Table 1. Major oncogenes and tumor suppressor genes in breast cancer.

Genes	Locus	Products	Function
Ongogenes (amplified)			
c-myc	8q24	DNA binding protein	Cell cycle control, Differentiation, Adhesion, Apoptosis
HER2 (c-erbB-2)	17q21	HER2 oncoprotein	Growth factor receptor
Cycline D1 (CCDN1)	11q13	Cell cycle regulator (G1-S phase)	Cell cycle regulator, Inactivation of PB1
Suppressor genes (mutated or loss)			
RB1	13q14	pRB	Cell cycle regulator (S phase)
TP53	17p13.1	p53	Guardian of genome
BRCA1	17q21	BRCA1	Care taker gene to maintain genomic integrity during DNA replication
BRCA2	13q12-13	BRCA2	DNA repair
E-cadherin (CDH1)	16q24	E-cadherin	Cell adhesion
Nm23 (NME1 and 2)	17q21.3	Nucleoside diphosphate kinase A, B	Catalysis of phosphorylation of nucleoside diphosphates to nucleoside triphosphates
PTEN (MMAC1)	10q23	Phosphatidylinositol phosphatase	Negative regulation of cell adhesion, Migration
p21	6q21	Inhibitor of cyclin-dependent kinase, p21	Mediation of G1 arrest
p16 (CDKN2A)	9p21	p16	Inhibition of cyclin dependent kinase (CDK)

Table 2. Hereditary breast cancer syndromes.

Hereditary cancer syndrome	Genes	Clinical association
Familial breast and ovarian cancer	*BRCA1*	Breast cancer, Ovarian cancer, Colon cancer, Prostate cancer
Familial breast cancer	*BRCA2*	Breast cancer, Prostate cancer, Male breast cancer
Li-Fraumeri syndrome	*TP53*	Multiple cancers in childhood, Soft tissue sarcoma, Brain tumors, Leukemia Adrenocortical carcinomas, Bilateral breast cancer in women
Cowden disease	*PTEN*	Breast cancer, Gastrointestinal cancers, Thyroid diseases, Mucocutaneous lesions, Trichilemmomas Acral keratosis, Papillomatous papules, Mucosal lesions, Macroencephaly Lhermitte-Duclos disease
Ataxia telangiectasia	*ATM*	Breast cancer, Telangiectasia, Leukemia, Lymphoma, Occulocutaneous telangiectasia Cerebellar ataxia, Immune deficiency
Hereditary nonpolyposis colorectal cancer	*MMR*	Gastrointestinal cancer, Basal cell carcinoma of genitourinary tract, Endometrial cancer Keratocanthomas, Colonic diverticula, Breast cancer in women

50 and 87% at age 70; for women with *BRCA2* mutations, it ranges between 40% and 80% at age 70 [3, 7]. In addition, the lifetime risk of ovarian cancer in female carriers with *BRCA1* mutations is around 30 – 40%; in *BRCA2* mutation carriers, it is approximately 20%.

As these estimates are based on studies of families at high risk of breast cancer, often including as many as 10 affected members, the study design may lead to very high estimates of penetrance. In fact, several population-based studies have demonstrated that the frequency of *BRCA1* and *BRCA2* mutations is much lower than expected from data based on these rare and large high-risk families [8-10]. However, meta-analysis has revealed that women with such genetic abnormalities and a strong family history of breast cancer are still likely to possess a much higher risk for breast cancer than are women with such abnormalities but without a strong family history [11]. In Japanese women with a modest to minimal family risk, defined as those having at least one breast cancer or ovarian cancer patient in their first degree relatives, the frequencies of *BRCA1* and *BRCA2* mutations were 13.3% and 18.6%, respectively [8]. Families at high risk for carrying *BRCA1* and *BRCA2* mutations could be predicted by family history profiles characterized by first-degree relatives with ovarian cancer or breast cancer along with young age at diagnosis (40 years old or less), bilateral occurrence and increased number of affected relatives [8].

Thus, mutations of *BRCA1* confer increased risk for breast cancer, but it is not clear why the mutations are associated with breast cancer. One study of particular interest indicated that BRCA1 could inhibit ER transcription [12], suggesting that wild-type BRCA1 suppresses estrogen-dependent transcriptional pathways related to mammary epithelial cell proliferation and that loss of this ability contributes to tumorigenesis. In agreement with these findings, bilateral prophylactic oophorectomy is associated with a reduced breast cancer risk in women who carry a *BRCA1* mutation [13].

Clinically, BRCA-associated hereditary breast cancer displays unique clinicopathological features compared with sporadic breast cancer [14-18]. For example, there is a higher frequency of medullary carcinoma in *BRCA1* mutations and lobular carcinoma in *BRCA2* mutations. Furthermore, breast cancer in patients with *BRCA1* mutations generally presents with high grade, highly proliferative, aneuploid tumors. These features have also been observed in patients with *BRCA2* mutations, although this association appears weak. *BRCA1*-related tumors are generally receptor-negative, whereas this does not appear to be the case with *BRCA2*-related tumors. Similarly, a lower frequency of HER2 overexpression was observed in tumors with *BRCA1* mutations compared with sporadic cases [19, 20], whereas there is no difference in frequency between *BRCA2* related tumors and sporadic cases.

On the other hand, a higher frequency of *TP53* mutation has been observed in *BRCA*-related breast cancer [21], but this was not found in another report [20]. With respect to prognosis, the majority of studies have reported either no difference or only slightly worse survival in patients with *BRCA1* mutations [20, 22], despite the pathologic features of *BRCA1*-related tumors that are associated with poor prognosis. In contrast, *BRCA2*-related breast cancer seems more heterogeneous than the *BRCA1* phenotype, and not clearly different from sporadic forms [23].

