REVIEW ARTICLE

WILEY Cancer Science

# Breast cancer: The translation of big genomic data to cancer precision medicine

Siew-Kee Low[1] (iD) | Hitoshi Zembutsu[1] | Yusuke Nakamura[2,3]

[1]Project for Development of Liquid Biopsy Diagnosis, Cancer Precision Medicine Center, Japanese Foundation for Cancer Research, Tokyo, Japan

[2]Department of Medicine, Center for Personalized Therapeutics, The University of Chicago, Chicago, IL, USA

[3]Department of Surgery, Center for Personalized Therapeutics, The University of Chicago, Chicago, IL, USA

**Correspondence**
Siew-Kee Low, Project for Development of Liquid Biopsy Diagnosis, Cancer Precision Medicine Center, Japanese Foundation for Cancer Research, Tokyo, Japan.
Email: siewkee.low@jfcr.or.jp

Cancer is a complex genetic disease that develops from the accumulation of genomic alterations in which germline variations predispose individuals to cancer and somatic alterations initiate and trigger the progression of cancer. For the past 2 decades, genomic research has advanced remarkably, evolving from single-gene to whole-genome screening by using genome-wide association study and next-generation sequencing that contributes to big genomic data. International collaborative efforts have contributed to curating these data to identify clinically significant alterations that could be used in clinical settings. Focusing on breast cancer, the present review summarizes the identification of genomic alterations with high-throughput screening as well as the use of genomic information in clinical trials that match cancer patients to therapies, which further leads to cancer precision medicine. Furthermore, cancer screening and monitoring were enhanced greatly by the use of liquid biopsies. With the growing data complexity and size, there is much anticipation in exploiting deep machine learning and artificial intelligence to curate integrative "−omics" data to refine the current medical practice to be applied in the near future.

**KEYWORDS**
breast cancer, clinical sequencing, genome-wide association study, liquid biopsy, next-generation sequencing

## 1 | INTRODUCTION

Cancer is the most common genetic disease that results from the accumulation of genetic alterations. These genetic alterations are divided into 2 major categories: germline and somatic. Germline alterations are found in the germ cell; hence, this type of alteration can be inherited from parents to offspring; somatic mutations are cellular alterations that are randomly acquired throughout the lifetime after exposure to various carcinogens or aging that damage the DNA. Both germline and somatic alterations play pivotal roles in predisposing individuals to cancer and to the initiation as well as to progression of cancer. Therefore, genetic alterations could serve as effective biomarkers for early detection, monitoring and prognosis of cancer.

Cancer precision medicine aims to provide the right dose of the right drug for the right patient at the right time, based on the genetic profiles of cancer and the individual.[1] To realize this vision, rapid advancement of bioinformatics and biotechnology that contribute to the great expansion of the "omics era" makes it possible to magnify screening from a single gene to the whole genome by using genome-wide genotyping for genome-wide association studies (GWAS) or next-generation sequencing (NGS) for cancer genome profiling studies. Equally, this also starts the requirement for a super-computer/high-performance cluster computing system and use of cloud space to accommodate the analyses and to store the rapidly accumulating big genomic data. Notably, just like finding a needle in a haystack, the greatest challenge of handling big genomic data is to curate and identify clinically significant variants that could be

implemented in clinical settings. Several open-access databases such as the GWAS catalog,[2] NCI Genomic Data Commons,[3] ClinVar[4] and ClinGen[5] have been established to allow researchers to access as well as to assist with the curation of large-scale genomic data.

The current review summarizes the accumulation and achievement of big genomic data and how these could be part of the vision of cancer precision medicine by using breast cancer as a disease model. Breast cancer is the most common malignancy among women worldwide. It is well known that breast cancer is a complex polygenic disease and besides the reported risk factors that include age, age at menarche, ethnicity, reproductive and menstrual history, oral contraceptive use, hormone therapy, radiation exposure, alcohol intake, dietary folate intake, physical activity and benign breast diseases, genetic factors play an important role in disease etiology and pathogenesis.[6]

## 2 | GERMLINE VARIATION: TRANSITION FROM THE CANDIDATE GENE APPROACH TO GWAS AND NGS

Identification of germline mutations began 3 decades ago by examining rare but highly penetrant mutations that inherit with cancer in large families that showed Mendelian modes of inheritance. These hereditary cancers account for approximately 5%-10% of all cancer.[7] The majority of these mutations are inherited in an autosomal dominant manner. Genetic linkage analysis has successfully localized *BRCA1* and *BRCA2* as highly penetrant cancer susceptibility genes for hereditary breast-ovarian cancers.[8,9] Individuals who possess mutations in these highly penetrant genes have a significantly higher risk of developing cancer than those in the general population. Nevertheless, mutations in high-penetrance genes explained only a fraction of the heritability of human cancers.[10] Even though the candidate gene approach, which focuses on the DNA damage response pathway and cancer-related genes, the mutations in *CHEK2*, *PALB2*, *PTEN* and *ATM*[11-14] confer moderate effects on breast cancer, indicating the necessity to uncover more genetic alterations that are associated with this complex cancer.

### 2.1 | Common Variants-Common Disease: Emergence of Gwas and NGS

Approximately 99.9% of the DNA sequence is identical in the 3.2 billion base pairs of the human genome across different individuals, the remaining 0.1% consists of mostly common variants showing significant inter-individual variability. These common variants are mostly represented by a single nucleotide polymorphism (SNP) that occurs in every 300-1000 nucleotides, and the human genome comprises approximately 10 million SNP. The common disease-common variants hypothesis presumed inheritance by the cumulative number of alleles that conferred a modest increase in disease risk (relative risk ranged from 1.1 to 1.5). Besides explaining inter-individual external phenotypic features such as eye color, height, size of the head and many others, these common variants could be used as predictive

markers for disease susceptibility, drug response and drug-induced toxicity.

To better understand genetic inheritance, linkage disequilibrium (LD), which is nonrandom association of alleles at nearby loci, was intensely studied. In the year 2001, the International HapMap Project was started to characterize the LD patterns of individuals from 4 major continents that included Caucasians, Africans, Chinese and Japanese.[15] The database that was established from the HapMap project enabled the selection of representative SNP (tagSNP) for LD blocks, which lead to the possibility of carrying out GWAS by evaluating a reduced number of tagSNP that represented the whole genome. To uncover all SNP for the whole genome, the untyped SNP could be inferred by referring typed SNP of GWAS to the whole genome reference sequence from the 1000 genomes reference database by using genotype imputation analysis.[16] Figure 1 summarizes the workflow of carrying out GWAS that includes phenotype selection, genome-wide genotyping, data quality control and visualization as well as subsequent post-GWAS analyses. GWAS arise as the initial stage of accumulating big data when study groups have started to increase the sample size up to hundreds of thousands through collaborative studies.

