We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Genomic Expression Profiles: From Molecular Signatures to Clinical Oncology Translation

Norfilza M. Mokhtar, Nor Azian Murad, Then Sue Mian and Rahman Jamal

Additional information is available at the end of the chapter

1. Introduction

Study related to diseases such as cancer has changed tremendously for a decade. For many years, the study was restricted largely to a single gene or a few genes in cancer cells. The studies have uncovered the roles of individual genes in the uncontrolled behavior of cancer cells. Studying the functional roles of genes in cancer cells has deepened our understanding not only the cancer cells as well as normal cells. Since 2003 onwards, the trend of publications was focusing on the analysis of thousands of genes with related molecular pathways. Steps taken from this analysis is then translated to clinical practice for the biological markers for an early detection, monitoring, prognosis of the disease and response to therapy.

The completion of the Human Genome Project in 2003 enabled a new era in biological sciences, in particular molecular medicine. The availability of the database of full sequences of approximately 3 billion base pairs and approximately 30,000 genes in human DNA will lead to a better understanding of physiological and pathophysiological changes in human body. Genome-wide expression technology allows the simultenous analysis of thousands of genes in a single experiment. The availability of the technology alters the way biological experiments can be designed. This has resulted of so called 'discovery biology'. The large amount of data produced by microarray resulted to new and unexpected features of cellular functions.

Since it was first introduced, microarrays are widely used for basic research, the development of prognostic tests, target discovery or toxicology researchs. The new form of cancer screening utilizes the molecular data generated from microarray studies. We will discuss the application of gene profiling data in the clinical screening of cancer. It is hopefully will give a broad picture the pipeline required to discover biomarkers of cancer.

The chapter is subdivided into a series of sections; each will discuss the scientific evidence on the molecular and cellular studies in selected cancers. We will try to critically assess the evidence upon which the theory on the cancer was built. The conversion of normal cells into cancer cells is a complex process and multistep processes. Scientists for many years tried to uncover the causes of cancer and emphasize certain oncogenes, or tumor suppressor genes or other groups of genes. Further information on how these findings were translated to the clinical settings will be provided. To date, with the massive gene expression profile data available to the researchers, there are still major hurdles in validating and reproducing the results. We will discuss the major drawbacks associated with the use of molecular signatures as the biomarkers or response to treatment.

2. Molecular signatures in colorectal carcinoma

Colorectal cancer (CRC) is a type of cancers that develops in the colon or the rectum of the human digestive system or gastrointestinal tract (1).Colorectal cancer is the third leading cause of death in both men and women in the US with 141,210 new cases and 49,380 death expected in 2011 (2). CRC progresses slowly over a period of time usually between 10 to 15 years (3, 4). The tumor begins with noncancerous polyps where the tissues that form the lining of the colon or rectum differentiate into cancerous tissues (5). Approximately, 96% of colorectal cancers are adenocarcinomas, which arise from the glandular tissue (6). It can grow along the lining of the epithelium into the wall of the colon and rectum and invade the digestive system (7). In addition, the cancerous cells can also penetrate into the circulating systems, the blood and lymphatic systems which known as metastasis (7). Typically, the cancerous cells will first spread into the nearby lymph nodes and subsequently penetrate into other organs such as liver, lungs and ovary through blood vessels (8, 9). Colorectal cancer can be classified as tumors/nodes/metastasis (TMN) staging and Dukes classification (12). The TMN assigns the number based on three categories, T, M and N, which are the degree of invasion of the intestinal wall, lymph node involvement and the degree of metastasis, respectively (10). The higher number of TNM system indicates the advanced stage of colorectal cancer (10).

Unhealthy lifestyles such as alcohol consumption, high intake of red meat, obesity, smoking and lack of physical activities are among the risk factors for CRC (1, 11). Age and gender also play significant role in the development of CRC as the risk is higher in male and elderly(7). People with inflammatory bowel disease such as ulcerative colitis and Crohn's disease are also at high risk of getting CRC (12). Among the patients with Crohn's disease, approximately, 2%, 8% and 18% of the patients will develop CRC after 10, 20 and 30 years, respectively (12). About 20% of patients with ulcerative colitis develop CRC within the first 10 years (13). Mutations in genes such as *KRAS*, *APC*, and *MMR* are the well-documented genetic factor that contributes to colorectal cancer (3, 14, 15). Individual with family history of CRC in two or more first degree relatives have 2 or 3-fold greater risk of getting CRC and this has accounted for 20% of all cases (7). Examples of CRC involving genetic mutations are hereditary nonpolyposis colorectal cancer (HNPCC or Lynch Syndrome), Gardner syndrome and Familial adenomatous polyposis (16).

Diagnosis of CRC is based on tumor biopsy performed during the sigmoidoscopy or colonoscopy (7). CT scan of chest, abdomen and pelvis could be performed to determine the metastasis state and in certain cases, PET or MRI may be used to assist in the diagnosis (7).Molecular testing for patients with a strong family history can be performed to identify mutation, thus initiate early diagnosis and screening in family members. In addition, molecular characterization of mutations involved in CRC may help doctors to plan a better treatment strategy for the patients. Managing our lifestyles can help us to reduce our risk of getting CRC, for example by improving lifestyle through regular exercise, increasing the consumption of whole grains, fruits and vegetables and reducing the red meat intake (17). The treatments for CRC include surgery, chemotherapy and radiotherapy.

2.1. Molecular biology of colorectal cancer

Colorectal cancer is a multistep process that includes accumulation of several genetic and epigenetic alterations (18, 19). It is well characterized that the adenoma to carcinoma sequence is due to accumulation of the genomic alteration, which is induced by genomic instability (4, 20). Genomic instability is an event, which will increase tendency of the genome to acquire mutations when several important processes in maintaining and replicating the genome are malfunction. It is a hallmark of many human cancers (20). There are three well-reported genomic instability pathways that could lead to colorectal cancer, which will be discussed in details below.

a. **Chromosomal instability (CIN)**

   Chromosomal instability lead to increase rate of losing or gaining chromosomes during cell division and accounts for 15% to 20% of sporadic CRC as well as Lynch Syndrome (Hereditary Non-Polyposis Colorectal Cancer) (21).There are three mechanisms involved in this process that includes structural chromosome instability, the chromosome breakage-fusion-bridge (BFB) cycles and numerical instability (22). Structural chromosome instability is caused by high incidences of DNA double-strand breaks, which may lead to abnormalities in chromosomal segregation during mitosis. Chromosomal damage may result in mitotically unstable chromosome, which may promote an event known as breakage-fusion-bridge (BFB) (22). An abnormal number of centrosome may be caused by abnormal mitotic polarity as well as unequal segregation of chromosomes during the anaphase stage (23). CIN promotes cancer progression by increasing clonal diversity (21). In the clinical perspective, large meta-analysis has shown that CIN is a marker of poor prognosis in colorectal cancer (20).

b. **Microsatelite instability (MIN)**

   Microsatellites are repetitive sequences of DNA, which is highly varied between individuals (24). The most common microsatellites in human is a dinucleotide repeat of CA (25). MIN is a condition, which is manifested by damaged DNA due to defective in the DNA repair mechanism. CRC with the presence of MIN have a better prognosis compared to CRC with CIN (26). MIN involves the inactivation of the DNA Mismatch

Repair (MMR) genes via aberrant methylation or somatic mutation (26). HNPCC or Lynch Syndrome is an example of CRC, which is caused by MIN with 15% occurrence (27). MIN could cause CRC in 2 mechanisms; 1) mutations in the MMR genes where error in the microsatellite repeat replication is unfixed. This leads to the inactivation of tumor suppressor genes (TSG), a group of genes which is crucial in maintaining cell cycle progression and apoptosis induction (20). Inactivation of these genes may lead to tumorigenesis through uncontrolled cell division 2) epigenetic changes that silence the MMR genes (20).

c. CpG Island Methylation and CpG Island Methylator Phenotype (CIMP)

Hypermethylation of the promoter region of a gene that contains CpG Island (CGI) and global DNA hypomethylation are associated with epigenetic instability in colorectal cancer (20). CGIs are short sequences rich in the CpG dinucleotides and are observed in the 5' region of almost half of all human genes (28). *In-vitro* study of BRAF in CRC cell lines showed no correlation between BRAF and CIMP (29).

2.2. Genome Wide Association Study (GWAS) in colorectal cancer

The completion of Human Genome Project in 2003 and the International HapMap Project in 2005 have opened up a new era in genetic and phenotype correlation study (30). The completion of these two projects has made the Genome wide association study (GWAS) possible. GWAS is considered as the most powerful tool to study the association between phenotypes and genotypes and also to identify common, low-penetrance susceptibility loci in a particular disease. In addition, GWAS can also be employed to investigate gene-environment interactions and the pooled analyses may also lead to the identification of novel modifying genes. Several GWAS studies have been performed in colorectal cancer and several loci were identified to be associated with CRC such as 8q24 (128.1-128.7 Mb, rs6983267) (31, 32). The *C-MYC (MYC)* oncogene is located approximately 300 kb from this region and is often over-expressed in CRC (33). Validation studies have confirmed that rs6983267 loci as the most promising variant in CRC, which has increased the chance of getting CRC by approximately 1.2 fold (33, 34). Recent publication has suggested that this variant is involved in enhancing the Wnt signaling and MYC regulation, which are known pathways in carcinogenesis (35). However, further functional analyses are still needed in order to determine the function of this variant. In the Japanese population, this variant leads to an increase risk of CRC with an allelic OR=1.22. Even after the adjustment for confounders, the OR remains significant (OR = 1.25). In the ARCTIC report, a locus at 9p24 was identified to be associated with CRC and was confirmed in the Colorectal Cancer Family Registry. Several numbers of loci that include 18q21:*SMAD7*; 15q13.3:*CRAC1*; 8q23.3: *E1F3H*; 14q22.2:*BMP4*; 16q22.1: *CDH1* and 19q13.1:*RHPN2* were also found to be associated with CRC. These genes have been shown to be involved in CRC progression. Studies conducted in Korean and Japanese patients with CRC have identified a novel susceptible locus in *SLC22A3*, which was significantly associated with distal colon cancer (36). The variant, rs7758229, was located on 6q26-q27 with OR=1.28. Three variants, rs7758229,

rs6983267 and rs4939827, in *SMAD7* together with alcohol consumption may increase the risk of CRC by approximately two-fold. Several variants including rs6983267, rs6695584, rs11986063, rs3087967, rs2059254 and rs72268855 showed evidence of association with CRC in Singaporean Chinese (31). sSNP rs3087967 at 11q23.1 was associated with increased risk of CRC in men (OR=1.34) compared to women (OR=1.07). The rs 10318 at locus 15q13 (GREM1) was also associated with CRC with OD =1.19 (37).

Almost half of the susceptibility loci in CRC are located nearby the transforming growth factor beta gene (*TGF-β1*), which is important in the carcinogenesis (38). An elevated level of *TGF-β1* was linked to tumor progression and recurrence in CRC. Germline mutations in components of *TGF-β1* signaling pathway such as SMAD4 is responsible for the high-penetrance juvenile polyposis syndrome. Other genes are *SMAD4, RHPN2, BMP4, BMP2* and *GREM1*.

2.3. Gene expression profiling in colorectal cancer

Gene expression profiling was performed to compare between colorectal adenomas and CRCs and the result showed that the level of six cancer-related gene sets were increased in CRCs compared to adenomas (FDR<0.05). These include genes that involved in chromosomal instability, proliferation, differentiation, angiogenesis, stroma activation and invasion. Changes in the activity of the chromosomal instability were the most significant gene set (FDR=0.004) (39). The key genes that are associated with colorectal adenoma to carcinoma progression are AURKA, TPX2 (Chromosomal instability), PLK1 (Proliferation), ADRM1 (Differentiation), SSCA1 (Stroma activation), SPARC and PDGFRB (Invasion). The expression levels of these genes were significantly higher in CRC compared to adenoma (p<1e-5). Overexpression of AURKA induces centrosome amplification, aneupploidy and cellular transformation *in vitro* (40). AURKA interacts with TPX2 and plays a role in centrosome maturation and spindle formation (41). The polo-like kinase 1 (PLK1) is important in spindle formation and cell cycle progression during the G2 and M phase (42).

Wu and colleagues showed that the extracellular matrix and metabolic pathways were activated and the genes related to cell homeostatsis were downregulated. In this study, they compared cancer transcriptome using massive parallel paired-end cDNA sequencing in 3 different tissues, CRC tissue (stage III), adjacent non-tumor tissue and normal tissue from a 57 years old female patient. They detected 1660, 1528 and 941 significant differential genes (DEGs) between the CRC and adjacent tissue, the CRC and normal tissue; and the adjacent and normal tissue respectively. 15-prostaglandin dehydrogenase (*15-PGDH*) was downregulated in cancer compared to normal tisssue, which is common oncogenic event in approximately 80% of CRC cases. The transition between adenoma and carcinoma processes involved inactivation of *TGFBR2*, thus progressive inactivation of this gene from cancer-adjacent and normal tissue was expected. In addition, *APC, MYH, CD133, IDH1* and *MINT2* were also dysregulated in CRC. They also identified many genes involved in extracellular matrix (ECM) receptor interactions were highly dysregulated in cancer. The findings showed that all collagen type proteins were overexpressed up to 1000-fold in cancer tissue.

In addition, members of MMP family, which degraded the ECM structures, were also induced significantly in tumor. These include MMP1, MMP3, MMP14 and MMP7. Other cell-cell adhesion-related molecules for examples laminins (LAMA4, LAMA5, LAMB1, LAMB2 and LAMC2) and integrins (ITGA5, ITGB5, ITGA11 and ITGBL1) were elevated in cancer tissues. It was suggested that "angiogenesis switch" was activated in tumor tissues since vascular endothelial growth factor (VEGF) was found to be upregulated. In conclusion, up-regulation of the ECM pathway and the angiogenic growth factors may lead to remodelling of the ECM pathways as well as expansion of the new vessel networks, which subsequently resulted in CRC progression. Since their results in concordance with previous studies that showed the ECM pathway was subjected to intensive epigenetic modification, therefore this ECM may be a good candidate as prognostic biomarkers in CRC (43).

3. Molecular signatures in ovarian cancer

Ovarian cancer is among the top ten leading cancers among women the United States. In this country alone, there are approximately 22,280 new cases and 15,500 estimated death in 2012 (44). At our local population, approximately 1627 women were diagnosed in 2003 to 2005 and the figure showed increasing trend in 2007(45).In Japan and Sweeden, the incidence of ovarian cancer per 100,000 women is 3.1 cases and 21 cases respectively (Green et al., 2012). Due to vague or absence of early signs and symptoms, patients suffer from this cancer seek late treatment (46). Therefore, the cancer is normally diagnosed late when the disease is not longer confined to the ovary. Based on different morphological characterisitcs of the cancer, it is divided into epithelial and nonepithelial types. The epithelial type is further subdivided into serous, mucinous, endometrioid and clear cells. On the other hand, the nonepithelial is granulosa cells, mixed germ cells tumour, immature teratoma, dysgerminoma and teratoma. The risk factor for this cancer is unclear, however the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort study has recently documented that women who smoke more than 10 cigarettes a day had doubled the risk to develop mucinous ovarian cancer (47). This has suggested that the effect of smoking differs based on different histological subtypes of ovarian cancer(47). On the other hand, a study has shown that long period of breastfeeding seems to have reduced risk of ovarian cancer (OR = 0.986, 95% CI 0.978-0.994 per month of breastfeeding) (48).This effect of breastfeeding was also varies between histological subtypes as there was no association between breastfeeding and borderline serous or mucinous cancer (48).

Ovarian cancer was initially divided based on molecular pathways involved in the development and progression of the subtypes (49). Type I is low-grade serous, low-grade endometrioid, mucinous and clear cells. They are believed to arise from benign lesions such as ovarian inclusion cyst or endometriotic lesions. These lesions follow the stepwise pattern, whereby it evolved from the benign adenoma to borderline and finally to malignant tumours (table 1).

Type II ovarian cancer is high-grade serous, high-grade endometrioid and undifferentiated. The common mutations that are found in these subtypes are p53, BRCA1/2, PIK3CA with chromosomal instability. They normally involve the peritoneum and grow rapidly.