2. Oncogenes

Among several oncogenes, the most important in breast cancer is *HER2/neu* (also known as *c-erbB-2*). This is located on 17q21, and encodes a transmembrane glycoprotein of 185 kilodaltons with intrinsic tyrosine kinase activity [24]. HER2 belongs to the epidermal growth factor (EGF) receptor (EGFR) family, which includes four known members (HER1, HER2, HER3, and HER4), and has molecular homology to EGFR. So far, no ligand has been identified for HER2; however, following ligand-dependent activation of HER1, HER3, or HER4 by EGF or heregulin, HER2 is activated by heterodimerization, resulting in activation of diverse subcellular signal transduction pathways [25-27], and thus HER2 appears to play a central role in HER signaling.

HER2 amplification is restricted to carcinoma of glandular epithelial origin, and is at least three times more common in breast cancer than in adenocarcinomas of other sites. Amplification or overexpression of HER2 has been found in 10-40% of primary breast cancer [25, 28]. Interestingly, HER2 has been identified in a high proportion (60-80%) of ductal carcinoma in situ of high grade, comedo-type, but is not common in the low nuclear grade forms [29, 30]. On the other hand, HER2 is very rarely expressed in lobular carcinoma in situ. HER2 has not been identified in benign proliferative disease including radial scar and atypical ductal hyperplasia [31]. However, low-level *HER2* amplification in benign breast biopsies was associated with increased risk of subsequent breast cancer [32]. In this study, HER2 overexpression was not detected in any of the benign breast tissues. Thus, HER2 may have a role in the inception of breast cancer.

Initially, *HER2* was identified as a marker of advanced-stage and poor prognosis, as it was associated with lymph nodes metastases, short relapse time, poor survival, and decreased response to endocrine therapy and chemotherapy [28, 33]. Because about one-third of ER and/or PR positive tumors do not respond to endocrine therapy, several retrospective studies have addressed the association of HER2 overexpression and decreased

efficacy of tamoxifen [34-36], but not all studies concur [37, 38]. Thus, the biologic classification of tumors by HER2 status as responsive or not to endocrine therapy has been difficult [39]. In this respect, however, it is interesting to note that tamoxifen can behave as an estrogen agonist in breast cancer cells expressing high levels of the co-activator, AIB1 and Her2, resulting in *de novo* tamoxifen resistance [40]. Moreover, in a study of preoperative endocrine therapy, ER positive, HER1 positive, and/or HER2 positive primary breast cancer responded well to letrozole, an aromatase inhibitor, but responded only infrequently to tamoxifen [41]. This result may be consistent with the observation that ER positive, HER2 positive tumors are highly estrogen dependent [42], and that MEKK1, a down-stream mediator of signaling by HER1 and HER2, activates the ER and stimulates the agonist activity of tamoxifen [43]. Aromatase inhibitors block the conversion of adrenal androgens to estrogens, which leaves the ER in target tissues deprived of ligand, and thus ER becomes inactive, monomeric and unable to bind DNA in ER positive, HER1 and/or HER2 positive tumors. As for chemotherapy, several studies have suggested that HER2 overexpression is predictive for the response to anthracycline-based adjuvant therapy [44-46]. Similarly, in one study using fluorescence in situ hybridization (FISH), anthracycline-based therapy was more effective than cyclophosphamide, methotrexate and 5-fluorouracil (CMF) in the subgroup of *HER2* amplified patients with node positive breast cancer [47]. In addition, the CALGB 8541 study found a significant dose-response effect of adjuvant cyclophosphamide, doxorubicin and 5-fluorouracil (CAF) in node-positive patients with amplified or overexpressed HER2 [44-46]. In this study, maximal benefit was achieved at 60 mg/m^2, which is now considered to be the standard dose of doxorubicin. Interestingly, in breast cancer cells, *HER2* amplification increased the activity of topoisomerase IIα (topo II), leading to increased sensitivity to topo II inhibitors such as anthracyclines [48]. The *topo II* gene is located on 17q12-q21, close to the *HER2* gene, and several studies have indicated that *topo II* amplification only occurs with concurrent *HER2* amplification [49-51]. Furthermore, co-amplification of *HER2* and *topo II* in breast cancer cells conferred extreme sensitive to anthracyclines, whereas *HER2* but not *topo II* amplified cells showed an intermediate degree of sensitivity to anthracyclines [49]. Thus, it is possible that the predictive value of HER2 regarding anthracycline-based therapy over ~~cyclophosphamide, methotrexate and 5-fluorouracil (~~CMF~~)~~ is explained by the concomitant amplification of topo II [47]. In contrast, in an anthracycline and taxane neoadjuvant chemotherapy study by Rody et al [52], HER-2, but not topo II, is highly predictive of tumor response. Based on these findings, anthracycline-based therapy seems to be preferred for HER2 positive

patients; however, when anthracyclines are contraindicated, CMF-like therapy should not be withheld from patients who overexpress HER2, as HER2 positivity does not convey absolute resistance to non-anthracycline-based regimens [39]. In contrast, in a recent study of anthracycline-based neoadjuvant chemotherapy in endocrine non-responsive Her2/neu-positive breast cancer [53], topo II gene status was predictive of pathological complete remission. Thus, further prospective studies should investigate the predictive value of co-amplification of *HER2* and *topo II*.

Importantly, *HER2* represents an excellent example of the translation of basic science to clinical practice. Its product, HER2, is a molecular target of the humanized monoclonal antibody trastuzumab (Herceptin) for the treatment of breast cancer [54, 55]. The activity of this agent is highly dependent on the HER2 status of the tumor. There are currently several different methods to evaluate HER2 status, the most common being immunohistochemical staining (IHC) and FISH [56]. Both results are often in agreement, but in cases with stain intensity of +1 or +2 by IHC, the interobserver agreement is poor and the predictive value is unsatisfactory for clinical use [57, 58]. Therefore, additional testing using FISH is recommended [58-60]. Similarly, equivocal result of FISH requires IHC assessment for final determination [56]. Otherwise, the proportion of phosphorylated HER2, which is indicative of activation, may provide more significant and additional information for the prediction of response to trastuzumab [61].