As the effect size of a variant varies among populations, GWAS are mostly carried out in a population-specific way, with the majority of GWAS reported from European descendants followed by East Asians, Africans and Latin Americans. Nevertheless, a great number of meta-analyses were carried out through international consortium networks with the purpose of identifying shared genetic susceptibilities among different populations for various complex diseases. For breast cancer, the Breast Cancer Association Consortium (BCAC) and Asia Breast Cancer Consortium were started to assess the associations of common genetic variations with breast cancer.

Breast cancer was the very first GWAS that was published among other cancer GWAS. A total of 28 studies reported and identified 70 loci for transethnic populations, 70 for Europeans, 8 for East Asians, 3 for Africans, 2 for Latinos and 1 for Ashkenazi Jews that were reported to be associated with susceptibility for breast cancer. Table S1 (with references) summarizes the genetic loci that are associated with breast cancer from different populations and meta-analyses with $P$-value $\leq 5.0 \times 10^{-8}$. Among the reported loci, 2 most consistently associated with breast cancer in various populations are fibroblast growth factor receptor 2 (*FGFR2*) on chromosome 10q26 (OR = 1.35, 95% CI = 1.31-1.40)[17] and TOX high mobility group box family member 3 (*TOX3*)-cancer susceptibility 16 (*CASC16*) on chromosome 16q12 (OR = 1.31, 95% CI = 1.22-1.41).[18] Association of *FGFR2* contributes up to 16% of all breast cancers indicating a significant disease burden. FGFR2 is overexpressed in 5%-10% of breast tumors.[19] System biology approach suggests a link between *FGFR2* germline variants could reduce a cell's ability to respond to estrogen activation.[20] In contrast, TOX3 expression is highly up-regulated in luminal breast cancer compared to normal breast tissues or basal-like tumors.[21] SNP rs4784227 alters the expression of TOX3 by disrupting the enhancer function through
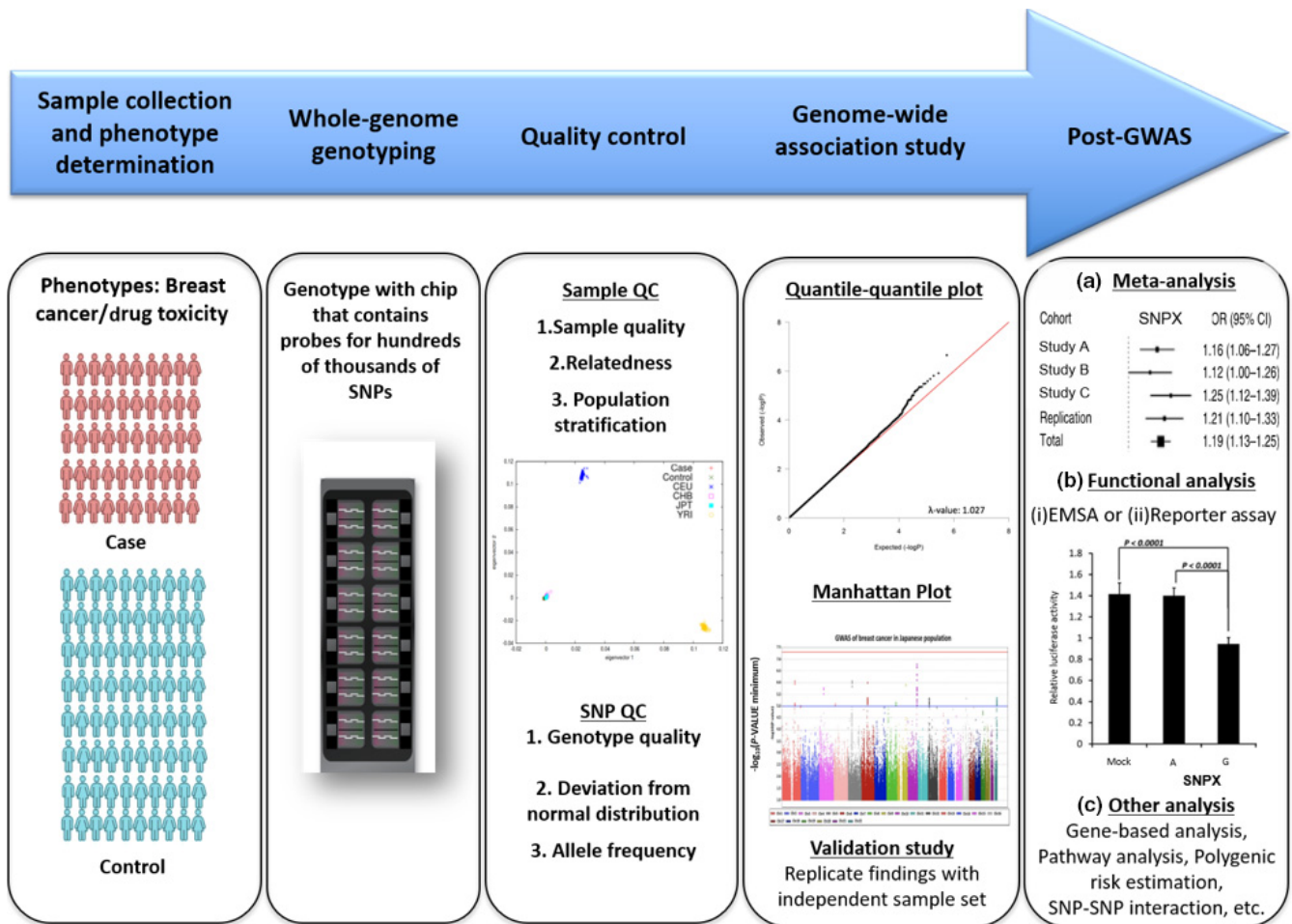
**FIGURE 1** Summary workflow of genome-wide association studies (GWAS). GWAS starts from the determination of phenotypes. Genomic DNA extracted from samples was genotyped with chips that contained up to hundreds of thousands of single nucleotide polymorphisms (SNP). Quality control (QC) was carried out on samples and SNP before association studies. Sample quality control includes: (1) sample quality to exclude poorly genotyped samples; (2) identity-by-state analysis to exclude close relatedness samples; and (3) principal component analysis to evaluate population stratification of the sample sets to obtain a homogeneous sample set before carrying out the association study. SNP QC were set to exclude SNP if: (1) they were of low genotype quality; (2) if SNP deviated from normal distribution by evaluating the Hardy-Weinberg equilibrium in control samples; and (3) if they contained non-polymorphic SNP (minor allele frequency = 0). To evaluate the association distribution, quantile-quantile plots (Q-Q plot) of observed $P$-value vs expected $P$-value and genomic inflation factor ($\lambda$ value) were evaluated to eliminate the possibility of population substructure. Manhattan plots of $P$-value ($-\log_{10}$) vs chromosome loci were used to depict an overview of the GWAS, with each dot representing a SNP and each color representing a chromosome. Post-GWAS included: (A) a meta-analysis that combined multiple studies to identify significantly associated SNP; and (B) functional analysis. Two of the most common functional analyses of the identified variants are: (i) electrophoretic mobility shift assay (EMSA) to check the existence of proteins, mainly transcription factors, binding to SNP-contained DNA fragments; and (ii) luciferase reporter assay (comparison of relative luciferase activity) to assess the associated SNP that could affect differential gene expression (as shown in the figure). (C) Other analyses, including gene-based analysis, pathway analysis, polygenic risk estimation, SNP-SNP interaction, SNP-environment interaction etc. could be carried out after GWAS

forkhead box A1 (FOXA1) affinity modulation in which FOXA1 is central to the establishment of transcriptional programs responding to estrogen stimulation in estrogen receptor 1 (ESR1)-positive breast cancer cells.[22] Notably, various GWAS also identified the association of variants on *ESR1* that encodes estrogen receptor alpha (ER-alpha) with breast cancer risk.[23] ER-alpha is known to act as a transcriptional regulator by interacting with estrogen. Taken together, the findings from different GWAS have successfully identified novel gene involvement in various pathways that are related to breast cancer carcinogenesis.