Characteristics of tumour	Type I	Type II
Type of tumor	Low-grade serous Low-grade endometrioid Mucinous Clear cell	High-grade serous High-grade endometrioid Undifferentiated
Common mutations and genetic modifications	KRAS BRAF PTEN CTNNB1 Microsatellite instability	p53 BRCA1 BRCA2 PIK3CA Chromosome instability

Table 1. Ovarian subtypes based on common mutations and genetic modifications

In clinical practice, the gyneoncologist still use CA125 as the biomarker to monitor treatment of this cancer. However, it is not sensitive and specific to detect the cancer in its early stage (46). It is of great demand to find new molecular marker for the ovarian cancer.

Ovarian cancer is treated by surgery, radiation or platinum-taxane based chemotherapy depending on the subtypes and extent of the cancer (50). Patients at stage I and II will undergo bilateral salphingo-oophorectomy. While for advanced cases, adjuvant chemotherapy combined with surgery is highly recommended. With the latest understanding on the mutational types of ovarian cancer, mitogen activated protein kinase (MEK) inhibitor such as CI-1040 was used to test the potential therapeutic agent in *in vitro* ovarian cancer cell line (51). This cell lines containing KRAS or BRAF mutations, which are known mutations for type I ovarian cancer. The targeted therapy for type II ovarian cancer encounters difficulty due to lack of common molecular pathways. In two cohort studies involving 16 international centers, women with BRCA1 or BRCA2 mutation were treated with two different doses of Olaparib (52). This drug is orally active poly(ADP-ribose) polymerase (PARP) inhibitor. The result showed a promising therapeutic indexin ovarian cancer patients with mutation of BRCA1 or BRCA2 (52).Based on this study, Olaparib has possible as therapeutic agent in type II ovarian cancer.

3.1. Molecular biology of ovarian cancer

Ovarian cancer is a heterogenous disease and thus, there is no clear molecular genetics involved in the transition of normal ovarian epithelial cells into cancer cells. Approximately 10 to 15% of ovarian cancer is thought to run in the families (53). It is closely related to BRCA1 and BRCA2 mutation (53). It was recently published that suggested screening of BRCA1/2 mutation in patients with ovarian cancer prior to chemotherapy treatment (54). This is because presence of such mutations may influence the treatment outcomes (54). Human DNA repair mismatch genes for example MLH1 and MSH2 accounts for 10% of patients with hereditary nonpolyposis colon cancer syndrome (55). Other related genes include glutathione S-transferase M1 (GSTM1) is associated with endometrioid or clear cells ovarian cancer.

Approximately 85% of ovarian cancer is regard as sporadic with no apparent hereditary factors. Accumulation of mutagenic genes and deregulation of signaling pathway frequently lead to the development of cancer. Different subtypes of ovarian cancer reveal different molecular pathways. Coagulation pathway was reported to be disturbed in clear cell ovarian carcinoma (56). Genes that stimulate or inhibit coagulation were noted to be dysregulated. Angiogenesis and glycolysis are two major activated pathways in clear cell ovarian carcinoma (56). Vascular endothelial growth factor (VEGF) and its receptor FLT1 were upregulated in this type of cancer and involved in angiogenesis. Earlier study by Yamaguchi et al 2010, reported molecular pathway related to clear cell ovarian cancer was related to hypoxia-inducible factor 1 (HIF1α) (57). HIF1α regulates ADM, which is related to angiogenesis. It also regulates genes that are linked to glucose metabolism including SLC2A1 in glucose transport and HK1/HK2 and ENO1/ENO2 in glycolysis. Both pathways could act as potential therapeutic target based on the small interfering RNA of genes related to these pathways combined with antiangiogenic drug, Sunitinib(56).

3.2. Gene expression profiling in ovarian cancer

In ovarian cancer study, microarray was used to classify 113 samples from five different histopathological subtypes; endometrioid, serous, mucinous, clear cell and mixed type according to the gene expression pattern (58). The results showed 95% of all samples were clustered within their expected groups. Gene expression profile in this study failed to distinguish between high-grade endometrioid and serous ovarian cancer. The result derived from the principal component analysis demonstrated the separation of celar cell, mucinous and endometrioid with serous ovarian cancer. This can be explained through the origin of these types of cancer, which is Mullerian epithelium. In contrast to serous ovarian cancer, which most likely arise directly from ovarian surface epithelium (58). Microarray was also used to distinguish between various grades of clear cells ovarian cancer from other subtypes of ovarian cancer including serous papillary (59). Among genes identified were E-cadherin and osteonidogen were detected at high level in clear cells. While discoidin domain receptor family member (DDR1), estrogen receptor 1 and cytochrome P450 4B1 were at a low level in clear cells ovarian cancer compared to other ovarian cancers (59).

A separate microarray study was done on 285 of various grades of endometrioid and serous ovarian cancer samples that were analysed together with low-grade serous and endometroid ovarian cancer (60). The result showed high-grade serous subtype was related to overexpression of Wnt/βcatenin and cadherin pathway genes including N-cadherin and P-cadherin but low E-cadherin protein expression. This finding demonstrated the high-grade serous ovarian cancer contained mesenchymal expression pattern. Also it has suggested there is epithelium-mesenchymal transition in this subtype of ovarian cancer. High expression of genes related to proliferation and extracellular matrix-related genes such as COL4A5, COL9A1 and CLDN6. Immune cell markers such as CD45, PTPRC and lymphocyte markers, CD2, CD3D and CD8A were expressed low in the high-grade serous subtype (60).

Gene expression profiling was also performed to detect genes that were differentially expressed in primary ovarian cells as compared to the neighboring metastatic tissue omentum (61). Among significant genes include hepsin (HPN), which is related to epithelial cells. Using immunohistochemistry technique, HPN protein was localised in epithelial cells, suggestive that it can be a marker of epithelia cells and not cancer (61). In advanced stage of ovarian cancer, predictive markers were suggested to be different. For example EZH2, PTTN and Lamin-B, were positively detected in primary as well as metastatic omental tissue. MGB2 is another biomarker that significantly overexpressed in primary as well as ovarian metastatic tissue. To characterize two different cancers; breast and ovarian cancers that involve serosal cavities, gene expression profiling was carried out (62). About 288 differentially expressed genes with at least 3.5-fold up-regulated in breast and ovarian/peritoneal serous cancers (62). These groups of genes may potentially used to distinguish both cancers for better therapeutic intervention.

Microarray of the nonepithelial ovarian cancer or type II ovarian cancer is still limited. Despite its rare incidence of this subtype of ovarian cancer, we have performed microarray assay on the formalin-fixed paraffin embedded tissues (63). About 804 differentially expressed genes with at least 2-fold change (P<0.005) (63). Among the significant genes were EEF1A2 and E2F2; which were up-regulated in nonepithelial ovarian cancer as compared to the normal ovarian cells. EEF1A2 may act as oncogene and play an important role in the progression of cancer (64). E2F2 plays a role in cell cycle and positive immunostaining in all subtypes of nonepithelial ovarian cancer may suggest its role as an oncogene (63).

4. Molecular signatures in endometrial cancer

Cancer of endometrium is cancer arises from the inner lining of the uterus. The cancer appears in multiple histologic subtypes as a result of müllerian differentiation. They are divided into two broad groups that include endometrioid and non-endometrioid (65). The recent surgical staging of endometrial cancer is based on the International Federation of Gynecology and Obstetrics in 2008 (66). Endometrial cancer is divided into two types based on the underlying pathical findings and clinical observations. There are endometrioid (type I) and nonendometrioid carcinoma (type II). The former is the commonest type (85% of total cancer) with history of estrogen exposure with underlying endometrial hyperplasia (67).Also the cancer cells expressed estrogen and progesterone receptor and typically of low histopathological grade (68). The majority of patients are relatively young with good prognosis. While the second type is less common and it is not related to estrogen. It presents with high histopathological grade with poor prognosis. The cancer has an underlying atrophic endometrium (69). Apart from this classification, there are still cancers that do not fit into these two categories, in particular endometrioid carcinoma with high histopathological grade (67).

Endometrial cancer is the most common malignancy of gynecological tract in the United States (44). The incidence is relatively high compared with Southeast countries such as

Malaysia where the cancer affects approximately 3.3% of women between the year 2003 to 2005 (70) and the figure increases to 4.6% in 2007 (45) . Among the main races in Malaysia, Chinese has the highest age-standardized incidence rate with 4.5 per 100,000 population, followed by Indians and Malays (45). Failure to control overweight problem, manage chronic anovulation and increased usage of estrogen, are most likely the reason for continued high incidence for this cancer.

Risk factors associated with endometrioid endometrial cancer include old age, unopposed exposure of estrogen as in estrogen replacement therapy, nulliparity and obesity. Also it is seen in diseases associated with high estrogen level, such as polycystic ovarian syndrome and estrogen-secreting ovarian cancer (71). Presence of estrogen increases the proliferative activity of endometrial cells, therefore causing higher chance to cause coding errors and somatic mutations (72). For nonendometrioid type, the risk factors are slightly different, which include additional history of primary cancers such as breast, colorectal and ovarian cancer (73). Combined oral contraceptives can interruptwith the menstrual cycle seems to have good benefits in reducing the risk of endometrial cancer (74). The current treatment for the disease is a combination of surgery with or without an adjuvant chemotherapy consisting of intravenous cisplatin, doxorubicin and cyclophosphamide (75). Diagnosis of this cancer is based on the clinical symptoms with underlying risk factors for endometrial cancer. Postmenopausal women under 50 years old presented with vaginal bleeding were reported to be free from endometrial cancer (76). This was based on the initial screening using transvaginal ultrasound scanning and endometrial biopsy procedure. The patients were follow-up between one to five years (76).

4.1. Molecular basis of endometrial cancer

Endometrial cancer can be divided based on its molecular change. Type 1 or endometrioid endometrial cancer was documented to have PTEN mutation(67).However, a recentcase control study investigating on the single nucleotide polymorphism in several cancer-related genes include PTEN, PIK3CA, AKT1, MLH1 and MSH2 failed to show any association with endometrial cancer (77).Approximately 20 to 40% of this type displayed mircosatellite instability or β-catenin mutations. Additionally, K-ras mutations occur in 15 to 30% of this cancer. Mutations in p53 and E-cadherin were detected in about 10 to 20% of cases and the lowest percentage of genetic alteration is in p16 inactivation. The genetic pattern in type II or nonendometrioid endometrial cancer is slightly different from the endometrioid type. This small percentage tumour comes from mesenchymal cells.The majority of this cancer (80 to 90%) has p53 mutations or E-cadherin alterations (78, 79). The type of cancer rarely contains mircosatellite instability, β-catenin or K-ras mutations (67).Sporadic endometrial cancer with positive microsatellite instability (MIN) was not associated with somatic mutations of mismatch repair genes such as MSH2 and MLH1 (80). Poor association was also observed between positive MIN with mutations in genes with coding region microsatellites repeats (80).

Genetic alterations	Endometrioid or type I	Nonendometrioid or type II
Microsatellite instability	20 - 40%	0 – 5%
K-ras mutations	15 - 30%	0 – 5%
p53 mutations	10 – 20%	90%
PTEN inactivation	35 – 50%	10%
β-catenin mutations	25 – 40 %	0 – 5%
p16 inactivations	10%	40%
E-cadherin alterations	10 – 20%	80 – 90%

Table 2. Molecular changes in both subtypes of endometrial cancer (67)

4.2. Molecular carcinogenesis of endometrial cancer

Endometrial cancer cells has the ability to proliferate without control or able to spread throughout the body following multistep processes.
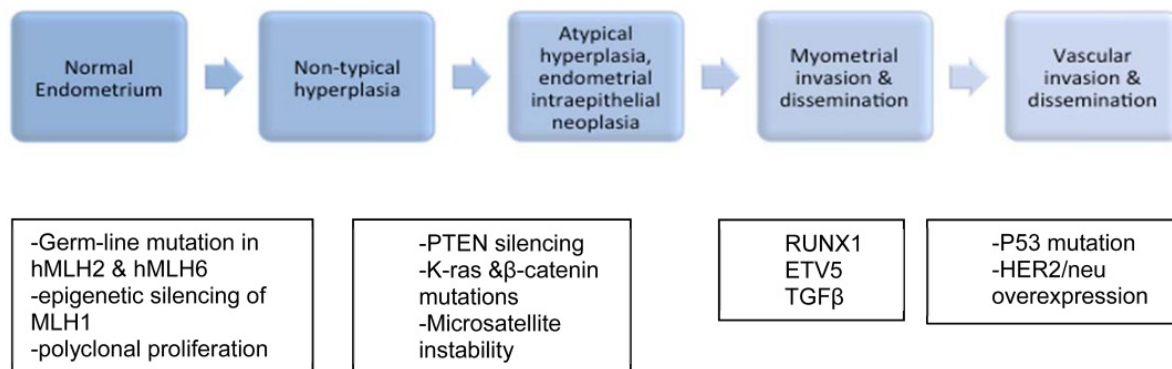

Figure 1. Figure 1: A model of endometrial cancer development. The genetic alterations at the early stage are different from the late stage of endometrial cancer (72).

4.3. Gene expression profiling in endometrial cancer

Earlier studies on the microarray in endometrial cancer tried to discriminate between different histologic types of endometrial cancer using the genomic expression profiling (81). The study analysed 119 endometrial cancer consisted of endometrioid, papillary serous, mixed mullerian tumor and normal cells. The result showed 151 genes that were significantly expressed with at least 2-fold change among endometrioid as compared to papillary serous cancer (P<0.001). Among the genes detected were BUB1, CCNB2 and Myc) (81). Comparing between mixed mullerian tumors and endometrioid revealed 1,132 genes that were significantly different with at least 2-fold change (81). High expression of IGF2 (somatomedin A) was reported in mixed mullerian tumor as compared to endometrioid and papillary serous tumour (81). Our local data showed low expression of IGF2 in endometrioid endometrial cancer compared with normal endometrium (82). Low expression of IGF2 was corresponds to an early stage of endometrial cancer (83). All reported results from these expression profiling studies have concluded that different histologic types of endometrial cancer displayed different expression profiles.

The use of microarray when combined with laser capture microdissection (LCM) tissues has presented reliable results (84). However, the decision whether to use the LCM technique is still relied on the ratio of stromal cells to the surrounding cancer cells. Pathways that are closely related to endometrial cancer were identified after isolation of microdissected cancerous cells was used (82). Among the significant pathways comprise of Wnt-β catenin, insulin action, cell cycle and NOTCH and B-cell pathways (82). The malignant potential of endometrial cancer cells was studied to identify gene signatures of vascular invasion (85). Total of 18-gene signatures were differentially expressed with at least fold change of 2. Among the genes were IL8, MMP3, COL8A1 and ANGPTL4, which were closely related to invasiveness, vascular biology and matrix remodelling (85). Microarray was also used to discrimate between different genetic backgrounds. As an example, molecular profiling was used to differentiate between self-described African-American with self-described Caucasian women (86). The result failed to differentiate the racial group using molecular background. This was probably due to limited sample size to represent the whole population.

5. Molecular signatures in breast cancer

Breast cancer is the most frequent cancer in women in most parts of the world (87). Approximately 1.1 million of women in the world were diagnosed with breast cancer every year and 410,100 died from the disease. Breast cancer can be divided into two main types; ductal carcinoma and lobularcarcinoma (88). The most common type is ductal carcinoma, which starts in the tubes or ducts that move milk from the breast to the nipple. Lobular carcinoma originates from lobules in the breast that produce milk. Breast cancer could become invasive where the cancerous cells may acquire the properties to escape from its primary sites into other tissues in the breast. Noninvasive or also known as 'in situ' indicates that the cancerous cells have not yet invaded other tissues within the breast. There are several grading systems used to classify breast cancer, which include histopathology, grade, stage and receptor status (89). Breast cancer staging uses TNM system, which is based on the size, the spreading and metastatic properties of the tumor to the other organs. There are 3 receptors on the surface as well as in the cytoplasm and nucleus of the breast cancer cells (90). The receptors are estrogen receptor (ER), progesteron receptor (PR) and HER2 receptor (90). Immunohistochemistry technique may be employed to differentiate whether the tumor has positive or negative ER, PR and HER2 receptors (90).

Risk factors of getting breast cancer in women include age and gender. The risk of getting breast cancer is increased in elderly (88). Women are 100 times more likely to get breast cancer compared to men. Genetic factors may also play a role in the development of breast cancer, although it is estimated that only 5-6% of breast cancer are hereditary (91). Mutations in the BRCA1 and BRCA2 genes account for 80% of hereditary breast cancer (92). Patient's positive for BRCA1 and/or BRCA2 may have 50% to 80% lifetime risk of developing breast cancer and 15% to 65% risk of developing ovarian cancer (92, 93). Other risk factors are high-fat-diet, alcohol intake, environmental factors such as tobacco smoking and radiation (94).

