating mutations succeeding kinase domain)), (prob(missense mutations preciding/succeeding kinase domain)). Probability (prob) of truncating (frameshift, nonsense, splice site) or missense mutation incidence was calculated by {[(#amino acids in specific domains)/(total #amino acids in gene)]* total # (truncating or missense) mutations detected}.

Tumor characteristics Tumor PR status, tumor stage and age of diagnosis of the patients were used as categorical variables to determine associations between incidence of mutations and patient/tumor characteristics. Fisher's exact test determined p-values by comparing different categories such as ER+/HER2- vs TNBC or germline vs somatic status for PR positivity and tumor stage while two-tailed Wilcoxon rank sum tests were used for continuous age differences.

Survival and Disease Progression analysis For univariate analyses, all survival data was obtained from MSKCC and METABRIC datasets. Outcome measures used were progression-free survival and recurrence-free survival. Only samples with survival metadata were included. Cox Proportional Hazards was used to calculate hazard ratios and p-values. Tumor stage, PR status, age at diagnosis and administration of endocrine therapy were included in multivariate analyses.

Statistical analysis Missing data were imputed with "NA" from mutation and survival data analysis. Samples classifying for more than one category were treated as separate set for statistical comparisons. Independent sample Student's T-tests were used for age comparisons and Fisher's Exact test was used for categorical data. Log rank test calculated p-values for survival analyses and Cox regression determined proportional hazards. For every analysis where multiple hypotheses were tested, p-values were adjusted using the Holm's method.

Results

ATR and CHEK2 mutations are enriched in ER+/HER2- breast cancer

To acquire a large enough sample size in which to study mutations in cell cycle checkpoint kinases, we combined six independent datasets of ER+/HER2- primary, ER+/HER2- metastatic and TNBC (**Figure 1**). Using this combined dataset, we analyzed the frequency of incidence of mutations (both germline and somatic) in each of the four cell cycle checkpoint kinase genes, *ATM*, *CHEK2*, *ATR*, and *CHEK1* in ER+/HER2- vs TNBC samples. We included known drivers, *ESR1* and *TP53* as positive controls for increased mutational frequency in ER+/HER2- and TNBC, respectively. As expected, the frequency of *ESR1* mutations was higher in ER+/HER2- vs TNBC and *TP53* mutation frequency was lower ($p<1.3e-15$ each, **Figure 2A**). The cumulative frequency of mutation incidence in all 4 cell cycle kinase genes was comparable between ER+/HER2- and TNBC samples (**Figure 2A**). However, when we considered the mutational frequency of each cell cycle checkpoint kinase gene individually, we observed a >5-fold enrichment for *CHEK2* mutations (29/3382 vs 0/640), but not for *ATR* or *CHEK1* mutations, in ER+/HER2- breast cancer samples relative to TNBCs ($p=0.001$, **Figure 2B**). Incidence of *ATM* mutations was 3-fold higher in ER+/HER2- over TNBC (33/3382 vs 2/640) but after adjustment for multiple comparisons, the p-value was no longer significant ($p=0.09$). Incidence of mutations in *CHEK1* was found to be extremely rare in both ER+/HER2- and TNBC and was therefore excluded from any further analyses. Of note, although we found no enrichment for *ATR* mutations by themselves, ER+/HER2- breast cancer samples were ~2-fold enriched for co-incident mutations in *ATR* and *TP53* while TNBC samples were not ($p=0.002$, **Figure 2C**).

