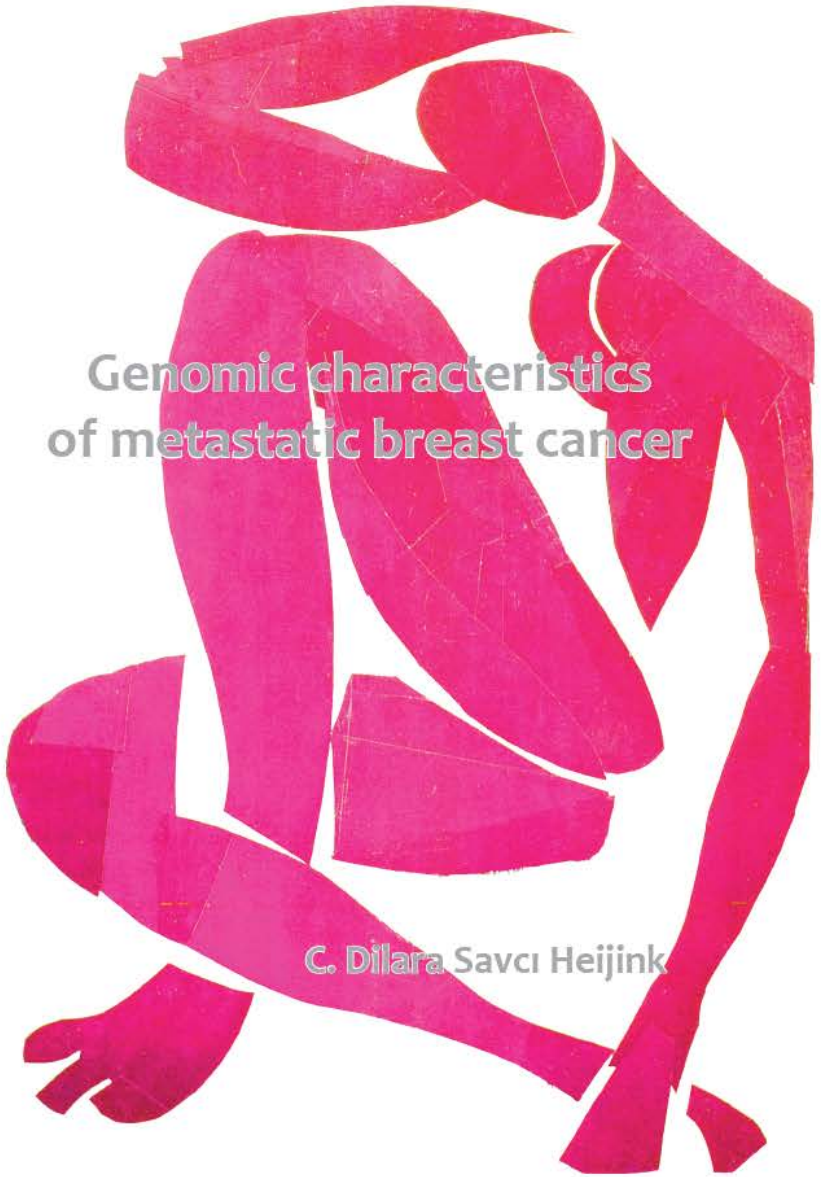
Genomic characteristics of metastatic breast cancer

Savcı Heijink, C.D.

Publication date
2019
Document Version
Final published version
License
Other

Genomic characteristics of metastatic breast cancer

C. Dilara Savcı Heijink



Genomic characteristics of metastatic breast cancer

C. Dilara Savcı Heijink

Genomic characteristics of metastatic breast cancer

Cemile Dilara Savcı Heijink

Genomic characteristics of metastatic breast cancer

PhD Thesis, University of Amsterdam, The Netherlands

The publication of this thesis was kindly supported by Center for Translational Molecular Medicine (BreastCARE), Amsterdam UMC (AMC), ChipSoft and Pfizer BV.

Cover design by Wim J.V. van Est , modified from Blue Nude II, Gouache-painted paper cut-out, by Henri Matisse completed in 1952.
Layout and design by David de Groot, persoonlijkproefschrift.nl.
Printed by Ipskamp Printing BV

ISBN/EAN: 978-94-028-1297-8

GENOMIC CHARACTERISTICS OF METASTATIC BREAST CANCER

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Agnietenkapel

op vrijdag 18 januari 2019, 14.00 uur

door Cemile Dilara Savcı Heijink

geboren te Adana, Turkije

Promotiecommissie

Promotor:	Prof. dr. M.J. van de Vijver	AMC-UvA
Copromotor:	Dr. H.M. Horlings	NKI-AvL
Overige leden:	Prof. dr. V.T.H.B.M Smit	Universiteit Leiden
	Prof. dr. P. J. van Diest	Universiteit Utrecht
	Prof. dr. L.F.A. Wessels	TU Delft
	Prof. dr. E.J.T. Rutgers	AMC-UvA
	Dr. A.D. Bins	AMC-UvA
	Dr. N. Bijker	AMC-UvA

Faculteit der Geneeskunde

For Şuşu,

The most beautiful sea,

hasn't been crossed yet

The most beautiful child,

hasn't grown up yet

Our most beautiful days,

we haven't seen yet

And the most beautiful words I wanted to tell you,

I haven't said yet...

Nâzım Hikmet Ran
September 24th 1945

Table of contents

Chapter 1 **9**

General introduction and outline of the thesis

Chapter 2 **25**

Translating the genomic architecture of breast cancer into clinical application

Horlings HM, Savci-Heijink CD, van de Vijver MJ

Sci Transl Med. 2010 Jun 30;2(38):38ps32.

Chapter 3 **41**

Retrospective analysis of metastatic behavior of breast cancer subtypes

Savci-Heijink CD, Halfwerk H, Hooijer GK, Horlings HM, Wesseling J, van de Vijver MJ

Breast Cancer Res Treat. 2015 Apr;150(3):547-57.

Chapter 4 **67**

A novel gene expression signature for bone metastasis in breast carcinomas

Savci-Heijink CD, Halfwerk H, Koster J, van de Vijver MJ

Breast Cancer Res Treat. 2016 Apr;156(2):249-59.

Chapter 5 **95**

A specific gene expression signature for visceral organ metastasis in breast cancer

Savci-Heijink CD, Halfwerk H, Koster J, Horlings HM, van de Vijver MJ

Submitted

Chapter 6 **115**

Association between gene expression profile of the primary tumor and chemotherapy response of metastatic breast cancer

Savci-Heijink CD, Halfwerk H, Koster J, van de Vijver MJ

BMC Cancer. 2017 Nov 13;17(1):755.

Chapter 7 **135**

Epithelial-mesenchymal transition status of primary breast carcinomas and its correlation with metastatic behavior

Savci-Heijink CD, Halfwerk H, Hooijer GK, Koster J, Horlings HM, Meijer SL, van de Vijver MJ

Accepted for publication in Breast Cancer Res Treat.

Chapter 8 **155**

General discussion and concluding remarks

Chapter 9 **171**

Summary in English and Dutch

Appendix **181**

Nawoord / Acknowledgement

About the author

Chapter 1

General introduction and outline of the thesis

Metastatic breast cancer

Breast cancer is the most common cancer in women in the western world and the most prevalent type of cancer in the Netherlands with a ten-year prevalence of 128.000 cases [1]. Notwithstanding the advances in early detection and improved cure rates, breast cancer remains a major cause of cancer related deaths for women [2-4]. Mortality rates are closely related to the development of distant metastases and associated complications [5, 6]. Approximately 5-10% of the patients initially present with metastatic breast cancer, while 20-30% of patients eventually develop distant metastases during the course of the disease [7, 8].

Metastatic progression in cancer is a heterogeneous process and encompasses stepwise sequential events initiating with invasion of tumor cells through the basal membrane followed by penetration into the bloodstream (intravasation). When cancer cells survive in the circulation, subsequent extravasation and eventual colonization leads to form a detectable macroscopic tumor. Considering that only less than 0.01% of tumor cells that reach the circulation give rise to an overt metastasis, metastasis has been regarded as a highly inefficient process [9, 10]. Therefore, accomplishment of this complex process requires multistep interactions between the circulating tumor cells and the microenvironment of the organ in which metastatic disease develops.

Organ-specific metastasis

Organotropism (Gk, *organon*, tool of the body + *tropism*, turning movement of a biological organism), i.e. organ-specific metastasis, depicts the concept of non-random involvement of the particular organ within a specific cancer type. Mechanisms and determinants of organotropism are intriguing and therefore are often subject of investigation for many researchers [5, 11-14]. The "*Anatomical/mechanical*" hypothesis, a rather conventional theory, proposes that the blood circulation pattern, the anatomy of the primary tumor and the surrounding vessels principally shape the metastatic spread pattern. This theory fails to fully relate the clinically observed metastasis pattern for many types of cancer. More than a century ago, *Stephen Paget* introduced a theory implicating that the essence of close interactions between the circulating tumor cells would constitute the "seed" and the microenvironment within the targeted area,

the “soil”. Since its introduction, *Paget*’s “*Seed and Soil*” hypothesis, along with the “*Anatomical/mechanical*” hypothesis, has been acknowledged by experimental and clinical research [10, 15-17].

Several experimental studies using animal models to decipher the underlying framework of these non-random distinct organ metastases, have shown that as well as extrinsic factors, tumor intrinsic factors play a substantial role in the development of metastatic disease [18-22]. These studies have investigated the underlying biology of organ-specific metastasis by using animal models of metastatic breast cancer, which were developed by injecting human breast cancer cell lines in immune compromised mice. The organ-tropic metastatic variants selected from these animal models were further analyzed by genomic profiling, which was subsequently combined with clinical genomic studies [23-30]. Some of these studies, carried out by *Massague* and colleagues, resulted in distinct gene expression profiling signatures associated with metastasis to bone, lung and brain [24-26, 30, 31]. Based on these site-specific signatures, several individual genes were further explored by means of validating these discovered genes in cohorts of primary breast tumors, to comprehend the biology of metastatic disease.

Clinical Genomics and its application in management of breast cancer

In addition to the histomorphologic features, such as histologic type and the grade of the primary tumor, tissue-based biomarkers and clinical characteristics play an important role in decision making in the management of breast cancer. Tissue-based biomarkers include hormone receptor status (Estrogen receptor, ER; Progesteron receptor, PR; Human epidermal growth factor, HER2 and proliferation index, Ki67) of the tumor and are most widely assessed in combination of immunohistochemistry and *in situ* DNA hybridization techniques. Despite compelling improvement in disease control, the heterogeneity of the breast tumors is still not fully reflected by this basic stratification approach.

In the past decades, extensive research applying genomics on breast cancers was carried out to identify the molecular characteristics of distinctive biological behavior of these tumors. *Genomics* is defined as the field of science concerned with applying high-throughput techniques to the genetic mapping and DNA sequencing of sets of genes or the complete genetic material and RNA expression profiling of selected organisms/samples. *Clinical genomics*, which employs genomic studies to improve patient care, has resulted in several molecular classifiers, particularly in the breast cancer field.

Several widely accepted molecular classifiers for breast cancer were developed based on gene expression profiling analyses of breast cancer [32-37]. Gene expression profiling experiments are based on measuring relative amounts of mRNA simultaneously for many genes and reflect the pattern of the transcription which is encoded in DNA sequences [38]. The data generated by these experiments, following a normalization step, can be analyzed in an unsupervised and a supervised manner. *Unsupervised classification* aims to group the genes or samples together with similar traits. *Hierarchical clustering* is one of the most common approaches used in unsupervised classification and operates by repetitively joining the two closest clusters from individual clusters or repetitively separating clusters starting with the initial dataset [39, 40]. Another frequently used unsupervised classification method is the *K-means clustering* algorithm which works by classifying the given data set into a certain number of clusters chosen a priori. This algorithm repeatedly calculates the center points for each cluster, following initial separation of the linear space into given *K* components. Using *unsupervised hierarchical clustering* analysis, several investigators have observed distinct gene expression traits of ER-positive and ER-negative tumors [34, 35]. These analyses have led to the discovery of several subgroups within breast carcinomas with distinct clinical behavior. Based on these studies different “intrinsic” subtypes of breast carcinomas are defined: luminal A and luminal B tumors (ER-positive tumors characterized by gene expression pattern similar to breast luminal cells), HER2-like tumors (ER-negative tumors with HER2 gene overexpression), basal-like tumors (ER-negative tumors with gene expression pattern overlapping with myoepithelial cells) and normal-like tumors. Even though a very similar classification of breast carcinomas into subgroups can be done by immunohistochemical analysis of ER, PR and HER2, the results are not consistently in concordance with the ‘intrinsic’ subtypes assigned by molecular classification of these tumors [41].

Supervised classification is described as a knowledge-driven classification of a data set to design a classifier with prognostic and predictive value. Usually a subset of tumors with known characteristics are used to train a model to classify samples and subsequently this classifier is tested in another dataset to create similar groups. Using the supervised classification approach, several gene sets have been described with the purpose of identifying breast tumors with distinctive clinical behavior [34, 35, 37, 42], eventual prediction of the metastatic potential and response to therapy of the tumors [24-26]. Some of these gene expression profiling studies have led to commercially available gene-expression-based molecular tests [32, 33, 36, 37, 43, 44]. The 21-gene recurrence classifier (Oncotype DX®) is one of these molecular tests that estimates the recurrence score (RS), which can be defined as a risk of developing distant metastasis at 10 years for the patients with ER-positive and lymph-node negative breast cancer. This estimation is based on the expression level of 21 genes (16 cancer related genes and 5 reference genes) and can be easily applied as a clinical assay on formalin-fixed, paraffin-embedded tumor tissue of breast cancer patients. This test is most commonly used (most widely in the USA) as a prognostic indicator in addition to tissue-based markers and clinicopathological characteristics of the primary tumor in decision-making regarding administration of chemotherapy [32, 45, 46]. Another well-known prognostic classifier and a companion diagnostic tool is the 70-gene signature (MammaPrint®). The 70-gene prognosis signature, classifies the breast tumors into 2 groups as a “good” or “poor” prognosis, and gives an estimation of developing distant metastasis, mainly in lymph-node negative breast cancer patients with tumor size <5 cm [37]. The prognostic value of this signature has been validated in large-scale prospective studies and has shown that this signature is also valid to predict the outcome in breast cancer patients with 1 to 3 lymph node metastases [47].

Clinical management of metastatic breast cancer

Metastatic breast cancer is considered to be an incurable disease. The main aim of treating metastatic breast cancer is to prolong survival of the patients with acceptable toxicity and to palliate the disease-related symptoms. The decisions of treating patients with metastatic disease and the choice of therapy (hormonal, chemotherapeutic and/or

targeted) depends on the patient's age/performance, site and the number of the distant metastases, hormone receptor status, HER2 status, menopausal status, type and extent of prior adjuvant therapy, and time between the last administered therapy [48-50].

Hormonal therapy

For patients with hormone receptor positive metastatic breast cancer, hormone therapy plays an important role. Even though a well-established consensus on the sequence is lacking, in the absence of visceral crisis (severe organ dysfunction as assessed by signs and symptoms, laboratory studies, and rapid progression of disease) two to four lines of hormonal therapy may be administered before considering to start chemotherapeutic therapy. For the patients who are premenopausal at the initial diagnosis, surgical (oophorectomy) or medical (gonotropin-releasing hormone analogs) introduction of menopause is indicated as an effective therapeutic action. First-line hormonal therapy options for the postmenopausal patients include aromatase inhibitors, selective estrogen receptor modulators (tamoxifen) or a selective estrogen receptor degrader (fulvestrant). Combination therapies with inclusion of cyclin-dependent kinase 4/6 inhibitors or everolimus can be considered in the treatment of metastatic breast cancer as second or further-line therapy [50-52].

Chemotherapy

In the metastatic setting, administration of chemotherapy is considered for the patients with hormone receptor negative tumors and for patients with hormone receptor positive disease resistant to hormonal therapy and rapidly progressive disease. Choice of chemotherapeutic agent depends prior (adjuvant) chemotherapeutic treatments and the status of the patient. For the first-line chemotherapy, anthracyline- and/or taxane-based regimens are the most commonly used therapy agents [49-51, 53]. In case of resistance to anthracyclines or taxanes, an antimetabolite cytotoxic agent (capecitabine) is usually used as a second-line chemotherapy agent [54].

Targeted therapy

HER2 is amplified and overexpressed in approximately 15% of breast carcinomas, and the tumors with HER2 overexpression are associated with shorter overall survival times [55-57]. For the patients with HER2-positive tumors and metastastic disease, targeted therapy including combination of HER2- receptor antagonists (trastuzumab,

pertuzumab) and taxanes is recommended [49, 50, 58]. As second-line targeted therapy option or in case of development of metastases within 6 months after completion of an anti-HER2 adjuvant therapy, administration of ado-trastuzumab emtansine (T-DMI) is indicated [59]. Dual inhibitors of epidermal growth factor 1 (EGFR1) and HER2 (lapatinib) can also be used as a single therapy agent or in combination with other chemotherapeutics in treatment of metastatic breast carcinoma [60].

Immunotherapy

The immune microenvironment of the tumor has been shown to play an important role in cancer progression. Cancer immunotherapy aims to boost antitumor immune response and has been historically implemented in passive forms (Interferons, Interleukin-2, Calmette-Guerin Bacillus and antibody dependent cell-mediated cytotoxicity) in various cancer types [61-65]. Immunotherapy has evolved to a therapeutic approach using an immune checkpoint blockade (ICB) with promising results in a variety of tumor types [66-70]. Several recent studies have pointed out the association between the immune-related genes and the tumor-infiltrating lymphocytes with prognosis in triple-negative and HER2-positive breast cancer subtypes [71-74]. Subsequent clinical trials investigating the usage of novel immunomodulatory agents in treatment of advanced breast cancer, have revealed that the blockading of programmed cell death protein 1(PD1)/programmed death ligand 1 (PDL1) by means of monotherapy or in combination with chemotherapy/anti-cytotoxic T lymphocyte antigen-4 (CTLA4) was paired with prolonged survival times in patients with triple negative tumors [75-77]. More recently, it has been showed that the PD-1/PD-L1 blockade in combination with chemotherapy resulted in increased pathologic response rates in the neoadjuvant setting for triple negative and HER2-negative breast tumors [78-80]. Despite the emerging role of immunotherapy in treatment of advanced breast cancer, refinement of patient selection and evaluation of response to these agents remain unsolved [81].

Evaluation of response to systemic therapy

Assessment of response to given therapeutic agents is determined by estimating the changes in tumor burden. To evaluate these changes, a widely accepted guideline, *Response Evaluation of Solid Tumors* (RECIST) *criteria*, which was introduced by the European Organization and Treatment of Cancer (EORTC), the National Cancer Institute of the United States and the National Cancer Institute of Canada Clinical Trials Group,

is commonly used [82]. According to the RECIST criteria, evaluation of an objective response is classified as; complete response (CR), partial response (PR), stable disease (SD) and progressive disease (PD). This assessment of therapy response in metastatic breast carcinoma is performed by means of imaging techniques and tumor marker determination [50].

Within the metastatic setting, response to combined chemotherapy agents varies between 50-70% [48, 83, 84]. To this day, no consensus exists on the optimal chemotherapy options or the sequence and duration of the given chemotherapeutics. As the current therapy approaches for metastatic breast cancer are mainly of a trial-and-error approach, predictors of response are needed for optimization the benefit and minimization of the toxic side-effects of the therapy.

Radiotherapy may be implemented to palliate the pain and discomfort for the patients with metastatic disease.

Rationale of this thesis

The aim of this thesis is to identify the characteristics of primary breast tumors that are predictive of metastatic behavior in terms of organ-specific metastasis, response to systemic therapy and associated patient outcomes.

Outline of this thesis

Chapter 1 provides a brief introduction to metastatic breast cancer and genomic research of breast cancer.

In **chapter 2,** a viewpoint on genomic alterations of breast cancer and their translation into clinical application is presented (in conjunction with article by *Russness* et al: “

Genomic architecture characterizes tumor progression paths and fate in breast cancer patients" [85]).

Chapter 3 describes a retrospective study, analyzing the clinicopathologic features of primary breast carcinomas focusing on the association between breast cancer subtypes and their metastatic behavior, including site-specific metastasis and metastasis specific survival outcomes.