3. Tumor suppressor genes

In contrast to hereditary breast cancer, few mutations of *BRCA1* and *BRCA2* have been found in sporadic breast cancer [62, 63]. However, it is interesting to note that *BRCA1* mRNA levels were decreased in sporadic breast cancer compared with normal breast tissues [64]. Furthermore, BRCA1 expression was decreased or undetectable in high-grade breast cancer [65]. Thus, BRCA1 may also be involved in the progression of sporadic breast cancer. Aberrant methylation of the *BRCA1* promoter may be one mechanism for decreased expression of BRCA1 [66], and hypermethylation of the *BRCA1* promoter was strongly associated with negativity of ER and PR [67]. In contrast, a role for *BRCA2* inactivation remains unclear in sporadic breast cancer.

The *TP53* tumor suppressor gene is located on chromosome 17p13.1, and encodes a nuclear phosphoprotein. p53 is a key component of a cellular emergency response mechanism that converts a variety of intra- and extracellular stress signals such as DNA damage, as well as conflicts between

positive and negative proliferation stimuli, into growth arrest or apoptosis. *TP53*-dependent mechanisms eliminating abnormal cells are highly efficient and provide a potent shield against proliferation and survival of cells that have been exposed to damage or have suffered faults in cell duplication and interaction with their neighbors. Therefore, without damage to, or reduced activity of, the p53 pathway, the carcinogenesis process is virtually impossible. Indeed, *TP53* mutations, most of which are missense mutations, are the most common genetic alteration in human cancer. The mutated *TP53* is able to oligomerize with the product of the wild-type allele, thus exerting a dominant negative effect [68]. In addition, missense mutations in *TP53* increase the stability of p53 protein, resulting in a protein that can be detected by IHC. Therefore, IHC has been used extensively in the study of *TP53* alterations in human cancer, and 25-50% of breast tumors are positive for p53 by IHC. Clinically, *TP53* mutations or protein overexpression are frequently found in medullary and ductal carcinomas, but not in lobular, mucinous or papillary carcinomas, and breast cancers with *TP53* mutations are in general highly aggressive and associated with negative hormone receptor status and high histological grade [69]. Until now, a number of studies have linked TP53 overexpression to poor prognosis of breast cancer, and a meta-analysis has also suggested that *TP53* mutations confer an independent relative risk of 1.7 for both disease-free survival and overall survival [70]. However, IHC for TP53 will miss *TP53* deletions. Moreover, TP53 may have functional duality, as p53 reactivation in p53-deficient tumors could induce either programmed cell death or cellular senescence, butTP53 inhibitory therapies could also be used to prevent the recovery of therapy-damaged tumor cells [71]. Thus, there is still a lot to be discovered about p53.

4. Cell survival and apoptosis related genes

Telomeres are long tandem repeat sequences at the ends of all eukaryotic chromosomes, and function to maintain chromosomal integrity during cell division. As the majority of normal somatic cells do not express the ribonucleoprotein enzyme telomerase, they are unable to maintain their telomere length, and thus with each cell division the ends of chromosomes, consisting of the telomeric repeats TTAGGG, progressively erode. In contrast, it is now well established that telomerase activation is one of the mechanisms by which cancer cells evade death and achieve immortalization. Interestingly, use of the telomeric repeat amplification protocol (TRAP) assay has detected telomerase activity in more than 95% of advanced breast cancers, whereas it was absent from 19-32% of less advanced cancers [72]. In this study, some benign tumors of breast, such as fibroadenoma, also had low

but detectable telomerase activity. To date, the TRAP assay, which is based on the polymerase chain reaction (PCR), is the most widely used strategy for detecting telomerase activity; however, quantitative analysis of telomerase activity by methods such as fluorescent based real time PCR appears to be useful to distinguish cancer from benign tumors [73]. In addition, telomerase activity was detectable in ductal carcinoma in situ lesions [74]. These lesions appear to be heterogeneous with respect to telomerase activity, and telomerase activation may precede the development of invasive cancer. If so, telomerase activity may be a useful adjunct in stratifying the risk of developing invasive breast cancer in patients with ductal carcinoma in situ.

So far, very few data are available regarding the prognostic significance of telomerase activity in breast cancer, although one study has demonstrated that levels of telomerase activity correlated with clinical outcomes and several prognostic indicators [75]. In contrast, another study failed to show any prognostic significance of telomerase activity [76]. Regulation of telomerase activity correlates with the expression of two major components, human telomerase RNA (hTERC/hTR), and human telomerase reverse transcriptase (hTERT), and therefore detection of both hTR, either by Northern blot analysis or by *in situ* hybridization, and hTERT, by reverse transcriptase(RT)-PCR, appears to be an important subject for further studies in breast cancer [77, 78].

Apoptosis (programmed cell death) is a physiological process that is crucial to the growth and development of multi-cellular organisms, and abnormality in the control of apoptosis plays an important role in tumorigenesis. Among a large number of genes implicated in the regulation of apoptosis, *BCL2* was the first shown to lead to prolonged survival of cells by inhibiting apoptosis. BCL2 belongs to the BCL2 family, which comprises at least 17 pro-apoptotic and anti-apoptotic members. Some of these family members, such as BCL2, BAX and BCLX, are expressed in normal breast, and their altered expression in breast cancer has been investigated. For example, loss of BAX, a pro-apoptotic protein, is a prognostic indicator of poor response to chemotherapy and shorter survival in metastatic breast cancer [79]. In a similar study, BCL2 expression was positively correlated with improved survival following endocrine therapy or chemotherapy with CMF [80]. Moreover, BCL2 showed an inverse relationship with HER2 and p53, whereas a strong positive relationship was seen between BCL2 and ER [81]. The prognostic role of BCL2 was supported by a recent meta-analysis [82]. Thus, BCL2 positive and BCL2 negative breast cancers have distinct phenotypes, and BCL2 expression seems to be indicative of a less aggressive phenotype.