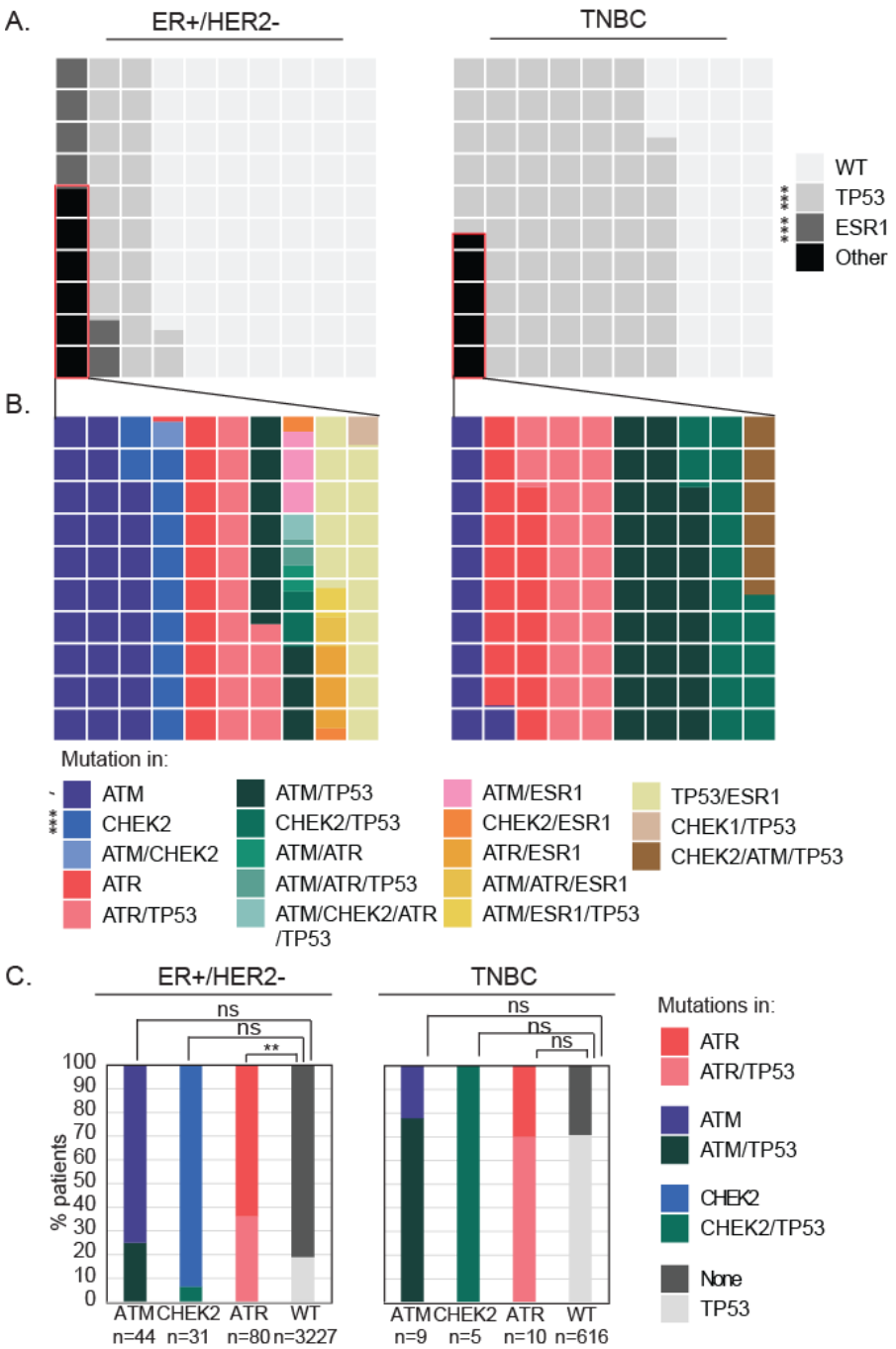

Figure 2. Mutational frequency of cell cycle checkpoint kinase genes based on breast cancer subtype. (A-B) Waffle chart charts showing *ESR1, TP53*, and cell cycle checkpoint kinase mutational frequencies in ER+/HER2- vs TNBC. Each square represents 1% of overall mutations in specified genes totaling 100% of our genes of interest. (C) Stacked column graphs representing the co-incidence of mutations in ATM/ATR/CHEK2 with mutations in TP53 in the indicated breast cancer subtypes. Fisher's Exact test determined p-values, and p-values were adjusted for multiple comparison using the Holm's method. $p \leq 0.1$', $p \leq 0.01$**, $p \leq 0.001$***, ns = not significant. Supporting data looking at mutational frequency of these genes in individual datasets in **Supplementary Figure 2**.

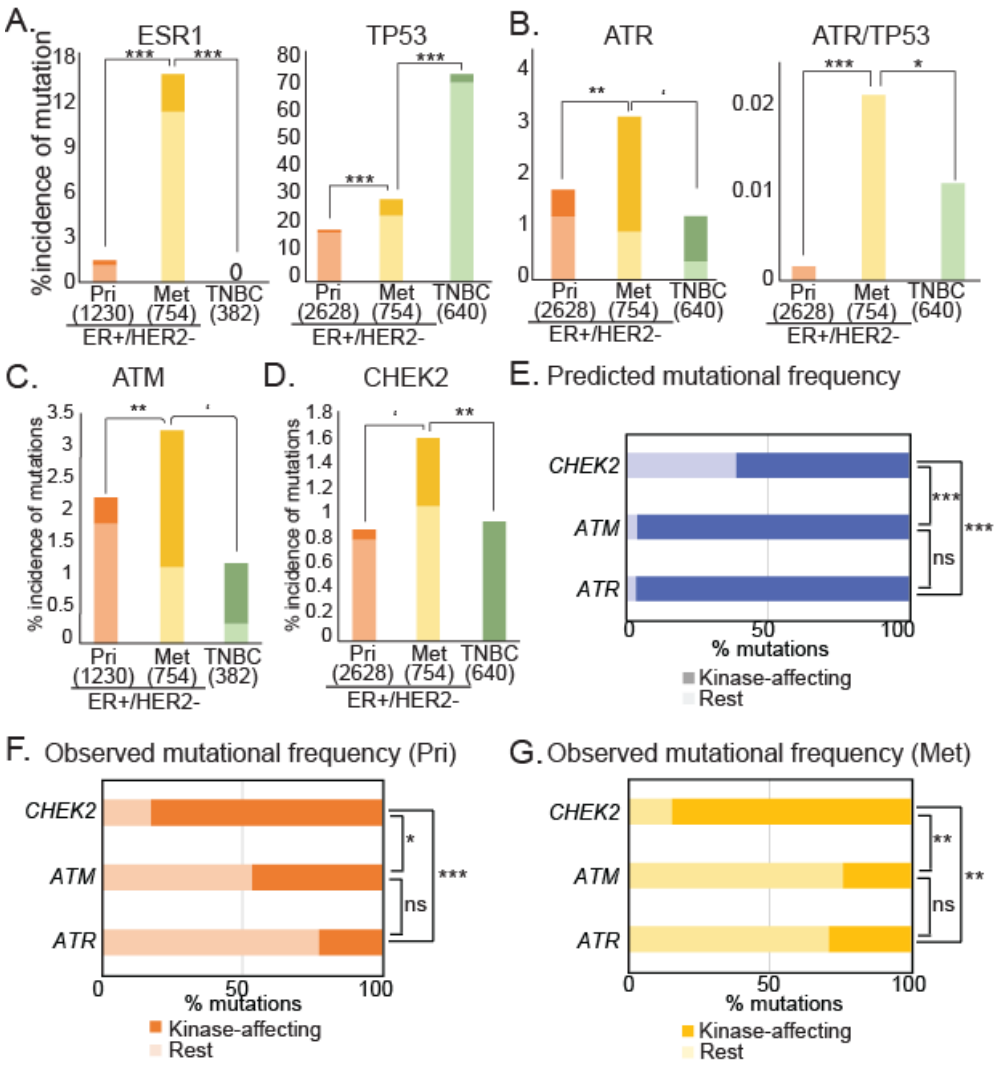

Figure 3. Distinct mutational patterns of cell cycle checkpoint kinase genes in primary vs metastatic ER+/HER2- breast cancer. (A-D) Bar graphs representing frequency of control genes: *ESR1* and *TP53* (A), *ATR*, and co-incident mutations in *ATR* and *TP53* (B), *ATM* (C), and *CHEK2* (D) mutations in ER+/HER2- primary (Pri) versus metastatic (Met) breast cancer and in TNBC. The darker shading indicates that there were mutations in multiple genes of interest while the lighter shading indicates that only the specified gene of interest was mutated in those samples. P-values were derived by comparing the incidence of mutations singly and in combination between the three breast cancer subtypes in a single Fisher's Exact test. Sample sizes in parentheses. (E-G) Stacked column graphs depicting predicted frequency of mutations (see Methods) in each cell cycle checkpoint kinase gene that can perturb the kinase domain (E) and the observed frequency in primary (Pri, F) and metastatic (Met, G) ER+/HER2- breast cancer samples. Fisher's Exact test determined all p-values. not significant, ns; $p \leq 0.1$'; $p \leq 0.05$*; $p \leq 0.01$**; $p \leq 0.001$***. P-values were adjusted for multiple comparison using the Holm's method. Supporting data with details of mutations presented as lolliplots in **Supplementary Figure 3**.

We next tested whether frequency of mutations in any cell cycle checkpoint kinase gene (except *CHEK1*) is enriched in metastatic relative to primary ER+/HER2- breast cancer. As expected, the incidence of *ESR1* mutations is highly enriched in metastatic vs primary ER+/HER2- or TNBC[43,44], while the incidence of *TP53* mutation is moderately enriched in metastatic ER+/HER2- breast cancer relative to primary ER+/HER2- breast cancer and significantly lower than in TNBCs (**Figure 3A**). We found 1.3-fold increase in mutational frequency of *ATR* singly in metastatic relative to primary ER+/HER2- (p=0.002) and >3-fold enrichment for co-incident mutations in *ATR* and *TP53* in metastatic ER+/HER2- disease (25/724) relative to both primary ER+/HER2- (27/2628, p=0.0002) and TNBC (7/640, p=0.02) (**Figure 3B**).