In **chapters 4** and **5**, we describe our results of gene expression profiling experiments of 157 primary breast carcinomas of patients who all developed distant metastases. The correlation of the gene expression profiling data to metastatic behavior is reported. Furthermore, specific gene expression profiling signatures, which were found to be associated with development of bone metastasis (in **chapter 4**) and visceral organ metastasis (in **chapter 5)**, are presented.

In **chapter 6**, we investigated the link between primary breast carcinoma features and the chemotherapy response in the frame of metastatic disease. We sought to develop genomic identifiers of chemotherapy responsiveness by comparing the gene expression profiling of primary tumors of the responders and non-responders.

In **chapter 7**, the concept of epithelial-to mesenchymal transition (EMT) in metastatic breast cancer is elaborated. Using a generic EMT-core signature, the EMT-status of each primary breast tumor in our data set was assessed and compared with tumor characteristics and their metastasis pattern. The concept to reconcile the EMT-status of the tumor and to identify the tumor cells with EMT-phenotype, using conventional immunohistochemistry, is also explored.

Chapter 8 provides a general discussion followed by concluding remarks, based on the findings generated in the abovementioned studies which form this thesis.

In **chapter 9**, a brief summary of this thesis, in the English and Dutch language, is presented.

References

1. Rijksinstituut voor gezondheid en milieu. *Public health info, breast cancer in the Netherlands*. 2018; Available from: https://www.rivm.nl/en/Topics/B/Breast_cancer_screening_programme/Breast_cancer_in_the_Netherlands.
2. Malvezzi, M., et al., *European cancer mortality predictions for the year 2012*. AnnOncol, 2012. 23(4): p. 1044-1052.
3. Sant, M., et al., *Breast carcinoma survival in Europe and the United States*. Cancer, 2004. 100(4): p. 715-22.
4. Thomssen, C., et al., *Follow-up after breast cancer diagnosis*. Breast Care (Basel), 2013. 8(6): p. 457-60.
5. Chambers, A.F., A.C. Groom, and I.C. MacDonald, *Dissemination and growth of cancer cells in metastatic sites*. NatRevCancer, 2002. 2(8): p. 563-572.
6. Chiang, A.C. and J. Massague, *Molecular basis of metastasis*. NEnglJMed, 2008. 359(26): p. 2814-2823.
7. Early Breast Cancer Trialists' Collaborative, G., *Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials*. Lancet, 2005. 365(9472): p. 1687-717.
8. O'Shaughnessy, J., *Extending survival with chemotherapy in metastatic breast cancer*. Oncologist, 2005. 10 Suppl 3: p. 20-9.
9. Fidler, I.J., *Metastasis: quantitative analysis of distribution and fate of tumor emboli labeled with 125 I-5-iodo-2'-deoxyuridine*. J Natl Cancer Inst, 1970. 45(4): p. 773-82.
10. Langley, R.R. and I.J. Fidler, *The seed and soil hypothesis revisited--the role of tumor-stroma interactions in metastasis to different organs*. IntJCancer, 2011. 128(11): p. 2527-2535.
11. Aslakson, C.J. and F.R. Miller, *Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor*. Cancer Res, 1992. 52(6): p. 1399-1405.
12. Ding, L., et al., *Genome remodelling in a basal-like breast cancer metastasis and xenograft*. Nature, 2010. 464(7291): p. 999-1005.
13. Hu, G., Y. Kang, and X.F. Wang, *From breast to the brain: unraveling the puzzle of metastasis organotropism*. JMolCell Biol, 2009. 1(1): p. 3-5.
14. Lorusso, G. and C. Ruegg, *New insights into the mechanisms of organ-specific breast cancer metastasis*. SeminCancer Biol, 2012. 22(3): p. 226-233.
15. Fidler, I.J., *The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited*. Nat Rev Cancer, 2003. 3(6): p. 453-8.
16. Lu, X. and Y. Kang, *Organotropism of breast cancer metastasis*. JMammaryGlandBiolNeoplasia, 2007. 12(2-3): p. 153-162.

17. Paget, S., *The distribution of secondary growths in cancer of the breast. 1889.* Cancer Metastasis Rev, 1989. 8(2): p. 98-101.

18. Chi, J.T., et al., *Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers.* PLoS Med, 2006. 3(3): p. e47.

19. Feinberg, A.P., R. Ohlsson, and S. Henikoff, *The epigenetic progenitor origin of human cancer.* Nat Rev Genet, 2006. 7(1): p. 21-33.

20. Fidler, I.J. and G. Poste, *The "seed and soil" hypothesis revisited.* Lancet Oncol, 2008. 9(8): p. 808.

21. Nuyten, D.S., et al., *Combining biological gene expression signatures in predicting outcome in breast cancer: An alternative to supervised classification.* Eur J Cancer, 2008. 44(15): p. 2319-29.

22. Semenza, G.L., *Targeting HIF-1 for cancer therapy.* Nat Rev Cancer, 2003. 3(10): p. 721-32.

23. Bos, P.D., D.X. Nguyen, and J. Massague, *Modeling metastasis in the mouse.* Curr Opin Pharmacol, 2010. 10(5): p. 571-7.

24. Bos, P.D., et al., *Genes that mediate breast cancer metastasis to the brain.* Nature, 2009. 459(7249): p. 1005-1009.

25. Kang, Y., et al., *A multigenic program mediating breast cancer metastasis to bone.* Cancer Cell, 2003. 3(6): p. 537-49.

26. Landemaine, T., et al., *A six-gene signature predicting breast cancer lung metastasis.* Cancer Res, 2008. 68(15): p. 6092-6099.

27. Minn, A.J., et al., *Lung metastasis genes couple breast tumor size and metastatic spread.* Proc Natl Acad Sci U S A, 2007. 104(16): p. 6740-5.

28. Minn, A.J., et al., *Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors.* J Clin Invest, 2005. 115(1): p. 44-55.

29. Zhang, X.H., et al., *Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma.* Cell, 2013. 154(5): p. 1060-73.

30. Zhang, X.H., et al., *Latent bone metastasis in breast cancer tied to Src-dependent survival signals.* Cancer Cell, 2009. 16(1): p. 67-78.

31. Harrell, J.C., et al., *Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse.* Breast Cancer ResTreat, 2012. 132(2): p. 523-535.

32. Paik, S., et al., *A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.* N Engl J Med, 2004. 351(27): p. 2817-26.

33. Parker, J.S., et al., *Supervised risk predictor of breast cancer based on intrinsic subtypes.* J Clin Oncol, 2009. 27(8): p. 1160-7.

34. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. 406(6797): p. 747-752.

35. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.* Proc Natl Acad Sci U S A, 2001. 98(19): p. 10869-74.

36. Sotiriou, C., et al., *Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis.* J Natl Cancer Inst, 2006. 98(4): p. 262-72.

37. van de Vijver, M.J., et al., *A gene-expression signature as a predictor of survival in breast cancer.* NEnglJMed, 2002. 347(25): p. 1999-2009.

38. Brazma, A. and J. Vilo, *Gene expression data analysis.* FEBS Lett, 2000. 480(1): p. 17-24.

39. Alon, U., et al., *Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays.* Proc Natl Acad Sci U S A, 1999. 96(12): p. 6745-50.

40. Eisen, M.B., et al., *Cluster analysis and display of genome-wide expression patterns.* Proc Natl Acad Sci U S A, 1998. 95(25): p. 14863-8.

41. van de Vijver, M.J., *Molecular tests as prognostic factors in breast cancer.* Virchows Arch, 2014. 464(3): p. 283-91.

42. van 't Veer, L.J., et al., *Gene expression profiling predicts clinical outcome of breast cancer.* Nature, 2002. 415(6871): p. 530-536.

43. Filipits, M., et al., *A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors.* Clin Cancer Res, 2011. 17(18): p. 6012-20.

44. Ma, X.J., et al., *A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen.* Cancer Cell, 2004. 5(6): p. 607-16.

45. Albain, K.S., et al., *Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial.* Lancet Oncol, 2010. 11(1): p. 55-65.

46. Goldhirsch, A., et al., *Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011.* AnnOncol, 2011. 22(8): p. 1736-1747.

47. Mook, S., 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): p. 295-302.

48. Johnston, S.R., *The role of chemotherapy and targeted agents in patients with metastatic breast cancer.* Eur J Cancer, 2011. 47 Suppl 3: p. S38-47.

49. Jones, S.E., *Metastatic breast cancer: the treatment challenge.* Clin Breast Cancer, 2008. 8(3): p. 224-33.

50. Oncoline Richtlijnen Oncologische zorg. *Gemetastaseerde borstkanker*. 2018; Available from: https://www.oncoline.nl/borstkanker.

51. Cardoso, F., et al., *3rd ESO-ESMO International Consensus Guidelines for Advanced Breast Cancer (ABC 3).* Ann Oncol, 2017. 28(12): p. 3111.

52. Rugo, H.S., et al., *Endocrine Therapy for Hormone Receptor-Positive Metastatic Breast Cancer: American Society of Clinical Oncology Guideline.* J Clin Oncol, 2016. 34(25): p. 3069-103.

53. Beslija, S., et al., *Second consensus on medical treatment of metastatic breast cancer.* Ann Oncol, 2007. 18(2): p. 215-25.

54. Karachaliou, N., et al., *A multicenter phase II trial of docetaxel and capecitabine as salvage treatment in anthracycline- and taxane-pretreated patients with metastatic breast cancer.* Cancer Chemother Pharmacol, 2012. 70(1): p. 169-76.

55. Heitz, F., et al., *Cerebral metastases in metastatic breast cancer: disease-specific risk factors and survival.* AnnOncol, 2011. 22(7): p. 1571-1581.

56. Kennecke, H., et al., *Metastatic behavior of breast cancer subtypes.* JClinOncol, 2010. 28(20): p. 3271-3277.

57. Wolff, A.C., et al., *American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer.* ArchPatholLab Med, 2007. 131(1): p. 18-43.

58. Lin, N.U., et al., *International guidelines for management of metastatic breast cancer (MBC) from the European School of Oncology (ESO)-MBC Task Force: Surveillance, staging, and evaluation of patients with early-stage and metastatic breast cancer.* Breast, 2013. 22(3): p. 203-10.

59. Dieras, V., et al., *Trastuzumab emtansine versus capecitabine plus lapatinib in patients with previously treated HER2-positive advanced breast cancer (EMILIA): a descriptive analysis of final overall survival results from a randomised, open-label, phase 3 trial.* Lancet Oncol, 2017. 18(6): p. 732-742.

60. de Azambuja, E., et al., *Lapatinib with trastuzumab for HER2-positive early breast cancer (NeoALTTO): survival outcomes of a randomised, open-label, multicentre, phase 3 trial and their association with pathological complete response.* Lancet Oncol, 2014. 15(10): p. 1137-46.

61. Achkar, T. and A.A. Tarhini, *The use of immunotherapy in the treatment of melanoma.* J Hematol Oncol, 2017. 10(1): p. 88.

62. Eggermont, A.M., et al., *Adjuvant therapy with pegylated interferon alfa-2b versus observation alone in resected stage III melanoma: final results of EORTC 18991, a randomised phase III trial.* Lancet, 2008. 372(9633): p. 117-26.

63. Ferris, R.L., E.M. Jaffee, and S. Ferrone, *Tumor antigen-targeted, monoclonal antibody-based immunotherapy: clinical response, cellular immunity, and immunoescape.* J Clin Oncol, 2010. 28(28): p. 4390-9.

64. Morales, A., *Treatment of superficial bladder cancer.* Can Med Assoc J, 1980. 122(10): p. 1133-8.

65. Ratta, R., et al., *Immunotherapy advances in uro-genital malignancies.* Crit Rev Oncol Hematol, 2016. 105: p. 52-64.

66. Ferris, R.L., et al., *Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck.* N Engl J Med, 2016. 375(19): p. 1856-1867.

67. Motzer, R.J., et al., *Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma.* N Engl J Med, 2015. 373(19): p. 1803-13.

68. Reck, M., et al., *Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer.* N Engl J Med, 2016. 375(19): p. 1823-1833.

69. Robert, C., et al., *Pembrolizumab versus Ipilimumab in Advanced Melanoma.* N Engl J Med, 2015. 372(26): p. 2521-32.

70. Rosenberg, J.E., et al., *Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial.* Lancet, 2016. 387(10031): p. 1909-20.

71. Bohling, S.D. and K.H. Allison, *Immunosuppressive regulatory T cells are associated with aggressive breast cancer phenotypes: a potential therapeutic target.* Mod Pathol, 2008. 21(12): p. 1527-32.

72. Ingold Heppner, B., et al., *Tumor-Infiltrating Lymphocytes: A Predictive and Prognostic Biomarker in Neoadjuvant-Treated HER2-Positive Breast Cancer.* Clin Cancer Res, 2016. 22(23): p. 5747-5754.

73. Salgado, R., et al., *The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014.* Ann Oncol, 2015. 26(2): p. 259-71.

74. Savas, P., et al., *Clinical relevance of host immunity in breast cancer: from TILs to the clinic.* Nat Rev Clin Oncol, 2016. 13(4): p. 228-41.

75. Adams, S., et al., *Phase 2 study of pembrolizumab as first-line therapy for PD-L1-positive metastatic triple-negative breast cancer (mTNBC): Preliminary data from KEYNOTE-086 cohort B.* Journal of Clinical Oncology, 2017. 35.

76. Adams, S., et al., *Phase 2 study of pembrolizumab (pembro) monotherapy for previously treated metastatic triple-negative breast cancer (mTNBC): KEYNOTE-086 cohort A.* Journal of Clinical Oncology, 2017. 35.

77. Nanda, R., et al., *Pembrolizumab in Patients With Advanced Triple-Negative Breast Cancer: Phase Ib KEYNOTE-012 Study.* J Clin Oncol, 2016. 34(21): p. 2460-7.

78. Nanda, R., et al., *Pembrolizumab plus standard neoadjuvant therapy for high-risk breast cancer (BC): Results from I-SPY 2.* Journal of Clinical Oncology, 2017. 35.

79. Pusztai, L., et al., *Safety of MED14736 (anti-PD-Li antibody) administered concomitant with weekly nab-paclitaxel and dose dense doxorubicin/cyclophosphamic (ddAC) as neoadjuvant chemotherapy for stage I-III triple negative breast cancer (TNBC): A Phase I/II trial.* Journal of Clinical Oncology, 2017. 35.

80. Schmid, P., et al., *Pembrolizumab (pembro) plus chemotherapy (chemo) as neoadjuvant treatment for triple negative breast cancer (TNBC): Preliminary results from KEYNOTE-173.* Journal of Clinical Oncology, 2017. 35.

81. Solinas, C., et al., *Targeting immune checkpoints in breast cancer: an update of early results.* ESMO Open, 2017. 2(5): p. e000255.

82. Eisenhauer, E.A., et al., *New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1).* EurJCancer, 2009. 45(2): p. 228-247.

83. El Saghir, N.S., et al., *Treatment of metastatic breast cancer: state-of-the-art, subtypes and perspectives.* Crit Rev Oncol Hematol, 2011. 80(3): p. 433-49.

84. Hortobagyi, G.N., *Treatment of breast cancer.* N Engl J Med, 1998. 339(14): p. 974-84.

85. Russnes, H.G., et al., *Genomic architecture characterizes tumor progression paths and fate in breast cancer patients.* Sci Transl Med, 2010. 2(38): p. 38ra47.

Chapter 2

Translating the genomic architecture of breast cancer into clinical application

Hugo M. Horlings, C. Dilara Savci Heijink, Marc J. van de Vijver

Sci Transl Med. 2010 Jun 30;2(38):38ps32

Abstract

The genetic in breast cancer have in recent years been studied through a variety of techniques: analysis of alterations in individual oncogenes and tumor suppressor genes; gene expression profiling of both messenger RNA and microRNA; global analysis of DNA copy number changes; and most recently, whole-genome sequence analysis. Analysis of the association between genetic alterations and gene expression profiles with prognosis and response to specific treatments will lead to improved possibilities for patient-tailored treatment. Russnes et al. now add an additional view on the complex genetic makeup of breast carcinomas by developing algorithms that can be used to subclassify tumors based on their patterns of genome-wide DNA copy number gains and losses, which vary from very simple (only a few gains and losses) to complex. The algorithms provide indices that can be used in conjunction with results from other genetic analyses to subclassify breast cancer, with the aim of defining subgroups of patients that differ with respect to prognosis and response to therapy.

Introduction

Breast cancer is markedly heterogeneous with respect to distinctive biological characteristics and clinical behavior; this attribute is also reflected by the heterogeneity in genetic alterations that have been identified by analyzing large series of tumors. Breast cancer, like all malignancies, arises from a multistep process of genetic alterations in oncogenes and tumor suppressor genes that affect the function of individual genes and cellular processes [1]. DNA copy number alterations are a reflection of the genetic aberrations of tumors: An increase from the two copies of a gene present in a normal diploid genome to several copies (usually 10 or more) represents gene amplification, producing a "gain of function" in the affected tumor cell, and is one way to activate a "normal" proto-oncogene to become an oncogene. Inactivation of a tumor suppressor gene by the mutation of one tumor suppressor gene allele and complete loss of the second allele, producing a "loss of function" in the tumor cell, is another step contributing to a malignant phenotype [2].

Because genetic alterations are the cause of cancer development, it is expected that the combination of specific genetic alterations in tumors will be predictive of clinical behavior [3-5]. The need to recognize genetically defined subtypes of breast cancer is enhanced by the increasing availability of specific and more effective therapy regimens. One important example of a genetically defined tumor type that guides treatment is human epidermal growth factor receptor 2 (HER2)–positive breast cancer; a positive HER2 status predicts the response to HER2 targeted therapy such as trastuzumab (Herceptin) [6]. Meticulous selection of patients for specific therapies could lead to improved treatment outcomes. Determining the molecular mechanism of primary or acquired drug resistance can be critical for identifying patients that will fail to respond to therapy and might help to design more efficient treatment protocols [7]. Genome-wide approaches for the identification of molecular genetic changes therefore provide powerful instruments to study cancer. Relevant molecular techniques include cytogenetic banding, spectral karyotyping, analysis of loss of heterozygosity, fluorescent and chromogenic in situ hybridization, and comparative genomic hybridization (CGH). In this issue of Science Translational Medicine, Russnes *et al.* [8] describe two algorithms

to measure changes in genomic architecture using data from CGH experiments; one measurement independently predicts breast cancer outcome.

CGH

Array-based CGH (aCGH) offers a good approach to screen whole genomes for a detailed analysis of DNA copy alterations [9]. In a CGH experiment, total genomic DNA is isolated from test cell populations (tumor tissue) and reference cell populations (normal tissue), differentially labeled with green and red fluorescent dyes, mixed in a 1:1 ratio, and hybridized to a microarray containing DNA fragments representing the whole genome, which allows the binding of sequences at different genomic locations to be distinguished [9]. Unlabeled human Cot-1 DNA (placental DNA that is enriched in repetitive DNA sequences) is contained in the mix to block nonspecific hybridization. Data processing of the scanned microarray slide includes signal intensity measurements with specialized image software and a fluorescent microscope. Deviation from 1:1 log-scaled intensity ratios (green/red) is counted as a change in DNA copy number. Normally, a threshold is set (gain, ratio 1.2; loss, 0.75), and statistical verification is applied (95 to 99% confidence intervals) [10].

The sensitivity of aCGH methods depends on the proportion of tumor cells in the tissue (a desirable proportion is > 70%) and the extent of the aberration. Smaller alterations in size [for example, a few hundred base pairs (bp)] and copy number are more difficult to detect than larger changes. The sensitivity of the technique has improved with the advent of high-resolution aCGH platforms. The resolution for the identification of genomic gains and losses is determined by the distance between two contiguous probes and varies depending on the type of probe: for example, bacterial artificial chromosome (BAC) clones (with a length of 100 to 200 kb), cDNA clones (~100 to 1000 bp), or oligonucleotides (30 to 100 bp). A drawback of aCGH is that it only recognizes physical changes in DNA copy number and is unable to identify smaller variations in DNA sequence or balanced chromosomal translocations (rearrangements that do not involve the loss or gain of any genetic material). aCGH has multiple applications; it can be used (i) to identify genomic regions that harbor oncogenes and tumor suppressor

genes that have been amplified or deleted, and (ii) to build class discovery tools for categorizing independent breast cancers [11].