The diagnosis of breast cancer is based on the microscopic analysis of breast biopsy, mammography and clinical breast exam (95). However, if the test is inconclusive, then Fine Needle Aspiration and Cytology (FNAC) may be used (96).Stage 1 breast cancer is treated with lumpectomy to remove a small part in the breast and usually have high prognosis. Stage 2 and 3 cancers are treated with lumpectomy or mastectomy, chemotherapy and radiation and usually have poor prognosis and high risk of recurrence. Stage 4 has poor prognosis and is treated by various combination of all treatments. Drugs used to treat breast cancer include hormone-blocking therapy for ER+ patients (tamoxifen, aromatase inhibitors), chemotherapy (cyclophosphamide and doxorubicin) and monoclonal antibodies (trastuzumab) for HER2+ breast patients (97).

5.1. Genome wide association study (GWAS) in breast cancer

A single nucleotide polymorphism (rs2046210, A/G allele) at 6q25.1 was identified in Chinese women. In a pooled analysis study performed in the East Asian, European, and African ancestries, this variant was also found to be associated with breast cancer risk in Chinese women (OR=1.3), Japanese women (OR=1.31), European (OR=1.07), and American women (OR=1.18) (98). However, there was no association observed in African American women (OR=0.81). This variant was found to be associated with increased breast cancer risk in all Chinese in Tianjin, Nanjing, Taiwan and Hong Kong. This was also in agreement with three studies conducted in Japanese women (Nagoya, MEC and Nagano) as well as studies performed in European women (NBHS, CBCS and LIBCSP). A putative functional variant, rs6913578 was identified at 1,440 downstream of rs2046210, which was associated with breast cancer risk in Chinese (r^2=0.91) and European ancestry (r^2=0.83), but not in Africans (r^2=0.57). Genes located at rs2046210 are *PLEKHG1*, *MTHFD1L*, *AKAP12*, *ZBTB2*, *RMND1*, *C6Orf211*, *C6orf97*, *ESR1*, *C6orf98*, *SYNE1* and *NANOGP11*. *In vitro* functional analysis on rs6913578 altered luciferase reporter activity hence may influence the DNA binding protein interactions, which subsequently lead to alteration of their neighboring genes expression. Electrophoretic mobility shift assay confirmed that the C allele of rs6913578 alter the DNA-nuclear protein interaction and could modify the expression of neighboring genes.

There was an association between an increased breast cancer risk with rs9397435 at the 6q25.1 locus in European, Chinese and African populations. This variant was located at 2,854 bp downstream of rs2046210 and 1,414 downstream of rs6913578. However, this variant was weakly correlated with rs2046210 in Europeans (r^2=0.087) and African (r^2=0.039) (99). Turnbull and colleagues conducted a GWAS in 3,659 European ancestry cases and 4,897 controls. They found that SNP rs3757318, which was located at 200kb upstream of ESR1 and 34,253bp of upstream of rs2046210 has the most significant association with breast cancer risk (OR=1.21). It was strongly correlated with rs2046210 in Chinese populations (r^2=0.48) but weakly correlated in Europeans (r^2=0.181) (100).

In Ashkenazi Jews population, Gold and colleagues performed three phases of GWAS in 249 familial breast cancer cases and 299 controls. In the first phase, they compared the allele frequencies of 150,080 SNPs in 249 high-risk, BRCA1-BRCA2 mutation-negative AJ familial

cases with control cases. In phase II, 343 SNPs were genotyped from 123 regions, which were most significantly associated with breast cancer including 4 SNPs in *FGFR2* region in other sets of 950 consecutive breast cancer cases. Major associations were replicated in third independent set of 243 breast cancer cases and 187 controls. The results showed a significant association at rs1078806 in the *FGFR2* region of chromosome 6q22.33 with OD=1.26 for all cases combined. Candidate genes in this locus such as *ECHDC1* and *RNF146*, which encode for mitochondrial fatty acid oxidation and ubiquitin protein ligase were among the known pathways in the pathogenesis of breast cancer (101). It is well known that results reported from GWAS could not be applied across all ethnicities. This is not surprising since most all variants are tagging SNPs, therefore they exist differently in the genetic make-up of different ethnic groups. Hence, it is important to determine the SNPs in breast cancer or any particular diseases in different populations to identify the risk of developing the disease in an individual.

5.2. Gene expression profiling in breast cancer

A research done to study bimodal gene expression profiles in breast cancer using 5 studies that used different microarray platforms including cDNA arrays, Affymetrix and Agilent (102). Bimodality is a conditional expression property of a particular gene and is associated with certain physiological conditions such as disease state and normal. They found 866 bimodal genes shared across all platforms. These genes were enriched in breast cancer-associated genes and involved in pathways related to carcinogenesis for example: *ERBB2*, *ESR1*, *CEACAM5* and *AR*. They also examined the close neighbor group and the analysis showed that 15 out of 23 bimodal genes were known and have been reported as breast cancer associated genes. These include *TCAP, PSMD3, GRB7 and CXCL10* (PMC2822536).

Microarray was also used to classify the differential gene expression in ER+ve and ER-ve breast cancer patients. A study showed that 67 genes were overexpressed in ER+ve tumors while 17 were overexpressed in ER-ve breast cancer. *ADCY1, ACOT4 AR, ATP2A3, DNAJA4* were examples of genes that overexpressed in ER+ve breast cancer. An example of genes that were overexpressed in ER –ve were *ACN9, EGFR, LYN* and *MALL* (103).

Gene expression profiling of tumor-associated stroma in breast cancer showed large changes during cancer progression (104). In this study, laser capture microdissection was used to dissect the normal epithelium, stroma, tumor epithelium and tumor-associated stroma samples followed by microarray and gene ontology analyses. Tumor-associated stroma undergoes massive changes in the expression profile of genes composed of the extracellular matrix, matrix metalloproteases and cell cycle-related protein. An increased in the mitochondrial ribosomal proteins and decreased in cytoplasmic ribosomal proteins were also observed in both, the tumor epithelium and stroma. The changes in expression profiles of the tumor-associated stroma were somewhat similar to tumor epithelium, which indicated that the tumorigenesis occured even before the tumor cells invaded into the stroma.

Gene expression profiling using whole genome oligonucleotide microarrays to catalog molecular variation in 52 widely used breast cancer cell lines. The cell lines were divided

into different categories including luminal with ER positive, basal and ER-ve, which subdivided into basal A (established at UT Southwestern including 2 BRCA1 mutant lines) and basal B (non-tumorigenic lines and several highly invasive cell lines). They identified 80 loci of high level of amplification in 35 different cell lines. These include increased expression of known oncogenes involved in breast cancer, for example *MYC* (8q24), *CCND1* (11q13) and *ERBB2* (17q12). Gain or losses resulting in increased or decreased expression of oncogenes or tumor supressor genes, which subsequently led to breast cancer. Using DR-Correlate, 3,511 genes were differentially expressed and correlated significantly with altered gene copy number (FDR<0.05). In total, 487 genes were resided in loci of high-amplitude CNA including known breast cancer genes such as *EGFR, FGFR1, ERBB2, PPMID* and *ZNF217*. In addition, several genes involved in oncogenesis such as cell proliferation, survival, migration/invasion, ER-signaling, maintenance of genome integrity were also upregulated in cancer cell lines. These include *E1F3H, CDC6, GAB2* (cell proliferation), *MCL1, APIP, MAP3K3* (survival), *ADAM9, CDD4* (migration/invasion), *MUC1, NCOA3* (ER-signaling), *RAD21, RAD9A* and *RAD51C* (maintanence of genome integrity) (105).

Gene expression profiling study was carried out on peripheral blood cells for an early detection of breast cancer in 121 females referred for mammography. Genome Survey Microarrays v2.0 that contains 32,878 probes representing 29,098 genes was used to determine the differentially expressed genes in breast cancer compared to normal. Genes that expressed higher in blood of breast cancer patients were *EEF1G, RPL14, RPLL15* (translation), *ATP5E, ETF1, ATP6V0B* (cellular biosynthetic process), *TIRAP, DEFA3* and *ANXA1* (response to external stimulus). Several genes involved in cellular lipid metabolic process, steroid metabolic process, catecholamine metabolic process and phenole metabolic processes were downregulated in breast cancer compared to normal control. These include *HDC, PEMT, HEXA, ACAT* and *SULT1A4* (106).

6. From lab to bedside: FDA approval

Advances in genomic research resulting in new molecular tools that serves as prognostic and predictive markers in cancer treatment. Particularly in breast cancer, surgeons know that early detection is one of the keys to successful treatment. If breast cancer is caught early, the tumor can be surgically removed and with an appropriate treatment, most patients can recover. However, within 5 to 10 years, 30% increase number of patients with early stage breast cancer develops metastases. The identification of patients with high risk of distant recurrence is essential for systematic adjuvant therapy to be most effective. At the same time, adjuvant therapies such as chemotherapy and hormonal therapy (e.g. Tamoxifen or aromatase inhibitors) may reduce the risk of distant metastases by approximately one-third for some patients. It is estimated that more than 70% of patients receiving such therapy may have survived without it –and may have safely avoided the harmful side effects (107-109).

Commercially available multigene molecular tests such as Oncotype DX® (Genomic Health, USA) and MammaPrint® (Agendia, Netherlands) have revolutionized the predictive and prognostic tools in clinic. Using the patients' own genetic expression patterns, it can provide

clinicians with more information on the treatment outcomes of using chemotherapy, endocrine therapy or combination therapies by stratifying the risk of recurrence for patients. Oncotype DX® and MammaPrint® provide clinical judgment as opposed to laboratory results that requireinterpretation by a clinician. Moreover, the algorithm used to reach this judgment is proprietary and thus inaccessible to the clinician. Therefore the arrival of the first generation of multigene molecular test involved a need for a paradigm shift in the configurations of persons and tools that marry genomic techniques to market, legal, and regulatory strategies in ways that reframe conceptions of risk, diagnosis, prognosis, therapy, discovery, utility, and validity. In addition, regulatory bodies need to handle these new advances without sacrificing patient's safety. These first generation multigene molecular tests are considered the first regulatory-scientific hybrid products (110).

The Oncotype DX® is a multigene panel which has been clinically validated to predict the risk of recurrence for those women with early stage (I, II, IIIa) invasive breast cancer that are estrogen-receptor positive (ER$^+$), human epidermal growth factor receptor negative (Her2$^-$), lymph node negative or positive, and predict who may or may not significantly benefit from adjuvant chemotherapy. While MammaPrint® analyzes 70 genes from an early-stage breast cancer tissue sample to determine if the cancer has a low or high risk of recurrence within 10 years after diagnosis. They claimed to be the first and only FDA-cleared IVDMIA breast cancer recurrence assay in their official website, http://www.agendia.com/pages/ mammaprint/21.php (110). The researchers at the Netherlands Cancer Institute (NKI) who discovered it, established a company to commercialize it as a test (111). Oncotype DXbegan as a commercial platform; the company (Genomics Health) that produced it did not discover a signature but rather constructed it by asking users at every step what clinical question they wanted the signature to answer and what data would be credible in that regard. The test has been designed to minimally disrupt existing clinical workflows (110, 112). MammaPrint requires a change in pathologists' and clinicians' routines in terms of specimen storage. MammaPrint requires specimen to be stored in RNARetain®, a proprietary RNA storage liquid instead of the standard FFPE block. Breast cancer classification was based on genomic signature instead of histopathology diagnosis as well as clinical judgement on the decision for chemotherapy treatment (113). Thus, while these two trials signify a new departure for clinical cancer trials on a number of levels – they both incorporate new models of interaction between biotech companies and public research. They also aim to establish the clinical relevance of genomic markers and also embody a different socio-technical direction. One attempts to accommodate established routines, while the other openly challenges prevailing evidential hierarchies and existing biomedical configurations (110).

The legal statute of the USA gives the US Food and Drug Administration (FDA) the power to regulate drugs and devices, with the multigene molecular tests fall under the less rigorous medical devices statute. The FDA has traditionally exercised 'enforcement discretion' by leaving the actual performance of 'in-house' tests to be regulated by a different mechanism defined by the Clinical Laboratory Improvement Amendments (CLIA). It is a set of federal regulatory standards that falls under the authority of the Centers for Medicare and Medicaid Services (114). The intention was to ensure the reliability and

accuracy of clinical laboratory testing. FDA regulators have suggested the development of translational medicine tests such as Oncotype DXand MammaPrint might constitute an entirely new regulatory category. In 2006 and 2007, the FDA published two versions of 'Draft Guidance', signaling the Agency's inclination to step in and take direct responsibility for the novel test category. In 2007, MammaPrint was submitted to the FDA and successfully obtained FDA clearance after only 30 days. An 'FDA cleared' button promptly appeared on all commercial MammaPrint material (http://www.agendia.com/pages/mammaprint/ 21.php). Given the non-binding nature of the FDA draft guidance, Genomic Health chose not to pursue this regulatory route. Instead they try to gain 'official' recognition from the clinicians via inclusion in the clinical practice guidelines of professional oncology organizations. The company viewed the pursuit of FDA clearance as much more costly and time-consuming than simply lobbying professional organizations of clinicians – many of whom the founders already knew through their previous works at Genentech (110). The American Society of Clinical Oncology (ASCO) included Onco*type* in its 2007 guidelines and the US National Comprehensive Cancer Network (NCCN) followed suit in its 2008 guidelines.

6.1. Study design of the multigenes panel

In cancer epidemiology, both retrospective case – control studies and prospective cohort studies are observational, rather than experimental, studies. Neither type of study involves random assignment of exposure hence; observed associations between exposures and disease do not provide as strong a basis for claims of causality as in experimental studies. The most serious limitation of epidemiological studies is their non-experimental nature, not whether they are retrospective or prospective. In therapeutics, many retrospective analyses are also non-experimental, with treatment selection based on patient factors and referral pattern rather than on randomization. Such studies are also often conducted without a written protocol and are unfocused, with numerous patient subsets and endpoints compared without control for the overall chance of a false-positive conclusion. In contrast, prospective randomized clinical trials contain internal control of treatment assignment, careful and proscribed data collection (including outcomes and endpoints), and a focused analysis plan that is developed before the data are examined (112).

Many biomarker studies are conducted with convenience samples of specimens, which just happen to be available and are assayed for the marker. They have not prospectively determined subject eligibility, power calculations, marker cut-point specification, or analytical plans. Such studies are more likely resulting in highly biased conclusions and truly deserved to be pejoratively labeled as "retrospective." However, if a "retrospective" study is designed to use archived specimens from a previously conducted prospective trial, and if certain conditions are prospectively delineated in a written protocol before the marker study is performed, it might be considered as a "prospective – retrospective" study. Such a study should carry considerably more weight toward determination of clinical utility of the marker than a simple study of convenience, in which specimens and assays were happened to be available. Multiple studies of different candidate biomarkers based on archived tissues from the same prospective trial would present a greater opportunity for false-positive

conclusions than a single fully prospective trial focused on a specific biomarker. Consequently, independent confirmation of findings for specific biomarkers in multiple prospective – retrospective study (115).

6.2. Oncotype DX breast cancer assay

The Oncotype DX® analyzes the expression of 21 genes (16 cancer-related and 5 reference genes) within a tumor to determine a recurrence score (RS) using reverse transcription PCR (RT-PCR) in formalin-fixed, paraffin-embedded (FFPE) breast cancer tissue samples. In the earlier stage, the researchers has to show that RNA extracted from FFPE tissues could match fresh tissue results in terms of producing a high concordance in the RT-PCR results (116, 117).To interpret the result, Oncotype DX test results assign a Recurrence Score (RS) – a number between 0 and 100 – to the early-stage breast cancer or DCIS as stated below:

- RS lower than 18: The cancer or DCIS has a low risk of recurrence. The benefit of chemotherapy for early-stage breast cancer or radiation therapy for DCIS is likely to be small and will not outweigh the risks of side effects.
- RS between 18 and 31: The cancer or DCIS has an intermediate risk of recurrence. It's unclear whether the benefits of chemotherapy for early-stage breast cancer or radiation therapy for DCIS outweigh the risks of side effects.
- RS greater than 31: The cancer or DCIS has a high risk of recurrence, and the benefits of chemotherapy for early-stage breast cancer or radiation therapy for DCIS are likely to be greater than the risks of side effects.