Similar to *ATR/TP53*, mutations in *CHEK2* are enriched in metastatic ER+ disease relative to TNBC (p=0.001), although enrichment did not attain statistical significance relative to primary ER+ disease (p=0.06 after adjustment for multiple comparisons) (**Figure 3C**). We also observed a 1.5-fold enrichment for *ATM* mutations in primary ER+/HER2- breast cancer relative to metastatic disease (p=0.003) and a~6-fold enrichment relative to TNBC (p=0.02) (**Figure 3C**). *ATM* is the only cell cycle checkpoint gene where we observed relative enrichment for mutations in the primary ER+ breast cancer setting.

We next investigated whether there were differences in the frequency of mutations that would affect kinase activity of these genes, since their main function is in phosphorylating downstream cell cycle regulators to activate them. We found that although the predicted frequency of kinase-affecting mutations is significantly higher in both *ATM* (p=0.0002) and *ATR* (p=0.00001) than in *CHEK2* (**Figure 3D**), mutations in both primary (ATM: 2-fold, p=0.04; ATR: 4-fold, p=0.0001) and metastatic (ATM: 3.5-fold, p=0.003; ATR: 3-fold, p=0.005) ER+/HER2- breast cancer samples were more likely to affect the kinase domain in *CHEK2* than in either of the other two kinases (**Figure 3E-F**). The overall mutational profile is comparable between primary and metastatic samples for all three genes, with roughly 30% of mutations being deleterious (frameshift, splice site or nonsense) and the majority being missense (**Figure S3**). Together, these data suggest that mutations in individual cell cycle checkpoint kinases are enriched in different breast cancer subtypes: *ATM* in primary, *CHEK2* in metastatic, and *ATR* in *TP53* mutant ER+/HER2- breast cancer.

CHEK2 mutations associate with higher PR positivity and younger age at diagnosis of metastatic ER+/HER2- breast cancer.

We next tested whether mutations in cell cycle checkpoint kinase genes associate with known prognostic factors in either primary or metastatic ER+/HER2- breast cancer samples. Primary ER+/HER2- breast cancer samples without mutations in any of the genes of interest are enriched for early tumor stages (I-II) relative to metastatic samples as expected (**Figure 4A**). While primary ER+/HER2- breast cancer samples with mutations in *TP53* or *ATR*, consistent with the general rule, are more likely to diagnosed at early tumor stages, this enrichment is no longer detectable in metastatic ER+/HER2- breast cancer samples with mutations in *CHEK2*, *ATM* or *ESR1* (**Figure 4A**). Further primary ER+/HER2- breast cancer without mutations in genes of interest are 1.3-fold enriched for PR positivity at diagnosis relative to metastatic disease as expected (p=3.2e-13, **Figure 4B**). However, mutations in *CHEK2* associate with higher PR positivity in metastatic ER+/HER2- breast cancer (10% increase in PR positivity in metastatic disease compared to the 30% decreased observed in *CHEK2* wildtype samples, p=0.03, **Figure 4B**).