Identifying aberrant chromosomal regions

An important challenge during the analysis of aCGH data is the detection of regions of concentrated high or low fluorescence ratios-that is, aberrant chromosomal regions specific to the problem under study. Broadly, there are two obstacles: (i) determining the statistical significance of the alteration and (ii) defining the boundaries of the alteration. To reach these goals, different approaches and algorithms have been used [12, 13].

The first approach uses only aCGH data. Amplifications and deletions in each sample are individually identified, and common aberrations between the samples are sought. The identification of amplifications and deletions can be simply done by setting a threshold and determining which DNA probes (which can be BAC clones, cDNA clones, or oligonucleotides) result in hybridization signals that exceed the threshold. These regions are considered to be amplified or deleted [14]. More complex algorithms employ the fact that copy number changes involve chromosome segments; ratios at contiguous sets of probes should be identical, except for occasional abrupt steps to another level (indicating a chromosomal breakpoint) [13]. Identification of these breakpoints is referred to as "segmentation" and produces "segmented data."

A second approach to detect aberrations across samples is to use gene expression data together with the chromosomal location of the genes [15]. This approach assumes that amplification directly affects gene expression. Therefore, the genes in an amplified region should have a detectable common overexpression. Similarly, the genes located in a deleted region would have a detectable underexpression. Because the alteration in expression may be caused by mechanisms other than a change in copy number, the potentially underlying chromosomal aberrations would need to be verified either by polymerase chain reaction or fluorescent in situ hybridization, if the number of loci to be tested is tractable; otherwise, the alterations would need to be confirmed by aCGH data.

A third approach combines aCGH and expression data to detect regions of chromosomal aberrations. The stepwise linkage analysis of microarray signatures (or SLAM) algorithm

[16] is an excellent example of this approach. First, significance analysis of microarrays (or SAM) analysis [17] is applied to the aCGH data to identify the DNA probes that distinguish tumor versus normal DNA. Then, the focus is on the DNA probes that display hybridization patterns that are correlated with the gene expression pattern. An algorithm to study this correlation is supervised identification of regions of aberration (or SIRAC) of aCGH data sets [12], which has been used to identify chromosomal regions associated with the classes of breast tumors defined by prognostic gene expression signatures or clinical and pathological characteristics.

Genomic alterations in breast cancer

Using these different approaches and algorithms, frequently observed genomic alterations in breast carcinomas include the gain of chromosomal regions 1q, 8q, 16p, 17q, and 20q; the loss of 16q and 17p; and DNA amplification in 8q12-24, 11q11-13, 17q12- 21, 17q22-24, and 20q13-ter (reviewed by Reis-Filho *et al.* [3]) . The histological grade of the tumor is strongly associated with the amount and complexity of the genomic aberrations; tumors with higher histological grades harbor more chromosomal alterations. Well-differentiated tumors (grade 1) often show only gain at 1q and loss of 16q, whereas poorly differentiated tumors (grade 3) exhibit more amplifications and less frequent loss of 16q [3].

aCGH data have also been correlated with the prognosis of patients with breast cancer. Tumors from patients with a poor prognosis exhibited significantly more changes than those from patients with a good prognosis. Additionally, DNA gains and losses have been shown to vary between tumors with different prognostic gene expression signatures and clinical and pathological features. For example, the previously identified 70-gene signature indicating poor prognosis is associated with the gain of 3q26-27, 8q22-24, and 17q24-25; the 70-gene good prognosis profile is associated with the loss of 16q11-24 [18].

The wealth of data derived from aCGH experiments has the potential to improve biological understanding of breast cancer development and progression, and in

combination with well-known clinical and pathological prognostic markers may also result in improved prediction of prognosis and response to breast cancer therapy [19]. Despite these ample data, translation into clinical practice remains a challenge. Until a resultant model can be of practical use, some limiting factors may hamper progress: (i) The size of sample sets available for microarray-based studies has so far been limited, (ii) studies often include a heterogeneous mix of patients with respect to clinical stage and treatment received, (iii)combining data sets to increase their size has been challenging because various types of array platforms have been used, (iv) validation of molecular classification in independent data sets has so far been limited, (v) analyses of high dimensional data are complex, and (vi) new genetic factors should demonstrate improved prediction accuracy over the combination of standard prognostic factors [20, 21].

Patterns of genomic aberrations

In the article published in this issue of Science Translational Medicine, a team of Norwegian, U.S., Swedish, and British scientists collaborated to overcome some of these obstacles. Russnes et *al.* (8) developed an objective estimate of genome-wide architectural distortion and investigated the ability of this marker of genomic complexity to provide prognostic information. They used aCGH data from 4 clinical cohorts, including 595 breast carcinomas. Their aim to develop such an estimate was inspired by three different patterns of segmented genome profiles “Simplex,” “Firestorm,” and “Sawtooth” (Fig. 1) that were previously visually recognized by Hicks et *al.* [22]. These patterns were presented as tools for distinguishing distinct processes of genomic rearrangement. Simplex has broad segments of duplications and deletions, usually comprising entire chromosomes or chromosome arms, with occasional isolated narrow peaks of amplification (Fig. 1A). Sawtooth is characterized by many narrow segments of duplications and deletions, more or less affecting all the chromosomes. Typically, the events in these tumors do not involve high copy number amplification, although little of the genome remains at the normal copy number (Fig. 1B). Firestorm resembles the Simplex type, except that the genomes contain at least one localized region of clustered, relatively narrow peaks of amplification, with each cluster confined to a single

chromosome arm. In these tumors, the amplifications often occur at high copy number (Fig. 1C) [22].

Russnes *et al*. [8] developed two new algorithms to estimate genomic complexity objectively using segmented aCGH data: (i) whole-arm aberration index (WAAI) and (ii) complex armwise aberration index (CAAI). This approach to classifying tumors is new, because the criteria include not only specific genomic regions but also the architectural type of rearrangement, such as the gain or loss of whole chromosome arms. Segmentation was performed on data from three different studies that each used slightly different methods to obtain DNA copy number data. Because the authors aimed to pool all the segmented aCGH profiles, they scaled different parameters to obtain roughly equal segmentation resolutions for the three studies.

The WAAI score was designed to capture events that involve whole chromosome arms rather than more localized gains and losses of DNA; this was done to reflect underlying defects in DNA maintenance, such as processes that lead to the formation of isochromosomes (which lack one arm and contain a duplication of the remaining arm) and translocations with a breakpoint close to the centromere. Russnes *et al*. [8] defined chromosomal arms with a WAAI score ≥ 0.8 as whole arms, and arms with a score of ≤–0.8 as whole-arm losses. In contrast, CAAI measures local distortion to recognize regions with structural complexity. For each breakpoint found by the segmentation algorithm, three scores were calculated: (i) the proximity to neighboring breakpoints, (ii) the magnitude of change, and (iii) a weight of importance. Areas of complex rearrangements were found by selecting chromosome arms with a CAAI score ≥ 0.5.

To define subgroups based on genomic architecture, the authors first distinguished four groups of tumors, based on previously identified genomic alterations that tend to occur in different breast cancer subtypes: (i) those with whole-arm gain of 1q and/or loss of 16q (group A), (ii) those with regional loss on 5q and/or gain on 10p (group B), (iii) those with both group A and group B alterations (group AB), and (iv) those with neither (group C). The subgroups displayed pronounced differences with respect to the number of whole–chromosome arm loss or gain events. To characterize these groups further, each was split into two CAAI sub- groups, depending on the level of complex rearrangement: those with CAAI < 0.5 for all arms (low-level CAAI; A1, B1, AB1, and

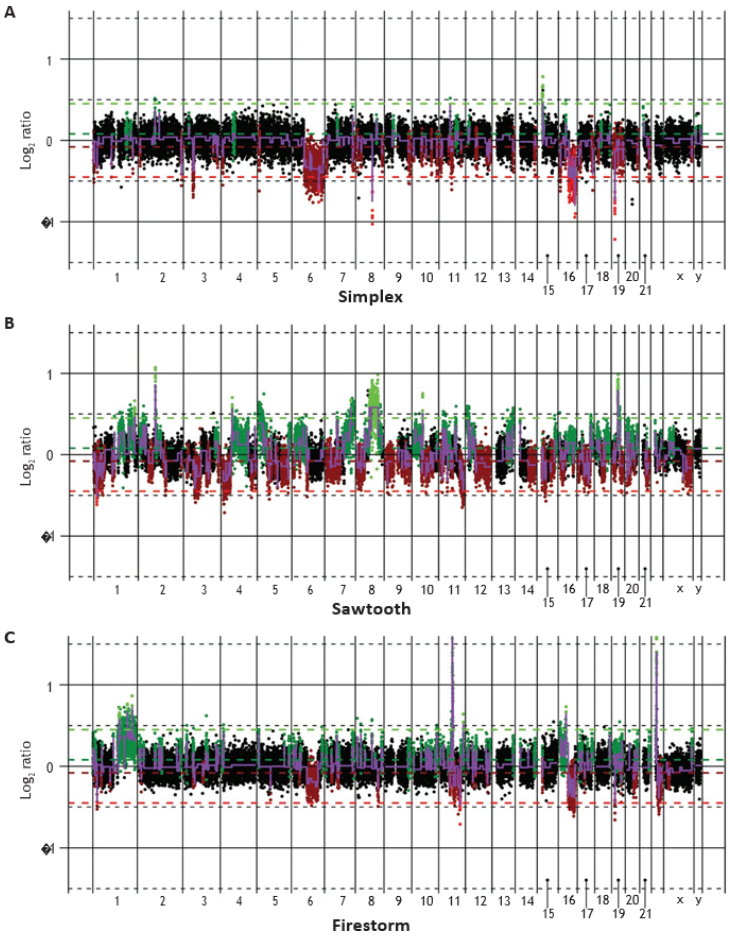


Figure 1. Segmented genome profiles.

Examples of patterns representing simplex (A), sawtooth (B), and Firestorm (C). the y axis displays the geometric mean value of two experiments on a log scale. chromosomes 1 to 22, plus X and Y, are displayed in order from left to right, according to the probe position.

C1) and those with CAAI ≥ 0.5 for at least one arm (high-level CAAI; A2, B2, AB2, and C2). The results show that all tumors with complex rearrangements had more whole arms affected than those without complex rearrangements.

These subgroups tend to have other characteristics in common. Breast tumors can be classified into several subtypes based on their hormone receptor [estrogen receptor (ER) and progesterone receptor (PR)] and HER2 status; based on gene expression profiling, luminal A, luminal B, HER2-enriched, basal-like, and normal-like tumors can be distinguished. The type A tumor class was dominated by ER-positive, luminal A tumors. These tumor genomes had high-magnitude WAAI scores, as well as 1q gain and 16q loss. As compared to A1 tumors, genomes from A2 tumors had chromosomes with more arms with high-magnitude WAAI scores and were more frequently aneuploid. A2 tumors tended to be of high grade and were associated with worse outcomes than A1 tumors. The B1 class was dominated by the basal-like subtype. Type B tumors had different and more heterogeneous genomic patterns; B1 tumor genomes were dominated by losses. A majority of HER2-enriched and normal-like tumors were classified as C tumors, and almost 30% of all basal-like tumors were classified as C tumors.

Next, the CAAI score was shown to have independent prognostic power in 451 breast carcinomas with clinical follow-up data. Patients with a B type tumor had a twofold increased risk of dying of breast cancer as compared with those with the A type, independent of lymph node status, tumor size, histological grade, and treatment. Individuals with tumors with a high-magnitude CAAI score had a twofold increased risk of dying of breast cancer as compared to those with a low-magnitude CAAI score, independent of lymph node status, tumor size, histological grade, and WAAI class.

Conclusion

The analyses presented by Russnes *et al*. [8] illustrate the complexity of the types of analysis that are needed to integrate large-scale genomic analysis with clinical and pathological parameters. In recent years, it has become possible to give tailored therapy

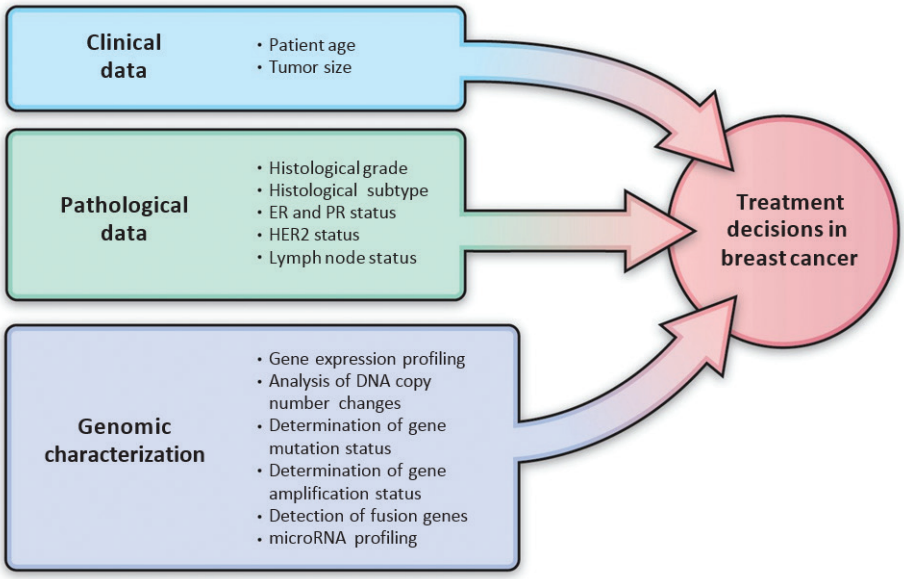


Figure 2. Better decisions.

Integration of clinical, pathological, and possible genetic factors to improve treatment decisions in breast cancer.

to many breast cancer patients as a result of the multitude of choices for surgical, radiotherapy, and adjuvant or neoadjuvant systemic treatment. (Adjuvant therapy is given after all detectable disease has been removed to reduce the risk of relapse; neoadjuvant therapy is given before the primary treatment.) The prognostic and predictive factors on which these choices are presently based include clinical and pathological factors (Fig. 2). Prognostic factors (defined as factors that predict the course of the disease) and predictive factors (defined as factors that predict the response to specific therapies) are used to guide treatment in individual patients [23]. With increasing knowledge of specific genetic alterations and gene expression profiles of tumors, and the prognostic and predictive value of these genetic tumor characteristics, more refined patient therapy is starting to be possible. Axillary lymph node status, tumor size, histological grade, histological subtype, HER2 status, and hormone receptor status are still the most important factors for determining treatment [23]. The St. Gallen [24], National Institutes of Health [25], and Nottingham Prognostic Index guidelines [26], as well as the Adjuvant Online! decision-making tool [27], use a combination of these prognostic

factors to guide decision-making about adjuvant systemic treatment of patients with early breast cancer. However, using these guidelines, a substantial proportion of breast cancer patients who would also survive without adjuvant systemic therapy undergo systemic therapy and suffer from its side effects without gaining any benefit [28, 29]. In addition, more (genetic) tests are urgently needed to predict the responsiveness of tumors to chemotherapy and targeted therapies.

To a progressively increasing extent, genetic factors are being added to clinical and pathological characteristics to derive individualized predictions of disease outcome and response to therapy. The information derived from gene expression profiling, aCGH, and more recently, massive parallel sequencing has been used for these reasons. Translation into the clinical area has so far been limited by (i) the heterogeneity of the studies, (ii) the complexity of breast cancer biology, (iii) the complex analyses, (iv) small sample sizes and lack of independent validation, and (v) the relatively rare occurrence of most genetic alterations in breast cancer (most alterations occur in only 10 to 20% of breast carcinomas); if subgroups of tumors are defined on the basis of combinations of genetic alterations, these subgroups become very small.

The study by Russnes *et al*. [28, 29] demonstrates a correlation between structural genomic alterations, molecular subtypes, and clinical outcomes and reveals that an objective score of genomic complexity can give independent prognostic information in breast cancer. In this way, an additional tool has been added to the arsenal of analysis algorithms for genetic data that is currently being used to decipher the association between the genetic makeup of breast carcinomas and their clinical behavior.

References

1. Hanahan D, Weinberg RA: The hallmarks of cancer. *Cell* 2000, 100(1):57-70.
2. Sherr CJ: Principles of tumor suppression. *Cell* 2004, 116(2):235-246.
3. Reis-Filho JS, Simpson PT, Gale T, Lakhani SR: The molecular genetics of breast cancer: the contribution of comparative genomic hybridization. *Pathol Res Pract* 2005, 201(11):713-725.
4. Sotiriou C, Piccart MJ: Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? *Nat Rev Cancer* 2007, 7(7):545-553.
5. Straehle C, Cardoso F, Azambuja E, Dolci S, Meirsman L, Vantongelen K, Ignatiadis M, Sotiriou C, Piccart-Gebhart MJ: Better translation from bench to bedside: breakthroughs in the individualized treatment of cancer. *Crit Care Med* 2009, 37(1 Suppl):S22-29.
6. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M *et al*: Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med* 2001, 344(11):783-792.
7. Berns K, Horlings HM, Hennessy BT, Madiredjo M, Hijmans EM, Beelen K, Linn SC, Gonzalez-Angulo AM, Stemke-Hale K, Hauptmann M *et al*: A functional genetic approach identifies the PI3K pathway as a major determinant of trastuzumab resistance in breast cancer. *Cancer Cell* 2007, 12(4):395-402.
8. Russnes HG, Vollan HKM, Lingjaerde OC, Krasnitz A, Lundin P, Naume B, Sorlie T, Borgen E, Rye IH, Langerod A *et al*: Genomic architecture characterizes tumor progression paths and fate in breast cancer patients. *Sci Transl Med* 2010, 2(38):38ra47.
9. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y *et al*: High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 1998, 20(2):207-211.
10. Weiss MM, Hermsen MA, Meijer GA, van Grieken NC, Baak JP, Kuipers EJ, van Diest PJ: Comparative genomic hybridisation. *Mol Pathol* 1999, 52(5):243-251.
11. van Beers EH, Nederlof PM: Array-CGH and breast cancer. *Breast Cancer Res* 2006, 8(3):210.
12. Lai C, Horlings HM, van de Vijver MJ, van Beers EH, Nederlof PM, Wessels LF, Reinders MJ: SIRAC: Supervised Identification of Regions of Aberration in aCGH datasets. *BMC Bioinformatics* 2007, 8:422.
13. Lai WR, Johnson MD, Kucherlapati R, Park PJ: Comparative analysis of algorithms for identifying amplifications and deletions in array CGH data. *Bioinformatics* 2005, 21(19):3763-3770.

14. Naylor TL, Greshock J, Wang Y, Colligon T, Yu QC, Clemmer V, Zaks TZ, Weber BL: High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization. *Breast Cancer Res* 2005, 7(6):R1186-1198.

15. Dressman MA, Baras A, Malinowski R, Alvis LB, Kwon I, Walz TM, Polymeropoulos MH: Gene expression profiling detects gene amplification and differentiates tumor types in breast cancer. *Cancer Res* 2003, 63(9):2194-2199.

16. Adler AS, Lin M, Horlings H, Nuyten DS, van de Vijver MJ, Chang HY: Genetic regulators of large-scale transcriptional signatures in cancer. *Nat Genet* 2006, 38(4):421-430.

17. Tusher VG, Tibshirani R, Chu G: Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America* 2001, 98(9):5116-5121.

18. Horlings HM, Lai C, Nuyten DS, Halfwerk H, Kristel P, van Beers E, Joosse SA, Klijn C, Nederlof PM, Reinders MJ *et al*: Integration of DNA copy number alterations and prognostic gene expression signatures in breast cancer patients. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2010, 16(2):651-663.

19. Faratian D, Clyde RG, Crawford JW, Harrison DJ: Systems pathology--taking molecular pathology into a new dimension. *Nat Rev Clin Oncol* 2009, 6(8):455-464.

20. Dunkler D, Michiels S, Schemper M: Gene expression profiling: does it add predictive accuracy to clinical characteristics in cancer prognosis? *Eur J Cancer* 2007, 43(4):745-751.

21. Subramanian J, Simon R: What should physicians look for in evaluating prognostic gene-expression signatures? *Nat Rev Clin Oncol* 2010, 7(6):327-334.

22. Hicks J, Krasnitz A, Lakshmi B, Navin NE, Riggs M, Leibu E, Esposito D, Alexander J, Troge J, Grubor V *et al*: Novel patterns of genome rearrangement and their association with survival in breast cancer. *Genome Res* 2006, 16(12):1465-1479.