Among the novel loci identified by breast cancer GWAS, 2 loci, 8q24.21 and 5p15.33, are of particular interest as they showed pleiotropic effects with multiple cancer susceptibility.[24,25] In locus 8q24.21, 14 independent significant associations were identified from cancer GWAS of breast, prostate, colon, ovarian, bladder, pancreatic cancers and chronic lymphocytic leukemia.[24] All of these variants are clustered within a large gene desert with 2 nearest genes, family with sequence similarity 84 member B (*FAM48B*) and MYC proto-oncogene (*MYC*). Several studies hypothesize that the risk regions possess a regulatory element to the well-known

proto-oncogene, *MYC*. Variants on this region showed differential binding to transcription factor 7-like 2 (*TCF7L2*) that physically interacts with *MYC* and Wnt-regulated transcription factor. Transcription factor 4 (*TCF4*) indicates the possibility of the enhancement of the Wnt signaling pathway that subsequently regulates *MYC*.[24] Nevertheless, the direct relationship between susceptibility SNP and MYC expression has not been elucidated. The second locus is telomerase reverse transcriptase (*TERT*)-cleft lip and palate transmembrane protein 1-like protein (*CLPTM1L*) on 5p15.33 that is found to be significantly associated with breast, lung, melanoma, prostate, pancreatic, ovarian, testicular cancers, chronic lymphocytic leukemia, glioma and glioblastoma.[25] TERT functions to maintain telomere length and integrity as well as promoting epithelial cell proliferation.[26] Telomerase is found to counteract the process of telomere shortening that is known to increase cancer risk and death.[27] Hence, associated variants that affect the expression of TERT are of significance in preventing or enhancing tumorigenesis. Thus far, there is no concrete evidence that suggests the role of *CLPTM1L* in predisposing individuals to various types of cancer. Pleiotropic effects in multiple cancers would not be able to be assessed without the accumulation of various data from several GWAS that enhance the contributions of big genome data.

Even though imputation analysis using the GWAS dataset could uncover most of the common variants for the whole genome with high accuracy, it still stands as a challenge to impute relatively rare variants. To gain information for all types of variants, NGS, also termed as massive parallel sequencing, has progressed rapidly in the last decade to allow simultaneous sequencing of up to millions of DNA fragments. NGS could generate data based on gene-panels that sequence only the number of genes of interest, whole exome sequencing (WES) that sequence only the exons, and whole genome sequencing (WGS) that cover the entire genome. The involvement of NGS in breast cancer is mostly focused on hereditary breast and ovarian cancer (HBOC) that uses cancer gene-panels because of its cost-efficiency and relatively straightforward bioinformatics pipeline. Table 1 summarizes the results from various gene-panels screened by NGS. All studies consistently found *BRCA1* and *BRCA2* confer a moderate to high risk, whereas *PALB2* and *CHEK2* confer an inconsistently low to moderate risk for HBOC. Rare mutations are also found in *TP53*, *ATM*, *RAD51C*, *MSH2*, *MSH6*, *PMS2*, *MRE11A*, *RAD50*, *NBS1*, *CDH1*, *BARD1*, *NBN*, *BRIP1*, *MUTYH*, *MLH1* and *CDKN2A*.[28-32]

Notably, Aloraifi et al[33] established a gene panel that consisted of 312 genes, which gathered the genes from hereditary, GWAS, somatic, methylated and gene candidates reported from other studies, and screened 104 subjects that were known to be *BRCA1/2* negative. As a result, the most frequently mutated genes in hereditary breast cancer are *ATM*, *RAD50*, *CHEK2*, *TP53*, *PALB2* and *MRE11A*, and among GWAS genes are *CDKN2B-AS1*, *AP3K1*, *CASP8*, *RAD51B*, *FGF10*, *NOTCH2*, *CCDC170*, *CYP1B1*, *LEPR*, *MYHFR* and *NQO1*. Taken together, a gene-panel that could screen for high penetrative mutations and common genetic alterations on a larger sample size are anticipated to identify individuals who are at risk for breast cancer.

Although individual common variants confer only moderate increased risk, the cumulative effects of these variants could be evaluated through a polygenic genetic risk score. In breast cancer, with the variants that are identified to date, women in the highest 1% of the distribution have a 3.5-fold greater risk of breast cancer compared to the population average. Such a risk prediction model could be informative as early detection and subsequent preventive measures could be carried out.[34]

## 2.2 | Pharmacogenomics Studies Identify Germline Variations for the Prediction of Drug Response and Drug-Induced Adverse Events

Pharmacogenomics is the study of how genetic variants in genes encoding drug metabolism and drug transporters that affect drug availability at the target site (drug pharmacokinetics) as well as drug target proteins, such as receptors, enzymes, and intracellular signaling proteins, affect a patient's sensitivity to a drug (drug pharmacodynamics). Hence, pharmacogenomics studies play an important role in cancer precision medicine, which provides the right medication to a patient with a good response and a low incidence of adverse drug reaction. In recent years, the US Food and Drug Administration has started to revise drug labels based on various pharmacogenomics studies; for example, information that includes the recommendation to carry out *TPMT* and *UGT1A1* genotyping before the use of 6-mercaptopurine and camptothecin, respectively, to predict the occurrence of severe adverse events before treatment.[35,36]

Similar to the identification of genetic variants associated with cancer susceptibility, a pharmacogenomics study for breast cancer

**TABLE 1** Various gene panels that are used for hereditary breast and ovarian cancer

| Study | No. patients | No. genes | >5% | 1%-5% | <1% |
|---|---|---|---|---|---|
| Castera et al., (2014)[28] | 708 | 27 | BRCA1/2 | PALB2 | TP53, CHEK2, ATM, RAD51C, MSH2, PMS2, MRE11A, RAD50, NBS1, CDH1, BARD1 |
| Tung et al., (2015) [29] | 1781 | 25 | BRCA1/2 | CHEK2 | ATM, PALB2, BRIP1, BARD1, NBN, TP53, PMS2, MSH6, MSH2, MUTYH |
| Schroeder et al., (2015)[30] | 620 | 10 | NA | BRCA1/2 | CHEK2, ATM, CDH1, PALB2, NBN, TP53 |
| Minion et al., (2015)[31a] | 353 | 21 | NA | CHEK2 | ATM, BRIP1, NBN, PALB2, BARD1, RAD50 |
| Lincoln et al., (2015)[32] | 1105 | 29 | BRCA1/2 | NA | ATM, PALB2, CHEK2, MLH1, MSH2, MSH6, PMS2, CDKN2A, MUTYH |