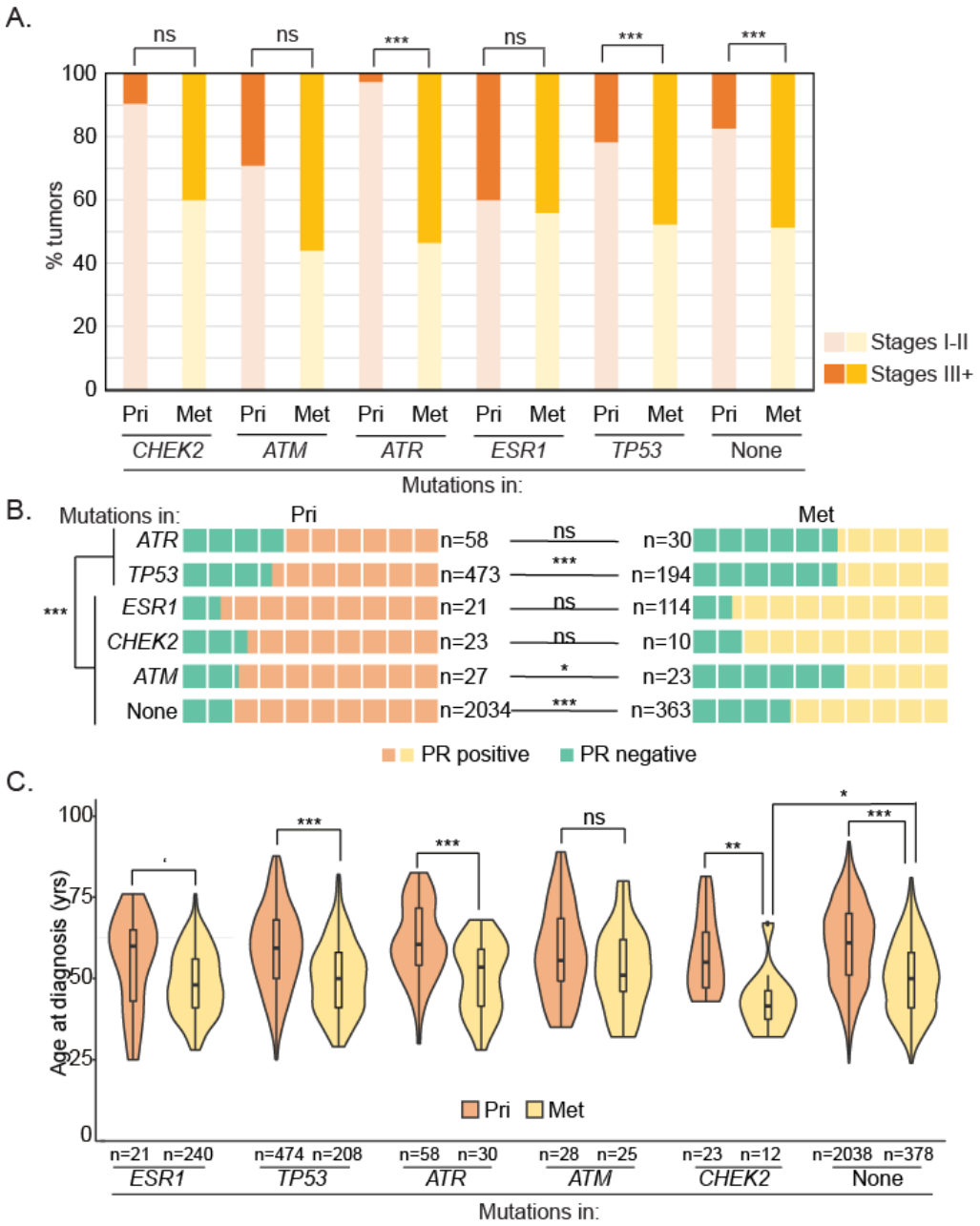

Figure 4. Comparison of primary versus metastatic samples based on PR status, tumor stage, and age at diagnosis in ER+/HER2- disease. (A) Stacked columns graphs representing tumor stage at diagnosis of primary vs metastatic ER+/HER2- breast cancer patients with mutations in any of the specified genes. Frequencies are compared for statistical significance using Fisher's Exact test. (B) Waffle chart showing proportion of PR positive and negative samples with mutations as indicated. Fisher's Exact test. (C) Violin plots representing age at diagnosis of primary vs metastatic ER+/HER2- breast cancer patients whose tumors either had mutations in *ATM*, *CHEK2* or *ESR1* or were WT. Student's T-test. not significant, ns; p≤0.1'; p≤0.05*; p≤0.01**; p≤0.001**. Holm's adjustment for multiple comparisons was conducted for all represented analyses.

This pattern is comparable to the associative effect of mutations in *ESR1* on PR positivity of metastatic ER+ breast cancer and is not observable in tumors with mutations in any of the other cell cycle checkpoint kinase genes or in *TP53* (**Figure 4B**). On the contrary, mutations in *TP53* and *ATR* associate with 1.7-fold higher likelihood of PR negativity in primary ER+/HER2- disease relative to tumors that are wildtype for both genes (p=2.7e-10, **Figure 4B**).

Finally, mutations in *CHEK2* associated with significantly younger age at diagnosis of metastatic ER+/HER2- breast cancer (41.5 years as median) relative to patients whose tumors are wildtype for *CHEK2* (median age of 50) (p=0.03, **Figure 4C**). There was no significant association with age at diagnosis for mutations in any of the other genes of interest including our control genes. This younger age at diagnosis seen in patients with *CHEK2* mutations could be driven by the incidence of germline, rather than somatic mutations. To test this hypothesis, we separated out germline and somatic mutations in each cell cycle checkpoint kinase gene. We observed no germline *ATR* mutations. However, we found 2-fold enrichment for germline *CHEK2* mutations in metastatic ER+/HER2- breast cancer (**Figure 5A**, p=0.02) and conversely, >7-fold enrichment for germline *ATM* mutations in primary ER+/HER2- breast cancer samples (p=0.04, **Figure 5B**).

We compared the landscape of germline and somatic *CHEK2* mutations and found 3-fold enrichment for deleterious (nonsense, frameshift or splice site) mutations in the germline group (12/18 germline vs 4/18 somatic, p=0.02) and no such enrichment for deleterious mutations in germline or somatic *ATM* mutations (p=0.13) (**Figure 5C**, **Figure S4**). We next assessed how somatic mutations in *CHEK2* compared to germline mutations in terms of association with PR positivity and age at diagnosis in primary and metastatic ER+/HER2- breast cancer. We found no statistically significant differences in PR positivity between primary and metastatic ER+/HER2- samples with either germline or somatic *CHEK2* mutations unlike the significant decrease in PR positivity in *CHEK2* wildtype samples (p>0.99 vs p<9e-16, **Figure 5D**). However, we detected a 3-fold increase in PR positivity in metastatic ER+ tumors with germline mutations in *CHEK2*, but not in those with somatic *CHEK2* mutations, relative to *CHEK2* wildtype tumors (p=0.02, **Figure 5D**). We also found a statistically significant association between younger age at diagnosis of metastatic ER+ disease and the incidence of germline (median age: 41, p=0.002), and not somatic (median age: 51, p>0.9), *CHEK2* mutations relative to *CHEK2* wildtype (median age: 49) metastatic ER+/HER2- breast cancer patients (**Figure 5E**). Interestingly, we did not observe a similar impact for germline *CHEK2* mutations on age at diagnosis for primary ER+/HER2- breast cancer patients. Overall, these data suggest that germline mutations in *CHEK2* may be implicated in the evolution of highly PR positive, metastatic ER+/HER2- breast cancer at a young age.

Mutations in cell cycle checkpoint kinase genes have prognostic and predictive value

Given these associations between *CHEK2* mutations and metastatic ER+/HER2- breast cancer, we next tested their association with patient outcome. We used two independent datasets, MSKCC[45] and METABRIC[38], to address this question. We found no significant association between mutations in *CHEK2* and disease-free or overall survival in either MSKCC or METABRIC. However, in MSKCC, we found that in metastatic ER+/HER2- breast cancer patients who progressed on their first two lines of endocrine therapy, both germline and somatic *CHEK2* mutations associate with shorter progression-free survival (average of 183 days) relative to patients with *CHEK2* wildtype disease (584 days) (p=0.03, **Figure 6A**, **Figure S5A**).