23. Chang JC: Diseases of the Breast. In., edn. Edited by Harris J. LME. Philadelphia: Lippincott Williams & Wilkins; 2004: 675-696.

24. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thurlimann B, Senn HJ, Panel m: Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. *Ann Oncol* 2009, 20(8):1319-1329.

25. Eifel P. AJA, Costa J., Crowley J, Curran W.J., Deshler A., Fulton S., Hendricks C.B., et al.: Adjuvant therapy for breast cancer. In: *National Institutes of Helath consensus development conference statement*. USA; 2000.

26. Blamey RW, Ellis IO, Pinder SE, Lee AH, Macmillan RD, Morgan DA, Robertson JF, Mitchell MJ, Ball GR, Haybittle JL *et al*: Survival of invasive breast cancer according to the Nottingham Prognostic Index in cases diagnosed in 1990-1999. *Eur J Cancer* 2007, 43(10):1548-1555.

27. Ravdin PM, Siminoff LA, Davis GJ, Mercer MB, Hewlett J, Gerson N, Parker HL: Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 2001, 19(4):980-991.

28. Early Breast Cancer Trialists' Collaborative G: Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005, 365(9472):1687-1717.

29. Smith I, Procter M, Gelber RD, Guillaume S, Feyereislova A, Dowsett M, Goldhirsch A, Untch M, Mariani G, Baselga J *et al*: 2-year follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. *Lancet* 2007, 369(9555):29-36.

Chapter 3

Retrospective analysis of metastatic behavior of breast cancer subtypes

C. Dilara Savci Heijink, Hans Halfwerk, Gerrit K. Hooijer, Hugo M. Horlings, Jelle Wesseling, Marc J. van de Vijver

Breast Cancer Res Treat. 2015 Apr;150(3):547-57

Abstract

Among breast cancer patients who develop distant metastases there is marked variability in the clinical course, including metastasis pattern. Here, we present a retrospective study of breast cancer patients who all developed distant metastases focusing on the association between breast cancer subtype and clinical course, including organ-specific metastasis. Tissue microarrays (TMAs) were assembled and stained for ER, PR, HER2, EGFR, CK5/6, CK14, E-Cadherin, TP53 and Ki67 for 263 breast cancer patients with metastatic disease. Tumors were classified into ER+/HER2-/Ki67high, ER+/HER2-/Ki67low, ER+/HER2+, ER-/HER2+ and ER-/HER2- groups. Relevant data related to metastasis pattern, metastasis timeline, systemic treatment and survival were retrieved. Associations between site-specific relapse and patient/tumor characteristics were assessed with multivariate models using logistic regression. Median time for development of distant metastasis was 30 months (range 0-15.3 years); 75.8% of the distance metastases developed in the first 5 years after treatment of the primary tumor. Patients with ER-/HER2- tumors had a median overall survival of 27 months; those with HER2+ tumors of 52 months; those with ER+/HER2-/Ki67high of 76 months and those with ER+/HER2-/Ki67low of 79 months. Bone was the most common site for distant metastasis (70.6%) followed by liver (54.5%) and lung (31.4%) respectively. Visceral metastasis was found in 76.8% of the patients. Patients with ER-/HER2- tumors developed visceral metastases in 81% and bone metastases in 55.2%; those with HER2+ tumors developed visceral metastases in 77.4% and bone metastases in 69. 8%; those with ER+/HER2-/Ki67high developed visceral metastases in 75.7% and bone metastases in 87.8% and those with ER+/HER2-/Ki67low developed visceral metastases in 76.9% and bone metastases in 73.1%. In metastatic breast cancer patients, tumor subtypes are associated with survival and pattern of distant metastases. These associations are of help in choices for surveillance and therapy in individual patients.

Introduction

Although the cure rate of breast cancer is increasing in the western world, breast cancer remains the leading cause of female cancer deaths [1]. Most breast cancer deaths are related to distant organ metastasis, which is considered to be essentially incurable. The development of metastatic breast cancer is a complex multi-step process manifesting with distinct patterns of distal organ involvement [2-6]. Using gene expression profiling studies, several molecular mechanisms associated with organ-specific metastasis patterns have been reported [4, 7-15]. Even though these gene expression signatures have already provided useful information in the characterization of novel molecular mediators of organ-specific metastasis, translation of these recently published data to clinical practice has not been accomplished. Moreover, the number of studies focusing on association of more conventional clinicopathologic findings to metastasis pattern is limited [3, 7, 16, 17].

The metastasis pattern of breast cancer varies by hormone receptor status. It has been shown that triple-negative tumors show increased incidence of visceral and cerebral distant metastasis, while hormone receptor-positive tumors have been shown to have a greater tendency to develop bone metastasis. HER2-positive tumors have been reported to metastasize to the brain more frequently than HER2-negative tumors [12, 16, 18-27].

Population-based studies suggest that the survival for metastatic breast cancer patients has been prolonged in recent years as a result of more effective systemic treatment [28-30]. However patients with triple negative breast cancer continue to have dismal outcome after the development of distant metastases [19, 22, 31-33] with a shorter median survival compared to hormone receptor and/or HER2 positive breast cancer [28].

To improve our understanding of the time course and pattern of distant metastases, a retrospective study was carried out using tissue microarrays of primary invasive breast carcinomas of patients who developed distant metastatic disease. Our objectives were

to compare the clinicopathologic findings with metastatic behavior of the breast tumors in terms of organ-specific metastasis and associated patient outcomes.

Material and Methods

Patients and tumor samples

Patients with metastatic breast cancer diagnosed between 1983 and 2009 were identified from the archives of the *Academic Medical Center* and the *Netherlands Cancer Institute* (total n=263) and relevant clinical information was abstracted from their clinical charts. This study material was strictly handled after coding of the data according to national ethical guidelines of 'Code for Proper Secondary Use of Human Tissue' developed by Federation of Medical Societies (FMWV) in the Netherlands [34]. Therefore the need for obtaining informed consent was waived by the Medical Ethical Committee of the Academic Medical Center.

Metastatic disease was defined as recurrence of breast cancer occurring beyond the confines of the ipsilateral breast, chest wall and regional lymph nodes. Metastatic site was classified as bone, lung, liver, pleura/peritoneum, brain, distant lymph nodes and other (including skin, spleen, ovary, eye and other organs). These individual metastasis sites were further used to separate patients in subgroups; for each metastatic site it was assessed whether patients developed metastases during follow-up (ever versus never for each organ site); when patients developed metastases to any organ site, it was recorded whether this was the first metastasis or a metastasis arising after metastases to other organ site arose (first/not first); and it was recorded when a patient developed metastases to one organ site only (only/not only). The presence of multiple metastases was also carefully recorded at the time of diagnosis of the first metastases as well as after the complete follow up. In instances where patients developed another distal organ involvement within less than two months after initial diagnosis of a metastasis, this was also considered as multiple organ metastases at first presentation.

Time from surgery to development of first metastasis, time from first metastasis to last event (metastasis specific survival, MSS) and overall survival (OS) time for each patient

was calculated. Last event date was recorded as most recent follow-up date for the patients who were alive and time of death for the others. Nineteen of the patients were lost to follow up.

Furthermore, data on systemic treatment (chemotherapy, hormonal therapy, HER2-targeted therapy) used to treat primary and metastatic disease was collected for a subset of the patients (n= 149 and n=124, respectively).

Morphological features and immunophenotypic analysis

From all tumors, hematoxylin-eosin stained-slides from paraffin embedded tissues were evaluated and tumor type, histologic grade according to Elston and Ellis [35] and the presence of lymfangioinvasion were assessed. Tissue microarrays (TMAs) were constructed by a manual tissue arrayer (Beecher Instruments, Silver Spring, MD, USA) from the selected representative blocks (n=263). Immunohistochemical staining for Estrogen Receptor (ER) [clone SP1, Ventana], Progesterone Receptor (PR) [clone 1E2, Ventana], Human Epidermal growth Factor receptor 2 (HER2) [clone SP3, Thermo Scientific], Epidermal Growth Factor Receptor (EGFR) [clone H11, Dako], Cytokeratin-5/6 (CK5/6) [clone D5/16 B4, Dako], Cytokeratin-14 (CK14) [clone LL002, Leica] E-Cadherin [clone HECD-1, Invitrogen], TP53 [clone DO-7 +BP53-12, Thermo Scientific], and Ki67 [clone SP6, Thermo Scientific] was performed using an automated slide preparation system (Benchmark XT, Ventana Medical Systems, Tucson Arizona, USA). On the same platform a Silver In Situ Hybridisation (SISH) was performed with INFORM HER2 DNA probe obtained from Ventana Medical Systems. The signal detection for IHC was performed with a biotine-free ultraview universal DAB detection Kit (Ventana Medical Systems) and for SISH with an ultraview SISH detection kit (Ventana Medical Systems).

The immunohistochemistry results were scored independently by two pathologists (C.D.S-H and MJvdV). ER and PR positivity were defined as nuclear staining in 10% or more of tumor cells. Scoring for HER2 immunohistochemistry and in situ hybridization was performed according to ASCO guidelines [36]. Briefly, HER2 staining was scored as 0, 1+, 2+ or 3+; a score of 3+ was considered to be HER2 positive and 0 or 1+ HER2 negative and 2+ scores were evaluated by silver enhanced in situ hybridization (SISH) to determine final HER2 status. For mono color SISH, the number of nuclear spots was

counted in 30 adjacent tumor cells and tumors with an average number of HER2 signals ≥ 6 were considered as HER2 amplified; all other tumors were considered as HER2 non-amplified. Tumors were further grouped by ER/HER2 expression pattern as: ER-positive/HER2-positive, ER-positive/HER2-negative, ER-negative/HER2-positive and ER-negative/HER2-negative tumors. ER-positive/HER2-negative tumors were further divided into 2 subgroups according to their Ki67 immunopositivity. For Ki-67 staining, the percentage of positively staining tumor cells was counted and a cut-off for low versus high of 13% was used according to the St Gallen consensus guidelines [37, 38]. Hormone receptor-negative group was also divided into 2 subgroups according to their so-called basal cell marker status. The hormone receptor-negative tumors which were positive for CK5/6 and/or CK14 and/or EGFR and/or C-kit was considered to be basal-like tumors whereas the others considered to be non-basal-like group of tumors [39, 40].

Samples were considered to be positive for TP53 if more than 50% of tumor cells showed positive staining in the nuclei. E-cadherin was scored as positive when there was any membranous staining. CK5/6, CK14, C-kit, E-cadherin, and EGFR were scored as positive if ≥10% of the tumor cells showed staining.

Statistical analysis

Association between immunophenotypic findings and metastatic behavior (including metastasis site and metastasis pattern) was assessed using either the Fisher Exact test (variable with two classes) or Chi-square test. To further explore this association, multivariate logistic regression analyses were applied to model the relationship between site-specific relapse and patient/tumor characteristics. All statistical tests were two sided and $p < 0.05$ was considered to be statistically significant. Survival analyses were estimated by the Kaplan-Meier method and were compared using the log-rank test. Analyses were performed using SPSS Statistics for Windows (Release version 21.0; IBM Corp. 2012, Armond, NY).

Results

Clinicopathologic features

For 263 patients treated for breast cancer who all developed distant metastases during follow-up, we have collected paraffin-embedded tumor tissue of the primary tumor; assessed the histopathological features; and performed immunohistochemical staining on tissue micro arrays (TMA's). Clinicopathologic data are shown in Table 1. The mean age at diagnosis was 50 years old (range 27-86). Median follow-up was 57 months for all patients (range 0.5-22.4 years) and 11.6 years (range 6.2 to 17.3 years) for patients (*n*=14) who were alive at last follow-up. The majority of the tumors (88.2%) was classified as invasive ductal carcinoma and 90.9% of the tumors were grade 2 or 3 [35]. Tumor size varied from 0,5 cm to 9 cm with a mean size of 3,2 cm.

Out of 149 patients with available adjuvant therapy data, 85 (57%) patients received chemotherapy, whereas 61 (40.9%) patients received hormonal therapy. More specifically 46.5% of patients with ER+ tumors was noted to receive hormonal therapy. Among 122 patients with available chemotherapy data for the metastatic disease, 50 patients received chemotherapy as first line treatment after the development of metastatic disease, whereas 66 patients received hormonal therapy (40.7% and 50.7%, respectively). Selective oestrogen receptor modulators were the most common (45.8%) administered first-line hormonal therapy regimen, followed by aromatase inhibitors (41.7%) and LH blockers (46.8%). Only 14 patients received Herceptin therapy for treating metastatic disease.

Results of immunohistochemical staining can be seen in Table 2. When grouped into subtypes according to ER/HER2 expression, 27.6% were ER-/HER2-, 24.8% were HER2 positive and 47.6% were ER+/HER2-. Of ER+/HER2- tumors, 93.7% were Ki67 high and 6.3% were Ki67 low, 31.8% were TP53 positive; 4.1% were EGFR positive; 9.4% were CK14 positive and 15.4% were CK5/6 positive. 61.1% of the ER-/HER2- tumors were positive for one of the so-called basal cell markers (CK5/6, CK14, EGFR or C-kit). Of note, within the hormone receptor negative group no significant difference was found between the tumors with and without basal-like markers regarding clinicopathological

Table 1. Clinical and pathological characteristics of metastatic breast cancer patients

		N	%
Age at diagnosis, years	<50	146	55.5
	>50	117	44.5
Lmphy node status	negative	54	34.4
	1-3 positive	52	33.1
	>3 positive	51	32.5
Histology	ductal	231	88.2
	lobular	20	7.6
	other	11	4.2
Tumor grade	1	23	9.1
	2	134	53.0
	3	96	37.9
Tumor size	0-2 cm	63	28.5
	2-5 cm	131	59.3
	>5 cm	27	12.2
Tumor subtype	ER(-) HER2(-)	59	27.6
	ER (+) HER2 (-) Ki67high	75	35.0
	ER (+) HER2 (-) Ki67low	27	12.6
	ER (+) HER2 (+)	28	13.1
	ER (-) HER2 (+)	25	11.7
Time to distant metastasis[a]	early	194	75.8
	late	62	24.2
Multiple metastasis sites at first presentation	no	160	62.7
	yes	95	37.3
Multiple metastasis sites during follow up	no	50	19.6
	yes	205	80.4

CT chemotherapy, HT hormonal therapy, ER, estrogen receptor, PR progesterone receptor; HER2 Human epidermal growth factor receptor.
[a] *Cut-off point 5 years.*

Table 2. Results of immunohistochemical staining in the primary tumors

		N	%*
ER	negative	84	34.1
	positive	162	65.9
PR	negative	93	37.1
	positive	158	62.9
HER2	negative	199	78.3
	positive	55	21.7
E-cadherin	negative	24	10.4
	positive	206	89.6
CK5/6	negative	193	84.6
	positive	35	15.4
CK14	negative	213	90.6
	positive	22	9.4
EGFR	negative	232	95.9
	positive	10	4.1
TP53	negative	163	69.1
	positive	73	30.9
c-kit	negative	233	96.7
	positive	8	3.3
Ki67	low	37	19.3
	high	155	80.7

ER estrogen receptor, PR progesterone receptor, HER2 Human epidermal growth factor receptor, CK cytokeratine, EGFR epidermal growth factor receptor.
** Valid percentages*

characteristics, metastatic behavior and survival outcomes. Therefore, we have chosen to proceed with hormone receptor-negative group as one group.

Findings of immunohistochemical staining for TP53, CK5/6, CK14, EGFR, c-kit and Ki67 in hormone receptor-negative tumors are displayed in Table 3. 10.4% of tumors were E-cadherin negative; 54. 2% of these were classified as invasive lobular carcinomas.

Time to distant metastasis

Median time to develop metastasis was 30 months (range 0-15.3 years) and median time from metastasis to death was 19 months and to last follow-up for patients alive was 64 months. Using the cut-off point of 5 years, 75.8% of the tumors were recorded as early metastasizing tumors. In table 4, the association between histologic and immunohistochemical variables and early versus late metastasis is shown.

Table 3. Results of immunohistochemical staining in ER-/HER2- tumors

		N	%*
TP53	negative	24	42.1
	positive	33	5.,9
CK5/6	negative	28	52,8
	positive	25	47.2
CK14	negative	39	70.9
	positive	16	29.1
EGFR	negative	47	82.5
	positive	10	17.5
c-kit	negative	52	91.2
	positive	5	8.8
Ki67	low	2	4.5
	high	42	95.5

CK Cytokeratine, EGFR Epidermal growth factor receptor.
**Valid percentages*

Table 4. Correlation between tumor characteristics and time to distant metastasis

Characteristics		Metastasis timeline				
		< 5 years		>5 years		
		N	%	N	%	p
Tumor size	<2cm	37	21.8	27	42.9	.001
	>=2	133	78.2	36	57.1	
Tumor grade	1	17	8.5	7	10.6	
	2	99	49.3	42	63.6	.056
	3	85	42.3.	17	25.8	
ER	negative	73	37.4.	10	15.2	
	positive	122	62.6.	56	84.8	.001
PR	negative	83	41.9	12	17.6	
	positive	115	58.1	56	82.4	<.001
Tumor subtype	ER (-) HER2 (-)	52	30.8	7	13.5	
	ER (+) HER2 (-) Ki67high	50	29.6	25	48.1	
	ER (+) HER2 (-) Ki67low	17	10.1	9	17.3	.003
	ER (+) HER2 (+)	29	17.2	8	15.4	
	ER (-) HER2 (+)	21	12.4	3	5.8	

ER estrogen receptor, PR progesterone receptor, HER2 Human epidermal growth factor receptor

Table 5. Tumor characteristics and interval to metastasis and last event

		Interval between surgery and metastasis, months		Interval between metastasis and last event, months		Overall survival	
		Mean	Median	Mean	Median	Mean	Median
Tumor subtype	ER (-) HER2 (-)	25	15	17	10	41	27
	ER (+) HER2 (-) Ki67high	51	37	36	25	86	76
	ER (+) HER2 (-) Ki67low	54	45	39	25	93	79
	ER (+) HER2 (+)	40	33	41	25	79	59
	ER (-) HER2 (+)	36	22	33	19	69	60
p value		<.001		.002		.020	

ER estrogen receptor, PR progesterone receptor, HER2 Human epidermal growth factor receptor

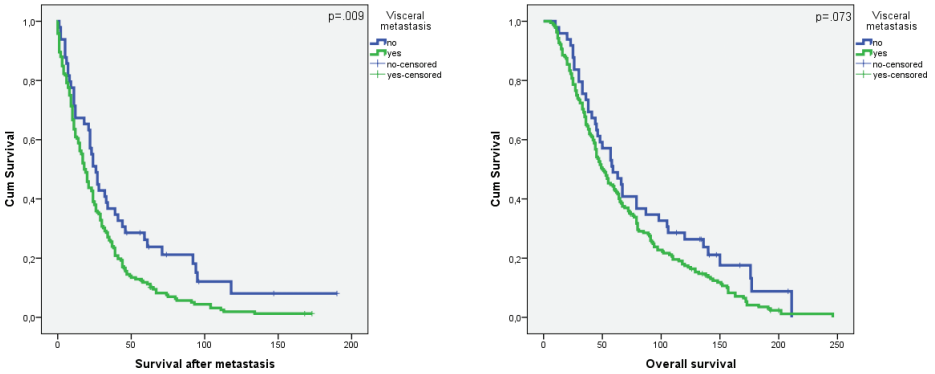


Figure 1. Metastasis specific (A) and overall (B) survival curves of breast cancer patients with and without visceral metastasis.

