



PIK3CA and TP53 Gene Mutations in Human Breast Cancer Tumors Frequently Detected by Ion Torrent DNA Sequencing

Xusheng Bai^{1,9}, Enke Zhang^{1,9}, Hua Ye², Vijayalakshmi Nandakumar³, Zhuo Wang¹, Lihong Chen¹, Chuanning Tang², Jianhui Li¹, Huijin Li¹, Wei Zhang¹, Wei Han¹, Feng Lou², Dandan Zhang², Hong Sun², Haichao Dong², Guangchun Zhang², Zhiyuan Liu², Zhishou Dong², Baishuai Guo², He Yan², Chaowei Yan², Lu Wang², Ziyi Su², Yangyang Li², Lindsey Jones³, Xue F. Huang³, Si-Yi Chen^{3*}, Jinglong Gao^{1*}

1 Central Laboratory, People's Hospital of Shan Xi Province, Xian, China, **2** San Valley Biotechnology Incorporated, Beijing, China, **3** Norris Comprehensive Cancer Center, Department of Molecular Microbiology and Immunology, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America

Abstract

Breast cancer is the most common malignancy and the leading cause of cancer deaths in women worldwide. While specific genetic mutations have been linked to 5–10% of breast cancer cases, other environmental and epigenetic factors influence the development and progression of the cancer. Since unique mutations patterns have been observed in individual cancer samples, identification and characterization of the distinctive breast cancer molecular profile is needed to develop more effective target therapies. Until recently, identifying genetic cancer mutations via personalized DNA sequencing was impractical and expensive. The recent technological advancements in next-generation DNA sequencing, such as the semiconductor-based Ion Torrent sequencing platform, has made DNA sequencing cost and time effective with more reliable results. Using the Ion Torrent Ampliseq Cancer Panel, we sequenced 737 loci from 45 cancer-related genes to identify genetic mutations in 105 human breast cancer samples. The sequencing analysis revealed missense mutations in PIK3CA, and TP53 genes in the breast cancer samples of various histologic types. Thus, this study demonstrates the necessity of sequencing individual human cancers in order to develop personalized drugs or combination therapies to effectively target individual, breast cancer-specific mutations.

Citation: Bai X, Zhang E, Ye H, Nandakumar V, Wang Z, et al. (2014) PIK3CA and TP53 Gene Mutations in Human Breast Cancer Tumors Frequently Detected by Ion Torrent DNA Sequencing. PLoS ONE 9(6): e99306. doi:10.1371/journal.pone.0099306

Editor: Todd W. Miller, Dartmouth, United States of America

Received: August 23, 2013; **Accepted:** May 13, 2014; **Published:** June 11, 2014

Copyright: © 2014 Bai et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by the grants from National Natural Science Foundation of China, the Wu Jieping Foundation, and the National Institute of Health (R01 CA90427 & R01 AI084811 to SY Chen). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Authors Hua Ye, Chuanning Tang, Feng Lou, Dandan Zhang, Hong Sun, Haichao Dong, Guangchun Zhang, Zhiyuan Liu, Zhishou Dong, Baishuai Guo, He Yan, Chaowei Yan, Lu Wang, Ziyi Su, and Yangyang Li are employees of San Valley Biotechnology, Inc. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: siyichen@usc.edu (SYC); gaojinglong126@126.com (JG)

† These authors contributed equally to this work.

Introduction

Breast cancer is both the leading cancer and cancer-related death in women, with nearly 1.7 million new cases diagnosed and over half a million deaths reported globally in 2012 [1]. The same year, China alone accounted for nearly 190,000 cases and roughly 48,000 deaths [1]. While prevalence in the US has been decreasing since the 2000's, the breast cancer incidence has been steadily increasing in Asia since the 1980's [2]. Thanks to technological advancements and improved screening methods, more cases are being diagnosed at earlier stages, and early detection is directly correlated with an increased chance of survival [3]. Despite this, the staggering incidence indicates that further screening, therapeutics, and preventative measures are necessary to reduce the rate of breast cancer and improve the prognosis of the disease.

There are a variety of factors which contribute to the development of breast cancer, the most significant of which being gender and old age. Additional etiologic agents include race, hormones, tobacco and alcohol consumption, obesity, lack of childbearing, and a combination of environmental and genetic factors [4–6]. Genetics are estimated to be the primary causal factor in 5–10% of breast cancers, while all others develop spontaneously with an accumulation of genetic and epigenetic changes [7]. Hereditary breast–ovarian cancer syndrome is the familial tendency to develop these cancers. The best characterized of these hereditary mutations are in BRCA1 and BRCA2 genes, which can interfere with repair of DNA cross links and DNA double strand breaks. These inherited mutations pose a lifetime risk of developing breast cancer between 40% and 80%, indicating cancer is not inevitable for carriers of these mutations [8,9]. However, only 2 to 3% of breast cancers have mutations in BRCA

genes [10], and an estimated 75–80% of hereditary breast cancers involve unknown genes [11]. Additionally characterized on the breast cancer cells are three important receptors: estrogen receptor (ER), progesterone receptor (PR), and ERBB2 (Her2), and the presence of these receptors can influence prognosis and treatment [12].

Despite ongoing efforts to improve screening and treatment of breast cancer, further research is needed to determine other unknown genetic mutations which are involved in the progression of the disease. Due to the variety of complex interactions between genetic and environmental factors, each tumor potentially exhibits a unique gene mutation profile. By profiling an individual's cancer genome it becomes possible to distinguish the oncogenic mechanisms that regulate the cancer. As such, there is accumulating evidence which suggests that individualized, tailored therapies are necessary for effective treatment against cancers. Until recently, individual genome sequencing for personalized medicine was impractical due to the cost and lengthy assay times; however, new semiconductor-based sequencing called Ion Torrent sequencing is tackling many of these issues associated with other sequencing methods [13]. In this study, we have used Ion Torrent sequencing to analyze 105 clinical breast cancer samples to identify the genetic mutations in 737 loci of 45 known cancer-related genes.