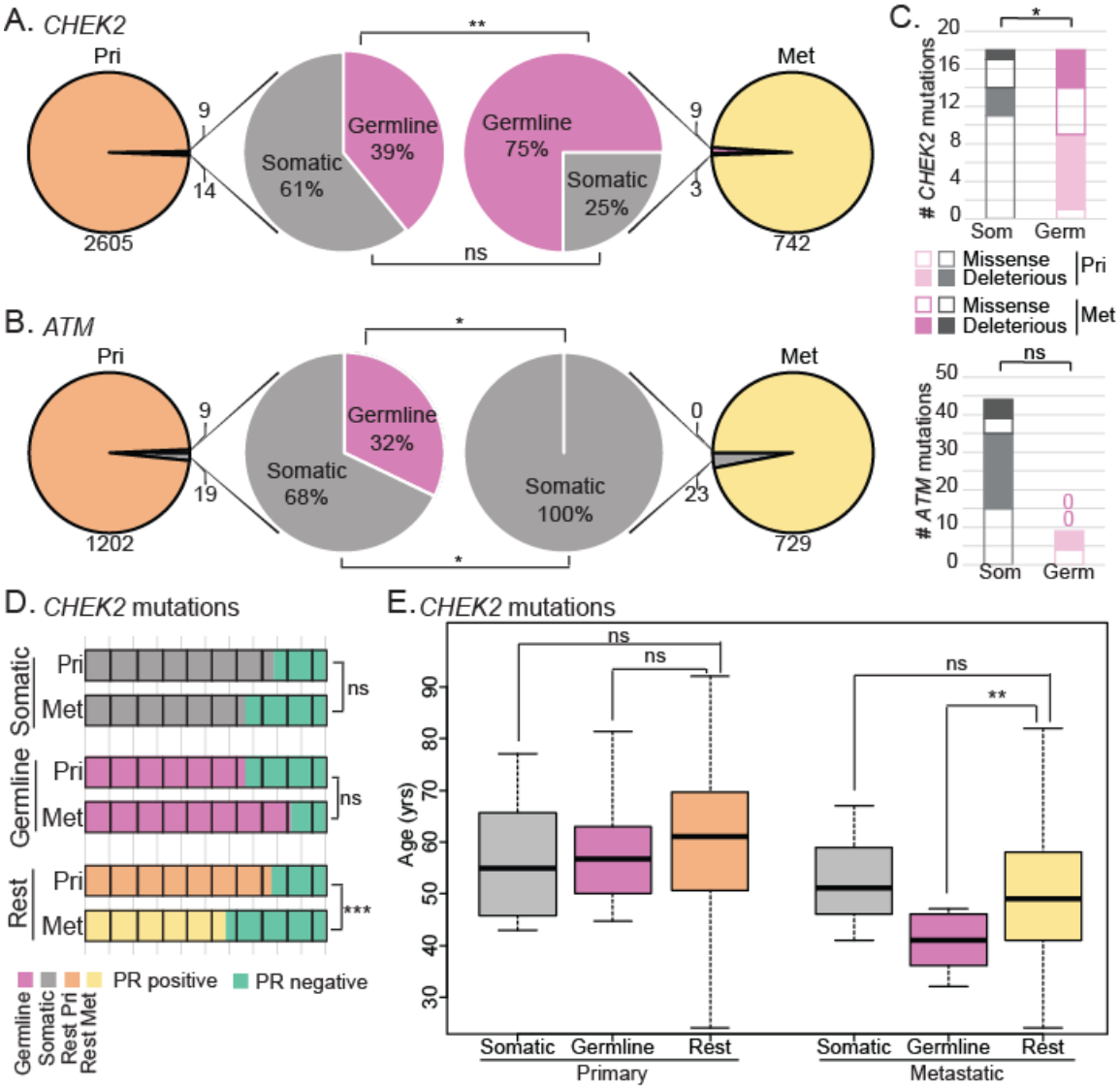

Figure 5. Germline, rather than somatic, mutations in *CHEK2* drive the association between *CHEK2* mutations and younger age at diagnosis of highly PR positive, metastatic ER+/HER2- breast cancer patients. (A-B) Pie charts demonstrating the percentage of *CHEK2* (A) and *ATM* (B) mutations that are somatic vs germline in primary (Pri) and metastatic (Met) ER+/HER2- breast cancer patients. Fisher's Exact test. (C) Stacked columns demonstrating the percentage of *CHEK2* and *ATM* mutations that are missense and deleterious (frameshift, nonsense or splice site) based on their somatic (Som) or germline (Germ) status in primary and metastatic ER+/HER2- breast cancer samples. Fisher's Exact test. (D) Waffle chart depicting percent ER+/HER2- primary and metastatic breast cancer samples with indicated mutations that are PR+ vs PR-. Fisher's Exact test. (E) Box plot indicating median age at diagnosis for indicated groups of breast cancer patients. Error bars indicate standard deviation. Two-tailed independent sample Student's T-test. Rest delineates patients with no *CHEK2* mutations. ns, not significant; $p \leq 0.05$*, $p \leq 0.01$**, $p \leq 0.001$***. Supporting data presented in **Supplementary Figure 4**.

The sample size of primary tumors with *CHEK2* mutations was too small to conduct similar analyses. However, as independent validation, we observed a statistically significant association between the incidence of germline *CHEK2* mutations and worse relapse-free survival in primary ER+/HER2- breast cancer patients in our second dataset (HR=6.15, p=0.01, **Figure 6B**, METABRIC). Of note, this association is specific to patients who were administered endocrine therapy. We did not observe a significant association between *CHEK2* mutations and poor outcome in primary ER+/HER2- breast cancer patients who were not given endocrine therapy (**Figure S5B**). In the samples from MSKCC, we observed no significant differences in deleteriousness of somatic vs germline *CHEK2* mutations in metastatic breast cancer patients. However, in METABRIC, we observed significant enrichment for deleterious mutations in the germline cohort of *CHEK2* mutated patients over the somatic (6/6 germline mutations were deleterious, i.e., nonsense, splice site or frameshift compared to 3/12 somatic mutations, p=0.009 by Fisher's Exact test). Therefore, the apparent specificity of the association of somatic vs germline *CHEK2* mutations with outcome in METABRIC could relate to the difference in the deleteriousness of mutations.

Since we observed an enrichment for co-incident mutations in *ATR* and *TP53* in ER+/HER2- breast cancer samples, we next tested whether this co-incidence associates with worse outcome in patients using MSKCC and METABRIC datasets. In MSKCC, we observed a trend towards better outcome in ER+/HER2- metastatic breast cancer patients with mutations in *ATR* alone (HR=0.59, p=0.08), and a trend towards worse outcome in patients with mutations in *TP53* alone (HR=1.17, p=0.08) but only co-incident mutation of *ATR* and *TP53* in primary ER+/HER2- breast cancer associated with significantly worse metastasis-free survival (HR=2.01, p=0.006, **Figure 6C**). Co-incident ATR/TP53 mutation also associated with significantly worse metastasis-free survival in ER+/HER2- breast cancer patients than mutations in TP53 alone (HR=1.67, p=0.03). A similar association was observed between co-incident mutation of *ATR* and *TP53* and recurrence-free survival in primary ER+/HER2- breast cancer patients in METABRIC (HR=2.97, p=0.02, **Figure 6D**). In both datasets, the association between co-incident mutations in ATR and TP53 and worse outcome remained significant in a multivariate Cox Proportional Hazards assessment which included age at diagnosis, PR status and tumor stage (HR=2.00, p=0.007 in MSKCC and HR=3.10, p=0.01 in METABRIC).

To test whether the worse outcome associated with co-incident mutations in *ATR* and *TP53* is universal or specific to ER+/HER2- breast cancer, we also assessed survival outcomes in ER- breast cancer patients in both METABRIC (**Figure 6E**) and MSKCC (**Figure S5C**) datasets. In both datasets, we observed that co-incident mutation of *ATR* and *TP53* associates with significantly worse recurrence-free survival irrespective of hormone receptor status (METABRIC: HR=1.89, p=0.004, MSKCC: HR=4.82, p=0.03). Overall, these data provide evidence of a dichotomy in the impact of CHK2 vs ATR inactivation on the type of breast cancer a patient develops and on clinical outcomes, i.e., response to initial endocrine therapies in ER+/HER2- breast cancer patients and metastatic potential in both ER+/HER2- and ER- breast cancer (**Figure 6F**).

Discussion

While a role for cell cycle checkpoint kinase inactivation in the process of tumorigenesis is well studied, we present here the first systematic analysis of the impact of mutations in individual cell cycle checkpoint kinase genes on tumor characteristics and patient outcome

(**Figure 6F**).