Kaplan-Meier plots of patients show that tumors with visceral metastasis had worse survival outcomes than the tumors without visceral metastasis.Patients who had visceral metastasis had shorter survival time from detection of metastasis to last event and from initial diagnosis of the disease to last event (p 0.009 and 0.073, respectively)

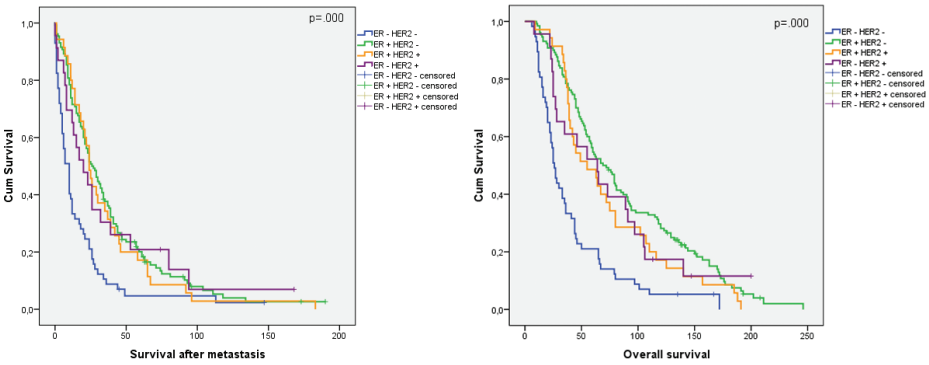


Figure 2. Metastasis specific (A) and overall (B) survival curves of breast cancer patients according to tumor subtypes. *ER* Estrogen receptor; *PR* progesterone receptor; *HER2* Human epidermal growth factor receptor type 2.

Kaplan-Meier plots of patients show that ER-/HER2- had worse survival outcomes compared to other tumor subtypes. Patients with hormone receptor-negative (ER-/HER2-) tumors had shorter survival time from detection of metastatic disease to last event and from the initial diagnosis of the disease to last event (p < 0.001).

As can be seen, ER-/HER2- and ER-/HER2+ tumors metastasized earlier than other subgroups of tumors (p 0.003). Almost 90% of hormone receptor-negative breast cancer patients developed distant metastases early; versus 66% of ER+/HER2-; within this group there was no significant difference between Ki67 high (66.7%) and Ki67 low (65.4%) tumors (p 0.54).

Survival after development of distant metastasis

Figure 1 shows that overall survival and survival after the detection of distant metastasis for patients who developed visceral metastases (*n*=198) is worse than for those who did not develop visceral metastases (p 0.073 and *p* 0.009, respectively). Figure 2 shows overall survival and survival after the detection of distant metastasis for the subgroups of patients defined by ER and HER2 status of the primary tumor.

Table 5 demonstrates the differences of time to develop metastasis and survival time after development of metastatic disease in various subgroups. As can be seen, patients with ER-/HER2- tumors had a median survival of 10 months after the detection of distant metastasis whereas ER-/HER2+ tumors had 19 months median survival (p 0.020). ER+/HER2- (Ki67 high as well as Ki67 low groups) and ER+/HER2+ tumors had a median survival time of 25 and 24 months, respectively (p 0.75).

Site of distant metastasis

Detailed information about metastatic behavior was available for 256 patients; 11(4.3%) patients presented with multiple metastasis simultaneously, while 205 (80.4%) developed multiple metastases during the course of follow-up. Bone was the most common site for metastasis (70.6%) followed by liver (54.5%) and lung (31.4%) respectively. Visceral (liver, lung and brain) metastases were found in 77.6% of the patients.

Twenty-five (9.8%) of the patients developed only bone metastasis and 29 (11.4%) of the patients developed only visceral metastasis during the course of the disease. Among these patients median time to develop bone metastasis and visceral metastasis differed and was 40 months and 23 months, respectively.

Multivariate analyses further revealed that patients who developed visceral metastasis had a higher prevalence of multiple metastases during follow-up ($p < 0.001$).

The metastasis pattern was similar for patients who received adjuvant systemic therapy compared to patients who did not undergo adjuvant systemic treatment.

Along with 81.3% of ER+ tumors, 88% of ER+/HER2- Ki67 high tumors noted to have bone metastasis. Contrarily, hormone receptor-negative (ER-/HER2-) tumors were associated with visceral organ metastasis, yet composing 55% of the tumors with only visceral metastasis.

ER status of the tumor was significantly positively correlated to bone metastasis in the univariate as well as in the multivariate analyses. Several immunohistochemical markers, such as E-Cadherin and Cytokeratine 14, were found to be correlated to visceral metastasis (p 0.013 and p 0.018). E-cadherin was found to be positive in primary tumors of patients who developed visceral organ metastasis and the ones with visceral metastasis as initial site of metastasis (p 0.028 and p 0.040). TP53 positive tumors developed brain metastasis with a rate of 38%, as opposed to 21.2% in TP53 negative ones (p 0.007).

Discussion

Breast cancer is a heterogeneous disease and this is also reflected in the clinical patterns of the development of distant metastases. There is marked variability in the time interval between treatment of the primary tumor and the occurrence of distant metastases; in the organs involved with distant metastases; and in the response to systemic treatment in patients with metastatic breast cancer. The concept of organotropism compasses the non-random distinct organ involvement of different cancer types as well as within a given type of cancer, which usually implies a more subtle intrinsic heterogeneity among organotropic cancer cells [4]. Along with the conventional metastatic model of "*anatomical/mechanical*" hypothesis, Paget's "*seed and soil*" hypothesis [41] are widely accepted models for site-specific metastasis. *Stephen Paget*' s century old theory that proposed the organ-preference patterns of tumor metastasis are the product of favourable interactions between metastatic tumor cells (the "*seed*") and their organ

microenvironment (the "*soil*") was confirmed by clinical and experimental research [4, 41-43]. Better understanding of this complex interaction between two compartments and consequently the mechanisms that lie beneath the site-specific metastasis may improve the clinical management, including developing novel therapeutic options, for metastatic disease.

Despite the increasing tendency to classify breast tumors into molecular subtypes based on gene expression profiles first described by *Perou* et al. [33], immunophenotypic characteristics of the tumor also remain an important cornerstone of defining subgroups of the disease. In the current study, we investigated the presence of site-specific metastasis and concomitant characteristics of the metastatic disease in a retrospective series of 263 breast cancer patients, focusing on the immunophenotypic features of the primary tumor.

Together with clinical observations, recent comprehensive molecular studies unveiled the considerable differences between ER-positive and ER-negative tumors. It has also been shown that ER status has a time-varying prognostic effect mainly pronounced in the early follow-up period [44-48]. ER-positive tumors are known for their tendency to relapse later with higher rate of bone recurrences than their ER-negative counterparts. In agreement with published literature, our data clearly indicates the close relation between ER-positive tumors and metastasis specific survival and bone metastasis [18, 45, 46, 49]. In addition to confirming the well-established prognostic markers in breast cancer, this study was also able to verify that ER-status is also an important factor for bone-only and bone-first metastasis. Likewise, in agreement with previously published data, ER-negative tumors showed a higher proportion of patients with visceral metastases [18, 20, 44, 45, 50]. HER2-positive tumors have been found as a risk factor for cerebral metastasis development (5,12,13,19,22,24,27,29,32,33,52,55) . However, in this study HER2 positivity was not identified as a strong predictive factor for site-specific metastasis and early metastatic disease; of note, in our study we did not find an association between HER2 positive status and brain metastasis. The fact that almost none of the patients included in this study received HER2-targeted therapy may play a role in the absence of a correlation between HER2 status and brain metastases (although we could not retrieve data on adjuvant systemic therapy for all patients, we

know that HER2-targeted therapy was not yet available as adjuvant therapy during large part of the period in which patients were treated).

A recent cohort study demonstrated that hormone receptor HR+/HER2+subtype was associated with the best prognosis after diagnosis of metastatic disease, with a median survival of 34.4 months even better than HR+/HER2- subgroup [51]. This subgroup was followed by HR+/HER2-, HR-/HER2+ and hormone receptor negative tumors. In our study, regardless of their HER2 status, hormone receptor-positive tumors had better survival after the diagnosis of distant metastasis. Similarly to *Lobbezoo* et al., ER-/HER2+ tumors had better survival than hormone receptor-negative tumors. Improved survival rates of HER2+ tumors have already been reported [28, 29, 52]. Even though our study includes patients before the implementation of HER2-targeted therapy for metastatic disease, our results are comparable to this recent cohort study. Additionally, further subgrouping of ER+/HER2- tumors according to their Ki67 status revealed that only minority (6.3%) of these tumors had low Ki67 status. Within this group of tumors there was also no significance regarding metastasis pattern. This result may suggest that if once metastatic event occurs, prognostic relevance of Ki67 might be limited.

Our analyses indicate the noticeable distinction between breast tumors with visceral metastasis and the ones without visceral metastasis. Additional to the remarkably shorter overall survival and metastasis-specific survival compared to the tumors without visceral metastasis, higher frequency of developing multiple metastasis during the course of disease, make this subgroup of tumors challenging and therefore worth to be recognized [49, 53]. Several immunohistochemical markers are known to be associated with hormone receptor-negative breast tumors especially the ones with basal-like features. It is also claimed that triple-negative status cannot be used as a surrogate for the basal cell phenotype [54-56]. In our study, we showed that a group of frequently registered immunohistochemistry markers such as CK 5/6, CK 14 and EGFR were related to the hormone receptor negative subgroup ($p < 0.001$), while c-kit was not found to be related to this subgroup of tumors (p 0.098). Further analyses within this group revealed no significant difference between basal-like group and non-basal-like group in relation to metastatic behavior and survival outcomes. Based on gene expression profiling studies, *Lehmann* et al showed that this aggressive type of breast cancer can be divided into seven subtypes as; *basal-like 1*, *basal-like 2*, *immunomodulatory*,

mesenchymal-*like*, *mesenchymal stem-like*, *luminal androgen receptor* and *unstable*. They also showed that independent analysis of five data sets based on triple-negative tumors identified by immunohistochemical staining had similar clustering [57]. In conjunction with this information, it is indicated that hormone receptor-negative group contains heterogeneous group of tumors with distinct phenotypes. We believe that further studies are indicated to explore the role of immunohistochemistry to portray these heterogeneous subgroups.

The role of E-cadherin in metastatic potential of the tumors has already been a topic of interest. The absence of E-cadherin expression as a result of genetic alterations in the E-cadherin gene is observed in the majority of lobular carcinomas. Reduced expression of E-cadherin has been reported in breast carcinoma cases with a frequency ranging from 4 5% to 63% of cases [58-60]. Several studies showed a higher metastatic potential for tumors with reduced E-cadherin expression [58, 61-65], whereas others were not able to prove such a relation [66, 67]. Interestingly, in the current study, we showed that immunostaining of E-cadherin was positively correlated with developing visceral metastases, also with developing visceral metastasis as first site of metastasis.

It has been shown that tumors with TP53 gene mutations are associated with brain metastases. [27, 68, 69]. Recently, *Lo et al.* have demonstrated that mutation of TP53 is the most common genetic change identified in brain metastases from breast cancer. They identified that 87% of CNS metastatic lesions in their study contained TP53 mutations compared to 25-34% mutations in all breast cancers. [70]. Consistent with the previous reports, we showed that TP53 immunopositivity is significantly associated with subsequent brain metastasis. The cohort of patients in our study was treated between 1983 and 2009; the median size of the primary tumors was 3.2 cm. It may well be that metastatic pattern will differ for patients who were treated more recntly and for patients who presented with smaller tumors. It will be therefore be of interest to perform a similar study to the one presented here in the future for a cohort of more recently treated patients.

Conclusions

This study demonstrates that subtypes of breast cancer mainly defined by ER, PR and HER2 and are strongly related to the metastasis pattern, in terms of site-specific relapse, early/late metastasis and survival outcomes. Hormone receptor-positive tumors have tendency to develop bone metastasis and they have better survival outcomes compared to hormone receptor-negative tumors with a tendency of developing visceral metastasis. HER2 status was not associated with pattern of distant metastases; in agreement with previous reports, P53-positive tumors were more likely to metastasize to the brain than P53-negative tumors. In addition, we show that tumors that develop visceral metastasis have worse prognosis than the ones without visceral metastasis and immunostaining for E-cadherin and Cytokeratine 14 can be of help to identify such tumors. These associations are of help in choices for surveillance and therapy in individual patients.

References

1. Malvezzi M, Bertuccio P, Levi F, La VC, Negri E: European cancer mortality predictions for the year 2012. *AnnOncol* 2012, 23(4):1044-1052.
2. Lorusso G, Ruegg C: New insights into the mechanisms of organ-specific breast cancer metastasis. *SeminCancer Biol* 2012, 22(3):226-233.
3. St RP, Madan R, Tawfik OW, Damjanov I, Fan F: Organotropism and prognostic marker discordance in distant metastases of breast carcinoma: fact or fiction? A clinicopathologic analysis. *HumPathol* 2012, 43(3):398-404.
4. Lu X, Kang Y: Organotropism of breast cancer metastasis. *JMammaryGlandBiolNeoplasia* 2007, 12(2-3):153-162.
5. Chiang AC, Massague J: Molecular basis of metastasis. *NEnglJMed* 2008, 359(26):2814-2823.
6. Chambers AF, Groom AC, MacDonald IC: Dissemination and growth of cancer cells in metastatic sites. *NatRevCancer* 2002, 2(8):563-572.
7. Altundag K, Bondy ML, Mirza NQ, Kau SW, Broglio K, Hortobagyi GN, Rivera E: Clinicopathologic characteristics and prognostic factors in 420 metastatic breast cancer patients with central nervous system metastasis. *Cancer* 2007, 110(12):2640-2647.
8. Bos PD, Zhang XH, Nadal C, Shu W, Gomis RR, Nguyen DX, Minn AJ, van de Vijver MJ, Gerald WL, Foekens JA *et al*: Genes that mediate breast cancer metastasis to the brain. *Nature* 2009, 459(7249):1005-1009.
9. Harrell JC, Prat A, Parker JS, Fan C, He X, Carey L, Anders C, Ewend M, Perou CM: Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer ResTreat* 2012, 132(2):523-535.
10. Landemaine T, Jackson A, Bellahcene A, Rucci N, Sin S, Abad BM, Sierra A, Boudinet A, Guinebretiere JM, Ricevuto E *et al*: A six-gene signature predicting breast cancer lung metastasis. *Cancer Res* 2008, 68(15):6092-6099.
11. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T *et al*: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *NEnglJMed* 2004, 351(27):2817-2826.
12. Slimane K, Andre F, Delaloge S, Dunant A, Perez A, Grenier J, Massard C, Spielmann M: Risk factors for brain relapse in patients with metastatic breast cancer. *AnnOncol* 2004, 15(11):1640-1644.
13. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ *et al*: A gene-expression signature as a predictor of survival in breast cancer. *NEnglJMed* 2002, 347(25):1999-2009.

14. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J *et al*: Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005, 365(9460):671-679.

15. Zhang XH, Wang Q, Gerald W, Hudis CA, Norton L, Smid M, Foekens JA, Massague J: Latent bone metastasis in breast cancer tied to Src-dependent survival signals. *Cancer Cell* 2009, 16(1):67-78.

16. Koo JS, Jung W, Jeong J: Metastatic breast cancer shows different immunohistochemical phenotype according to metastatic site. *Tumori* 2010, 96(3):424-432.

17. Lin Y, Yin W, Yan T, Zhou L, Di G, Wu J, Shen Z, Shao Z, Lu J: Site-specific relapse pattern of the triple negative tumors in Chinese breast cancer patients. *BMCCancer* 2009, 9:342.

18. Campbell FC, Blamey RW, Elston CW, Nicholson RI, Griffiths K, Haybittle JL: Oestrogen-receptor status and sites of metastasis in breast cancer. *BrJCancer* 1981, 44(3):456-459.

19. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P, Narod SA: Triple-negative breast cancer: clinical features and patterns of recurrence. *ClinCancer Res* 2007, 13(15 Pt 1):4429-4434.

20. Dent R, Hanna WM, Trudeau M, Rawlinson E, Sun P, Narod SA: Pattern of metastatic spread in triple-negative breast cancer. *Breast Cancer ResTreat* 2009, 115(2):423-428.

21. Fuqua SA: The role of estrogen receptors in breast cancer metastasis. *JMammaryGlandBiolNeoplasia* 2001, 6(4):407-417.

22. Haffty BG, Yang Q, Reiss M, Kearney T, Higgins SA, Weidhaas J, Harris L, Hait W, Toppmeyer D: Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. *JClinOncol* 2006, 24(36):5652-5657.

23. Heitz F, Harter P, Lueck HJ, Fissler-Eckhoff A, Lorenz-Salehi F, Scheil-Bertram S, Traut A, du BA: Triple-negative and HER2-overexpressing breast cancers exhibit an elevated risk and an earlier occurrence of cerebral metastases. *EurJCancer* 2009, 45(16):2792-2798.

24. Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, Speers CH, Nielsen TO, Gelmon K: Metastatic behavior of breast cancer subtypes. *JClinOncol* 2010, 28(20):3271-3277.

25. Largillier R, Ferrero JM, Doyen J, Barriere J, Namer M, Mari V, Courdi A, Hannoun-Levi JM, Ettore F, Birtwisle-Peyrottes I *et al*: Prognostic factors in 1,038 women with metastatic breast cancer. *AnnOncol* 2008, 19(12):2012-2019.

26. Lin NU, Claus E, Sohl J, Razzak AR, Arnaout A, Winer EP: Sites of distant recurrence and clinical outcomes in patients with metastatic triple-negative breast cancer: high incidence of central nervous system metastases. *Cancer* 2008, 113(10):2638-2645.

27. Tham YL, Sexton K, Kramer R, Hilsenbeck S, Elledge R: Primary breast cancer phenotypes associated with propensity for central nervous system metastases. *Cancer* 2006, 107(4):696-704.

28. Chia SK, Speers CH, D'yachkova Y, Kang A, Malfair-Taylor S, Barnett J, Coldman A, Gelmon KA, O'reilly SE, Olivotto IA: The impact of new chemotherapeutic and hormone agents on

survival in a population-based cohort of women with metastatic breast cancer. *Cancer* 2007, 110(5):973-979.

29. Dawood S, Broglio K, Gonzalez-Angulo AM, Buzdar AU, Hortobagyi GN, Giordano SH: Trends in survival over the past two decades among white and black patients with newly diagnosed stage IV breast cancer. *JClinOncol* 2008, 26(30):4891-4898.

30. Giordano SH, Buzdar AU, Smith TL, Kau SW, Yang Y, Hortobagyi GN: Is breast cancer survival improving? *Cancer* 2004, 100(1):44-52.

31. Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V: Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer* 2007, 109(9):1721-1728.

32. Foulkes WD, Brunet JS, Stefansson IM, Straume O, Chappuis PO, Begin LR, Hamel N, Goffin JR, Wong N, Trudel M *et al*: The prognostic implication of the basal-like (cyclin E high/p27 low/p53+/glomeruloid-microvascular-proliferation+) phenotype of BRCA1-related breast cancer. *Cancer Res* 2004, 64(3):830-835.

33. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA *et al*: Molecular portraits of human breast tumours. *Nature* 2000, 406(6797):747-752.

34. Vermeulen E, Geesink I, Schmidt MK, Steegers C, Verhue D, Brom FW, Aaronson NK, van Leeuwen FE: [Secondary use of human tissue: consent and better information required]. *NedTijdschrGeneeskd* 2009, 153:A948.

35. Elston CW, Ellis IO: Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. C. W. Elston & I. O. Ellis. Histopathology 1991; 19; 403-410. *Histopathology* 2002, 41(3A):151-152, discussion.

36. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A *et al*: American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *ArchPatholLab Med* 2007, 131(1):18-43.

37. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS *et al*: Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *JNatlCancer Inst* 2009, 101(10):736-750.

38. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ: Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *AnnOncol* 2011, 22(8):1736-1747.

39. Kim MJ, Ro JY, Ahn SH, Kim HH, Kim SB, Gong G: Clinicopathologic significance of the basal-like subtype of breast cancer: a comparison with hormone receptor and Her2/neu-overexpressing phenotypes. *Human pathology* 2006, 37(9):1217-1226.

40. Kreike B, van Kouwenhove M, Horlings H, Weigelt B, Peterse H, Bartelink H, van de Vijver MJ: Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas. *Breast Cancer Res* 2007, 9(5):R65.

41. Paget S: The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev* 1989, 8(2):98-101.

42. Fidler IJ: Selection of successive tumour lines for metastasis. *NatNew Biol* 1973, 242(118):148-149.

43. Langley RR, Fidler IJ: The seed and soil hypothesis revisited--the role of tumor-stroma interactions in metastasis to different organs. *IntJCancer* 2011, 128(11):2527-2535.