Results

Breast Cancer Mutation Spectrum in Chinese Patients

We analyzed 105 breast cancer samples from Chinese patients ranging from 21–100 years of age (**Table 1**). The patients were categorized based on their age, menopausal states, receptor status (ER, PR, and Her), and AJCC/TNM cancer staging system (**Tables S1–4**).

This Personalized Cancer Mutation Panel is designed to target 737 mutations in the following 45 key cancer genes: ABL1, AKT1, ALK, APC, ATM, BRAF, CDH1, CDKN2A, CSF1R, CTNNA1, EGFR, ERBB2, ERBB4, FBXW7, FGFR1, FGFR2, FGFR3, FLT3, GNAS, HNF1A, HRAS, IDH1, JAK3, KDR, KIT, KRAS, MET, MLH1, MPL, NOTCH1, NPM1, NRAS, PDGFRA, PIK3CA, PTEN, PTPN11, RB1, RET, SMAD4, SMARCB1, SMO, SRC, STK11, TP53, and VHL. The mean read length was 76 bp, and the average sequence per sample was

approximately 24 Mb. With normalization to 300,000 reads per specimen, there was an average of 1639 reads per amplicon (range: 28 to 4732) (**Fig. S2A**), 176/189 (93.1%) amplicons averaged at least 100 reads, and 168/189 (88.9%) amplicons averaged at least 300 reads (**Fig. S2B**).

Of the 45 oncogenes and tumor suppressor genes sequenced in the 105 breast cancers, only PIK3CA (35.2%), TP53 (15.2%), and ERBB2 (1%) incurred missense mutations (**Table 2**). Immunohistochemical staining revealed different states of mutation in the ERBB2 (v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2), ER (Estrogen), and PR (Progesterone) receptors of these patient samples in addition to their above incurred mutations (**Tables S1–3**). Frequencies of these incurred mutations at different stages of breast cancer development observed in our sample set according to AJCC Cancer Staging is shown in **Table S4**. The detailed list of missense, point mutations, insertions and deletions profiled on the 737 loci of 45 tumor suppressor and oncogenes in 105 breast cancer samples is listed in the **Table S5**.

Missense Mutation Distribution in the Exons and Functional Domains of PIK3CA

PIK3CA mutations were identified in 35.2% of 105 tumors and most of these mutations were focused in exons 4 (2.6%), 9 (42.0%), and 20 (55.3%) (**Fig. 1**). These exons encode the helical and kinase domains, and mutations in these domains are associated with increased lipid kinase activity and oncogenicity [14,15].

The phosphatidylinositol 3-kinase (PI3K) pathway has been identified as an important player in cancer development and progression. Upon receptor tyrosine kinase activation, the PI3K kinase phosphorylates inositol lipids to phosphatidylinositol-3,4,5-trisphosphate. PI3K is a heterodimeric enzyme composed of a p110 α catalytic subunit encoded by the PIK3CA gene and a p85 regulatory subunit encoded by the PIK3R1 gene. Phosphatidylinositol-3,4,5-trisphosphate activates the serine/threonine kinase AKT, which in turn regulates several signaling pathways controlling cell survival, apoptosis, proliferation, motility, and adhesion [16,17].

Immunohistochemical staining revealed different states in the ERBB2, ER, and PR receptors and the frequency of PIK3CA mutations differed markedly at different states of these receptors

Table 1. Patient info for 105 female breast cancer samples.

	Subgroups of samples	No. of patients
Patients with age info	Total	105
	Age 21–40	19
	Age 41–60	55
	Age 61–80	28
	Age 81–100	3
	Unknown Age	1
Patients with AJCC/TNM info	I	24
	IIa	29
	IIb	14
	IIIa	20
	IIIC	8
	Unknown	2

doi:10.1371/journal.pone.0099306.t001

Table 2. Missense mutation frequencies (including coding silent/deletion/insertion) of 45 genes (737 loci) at different ages in 105 female breast cancer patients.

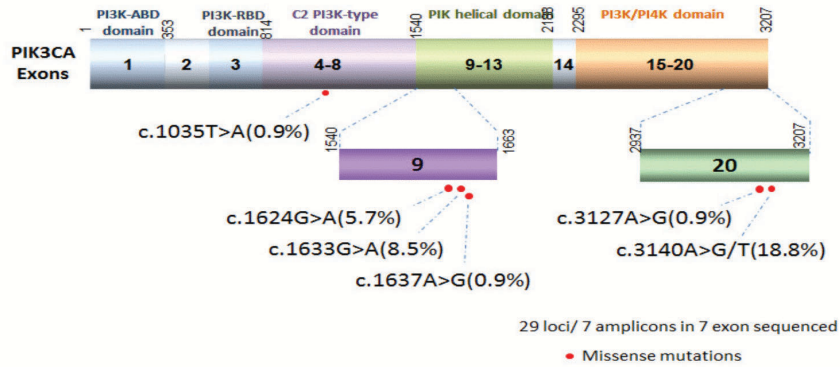
Genes	Number of Samples (Mutation Frequency)	21–40 years old (Pre-Menopausal)	41–60 years old (Pre-Menopausal)	41–60 years old (Post-Menopausal)	61–80 years old (Post-Menopausal)	81–100 years old (Post-Menopausal)
ABL1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
AKT1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
ALK	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
APC	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
ATM	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
BRAF	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
CDH1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
CDKN2A	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
CSF1R	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
CTNNB1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
EGFR	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
ERBB2	1(1.0%)	0(0.0%)	0(0.0%)	1(4.5%)	0(0.0%)	0(0.0%)
ERBB4	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
FBXW7	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
FGFR1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
FGFR2	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
FGFR3	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
FLT3	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
GNAS	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
HNF1A	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
HRAS	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
IDH1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
JAK3	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
KDR	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
KIT	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
KRAS	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
MET	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
MLH1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
MPL	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
NOTCH1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
NPM1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
NRAS	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
PDGFRA	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
PIK3CA	37(35.2%)	5(26.3%)	12(36.4%)	11(50.0%)	7(25.0%)	2(66.7%)
PTEN	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
PTPN11	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
RB1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
RET	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
SMAD4	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
SMARCB1	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
SMO	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
SRC	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
STK11	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)
TP53	16(15.2%)	2(10.5%)	6(18.2%)	4(18.2%)	4(14.3%)	0(0.0%)
VHL	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)	0(0.0%)