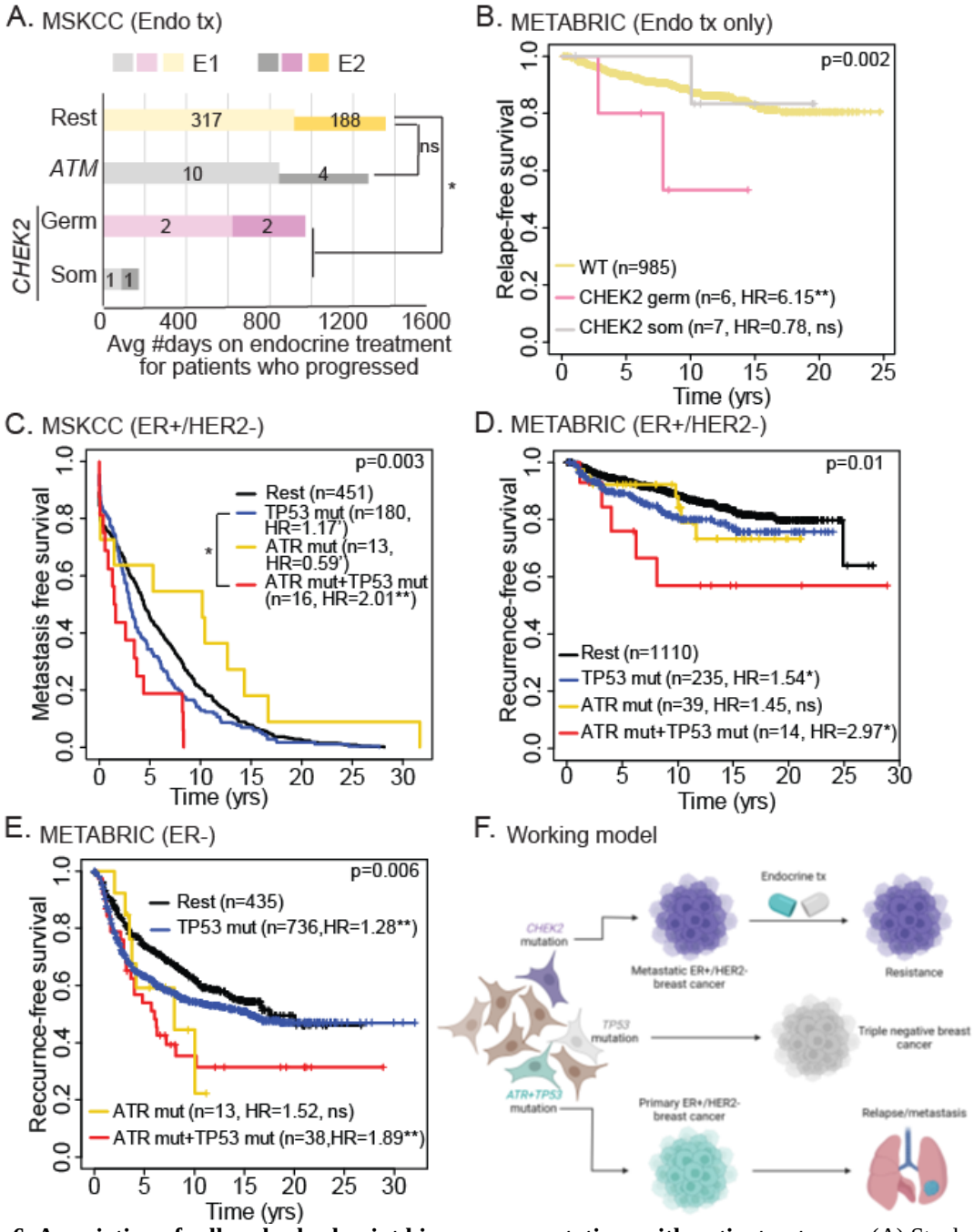

Figure 6. Association of cell cycle checkpoint kinase gene mutations with patient outcome. (A) Stacked column graph depicting the length of time for which metastatic ER+/HER2- breast cancer patients stayed on their first (E1) and second (E2) endocrine treatments before progressing categorized based on the presence of somatic (Som) or germline (Germ) mutations in *ATM* or *CHEK2*. Numbers indicate the number of patients who were represented for each endocrine intervention. Fisher's Exact test. (B-E) Kaplan-Meier survival curves measuring the specified outcomes in patients whose cancer has mutations in specified cell cycle checkpoint kinase genes. Log rank test. HR, hazard ratios. (F) Working model depicting the impact of mutations in different cell cycle checkpoint kinase genes on the type of breast cancer a patient may develop and the clinical consequences in terms of treatment response and metastatic recurrence. ns, not significant; p≤0.05*, p≤0.01**, p≤0.001***. Associated data presented in **Supplementary Figure 5**.

Most previous studies of mutations in these genes have evaluated the impact of germline variants on breast cancer risk[20,23,46–48] but have not assessed the impact of these mutations, germline or somatic, on disease progression or treatment response. It is becoming increasingly clear that combined analysis of germline and somatic mutations has the potential to improve personalized medicine by tailoring therapeutic strategies more closely to each patient[35,49,50]. Our analysis of the impact of germline and somatic mutations in *ATM*, *CHEK2* and *ATR* on tumor characteristics sheds new insight into this avenue of research.

A few previous studies have analyzed associations between the incidence of *ATM* variants and the frequency of contralateral breast cancer incidence after radiation therapy[51,52]. These studies suggest that some mutations in *ATM* may associate with a decreased risk of contralateral breast cancer recurrence[22]. Conversely, several studies have also identified that mutations in *ATM* may occur with higher frequency in metastatic breast cancer patients[53]. In this analysis of patient samples from 6 different breast cancer datasets, we did not find a statistically significant association between mutations in *ATM*, whether missense or deleterious, and the incidence of metastatic ER+/HER2- breast cancer. Rather we found enrichment for these mutations in primary ER+/HER2- tumors. Further investigation into whether *ATM* mutations impact endocrine therapy responsiveness or metastatic potential of primary ER+ breast cancer and to identify other risk factors that might modulate this effect is warranted.

As with *ATM*, germline variants in *CHEK2* have been associated with the incidence of ER+/HER2- breast cancer linked to worse outcome[54,55]. Our data support these previous findings. In addition, we demonstrate a further enrichment for the diagnosis of metastatic ER+/HER2- breast cancer at a younger age in women with germline variants in *CHEK2*. Our identification of increased PR positivity in *CHEK2* mutant metastatic ER+/HER2- breast cancer samples supports previous observations that loss of Chk2 activation can potentially delink ER signaling from cell cycle modulation, resulting in highly ER+ tumors that are resistant to endocrine therapy[15,17,18]. Our results here further support this finding by demonstrating that deleterious *CHEK2* mutations, whether germline or somatic, associate with intrinsic resistance to endocrine therapies. An understanding of which endocrine therapies may be effective in patients with *CHEK2* mutations, or whether combinations of chemotherapy or CDK4/6 inhibitors with endocrine therapy might be stronger therapeutic options is required.