44. Andry G, Suciu S, Pratola D, Sylvester R, Leclercq G, da Costa PM, Legros N, Andry-t'Hooft M, Verhest A, Mattheiem W *et al*: Relation between estrogen receptor concentration and clinical and histological factors: their relative prognostic importance after radical mastectomy for primary breast cancer. *EurJCancer ClinOncol* 1989, 25(2):319-329.

45. Hess KR, Pusztai L, Buzdar AU, Hortobagyi GN: Estrogen receptors and distinct patterns of breast cancer relapse. *Breast Cancer ResTreat* 2003, 78(1):105-118.

46. Koenders PG, Beex LV, Langens R, Kloppenborg PW, Smals AG, Benraad TJ: Steroid hormone receptor activity of primary human breast cancer and pattern of first metastasis. The Breast Cancer Study Group. *Breast Cancer ResTreat* 1991, 18(1):27-32.

47. Raemaekers JM, Beex LV, Koenders AJ, Pieters GF, Smals AG, Benraad TJ, Kloppenborg PW: Disease-free interval and estrogen receptor activity in tumor tissue of patients with primary breast cancer: analysis after long-term follow-up. *Breast Cancer ResTreat* 1985, 6(2):123-130.

48. Winstanley J, Cooke T, Murray GD, Platt-Higgins A, George WD, Holt S, Myskov M, Spedding A, Barraclough BR, Rudland PS: The long term prognostic significance of c-erbB-2 in primary breast cancer. *BrJCancer* 1991, 63(3):447-450.

49. Solomayer EF, Diel IJ, Meyberg GC, Gollan C, Bastert G: Metastatic breast cancer: clinical course, prognosis and therapy related to the first site of metastasis. *Breast Cancer ResTreat* 2000, 59(3):271-278.

50. Hicks DG, Short SM, Prescott NL, Tarr SM, Coleman KA, Yoder BJ, Crowe JP, Choueiri TK, Dawson AE, Budd GT *et al*: Breast cancers with brain metastases are more likely to be estrogen receptor negative, express the basal cytokeratin CK5/6, and overexpress HER2 or EGFR. *AmJSurgPathol* 2006, 30(9):1097-1104.

51. Lobbezoo DJ, van Kampen RJ, Voogd AC, Dercksen MW, van den Berkmortel F, Smilde TJ, van de Wouw AJ, Peters FP, van Riel JM, Peters NA *et al*: Prognosis of metastatic breast cancer subtypes: the hormone receptor/HER2-positive subtype is associated with the most favorable outcome. *Breast cancer research and treatment* 2013, 141(3):507-514.

52. Planchat E, Durando X, Abrial C, Thivat E, Mouret-Reynier MA, Ferriere JP, Pomel C, Kwiatkowski F, Chollet P, Nabholtz JM: Prognostic value of initial tumor parameters after metastatic relapse. *Cancer Invest* 2011, 29(9):635-643.

53. Rosa Mendoza ES, Moreno E, Caguioa PB: Predictors of early distant metastasis in women with breast cancer. *JCancer ResClinOncol* 2013.

54. Badve S, Dabbs DJ, Schnitt SJ, Baehner FL, Decker T, Eusebi V, Fox SB, Ichihara S, Jacquemier J, Lakhani SR *et al*: Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2011, 24(2):157-167.

55. Keam B, Im SA, Lee KH, Han SW, Oh DY, Kim JH, Lee SH, Han W, Kim DW, Kim TY *et al*: Ki-67 can be used for further classification of triple negative breast cancer into two subtypes with different response and prognosis. *Breast Cancer Res* 2011, 13(2):R22.

56. Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, Hernandez-Boussard T, Livasy C, Cowan D, Dressler L *et al*: Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2004, 10(16):5367-5374.

57. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietenpol JA: Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *The Journal of clinical investigation* 2011, 121(7):2750-2767.

58. Oka H, Shiozaki H, Kobayashi K, Inoue M, Tahara H, Kobayashi T, Takatsuka Y, Matsuyoshi N, Hirano S, Takeichi M *et al*: Expression of E-cadherin cell adhesion molecules in human breast cancer tissues and its relationship to metastasis. *Cancer Res* 1993, 53(7):1696-1701.

59. Rimm DL, Sinard JH, Morrow JS: Reduced alpha-catenin and E-cadherin expression in breast cancer. *Lab Invest* 1995, 72(5):506-512.

60. Siitonen SM, Kononen JT, Helin HJ, Rantala IS, Holli KA, Isola JJ: Reduced E-cadherin expression is associated with invasiveness and unfavorable prognosis in breast cancer. *AmJClinPathol* 1996, 105(4):394-402.

61. Chen WC, Obrink B: Cell-cell contacts mediated by E-cadherin (uvomorulin) restrict invasive behavior of L-cells. *JCell Biol* 1991, 114(2):319-327.

62. Frixen UH, Behrens J, Sachs M, Eberle G, Voss B, Warda A, Lochner D, Birchmeier W: E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. *JCell Biol* 1991, 113(1):173-185.

63. Mareel MM, Behrens J, Birchmeier W, De Bruyne GK, Vleminckx K, Hoogewijs A, Fiers WC, Van Roy FM: Down-regulation of E-cadherin expression in Madin Darby canine kidney (MDCK) cells inside tumors of nude mice. *IntJCancer* 1991, 47(6):922-928.

64. Pierceall WE, Woodard AS, Morrow JS, Rimm D, Fearon ER: Frequent alterations in E-cadherin and alpha- and beta-catenin expression in human breast cancer cell lines. *Oncogene* 1995, 11(7):1319-1326.

65. Vleminckx K, Vakaet L, Jr., Mareel M, Fiers W, van RF: Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. *Cell* 1991, 66(1):107-119.

66. Lipponen P, Saarelainen E, Ji H, Aaltomaa S, Syrjanen K: Expression of E-cadherin (E-CD) as related to other prognostic factors and survival in breast cancer. *JPathol* 1994, 174(2):101-109.

67. Parker C, Rampaul RS, Pinder SE, Bell JA, Wencyk PM, Blamey RW, Nicholson RI, Robertson JF: E-cadherin as a prognostic indicator in primary breast cancer. *BrJCancer* 2001, 85(12):1958-1963.

68. Ding L, Ellis MJ, Li S, Larson DE, Chen K, Wallis JW, Harris CC, McLellan MD, Fulton RS, Fulton LL *et al*: Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* 2010, 464(7291):999-1005.

69. Piccirilli C, Saxena A, Robertson J, Clark W, Ikejiri B, Oldfield E, Ali I: Allelic deletions on chromosome-17 and mutations in the p53 gene in tumors metastatic to brain. *IntJOncol* 1994, 4(1):37-42.

70. Lo NC, Vivenza D, Monteverde M, Lattanzio L, Gojis O, Garrone O, Comino A, Merlano M, Quinlan PR, Syed N *et al*: High frequency of complex TP53 mutations in CNS metastases from breast cancer. *BrJCancer* 2012, 106(2):397-404.

3

Chapter 4

A novel gene expression signature for bone metastasis in breast carcinomas

C. Dilara Savci Heijink, Hans Halfwerk, Jan Koster, Marc J. van de Vijver

Breast Cancer Res Treat. 2016 Apr;156(2):249-59

Abstract

Metastatic cancer remains the leading cause of death for patients with breast cancer. To understand the mechanisms underlying the development of distant metastases to specific sites is therefore important and of potential clinical value. From 157 primary breast tumors of the patients with known metastatic disease, gene expression profiling data were generated and correlated to metastatic behavior including site-specific metastasis, metastasis pattern and survival outcomes. We analyzed gene expression signatures specifically associated with the development of bone metastases. As a validation cohort, we used a published dataset of 376 breast carcinomas for which gene expression data and site specific metastasis information were available. 80.5% of luminal-type tumors developed bone metastasis as opposed to 41.7% of basal and 55.6% of HER2-like tumors. A novel 15-gene signature identified 82.4% of the tumors with bone metastasis, 85.2% of the tumors which had bone metastasis as first site of metastasis and 100% of the ones with bone metastasis only (p 9.99e-09), in the training set. In the independent data set, 81.2% of the positive tested tumors had known metastatic disease to the bone (p 4.28e-10). This 15-gene signature showed much better correlation with the development of bone metastases than previously identified signatures and was predictive in both ER- positive as well as in ER-negative tumors. Multivariate analyses revealed that together with the molecular subtype, our 15-gene expression signature was significantly correlated to bone metastasis status (p <0.001, 95% CI 3.86-48.02 in the training set; p 0.001, 95% CI 1.54-5.00 in the independent set). The 15 genes, *APOPEC3B*, *ATL2*, *BBS1*, *C6orf61*, *C6orf167*, *MMS22L*, *KCNS1*, *MFAP3L*, *NIP7*, *NUP155*, *PALM2*, *PH-4*, *PGD5*, *SFT2D2* and *STEAP3*, encoded mainly membrane-bound molecules with molecular function of protein binding. The expression levels of the upregulated genes (*NAT1*, *BBS1* and *PH-4*) were also found to be correlated to epithelial to mesenchymal transition (EMT) status of the tumor. We have identified a novel 15-gene expression signature associated with the development of bone metastases in breast cancer patients. This bone metastasis signature is the first to be identified using a supervised classification approach in a large series of patients and will help forward research in this area towards clinical applications.

Introduction

After the initial treatment of primary breast cancer, 20-30% of patients develop distant metastases [1, 2]. The survival outcomes and sites at which distant metastases develop differ greatly among patients [3-6]. Several studies have already reported gene expression profiles correlated with risk of distant metastasis, which are in the process of being validated with prospective studies [7-9]. Moreover, breast cancer's propensity to spread to certain organs, so called "non-random organ-specific metastasis" has also been investigated [10-14]. There have been several important studies using animal models to unravel the mechanism of site specific distant metastases in breast cancer [15-22]. These studies focusing on organotropism of metastatic breast cancer have used human breast cancer cell lines which were injected in immune compromised mice. By combining genomic profiling of organ-tropic metastatic variants selected *in vivo* from the animal models of metastatic disease with clinical genomic studies, *Massague* and his colleagues, were able to identify gene expression signatures that were associated with metastasis to bone, lung and brain [16, 17, 19]. They have further explored the association between specific patterns of gene expression and metastatic pattern. The discovered candidate genes were then further investigated and their metastatic role was confirmed by means of overexpressing or inactivating their expression. Hereafter they have validated these gene expression signatures in several cohorts of primary breast tumors with known metastatic disease.

We have recently described the metastatic behavior (organ-specific metastasis) related immunophenotypic findings of the primary tumors in a retrospective study including 263 primary breast tumors with known metastatic disease [23]. We have shown that the time to distant metastasis was less than 5 year in 90% of the hormone receptor negative breast cancer patients as compared to 66% of hormone receptor positive patients. The role of estrogen receptor (ER) positivity was found to be closely associated to the development of bone metastasis including bone-only and bone-first metastasis in the course of the disease; whereas ER negativity was found to be related to visceral (liver, lung or brain) metastasis. Along with the hormone status, tumor size and tumor grade, we found that patients who developed visceral metastasis had worse survival outcome, in terms of metastasis specific survival and overall survival and additionally

they frequently developed multiple metastasis during the course of the disease. We have concluded that tumor types were associated with survival and pattern of distant metastasis during the course of the disease. Gene expression profiling patterns predicting site-specific metastasis may aid in better understanding the mechanisms for the development of distant metastases.

In this study, we analysed the gene expression profile of 157 primary tumors that all metastasized. In order to identify and validate tumor factors of metastatic breast cancer that are predictive of metastatic behavior; gene expression profiling of primary tumors are correlated to metastasis pattern and subsequently gene expression signatures are investigated for prediction of the site specific distinct metastasis.

Materials and Methods

Patients and Tumor samples

The study was carried out according to the national ethical guidelines of 'Code for Proper Secondary Use of Human Tissue' developed by Federation of Medical Societies (FMWV) in the Netherlands [24]. Selection of patients and tumor samples and definition of metastatic disease has been previously described [23].

Briefly, metastatic breast cancer patients were identified and a subset of 157 of tumors with available frozen material was selected for further gene expression profiling experiments. Subsequently relevant clinical data were abstracted from the clinical charts. Detailed information on metastasis site, metastatic behavior and survival (metastasis specific and overall) data were collected for 151 patients.

The individual metastatic sites were recorded and used to separate tumors in subgroups. For each metastatic site it was assessed whether patients developed metastases during follow-up (ever versus never for each organ site); when patients developed metastases to any organ site it was recorded whether this was the first metastasis or a metastasis arising after metastases to other organ site arose (first/not first); and it was recorded when a patient developed metastases to one organ site only (only/not only). Presence

of multiple metastases was also carefully recorded at the time of diagnosis of the first metastases as well as after complete follow up. In instances where patients developed another distal organ involvement within less than two months after initial diagnosis of a metastasis, this was also considered as multiple organ metastases at first presentation.

Time from surgery to development of first metastasis, time from first metastasis to last event (metastasis specific survival, MSS) and overall survival (OS) time for each patient was calculated. Last event date was recorded as most recent follow-up date for the patients who were alive and time of death for the others.

Additionally, data on systemic treatment (chemotherapy, hormonal therapy and targeted therapy) used to treat primary and metastatic disease was collected (n=142 and n=122, respectively).

Hemotoxylin and eosin stained slides from all tumors were evaluated and histologic subtype, tumor grade and hormone receptor status were immunohistochemically assessed as previously described [23].

Gene expression profiling

From primary tumors of 157 patients 30 sections of 20-µm were used for isolation of RNA. The first and last sections (5 µm) were stained with hematoxylin and eosin; only samples containing an average of at least 50% tumor cells were used in this analysis. Total RNA was isolated with RNA-Bee (Tel-Test) and dissolved in RNase-free water. Then, total RNA was treated with DNase with use of the Qiagen RNase-Free DNase Kit and RNeasy spin columns and dissolved in RNase-free water. The gene expression microarrays used in this study were the HumanHT-12 v4 Expression BeadChip arrays (Illumina, Inc.) containing more than 47,000 probes. Details of RNA amplification, labeling, and hybridization are available on the Illumina website (http://www.illumina.com). The arrays were processed in the Central Microarray Facility of the Netherlands Cancer Institute. The data was normalized using robust spline normalization (rsn) and log2 transformed, followed by ComBat (http://www.bu.edu/jlab/wp-assets/ComBat/Abstract.html) to adjust for batch effects.

Human breast tumor microarray datasets

Additional to the gene expression profiling data set of 157 primary breast tumors generated herein with this study, a second data set which was a combined data set of four studies captured from the public domain was analysed. This large data set was employed as a combined data set as previously described by *Harrell* et al [25]. Aforementioned combined data set included total 855 tumors of which 376 with known metastatic disease. Analyses in order to validate the gene signatures were performed using this subset of tumors with known metastatic disease (n= 376) and whole data set (n= 855). Four microarray sets are listed as: GSE2034, GSE12276, GSE2603 and the NKI295 (*microarray-pubs.stanford.edu/wound_NKI/Clinical_Data_Supplement.xls*).

Microarray data analysis/ bioinformatics

All data were analyzed using the R2 (Microarray Analysis and Visualization Platform) web application, which is publicly available at http://r2.amc.nl.

Samples were classified into 5 intrinsic breast cancer subtypes by using the PAM50 classifier [26]. To assess the prognostic status of the tumors according to the 70-gene prognostic signature [9], these 70 genes were mapped via Gene Symbol ID to the Illumina platform. Out of the 70 genes 62 genes were found to be present on the Illumina platform corresponding to 65 Illumina probes. In case of presence of multiple probes for one gene, the probe with the highest variance across the samples was selected. We calculated the Pearson correlation coefficient between the centroids of the original good prognosis template and the gene expression levels of each sample with regard to these mapped 62 gene to assign tumors in good or poor prognostic group.

To validate already published gene expression signatures for site specific metastasis for bone metastasis [17] and epithelial to mesenchymal transition (EMT) gene signature [27], the specified genes from bone metastasis signature and from EMT signature were first mapped to the Illumina platform via Gene Symbol ID. As previously mentioned, in the instance of presence of multiple probes for one gene, the probe with the highest variance across the samples were selected. Respectively, a K-means method was used to cluster the data in 2 groups and a t-test revealed the performance of these signatures (bone metastasis and EMT signature) in our dataset.

Identification and validation of site-specific metastasis signature

In order to identify a gene expression signature associated an organ specific metastasis, we used the one-way ANOVA function in R2. Samples were split into 2 groups; one group in which the patient developed a metastasis and another group in which a patient never developed a certain organ metastasis. The genes with an expression level above the background expression level were selected (total 16036 genes. The calculated p-values (<0.01) were corrected for multiple testing using Benjamini Hochberg False Discovery Rate (FDR) calculation. The metastatic signature was validated in multiple datasets using the K-means and t-test function in R2. To further investigate the association between this gene signature and clinical variables multivariate logistic regression tests were applied using SPSS Statistics for Windows (Release version 21.0; IBM Corp.2012, Armond, NY). All statistical tests were two sided and $p < 0.05$ was considered to be statistically significant.

Results

For 157 primary invasive breast carcinomas from patients who all developed metastatic disease, mRNA expression signatures were assessed using microarray analysis. The patient characteristics and metastasis patterns are described in Table 1. Tumors were subdivided into 5 molecular subtypes using the PAM50 classifier [26]. Out of 157 cases, 67 (42.7%) were identified as Luminal A, 46 (29.3%) as Luminal B, 18 (11.5%) as HER2-like and 25 (15.9%) as basal type. One (0.6%) of these tumors was identified as normal-like. For statistical purposes, the normal like breast tumor was excluded from the multivariate analysis. Median follow-up time for patients who were alive was 11.5 years (range 6.2-17.3 years). 79.4% of the patients with Luminal A, 72.5% of Luminal B, 78.6% of HER2-like and 87.5% of basal-type tumors received adjuvant therapy. None of the patients received trastuzumab as adjuvant therapy; a subgroup of patients (*n*=10) received trastuzumab for treatment of metastatic disease.

Bone was the most frequent site of distant metastasis (71.5%) followed by liver (51.7%) and lung (34.4%). 74.2% of the patients developed visceral organ metastasis (lung, liver or brain).

Table 1. Clinical and pathological characteristics of metastatic breast cancer patients

		N	%
Age at diagnosis, years	<50	83	52.9
	>50	74	47.1
Surgical procedure	none	4	2.8
	mastectomy	73	51.8
	breast conserving	64	45.4
Adjuvant therapy	none	30	21.1
	only CT	50	35.2
	only HT	17	12.0
	CT+HT	45	31.7
Lymph node status	none	43	29.3
	1-3 positive	48	32.7
	>3 positive	56	38.1
Histology	Ductal	134	86.5
	Lobular	14	9.0
	Other	7	4.5
Tumor grade	1	13	8.6
	2	84	55.3
	3	55	36.2
Time to distant metastasis[a]	early	117	77.0
	late	35	23.0
Metastasis at first presentation	no	141	92.8
	yes	11	7.2
Multiple metastasis sites at first presentation	no	97	64.2
	yes	54	35.8
Multiple metastasis sites during follow up	no	37	24.5
	yes	114	75.5

CT chemotherapy, HT hormonal therapy.

[a] *Cut-off point 5 years.*

Survival analysis revealed that luminal-type tumors had better outcomes in terms of metastasis specific and overall survival compared to basal-type tumors and HER2-like tumors ($p < 0.001$). Median time to develop metastasis was 37, 27, 19 and 15 months for Luminal A, Luminal B, HER2-like tumors and basal-type tumors, respectively. 88.3% of basal-type and HER2-like tumors developed metastases within 5 years, versus 72.7% of luminal A and 76.7% of Luminal B tumors.

Among luminal subtype 80.5% of the tumors developed bone metastasis as opposed to, respectively, 41.7 and 55.6% of basal-type and HER2-like tumors (p 0.001). This group of tumors also composed the 81.8% of the tumors which metastasized to bone as initial site of metastasis (p 0.001). The rate of development of visceral metastasis were 70.4% in luminal-type tumors, 87.5% in basal-type tumors and 77.8% in HER2-like tumors. Of basal-type tumors 66.7% developed visceral metastasis as first metastasis site and 29.2% of these tumors had only visceral site metastasis during the course of disease (p 0.061 and p 0.034).

The tumor samples from all patients were assigned to the poor prognostic group according to the 70-gene signature [9]. Based on recently published epithelial mesenchymal transition (EMT) gene classifiers [27], 100 of the tumors allocated as EMT-activated and the rest, *n*=51, as EMT-non activated.