doi:10.1371/journal.pone.0099306.t002

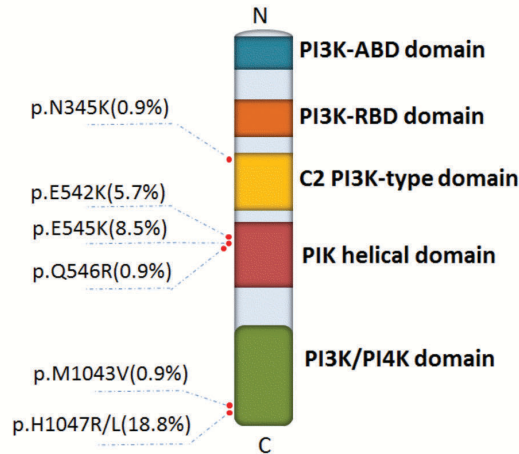
A

Gene	Exon	All mutation counts	Missense mutations(including coding silent/deletion/insertion counts (Percentage in all mutation counts of this gene)	Missense mutations(including coding silent) counts (Percentage in all mutation counts of this gene)	Synonymous mutations counts (Percentage in all mutation counts of this gene)	Deletion counts (Percentage in all mutation counts of this gene)	Insertion counts (Percentage in all mutation counts of this gene)
PIK3CA	4	1(2.6%)	1(2.6%)	1(2.6%)	0(0.0%)	0(0.0%)	0(0.0%)
PIK3CA	9	16(42.0%)	16(42.0%)	16(42.0%)	0(0.0%)	0(0.0%)	0(0.0%)
PIK3CA	20	21(55.3%)	21(55.3%)	21(55.3%)	0(0.0%)	0(0.0%)	0(0.0%)

B



C



D

		menopausal		ER		PR		HER	
		pre	post	-	+	-	+	-	+
menopausal	pre	32.7%	---	23.8%	37.9%	18.2%	42.9%	28.6%	50.0%
	post	---	37.7%	23.1%	50.0%	26.7%	50.0%	37.0%	33.3%
ER	-	23.8%	23.1%	23.4%	---	23.8%	20.0%	22.5%	0.0%
	+	37.9%	50.0%	---	43.6%	20.0%	48.9%	41.7%	57.1%
PR	-	18.2%	26.7%	23.8%	20.0%	23.1%	---	23.8%	20.0%
	+	42.9%	50.0%	20.0%	48.9%	---	46.0%	41.3%	100.0%
HER	-	28.6%	37.0%	22.5%	41.7%	23.8%	41.3%	33.0%	---
	+	50.0%	33.3%	28.6%	57.1%	20.0%	100.0%	---	42.9%

Figure 1. Missense mutation distribution in the exons and functional domains of PIK3CA. A. Frequencies of detected mutations in different exons. B. Mutation distribution in exons. C. Mutation distribution in functional domains. D. PIK3CA mutation distribution in correlation with the hormone receptor status in pre- and post-menopausal women.
 doi:10.1371/journal.pone.0099306.g001

(**Tables S1–3**). The frequencies of PIK3CA mutation occurring at different receptor states and in pre- and post-menopausal women is illustrated in **Fig. 1D**. For example, 37.7% of post-menopausal women carrying PIK3CA mutations were 50% ER+ and PR+, and 33.3% of post-menopausal women were Her+; however, PIK3CA mutation frequencies in women positive for these receptors were slightly less in pre-menopausal women who were ER+ and PR+. Not only that, women carrying PIK3CA mutations and who were ER+ were 43.9% PR+, 57.1% Her+; who were PR+ were 48.9% ER+ and 100% Her+; who were Her+ were 57.1% ER positive and 100% PR+. Also, PIK3CA mutations associate with older age at diagnosis with 66.7% of those in the age range of 81–100 years, 25.0% in the age range of 61–80 years, 36.4% in the age range of 41–60 years, and 26.3% in the age range of 21–40 years (**Table 2**).

Missense Mutation Distribution in the Exons and Functional Domains of TP53

The p53 tumor suppressor gene is located on 17p13 chromosome and spans 20 kb genomic DNA encompassing 11 exons that encodes for 53 KD phosphoprotein [18]. The phosphoprotein is a transcription factor which regulates apoptosis, genomic stability, and angiogenesis. Functional loss of p53 can lead to defective DNA replication and malignant transformation, common in the dysplasias of breast cancers [18]. The p53 gene exhibits numerous genetic alterations in patients with breast cancer [18]. This was indeed true with our sample set as well, constituting several missense mutations throughout the p53 coding region (**Fig. 2**). The incidence of p53 abnormalities varies with the degree of dysplasia and patients features [18]. This highlights the need for the administration of effective treatments such as cell-cycle inhibitors in the form of target therapies and combinatorial target therapies against the wide range of p53 mutations accumulated in that locus.