The RS corresponds to a specific likelihood of breast cancer recurrence within 10 years of the initial diagnosis, as well as response to adjuvant treatment. Using recurrence score, it may be possible for healthcare providers and patients to determine whether adjuvant chemotherapy is needed following primary therapy for breast cancer (118, 119).

i. NSABP Study B-14

The Oncotype *DX* was developed and clinically validated on the basis of a retrospective analysis of the existing material from two randomized clinical trials (NSABP-B-20 and NSABP-B-14). The signature is based on the expression of genes that are associated with proliferation, ER signaling, HER2, and invasion (118). The 21 multigene chosen were always at the top of the list in published literature. The developers used the samples from 447 patients as the 'discovery' or 'training' set to select the 21 genes eventually included in the Onco*type* test. Company researchers then applied an algorithm to the results of the tests and developed the aforementioned RS score. They believe the score is one of the strengths of the Onco*type* test: as a single number on a continuous 0–100 scale and not a category (that is, yes/no, good/poor). It is supposed to provide clinicians with 'useful' information as a basis on which to act, while preserving clinical decision-making as a clinician's prerogative, since by not providing a categorical answer it does not entail a specific intervention (110). Results from this study demonstrated that Onco*type DX* is an accurate and reliable predictor of breast cancer recurrence. (120). The

study also concluded that the RS has been validated as quantifying the likelihood of distant recurrence in tamoxifen-treated patients with node-negative, estrogen receptor-positive breast cancer (118).

ii. NSABP Study B-20

About 668 samples of cancer tissue from a clinical trial called NSABP B-20 ("A Clinical Trial to Assess Tamoxifen in Patients with Primary Breast Cancer and Negative Axillary Nodes Whose Tumors Are Positive for Estrogen Receptors) were used to show that Onco*type DX* can predict chemotherapy benefit (119). The study concluded that the RS of the assay not only quantifies the likelihood of breast cancer recurrence in women with node-negative, estrogen receptor-positive breast cancer, but also predicts the magnitude of chemotherapy benefit (118).

iii. Kaiser Permanente study

A large clinical study of 234 cases and 631 controls available for pathology studies (after screening of 4964 patients) conducted by Kaiser Permanente confirmed in a community setting that Onco*type DX* helps to predict the likelihood of breast cancer survival at 10 years (121). The primary objective of this study was to determine whether the proportion of patients who were free of a distant recurrence for more than 10 years after surgery was significantly greater in the low-risk group than in the high-risk group. The second primary objective was to determine whether there was a statistically significant relation between the RS and the risk of distant recurrence. The cutoff points were prespecified to classify patients into the following categories: low risk, intermediate risk and high risk. The cutoff points were chosen on the basis of the results of NSABP trial B-20. The study concluded that in a large, population-based study of lymph node-negative patients not treated with chemotherapy, the RS value was strongly associated with risk of breast cancer death among ER-positive, tamoxifen-treated and -untreated patients.

iv. SWOG 8814 study

SWOG-8814 was a randomized phase III clinical trial of 1,477 postmenopausal women, all of whom had estrogen receptor-positive (ER+) breast cancer that had spread to the axillary lymph nodes. All women in the trial got daily tamoxifen for up to five years, longer than the standard therapy for treating ER+ breast cancer. One arm of 361 patients got only tamoxifen. The rest got tamoxifen plus a three-drug chemotherapy regimen of cyclophosphamide, Adriamycin®, and 5-fluorouracil, a combination known as CAF. Investigators retrospectively analyzed tumor specimens from this trial using the Oncotype *DX®* in 367 women with ER-positive, mainly tamoxifen-treated lymph node-positive, the RS assay quantified the likelihood of breast cancer recurrence and also predicted the magnitude of chemotherapy benefit (122).

v. Oncotype DX TAILORx Trial

Following the development of the specialized translational research program from National Cancer Institute (NCI), the Program for the Assessment of Clinical Cancer

Tests (PACCT) launched the TAILORx trial (123).Since the validation of the Oncotype DX Breast Cancer Assay Recurrence Score were able to clearly show that the multigene panel were able to predict chemotherapy with hormonal treatment benefit for patients with high Recurrence Score while patients with low Recurrence Score do not benefit from chemotherapy. However as high as 37% of patients fall into the intermediate range, which do not show a clear outcome of the benefit of chemotherapy (122). A randomized prospective clinical trial is currently ongoing to further validate a group of node-negative, ER+ breast cancer patients with a RS in the intermediate range, which is known as **T**rial **A**ssigning **I**ndividua**L**ized **O**ptions for Treatment (**Rx**) TAILORx conducted by the North American Breast Cancer Group (http://www.cancer.gov/clinicaltrials/noteworthy-trials/tailorx). Since 2006, the trial enrolled 10,000 patients (of which 4500 were to be in the randomized arm) in 900 participating centers (110). Patients with mid-range RS will be randomized for chemotherapy while patients with low and high RS will not be randomized as the outcome has been clearly defined in previous studies.

6.3. Recommendation of use as tumor marker

Because Oncotype *DX* was able to achieve level II evidence to support it's prognostic role, Oncotype DX has received approval from the American Society of Clinical Oncology (ASCO) in the2007 guidelines (124). It was included in the National Comprehensive Cancer Network (NCCN) 2008 guidelines (Breast Cancer version 1.2011 [http://www.nccn.org].) as an option to evaluate prognosis and as a complement to clinicopathological features to predict response to chemotherapy for patients with ER-positive, node-negative breast cancer. None of the microarray-based prognostic signatures has been endorsed by these professional bodies.

6.4. MammaPrint

MammaPrint (initially known as the 70 Gene Amsterdam Signature) was originally developed as an academic/scientific endeavor using whole genome microarray technology. The objective was to develop a gene expression signature that could accurately identify early stage breast cancer patients who were either at high risk or at low risk of recurrence and, therefore, enable more individualized treatment. The MammaPrint investigators from the NKI-AVL in collaboration with the Rosetta Inpharmatics (a Seattle company) procured and analyzed 78 tumors with the whole-genome microarray. Out of the 25,000 genes in the human genome, 231 genes were selected according to its association with the disease outcome. Further bioinformatic analysis using 2-D cluster analysis followed by a leave-one-out cross validation procedure produced 70 critical genes that were shown to correlate best with the likelihood of distant recurrence. These 70 genes affect all steps known to be important for metastasis including cell cycle regulation, angiogenesis, invasion, cell migration and signal transduction (111).The resulting 70-gene signature profile classifies tumors as either high risk or low risk of recurrence. If it is used in conjunction with other risk factors, it helps to identify patients who will benefit from the adjuvant therapy. The 70-

gene signature was constructed as a dichotomy as the discussions between the research team and clinicians, who insisted that the main goal of the test should be to avoid overtreatment of the disease. To accomplish this end, the low-risk group had to be defined inclusively. At the same time, the test developers felt that clinicians expected a clear answer (good/poor signature) from the test, hence the dichotomy (111). This position, once again, contrasts with the Genomic Health's decision to report their Oncotype DX data analysis as a continuous variable that leaves room for clinical judgment (110).

With this 70-gene signature, further validation was needed on a larger, independent patient population. The primary validation was thus carried out via another retrospective study that used samples from 295 patients held in the same NKI bio-bank. The first validation for the 70-gene signature was undertaken in a series of 295 consecutive women with breast cancer. The proportion of patients who remained free from distant metastases at ten years was 87% in the low-risk group and 44% in the high-risk group. The profile was a statistically independent predictor and added to the power of standard clinico-pathologic parameters (125). A research network team called TRANSBIG, an abbreviation for "Translating molecular knowledge into early breast cancer management: building on the BIG network for improved treatment tailoring", used the 70-gene signature as retrospective study in 2006 using 307 pa54tient samples from five European institutions. The results showed that the proportion of patients who remained free from distant metastases at 10 years was 90% in the low-risk group and 71% in the high-risk group. The 70-gene signature was found to provide prognostic information more than what could be determined from patient age, tumor grade, tumor size, and ER status in a population of lymph node negative patients without adjuvant chemotherapy (113). Although they initially favored licensing the technology, the NKI team found no viable taker. So, in 2003 the original researchers, in consultation with the NKI board of directors, established a spin-off company using private venture capital and European Union (EU) funding, and convinced the director of oncology at a leading diagnostic company, Agendia, for Amsterdam Genetic Diagnostics Amersham (110). The Agendia team had a signature but they did not have a test. In other words, it was not immediately obvious how to convert the 70-gene signature into what eventually became MammaPrint, a 'high-throughput diagnostic test' (126). The original signature had been developed using microarrays containing 25,000 oligonucleotides, a highly impractical platform for routine use. The company therefore developed a customized microarray containing a reduced set of probes, whose production was entrusted to Agilent, to whom Rosetta had, in the meantime, sold its technology.

The TRANSBIG Consortium performed another independent validation study of 302 adjuvantly untreated patients with at least ten years of follow-up. For the NKI researchers, the problem was less RNA extraction than the microarray analysis itself. Compared with RT-PCR, microarray analysis was a relatively novel, non-standardized technology and as such, it raised a number of logistic and statistical challenges (127, 128). As a result, in addition to the validation studies of the signature per se, researchers conducted a number of other studies to show that sample collection for the test (as distinct from the centrally

performed test itself) was feasible and reproducible in community-based settings (129). Additional studies demonstrated that the MammaPrint classifies greater than 95% of ER-negative cancers as poor prognosis and there was a strong correlation between 70-gene signature-defined poor prognosis and high histological grade (130, 131). Furthermore, the studies demonstrated that the 70-gene signature would outperform the current methods based on clinicopathological parameters for chemotherapy use.

One study revealed that MammaPrint validates in older American breast cancer patients (132). While another study demonstrated that MammaPrint has strong prognostic value in patients with 1-3 positive lymph nodes (133). With more than 14,000 patient results reported to date, the technical robustness and reliability of MammaPrint is well established. MammaPrint is a considerable a step forward in the advancement of personalized cancer treatment. Several other prognostic signatures including the 76-gene signature (134, 135) and genomic grade index (136-139) were also shown to be independent predictors for the cancer outcomes.

i. MicroarRAy PrognoSTics in Breast CancER (RASTER) study

To evaluate whether the prognostic signature is suitable for the use in clinical practice, the MammaPrint was used to assess feasibility of implementation of the test as a diagnostic test in community hospitals in the Netherlands. The study aimed to test the effect of the signature on the use of adjuvant systemic treatment; proportion of patients with "poor" versus "good" prognosis in a series of unselected patients with node-negative breast cancer; and finally to examine the concordance between risk predicted by the prognosis signature and risk predicted by commonly used clinicopathological guidelines. The findings of this study show that implementation of the 70-gene prognosis signature as a diagnostic test is feasible in community hospitals in the Netherlands (129).

ii. MINDACT Trials

MammaPrint is currently being tested in the MINDACT (Microarray In Node-negative and 1-3 positive lymph-node Disease may Avoid ChemoTherapy) trial (140). This is to determine whether this signature can actually replace clinicopathological parameters for the identification of patients who could be spared from the use of chemotherapy. The more 'confrontational attitude' of the MINDACT leaders toward traditional clinico-pathological tools has resulted in a very different trial design compared to the Oncotype DX TAILORx Trial.In the MINDACT trial, women recruited into the trial are assigned to high- and low-risk categories using both standard clinical-pathological features and the 70-gene MammaPrint test results. An open-access computer program, Adjuvant! Online, developed in the US and widely used by breast cancer clinicians to estimate the outcome in terms of relapse and survival with or without chemotherapy. By confronting the predictions of MammaPrint and Adjuvant! Online, the trial directly compares between these two prognostic tools: women whose Adjuvant! Online and MammaPrint results are discordant (when clinical/pathological features indicate high

risk of recurrence and when MammaPrint indicates low risk, or vice versa) are then randomized for chemotherapy.

6.5. Conclusion

Based on the recommendation by the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group, the general consensus was that retrospective study of samples and data from prospective studies were insufficient, although these studies were superior to studies using 'convenience samples', such as those contained in a general-purpose bio-bank. However, potential for patient selection bias cannot be excluded (141), therefore the working group recommend prospective studies such as TAILORx and MINDACT as the gold standard for testing the value of a multigene molecular test such as Onco*type* or MammaPrint. From their review on these multigene molecular tests, they found insufficient evidence to make a recommendation for or against the use of tumor gene expression profiles to improve outcomes indefined populations of women with breast cancer. The working group found preliminary evidence on the potential benefit of the Oncotype *DX* testing results to some women who face decisions about treatment options (reduced adverse events due to low risk women avoiding chemotherapy) but could not rule out the potential harm for others (breast cancer recurrence that might have been prevented). The evidence is insufficient to assess the balance of benefits and harms of the proposed uses of the tests. The working group therefore encourages further development and evaluation of these technologies. There are still limitations that prevent these multigene molecular test such as the Oncotype *DX*, MammaPrint and other genomic prognostic markers from replacing the microscope for diagnosis, prognosis and treatment of an early breast cancer. However, additional important clinical information from this test has added to traditional histology and IHC determination of ER, PR and HER2 in terms of prognostic and predictive power.

7. Challenges associated with the clinical translation

Advances in laboratory and clinical science has propelled to a transitional period, which requires a redefinition of biology, genomics, and medicine in relation to one another. "Molecular gene signatures" is a new buzz word within the field of personalized medicine in the treatment of breast cancer (111, 118), thyroid cancer (142), endometrial cancer (143), ovarian cancer (144) and other cancers as well. However, the road from the scientific discovery of molecular signatures associated with cancer until it can be translated to clinical application is long and arduous. A recent review on the current status of translational research in cancer genetics has analyzed the extramural grant portfolio of the National Cancer Institute (NCI) from Fiscal Year of 2007. From the study, the funded grants and publications were classified as follows: T_0 as discovery research; T_1 as research to develop a candidate health application (e.g., test or therapy); T_2 as research that evaluates a candidate application and develops evidence-based recommendations; T_3 as research that assesses how to integrate an evidence-based recommendation into cancer care and prevention; and T_4 as research that assesses health outcomes and population impact (145). An "explosion" in gene expression research during the

last few years has already led to the development of several genetic classifiers in the genomic discovery (T_0) stage and T_1 stage (which bridges discovery to candidate health application, or "bench to bedside"). However, less genomic research was conducted and published in T_2 and above, with only 1.8% of the grant portfolio and 0.6% of the published literatures in these categories. In addition to discovery research in cancer genetics, a translational research infrastructure is urgently needed to methodically evaluate and translate gene discoveries for cancer care and prevention (146, 147).

	POPULATION ONCOLOGY	TRANSITION	PERSONALIZED ONCOLOGY
Screening	Population-wide risk reduction	Population-wide approaches modified for at-risk subpopulations	Individualized risk estimation and programs adapted to individual risk
Diagnosis	Organ-of-origin-/histology-based	Organ of origin, histology, and some molecular markers	Primarily molecular marker-based
Staging	Anatomic extent of disease	Anatomic extent with some molecular risk profiling	Primarily molecular risk-based
Treatment determination	Typically organ-of-origin- and stage-based	Organ-of-origin- and stage-based with some implementation of molecular markers	Primarily molecular marker-based
Assessment intervals	Based on clinical evaluation/ examination findings	Based on routine interval imaging	Early, frequent serial assessments by imaging, circulating tumor cells, and other marker assessments
Early phase clinical trials	Oriented to maximum tolerated dose	Oriented to "optimum biologic dose"	Determine range of tolerable and active doses
Mid-phase clinical trials	Histology and prior treatment-based eligibility; typically single-arm, noncomparator trials	Histology and prior treatment-based eligibility; some marker-based screening; some randomized controlled trials	Some trials, histology, and prior treatment-based eligibility with rapid, serial assessments; many with eligibility restricted to tumor marker subsets

Table 3. Systemic therapy options for the treatment of invasive breast cancer in the adjuvant and advanced disease settings. Among solid tumors, breast cancer treatment arguably has made some of the greatest advances during the previous 3 decades (148). Advances in laboratories and clinical science have propelled us into the current transitional period and how clinical trials must evolve to lead us into the era of personalized oncology (148)

7.1. Challenges of gene expression profiling studies

In order to understand challenges associated with the clinical translation of molecular gene signature obtained from microarray studies, we must understand the challenges and limitations of gene expression profiling. Although gene expression profiling seems to have value in the discovery of molecular markers for potential use in diagnosis or as a therapeutic target, translating this technology into genomic medicine is still a work in progress. For a better understanding in terms of strengths and limitations of gene expression profiling techniques, we need to understand biological, technological, statistical, and informatics challenges and caveats.