Lastly, our study demonstrates for the first time that co-incident mutations in *ATR* and *TP53* are specifically enriched in ER+/HER2- breast cancer, both in the primary and metastatic setting. Further, we show that these co-incident mutations strongly associate with increased metastasis and relapse, even more so than *TP53* mutations alone. While ATM, and to some extent CHK2, are most active in inducing G1 cell cycle checkpoint, ATR and CHK1 closely regulate the S/G2-M and mitotic checkpoints[1]. The tumor suppressor, TP53 regulates all of these cell cycle checkpoints but is dispensable for G1 checkpoint regulation[56,57]. It is possible that the TP53-independent link between ATM/CHK2 and G1 progression vs the TP53-dependence of ATR/CHEK1 S/G2-M regulation explains the dichotomous impact of *TP53* mutations on the associations between mutations in ATR and CHK2 with poor prognosis of ER+/HER2- breast cancer patients. It is important to note that this association, although enriched in ER+/HER2- breast cancer, is not exclusive to these patients and seems to play a similar role in ER- breast cancer as well.

Overall, the systematic study of ATM and ATR signaling node mutations described here is the first to investigate the impact of germline and somatic events in the context of different

breast cancer subtypes. The results of this study suggest a demarcation of the roles of each of these cell cycle checkpoint kinases. Mutations in *ATM* are enriched in primary ER+/HER2- breast cancer samples, while mutations in *CHEK2* are enriched in metastatic disease. Conversely, *ATR* mutations are not enriched in ER+/HER- breast cancer except when co-incident with *TP53* mutations. Additionally, *CHEK2* mutations increase PR positivity in metastatic ER+ breast cancer while *ATR* and *TP53* mutations decrease PR positivity in primary ER+ disease. Finally, we find a specific association between deleterious *CHEK2* mutations and progression on first line endocrine therapies in both primary and metastatic ER+/HER2- breast cancer patients. While there is no association between mutations in either *ATM* or *ATR* alone and endocrine therapy responsiveness, we find associations between co-incident *ATR*/*TP53* mutations and increased metastatic potential for ER+ and ER- breast cancer patients. The early loss of specific cell cycle checkpoint kinases could serve as a decision point for the evolution of different subtypes of breast cancer, arguing for an alternate system of breast cancer classification based on DNA repair defects and associated cell cycle checkpoint inactivation. Development of such prognostic and predictive stratifiers could provide new strategies to match CDK inhibitors or cell cycle checkpoint activators to the individual cell cycle dependencies of each breast tumor.

Author contributions: Data collection, curation and analysis (SP, NP, AM, NF, MNB, SH). Conceptual oversight (MNB, SH). Preparing manuscript (SP, NP, SC, SH). Writing and editing of manuscript (SP, SC, NF, MNB, SH).

References

1. Abraham, R. T. Cell cycle checkpoint signaling through the ATM and ATR kinases. *Genes Dev.* **15**, 2177–2196 (2001).
2. Bartek, J. & Lukas, J. Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell* **3**, 421–429 (2003).
3. Smith, H. L., Southgate, H., Tweddle, D. A. & Curtin, N. J. DNA damage checkpoint kinases in cancer. *Expert Reviews in Molecular Medicine* **22**, (2020).
4. Sancar, A., Lindsey-Boltz, L. A., Unsal-Kaçmaz, K. & Linn, S. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem* **73**, 39–85 (2004).
5. Aliouat-Denis, C.-M. *et al.* p53-independent regulation of p21Waf1/Cip1 expression and senescence by Chk2. *Mol. Cancer Res.* **3**, 627–634 (2005).
6. de Toledo, S. M., Azzam, E. I., Dahlberg, W. K., Gooding, T. B. & Little, J. B. ATM complexes with HDM2 and promotes its rapid phosphorylation in a p53-independent manner in normal and tumor human cells exposed to ionizing radiation. *Oncogene* **19**, 6185–6193 (2000).
7. Jones, G. G., Reaper, P. M., Pettitt, A. R. & Sherrington, P. D. The ATR-p53 pathway is suppressed in noncycling normal and malignant lymphocytes. *Oncogene* **23**, 1911–1921 (2004).
8. Fagan-Solis, K. D. *et al.* A P53-Independent DNA Damage Response Suppresses Oncogenic Proliferation and Genome Instability. *Cell Rep* **30**, 1385-1399.e7 (2020).
9. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study | Breast Cancer | JAMA Oncology | JAMA Network. https://jamanetwork.com/journals/jamaoncology/fullarticle/2752381.
10. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
11. Lei, J. T., Anurag, M., Haricharan, S., Gou, X. & Ellis, M. J. Endocrine therapy resistance: new insights. *Breast* **48 Suppl 1**, S26–S30 (2019).
12. Huang, B., Warner, M. & Gustafsson, J.-Å. Estrogen receptors in breast carcinogenesis and endocrine therapy. *Mol. Cell. Endocrinol.* (2014) doi:10.1016/j.mce.2014.11.015.
13. Hanker, A. B., Sudhan, D. R. & Arteaga, C. L. Overcoming Endocrine Resistance in Breast Cancer. *Cancer Cell* **37**, 496–513 (2020).
14. Razavi, P. *et al.* The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. *Cancer Cell* **34**, 427-438.e6 (2018).
15. Anurag, M., Haricharan, S. & Ellis, M. J. CDK4/6 Inhibitor Biomarker Research: Are We Barking Up the Wrong Tree? *Clin. Cancer Res.* **26**, 3–5 (2020).
16. Musgrove, E. A., Hui, R., Sweeney, K. J., Watts, C. K. & Sutherland, R. L. Cyclins and breast cancer. *J Mammary Gland Biol Neoplasia* **1**, 153–162 (1996).
17. Haricharan, S. *et al.* Loss of MutL Disrupts Chk2-dependent Cell Cycle Control Through CDK4/6 to Promote Intrinsic Endocrine Therapy Resistance in Primary Breast Cancer. *Cancer Discov* (2017) doi:10.1158/2159-8290.CD-16-1179.
18. Anurag, M. *et al.* Comprehensive Profiling of DNA Repair Defects in Breast Cancer Identifies a Novel Class of Endocrine Therapy Resistance Drivers. *Clin. Cancer Res.* **24**, 4887–4899 (2018).