Most TP53 mutations detected in our sample set by Ion Torrent sequencing were missense (16/105, 15.2%), one of the frequently occurring mutation in TP53 [19]. The mutations were along exons 4–10, encoding the DNA-binding and oligomerization domain; specifically, the missense mutations were concentrated along the domain required for interactions with FBX042, HIPK1, AXIN1, the DNA major groove, and the domain that contains the nuclear export signal.

The frequency of TP53 mutations varied widely at different states of the hormone receptors (**S1–3**). The frequencies of TP53 mutation occurring at different receptor states and in pre- and post-menopausal women is illustrated in **Fig. 2D**. For example, 15.1% of post-menopausal women carrying TP53 mutations were 11.5% ER+, 9.1% PR+, and 33.3% Her+; however, TP53 mutation frequencies in women positive for these receptors were slightly more in pre-menopausal women. Also, TP53 mutations associate with older age at diagnosis with 53.67% in the age range of 81–100 years, 34.50% in the age range of 61–80 years, 15.8% in the age range of 41–60 years, and 34.9% in the age range of 21–40 years (**Table 2**).

Multiple Mutations and Mutation Hot Spots in Human Breast Cancers

Clinical success with individualized combination therapy relies on the identification of mutational combinations and patterns for co-administration of a single or combination of target agents against the detected mutational combinations. Some of the mutations detected in our tumor group through sequencing analysis were not only recurrent and frequent but also occurred in

combination with other mutations. Breast cancers in our sample set contained the following: 72.6% of samples had at least one or more missense mutations, 34.0% had at least two or more missense mutations, 7.5% had at least three or more missense mutations, 1.9% had at least four or more missense mutations, and 27.4% of samples incurred no deleterious mutations in any of the screened 737 loci of the potential tumor suppressor and oncogenes (**Table 3**).

Discussion

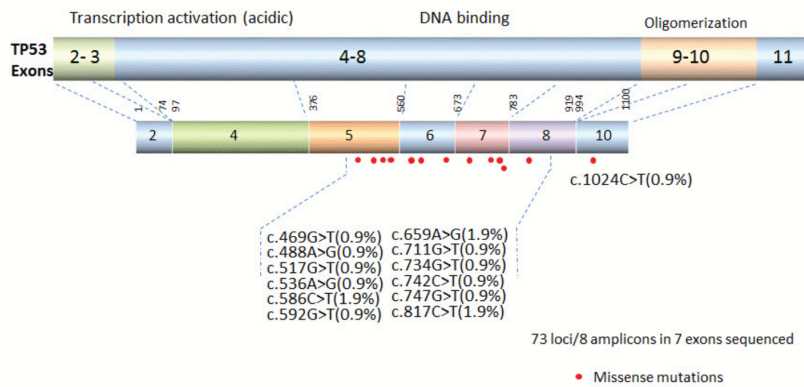
In this study we have performed a high-resolution genomic sequencing on 105 breast cancers in Chinese patients using the high throughput Ion Torrent sequencing technology. We mainly identified mutations focused along two hotspot loci, PIK3CA and TP53 in the breast cancer genomes of our sample set. In comparison with traditional Sanger sequencing and other sequence analysis methods, our analysis was at much faster rates and were of reduced sequencing costs per base [20]. The cost and complexity associated with the 4-color optical detection used in all other NGS platforms is evaded through the use of Post Light sequencing technology employed in the Ion Torrent sequencing. Despite these benefits, there is less awareness about the use and availability of this platform; however, it's starting to reach clinical investigators in recent times [21,22].

In this study, we identified PIK3CA mutations in 35.2% and TP53 mutations in 15.2% of breast tumors. Previous studies have identified PIK3CA mutation in 10.3–37.5% of the HER2-positive breast cancer cases [23–26] and p53 mutations in 18%–25% of primary breast carcinomas [27]. TP53 mutations generally have a poor prognostic power, which may be due to the screening approach used [28]. TP53 mutations are commonly detected through IHC, which detects only mutations that induce protein accumulation, missing frameshift, nonsense, and splice mutations. The Ion Torrent sequencers helped us detect robustly coding, silent mutations, insertions, and deletions more precisely. Not only that, breast cancer is a heterogeneous disease. There is a high degree of diversity between and within tumors as well as among cancer-bearing individuals, and all of these factors together determine the risk of disease progression [29]. Due to these various levels of heterogeneity, generalized treatments may be less effective. Instead targeted therapy, which involves the usage of specially designed drugs to selectively target molecular pathways correlated with the malignant phenotype of breast cancer cells, may be more useful [30]. This indicates the necessity of sequencing individual human breast cancers in order to match the use of a single targeted drug or two or more targeted drugs in combination against individual breast cancer-specific mutations. It is also critical to examine the biological features associated with each of these individual tumors to assign an appropriate treatment response. For example, trastuzumab, a humanized monoclonal antibody targeting the extracellular domain of the HER2 receptor that blocks the ligand-independent HER2 signaling, is affected by the status of PIK3CA gene mutations [26,31,32]. Similarly, prognostic studies focusing on breast cancer in the absence of p53 mutations predicts longer survival following primary therapy. However the clinical course of metastatic breast cancers and p53 mutations have not been thoroughly investigated and it remains somewhat controversial whether p53 has any significance in prediction of the clinical outcome of breast cancer [33]. Our results were somewhat consistent to the above studies in terms of the observed mutation frequencies in the PIK3CA and TP53 loci, however none of the previous reported studies compared the correlation of these mutations at different intensities and in the