It is generally accepted that BCL2 mainly exerts its anti-apoptotic effect by heterodimerization with pro-apoptotic members of the BCL2 family such as BAX [83]. Interestingly, phosphorylation of BCL2 reduces heterodimer formation with BAX, decreasing or inactivating the antiapoptotic function of BCL2 [84, 85]. In agreement with these studies, Shitashige et al [86] demonstrated that the phosphorylation of BCL2 at serine 70 destroyed its binding affinity for BAX, and that BCL2 kinases or their activators were synthesized *de novo* after treatment with paclitaxel, leading to phosphorylation of BCL2 in a time- and dose-dependent manner. In this study, using polyclonal antibody against phosphorylated S70-BCL2 (pS70-Bcl2), the expression of pS70-Bcl2 correlated with the response to paclitaxel and docetaxel in breast cancer, suggesting that phosphorylation of BCL2 might be predictive for the response to taxanes.

Among the inhibitors of apoptosis (IAP), *survivin* is characterized by a unique structure with a single baculovirus IAP repeat and no zinc-binding domain (RING finger), and by selective tissue distribution in cancer but not in normal adjacent tissues [87]. Survivin is repressed in the G1 phase of the cell cycle, but is highly expressed in G2/M, and the expression of survivin is negatively regulated by p53 [88]. By binding the mitotic spindle during mitosis, survivin may counteract a default induction of apoptosis in G2/M phase [89, 90]. Thus, the expression of survivin is cell cycle-dependent, and the overexpression of survivin in cancer may overcome this apoptotic checkpoint and favor aberrant progression of transformed cells through mitosis. In fact, silencing survivin gene expression could promote apoptosis of human breast cancer cells through a caspase-independent pathway [91], and expression of survivin was positive in 70% of tumors and was strongly correlated with BCL2 expression and reduced apoptotic index in breast cancer [92].

5. Metastasis suppressor genes

E-cadherin is a transmembrane glycoprotein that functions as a cell adhesion molecule through its attachment to the actin cytoskeleton via α-, β -, γ-catenin. Therefore, loss of function of cadherin or catenin will promote metastatic progression. In fact, a high frequency of LOH at the *E-cadherin (CDH1)* locus on 16q24 has been reported in lobular carcinoma [93]. In contrast, although expression of E-cadherin was reduced in invasive breast cancer [94], inactivating point mutations in the *CDH1* gene have never been found in invasive ductal carcinoma [93, 95]. These findings may offer a molecular explanation for the typical scattered tumor cell growth in invasive lobular carcinoma. As for the clinical significance of E-cadherin expression,

it is interesting to note that a recent systematic review by Gould Rothberg et al [96] has demonstrated that reduced or absent E-cadherin expression significantly increase the risk of all-cause mortality.

Other important genes are *nm23-H1* (*NME1*) and *nm23-H2* (*NME2*) located on 17q21.3. These encode nucleoside diphosphate kinase A and B, respectively. Decreased expression is the key parameter that determines metastatic potential, and may occur through a variety of mechanisms, such as LOH or methylation [97, 98]. However, *NME1* is rarely inactivated by point mutations [99]. Reduced expression of *NME1* was associated with a highly metastatic phenotype in many epithelial tumors [100, 101], and high *NME1* mRNA levels were correlated with LN negative status, less aggressive disease and longer disease-free and overall survival in breast cancer [102]. However, association with LN metastasis and the long-term prognostic value of this gene has been questioned in other studies [103-106].

6. Multidrug resistance genes

Multidrug resistance (MDR) is one of the major reasons for the failure of cancer chemotherapy. MDR can be intrinsic or acquired, and is a multifactorial phenomenon involving several specific and nonspecific mechanisms. For example, increased level of thymidylate synthetase, altered expression of Topo II, and enhanced detoxification by glutathione-linked enzyme systems can cause clinical drug resistance [107-111]. Moreover, recent molecular investigations in MDR have identified several genes coding for multidrug resistance 1 (MDR1)/P-glycoprotein (P-gp), multidrug resistance-related protein 1 (MRP1), breast cancer resistant protein (BCRP), and lung resistance-related protein (LRP). These transmembrane proteins cause MDR either by decreasing the total intracellular retention of drugs or by redistributing intracellularly accumulated drugs away from target organelles [112].

The best characterized genes responsible for the MDR phenotype are *MDR1* and *MRP1*. Both belong to the ATP binding cassette (ABC) transporter gene family, and act as a membrane-associated ATP-dependent efflux pump or an ATP-dependent organic ion transporter, respectively [112]. Overexpression of MDR1 reduces the accumulation of chemotherapeutic drugs such as anthracyclines and taxanes inside cells, and results in the MDR phenotype. However, the contribution of these MDR transporters to chemotherapeutic failure in human breast cancer remains to be proven [113-116]. Several methods are available to investigate MDR. One approach to establishing the clinical significance of these transporters is to investigate the effects of specific inhibitors of each system. Furthermore, ^{99m}Tc-Sestamibi

scintigraphy appears to be a novel technique to characterize the function of MDR transporters *in vivo* [117, 118].

II. Hormone receptors

Estrogen exposure is the most important factor in breast tumorigenesis: risk for developing breast cancer increases with extended lifetime exposure to estrogens, whereas it decreases following oophorectomy. Furthermore, it has become clear that estrogen influences the expression of the PR through the ER, and that breast cancers are dependent for growth on estrogen or progesterone, acting through the ER or the PR, respectively. Both receptors belong to a superfamily of nuclear hormone receptors including other steroid hormone receptors, thyroid hormone receptors, vitamin D receptors and retinoic acid receptors. These receptors share a common structural and functional organization, and function as transcription factors when they are bound to their respective ligand. Fig. 1 shows the functional domains of the ER, designated A through F.