[a]BRCA1/2 are not included in the gene list.

began with candidate gene approaches. For instance, the associations of variants on *CYP2D6*, the enzyme that activates tamoxifen, and *ABCC2*, the transporter that may be involved in transporting tamoxifen and/or its metabolites, significantly affect recurrence-free survival of breast cancer after tamoxifen treatment.[37] *CYP2D6* variants were found to affect Ki-67 response in breast cancer tissues after tamoxifen therapy through a multicenter prospective study.[38] Even though the candidate gene approach in pharmacogenomics studies has successfully identified genes involved in drug response or drug-induced adverse events, it could only explain a small proportion of the mechanism. Genes that are involved in the immunological pathway might play a role in drug-induced adverse events, such as the association of human leukocyte antigen (HLA) alleles with various drug-induced toxicities.[39,40] GWAS that provides a free-hypothesis approach facilitates the identification of

novel genes that are responsible for drug response or drug-induced toxicity.

The GWAS of breast cancer pharmacogenomics is mostly divided into 2 categories: drug response that influences breast cancer survival, and breast cancer therapy-induced toxicity. Table 2 summarizes the up-to-date pharmacogenomics studies by using GWAS in breast cancer.[41-51] Most of the studies could not identify variants that surpassed the genome-wide significance threshold, $5.0 \times 10^{-8}$ (multiple testing of $P = .05$ divided by a million independent test SNP), owing to the relatively low statistical power studies that contributed from the small sample size. It is always a challenge for cancer pharmacogenomics studies to recruit an adequate number of samples with uniform therapy for a specific phenotype as the incidence of drug-induced adverse events is often low, and cancer therapy often varies in drug combinations, dosing regimens and treatment duration.

**TABLE 2** Breast cancer pharmacogenomics study

| Treatment | Phenotype | SNP (*P*-value, OR/HR) | Gene | References |
|---|---|---|---|---|
| **Drug response** | | | | |
| Tamoxifen | Recurrence-free survival | rs10509373 ($P = 1.26E^{-10}$, HR = 4.53) | *C10orf11* | 41 |
| Anastrozole exemestane (aromatase inhibitor) | Breast cancer-free interval | rs13260300 ($P = 2.0E^{-07}$, HR = 1.64) | Intergenic region of chr8q21.11 | 42 |
| Endocrine therapy | Survival | rs8113308 ($P = 6.3E^{-07}$, HR = 1.69) | *ZNF613* | 43 |
| **Drug-induced adverse events** | | | | |
| Anastrozole exemestane (aromatase inhibitor) | Musculoskeletal adverse events | rs11849538 ($P = 6.67E^{-07}$, OR = 2.21) | *TCL1A* | 44 |
| Paclitaxel | Sensory neuropathy | rs7349683 ($P = 9.60E^{-07}$, HR = 1.63) | *EPHA5* | 45 |
| | | rs10771973 ($P = 2.6E^{-06}$, HR = 1.57) | *FGD4* | |
| Combinations of chemotherapy | Alopecia | rs3820706 ($P = 1.85E^{-09}$, OR = 2.38) | *CACNB4* | 46 |
| Anthracycline | Congestive heart failure | rs28714259 (Discovery: $P = 9.25E^{-06}$; OR = 2.1; Rep1: $P = .04$; OR = 1.9; and Rep2: $P = .02$; OR = 4.2) | Intergenic region of chr15q11.2 | 48 |
| Bevacizumab | Hypertension | rs6453204 ($P = 6.0E^{-08}$, OR = 3.3) | *SV2C* | 47 |
| Trastuzumab | Cardiotoxicity (decline in LVEF) | rs55756123 ($P = 9.0E^{-08}$, β = 6.11 unit decrease) | *LDB2* | 49 |
| | | rs10117876 ($P = 6.0E^{-07}$, β = 7.79 unit decrease) | *BRINP1* | |
| | | rs4305714 ($P = 1.0E^{-06}$, β = 1.87 unit decrease) | Intergenic region of chr6p22.3 | |
| | | rs707557 ($P = 6.0E^{-06}$, β = 1.46 unit decrease) | *RAB22A* | |
| | | rs77679196 ($P = 8.0E^{-06}$, β = 7 unit decrease) | *TRPC6* | |
| Lapatinib | Hepatotoxicity | HLA-DRB1*07:01 ($P = 2.0E^{-18}$, OR = NA) | *HLA-DRB1* | 50,51 |
| | | rs7828135 ($P = 4.5E^{-08}$, OR = NA) | *TPD52* | 50 |

HR, hazard ratio; LVEF, left ventricular ejection fraction; OR, odds ratio; SNP, single nucleotide polymorphism.

Importantly, dosage adjustment and change of therapy are often carried out according to the patient's physical condition.[52]

Nevertheless, GWAS has successfully identified some novel genes that might deepen our understanding of pathophysiology drug-induced toxicity. For example, Ingle et al[44] reported a variant on the T-cell leukemia 1A (TCL1A) gene to be associated with aromatase inhibitor induced musculoskeletal adverse events and suggested that the mechanism of adverse events involved cytokine receptor genes that are related to the inflammatory response. Chung et al[46] reported the variants of the calcium voltage-gated channel auxiliary subunit β 4 (CACNB4) gene to be significantly associated with drug-induced alopecia in breast cancer. Since minoxidil, a potassium channel opener, was approved by the FDA for the treatment of alopecia,[53] the finding from this GWAS suggested the involvement of ion channels in the pathogenesis of alopecia.[46]

# 3 | SOMATIC MUTATION: MOLECULAR PORTRAITS OF BREAST CANCER AND PRECISION MEDICINE

Germline alterations are mostly identified through GWAS and NGS gene-panels, whereas somatic mutations for tumors are mostly uncovered through targeted re-sequencing, WES or WGS by sequencing surgically resected tumor tissue that subsequently contributes to big genomic data. Two main collaborative efforts, the International Cancer Genome Consortium (ICGC, http://icgc.org/) and The Cancer Genome atlas (TCGA, https://cancergenome.nih.gov/) projects, and numerous independent research groups have successfully used NGS technology to characterize the landscape of various common cancer types. The mutation information that was collected from these collaborative efforts was deposited in the COSMIC database, which is an open access platform to provide mutation information for other researchers.[54] The results obtained from NGS have deepened understanding in the era of cancer biology through identification of genetic alterations that play roles in tumor initiation, development and metastasis as well as enabling the possibility of studying tumor evolution. It also showed the clues of driver vs passenger mutations, which are important criteria for treatment selection. The application of NGS could help to improve patient classification, predict prognosis, evaluate drug resistance and identify drug targets. More importantly, current advancement in cancer genomics discoveries could be translated into therapeutic advances. For instance, targeted therapy epidermal growth factor receptor (EGFR) inhibitor could be paired with genomic mutation information to distinguish responsive patients and to monitor resistance occurrence throughout the course of treatment in non-small cell lung cancer.[55]