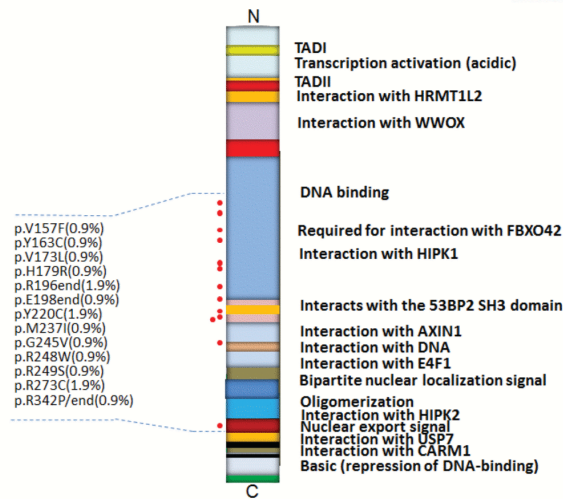
A

Gene	Exon	All mutation counts	Missense mutations(including silent/deletion/insertion counts (Percentage in all mutation counts of this gene)	Missense mutations(including coding silent) counts (Percentage in all mutation counts of this gene)	Synonymous mutations counts (Percentage in all mutation counts of this gene)	Deletion counts (Percentage in all mutation counts of this gene)	Insertion counts (Percentage in all mutation counts of this gene)
TP53	5	4(25.0%)	4(25.0%)	4(25.0%)	0(0.0%)	0(0.0%)	0(0.0%)
TP53	6	5(31.3%)	5(31.3%)	5(31.3%)	0(0.0%)	0(0.0%)	0(0.0%)
TP53	7	4(25.0%)	4(25.0%)	4(25.0%)	0(0.0%)	0(0.0%)	0(0.0%)
TP53	8	2(12.5%)	2(12.5%)	2(12.5%)	0(0.0%)	0(0.0%)	0(0.0%)
TP53	10	1(6.3%)	1(6.3%)	1(6.3%)	0(0.0%)	0(0.0%)	0(0.0%)

B



C



D

		menopausal		ER		PR		HER	
		pre	post	-	+	-	+	-	+
menopausal	pre	15.4%	---	19.0%	13.8%	18.2%	14.3%	11.9%	37.5%
	post	---	15.1%	15.4%	11.5%	16.7%	9.1%	10.9%	33.3%
ER	-	19.0%	15.4%	17.0%	---	14.3%	40.0%	15.0%	28.6%
	+	13.8%	11.5%	---	12.7%	30.0%	8.9%	8.3%	42.9%
PR	-	18.2%	16.7%	14.3%	30.0%	17.3%	---	11.9%	40.0%
	+	14.3%	9.1%	40.0%	8.9%	---	12.0%	10.9%	25.0%
HER	-	11.9%	10.9%	15.0%	8.3%	11.9%	10.9%	11.4%	---
	+	37.5%	33.3%	28.6%	42.9%	40.0%	25.0%	---	35.7%

Figure 2. Missense mutation distribution in the exons and functional domains of TP53. A. Frequencies of detected mutations in different exons. B. Mutation distribution in exons. C. Mutation distribution in functional domains. D. TP53 mutation distribution in correlation with the hormone receptor status in pre- and post-menopausal women. doi:10.1371/journal.pone.0099306.g002

Table 3. Single and multiple mutations in 105 human breast cancers.

Missense mutations (including coding silent /deletion/insertion combination)	Number of samples with mutation combination	Percentage in all sequenced samples
2	8	7.6%
1	39	37.1%
0	58	55.2%

doi:10.1371/journal.pone.0099306.t003

presence and absence of all three hormone receptors in pre- and post-menopausal Chinese patients. In this study we have evaluated and compared the relationships between TP53 and PIK3CA status with different subgroups of Chinese patients. Our current study is more of a feasibility test aiming to validate the applicability of this advanced tool in categorizing the breast cancers into different subgroups based on the identified features in their cancer genome and proteome. We further aim to use this information in a prospective clinical study to test the response of a personalized treatment regimen in different subgroups of Chinese patients.

The Ion Torrent sequencing platform helped us identify distinct mutation combinations as listed in **Table 3, Figs. 1D and 2D**, rendering potential possibilities for developing personalized combinatorial therapies. For example, depending on patients' mutated loci, their accompanied ER, PR, and Her receptor states, combined with the knowledge of patients menopausal status, personalized combinatorial therapies can be rendered as a more effective and specific treatment option over those that are currently available. Breast cancer treatment options include surgery, radiation therapy, and chemotherapy, but often depend on the stage of the disease. Targeted therapies, including anti-hormone therapies have become standard treatment for breast cancers expressing the targets of these drugs. Clinical trials are underway to evaluate targeted therapies and investigate how best to use these drugs in combination with each other and with other standard therapies.

As there is increasing information about the changes in breast cancer cells in recent times, newer drugs that specifically target these changes have been developed. These targeted drugs either work synergistically with the chemo drugs or by themselves with less toxicity due to a selective effect to a more systemic modulation of proteins associated with oncogenesis. Briefly, trastuzumab, as mentioned earlier, is a humanized monoclonal antibody targeting the extracellular domain of the HER2 receptor that blocks the ligand-independent HER2 signaling. It was initially approved by the FDA for metastatic breast cancer in 1998 [34]. Lapatinib is a dual EGFR/ErbB2 reversible tyrosine kinase inhibitor blocking both HER1 and HER2 and consequently the downstream pathways of MAPK/Erk1/2 and PI3K/Akt pathways [35]. Two different types of antihormone therapies are used to treat women with ER-positive breast tumors: selective estrogen receptor modulators and aromatase inhibitors [36]. Aromatase inhibitors are used mainly in postmenopausal women because they do not work very efficiently in premenopausal women, whose ovaries make too much aromatase [37]. Anti-hormone therapies such as tamoxifen and anastrozole can be used to treat most stages of breast cancer [38]. Women with early-stage breast cancer are usually treated with surgery followed by antihormone therapy and radiation therapy, and sometimes chemotherapy. Antihormone therapies can also be used to reduce the risk of developing breast cancer. Women at high risk for the disease are usually treated with

Tamoxifen and another selective estrogen receptor modulator, raloxifene, and several other aromatase inhibitors [39].