Validation of a previously identified gene signature for bone-specific metastasis

First we have studied the predictive value of the previously published bone metastasis signature of *Kang* et al [17]. This signature was assessed as positive in 110 of the tumors in the current study set. All (100%) Luminal A tumors and 90,7% of the Luminal B tumors were found be positive for the signature, whereas 33% of the HER2-like tumors were positive. None of the basal type tumors were found be positive for this site specific metastasis signature. Within this site-specific signature positive subgroup of tumors, 80% had clinically identified bone metastasis (*n*=88, p 4.26e-04). *Kang* et al' s 102-gene expression signature for bone metastasis was able to identify 81.5% of the tumors with bone metastasis, 84.1% of the tumors which had bone as initial site of metastasis and 100% (*n*=18) of the tumors which had bone only metastasis in the training set (p values <0.001, <0.001 and 0.002, respectively. Sensitivity: 81.5%

and specificity: 48.8%). When tested in ER-positive (*n*=108) and ER-negative (*n*=43) groups separately, 61.1% (*n*=66) of the ER-positive tumors and 60.4% (*n*=26) of the ER-negative tumors were tested to be positive with this 102-gene expression signature. Out of positively tested ER-positive tumors (*n*=66) 83.3% had clinically evident bone metastasis (p 0.456). Of the 26 bone signature positive tested ER-negative tumors, 50% had bone metastatic disease (p 1.000).

Supervised classification of bone (specific) metastasis-related genes

To identify site-specific metastasis genes, differentially expressed genes between tumors with bone metastasis and the ones without bone metastasis were explored. A t test was conducted with a p value of <0.01. After application of filtering criteria, differentially expressed genes were identified between two subgroups of tumors with and without bone metastasis. The group of differentially expressed genes were subsequently validated in the training data set as well as in the independent data set with the help of *K*-means and t testing.

We identified 15 differentially expressed genes between tumors with bone metastasis and the ones without bone metastasis (Table 2). The heat map with gene expression pattern of these 15 genes is displayed in Figure 1. None of the genes in this set overlapped with the bone signature of *Kang* et al. Three genes, namely *NAT1*, *PH-4* and *BBS1*, were up-regulated and the other genes were found to be down-regulated. Mapping into the Gene Ontology and Kyoto Encyclopaedia of Genes and Genomes databases showed an overrepresentation of membrane-bound molecules with molecular function of protein binding (*APOPEC3B*, *ATL2*, *BBS1*, *MMS22L*, *KCNS1*, *MFAP3L*, *NIP7*, *NUP155*, *PALM2*, *PH-4* and *STEAP3*).

In order to validate this gene expression signature, conjointly with our training set, an independent large combined microarray data set of four studies was analysed. This combined previously published by *Harrell* et al [25]. With the help of *K*-means clustering method, we have grouped our training data set and independent data set into two groups based on their expression levels for our newly developed bone metastasis gene expression signature and subsequently these two groups were compared using a t test.

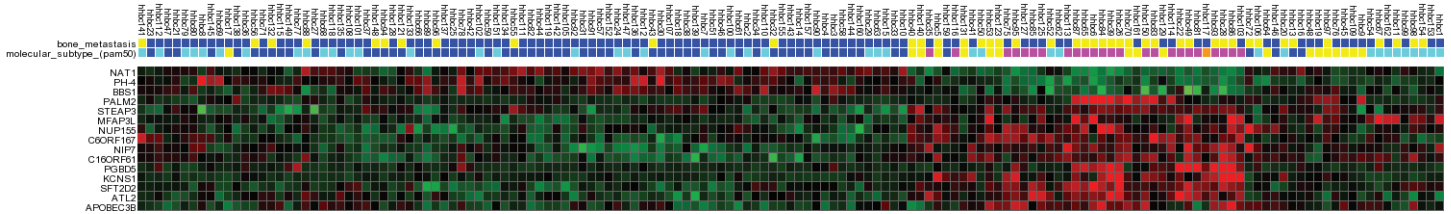


Figure 1. The gene expression pattern of 15 genes of bone metastasis gene signature.

Heat map shows the gene expression pattern of 15-genes among 151 patients. Primary tumors with clinically evident bone metastasis are illustrated in blue and the ones without bone metastasis are in yellow . For each primary tumor, the expression level of the specific gene is exhibited as red, if up-regulated and green, if down-regulated. Molecular subtypes of primary tumors are also demonstrated as; dark blue for luminal A, light blue for luminal B, yellow for HER2-like, pink for basal type and orange for normal-like tumors.

Table 2. The list of differentially expressed genes in bone metastatic disease

	Accession Number	HUGO	Description	R-value	p-value	Level of expression*
1	NM_004900	APOBEC3B	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B (APOBEC3B), mRNA.	-0,407	3,55e-03	<
2	NM_153485	NUP155	Nucleoporin 155kDa (NUP155), transcript variant 1, mRNA.	-0,385	8,43e-03	<
3	NM_021647	MFAP3L	Microfibrillar-associated protein 3-like (MFAP3L), transcript variant 1, mRNA.	-0,382	6,77e-03	<
4	NM_016101	NIP7	Nuclear import 7 homolog (S. cerevisiae) (NIP7), mRNA.	-0,375	8,67e-03	<
5	NM_198468	C6orf167	Chromosome 6 open reading frame 167 (C6orf167), mRNA.	-0,371	7,22e-03	<
6	NM_002251	KCNS1	Potassium voltage-gated channel, delayed-rectifier, subfamily S, member 1 (KCNS1), mRNA.	-0,368	7,41e-03	<
7	NM_001258311	PGBD5	PiggyBac transposable element derived 5 (PGBD5), mRNA.	-0,364	7,88e-03	<
8	NM_182915	STEAP3	STEAP family member 3 (STEAP3), transcript variant 1, mRNA.	-0,364	8,77e-03	<
9	NM_020188	C16orf61	Chromosome 16 open reading frame 61 (C16orf61), mRNA.	-0,357	9,84e-03	<
10	NM_053016	PALM2	Paralemmin 2 (PALM2), transcript variant 2, mRNA.	-0,356	9,02e-03	<
11	NM_022374	ATL2	Atlastin GTPase 2 (ATL2), mRNA.	-0,354	9,68e-03	<
12	NM_199344	SFT2D2	SFT2 domain containing 2 (SFT2D2), mRNA.	-0,353	9,66e-03	<
13	NM_001160170	NAT1	N-acetyltransferase 1 (arylamine N-acetyltransferase) (NAT1), mRNA.	0,352	9,24e-03	>
14	NM_177938	PH-4	Hypoxia-inducible factor prolyl 4-hydroxylase (PH-4), transcript variant 2, mRNA.	0,357	9,20e-03	>
15	NM_024649	BBS1	Bardet-Biedl syndrome 1 (BBS1), mRNA.	0,372	8,29e-03	>

* *> up-regulated, < down-regulated*

Table 3. Performance of the gene expression signatures

Gene expression signatures		Training data set				Independent data set		
			Bone metastasis			Bone metastasis		
		Signature	yes	no	p	yes	no	p
102-gene expression signature[a]	All	present	88	22	4.26e-04	146	55	6.92e-05
		absent	20	21		92	83	
	ER-positive	present	55	11	0.456	105	34	0.466
		absent	32	10		75	31	
	ER-negative	present	13	13	1.000	27	45	0.051
		absent	8	9		31	25	
15-gene expression signature[b]	All	present	89	14	9.99e-09	130	30	4.28e-10
		absent	19	29		108	108	
	ER-positive	present	69	9	1.99e-03	113	23	2.38e-04
		absent	18	12		67	42	
	ER-negative	present	14	6	0.015	42	32	3.83e-03
		absent	7	16		16	38	

[a] *The 102-gene signature by Kang et al.*
[b] *The 15-gene expression signature developed in this study.*
ER, estrogen receptor.

The 15-gene bone metastasis gene signature was found to be present in 103 tumors in the training data set. With the help of this signature, 82.4% of the tumors with known metastatic disease, 85.2% of the tumors which had bone metastasis as first metastasis site and 100% of the ones with bone metastasis only were identified (p 9.99e-09, sensitivity: 82.4% and specificity: 67.4%). When analysed in the independent data set, the 15-gene expression signature was found to be present in 160 tumors (total *n*=376) and 81.2% of these positive tested tumors had also clinically evident bone metastatic disease (p 4.28e-10, sensitivity: 54.6% and specificity: 78.2%). The independent data base of *Harrell* et al was also utilized to test the bone-specific metastasis of *Kang* et al. The 102-gene expression signature was assessed as present in 201 tumors (total *n* = 376) and 72.6% of these tumors reported to have bone metastasis (p 6.92e-05, sensitivity: 61.3% and specificity: 60.1%).

In addition, the independent data set was analysed separately in ER-positive and ER-negative tumors. Among ER-positive tumors (*n* = 245) the 15-gene expression signature was found to be present in 136 tumors and 83.1% of these tumors had known bone metastasis; 38.5% of the negatively tested tumors had no bone metastasis (p 2.38e-04, sensitivity: 79.3% and specificity: 57.1%). Out of 139 ER-positive tumors which were tested to be positive for the 102-gene expression signature, 75.5% had bone metastatic disease and 29.2% of the negatively tested tumors had no bone metastasis (p 0.466, sensitivity: 63.2% and specificity: 47.6%). Within the ER-negative subgroup (*n*=128) 74 tumors were tested positive for the 15-gene expression signature and 56.8% these tumors had bone metastasis; 70.4% of negatively tested tumors had no evidence of bone metastasis (p 3.83e-03, sensitivity: 72.4% and specificity: 56.8%). Out of 56 ER-negative tumors which were tested positive for 102-gene expression signature, 55.4% had clinically bone metastasis; 62.5% of negatively tested tumors had no bone metastasis (p 0.05, sensitivity: 53.5% and specificity: 64.3%). Table 3 summarizes the validation of gene signatures in training and independent data sets.

In addition, in a subsequent study a subset of 50-genes (out of initially identified 102 genes) was selected by *Massague*'s group [20]; this subset of 50 genes was also analysed in our training and in the independent data sets for its predictive value for bone-specific metastasis The 50-gene signature was able to identify the patients with bone metastasis in the training set (p 1.14e-03) and the independent data set (p 0.014).

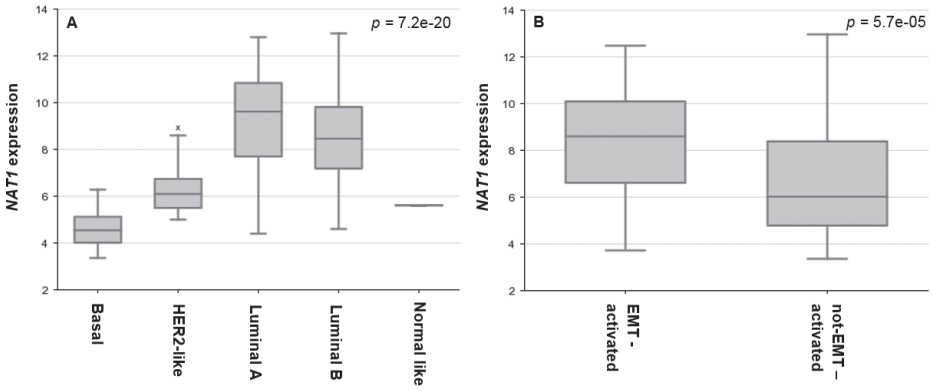


Figure 2. The expression levels (log2) of NAT1 among molecular subtypes (A) and in EMT-activated and not-EMT-activated group (B).

The box plots show that NAT1 expression was higher in Luminal-type tumors compared to the other molecular subtypes (p 7.2e-20). NAT1 expression was also found to be higher in the EMT (epithelial to mesenchymal transition)-activated group (p 5.7e-05)

When tested in the ER-positive and the ER-negative tumors separately, this 50-gene signature was not predictive for bone metastatic disease.

When tested among all patients with metastatic and not-metastatic disease in the independent data set (*n*=855), the 15-gene signature was able to identify the patients with bone metastasis (p 5.48e04, sensitivity: 54.6% and specificity: 58.7%). This gene expression signature remained statistically significant for identification of bone metastasis when separately analysed in ER-positive and ER-negative tumors (p 3.45e-04, sensitivity: 63.9% and specificity: 52.2%; p 3.82e-03, sensitivity: 75.9% and specificity: 45.5%, respectively).

The up-regulated genes and their correlation with molecular subtypes and known prognostic gene signatures were further explored. *NAT1* was identified to be expressed at the highest levels in Luminal A followed by Luminal B, HER2-like group and being least expressed in the basal-type group. *NAT1* expression was also correlated with the EMT-activated group, being overexpressed in this group of tumors compared to the not-EMT-activated group (p 5.7e-05) (Figure 2). Similarly the other up-regulated genes,

Table 4. Multivariate analyses results of predictive factors among the gene datasets

	Training data set				
	B	**Wald x^2**	**p**	**Odds ratio**	**95% CI**
ER status	-0.48	0.53	.468	0.620	0.17 – 2.25
Molecular subtype	0.53	0.07	.793	1.05	0.71 - 1.57
15-gene signature[a]	2.61	16.49	< .000	13.61	3.86 – 48.03

	Independent data set				
	B	**Wald x^2**	**p**	**Odds ratio**	**95% C.I.**
ER status	0.25	0.06	.939	1.02	0.54 - 1.96
Molecular subtype	0.30	10.70	.001	1.36	1.13 - 1.64
15-gene signature[a]	2.62	11.54	.001	2.78	1.54 – 5.00

[a] *Novel gene expression signature*
ER estrogen receptor, CI confidence interval

BBS1 and *PH-4*, were also found to be significantly correlated with the EMT-activated group of tumors (p 5.8e-04 and p 0.01, respectively).

The 15-gene bone metastasis signature was positive in 96.9% of the Luminal A tumors, in 79.7% of Luminal B tumors and in 38.9% of HER2-like tumors. Similar to *Kang*'s bone metastasis signature, none of the basal like tumors were found to be positive for this signature.

Univariate analyses showed that our bone metastasis signature was significantly correlated to the development of bone metastasis especially in the group of patients who developed only bone metastasis in the course of their disease ($p < 0.001$). As expected, ER status and molecular subtypes were the parameters that were closely related to bone metastasis status ($p < 0.001$). Subsequently, multivariate analyses were applied in order to further explore the link between our gene signature and these parameters. Table 4 displays the multivariate analyses results for ER status, molecular subtypes and two separate gene data sets (training and independent) for bone specific metastasis. As shown, the 15-gene signature was the only parameter that was significantly correlated to bone metastasis status in the training data set (p <0.001, 95% CI 3.86-48.02). In the independent data set, together with the molecular subtype, the 15-gene signature was significantly correlated to bone metastasis status (p 0.001, 95% CI 1.54-5.00).

Discussion

The metastatic potential of the primary tumor revolves around multistep biological processes within host tissue and microenvironment of the distant organ site [28]. In addition to the early origin of genetic instability [28-30] and hence the metastatic potential of the tumor cells, several intrinsic and extrinsic factors are recognised as potential promoters of metastatic relapse [31-33]. Upon sustaining the elementary steps of dissemination, the circulating tumor cells can colonize a new organ, forming a detectable metastasis [11, 28].

Experimental models of metastasis yielded distinct sets of genes that mediated site-specific metastasis in breast cancer [16-19]. *Kang* et al. identified a bone metastasis signature composed of 102 genes mostly encoding cell surface and secretory proteins, with functions including bone marrow homing and extravasation, pericellular proteolysis and invasion, angiogenesis, osteoclastogenesis, growth factor regulation and extracellular matrix alteration [17]. The authors concluded that this gene set was superimposed on a poor-prognosis gene signature to provide additional functions in order to achieve an overt bone-specific metastasis.

Despite these interesting findings from mouse model system and validation of the results from the mouse models in human breast cancer, no clinical application or follow-up research has emerged since these first findings. Here we present results of the largest study to date on the association between gene expression profiling of primary breast cancer and the development of bone metastases, and the first study in which supervised classification has been used to identify a bone metastasis associated gene expression signature. This gene expression signature was composed of 15 genes, with 3 (*NAT1, PH-4 and BBS1*) of them being up-regulated in the primary tumor samples. The overexpressed genes in this bone-specific metastasis signature were associated with metabolic (*NAT1*) and oxidation-reduction (*PH-4*) processes, and protein transport (*BBS-1*), in agreement with previous works hypothesizing their potential role in altering the host tissue environment in order to achieve a bone metastasis [31, 33-35].

N-acetyltransferase 1 (*NAT1*) was first reported to be associated with enhanced growth and survival of breast epithelial cells by *Adam* et al. [36], and later reported to be a potential biomarker for breast cancer [37-41]. In several studies, inhibiting *NAT1* resulted in cell morphology change, a loss of surface filopodia and subsequent reduction of invasive potential both in vitro and in vivo [40]. Likewise, knockdown of this gene led to inhibition of invasion and metastasis, by means of modification/rearrangement of filopodial (intracellular) actin [41, 42]. In agreement with other gene expression profiling studies in human cancer samples, here we showed that *NAT1* clusters close to the estrogen receptor with higher expression levels in luminal-type tumors [36, 43, 44]. *Tiang* et al. also showed that the loss of *NAT1* resulted in alteration of cell-to-cell contact and up-regulation of E-cadherin. Based on aforementioned cell-line studies a possible association between this gene and EMT/MET has been speculated [42]. Interestingly, in our data set overexpression of this gene was significantly correlated to the so-called EMT-activated group (p 5.7e-05). To our knowledge, this is the first study pointing to the association between *NAT1* and EMT in human female breast cancer samples. Along with the considerations of the potentiality of this gene as a drug target [42, 45], we believe that further studies in human breast cancer samples are indicated to explore this link.

The extracellular matrix (ECM) plays important role in diverse pathological and physiological processes, including cancer invasion and metastasis [46, 47]. Collagens compose the major component of ECM. Increased expression of collagens, thereupon increase in deposition and stiffening in ECM, is associated with tumor progression [48, 49]. Collagen prolyl 4-hydroxylase (*PH-4*), a member of post-transcription modification enzyme family, is required in collagen biosynthesis and angiogenesis. Hypoxia-induced collagen prolyl 4-hydroxylase expression is reported to be associated with increased progression and mortality in breast cancer [49-51]. Indeed, animal studies showed that knockdown of *PH-4* resulted in inhibition of tumor growth and lung metastasis [52, 53]. With gene expression profiling of breast cancer samples we have found that *PH-4* was positively correlated with site-specific metastasis to bone. This finding confirms the observations by others [46-49] and advocates for the importance of extracellular matrix alterations in disease progression.

Twelve out of 15 genes were found to be down-regulated in the primary tumors of breast cancer patients who developed bone metastasis. One of these genes, apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like-3B (*APOBEC3B*), is reported to be up-regulated in a large proportion of breast tumors and high levels of *APOBEC3B* were found to be associated with worse disease-free and overall survival [54-57]. Recently, several independent genome-wide association studies have shown a deletion resulting in complete elimination of the *APOBEC3B* gene-encoding region [58-60]. This deletion has been indicated to be associated with decreased expression of *APOBEC3B* in breast cancer cells [58]. In this study, we have also shown that *APOBEC3B* was significantly down-regulated in the group of tumors with bone metastatic disease (p 3.55e-03). We believe that further copy number variations studies are required to explore such an association between *APOBEC3B* deletion and site-specific metastasis. Six-transmembrane epithelial antigen of prostate 3 (*STEAP3*), which is thought be involved in apoptosis and cell-cycle progression [61-63], is also found to be down-regulated in the bone metastatic group of primary breast tumors in our study. *STEAP3* expression is shown to be diminished in hepatocellular carcinoma nodules compared to cirrhotic peritumoral tissue and healthy liver [64]. Another family member of these proteins, *STEAP1*, has already shown to be overexpressed in breast cancer cells [65-67]. However, we could not retrieve any similar data pointing *STEAP3* expression levels in breast cancer tissues.