In the context of breast cancer, comprehensive sequencing of breast tumor tissues identified frequently mutated genes that include PIK3CA, TP53, GATA3, PTEN, AKT1, CDH1, ARID1B, CASP8, BRCA1, RB1, MLL3, MAP3K1, MAP3K13, NCOR1, SMARCD1, CDKN1B, TBX3, RUNX1, CBFB, AFF2, PIK3R1, PTPN22, PTPRD, NF1, SF3B1 and CCND3 as well as copy number alterations in PIK3CA, ERBB2, TP53,

MAP2K4, MLL3, CDKN2A, PTEN and RB1.[56,57] Somatic mutations in TP53, PIK3CA and GATA3 occurred at >10% incidence across all breast cancer.[56] Currently, the classification of breast tumor is based mainly on the expression of estrogen receptor (ER), progesterone receptor (PgR) and the overexpression or amplification of oncogenic human epidermal growth factor receptor 2 (HER2). The different molecular subtypes are:

1. ER-positive group is divided into
   a. luminal A: PgR high, HER2 negative
   b. luminal B: PgR low, HER2 negative
2. HER2 type: HER2 positive (particularly aggressive)
3. Basal like: often referred to as triple negative breast cancer (TNBC): ER negative, PgR negative, HER2 negative.

Even though the classification is based mainly on the expression of hormonal receptors and HER2 amplification, these molecular subtypes of breast cancer showed different gene mutation patterns that could further characterize the different types of breast cancer.[56]

Approximately 60% of breast cancers are luminal subtype (luminal A and B) and could be treated with endocrine therapy according to the St Gallen 2015 recommendations.[58] The majority of luminal A show good prognosis and require no chemotherapy except those with high risk of relapse. Luminal B/HER2 negative tumors require both endocrine therapy and chemotherapy. Aromatase inhibitor that suppresses estrogen production is 1 of the recommended endocrine therapies for metastatic breast cancer patients with ER positivity.[59] Nonetheless, more than one-third of patients do not benefit from endocrine therapy owing to intrinsic resistance.[60] Even if therapy is shown to be effective initially, prolonged exposure causes resistance towards the therapy. One of the proposed mechanisms that causes resistance are mutations on the ESR1 gene that render the estrogen receptor from constitutive activation.[61,62] Hence, monitoring ESR1 mutations could help to assess the resistance of endocrine therapy. From the perspective of cancer genomics, luminal A subtypes are shown to have the most mutated genes, with the most frequent in PIK3CA, followed by MAP3K1, GATA3, TP53, CDH1 and MAP2K4. Notably, luminal A tumors harbored inactivating mutations in MAP3K1 and MAP2K4, which represent 2 contiguous steps in the p38-JNK1 stress kinase pathway.[56] Luminal B subtypes showed a diversity of mutated genes that include TP53 and PIK3CA being the most frequent. Significantly, the TP53 pathway remains largely intact in luminal A cancers but is often inactivated in more aggressive luminal B cancers.[56] The high frequency of PIK3CA mutations in this luminal subgroup indicates that inhibitors of this kinase and its related signaling pathway may act as potential druggable targets.[56]

Approximately 25% of breast cancer patients are HER2 positive which is associated with decreased overall survival and increased risk of metastasis.[63] HER2 protein (human epidermal growth factor receptor 2) regulates cell growth, proliferation and differentiation. The establishment of trastuzumab, a humanized monoclonal antibody against HER2-extracellular domain, and lapatinib, an intracellular tyrosine kinase inhibitor that blocks both HER2 and EGFR activation,

represent a prominent therapeutic advance for HER-2 breast cancer patients.[64,65] Findings from a cancer genomics study suggested EGFR, FGFR, CDK4 and cyclin D1 as possible druggable targets.[56]

Among the subtypes of breast cancer, TNBC is the most heterogeneous subtype and, hence, treatment of this subtype is extremely ineffective and challenging. Owing to this subgroup's lack of hormonal receptor expression and HER2 amplification as drug target, chemotherapy remains the first-line standard treatment; hence, identification of new drug targets for this group is essential. In the context of cancer genomics, unlike luminal and HER2 subgroups, basal-like breast cancer showed a high frequency of TP53 mutations (80%) and these mutations caused the loss of TP53 function in almost all basal-like breast cancer.[56] In addition, basal-like breast cancer showed similar genetic characteristics with serous ovarian cancers that included TP53, RB1 and BRCA1 loss, with MYC amplification, which strongly suggests that common therapeutic approaches, such as platinum analogues and taxanes could be considered.[56] Besides, approximately 20% of basal-like breast cancer was shown to carry BRCA1/2 mutations, which suggests these patients might benefit from poly ADP ribose polymerase (PARP) inhibitors and/or platinum compounds.[66,67] Several clinical trials were conducted to evaluate the efficacy and response rate of different PARP inhibitors in metastatic breast cancer with BRCA1/2 mutations that included basal-like breast cancer. In the proof-of-concept trial, the response rate of TNBC carrying BRCA1/2 mutations was 54% treated with 400 mg olaparib, a PARP inhibitor, and 25% treated with 100 mg olaparib.[68] A phase III study that evaluated the additional iniparib, a PARP inhibitor, with gemcitabine and carboplatin, reported that even though there was no statistically significant difference between the combination of regimen vs gemcitabine and carboplatin as first-line treatment, exploratory analysis indicated that patients in the second/third line showed improved overall survival.[69] Notably, a randomized, open-label phase III trial reported that HER2-negative metastatic breast cancer patients with germline BRCA mutation who received olaparib showed significantly longer progression-free survival (7.0 vs 4.2 months) and a 42% lower risk of disease progression or death compared with standard therapy.[70]

## 4 | GENOME-BASED MEDICINE THAT LEADS TO CANCER PRECISION MEDICINE

Undeniably, the advancement of big genomic data that evolves from single gene to whole-exome, whole-genome and whole-transcriptome sequencing has offered an unbiased approach for discovery and provided a great amount of useful information to enhance the progression of cancer precision medicine. Nevertheless, considering its practical use in clinical settings, many academic cancer centers and commercial testing laboratories have focused on a fraction of the frequently mutated genes that provide better cost-effectiveness, lower burden of data analysis and rapid turnaround time for making clinical decisions.