In this study we have used the Ion Ampliseq Cancer Panel to sequence 737 loci in 45 cancer-related genes, mainly oncogenes and tumor suppressor genes, of 105 human breast cancer samples [30]. Having gained more knowledge and experience through next generation technologies, it is necessary to expand our understanding of specific mutations to enhance individualized therapies. Therefore, gathering a complete profile of mutations in breast cancers for the application of personalized and tailored targeted therapy is critical to develop future cancer treatments. We believe a faster and more cost effective genotyping tool such as Ion Torrent sequencing technology will be greatly beneficial to assign such specific therapeutics for breast cancer patients in the near future.

Materials and Methods

Ethics Statement

The study has been approved by the Human Research Ethics Committee of the People's Hospital of Shan Xi Province, Xian, China. For Formalin fixed and paraffin embedded (FFPE) tumor samples from the tumor tissue bank at the Department of Pathology of the hospital, the institutional ethics committee waived the need for consent. All samples and medical data used in this study have been irreversibly anonymized.

Patient Information

Tumor samples used in the study were collected from the People's Hospital of Shan Xi Province, Xian, China. A total of 105 FFPE tumor samples from female breast cancer patients were analyzed. Patients were classified by age ranges as follows: 19 were between 21–40 years, 55 were 41–60 years, 28 were 61–80, 3 were 81–100, and one patient was of unknown age (**Table 1**). AJCC/TNM cancer staging is as follows: 24 patients at stage 1, 29 at stage 2a, 14 at stage 2b, 20 at stage 3a, 8 at stage 3c, and 2 of unknown stage (**Table S4**). Tumor samples were also analyzed for immunohistochemical status of estrogen receptor (ER), progesterone receptor (PR), and ERBB2 (**Tables S1–3**).

DNA Preparation

Formalin-fixed, paraffin embedded (FFPE) tissue samples were deparaffinized in xylene and 3–5 μ m thick sections were extracted. DNA was then isolated using the QIAamp DNA Mini Kit (Qiagen) following manufacturer's instructions.

Ion Torrent PGM Library Preparation and Sequencing

An Ion Torrent adapter-ligated library was constructed with the Ion AmpliSeq Library Kit 2.0 (Life Technologies, Part #4475345 Rev. A) as per manufacturer's protocol. Briefly, 50 ng of pooled

amplicons were end-repaired, and Ion Torrent adapters P1 and A were ligated with DNA ligase. Adapter-ligated products were then purified with AMPure beads (Beckman Coulter, Brea, CA, USA), nick-translated, and PCR-amplified for a total of 5 cycles. The resulting library was purified with AMPure beads (Beckman Coulter), and the concentration and size of the library was determined by Agilent 2100 BioAnalyzer and Agilent BioAnalyzer DNA High-Sensitivity LabChip (Agilent Technologies).

Sample emulsion PCR, emulsion breaking, and enrichment were performed using the Ion PGM 200 Xpress Template Kit (Life Technologies, Part #4474280 Rev. B), according to the manufacturer's instructions. Briefly, an input concentration of one DNA template copy/Ion Sphere Particles (ISPs) was added to emulsion PCR master mix and an IKADT-20 mixer (Life Technologies) was used to generate the emulsion. Next, ISPs were recovered and template-positive ISPs were enriched for use with Dynabeads MyOne Streptavidin C1 beads (Life Technologies). The Qubit 2.0 fluorometer (Life Technologies) was used to confirm ISP enrichment. 316 chips were used to sequence barcoded samples on the Ion Torrent PGM for 65 cycles, and an Ion PGM 200 Sequencing Kit (Life Technologies, Part #4474004 Rev. B) was used for sequencing reactions, as per the recommended protocol.

Variant Calling

Data from the PGM runs were processed initially using the Ion Torrent platform-specific pipeline software Torrent Suite to generate sequence reads, trim adapter sequences, filter, and remove poor signal-profile reads. Initial variant calling from the Ion AmpliSeq sequencing data was generated using Torrent Suite Software v3.2 with a plug-in "variant caller v3.2" program. In order to eliminate errors in base calling, several filtering steps were used to generate final variant calling (**Fig. S1**). The first filter was set at an average depth of total coverage of >100, an each variant coverage of >20, a variant frequency of each sample >5, and P-value < 0.01. The second filter was employed by visually examining mutations using Integrative Genomics Viewer (IGV) software (<http://www.broadinstitute.org/igv>) or Samtools software SAMtools software (<http://samtools.sourceforge.net>), as well as by filtering out possible strand-specific errors, such as a mutation detected in either "+" or "-" strand, but not in both strands of DNA. The third filtering step was set as variants within 727 hotspots, according to the manufacturer's instructions. The last filter step was eliminate variants in amplicon AMPL339432 (PIK3CA, exon13, chr3:178938822–178938906), which is not uniquely matched in human genome. From our sequencing runs using the Ion AmpliSeq Cancer Panel, false deletion data were generated from the JAK2 gene locus and thus the sequencing data from this locus were excluded from further analysis.

Bioinformatical and Experimental Validation

We used the COSMIC3 (version 64), MyCancerGenome database (<http://www.mycancergenome.org/>) and some publications to assess mutations reappearing in lung cancer (**Table S5**). Additionally, some detected missense mutations were confirmed by Sanger's sequencing (**Table S6**).

Statistical Analysis

We selected reappearing somatic missense/insertion-deletion mutations of breast cancer to perform the statistical analysis.