On estrogen binding at the ligand binding domain (LBD), ER forms a homodimer and binds to DNA through its DNA binding domain (DBD) with high affinity at a specific site, the estrogen response element (ERE), which exists in the enhancer region of target genes. Thus, the ER influences the expression of estrogen-responsive genes, including the PR, which are important in mitogenic signaling [119] (Fig. 2). These processes are regulated by receptor coregulating proteins. In the inactivated state, corepressor proteins bound to the receptor inhibit receptor function. When estrogen binds to the ER, corepressor proteins are replaced by coactivators which relax DNA and facilitate the recruitment of transcription factors to the complex, all of which serves to enhance transcription. Importantly, transcription is stimulated

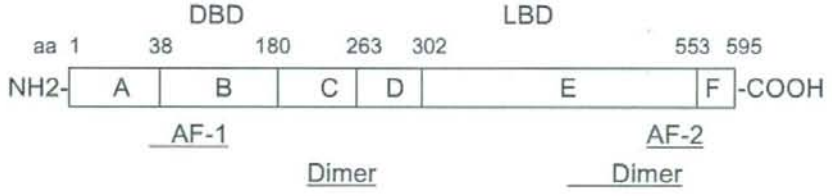

Figure 1. Schematic diagram of ERα. ER contains 595 amino acids with functional domains labeled A through F. DBD and LBD indicate DNA binding domain and ligand binding domain, respectively, and the regions important for dimerization (Dimer) and transactivation function (AF-1, AF-2) are shown.

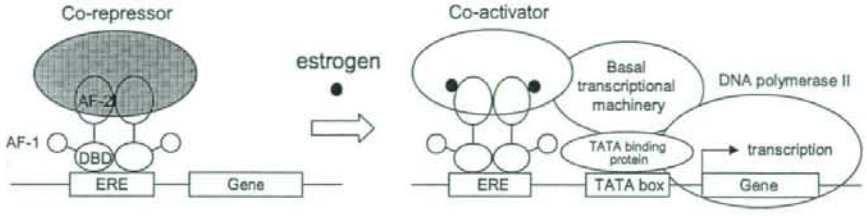

Figure 2. Model of agonist activity of ERα. In the inactivated state, DNA-bound histone deacetylase results in tight coiling of DNA, and co-repressor proteins bound to receptor inhibit receptor function. When ligand (estrogen) binds receptor, corepressor proteins are replaced by co-activators, and co-activators relax DNA and facilitate the recruitment of transcription factors to the complex, all of which serves to enhance transcription.

through at least two distinct transactivation domains, activation function (AF)-1, a hormone-independent amino-terminal ER activation domain located in the A/B region, and AF-2, a hormone-dependent, carboxy-terminal activation domain located in the E region. Although both AF-1 and AF-2 are required for maximal ER transcriptional activity in most cellular environments, AF-1and AF-2 can function independently [120]. When AF-2 is required for transcriptional activity, antiestrogens such as tamoxifen function as antagonists; however, where AF-2 is not required and AF-1 is sufficient for ER transcriptional activity, tamoxifen can function as a partial agonist on an ERE based reporter gene [121]. Thus, the function of AF-1 or AF-2 depends strongly upon promoter context, and the antagonist activity of tamoxifen arises from their intrinsic inability to activate AF-2. These observations provide a possible molecular explanation for the tissue-specific partial agonist properties of tamoxifen.

Clinically, approximately two-thirds of breast cancers express ER protein as assayed by ligand binding or IHC; however, IHC analysis of ER and PR is now becoming the predominant method [122]. Currently, ER and PR status are the only firmly established predictive factors for endocrine therapy. Endocrine therapy contributes significantly to prolongation of the disease-free survival, whereas breast cancers lacking ER rarely respond to endocrine therapy and are associated with a lower grade of histological differentiation, higher growth fraction and worse clinical outcome. Therefore, hormone resistance is a challenging problem in treatment of breast cancer, and loss of ER expression is an important cause of hormone resistance. The absence of PR from ER positive tumors generally implies that the ER is not active *in vivo*. In contrast, loss of ER may be due to a genomic deletion, mutations or rearrangement of the gene, or down-regulation of transcription at the protein

level; however, ER loss seems to be caused primarily by transcriptional inactivation of the ER gene as a result of abnormal DNA methylation within the coding domain or promoter region [123, 124]. In addition, loss of critical transcriptional factors such as ERBF-1 may also be involved in some ER-negative phenotypes [124], as may be hypoxia, which has been shown to down-regulate ER expression as well as ER function in breast cancer cells [125]. As hypoxia stimulates the hypoxia-inducible factor(HIP)-1α signaling pathway and can upregulate vascular endothelial growth factor (VEGF), an important angiogenic molecule [126, 127], it is interesting to note that HIP-1α expression was negatively correlated with ER expression level [125], and that VEGF was inversely associated with ER and PR in breast cancer patients [128]. A recent study has revealed that estrogen can directly regulate VEGF expression through the ERE [129], indicating that estrogens and antiestrogens may play a role in the regulation of VEGF. Thus, patients treated with endocrine therapy might benefit from antiangiogenic therapy aimed at suppressing or neutralizing the production of VEGF.

Furthermore, several ER variants, including base-pair insertion, transition and deletion as well as alternative splicing, may result in outlaw functions such as dominant-negative, which is itself transcriptionally inactive but prevents the function of normal ER or dominant-positive, which is transcriptionally active even in the absence of estrogen [130]. As missense mutations of the ER are not common in primary breast cancer, alternative RNA splicing of the ER generating truncated forms of the receptor protein, seems to be important. For example, a deletion of exon 5, which involves the hormone-binding domain, was cloned from an ER-negative, PR-positive tumor [130, 131]. To date, various alternatively spliced forms of the ER have been reported, but further studies are needed to assess whether these variants are expressed as proteins or have any major biological role.