7.2. Biological challenges

A microarray experiment presents a snapshot of the gene expression of the biological system that is dynamic and constantly changing at a given time point, which may not provide the complete picture or accurately depict of what is really happening at cellular level. Thus, the presence of mRNA does not explicitly mean that it was just synthesized. Likewise, the

inability to detect an unstable transcript may be due to its high degradation rate (149). The expression of some genes ("housekeeping genes") is thought to be more stable, and these genes are often used as controls for the normalization of expression levels of other genes. However, the expression of traditionally used controls such as ribosomal RNA genes, also changes across different tissues and experimental conditions making it difficult to select "gold standards" (150). Sampling issues such as biopsy method (151), contamination from neighboring tissues may seriously affect in different expression profiles as microarray technology is very sensitive to such variations (152). RNA quality is a critical issue in genome-wide analysis of gene expression. RNA is less stable than DNA and care should be taken and adequate protocols followed to preserve the quality of biological material. This is particularly important in clinical setting. Another limitation in prognostic or predictive markers from gene expression profiling is that microarray covers only part of the whole picture. Most of cellular functions are performed by proteins and physiological changes can be modulated by not only changes in protein levels but also by protein modifications such as glycosalation, methylation, acethylation, and phosphorylation. These modifications could change protein conformation and lead to changes in activity, which is not detectable by gene expression profiling (152).

7.3. Technological challenges

All of the microarray platforms available in the market are proprietary, a general concern for the inter-platform variability in the gene expression profiles has been addressed by the MicroArray Quality Consortium II (MAQC) (153). Despite the high variability in gene expression attributed to differences in microarray platforms, studies have demonstrated that reproducibility across platforms can be dramatically improved when standardized protocols are implemented for RNA labeling, hybridization, data processing, data acquisition, and data normalization. When these technical variables are standardized, different microarray platforms can produce comparable outcomes (154, 155). Nevertheless, the results from comparison across different platforms can be misleading and should be interpreted with great caution (156). Technicalities of the microarray platforms deals with binding efficiency of labeled target to the respective probe as well as technical variation during experiments also may affect the reproducibility of the gene expression profiles (152). With regards to prospective experiments, the uniformity of experimental conduct will help to minimize potential bias and thus improve the validity of a study. The establishment of the Microarray Quality Control (MAQC) project in 2005 to develop procedural guidelines and quality control metrics in the first phase and the second phase aims to evaluate various data analysis method and predictive models (153). One of the serious problems has been a wide diversity of data formats used in microarray experiments. As a result, the Microarray GeneExpression Database Society (MGED) was created in 1999 to develop a common standard for data input and reporting that could be shared among scientists in the microarray field. In 2001, the MGED created the Minimum Information About a Microarray Experiment (MIAME) guidelines, which serve as a template for researchers to report an adequate description of how microarray data were obtained (157).

7.4. Statistical and and bioinformatic challenges

The experimental design of the microarray studies is of paramount importance, as it should have a clear goal and a specific hypothesis to test. In the design of a microarray experiment, all potential sources of variation should be taken into account to avoid any systematic bias. Researchers should adhere to the sound principles of study and match the experimental variables of cases and controls to the fullest extent possible. It is important to select biologically homogenous sample populations, balancing a design with respect to all factors that can confound results among the comparison groups, and handling samples uniformly through the course of the entire experiment when designing a microarray study (158). Randomization of samples will assure baseline equality between the groups being compared. Violation of these principles will lead to biased results and can cause a loss in power. It should be pointed out that statistical analysis of data couldn't solve fundamental problems of study design. Significantly, the validity of gene expression profiles depends on the characteristics of samples and selection bias, eligible criteria of participation and other confounding factors. An adequate sample size is necessary to achieve sufficient power to demonstrate significance of findings, especially in microarray studies where thousands of genes are tested simultaneously (159). Appropriate preprocessing of microarray data, known as "normalization" prior to analysis is critical for identifying differentially expressed genes. Normalization attempts to remove variability among chips and other systematic biases that are unrelated to biological variation so that a meaningful biological comparison can be made. Transformation is used for multiple purposes, including stabilizing variance in data so that underlying assumptions required for the statistical analysis method are met. Although it is expected that the choice of a preprocessing procedure does not affect the core results of microarray data, different normalization and/or transformation methods may result in different outcomes (160).

Application of appropriate analysis methods to the microarray data, for example classification and cluster analysis are typical analytical approaches to categorized microarray data into manageable classes. However, there is no standard 'method' to how to best analyze the genomic data and it's very tempting to present / published the best-looking result, leading to biased evaluation of the statistical prediction rule. Another issue of classification is "overfitting", which occurs when a classifier is made to perfectly fit a set of data that was used in the model development, but has no discriminatory power so that the results cannot be reproduced in a set of completely independent samples (161). This may lead to insufficient evidence of accuracy and reproducibility of multigene signature from gene expression profiles for clinical use, although it showed initial promising and reproducible results in class discovery studies and preclinical analysis (162). An adequate sample size is essential for any cross-validation technique to be effective. Another significant challenge for researchers is to reconstruct network structure from available expression data. Many different methods for network inference have been proposed (163). A common problem of such models is exponential complexity: the number of parameters increases exponentially with the number of variables. Thus, many alternative and equally probable network structures may be constructed from a given dataset. Dupuy and Simon (164)

reviewed the cancer literature of studies relating gene expression profiles to patient outcome, either response to treatment, survival or disease-free survival and found that 50% of the publications had at least one flaw so serious as to raise questions about the validity of the conclusions. The three most common serious flaws they found were: misleading use of cluster analysis, lack of adjustment for the multiplicity of analyzing thousands of genes, and erroneous use of partial cross-validation. They pointed out that cluster analysis rarely has a valid role in the development of predictive classifiers. Its wide use in the literature reflects a lack of proper statistical guidance or collaboration in the conduct of expression profiling studies (164).Therefore, cancer research organization need to better appreciate the fundamental changes occurring in the nature of biomedical research and make major commitments to departments for providing professional biostatistical collaboration as an integral part of translational research.

7.5. Challenges in incorporating molecular profiling assays into routine clinical practice

While the first-generation prognostic multigene classifiers, such as the MammaPrint assay and the Oncotype *DX* breast cancer assay, are the closest to clinical practice, the second-generation prognostic multigene assays have not been commercialized. This includes the assessment of breast cancer microenvironment or host immune response. The assay requires further external validation studies to determine their clinical utilities (165). Despite several studies, the translation of predictive multigene classifiers into the clinic is even more challenging than that of prognostic multigene classifiers (166). Most of the predictive assays are derived mainly from cell lines. Microarray as the assay platform is not as quantitative as using a qRT-PCR assay. Therefore, subtle changes in gene expression may not be reflected in microarray-based assays, although these subtle differences may be sufficient to cause resistance to chemotherapeutics. Furthermore, resistance may occur due to low penetrance of the drug being administered and may be unrelated to tumor tissue. To incorporate prognostic and/or predictive multigene classifiers into clinical practice, the following key criteria need to be fulfilled:

First, the platform on which the classifier is based should be suitable for broad clinical application and ensure that the classifier is stable under a variety of operating conditions. If not, the classifier needs to be translated to a clinically applicable platform (167). The assay protocols should be standardized to achieve satisfactory inter-laboratory and intra-laboratory reproducibility, thereby establishing analytic validity. Assay standardization includes pre-analytic parameters, such as sample storage and preparation, and analytic performance parameters, such as the sensitivity and specificity of the system as well as assay reproducibility. The Clinical Laboratory Improvement Amendments of 1988 (CLIA) requires laboratories to independently establish analytic validity and improve assay standardization. To venture from scientific discovery to the beginning of clinical translational research is a challenge as academic scientist are usually funded and rewarded for discovery, rather than to pursue focused translational research as members of a large interdisciplinary team. Funding agencies may not be experience in funding and monitoring

focus translational research. In some other developing countries, to fund such large interdisciplinary and multicenter translational research is prohibitively expensive. Because of these limitations in conducting and funding focused translational research, a defined discovery to a product for use in a defined medical context goes untranslated unless they are of interest to the industry (168, 169).

Second, it is critical to classify studies as developmental or validation studies in order to increase the clinical validity of the classifier. For assays that purport to elucidate predictive significance, this strategy needs to be applied to determine the clinical utility of the classifier (167, 170). Developmental studies need to include internal clinical validation; this can be accomplished either by splitting the study population into two populations (the training model and the testing model or by cross-validation based on repeated model development and testing on random data partitions. These approaches will increase the accuracy of the classifier, which in turn makes its further development possible. Independent validation studies are critical to further evaluate the predictive accuracy and usefulness of the classifier in clinical practice. The studies should be prospectively designed, and should verify both clinical validity and clinical utility. Pusztai et al (171) identified out of the 939 publications over twenty years period on prognostic factors for patients with breast cancer, only estrogen receptor, progesterone receptor and HER2 amplification and *Oncotype DX* RS were included alongside the traditional staging variables recommended by the ASCO guidelines. The pitfall for most of these genomics discovery researches is that only a few of the markers studied were properly validated in a cohort. However, most of the studies were performed using convenience sampling of heterogenous collection of patients and difficult to use such results in therapeutic decision making for individual patients. Finally, most of the publications were based on research assays without demonstration of robustness or analytical validity. Without a diagnostic company to develop a robust assay for a test with a clear and important medical application, the publication is unlikely to be part of successful translational research (169).

Third, does the classifier only assess prognosis? Or does it help with selection of a certain type of therapy? What is the therapeutic relevance of the classifier? Prognostic multigene classifiers assess the likelihood of disease recurrence, whereas predictive multigene classifiers evaluate the potential benefit from certain types of chemotherapy or anti-estrogen therapy. However, a prognostic classifier may also exhibit predictive significance. If a classifier is a predictive classifier, the bar for utility is often quite low. For example, approximately half of patients with HER2 positivity respond to trastuzumab. However, if the assay assesses low likelihood for recurrence or metastases (a prognostic assay), patients classified as low risk need to have such a low risk that they can be spared from adjuvant therapy without affecting their long-term prognosis (172).

Fourth, the incorporation of the classifier into the clinic might be more beneficial if it outperforms or adds predictive power to existing prognostic methods; this would help justify the money and time invested in its external validation in a trial of a much larger scale. In other words, it is important to determine cost-effectiveness. The "intrinsic" classification was the first assay to use modern molecular tools to classify breast cancers. MammaPrint

(111) and the *Oncotype DX*(118) have been tested in more than one validation cohort and are being tested for further clinical utility in large prospective trials in Europe (MINDACT; MammaPrint assay) (140) and in the United States (TAILORx; *Oncotype DX* assay)(173). Both assays have completed a cost-benefit analysis on the utility of the assay in clinical practice (174-178). Both assays demonstrate cost effectiveness in guiding adjuvant chemotherapy treatment in patients with early-stage breast cancer. Another assay in an advanced stage of development is a 50-gene assay (PAM50) (179), although the clinically applicable platform of intrinsic subtype classification is still a long away from clinical application.

The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group (EWG) assessed the value of the *Oncotype DX* and MammaPrint assay. The EWG found insufficient evidence to make a recommendation for or against the use of tumor gene expression profiles to improve outcomes in defined populations of women with breast cancer (180). The EWG encouraged further development and evaluation of these technologies. It is clear that the molecular profiling tests have a great potential to improve clinical decision making, since they address the complexity of breast cancer. It was suggested that the combinatorial use of these assays with the existing traditional clinicopathologic parameters to be more favorable, as clinicians are hesitant to do away with the existing clinicopathologic parameters. Indeed, a recent study used a similar combinatorial approach in which the *Oncotype DX* RS was integrated with clinicopathological parameters to develop a tool, the RS-Pathology-Clinical (RSPC) assessment (181). This model although requires validation, might have the greatest predictive and/or prognostic utility in cases classified as "intermediate risk" by the *Oncotype DX* (182).These studies highlight the difficulties in prognostication in patients with breast cancer and the need to use anatomical, histological, and biological approaches to assist with clinical decision-making. It is indisputable that multigene classifiers cannot replace, but rather strengthen, prognostication and prediction in combination with clinicopathological parameters. They do not have a role in cases in which the patient (or the clinician) has already made the decision to proceed with systemic adjuvant therapy. However, these tests have a role to play in those patients who are undecided or for whom a definite decision cannot be made based on clinicopathological findings. No test should be ordered if its results are not going to influence clinical decisions (168).

i. **Problems related to early detection**

Scientists postulate the basic underlying prognostic microarray studies is that all tumors acquire a metastasis phenotype through the same unique mechanism, and that gene expression data in tumor tissue obtained at resection of the primary tumor can be used to clearly distinguish between tumors that will relapse or will not relapse. The results of the pioneering prognostic microarray study concerning breast cancer (111) are considered proof of concept and have led to general acceptance of the postulate. However, the performances of microarray studies are poorer than initially thought and published gene signature lists are unstable (161). Some of the multi-biomarker scores do show consistent prognostic value such as in breast cancer, but until the recent advent of large validation studies, microarray studies are not significantly better prognostic

classification than conventional prognostic models (113, 122). In addition, it has been shown that almost all first-generation gene signatures in breast cancer provide a quantitative read-out of the same biological pathway of proliferation (183, 184). As of today we are still in need of a precise estimation of the incremental value (185-187). Moreover, by assuming a unique mechanism for the metastasis phenotype, the postulation contradicts with the concept of cancer heterogeneity and consequently with personalized treatments. The potential interest of microarrays could not be rejected provided true critical consideration, incorporating, and not opposed to, full clinical evidence is now necessary.

ii. Problems related to prognosis indicator

The validation of "first-generation" prognostic signatures, usually based exclusively on gene expression profiling, has proven particularly challenging (188). It has been even more difficult to identify and validate predictors of response to nontargeted therapies (radiotherapy and chemotherapy), although analysis of large sample sets from clinical trials have already provided preliminary evidence of novel markers (189).

Limitations to the current prognostic multigene signatures

The ability of the *Oncotype Dx* and MammaPrint, to determine prognosis seems to be directly correlated to the assessment of proliferation/cell cycle-related genes (183, 190). The fact that these multigene signatures are mere surrogates of proliferation poses some important problems for their uses. First, given that proliferation has been shown to be prognostic in ER-positive disease and not in ER-negative cancers, first-generation signatures are applicable only for the prognostication of patients with ER-positive and HER2-negative breast cancers (190, 191). As the expression level of proliferation related genes in ER-positive cancers has been demonstrated to follow a continuum rather than a bimodal distribution, the subdivision of ER-positive cancers into good-prognosis (luminal A) and poor-prognosis (luminal B) groups is considered artificial (183, 190). In fact, the continuous nature of the *Oncotype DX* RS is more representative of the ranges of prognosis of patients with ER-positive disease. It should be noted, however, that this approach for clinical decision-making might be problematic. For instance, the prognostication and management of patients with an intermediate RS remain unclear, and up to 40% to 60% of clinically intermediate-risk patients (that is, breast cancers combining ER-positive, HER2-negative, and grade II status) are allocated to the intermediate-risk RS group (175). Therefore, the actual contribution of *Oncotype DX* to the management of this particular group of patients remains to be elucidated, and is currently being examined in the TAILORx trial (173, 175). Lack of prognostic power of first-generation prognostic signatures in ER-negative breast cancer and their associations with proliferation in ER-positive breast cancer have brought to the forefront of cancer research the limitations of histological grading. Classical histological grade is not prognostic in ER-negative disease and is strongly associated with proliferation (190, 192). It should be noted, however, that the levels of intra- and inter-observer agreement of histological grade remain suboptimal, despite the numerous

efforts to implement a standardized histological grading system (192). It could be argued, on the basis of the above observations that the major contribution of first-generation prognostic gene signatures is to provide a standardized proliferation assay for breast cancer. A second limitation of the first-generation prognostic signatures stems from the fact that most of them were developed to predict short-term distant recurrence (<5 years) and were shown to have a strong 'time dependence' and a reduced prognostic value after 5 to 10 years of follow-up (113, 193). Hence, these signatures may represent merely early distant recurrence surrogates and are unable to predict late relapses with the same accuracy. Thus, there is still a need to develop signatures that could identify patients who have a higher risk of late relapse and who may benefit from prolonged therapy.

iii. Problems related to therapeutic response

There is also increasing evidence that better classifiers and improved prognostication can be derived from combined analysis that profile both tumour DNA and RNA (194-196). Neoadjuvant therapy trials hold great promise as the right framework to identify these predictive biomarkers for chemotherapy (and targeted therapies) response. ER and Her2 are predictors of a lack of benefit from targeted therapies, hormone therapy and anti-Her2-targeted agents, when the cancers do not express the markers. These predictors, however, fail to identify tumours that despite expressing the biomarkers still fail to respond to the targeted therapies (197).