Supporting Information

Figure S1 Filter process of variants. (a) Strand-biased variants were eliminated using Integrative Genomics Viewer (IGV) software (<http://www.broadinstitute.org/igv>); (b) Variants in AMPL339432 should be eliminated, because this amplicon is not uniquely matched to PIK3CA in human genome; (c) All of our statistical analysis is based on the data in blue box. (DOCX)

Figure S2 Sequence read distribution across 189 amplicons generated from 105 FFPE specimens, normalized to 300,000 reads per sample. A. Distribution of average coverage of each amplicon. Data are showed as mean \pm SD. The top four amplicons, sixth amplicon, and seventh amplicon (in red box with arrow) which have high standard deviation bar target ERBB2 gene and FGFR1. B. Number of amplicons with a given read depth, sorted in bins of 100 reads. (blue bars present number of target amplicons within read depth, red line presents % of target amplicons \geq read depth). (DOCX)

Table S1 Mutations (including missense point mutations/deletion/insertion) frequencies in 45 genes (737 loci) in 105 HER2+ and HER2- breast cancer patients. The p-value of Fisher's exact of PIK3CA in HER2+ and HER2- is 0.59. The p-value of Fisher's exact of TP53 in HER2+ and HER2- is 0.07. (DOCX)

Table S2 Mutations (including missense point mutations/deletion/insertion) frequencies in 45 genes (737 loci) in 105 ER+ and ER- breast cancer patients. The p-value of Fisher's exact of PIK3CA in ER+ and ER- is 0.11. The p-value of Fisher's exact of TP53 in ER+ and ER- is 0.59. (DOCX)

Table S3 Missense mutation frequencies (including coding silent/deletion/insertion) of 45 genes (737 loci) with different immunohistochemical results of progesterone receptor (PR) in 105 female breast cancer patients. The p-value of Fisher's exact of PIK3CA in PR+ and PR- is 0.076. The p-value of Fisher's exact of TP53 in PR+ and PR- is 0.59. (DOCX)

Table S4 Missense mutation frequencies (including coding silent/deletion/insertion) of 45 genes (737 loci) of 105 female breast cancer patients according to AJCC Cancer Staging. (DOCX)

Table S5 Frequencies of missense point mutations, insertion and deletion mutations in 737 loci of 45 genes in 105 breast cancer samples. (DOCX)

Table S6 Confirmation of missense mutations by Sanger sequencing. (DOCX)

Acknowledgments

We would like to thank Rong Shi at the WuJieping Foundation, Dr. Haibo Wang, Zhi Yu, Ying Li, and other members of San Valley Biotechnology Inc. Beijing for their assistance in sample and data collection. We would also like to thank the staff at the Beijing Military Hospital for their generous support for DNA sequencing and data collection.

Author Contributions

Conceived and designed the experiments: XB EZ SYC JG. Performed the experiments: HY ZW LC CT JL HL WZ WH FL DZ HS HD GZ ZL ZD

ZD BG HY CY LW ZS YL. Analyzed the data: XB EZ CT SYC JG. Contributed reagents/materials/analysis tools: XB EZ SYC JG. Wrote the paper: XB EZ VN CT LJ XFH SYC JG. RS HW ZY YL.