Gene panel-based targeted sequencing that is used in clinical cancer genomic profiling is designed to sequence genomic "hot spot" regions that are frequently mutated in human cancer genes or druggable targets by using DNA from the formalin-fixed paraffin-embedded (FFPE) tumor or frozen tissue samples. Clinical cancer genomic profiling has facilitated the establishment of a "basket" trial in which enrollment of patients is based on particular mutations regardless of tumor histology. Frampton et al[71] reported a clinical cancer genomic profiling test that included 287 cancer-related genes and, among 2221 clinical cases with various cancers, showed actionable alterations in 76% of tumors. In a large-scale study of 2000 patients with advanced cancer, 789 (39%) harbored at least 1 mutation in potentially actionable genes; of the 230 patients with PIK3CA/AKT1/PTEN/BRAF mutations, 116 (50%) received a genotype-matched drug; 40 (17%) were treated in a genotype-selected trial requiring a mutation for eligibility and 40 (17%) received a genotype-relevant drug off trial.[72] Notably, a phase I program that was conducted at The University of Texas MD Anderson Cancer Center showed that patients who received therapy that matched with the alteration (n = 143) showed a higher objective response rate (12% vs 5%; P < .0001), longer PFS (median, 3.9 vs 2.2 months; P = .001), and longer overall survival (median, 11.4 vs 8.6 months; P = .04) compared with treatment without matching (n = 236).[73] Recently, genomic alterations in druggable targets, such as EGFR, BRAF, RET, ALK, ROS1, CDK4/6, MET, FGFR, were included in trials to match patients to therapies.[74-81] Taken together, comprehensive genomic profiling by clinical sequencing could categorize cancer patients based on genetic alterations and therapy could be provided according to the genetic alteration, thus implementing cancer precision medicine. Particularly, clinical cancer genomic profiling tests are also extremely useful for cancer patients with unknown primary origin, in which the genomic profile could provide clues for therapy selection according to the genetic alteration.[82] Nonetheless, there are several challenges with this approach that include insufficient amount of DNA obtained from the FFPE tumor for genomic profiling, treating patients with mutations of unknown significance (driver vs passenger mutations), limited clinical trials in the institutions, availability of off-label approved drugs as well as social acceptability of this approach.[83]

Owing to the invasive method of obtaining biopsy samples from primary or metastatic lesions, there is growing interest in the field of blood-based biomarkers that includes circulating tumor cells (CTC), exosomes and circulating tumor DNA (ctDNA), which, together, is termed liquid biopsy. CTC that are released from the primary tumor and survive in the bloodstream have the potential to seed at secondary sites to form metastases.[84] Importantly, CTC could be used to generate CTC-derived explants (CDX) that show broad similarity with the primary tumor; hence, the response of CDX to therapeutic agents mirrors the patient's response to the same treatment.[85] The major demerit of CTC is that there are only low-input amounts of CTC from the bloodstream that could be isolated using limited capture techniques.[86] Exosomes are small membrane vesicles released from diverse cell types that transfer functional molecules such as DNA, miRNA, proteins and lipid to the recipient cells.[87,88] Lately, much research has focused on specific exosomal miRNA that reflect pathological changes in cancer and suggests that exosomes are

promising biomarkers.[89,90] Nevertheless, the isolation and purification of exosomes presently remains a challenge.[86] ctDNA are short DNA fragments that originate from apoptosis and necrosis of normal and tumor cells.[91] ctDNA can be detected in the cell-free fraction of the blood and provide a snapshot of the genetic profiles of primary and metastatic tumor site(s).[92,93] Mutation status in ctDNA is highly concordant with the corresponding tumor tissue and the level of ctDNA increase corresponds to the stage of cancer.[94] In addition, it is known that ctDNA has superior sensitivity compared to conventional biomarkers, such as CA-153, and has a greater dynamic range that correlates with changes in tumor burden.[95] Because half-time of ctDNA is quite short (about 2 hours), they could reflect the current status of both primary tumors and secondary deposits accurately and sensitively. Applications of ctDNA could be used throughout the course of cancer management that includes: (i) early cancer detection; (ii) molecular profiling (prognostication); (iii) detection of residual disease; (iv) monitoring therapeutic responses; and (v) monitoring clonal evolution of the tumor.[96] In metastatic breast cancer, *ESR1* mutations in ctDNA were used to monitor the resistance of aromatase inhibitor treatment in which patients with *ESR1* mutations detected in the plasma had a markedly shorter progression-free survival when treated with aromatase inhibitors.[97]

## 5 | CONCLUSION AND PERSPECTIVE

Cross-talk genetic screening by incorporating germline alterations to stratify individuals who are predisposed to a higher risk of cancer and somatic mutations to profile tumor characteristic for precise therapy selection can importantly aid patient care. In addition to genomics, the advancement of other "−omics" that includes transcriptomics, epigenomics, proteomics, and metabolomics have further increased the complexity of datasets that require sophisticated analytical tools. Hence, there is much anticipation about the involvement of deep machine learning and artificial intelligence to examine the integrative clinical and -omics datasets in order to inform, educate and help cancer treatment and research.

### REFERENCES

1. Peck RW. The right dose for every patient: a key step for precision medicine. *Nat Rev Drug Discov*. 2016;15:145-146.
2. MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res*. 2017;45:D896-D901.
3. Grossman RL, Heath AP, Ferretti V, et al. Toward a shared vision for cancer genomic data. *N Engl J Med*. 2016;375:1109-1112.
4. Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016;44:D862-D868.
5. Strande NT, Riggs ER, Buchanan AH, et al. Evaluating the clinical validity of gene-disease associations: an evidence-based framework developed by the clinical genome resource. *Am J Hum Genet*. 2017;100:895-906.
6. McPherson K, Steel CM, Dixon JM. ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *BMJ*. 2000;321:624-628.
7. Nagy R, Sweet K, Eng C. Highly penetrant hereditary cancer syndromes. *Oncogene*. 2004;23:6445-6470.
8. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*. 1994;266:66-71.
9. Wooster R, Bignell G, Lancaster J, et al. Identification of the breast cancer susceptibility gene BRCA2. *Nature*. 1995;378:789-792.
10. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer−analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med*. 2000;343:78-85.
11. Consortium CBCC-C. CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet*. 2004;74:1175-1182.
12. Hofstatter EW, Domchek SM, Miron A, et al. PALB2 mutations in familial breast and pancreatic cancer. *Fam Cancer*. 2011;10:225-231.
13. Liaw D, Marsh DJ, Li J, et al. Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nat Genet*. 1997;16:64-67.
14. Renwick A, Thompson D, Seal S, et al. ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. *Nat Genet*. 2006;38:873-875.
15. International HapMap C. The international HapMap project. *Nature*. 2003;426:789-796.
16. Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet*. 2010;11:499-511.
17. Zhang Y, Zeng X, Liu P, et al. Association between FGFR2 (rs2981582, rs2420946 and rs2981578) polymorphism and breast cancer susceptibility: a meta-analysis. *Oncotarget*. 2017;8:3454-3470.
18. Zhang L, Long X. Association of three SNPs in TOX3 and breast cancer risk: evidence from 97275 cases and 128686 controls. *Sci Rep*. 2015;5:12773.
19. Penault-Llorca F, Bertucci F, Adelaide J, et al. Expression of FGF and FGF receptor genes in human breast cancer. *Int J Cancer*. 1995;61:170-176.
20. Campbell TM, Castro MAA, de Santiago I, et al. FGFR2 risk SNPs confer breast cancer risk by augmenting oestrogen responsiveness. *Carcinogenesis*. 2016;37:741-750.
21. Han YJ, Zhang J, Zheng Y, Huo D, Olopade OI. Genetic and epigenetic regulation of TOX3 expression in breast cancer. *PLoS ONE*. 2016;11:e0165559.
22. Cowper-Sal lari R, Zhang X, Wright JB, et al. Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. *Nat Genet*. 2012;44:1191-1198.
23. Li T, Zhao J, Yang J, et al. A meta-analysis of the association between ESR1 genetic variants and the risk of breast cancer. *PLoS ONE*. 2016;11:e0153314.
24. Grisanzio C, Freedman ML. Chromosome 8q24-associated cancers and MYC. *Genes Cancer*. 2010;1:555-559.
25. Rafnar T, Sulem P, Stacey SN, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat Genet*. 2009;41:221-227.