7.6. Gene expression signatures and response to chemotherapy

With the clinical need for predictive markers for specific chemotherapy agents and multidrug regimens, several groups have developed multigene signatures specifically designed to predict response in patients receiving either chemotherapy or endocrine therapy. Using supervised approaches, several studies have attempted to identify multigene signatures of response to chemotherapy by comparing gene expression profiles between high sensitivity and low-responsiveness tumors (198-201). The majority of the studies focused on neoadjuvant chemotherapy and analyzed tumor samples obtained from biopsies taken at diagnosis before initiation of chemotherapy by microarrays or RT-PCR. Chemotherapy sensitivity usually was estimated with rate of pathological complete response to neoadjuvant therapy (pCR) as a surrogate of long-term benefit from the treatment. For example, a 30-gene signature was developed by the MD Anderson Cancer Center group in 82 breast cancer patients receiving T/FAC chemotherapy (paclitaxel, fluorouracil, doxorubicin, cyclophosphamide). This predictor signature was then validated in 51 independent patients and predicted pCR probability with higher sensitivity and negative predictive value than clinical variables based on age, grade, and ER status (198, 200), which were later confirmed in an independent study (202). Despite these interesting preliminary results, the accuracy of the 30-gene predictor was not found in a recent study in which it was not an independent predictor of pCR after multivariate analysis and did not perform better than clinical variables (203). A similar 78-gene signature to MammaPrint that

was developed from a dataset of metastatic breast cancerpatients who did and did not respond to tamoxifen treatment was identified as truly predictive of tamoxifen response. They found that their signatures seemed to be more predictive than prognostic compared with the RS in an independent set of tamoxifen-treated ER-positive metastatic breast cancer patients (204). Whilst the metastatic setting may be the most logical way to investigate the true predictive ability of a biomarker, it remains plausible that metastatic breast cancer patients have different disease biology compared with those having early-stage disease. Miller et al (205) used the neoadjuvant or preoperative setting to uncover gene profiles for which baseline expression and relative change with 14 days of treatment differed between breast cancers that were clinically responsive or resistant toletrozole therapy. The advantage of the neoadjuvant settingis that it allows multiple ways of assessment of response to therapy, eg, monitoring of changes in tumor size during the first months of treatment and sequential tumor biopsiesbefore and after neoadjuvant treatment with letrozole. Gene expression profiles were then related to clinical responses as assessed from tumor volume measurements after three months of treatment. This study underscores the potential of the neoadjuvant setting for high-level correlative science, but also supports the need for biologically driven hypotheses and stratification of luminal subtypes, and also highlights the difficulties of serial analyses using high-dimensional data.

An alternative attempt to predict chemosensitivity to specific chemotherapy regimens was developed with the use of *in vitro* models. Using a combination of *in vitro* signatures associated with drug sensitivity in cell lines, a composite signatures that could predict response to multidrug regimens were derived and translated to patients receiving multidrug chemotherapy (206). These 'regimen-specific' signatures tested in patients who, as participants in the European Organization for Research and Treatment of Cancer (EORTC) BIG00-01 clinical trial, received TET (docetaxel, epirubicin-docetaxel) or FEC (fluorouracil, epirubicin, and cyclophosphamide) chemotherapy resulted in a validation study (207). Importantly, problems with the methodology of these studies have been identified (208) and serious concerns about the validity of the published results were raised. Subsequently, after a series of investigations, the findings derived from *in vitro* studies were considered invalid, and this led to the discontinuation of the clinical trials based on these prediction models (166, 209).

Another method to develop multigene classifiers of chemosensitivity is based on the use of metagenes, groups of co-expressed genes associated with a small number of biological processes. A retrospective microarray analysis of prospectively collected ER-negative breast cancer samples demonstrated that increased stromal gene expression predicted resistance to FEC chemotherapy, which was subsequently validated in two independent cohorts (210). Despite the promising initial results, the signatures of chemotherapy sensitivity have so far had limited use in clinical practice. Most of them have been developed in small, convenience cohorts and require further external validation. None of the different predictors of chemosensitivity is commercially available, and additional evidence is still required before they can be implemented in clinical practice. A recent review has discussed the reasons for the limited success of the predictive signatures available to date (166). On the basis of the

design employed in most of the studies, the predictive signatures for multidrug regimens are likely to capture the transcriptomic features of sensitivity/resistance to cytotoxic agents in general. These mechanisms may constitute convergent phenotypes, that are multiple genetic/epigenetic aberrations that may lead to resistance to cytoxic agents (211).

8. Conclusion

Cancer is a multi-factorial disease that involves multiple genes and distinct pathways. The ultimate objective in the high throughput gene expression study approach is to fill the gap in the early biomarker detection, prognostication improvement and gene-targetted therapy. Outcomes from these studies can be obtained from the literatures and some are available as open public databases. Scientists have taken steps forward by using the data either as a single gene studies or multiple genes with related molecular pathways to investigate further on an individual cancer. However, there is a great challenge to devise the suitable gene lists from heterogenous data especially for drug discovery studies. With a great amount of genomic data avaiable, nearly all cancers faced the same setbacks of unable to pick the right genes for the right cancer. Among all cancer, breast cancer has the most advance experience in translating the lab findings into the clinical practice with the emergence of multigene signatures. The current array data can provide a platform for future scientists to explain the complexity of cancer in combination with the latest advancement in deep sequencing technology,

Author details

Norfilza M. Mokhtar
Department of Physiology, Faculty of Medicine
UKM Medical Molecular Biology Institute, Universiti Kebangsaan Malaysia, Cheras,
Kuala Lumpur, Malaysia

Nor Azian Murad, Then Sue Mian and Rahman Jamal
UKM Medical Molecular Biology Institute, Universiti Kebangsaan Malaysia, Cheras,
Kuala Lumpur, Malaysia

Acknowledgement

9. References

[1]   Abelev GI, Eraiser TL. On the path to understanding the nature of cancer. Biochemistry (Mosc). 2008 May;73(5):487-97.

[2] Deschoolmeester V, Baay M, Lardon F, Pauwels P, Peeters M. Immune Cells in Colorectal Cancer: Prognostic Relevance and Role of MSI. Cancer Microenviron. 2011 Dec;4(3):377-92.

[3] Allen JI. Molecular biology of colon polyps and colon cancer. Semin Surg Oncol. 1995 Nov-Dec;11(6):399-405.

[4] Al-Sohaily S, Biankin A, Leong R, Kohonen-Corish M, Warusavitarne J. Molecular pathways in colorectal cancer. J Gastroenterol Hepatol. 2012 Jun 13.

[5] Blum HE. [Colon carcinoma: molecular diagnosis and therapy]. Praxis (Bern 1994). 1997 Sep 24;86(39):1504-9.

[6] Bodmer WF. Cancer genetics: colorectal cancer as a model. J Hum Genet. 2006;51(5):391-6.

[7] Cunningham D, Atkin W, Lenz HJ, Lynch HT, Minsky B, Nordlinger B, et al. Colorectal cancer. Lancet. 2010 Mar 20;375(9719):1030-47.

[8] Jin K, Gao W, Lu Y, Lan H, Teng L, Cao F. Mechanisms regulating colorectal cancer cell metastasis into liver (Review). Oncol Lett. 2012 Jan;3(1):11-5.

[9] Salah S, Watanabe K, Welter S, Park JS, Park JW, Zabaleta J, et al. Colorectal cancer pulmonary oligometastases: pooled analysis and construction of a clinical lung metastasectomy prognostic model. Ann Oncol. 2012 Apr 29.

[10] Horton JK, Tepper JE. Staging of colorectal cancer: past, present, and future. Clin Colorectal Cancer. 2005 Jan;4(5):302-12.

[11] Aarts MJ, Lemmens VE, Louwman MW, Kunst AE, Coebergh JW. Socioeconomic status and changing inequalities in colorectal cancer? A review of the associations with risk, treatment and outcome. Eur J Cancer. 2010 Oct;46(15):2681-95.

[12] Fornaro R, Frascio M, Denegri A, Stabilini C, Impenatore M, Mandolfino F, et al. [Chron's disease and cancer]. Ann Ital Chir. 2009 Mar-Apr;80(2):119-25.

[13] Triantafillidis JK, Nasioulas G, Kosmidis PA. Colorectal cancer and inflammatory bowel disease: epidemiology, risk factors, mechanisms of carcinogenesis and prevention strategies. Anticancer Res. 2009 Jul;29(7):2727-37.

[14] Abdul Murad NA, Othman Z, Khalid M, Abdul Razak Z, Hussain R, Nadesan S, et al. Missense Mutations in MLH1, MSH2, KRAS, and APC Genes in Colorectal Cancer Patients in Malaysia. Dig Dis Sci. 2012 Jun 6.

[15] Hewish M, Lord CJ, Martin SA, Cunningham D, Ashworth A. Mismatch repair deficient colorectal cancer in the era of personalized treatment. Nat Rev Clin Oncol. 2010 Apr;7(4):197-208.

[16] Claes K, Dahan K, Tejpar S, De Paepe A, Bonduelle M, Abramowicz M, et al. The genetics of familial adenomatous polyposis (FAP) and MutYH-associated polyposis (MAP). Acta Gastroenterol Belg. 2011 Sep;74(3):421-6.

[17] Hanks H, Veitch C, Harris M. Colorectal cancer management - the role of the GP. Aust Fam Physician. 2008 Apr;37(4):259-61.

[18] Finlay GJ. Genetics, molecular biology and colorectal cancer. Mutat Res. 1993 Nov;290(1):3-12.

[19] Fearon ER. Molecular genetics of colorectal cancer. Annu Rev Pathol. 2011;6:479-507.

[20] Pritchard CC, Grady WM. Colorectal cancer molecular biology moves into clinical practice. Gut. 2011 Jan;60(1):116-29.

[21] Hendry JH. Genomic instability: potential contributions to tumour and normal tissue response, and second tumours, after radiotherapy. Radiother Oncol. 2001 May;59(2):117-26.

[22] Geigl JB, Obenauf AC, Schwarzbraun T, Speicher MR. Defining 'chromosomal instability'. Trends Genet. 2008 Feb;24(2):64-9.

[23] Gagos S, Irminger-Finger I. Chromosome instability in neoplasia: chaotic roots to continuous growth. Int J Biochem Cell Biol. 2005 May;37(5):1014-33.

[24] Guastadisegni C, Colafranceschi M, Ottini L, Dogliotti E. Microsatellite instability as a marker of prognosis and response to therapy: a meta-analysis of colorectal cancer survival data. Eur J Cancer. 2010 Oct;46(15):2788-98.

[25] Iacopetta B, Grieu F, Amanuel B. Microsatellite instability in colorectal cancer. Asia Pac J Clin Oncol. 2010 Dec;6(4):260-9.

[26] Des Guetz G, Uzzan B, Nicolas P, Schischmanoff O, Perret GY, Morere JF. Microsatellite instability does not predict the efficacy of chemotherapy in metastatic colorectal cancer. A systematic review and meta-analysis. Anticancer Res. 2009 May;29(5):1615-20.

[27] Vilar E, Gruber SB. Microsatellite instability in colorectal cancer-the stable evidence. Nat Rev Clin Oncol. 2010 Mar;7(3):153-62.

[28] Barault L, Charon-Barra C, Jooste V, de la Vega MF, Martin L, Roignot P, et al. Hypermethylator phenotype in sporadic colon cancer: study on a population-based series of 582 cases. Cancer Res. 2008 Oct 15;68(20):8541-6.

[29] Hinoue T, Weisenberger DJ, Pan F, Campan M, Kim M, Young J, et al. Analysis of the association between CIMP and BRAF in colorectal cancer by DNA methylation profiling. PLoS One. 2009;4(12):e8357.

[30] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001 Feb 15;409(6822):860-921.

[31] Thean LF, Li HH, Teo YY, Koh WP, Yuan JM, Teoh ML, et al. Association of caucasian-identified variants with colorectal cancer risk in singapore chinese. PLoS One. 2012;7(8):e42407.

[32] Gerber MM, Hampel H, Schulz NP, Fernandez S, Wei L, Zhou XP, et al. Evaluation of allele-specific somatic changes of genome-wide association study susceptibility alleles in human colorectal cancers. PLoS One. 2012;7(5):e37672.

[33] Stadler ZK, Thom P, Robson ME, Weitzel JN, Kauff ND, Hurley KE, et al. Genome-wide association studies of cancer. J Clin Oncol. 2010 Sep 20;28(27):4255-67.

[34] Tenesa A, Farrington SM, Prendergast JG, Porteous ME, Walker M, Haq N, et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. Nat Genet. 2008 May;40(5):631-7.

[35] Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet. 2009 Aug;41(8):882-4.

[36] Matsuo K, Suzuki T, Ito H, Hosono S, Kawase T, Watanabe M, et al. Association between an 8q24 locus and the risk of colorectal cancer in Japanese. BMC Cancer. 2009;9:379.

[37] Jaeger E, Webb E, Howarth K, Carvajal-Carmona L, Rowan A, Broderick P, et al. Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. Nat Genet. 2008 Jan;40(1):26-8.

[38] Roberts AB, Wakefield LM. The two faces of transforming growth factor beta in carcinogenesis. Proc Natl Acad Sci U S A. 2003 Jul 22;100(15):8621-3.

[39] Sillars-Hardebol AH, Carvalho B, de Wit M, Postma C, Delis-van Diemen PM, Mongera S, et al. Identification of key genes for carcinogenic pathways associated with colorectal adenoma-to-carcinoma progression. Tumour Biol. 2010 Apr;31(2):89-96.

[40] Zhou H, Kuang J, Zhong L, Kuo WL, Gray JW, Sahin A, et al. Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation. Nat Genet. 1998 Oct;20(2):189-93.

[41] De Luca M, Lavia P, Guarguaglini G. A functional interplay between Aurora-A, Plk1 and TPX2 at spindle poles: Plk1 controls centrosomal localization of Aurora-A and TPX2 spindle association. Cell Cycle. 2006 Feb;5(3):296-303.

[42] Smits VA, Klompmaker R, Arnaud L, Rijksen G, Nigg EA, Medema RH. Polo-like kinase-1 is a target of the DNA damage checkpoint. Nat Cell Biol. 2000 Sep;2(9):672-6.

[43] Wu Y, Wang X, Wu F, Huang R, Xue F, Liang G, et al. Transcriptome profiling of the cancer, adjacent non-tumor and distant normal tissues from a colorectal cancer patient by deep sequencing. PLoS One. 2012;7(8):e41001.

[44] Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA Cancer J Clin. 2012 Jan-Feb;62(1):10-29.

[45] Zainal Ariffin O, Nor Saleha IT. National Cancer Registry Report 2007. Kuala Lumpur: Ministry of Health, Malaysia2011.

[46] Menon U, Jacobs IJ. Ovarian cancer screening in the general population: current status. Int J Gynecol Cancer. 2001;11 Suppl 1:3-6.

[47] Gram IT, Lukanova A, Brill I, Braaten T, Lund E, Lundin E, et al. Cigarette smoking and risk of histological subtypes of epithelial ovarian cancer in the EPIC cohort study. Int J Cancer. 2012 May 1;130(9):2204-10.

[48] Jordan SJ, Siskind V, A CG, Whiteman DC, Webb PM. Breastfeeding and risk of epithelial ovarian cancer. Cancer Causes Control. 2010 Jan;21(1):109-16.

[49] Fekete T, Raso E, Pete I, Tegze B, Liko I, Munkacsy G, et al. Meta-analysis of gene expression profiles associated with histological classification and survival in 829 ovarian cancer samples. Int J Cancer. 2012 Jul 1;131(1):95-105.

[50] Cannistra SA. Cancer of the ovary. N Engl J Med. 2004 Dec 9;351(24):2519-29.