19. Mazumder, A. *et al.* The DNA damage repair landscape in Black women with breast cancer. *Ther Adv Med Oncol* **14**, 17588359221075458 (2022).
20. Southey, M. C. *et al.* PALB2, CHEK2 and ATM rare variants and cancer risk: data from COGS. *J Med Genet* **53**, 800–811 (2016).
21. Decker, B. *et al.* Rare, protein-truncating variants in ATM, CHEK2 and PALB2, but not XRCC2, are associated with increased breast cancer risks. *J. Med. Genet.* (2017) doi:10.1136/jmedgenet-2017-104588.
22. Stucci, L. S. *et al.* The ATM Gene in Breast Cancer: Its Relevance in Clinical Practice. *Genes* **12**, 727 (2021).
23. Tung, N. *et al.* Frequency of Germline Mutations in 25 Cancer Susceptibility Genes in a Sequential Series of Patients With Breast Cancer. *J Clin Oncol* **34**, 1460–1468 (2016).
24. Choi, M., Kipps, T. & Kurzrock, R. ATM Mutations in Cancer: Therapeutic Implications. *Mol Cancer Ther* **15**, 1781–1791 (2016).
25. Abdel-Fatah, T. M. A. *et al.* Clinicopathological Significance of ATM-Chk2 Expression in Sporadic Breast Cancers: a Comprehensive Analysis in Large Cohorts. *Neoplasia* **16**, 982–991 (2014).
26. Krug, K. *et al.* Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. *Cell* **183**, 1436-1456.e31 (2020).
27. Day, F. R. *et al.* Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. *Nat. Genet.* **49**, 834–841 (2017).
28. Mandelker, D. *et al.* The Landscape of Somatic Genetic Alterations in Breast Cancers from CHEK2 Germline Mutation Carriers. *JNCI Cancer Spectr* **3**, pkz027 (2019).
29. Cortes-Urrea, C. *et al.* Exomes of Ductal Luminal Breast Cancer Patients from Southwest Colombia: Gene Mutational Profile and Related Expression Alterations. *Biomolecules* **10**, E698 (2020).
30. da Costa E Silva Carvalho, S. *et al.* Germline variants in DNA repair genes associated with hereditary breast and ovarian cancer syndrome: analysis of a 21 gene panel in the Brazilian population. *BMC Med Genomics* **13**, 21 (2020).
31. Durocher, F. *et al.* Mutation analysis and characterization of ATR sequence variants in breast cancer cases from high-risk French Canadian breast/ovarian cancer families. *BMC Cancer* **6**, 230 (2006).
32. Lee, K. J. *et al.* Exploiting DNA repair defects in triple negative breast cancer to improve cell killing. *Ther Adv Med Oncol* **12**, 1758835920958354 (2020).
33. Ma, C. X. *et al.* Targeting Chk1 in p53-deficient triple-negative breast cancer is therapeutically beneficial in human-in-mouse tumor models. *J Clin Invest* **122**, 1541–1552 (2012).
34. Wilson, Z. *et al.* ATR inhibitor AZD6738 (ceralasertib) exerts antitumor activity as a monotherapy and in combination with chemotherapy and the PARP inhibitor olaparib. *Cancer Res* canres.2997.2021 (2022) doi:10.1158/0008-5472.CAN-21-2997.
35. Stadler, Z. K. *et al.* Therapeutic Implications of Germline Testing in Patients With Advanced Cancers. *JCO* **39**, 2698–2709 (2021).
36. Network, T. C. G. A. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
37. Cerami, E. *et al.* The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery* **2**, 401–404 (2012).

38. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
39. Miller, C. A. *et al.* Aromatase inhibition remodels the clonal architecture of estrogen-receptor-positive breast cancers. *Nature Communications* **7**, 12498 (2016).
40. Banerji, S. *et al.* Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* **486**, 405–409 (2012).
41. The Metastatic Breast Cancer Project. https://mbcproject.org/data-release.
42. Shah, S. P. *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395–399 (2012).
43. Jeselsohn, R., Buchwalter, G., De Angelis, C., Brown, M. & Schiff, R. ESR1 mutations—a mechanism for acquired endocrine resistance in breast cancer. *Nat Rev Clin Oncol* **12**, 573–583 (2015).
44. Robinson, D. R. *et al.* Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat. Genet.* **45**, 1446–1451 (2013).
45. Buendia-Buendia, J. *et al.* Characterization of mutational processes in ER+ metastatic breast cancer. *JCO* **37**, 1019–1019 (2019).
46. Couch, F. J. *et al.* Associations Between Cancer Predisposition Testing Panel Genes and Breast Cancer. *JAMA Oncol* **3**, 1190–1196 (2017).
47. Hu, C. *et al.* A Population-Based Study of Genes Previously Implicated in Breast Cancer. *New England Journal of Medicine* **0**, null (2021).
48. Gao, C. *et al.* Risk of Breast Cancer Among Carriers of Pathogenic Variants in Breast Cancer Predisposition Genes Varies by Polygenic Risk Score. *J Clin Oncol* **39**, 2564–2573 (2021).
49. Akcakanat, A. *et al.* Genomic, Transcriptomic, and Proteomic Profiling of Metastatic Breast Cancer. *Clin Cancer Res* **27**, 3243–3252 (2021).
50. D'Oronzo, S. *et al.* Dissection of major cancer gene variants in subsets of circulating tumor cells in advanced breast cancer. *Sci Rep* **9**, 17276 (2019).
51. Reiner, A. S. *et al.* Radiation Treatment, ATM, BRCA1/2, and CHEK2*1100delC Pathogenic Variants and Risk of Contralateral Breast Cancer. *J Natl Cancer Inst* **112**, 1275–1279 (2020).
52. Concannon, P. *et al.* Variants in the ATM gene associated with a reduced risk of contralateral breast cancer. *Cancer Res* **68**, 6486–6491 (2008).
53. Davis, A. A. *et al.* Association of a novel circulating tumor DNA next-generating sequencing platform with circulating tumor cells (CTCs) and CTC clusters in metastatic breast cancer. *Breast Cancer Res* **21**, 137 (2019).
54. de Bock, G. H. *et al.* Tumour characteristics and prognosis of breast cancer patients carrying the germline CHEK2*1100delC variant. *J Med Genet* **41**, 731–735 (2004).
55. Toss, A. *et al.* Clinicopathologic Profile of Breast Cancer in Germline ATM and CHEK2 Mutation Carriers. *Genes (Basel)* **12**, 616 (2021).
56. Morgan, S. E. & Kastan, M. B. p53 and ATM: cell cycle, cell death, and cancer. *Adv Cancer Res* **71**, 1–25 (1997).
57. Luo, Y., Lin, F.-T. & Lin, W.-C. ATM-Mediated Stabilization of hMutL DNA Mismatch Repair Proteins Augments p53 Activation during DNA Damage. *Mol Cell Biol* **24**, 6430–6444 (2004).