26. Baird DM. Variation at the TERT locus and predisposition for cancer. *Expert Rev Mol Med*. 2010;12:e16.

27. Shammas MA. Telomeres, lifestyle, cancer, and aging. *Curr Opin Clin Nutr Metab Care*. 2011;14:28-34.

28. Castera L, Krieger S, Rousselin A, et al. Next-generation sequencing for the diagnosis of hereditary breast and ovarian cancer using genomic capture targeting multiple candidate genes. *Eur J Hum Genet*. 2014;22:1305-1313.

29. Tung N, Battelli C, Allen B, et al. Frequency of mutations in individuals with breast cancer referred for BRCA1 and BRCA2 testing using next-generation sequencing with a 25-gene panel. *Cancer*. 2015;121:25-33.

30. Schroeder C, Faust U, Sturm M, et al. HBOC multi-gene panel testing: comparison of two sequencing centers. *Breast Cancer Res Treat*. 2015;152:129-136.

31. Minion LE, Dolinsky JS, Chase DM, Dunlop CL, Chao EC, Monk BJ. Hereditary predisposition to ovarian cancer, looking beyond BRCA1/BRCA2. *Gynecol Oncol*. 2015;137:86-92.

32. Lincoln SE, Kobayashi Y, Anderson MJ, et al. A systematic comparison of traditional and multigene panel testing for hereditary breast and ovarian cancer genes in more than 1000 patients. *J Mol Diagn*. 2015;17:533-544.

33. Aloraifi F, McDevitt T, Martiniano R, et al. Detection of novel germline mutations for breast cancer in non-BRCA1/2 families. *FEBS J*. 2015;282:3424-3437.

34. Michailidou K, Lindstrom S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. 2017;551:92-94.

35. Relling MV, Hancock ML, Rivera GK, et al. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. *J Natl Cancer Inst*. 1999;91:2001-2008.

36. Iyer L, Das S, Janisch L, et al. UGT1A1*28 polymorphism as a determinant of irinotecan disposition and toxicity. *Pharmacogenomics J*. 2002;2:43-47.

37. Kiyotani K, Mushiroda T, Imamura CK, et al. Significant effect of polymorphisms in CYP2D6 and ABCC2 on clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients. *J Clin Oncol*. 2010;28:1287-1293.

38. Zembutsu H, Nakamura S, Akashi-Tanaka S, et al. Significant effect of polymorphisms in CYP2D6 on response to tamoxifen therapy for breast cancer: a prospective multicenter study. *Clin Cancer Res*. 2017;23:2019-2026.

39. Illing PT, Vivian JP, Purcell AW, Rossjohn J, McCluskey J. Human leukocyte antigen-associated drug hypersensitivity. *Curr Opin Immunol*. 2013;25:81-89.

40. Nicoletti P, Aithal GP, Bjornsson ES, et al. Association of liver injury from specific drugs, or groups of drugs, with polymorphisms in HLA and other genes in a genome-wide association study. *Gastroenterology*. 2017;152:1078-1089.

41. Kiyotani K, Mushiroda T, Tsunoda T, et al. A genome-wide association study identifies locus at 10q22 associated with clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients in Japanese. *Hum Mol Genet*. 2012;21:1665-1672.

42. Ingle JN, Xie F, Ellis MJ, et al. Genetic polymorphisms in the long noncoding RNA MIR2052HG offer a pharmacogenomic basis for the response of breast cancer patients to aromatase inhibitor therapy. *Cancer Res*. 2016;76:7012-7023.

43. Khan S, Fagerholm R, Rafiq S, et al. Polymorphism at 19q13.41 predicts breast cancer survival specifically after endocrine therapy. *Clin Cancer Res*. 2015;21:4086-4096.

44. Ingle JN, Schaid DJ, Goss PE, et al. Genome-wide associations and functional genomic studies of musculoskeletal adverse events in women receiving aromatase inhibitors. *J Clin Oncol*. 2010;28:4674-4682.

45. Baldwin RM, Owzar K, Zembutsu H, et al. A genome-wide association study identifies novel loci for paclitaxel-induced sensory peripheral neuropathy in CALGB 40101. *Clin Cancer Res*. 2012;18:5099-5109.

46. Chung S, Low SK, Zembutsu H, et al. A genome-wide association study of chemotherapy-induced alopecia in breast cancer patients. *Breast Cancer Res*. 2013;15:R81.

47. Schneider BP, Li L, Shen F, et al. Genetic variant predicts bevacizumab-induced hypertension in ECOG-5103 and ECOG-2100. *Br J Cancer*. 2014;111:1241-1248.

48. Schneider BP, Shen F, Gardner L, et al. Genome-wide association study for anthracycline-induced congestive heart failure. *Clin Cancer Res*. 2017;23:43-51.

49. Serie DJ, Crook JE, Necela BM, et al. Genome-wide association study of cardiotoxicity in the NCCTG N9831 (Alliance) adjuvant trastuzumab trial. *Pharmacogenet Genomics*. 2017;27:378-385.

50. Parham LR, Briley LP, Li L, et al. Comprehensive genome-wide evaluation of lapatinib-induced liver injury yields a single genetic signal centered on known risk allele HLA-DRB1*07:01. *Pharmacogenomics J*. 2016;16:180-185.

51. Spraggs CF, Budde LR, Briley LP, et al. HLA-DQA1*02:01 is a major risk factor for lapatinib-induced hepatotoxicity in women with advanced breast cancer. *J Clin Oncol*. 2011;29:667-673.

52. Low SK, Takahashi A, Mushiroda T, Kubo M. Genome-wide association study: a useful tool to identify common genetic variants associated with drug toxicity and efficacy in cancer pharmacogenomics. *Clin Cancer Res*. 2014;20:2541-2552.