Recently, a second ER gene, ERβ, was identified from rat prostate [132], and therefore the classical ER was renamed as ERα. ERα and ERβ are the products of independent genes located at 6q15.1 and 14q22-24, respectively; however, they are highly conserved at the DBD (96%), and LBD (60%) with reduced homology between other parts of the molecules, especially at the N-terminus [133]. ERβ binds estrogens with a similar affinity to ERα and activates the expression of reporter genes containing ERE in an estrogen-dependent manner [132, 134]. On the other hand, ERβ lacks a large portion of the C-terminal region of the ERα [133], domain F, which is important both in modulating the magnitude of gene transcription by estrogens and antiestrogens, and in determining the effectiveness of antiestrogens in suppressing estrogen-stimulated gene transcription [135]. In certain cells,

antiestrogens are less effective if ERα lacks the F domain, and thus ERβ may contribute to resistance to antiestrogens by interfering with their inhibition of tumor growth.

ERα and ERβ differ significantly in their tissue distribution and physiological function [136, 137]. ERβ is widely distributed in the organism, including prostate, ovary, central nervous system, cardiovascular system, lung, kidney, urogenital tract, mammary gland, colon, and immune system. In normal human mammary gland, ERα expression is restricted to the cell nuclei of epithelial cells lining ducts and lobules, whereas ERβ is the predominant form, and additional strong expression of ERβ is detected in the cell nuclei of stromal cells including myoepithelial cells [138]. Under normal conditions, the ratio of ERα to ERβ in the breast depends on plasma concentrations of estradiol, and elevated expression of ERα (as occurs in postmenopausal women) is a normal response to loss of estradiol and indicates nonproliferating cells [139]. In contrast, ERα is usually the predominant form in most breast cancers, although both ERα and ERβ are present [140]. This may be important in some breast cancers, as ERα and ERβ can form heterodimers when both are present in the same cell, resulting in receptor activity distinct from that of ER homodimers [141-143]. Interestingly, the ratio of ERα/ERβ gene expression appears to alter during carcinogenesis [140], suggesting that ERα-specific and ERβ-specific pathways may have distinct role in this process. In addition, a decrease in ERβ expression has been reported in the transition from benign breast lesion to carcinoma *in situ* [144]. Thus, the relative expression level of ERα and ERβ will be a key determinant of cellular responses to agonists and antagonists [141].

With respect to the clinical significance of ERβ, controversies have arisen regarding the function of ERβ in breast cancer as an indicator of good or poor prognosis [140, 145-147]. These studies were conducted at the mRNA level, however, it has been reported that mRNA levels of ERβ do not correlate with ERβ protein level [148], and using IHC, ERβ showed strong association with PR expression and well-differentiated carcinoma [147, 149, 150]. Another study using IHC found a significant association between ERβ positivity and increased disease-free survival [151]. So far, the differences in study population and methodology (e.g. antibodies or isotypes) make it difficult to compare the results, however, it is interesting to note that a large and systematic study by Homma et al [152] has found an association between positive staining for ERβ1-3 or ERβ1 and survival in breast cancer patients treated with tamoxifen. In this study, ERβ1 positivity was also associated with better survival in patients with ERα negative/PR negative/HER2 negative (triple negative) tumors.

Among several variant forms of ERβ, ERβcx, which is truncated at the C-terminal region but has an extra 26 amino acids due to alternative splicing, shows very poor ability to bind to estrogen and has not or only low ERE binding ability [153]. Importantly, ERβcx is capable of heterodimerization with ERα and has a dominant negative effect on ERα function [153, 154]. Moreover, Saji et al [155] found that ERβcx might be one cause of the ERα positive/PR negative phenotype of cancer cells, and that ERβcx expression along with a low expression of PR correlated with poor response to tamoxifen.

In addition to activity at the ERE, the ER can mediate gene transcription through binding to AP-1 protein complexes at AP-1 enhancer sites. Interestingly, differential ligand activation of ERα and ERβ can occur at AP-1 sites [156]. At AP-1 site, ligands including 17β-estradiol, tamoxifen, and raloxifene activated transcription with ERα, whereas 17β-estradiol did not with ERβ. Moreover, tamoxifen and raloxifene were potent transcriptional activators with ERβ at the AP-1 site. On the other hand, at the ERE site, both ERα and ERβ showed the same activation profiles with these ligands. Thus, the two ERs signal in different ways depending on ligand and response element. These findings suggesting a possible role for ERβ at the AP-1 site warrant further investigations.

PR are ligand-activated transcription factors that act in concert with intracellular signaling pathways as "sensors" of multiple growth factor inputs to hormonally regulated tissues, such as the breast. The biological response to progesterone is mediated by two, nearly identical, forms of PR, PR-A and PR-B. These two isoforms are transcribed from distinct estrogen-inducible promoters within a single-copy PR gene, and the only difference between them is that the first 164 amino acids of PR-B are absent from PR-A. PR-A functions as a transcriptional repressor of progesterone-responsive promoters, whereas PR-B functions as a transcriptional activator of the same genes [157], indicating that their balance may affect endocrine response in breast cancer. To our knowledge, only a few studies have investigated the clinical significance of the expression of these PRs [158, 159]. In a study by Bamberger et al [158], expression of both isoforms was correlated with ER expression, and most breast tumors expressed PR-A at levels equal to or higher than those of PR-B [158]. Interestingly, PR-A expression at a level less than or equal to that of PR-B was associated with G1/G2 grading, whereas a PR-A expression greater than that of PR-B was associated with a more undifferentiated phenotype (G3 grading). Another study investigated PR-A and PR-B expression using immunoblot analysis in ER-positive, PR-positive patients with node-positive breast cancer, and found that tamoxifen-treated patients with high PR-A/PR-B ratios were more likely to relapse than

those with lower ratios [159]. Thus, PR-A/PR-B ratio may identify a subgroup that benefit poorly from endocrine therapy.