[51] Nakayama N, Nakayama K, Yeasmin S, Ishibashi M, Katagiri A, Iida K, et al. KRAS or BRAF mutation status is a useful predictor of sensitivity to MEK inhibition in ovarian cancer. Br J Cancer. 2008 Dec 16;99(12):2020-8.

[52] Tutt A, Robson M, Garber JE, Domchek SM, Audeh MW, Weitzel JN, et al. Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. Lancet. 2010 Jul 24;376(9737):235-44.

[53] Christie M, Oehler MK. Molecular pathology of epithelial ovarian cancer. J Br Menopause Soc. 2006 Jun;12(2):57-63.

[54] Alsop K, Fereday S, Meldrum C, Defazio A, Emmanuel C, George J, et al. BRCA Mutation Frequency and Patterns of Treatment Response in BRCA Mutation-Positive Women With Ovarian Cancer: A Report From the Australian Ovarian Cancer Study Group. J Clin Oncol. 2012 Jul 20;30(21):2654-63.

[55] Aarnio M, Sankila R, Pukkala E, Salovaara R, Aaltonen LA, de la Chapelle A, et al. Cancer risk in mutation carriers of DNA-mismatch-repair genes. Int J Cancer. 1999 Apr 12;81(2):214-8.

[56] Stany MP, Vathipadiekal V, Ozbun L, Stone RL, Mok SC, Xue H, et al. Identification of novel therapeutic targets in microdissected clear cell ovarian cancers. PLoS One. 2011;6(7):e21121.

[57] Yamaguchi K, Mandai M, Oura T, Matsumura N, Hamanishi J, Baba T, et al. Identification of an ovarian clear cell carcinoma gene signature that reflects inherent disease biology and the carcinogenic processes. Oncogene. 2010 Mar 25;29(12):1741-52.

[58] Schwartz DR, Kardia SL, Shedden KA, Kuick R, Michailidis G, Taylor JM, et al. Gene expression in ovarian cancer reflects both morphology and biological behavior, distinguishing clear cell from other poor-prognosis ovarian carcinomas. Cancer Res. 2002 Aug 15;62(16):4722-9.

[59] Schaner ME, Ross DT, Ciaravino G, Sorlie T, Troyanskaya O, Diehn M, et al. Gene expression patterns in ovarian carcinomas. Mol Biol Cell. 2003 Nov;14(11):4376-86.

[60] Tothill RW, Tinker AV, George J, Brown R, Fox SB, Lade S, et al. Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. Clin Cancer Res. 2008 Aug 15;14(16):5198-208.

[61] Shridhar V, Lee J, Pandita A, Iturria S, Avula R, Staub J, et al. Genetic analysis of early-versus late-stage ovarian tumors. Cancer Res. 2001 Aug 1;61(15):5895-904.

[62] Davidson B, Stavnes HT, Holth A, Chen X, Yang Y, Shih Ie M, et al. Gene expression signatures differentiate ovarian/peritoneal serous carcinoma from breast carcinoma in effusions. J Cell Mol Med. 2011 Mar;15(3):535-44.

[63] Vui-Kee K, Mohd Dali AZ, Mohamed Rose I, Ghazali R, Jamal R, Mokhtar NM. Molecular markers associated with nonepithelial ovarian cancer in formalin-fixed, paraffin-embedded specimens by genome wide expression profiling. Kaohsiung J Med Sci. 2012 May;28(5):243-50.

[64] Tomlinson VA, Newbery HJ, Wray NR, Jackson J, Larionov A, Miller WR, et al. Translation elongation factor eEF1A2 is a potential oncoprotein that is overexpressed in two-thirds of breast tumours. BMC Cancer. 2005;5:113.

[65] Sherman ME, Bur ME, Kurman RJ. p53 in endometrial cancer and its putative precursors: evidence for diverse pathways of tumorigenesis. Hum Pathol. 1995 Nov;26(11):1268-74.

[66] Creasman W. Revised FIGO staging for carcinoma of the endometrium. Int J Gynaecol Obstet. 2009 May;105(2):109.

[67] Lax SF. Molecular genetic pathways in various types of endometrial carcinoma: from a phenotypical to a molecular-based classification. Virchows Arch. 2004 Mar;444(3):213-23.

[68] Lax SF, Pizer ES, Ronnett BM, Kurman RJ. Comparison of estrogen and progesterone receptor, Ki-67, and p53 immunoreactivity in uterine endometrioid carcinoma and endometrioid carcinoma with squamous, mucinous, secretory, and ciliated cell differentiation. Hum Pathol. 1998 Sep;29(9):924-31.

[69] Moreno-Bueno G, Sanchez-Estevez C, Cassia R, Rodriguez-Perales S, Diaz-Uriarte R, Dominguez O, et al. Differential gene expression profile in endometrioid and

nonendometrioid endometrial carcinoma: STK15 is frequently overexpressed and amplified in nonendometrioid carcinomas. Cancer Res. 2003 Sep 15;63(18):5697-702.

[70] Kok Ying N, Ganesalingam M, Sabaratnam S. Cancer Incidence in Peninsular Malaysia. Kuala Lumpur2008.

[71] Doll A, Abal M, Rigau M, Monge M, Gonzalez M, Demajo S, et al. Novel molecular profiles of endometrial cancer-new light through old windows. J Steroid Biochem Mol Biol. 2008 Feb;108(3-5):221-9.

[72] Llaurado M, Ruiz A, Majem B, Ertekin T, Colas E, Pedrola N, et al. Molecular bases of endometrial cancer: new roles for new actors in the diagnosis and the therapy of the disease. Mol Cell Endocrinol. 2012 Jul 25;358(2):244-55.

[73] Felix AS, Weissfeld JL, Stone RA, Bowser R, Chivukula M, Edwards RP, et al. Factors associated with Type I and Type II endometrial cancer. Cancer Causes Control. 2010 Nov;21(11):1851-6.

[74] Deligeoroglou E, Michailidis E, Creatsas G. Oral contraceptives and reproductive system cancer. Ann N Y Acad Sci. 2003 Nov;997:199-208.

[75] Aoki Y, Watanabe M, Amikura T, Obata H, Sekine M, Yahata T, et al. Adjuvant chemotherapy as treatment of high-risk stage I and II endometrial cancer. Gynecol Oncol. 2004 Aug;94(2):333-9.

[76] Burbos N, Musonda P, Crocker SG, Morris EP, Duncan TJ, Nieto JJ. Outcome of investigations for postmenopausal vaginal bleeding in women under the age of 50 years. Gynecol Oncol. 2012 Apr;125(1):120-3.

[77] Lacey JV, Jr., Yang H, Gaudet MM, Dunning A, Lissowska J, Sherman ME, et al. Endometrial cancer and genetic variation in PTEN, PIK3CA, AKT1, MLH1, and MSH2 within a population-based case-control study. Gynecol Oncol. 2011 Feb;120(2):167-73.

[78] Seeger A, Kolbl H, Petry IB, Gebhard S, Battista MJ, Bohm D, et al. p53 is correlated with low BMI negative progesterone receptor status and recurring disease in patients with endometrial cancer. Gynecol Oncol. 2012 Apr;125(1):200-7.

[79] Singh M, Darcy KM, Brady WE, Clubwala R, Weber Z, Rittenbach JV, et al. Cadherins, catenins and cell cycle regulators: impact on survival in a Gynecologic Oncology Group phase II endometrial cancer trial. Gynecol Oncol. 2011 Nov;123(2):320-8.

[80] Gurin CC, Federici MG, Kang L, Boyd J. Causes and consequences of microsatellite instability in endometrial carcinoma. Cancer Res. 1999 Jan 15;59(2):462-6.

[81] Maxwell GL, Chandramouli GV, Dainty L, Litzi TJ, Berchuck A, Barrett JC, et al. Microarray analysis of endometrial carcinomas and mixed mullerian tumors reveals distinct gene expression profiles associated with different histologic types of uterine cancer. Clin Cancer Res. 2005 Jun 1;11(11):4056-66.

[82] Mokhtar NM, Ramzi NH, Yin-Ling W, Rose IM, Hatta Mohd Dali AZ, Jamal R. Laser capture microdissection with genome-wide expression profiling displayed gene expression signatures in endometrioid endometrial cancer. Cancer Invest. 2012 Feb;30(2):156-64.

[83] Pavelic J, Radakovic B, Pavelic K. Insulin-like growth factor 2 and its receptors (IGF 1R and IGF 2R/mannose 6-phosphate) in endometrial adenocarcinoma. Gynecol Oncol. 2007 Jun;105(3):727-35.

[84] Sugiyama Y, Sugiyama K, Hirai Y, Akiyama F, Hasumi K. Microdissection is essential for gene expression profiling of clinically resected cancer tissues. 2002 [updated Jan; cited 117 1]; 2002/01/16:[109-16]. Available from:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11789716.

[85] Mannelqvist M, Stefansson IM, Bredholt G, Hellem Bo T, Oyan AM, Jonassen I, et al. Gene expression patterns related to vascular invasion and aggressive features in endometrial cancer. Am J Pathol. 2011 Feb;178(2):861-71.

[86] Ferguson SE, Olshen AB, Levine DA, Viale A, Barakat RR, Boyd J. Molecular profiling of endometrial cancers from African-American and Caucasian women. Gynecol Oncol. 2006 May;101(2):209-13.

[87] Metcalfe K, Lubinski J, Lynch HT, Ghadirian P, Foulkes WD, Kim-Sing C, et al. Family history of cancer and cancer risks in women with BRCA1 or BRCA2 mutations. J Natl Cancer Inst. 2010 Dec 15;102(24):1874-8.

[88] Sariego J. Breast cancer in the young patient. Am Surg. 2010 Dec;76(12):1397-400.

[89] Ross JS. Multigene classifiers, prognostic factors, and predictors of breast cancer clinical outcome. Adv Anat Pathol. 2009 Jul;16(4):204-15.

[90] Ross JS, Fletcher JA, Linette GP, Stec J, Clark E, Ayers M, et al. The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy. Oncologist. 2003;8(4):307-25.

[91] Eroles P, Bosch A, Alejandro Perez-Fidalgo J, Lluch A. Molecular biology in breast cancer: Intrinsic subtypes and signaling pathways. Cancer Treat Rev. 2012 Oct;38(6):698-707.

[92] This P, de la Rochefordiere A, Savignoni A, Falcou MC, Tardivon A, Thibault F, et al. Breast and ovarian cancer risk management in a French cohort of 158 women carrying a BRCA1 or BRCA2 germline mutation: patient choices and outcome. Fam Cancer. 2012 Jun 19.

[93] de Bruin MA, Kwong A, Goldstein BA, Lipson JA, Ikeda DM, McPherson L, et al. Breast cancer risk factors differ between Asian and white women with BRCA1/2 mutations. Fam Cancer. 2012 May 26.

[94] Keogh RH, Park JY, White IR, Lentjes MA, McTaggart A, Bhaniani A, et al. Estimating the alcohol-breast cancer association: a comparison of diet diaries, FFQs and combined measurements. Eur J Epidemiol. 2012 May 29.

[95] Saslow D, Hannan J, Osuch J, Alciati MH, Baines C, Barton M, et al. Clinical breast examination: practical recommendations for optimizing performance and reporting. CA Cancer J Clin. 2004 Nov-Dec;54(6):327-44.

[96] McDonald S, Saslow D, Alciati MH. Performance and reporting of clinical breast examination: a review of the literature. CA Cancer J Clin. 2004 Nov-Dec;54(6):345-61.

[97] Murphy CG, Morris PG. Recent advances in novel targeted therapies for HER2-positive breast cancer. Anticancer Drugs. 2012 Sep;23(8):765-76.

[98] Cai Q, Wen W, Qu S, Li G, Egan KM, Chen K, et al. Replication and functional genomic analyses of the breast cancer susceptibility locus at 6q25.1 generalize its importance in women of chinese, Japanese, and European ancestry. Cancer Res. 2011 Feb 15;71(4):1344-55.

[99] Stacey SN, Sulem P, Zanon C, Gudjonsson SA, Thorleifsson G, Helgason A, et al. Ancestry-shift refinement mapping of the C6orf97-ESR1 breast cancer susceptibility locus. PLoS Genet. 2010 Jul;6(7):e1001029.

[100] Turnbull C, Ahmed S, Morrison J, Pernet D, Renwick A, Maranian M, et al. Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet. 2010 Jun;42(6):504-7.

[101] Gold B, Kirchhoff T, Stefanov S, Lautenberger J, Viale A, Garber J, et al. Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. Proc Natl Acad Sci U S A. 2008 Mar 18;105(11):4340-5.

[102] Bessarabova M, Kirillov E, Shi W, Bugrim A, Nikolsky Y, Nikolskaya T. Bimodal gene expression patterns in breast cancer. BMC Genomics. 2010;11 Suppl 1:S8.

[103] Sun Z, Asmann YW, Kalari KR, Bot B, Eckel-Passow JE, Baker TR, et al. Integrated analysis of gene expression, CpG island methylation, and gene copy number in breast cancer cells by deep sequencing. PLoS One. 2011;6(2):e17490.

[104] Ma XJ, Dahiya S, Richardson E, Erlander M, Sgroi DC. Gene expression profiling of the tumor microenvironment during breast cancer progression. Breast Cancer Res. 2009;11(1):R7.

[105] Kao J, Salari K, Bocanegra M, Choi YL, Girard L, Gandhi J, et al. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. PLoS One. 2009;4(7):e6146.

[106] Aaroe J, Lindahl T, Dumeaux V, Saebo S, Tobin D, Hagen N, et al. Gene expression profiling of peripheral blood cells for early detection of breast cancer. Breast Cancer Res. 2010;12(1):R7.

[107] Iwamoto T, Bianchini G, Booser D, Qi Y, Coutant C, Ya-Hui Shiang C, et al. Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer. J Natl Cancer Inst. 2011;103:264-72.

[108] Goldhirsch A, Wood WC, Gelber RD, Coates AS, Thürlimann B, Senn HJ, 10th St. Gallen conference. Meeting highlights: updated international expert consensus on the primary therapy of early breast cancer. J Clin Oncol. 2003;21:3357-65.

[109] NCCN Guidelines for Breast Cancer, NCCN Task Force Report: adjuvant therapy for breast cancer. . J Natl Compr Canc Netw.; 2006. p. S1-26.

[110] Kohli-Laven N, Bourret P, Keating P, Cambrosio A. Cancer clinical trials in the era of genomic signatures: Biomedical innovation, clinical utility, and regulatory scientific hybrids. Soc Stud Sci. 2011;41(4):487-513.

[111] van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature. 2002;415(6871):530-6.

[112] Simon RM, Paik S, Hayes DF. Use of Archived Specimens in Evaluation of Prognostic and Predictive Biomarkers. J Natl Cancer Inst. 2009;101:1446-52.

[113] Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst. 2006;98:1183-92.

[114] Bourreta P, Keating P, Cambrosio A. Regulating diagnosis in post-genomic medicine: Re-aligning clinical judgment? Soc Stud Sci. 2011;73(6):816-24.

[115] Glück S, Yip AYS, Ng ELY. Can we replace the microscope with microarrays for diagnosis, prognosis and treatment of early breast cancer? Expert Opin Ther Targets. 2012;16((Suppl.1)):S17-S22.

[116] Cronin M, Ghosh K, Sistare F, Quackenbush J, Vilker V, O'Connell C. Universal RNA reference materials for gene expression. Clinical Chemistry 2004;50:1464–71.

[117] Cronin MT, Dutta D, Pho M, Nguyen A, Jeong J, Liu ML. Tumor marker discovery by expression profiling RNA from formalin fixed paraffin embedded tissues. . Methods in Molecular Biology. 2009;520:177–93.

[118] Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004;351:2817-26.

[119] Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, et al. Gene Expression and Benefit of Chemotherapy in Women with Node-Negative, Estrogen Receptor-Positive Breast Cancer. J Clin Oncol 2006;24(23):3726-34

[120] Ross JS, Hatzis C, Symmans WF, Pusztai L, Hortobágyi GN. Commercialized multigene predictors of clinical outcome for breast cancer. The Oncologist. 2008;13:477-93.

[121] Habel LA, Shak S, Jacobs M, Capra A, Alexander C, Pho M, et al. A Population-Based Study of Tumor Gene Expression and Risk of Breast Cancer Death Among Lymph Node-Negative Patients. Breast Cancer Res 2006;8(3):R25.