53. Price VH. Treatment of hair loss. *N Engl J Med*. 1999;341:964-973.

54. Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. *Nucleic Acids Res*. 2017;45:D777-D783.

55. Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer*. 2007;7:169-181.

56. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490:61-70.

57. Stephens PJ, Tarpey PS, Davies H, et al. The landscape of cancer genes and mutational processes in breast cancer. *Nature*. 2012;486:400-404.

58. Coates AS, Winer EP, Goldhirsch A, et al. Tailoring therapies–improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol*. 2015;26:1533-1546.

59. Pritchard KI, Burris HA 3rd, Ito Y, et al. Safety and efficacy of everolimus with exemestane vs. exemestane alone in elderly patients with HER2-negative, hormone receptor-positive breast cancer in BOLERO-2. *Clin Breast Cancer*. 2013;13:421-432 e8.

60. Peng J, Sengupta S, Jordan VC. Potential of selective estrogen receptor modulators as treatments and preventives of breast cancer. *Anticancer Agents Med Chem*. 2009;9:481-499.

61. Toy W, Shen Y, Won H, et al. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat Genet*. 2013;45:1439-1445.

62. Robinson DR, Wu YM, Vats P, et al. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat Genet*. 2013;45:1446-1451.

63. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*. 1987;235:177-182.

64. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med*. 2005;353:1659-1672.

65. Cameron D, Casey M, Press M, et al. A phase III randomized comparison of lapatinib plus capecitabine versus capecitabine alone in women with advanced breast cancer that has progressed on trastuzumab: updated efficacy and biomarker analyses. *Breast Cancer Res Treat*. 2008;112:533-543.

66. Fong PC, Boss DS, Yap TA, et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. *N Engl J Med*. 2009;361:123-134.

67. Audeh MW, Carmichael J, Penson RT, et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and recurrent ovarian cancer: a proof-of-concept trial. *Lancet*. 2010;376:245-251.

68. Tutt A, Robson M, Garber JE, et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. *Lancet*. 2010;376:235-244.

69. O'Shaughnessy J, Schwartzberg L, Danso MA, et al. Phase III study of iniparib plus gemcitabine and carboplatin versus gemcitabine and carboplatin in patients with metastatic triple-negative breast cancer. *J Clin Oncol*. 2014;32:3840-3847.

70. Robson M, Im SA, Senkus E, et al. Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. *N Engl J Med*. 2017;377:523-533.

71. Frampton GM, Fichtenholtz A, Otto GA, et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat Biotechnol*. 2013;31:1023-1031.

72. Meric-Bernstam F, Brusco L, Shaw K, et al. Feasibility of large-scale genomic testing to facilitate enrollment onto genomically matched clinical trials. *J Clin Oncol*. 2015;33:2753-2762.

73. Tsimberidou AM, Wen S, Hong DS, et al. Personalized medicine for patients with advanced cancer in the phase I program at MD Anderson: validation and landmark analyses. *Clin Cancer Res*. 2014;20:4827-4836.

74. Melosky B. Review of EGFR TKIs in metastatic NSCLC, including ongoing trials. *Front Oncol*. 2014;4:244.

75. Luke JJ, Flaherty KT, Ribas A, Long GV. Targeted agents and immunotherapies: optimizing outcomes in melanoma. *Nat Rev Clin Oncol*. 2017;14:463-482.

76. Yoh K, Seto T, Satouchi M, et al. Vandetanib in patients with previously treated RET-rearranged advanced non-small-cell lung cancer (LURET): an open-label, multicentre phase 2 trial. *Lancet Respir Med*. 2017;5:42-50.

77. Esfahani K, Agulnik JS, Cohen V. A systemic review of resistance mechanisms and ongoing clinical trials in ALK-rearranged non-small cell lung cancer. *Front Oncol*. 2014;4:174.

78. Bubendorf L, Buttner R, Al-Dayel F, et al. Testing for ROS1 in non-small cell lung cancer: a review with recommendations. *Virchows Arch*. 2016;469:489-503.

79. Xu H, Yu S, Liu Q, et al. Recent advances of highly selective CDK4/6 inhibitors in breast cancer. *J Hematol Oncol*. 2017;10:97.

80. Sharma N, Adjei AA. In the clinic: ongoing clinical trials evaluating c-MET-inhibiting drugs. *Ther Adv Med Oncol*. 2011;3:S37-S50.

81. Babina IS, Turner NC. Advances and challenges in targeting FGFR signalling in cancer. *Nat Rev Cancer*. 2017;17:318-332.

82. Ross JS, Wang K, Gay L, et al. Comprehensive genomic profiling of carcinoma of unknown primary site: new routes to targeted therapies. *JAMA Oncol*. 2015;1:40-49.

83. Kou T, Kanai M, Matsumoto S, Okuno Y, Muto M. The possibility of clinical sequencing in the management of cancer. *Jpn J Clin Oncol*. 2016;46:399-406.

84. Pantel K, Speicher MR. The biology of circulating tumor cells. *Oncogene*. 2016;35:1216-1224.

85. Hodgkinson CL, Morrow CJ, Li Y, et al. Tumorigenicity and genetic profiling of circulating tumor cells in small-cell lung cancer. *Nat Med*. 2014;20:897-903.

86. Jia S, Zhang R, Li Z, Li J. Clinical and biological significance of circulating tumor cells, circulating tumor DNA, and exosomes as biomarkers in colorectal cancer. *Oncotarget*. 2017;8:55632-55645.

87. Raposo G, Stoorvogel W. Extracellular vesicles: exosomes, microvesicles, and friends. *J Cell Biol*. 2013;200:373-383.

88. He M, Zeng Y. Microfluidic exosome analysis toward liquid biopsy for cancer. *J Lab Autom*. 2016;21:599-608.

89. Ogata-Kawata H, Izumiya M, Kurioka D, et al. Circulating exosomal microRNAs as biomarkers of colon cancer. *PLoS ONE*. 2014;9:e92921.

90. Matsumura T, Sugimachi K, Iinuma H, et al. Exosomal microRNA in serum is a novel biomarker of recurrence in human colorectal cancer. *Br J Cancer*. 2015;113:275-281.

91. Jahr S, Hentze H, Englisch S, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res*. 2001;61:1659-1665.

92. Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer*. 2011;11:426-437.

93. De Mattos-Arruda L, Cortes J, Santarpia L, et al. Circulating tumour cells and cell-free DNA as tools for managing breast cancer. *Nat Rev Clin Oncol*. 2013;10:377-389.

94. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med*. 2014;6:224ra24.

95. Dawson SJ, Tsui DW, Murtaza M, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med*. 2013;368:1199-1209.

96. Wan JCM, Massie C, Garcia-Corbacho J, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat Rev Cancer*. 2017;17:223-238.

97. Schiavon G, Hrebien S, Garcia-Murillas I, et al. Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. *Sci Transl Med*. 2015;7:313ra182.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.