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, et al. (2013) GLOBOCAN 2012 v1.0. Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. International Agency for Research on Cancer.
2. Jemal A, Center MM, DeSantis C, Ward EM (2010) Global Patterns of Cancer Incidence and Mortality Rates and Trends. *Cancer Epidemiology Biomarkers & Prevention* 19: 1893–1907.
3. Etzioni R, Urban N, Ramsey S, McIntosh M, Schwartz S, et al. (2003) The case for early detection. *Nat Rev Cancer* 3: 243–252.
4. Blackburn GL, Wang KA (2007) Dietary fat reduction and breast cancer outcome: results from the Women's Intervention Nutrition Study (WINS). *Am J Clin Nutr* 86: 878–881.
5. Chen WY, Colditz GA (2007) Risk factors and hormone-receptor status: epidemiology, risk-prediction models and treatment implications for breast cancer. *Nat Clin Pract Oncol* 4: 415–423.
6. Brody JG, Rudel RA, Michels KB, Moysich KB, Bernstein L, et al. (2007) Environmental pollutants, diet, physical activity, body size, and breast cancer. *Cancer* 109: 2627–2634.
7. Fackenthal JD, Olopade OI (2007) Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nat Rev Cancer* 7: 937–948.
8. Easton DF, Ford D, Bishop DT (1995) Breast and ovarian cancer incidence in BRCA1-mutation carriers. *Breast Cancer Linkage Consortium*. *Am J Hum Genet* 56: 265–271.
9. Levin B, Lech D, Friedenson B (2012) Evidence that BRCA1- or BRCA2-associated cancers are not inevitable. *Mol Med* 18: 1327–1337.
10. Wooster R, Weber BL (2003) Breast and Ovarian Cancer. *New England Journal of Medicine* 348: 2339–2347.
11. Easton DF (1999) How many more breast cancer predisposition genes are there? *Breast Cancer Res* 1: 14–17.
12. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, et al. (2000) Molecular portraits of human breast tumours. *Nature* 406: 747–752.
13. Hadd AG, Houghton J, Choudhary A, Sah S, Chen L, et al. (2013) Targeted, high-depth, next-generation sequencing of cancer genes in formalin-fixed, paraffin-embedded and fine-needle aspiration tumor specimens. *J Mol Diagn* 15: 234–247.
14. Bader AG, Kang S, Vogt PK (2006) Cancer-specific mutations in PIK3CA are oncogenic in vivo. *Proceedings of the National Academy of Sciences of the United States of America* 103: 1475–1479.
15. Kang S, Bader AG, Zhao L, Vogt PK (2005) Mutated PI 3-Kinases: Cancer Targets on a Silver Platter. *Cell Cycle* 4: 571–574.
16. Zhao L, Vogt PK (2008) Class I PI3K in oncogenic cellular transformation. *Oncogene* 27: 5486–5496.
17. Faber AC, Corcoran RB, Ebi H, Sequist LV, Waltman BA, et al. (2011) BIM Expression in Treatment-Naïve Cancers Predicts Responsiveness to Kinase Inhibitors. *Cancer Discovery* 1: 352–365.
18. Levine AJ (1997) p53, the Cellular Gatekeeper for Growth and Division. *Cell* 88: 323–331.
19. Langerod A, Zhao H, Borgan O, Nesland J, Bukholm I, et al. (2007) TP53 mutation status and gene expression profiles are powerful prognostic markers of breast cancer. *Breast Cancer Research* 9: R30.
20. Roukos D, Ku CS (2012) Clinical Cancer Genome and Precision Medicine. *Annals of Surgical Oncology* 19: 3646–3650.
21. Singh RR, Patel KP, Routbort MJ, Reddy NG, Barkoh BA, et al. (2013) Clinical Validation of a Next-Generation Sequencing Screen for Mutational Hotspots in 46 Cancer-Related Genes. *The Journal of Molecular Diagnostics* 15: 607–622.
22. Beadling C, Neff TL, Heinrich MC, Rhodes K, Thornton M, et al. (2013) Combining Highly Multiplexed PCR with Semiconductor-Based Sequencing for Rapid Cancer Genotyping. *The Journal of Molecular Diagnostics* 15: 171–176.
23. Pérez-Tenorio G, Alkhori L, Olsson B, Waltersson MA, Nordenskjöld B, et al. (2007) PIK3CA Mutations and PTEN Loss Correlate with Similar Prognostic Factors and Are Not Mutually Exclusive in Breast Cancer. *Clinical Cancer Research* 13: 3577–3584.
24. Berns K, Horlings HM, Hennessy BT, Madiredjo M, Hijmans EM, et al. (2007) A Functional Genetic Approach Identifies the PI3K Pathway as a Major Determinant of Trastuzumab Resistance in Breast Cancer. *Cancer cell* 12: 395–402.
25. Toi M, Iwata H, Fujiwara Y, Ito Y, Nakamura S, et al. (2009) Lapatinib monotherapy in patients with relapsed, advanced, or metastatic breast cancer: efficacy, safety, and biomarker results from Japanese patients phase II studies. *Br J Cancer* 101: 1676–1682.
26. Xu BH, Jiang ZF, Chua D, Shao ZM, Luo RC, et al. (2011) Lapatinib plus capecitabine in treating HER2-positive advanced breast cancer: efficacy, safety, and biomarker results from Chinese patients. *Chin J Cancer* 30: 327–335.
27. Alsner J, Yilmaz M, Guldberg P, Hansen LL, Overgaard J (2000) Heterogeneity in the Clinical Phenotype of TP53 Mutations in Breast Cancer Patients. *Clinical Cancer Research* 6: 3923–3931.
28. Pharoah PDP, Day NE, Caldas C (1999) Somatic mutations in the p53 gene and prognosis in breast cancer: a meta-analysis. *Br J Cancer* 80: 1968–1973.
29. Bertucci F, Birnbaum D (2008) Reasons for breast cancer heterogeneity. *Journal of Biology* 7: 6.
30. Nielsen DL, Kümler I, Palshof JAE, Andersson M (2013) Efficacy of HER2-targeted therapy in metastatic breast cancer. Monoclonal antibodies and tyrosine kinase inhibitors. *The Breast* 22: 1–12.
31. Scaltriti M, Rojo F, Ocaña A, Anido J, Guzman M, et al. (2007) Expression of p95HER2, a Truncated Form of the HER2 Receptor, and Response to Anti-HER2 Therapies in Breast Cancer. *Journal of the National Cancer Institute* 99: 628–638.
32. Xia W, Husain I, Liu L, Bacus S, Saini S, et al. (2007) Lapatinib Antitumor Activity Is Not Dependent upon Phosphatase and Tensin Homologue Deleted on Chromosome 10 in ErbB2-Overexpressing Breast Cancers. *Cancer Research* 67: 1170–1175.
33. Nakopoulou LL, Alexiadou A, Theodoropoulos GE, Lazaris AC, Tzonou A, et al. (1996) PROGNOSTIC SIGNIFICANCE OF THE CO-EXPRESSION OF p53 AND c-erbB-2 PROTEINS IN BREAST CANCER. *The Journal of Pathology* 179: 31–38.
34. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, et al. (2001) Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2. *New England Journal of Medicine* 344: 783–792.
35. Montemurro F, Valabrega G, Aglietta M (2007) Lapatinib: a dual inhibitor of EGFR and HER2 tyrosine kinase activity. *Expert Opin Biol Ther* 7: 257–268.
36. Obiorah I, Jordan VC (2011) Progress in endocrine approaches to the treatment and prevention of breast cancer. *Maturitas* 70: 315–321.
37. Fabian CJ (2007) The what, why and how of aromatase inhibitors: hormonal agents for treatment and prevention of breast cancer. *International Journal of Clinical Practice* 61: 2051–2063.
38. Lazzeroni M, DeCensi A (2013) Breast Cancer Prevention by Antihormones and Other Drugs: Where Do We Stand? *Hematology/Oncology Clinics of North America* 27: 657–672.
39. Cauley J, Norton L, Lippman M, Eckert S, Krueger K, et al. (2001) Continued Breast Cancer Risk Reduction in Postmenopausal Women Treated with Raloxifene: 4-Year Results from the MORE Trial. *Breast Cancer Research and Treatment* 65: 125–134.