III. Multiparameter gene expression analysis

Gene expression profiling recently has been introduced into the clinical literature during the last decade. The molecular signatures has provided us with insight into the heterogeneity of breast cancer and hold promise for improving diagnosis, for the prediction of recurrence and in aiding selection of therapies for individual patients [160, 161]. For example, using complementary DNA microarray, Perou et al [162, 163] classified breast cancer into distinct sub-categories; luminal A (ER+, HER2-), luminal B (ER+, HER2+), HER2-overexpressing (ER-, HER2+), basal-like (ER-, HER2-, EGFR+, and/or cytokeratin 5/6+), and normal breast-like subgroups. Luminal subgroups share features with luminal epithelial cells arising from the inner layer of the duct lining, normal breast-like subgroup is characterized by high expression of basal epithelial-cell genes and low expression of luminal epithelial-cell genes. Basal-like subgroup shares features with normal breast basal epithelial cells, and often exhibits p53 mutation and low expression of BRCA1 tumor suppressor genes. This phenotype is common among BRCA1 carriers and sporadic triple-negative tumors. Importantly, when analyzed in a similarly treated group of patients, a significant difference in overall survival was found between these subtypes [163, 164].

Until now, several technologies have been developed to generate molecular signatures, including cDNA and oligonucleotide arrays and multiplex PCR technologies, and a number of studies specifically investigated molecular signatures in breast cancer, primarily focused on associations between particular sets of genes with altered expression and survival [161]. Here, we introduce several gene-expression profiling models including the 21-gene recurrence score (Oncotype DX), the 70-gene prognostic signature (MammaPrint).

The 21-gene recurrence score is a RT-PCR assay that measures the expression of 21 genes—sixteen cancer-related genes and five reference genes—in RNA extracted from formalin-fixed paraffin-embedded tissue samples from primary breast cancer [165]. These genes were selected from a much larger set of genes following the analysis of retrospective test sets of clinical material from several sources, including specimens from a cooperative group trial in which patients with ER-positive, node-negative breast cancer received tamoxifen .versus tamoxifen plus chemotherapy (NSABP B-20) [166]. The levels of expression of the 21 genes are manipulated by an empirically derived, prospectively defined mathematical

algorithm to calculate an recurrence score (RS), which is then used to assign a patient to one of three groups by estimated risk of distant recurrence: low, intermediate, and high. The assay is intended to estimate risk of recurrence of patients with hormone receptor–positive, node-negative, stage I or II breast cancer. It has been suggested that tamoxifen-treated patients with an excellent estimated prognosis may be spared adjuvant chemotherapy. In addition, patients with a high RS appear to achieve a higher proportional benefit from adjuvant CMF/MF chemotherapy than those with low or intermediate RSs. This model was validated by the analysis of specimens and data from a second set of patients with node-negative, ER-positive breast cancer treated only with tamoxifen, who were enrolled in the NSABP clinical trial B-14 [167].

The 70-gene signature, so called Netherlands signature (MammaPrint) is a gene expression profiling platform, requiring a fresh sample of tissue that is composed of a minimum of 30% malignant cells [168]. In the development, primary tumors from patients with node-negative primary breast cancer were analyzed on oligonucleotide microarrays, and the data were subjected to supervised classification to establish a 70-gene RNA expression profile that correlated with a relatively short interval to distant metastases. The signature largely consists of genes regulating proliferation plus those involved in invasion, metastasis, stromal integrity, and angiogenesis. This assay was then validated in stage I, II primary breast cancer patients younger than age 53 years [168].

The Rotterdam 76-gene signature consists of a 76-gene microarray assay that does not overlap with either the 21-gene recurrence score or 70-gene signature [169, 170], and was specifically studied in node-negative breast cancer patients. The breast cancer gene expression ratio (two-gene expression ratio) is a quantitative RT-PCR–based assay that measures the ratio of the HOXB6 and IL17BR genes [171, 172]. This assay was developed to predict poor outcome in ER-positive patients treated with tamoxifen [173, 174]. The wound-response gene signature model is developed based on the hypothesis that fibroblast play a similar role in wound healing and tumor progression [175].

IV. Molecular detection of breast cancer cells

To date, morphological examination remains the gold-standard for cancer diagnosis. However, inaccurate diagnosis may result from inadequate preparation of the specimen, unskilled evaluation and differences in interpretation of morphological criteria. Therefore, it would be of great value to have a more sensitive method for cancer diagnosis that is less subject to evaluation bias. Metastatic spread of breast cancer cells occurs through the

blood and lymphatic circulation, and these cells settle in a specific organ, and develop into secondary tumors [176]. Clinically, micrometastasis may not be detectable at the time of diagnosis of primary breast cancer, even though it may have already occurred. This includes mini-residual and quiescent, dormant cells or stem cell [176, 177]. Currently, tumor metastasis is diagnosed by clinical manifestations and imaging studies, such as bone scans, chest radiography, liver ultrasonographys, or computed tomography scans, together with serum marker assays. However, clinical data are not sufficient to support the use of circulating tumor markers such as carcinoembryonic antigen (CEA) and carbohydrate antigen 15-3, p53 or circulating extracellular domain of HER2 for routine clinical decision-making [160]. Therefore, the ability to rule out the presence of micrometastasis at any stage of the clinical management protocol, whether before, during, or after therapy, would provide a useful monitoring and diagnostic tool.

In this respect, PCR and RT-PCR, which exponentially amplify a previously undetectable amount of nucleic acid to a detectable level, are promising high-sensitivity techniques. The use of these methods offers a number of advantages compared with cytological and IHC techniques, as both suffer from poor sensitivity [177, 178]. In addition, PCR-based methods can be designed to detect a cell using a specific target. Today, gene-expression profiling of single cells was reported to be feasible [179, 180].