[122] Albain KS, Barlow W, Shak S, Hortobagyi G, Livingston R, Yeh I-T, et al. Prognostic and predictive value of the 21-gene recurrence score assay in post-menopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol. 2010;11(1):55-65.

[123] Jessup JM, Lively TG, Taube SE. Program for the assessment of cancer clinical tests (PACCT): Implementing promising assays into clinical practice. Expert Review of Molecular Diagnostics. 2005;5:271–3.

[124] Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, et al. American Society of Clinical Oncology: Update of recommendations for the use of tumor markers in breast cancer. J Clin Oncol. 2007;25:5287-312.

[125] van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med. 2002;347:1999-2009.

[126] Glas AM, Floore A, Delahaye LJ, Witteveen AT, Pover RC, Bakx N, et al. Converting a breast cancer microarray signature into a high-throughput diagnostic test. BMC Genomics. 2006;7:278.

[127] Keating P, Cambrosio A. Too many numbers: Microarrays in clinical cancer research. Stud Hist Philos Biol Biomed Sci. 2012;43:37-51.

[128] Rogers S, Cambrosio A. Making a new technology work: The standardization and regulation of microarrays. Yale Journal of Biology and Medicine. 2007;80:165–78.

[129] Bueno-de-Mesquita JM, van Harten WH, Retel VP, van't Veer LJ, van Dam FS, Karsenberg K, et al. Use of 70-gene signature to predict prognosis of patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). Lancet Oncol. 2007;8:1079-87.

[130] Beuno-de-Mesquita JM, Linn SC, Keijzer R, Wesseling J, Nuyten DS, van Krimpen C, et al. Validation of 70-gene prognosis signature in node-negative breast cancer. Breast Cancer Res Treat. 2009;117(3).

[131] Wittner BS, Sgroi DC, Ryan PD, Bruinsma TJ, Glas AM, Male A, et al. Analysis of the MammaPrint breast cancer assay in a predominantly postmenopausal cohort. Clin Cancer Res. 2008 May 15;14(10):2988-93.

[132] Wittner BS, Sgroi DC, Ryan PD, Bruinsma TJ, Glas AM, Male A, et al. Analysis of the MammaPrint breast cancer assay in a predominantly postmenopausal cohort. Clin Cancer Res. 2008;14(10):2988-93.

[133] Mook S, Schmidt MK, Viale G, Pruneri G, Eekhout I, Floore A, et al. The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1-3 positive lymph nodes in an independent validation study. Breast Cancer Res Treat. 2009;116(2):295-302.

[134] Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, et al. Gene-expression profi les to predict distant metastasis of h-node-negative primary breast cancer. Lancet Oncol. 2005;365:671-9.

[135] Foekens JA, Atkins D, Zhang Y, Sweep FC, Harbeck N, Paradiso A, et al. Multicenter validation of a gene expression-based prognostic signature in h node-negative primary breast cancer. J Clin Oncol. 2006;24:1665-71.

[136] Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst 2006;98:262-72.

[137] Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. J Clin Oncol. 2007;25:1239-46.

[138] Liedtke C, Hatzis C, Symmans WF, Desmedt C, Haibe-Kains B, Valero V, et al. Genomic grade index is associated with response to chemotherapy in patients with breast cancer. . J Clin Oncol 2009;27:3185-91.

[139] Desmedt C, Giobbie-Hurder A, Neven P, Paridaens R, Christiaens MR, Smeets A, et al. The Gene expression Grade Index: a potential predictor of relapse for endocrine-treated breast cancer patients in the BIG 1-98 trial. . BMC Med Genomics 2009;2:40.

[140] Cardoso F, Van't Veer L, Rutgers E, Loi S, Mook S, Piccart-Gebhart MJ. Clinical application of the 70-gene profile: the MINDACT trial. J Clin Oncol. 2008;26:729-35.

[141] Cardoso F. Show me the genes - I will tell you who/how to treat! Breast Cancer Research. 2005;7:77-9.

[142] Handkiewicz-Junak D, Czarniecka A, Jarzab B. Molecular prognostic markers in papillary and follicular thyroid cancer: Current status and future directions. . Mol Cell Endocrinol. 2010;322:8-28.

[143] Risinger JI, Maxwell GL, Chandramouli GVR, Jazaeri A, Aprelikova O, Patterson T, et al. Microarray Analysis Reveals Distinct Gene Expression Profiles among Different Histologic Types of Endometrial Cancer. Cancer Res. 2003;63:6-11.

[144] Schwartz DR, Kardia SLR, Shedden KA, Kuick R, Michailidis G, Taylor JMG, et al. Gene Expression in Ovarian Cancer Reflects Both Morphology and Biological Behavior, Distinguishing Clear Cell from Other Poor-Prognosis Ovarian Carcinomas. Cancer Res. 2002;62:4722-9.

[145] Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L. The continuum of translation research in genomic medicine how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? . Genet Med. 2007;9:665-74.

[146] Khoury MJ, Coates RJ, Fennell ML, Glasgow RE, Scheuner MT, Schully SD, et al. Multilevel research and the challenges of implementing genomic medicine. J Natl Cancer Inst Monogr. 2012;44:112-20.

[147] Schully SD, Benedicto CB, Gillanders EM, Wang SS, Khoury MJ. Translational Research in Cancer Genetics: The Road Less Traveled. Public Health Genomics 2010;14:1-8.

[148] Maitland ML, Schilsky RL. Clinical trials in the era of personalized oncology. CA Cancer J Clin. 2011;61:365-81.

[149] Cheadle C, Fan J, Cho-Chung YS, Werner T, Ray J, Do L, et al. Stability regulation of mRNA and the control of gene expression. Ann N Y Acad Sci. 2005;1058:196–204.

[150] Thorrez L, Van Deun K, Tranchevent LC, Van Lommel L, Engelen K, Marchal K, et al. Using ribosomal protein genes as reference: a tale of caution. . PLoS ONE. 2008;3:e1854.

[151] Mutch DM, Tordjman J, Pelloux V, Hanczar B, Henegar C, Poitou C, et al. Needle and surgical biopsy techniques differentially affect adipose tissue gene expression profiles. Am J Clin Nutr. 2009;89:51-7.

[152] Kim K, Zakharkin SO, Allison DB. Expectations, validity and reality in gene expression profiling J Clin Epidemiol. 2010;63:950–9.

[153] Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, et al. The MicroArra Quality Control (MAQC) projects shoes inter- and intraplatform reproducibility of gene expression measureament. Nat Biotechnol. 2006;24:1151-61.

[154] Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J. Independence and reproducibility across microarray platforms. Nat Methods. 2005;2:337–44.

[155] Irizarry RA, Warren D, Spencer F, Kim IF, Biswal S, Frank BC, et al. Multiple-laboratory comparison of microarray platforms. Nat Methods. 2005;2:345–50.

[156] Standardizing global gene expression analysis between laboratories and across platforms, (2005).

[157] Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet. 2001 Dec;29(4):365-71.

[158] Ransohoff DF. How to improve reliability and efficiency of research about molecular markers: role of phases, guidelines and study design. J Clin Epidemiol. 2007;60:1205-19.

[159] Zakharkin SO, Kim K, Mehta T, Chen L, Barnes S, Scheirer KE, et al. Sources of variation in Affymetrix microarray experiments. BMC Bioinformatics. 2005;6:214.

[160] Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. 2003;19:185–93.

[161] Michiels S, Kramar A, Koscielny S. Multidimentionality of microarrays: Statiscal challenges and (im)possible solutions. Mol Onco. 2011;5:190-6.

[162] Kim K, Zakharkin SO, Allison DB. Expectations, validity, and reality in gene expression profiling. J Clin Epidemiol. 2010 Sep;63(9):950-9.

[163] Lee WP, Tzou WS. Computational methods for discovering gene networks from expression data. Brief Bioinform. 2009;10:408-23.

[164] Dupuy A, Simon RM. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. J Natl Cancer Inst. 2007;99:147-57.

[165] Reis-Filho JS, Weigelt B, Fumagalli D, Sotiriou C. Molecular profiling: moving away from tumor philately. Sci Transl Med. 2010;2:47ps3.

[166] Borst P, Wessels L. Do predictive signatures really predict response to cancer chemotherapy? . Cell Cycle. 2010;9:4836-40.

[167] Simon R. Roadmap for developing and validating therapeutically relevant genomic classifiers. J Clin Oncol. 2005;23:7332-41.

[168] Gökmen-Polar YBS. Molecular profiling assays in breast cancer: are we ready for prime time? Oncology (Willston Park). 2012;26(4):350-7.

[169] Simon R. Lost in translation problems and pitfalss in translating laboratory observation to clinical utility. Eur J Cancer. 2008;44:27072-2713.

[170] Simon R. Development and validation of therapeutically relevant multi-gene biomarker classifiers. J Natl Cancer Inst. 2005;97:866-7.

[171] Pusztai L, Ayers M, Stec J, Hortobagyi GN. Clinical application of cDNA microarrays in oncology. The Oncologist. 2003;8:252-8.

[172] Sparano JA, Solin LJ. Defining the clinical utility of gene expression assays in breast cancer: the intersection of science and art in clinical decision making. J Clin Oncol. 2010;28:1625-7.

[173] Sparano JA, Paik S. Development of the 21-gene assay and its application in clinical practice and clinical trials. J Clin Oncol. 2008;26:721-8.

[174] Lyman GH, Cosler LE, Kuderer NM, Hornberger J. Impact of a 21-Gene RT-PCR Assay on Treatment Decisions in Early-Stage Breast Cancer: An Economic Analysis Based on Prognostic and Predictive Validation Studies. Cancer. 2007;109(6):1011-8.

[175] Kelly CM, Krishnamurthy S, Bianchini G, Litton JK, Gonzalez-Angulo AM, Hortobagyi GN, et al. Utility of oncotype DX risk estimates in clinically intermediate risk hormone receptor-positive, HER2-normal, grade II, lymph node-negative breast cancers. Cancer. 2010;116:5161-7.

[176] Retèl VP, Joore MA, Knauer M, Linn SC, Hauptmann M, Harten WH. Cost-effectiveness of the 70-gene signature versus St. Gallen guidelines and Adjuvant Online for early breast cancer. Eur J Cancer. 2010;46:1382-91.

[177] Chen E, Tong KB, Malin JL. Cost-effectiveness of 70-gene MammaPrint signature in node-negative breast cancer. Am J Manag Care. 2010;16:e333-42.

[178] Hornberger J, Cosler LE, Lyman GH. Economic Analysis of Targeting Chemotherapy Using a 21-Gene RT-PCR Assay in Lymph-Node–Negative, Estrogen-Receptor–Positive, Early-Stage Breast Cancer. Am J Manag Care. 2005;11:313-24.

[179] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol. 2009;27:1160-7.

[180] Recommendations from the EGAPP Working Group: can tumor gene expression profiling improve outcomes in patients with breast cancer?, 11 (2009).

[181] Tang G, Cuzick J, Costantino JP, al. e. Risk of recurrence and chemotherapy benefit for patients with node-negative, estrogen receptor-positive breast cancer: recurrence score alone and integrated with pathologic and clinical factors. J Clin Oncol. 2011;29:4365-72.

[182] Pusztai L. Anatomy and biology: two complementary sides of breast cancer prognostication. J Clin Oncol. 2011;29:4347-8.

[183] Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, et al. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. Breast Cancer Res Treat. 2008;10:R65.

[184] Haibe-Kains B, Desmedt C, Sotiriou C, Bontempi G. A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all? Bioinformatics [serial on the Internet]. 2008; 24.

[185] Recommendations from the EGAPP Working Group: can tumor gene expression profiling improve outcomes in patients with breast cancer?, 11 (2009).

[186] Koscielny S, Michiels S. Clinical usefulness of microarrays for cancer prognosis in 2010--letter. n Cancer Res. 2010;16:6180-1.

[187] Marchionni L, Wilson RF, Marinopoulos SS, Wolff AC, Parmigiani G, Bass EB, et al. Impact of gene expression profiling tests on breast cancer outcomes. Evid Rep Technol Assess (Full Rep). 2007;160:1-105.

[188] Pharoah PD, Caldas C. How to validate a breast cancer prognosticature. Nat Rev Clin Oncol. 2010;7:615–6.

[189] Bartlett JM, Munro AF, Dunn JA, McConkey C, Jordan S, Twelves CJ, et al. Predictive markers of anthracycline benefit: a prospectively planned analysis of the UK National Epirubicin Adjuvant Trial (NEAT/BR9601). . Lancet Oncol. 2010;11:266–74.

[190] Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G, et al. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. Clin Cancer Res. 2008;14:5158-65.

[191] Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet Oncol. 2005;365:671-9.

[192] Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. Breast Cancer Res. 2010;12:207.

[193] Desmedt C, Piette F, Loi S, Wang Y, Lallemand F, Haibe-Kains B, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clin Cancer Res. 2007;13:3207-14.

[194] Chin S-F, Teschendorff AE, Marioni JC, Wang Y, Barbosa-Morais NL, Thorne NP, et al. High-resolution array-CGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. Genome Biol. 2007;8:R215.

[195] Blenkiron C, Goldstein LD, Thorne NP, Spiteri I, Chin SF, Dunning MJ, et al. MicroRNA expression profiling of human breast cancer identifies new markers of tumour subtype. . Genome Biol. 2007;8:R214.

[196] Russnes HG, Vollan HK, Lingjaerde OC, Krasnitz A, Lundin P, Naume B, et al. Genomic architecture characterizes tumor progression paths and fate in breast cancer patients. . Sci Transl Med. 2010;2:38ra47.

[197] Berns K, Horlings HM, Hennessy BT, Madiredjo M, Hijmans EM, Beelen K, et al. A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. . Cancer Cell. 2007;12:395-402.

[198] Ayers M, Symmans WF, Stec J, Damokosh AI, Clark E, Hess K, et al. Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. J Clin Oncol. 2004;22:2284-93.

[199] Iwao-Koizumi K, Matoba R, Ueno N, Kim SJ, Ando A, Miyoshi Y, et al. Prediction of docetaxel response in human breast cancer by gene expression profiling. J Clin Oncol. 2005;23:422-31.

[200] Hess KR, Anderson K, Symmans WF, Valero V, Ibrahim N, Mejia JA, et al. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. J Clin Oncol. 2006;24:4236-44.

[201] Thuerigen O, Schneeweiss A, Toedt G, Warnat P, Hahn M, Kramer H, et al. Gene expression signature predicting pathologic complete response with gemcitabine, epirubicin, and docetaxel in primary breast cancer. J Clin Oncol 2006;24:1839-45.

[202] Peintinger F, Anderson K, Mazouni C, Kuerer HM, Hatzis C, Lin F, et al. Thirty-gene pharmacogenomic test correlates with residual cancer burden after preoperative chemotherapy for breast cancer. . Clin Cancer Res. 2007;13:4078-82.

[203] Tabchy A, Valero V, Vidaurre T, Lluch A, Gomez H, Martin M, et al. Evaluation of a 30-gene paclitaxel, fluorouracil, doxorubicin, and cyclophosphamide chemotherapy response predictor in a multicenter randomized trial in breast cancer. Clin Cancer Res. 2010;16:5351-61.

[204] Kok M, Linn S, Laar R. Comparison of gene expression profiles predicting progression in breast cancer patients treated with tamoxifen. Breast Cancer Res Treat. 2009;113:275–83.

[205] Miller WR, Larionov A, Renshaw L, Anderson TJ, Walker JR, Krause A, et al. Gene expression profiles differentiating between breast cancer clinically responsive or resistant to letrozole. J Clin Oncol. 2009;27:1382–7.

[206] Potti A, Dressman HK, Bild A, Riedel RF, Chan G, Sayer R, et al. Genomic signatures to guide the use of chemotherapeutics. Nat Med. 2006;12:1294-300.

[207] Bonnefoi H, Potti A, Delorenzi M, Mauriac L, Campone M, Tubiana-Hulin M, et al. Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. Lancet Oncol. 2007;8:1071-8.

[208] Liedtke C, Wang J, Tordai A, Symmans WF, Hortobagyi GN, Kiesel L, et al. Clinical evaluation of chemotherapy response predictors developed from breast cancer cell lines. Breast Cancer Res Treat. 2010;121:301-9.

[209] Collingridge D. Expression of concern--validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. Lancet Oncol. 2010;11:813-4.

[210] Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, et al. A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. Nat Med. 2009;15:68-74.

[211] Ashworth A, Lord CJ, Reis-Filho JS. Genetic interactions in cancer progression and treatment. Cell. 2011;145:30-8.