1. Molecular diagnosis of breast cancer

To detect small numbers of cancer cells, not only in samples from fine needle aspiration or core needle, but also in nipple discharge or in samples that are cytologically undetermined to be malignant, PCR-based methods seem to be promising for accurate diagnosis. For example, one study investigated *HER2* mRNA in nipple discharge from breast cancer [181]. In this study, amplification of the *HER2* gene was more frequent in carcinoma *in situ* than in invasive types. However, mRNA in discharge may not be a suitable target for molecular detection, as it may be rapidly degraded in the collected fluid. Therefore, the best approach for nipple discharge seems to be PCR. In a recent study using PCR for seven markers of LOH, at least one LOH was found in either supernatant or cell pellet in 70% of breast cancer patients with nipple discharge [182]. This method showed higher sensitivity than smear cytology or quantification of CEA concentration of the discharge. However, some papillomas also showed LOH [183]. In a similar study using methylation-specific PCR, methylated alleles of the *cyclin D2*, *RAR-β* and *Twist* genes were frequently detected in cells collected from breast ducts, but rarely in ductal lavage fluid from healthy ducts [184]. These genes are

hypermethylated in more than 30% of breast cancers, but unmethylated in healthy mammary epithelial cells, mammary stroma, and white blood cells [185-187]. In association with this, it is interesting to note that in the study by Fackler et al [188], the quantitative multiplex methylation-specific PCR doubled the sensitivity of detection of cancer compared with cytology. In addition, telomerase detection in cells derived from fine needle aspirates of breast showed high sensitivity and specificity for cancer detection [73, 189]. Among the several gene alterations in cancer, telomerase is unique as it is detected in almost all kinds of cancer tissues with a very high positive rate. Thus, molecular analysis of nipple fluid appears to be promising, not only for diagnosis of breast cancer but also for screening of high risk women [187, 190].

Increased quantities of DNA or nucleosomes have been found in the plasma or serum of breast cancer patients [191-200] (Table 3). Interestingly, this circulating extracellular DNA may exhibit tumor-related alterations. PCR for LOH, p53 mutation, and methylation of the first exon of $p16^{INK4a}$ found that 66% of breast cancer patients had molecular alterations in plasma DNA [192]. In a similar study using different markers for microsatellite instability and LOH, two studies independently demonstrated that plasma or serum DNA of breast cancer patients displayed the same gene alterations as those found within the primary tumor [191, 193]. Of clinical importance, the presence of plasma DNA of tumor origin correlated with pathological parameters associated with a poor prognosis, histological features of highly malignant lesions [192]. Moreover, circulating DNA was shown to undergo quantitative changes in cancer patients after radiation or chemotherapy [194-196]. Thus, DNA derived from plasma or serum could be used for molecular diagnosis of cancer, and monitoring the disease course [197, 201]. Moreover, it is interesting to note that telomerase RNA (hTERT) could be detectable in plasma from cancer patients [202].

2. Clinical significance of micrometastasis in breast cancer

In contrast to hematological malignancies, solid tumors rarely have specific genetic changes. To overcome this limitation, the majority of studies have used RT-PCR for molecular detection of micrometastasis [177, 203]. Table 4 lists the molecular markers used most frequently to detect micrometastasis in breast cancer. The major tissues of interest for molecular detection of breast cancer micrometastasis and disseminated or circulating tumor cells (DTCs, CTCs, respectively) are LNs [204-249], BM [193, 213, 245, 250-288] and PB [193, 207, 212, 218, 228, 235, 242, 245, 250, 254, 255, 257, 258, 260-273, 276, 278, 282, 285, 287, 289-352]. In addition, cancer cells are detected in body fluids such as cerebrospinal fluid (CSF) and pleural

Table 3. Representative studies of circulating extracellular DNA.

Author ,Year	Targets	Sample source	Positive rates in breast cancer		Positive rates in controls	Remarks
			Primary, NMD	Metastatic disease		
RT-PCR based analysis of plasma/serum sample						
Chen, 1999 [191]	LOH	P, S	18.4%*(49)	66.7%* (3)	0%* (>10[1])	Useful
Silva, 1999 [192]	LOH	P	61% (38)	-	0% (13)	Identical to tumor DNA
	MI	P	5% (38)	-	0% (13)	Not present in tumor DNA
	Methylation of p16 ^{NK4a}	P	14% (43)	-	0% (3)	
	p53	P	5% (62)	-	0% (5)	
Shaw, 2000 [193]	LOH	P	15% (39)	63% (16)	0% (9)	Sensitive
	MI	P	10% (39)	38% (16)	0% (9)	Sensitive
Rykova, 2004 [200]	Methylation[2]	P	33-47%(17)		0-1%(6)	Promising
RIA, ELISA based analysis of plasma/serum sample						
Leon, 1977 [194]	Free DNA(RIA)	S	Increased**		7% (55)	Increased, especially in metastatic diseases.
Holdenrieder, 1999 [195]	Nucleosome (ELISA)	S	84% (43)***		4% (50)	Sensitive, useful for monitoring
Kuroi, 2001 [196]	Nucleosome (ELISA)	P	68.7% (99)	84.6% (26)	4.5% (111)	Predictive of response

Abbreviations: ELISA, enzyme linked immunosorbent assay; LOH, loss of heterozygosity; MI, microsatellite instability; NMD, Non metastatic disease; P, plasma; RIA, radioimmunoassay; S, serum.

() Number of patients.

1) At least 10 controls/microsatellite markers were investigated.

2) APC, RASSF1A, RAFβ, CDH1, CDH13 gene promoter methylation.

*Percentage of positive cases for at least one locus, **Metastatic disease was included, ***Breast cancer, NOS.