In order to determine the validity of the experimentally derived 102 gene bone metastasis signature, *Kang* et al. have utilized a cohort of 63 primary breast carcinomas to test this signature. The authors have selected a subset of 50 genes to carry on their validation studies and they have shown that this gene set was not able to identify the group of tumors with bone metastasis. When the authors restricted their analyses to 25 breast tumors with known metastatic disease, they were able distinguish the tumors preferentialy metastasized to bone rather than other distant organs [20]. In this current study along with new identified 15-gene expression signature, we have shown that the 102-gene expression signature and the subset of 50 genes as reported by *Kang* et al. were informative in identifying likelihood of developing bone metastasis in the training and the independent data sets. However, when datasets subdivided into 2 groups according to their ER-status, the 102-gene expression signature as well as the 50-gene signature were not effective in predicting bone metastasis, whereas herein

identified 15-gene expression signature remained associated with the likelihood of bone metastasis development in ER-positive and ER-negative tumor groups.

Notably, the bone-specific metastasis signature presented in this study did not include any of the genes from already published *Kang' s* bone signature [17]. The absence of overlap between these gene sets could be justified with the fact that in the former study tumor cells from the metastasis site were utilized to generate gene signatures in contrast to primary tumors in the current study. Considering that tumor progression and development of metastasis requires compiled steps of modification, we may assume that these two different gene signature sets play a complementary role in separate levels of this multi-complex process.

Notwithstanding several well received studies focusing on the biology of metastasic breast cancer, little progress has been made over the past years to identify a robust gene expression signature for site-specific metastasis. Moreover, the experimentally derived gene expression signatures when tested in human breast carcinomas were not as strongly associated with site-specific metastasis as in the experimental conditions. A reproducible gene expression signature associated with the development of bone metastases in breast cancer will have clinical utility in two ways: first, the knowledge of the specific gene expressed at higher or lower levels in the metastatic disease will lead to the investigation of targeted therapy options directed to the altered mechanism related to this gene and second, reliable identification of the patients at high risk of developing bone metastases may lead to therapeutic interventions specifically aimed at preventing the development of bone metastases, for example treatment with bisphosphonates.

In summary, we present the largest study to date revealing the association between the gene expression profiling patterns and bone specific metastasis in breast carcinomas. The identification of novel 15-gene expression signature will forward this area of research, including subsequent exploration of the underlying mechanisms of metastatic behavior and ultimately help improve outcome for breast cancer patients.

References

1. Early Breast Cancer Trialists' Collaborative G: Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005, 365(9472):1687-1717.
2. O'Shaughnessy J: Extending survival with chemotherapy in metastatic breast cancer. *Oncologist* 2005, 10 Suppl 3:20-29.
3. Giordano SH, Buzdar AU, Smith TL, Kau SW, Yang Y, Hortobagyi GN: Is breast cancer survival improving? *Cancer* 2004, 100(1):44-52.
4. Greenberg PA, Hortobagyi GN, Smith TL, Ziegler LD, Frye DK, Buzdar AU: Long-term follow-up of patients with complete remission following combination chemotherapy for metastatic breast cancer. *J Clin Oncol* 1996, 14(8):2197-2205.
5. Hayes DF, Van Zyl JA, Hacking A, Goedhals L, Bezwoda WR, Mailliard JA, Jones SE, Vogel CL, Berris RF, Shemano I *et al*: Randomized comparison of tamoxifen and two separate doses of toremifene in postmenopausal patients with metastatic breast cancer. *J Clin Oncol* 1995, 13(10):2556-2566.
6. Sledge GW, Jr., Hu P, Falkson G, Tormey D, Abeloff M: Comparison of chemotherapy with chemohormonal therapy as first-line therapy for metastatic, hormone-sensitive breast cancer: An Eastern Cooperative Oncology Group study. *J Clin Oncol* 2000, 18(2):262-266.
7. Drukker CA, van Tinteren H, Schmidt MK, Rutgers EJ, Bernards R, van de Vijver MJ, Van't Veer LJ: Long-term impact of the 70-gene signature on breast cancer outcome. *Breast cancer research and treatment* 2014, 143(3):587-592.
8. Mook S, Knauer M, Bueno-de-Mesquita JM, Retel VP, Wesseling J, Linn SC, Van't Veer LJ, Rutgers EJ: Metastatic potential of T1 breast cancer can be predicted by the 70-gene MammaPrint signature. *Annals of surgical oncology* 2010, 17(5):1406-1413.
9. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ *et al*: A gene-expression signature as a predictor of survival in breast cancer. *NEnglJMed* 2002, 347(25):1999-2009.
10. Aslakson CJ, Miller FR: Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. *Cancer Res* 1992, 52(6):1399-1405.
11. Chambers AF, Groom AC, MacDonald IC: Dissemination and growth of cancer cells in metastatic sites. *NatRevCancer* 2002, 2(8):563-572.
12. Ding L, Ellis MJ, Li S, Larson DE, Chen K, Wallis JW, Harris CC, McLellan MD, Fulton RS, Fulton LL *et al*: Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* 2010, 464(7291):999-1005.
13. Hu G, Kang Y, Wang XF: From breast to the brain: unraveling the puzzle of metastasis organotropism. *JMolCell Biol* 2009, 1(1):3-5.

14. Lorusso G, Ruegg C: New insights into the mechanisms of organ-specific breast cancer metastasis. *SeminCancer Biol* 2012, 22(3):226-233.

15. Bos PD, Nguyen DX, Massague J: Modeling metastasis in the mouse. *Curr Opin Pharmacol* 2010, 10(5):571-577.

16. Bos PD, Zhang XH, Nadal C, Shu W, Gomis RR, Nguyen DX, Minn AJ, van de Vijver MJ, Gerald WL, Foekens JA *et al*: Genes that mediate breast cancer metastasis to the brain. *Nature* 2009, 459(7249):1005-1009.

17. Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Cordon-Cardo C, Guise TA, Massague J: A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* 2003, 3(6):537-549.

18. Landemaine T, Jackson A, Bellahcene A, Rucci N, Sin S, Abad BM, Sierra A, Boudinet A, Guinebretiere JM, Ricevuto E *et al*: A six-gene signature predicting breast cancer lung metastasis. *Cancer Res* 2008, 68(15):6092-6099.

19. Minn AJ, Gupta GP, Padua D, Bos P, Nguyen DX, Nuyten D, Kreike B, Zhang Y, Wang Y, Ishwaran H *et al*: Lung metastasis genes couple breast tumor size and metastatic spread. *Proceedings of the National Academy of Sciences of the United States of America* 2007, 104(16):6740-6745.

20. Minn AJ, Kang Y, Serganova I, Gupta GP, Giri DD, Doubrovin M, Ponomarev V, Gerald WL, Blasberg R, Massague J: Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors. *The Journal of clinical investigation* 2005, 115(1):44-55.

21. Zhang XH, Jin X, Malladi S, Zou Y, Wen YH, Brogi E, Smid M, Foekens JA, Massague J: Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. *Cell* 2013, 154(5):1060-1073.

22. Zhang XH, Wang Q, Gerald W, Hudis CA, Norton L, Smid M, Foekens JA, Massague J: Latent bone metastasis in breast cancer tied to Src-dependent survival signals. *Cancer Cell* 2009, 16(1):67-78.

23. Savci-Heijink CD, Halfwerk H, Hooijer GK, Horlings HM, Wesseling J, van de Vijver MJ: Retrospective analysis of metastatic behaviour of breast cancer subtypes. *Breast cancer research and treatment* 2015, 150(3):547-557.

24. Vermeulen E, Geesink I, Schmidt MK, Steegers C, Verhue D, Brom FW, Aaronson NK, van Leeuwen FE: [Secondary use of human tissue: consent and better information required]. *NedTijdschrGeneeskd* 2009, 153:A948.

25. Harrell JC, Prat A, Parker JS, Fan C, He X, Carey L, Anders C, Ewend M, Perou CM: Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer ResTreat* 2012, 132(2):523-535.

26. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z *et al*: Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009, 27(8):1160-1167.

27. Groger CJ, Grubinger M, Waldhor T, Vierlinger K, Mikulits W: Meta-analysis of gene expression signatures defining the epithelial to mesenchymal transition during cancer progression. *PLoS One* 2012, 7(12):e51136.

28. Fidler IJ: The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer* 2003, 3(6):453-458.

29. Bartkova J, Horejsi Z, Koed K, Kramer A, Tort F, Zieger K, Guldberg P, Sehested M, Nesland JM, Lukas C *et al*: DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* 2005, 434(7035):864-870.

30. Feinberg AP, Ohlsson R, Henikoff S: The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 2006, 7(1):21-33.

31. Chi JT, Wang Z, Nuyten DS, Rodriguez EH, Schaner ME, Salim A, Wang Y, Kristensen GB, Helland A, Borresen-Dale AL *et al*: Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. *PLoS Med* 2006, 3(3):e47.

32. Nuyten DS, Hastie T, Chi JT, Chang HY, van de Vijver MJ: Combining biological gene expression signatures in predicting outcome in breast cancer: An alternative to supervised classification. *Eur J Cancer* 2008, 44(15):2319-2329.

33. Semenza GL: Targeting HIF-1 for cancer therapy. *Nat Rev Cancer* 2003, 3(10):721-732.

34. Harris AL: Hypoxia--a key regulatory factor in tumor growth. *Nat Rev Cancer* 2002, 2(1):38-47.

35. Pennacchietti S, Michieli P, Galluzzo M, Mazzone M, Giordano S, Comoglio PM: Hypoxia promotes invasive growth by transcriptional activation of the met protooncogene. *Cancer Cell* 2003, 3(4):347-361.

36. Adam PJ, Berry J, Loader JA, Tyson KL, Craggs G, Smith P, De Belin J, Steers G, Pezzella F, Sachsenmeir KF *et al*: Arylamine N-acetyltransferase-1 is highly expressed in breast cancers and conveys enhanced growth and resistance to etoposide in vitro. *Mol Cancer Res* 2003, 1(11):826-835.

37. Dolled-Filhart M, Ryden L, Cregger M, Jirstrom K, Harigopal M, Camp RL, Rimm DL: Classification of breast cancer using genetic algorithms and tissue microarrays. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2006, 12(21):6459-6468.

38. Endo Y, Yamashita H, Takahashi S, Sato S, Yoshimoto N, Asano T, Hato Y, Dong Y, Fujii Y, Toyama T: Immunohistochemical determination of the miR-1290 target arylamine N-acetyltransferase 1 (NAT1) as a prognostic biomarker in breast cancer. *BMC Cancer* 2014, 14:990.

39. Laurieri N, Crawford MH, Kawamura A, Westwood IM, Robinson J, Fletcher AM, Davies SG, Sim E, Russell AJ: Small molecule colorimetric probes for specific detection of human arylamine N-acetyltransferase 1, a potential breast cancer biomarker. *J Am Chem Soc* 2010, 132(10):3238-3239.

40. Tiang JM, Butcher NJ, Minchin RF: Small molecule inhibition of arylamine N-acetyltransferase Type I inhibits proliferation and invasiveness of MDA-MB-231 breast cancer cells. *Biochem Biophys Res Commun* 2010, 393(1):95-100.

41. Tiang JM, Butcher NJ, Minchin RF: Effects of human arylamine N-acetyltransferase I knockdown in triple-negative breast cancer cell lines. *Cancer Med* 2015, 4(4):565-574.

42. Tiang JM, Butcher NJ, Cullinane C, Humbert PO, Minchin RF: RNAi-mediated knock-down of arylamine N-acetyltransferase-1 expression induces E-cadherin up-regulation and cell-cell contact growth inhibition. *PLoS One* 2011, 6(2):e17031.

43. Bieche I, Girault I, Urbain E, Tozlu S, Lidereau R: Relationship between intratumoral expression of genes coding for xenobiotic-metabolizing enzymes and benefit from adjuvant tamoxifen in estrogen receptor alpha-positive postmenopausal breast carcinoma. *Breast Cancer Res* 2004, 6(3):R252-263.

44. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS *et al*: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the United States of America* 2001, 98(19):10869-10874.

45. Sugamori KS, Wong S, Gaedigk A, Yu V, Abramovici H, Rozmahel R, Grant DM: Generation and functional characterization of arylamine N-acetyltransferase Nat1/Nat2 double-knockout mice. *Mol Pharmacol* 2003, 64(1):170-179.

46. Frantz C, Stewart KM, Weaver VM: The extracellular matrix at a glance. *J Cell Sci* 2010, 123(Pt 24):4195-4200.

47. Hynes RO: The extracellular matrix: not just pretty fibrils. *Science* 2009, 326(5957):1216-1219.

48. Levental KR, Yu H, Kass L, Lakins JN, Egeblad M, Erler JT, Fong SF, Csiszar K, Giaccia A, Weninger W *et al*: Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* 2009, 139(5):891-906.

49. Provenzano PP, Inman DR, Eliceiri KW, Knittel JG, Yan L, Rueden CT, White JG, Keely PJ: Collagen density promotes mammary tumor initiation and progression. *BMC Med* 2008, 6:11.

50. Conklin MW, Eickhoff JC, Riching KM, Pehlke CA, Eliceiri KW, Provenzano PP, Friedl A, Keely PJ: Aligned collagen is a prognostic signature for survival in human breast carcinoma. *Am J Pathol* 2011, 178(3):1221-1232.

51. Fox SB, Generali D, Berruti A, Brizzi MP, Campo L, Bonardi S, Bersiga A, Allevi G, Milani M, Aguggini S *et al*: The prolyl hydroxylase enzymes are positively associated with hypoxia-inducible factor-1alpha and vascular endothelial growth factor in human breast cancer and alter in response to primary systemic treatment with epirubicin and tamoxifen. *Breast Cancer Res* 2011, 13(1):R16.

52. Wottawa M, Leisering P, Ahlen M, Schnelle M, Vogel S, Malz C, Bordoli MR, Camenisch G, Hesse A, Napp J *et al*: Knockdown of prolyl-4-hydroxylase domain 2 inhibits tumor growth

of human breast cancer MDA-MB-231 cells by affecting TGF-beta1 processing. *Int J Cancer* 2013, 132(12):2787-2798.

53. Gilkes DM, Bajpai S, Chaturvedi P, Wirtz D, Semenza GL: Hypoxia-inducible factor 1 (HIF-1) promotes extracellular matrix remodeling under hypoxic conditions by inducing P4HA1, P4HA2, and PLOD2 expression in fibroblasts. *J Biol Chem* 2013, 288(15):10819-10829.

54. Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, Refsland EW, Kotandeniya D, Tretyakova N, Nikas JB *et al*: APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 2013, 494(7437):366-370.

55. Burns MB, Temiz NA, Harris RS: Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat Genet* 2013, 45(9):977-983.

56. Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, Kiezun A, Kryukov GV, Carter SL, Saksena G *et al*: An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat Genet* 2013, 45(9):970-976.

57. Sieuwerts AM, Willis S, Burns MB, Look MP, Meijer-Van Gelder ME, Schlicker A, Heideman MR, Jacobs H, Wessels L, Leyland-Jones B *et al*: Elevated APOBEC3B correlates with poor outcomes for estrogen-receptor-positive breast cancers. *Horm Cancer* 2014, 5(6):405-413.

58. Komatsu A, Nagasaki K, Fujimori M, Amano J, Miki Y: Identification of novel deletion polymorphisms in breast cancer. *Int J Oncol* 2008, 33(2):261-270.

59. Long J, Delahanty RJ, Li G, Gao YT, Lu W, Cai Q, Xiang YB, Li C, Ji BT, Zheng Y *et al*: A common deletion in the APOBEC3 genes and breast cancer risk. *J Natl Cancer Inst* 2013, 105(8):573-579.

60. Xuan D, Li G, Cai Q, Deming-Halverson S, Shrubsole MJ, Shu XO, Kelley MC, Zheng W, Long J: APOBEC3 deletion polymorphism is associated with breast cancer risk among women of European ancestry. *Carcinogenesis* 2013, 34(10):2240-2243.

61. Aerbajinai W, Giattina M, Lee YT, Raffeld M, Miller JL: The proapoptotic factor Nix is coexpressed with Bcl-xL during terminal erythroid differentiation. *Blood* 2003, 102(2):712-717.

62. Liu F, Stanton JJ, Wu Z, Piwnica-Worms H: The human Myt1 kinase preferentially phosphorylates Cdc2 on threonine 14 and localizes to the endoplasmic reticulum and Golgi complex. *Mol Cell Biol* 1997, 17(2):571-583.

63. Yussman MG, Toyokawa T, Odley A, Lynch RA, Wu G, Colbert MC, Aronow BJ, Lorenz JN, Dorn GW, 2nd: Mitochondrial death protein Nix is induced in cardiac hypertrophy and triggers apoptotic cardiomyopathy. *Nat Med* 2002, 8(7):725-730.

64. Coulouarn C, Derambure C, Lefebvre G, Daveau R, Hiron M, Scotte M, Francois A, Daveau M, Salier JP: Global gene repression in hepatocellular carcinoma and fetal liver, and suppression of dudulin-2 mRNA as a possible marker for the cirrhosis-to-tumor transition. *J Hepatol* 2005, 42(6):860-869.

65. Hubert RS, Vivanco I, Chen E, Rastegar S, Leong K, Mitchell SC, Madraswala R, Zhou Y, Kuo J, Raitano AB *et al*: STEAP: a prostate-specific cell-surface antigen highly expressed

in human prostate tumors. *Proceedings of the National Academy of Sciences of the United States of America* 1999, 96(25):14523-14528.

66. Korkmaz CG, Korkmaz KS, Kurys P, Elbi C, Wang L, Klokk TI, Hammarstrom C, Troen G, Svindland A, Hager GL *et al*: Molecular cloning and characterization of STAMP2, an androgen-regulated six transmembrane protein that is overexpressed in prostate cancer. *Oncogene* 2005, 24(31):4934-4945.

67. Maia CJ, Socorro S, Schmitt F, Santos CR: STEAP1 is over-expressed in breast cancer and down-regulated by 17beta-estradiol in MCF-7 cells and in the rat mammary gland. *Endocrine* 2008, 34(1-3):108-116.

4

Chapter 5

A specific gene expression signature for visceral organ metastasis in breast cancer

C. Dilara Savci Heijink, Hans Halfwerk, Jan Koster, Hugo M. Horlings, Marc J. van de Vijver

Submitted

Abstract

Visceral organ metastasis is associated with poor survival outcomes in terms of metastasis free survival and overall survival in breast carcinomas. Identification of a gene expression profile in tumors that selects a subpopulation of patients that is more likely to develop visceral organ metastases will help elucidate mechanisms for the development of distant metastases and could be of clinical value. Gene expression profiling data of 157 primary tumors from breast cancer patients, who during follow-up developed distant metastases were analyzed and differentially expressed genes between the group of tumors with visceral metastasis (liver, lung and brain) and the those without visceral metastases were identified. Published data were used to validate our findings. Multivariate logistic regression tests were applied to further investigate the association between the gene expression signature and clinical variables. Survival analyses were performed by the Kaplan-Meier method. 14 differentially expressed genes (*WDR6*, *CDYL*, *ATP6V0A4*, *CHAD*, *IDUA*, *MYL5*, *PREP*, *RTN4IP1*, *BTG2*, *TPRG1*, *ABHD14A*, *KIF18A*, *S100PBP* and *BEND3*) were identified between the group of tumors with and without visceral metastatic disease. Five of these genes (*CDYL*, *ATP6V0A4*, *PREP*, *RTN4IP1* and *KIF18A*) were up-regulated and the other genes were down-regulated. This gene expression signature was validated in the training and in the independent data set (p 2.13e-08 and p 9.68e-06, respectively). Multivariate analyses revealed that the 14-gene expression signature was associated with visceral metastatic disease (p 0.001, 95% CI 1.43-4.27), independent of other clinicopathologic features. This signature has been also found to be associated with survival status of the patients ($p < 0.001$). We have identified an unique gene expression signature which is specific to visceral metastasis. This 14-gene expression signature may play a role in identifying the subgroup of patients with potential to develop visceral metastasis.

Introduction

The implementation of breast cancer screening programmes and improved options for the treatment of patients with early breast cancer have contributed to the improved outcome in breast cancer [1]. However, once metastatic disease develops, breast cancer is still a deadly disease [2, 3]. Predicting the likelihood of metastatic behavior and organ-specific metastasis of the primary tumors could help to improve the modalities for the treatment of the primary tumor and of metastatic disease.

The relationship between primary tumor and their metastases has been an important area in cancer research since the "seed and soil' theory proposed by *Stephen Paget* [4]. Several studies have investigated the predictors for metastatic potential [5-7] of the primary tumors and site- specific distant organ metastasis [8-11] in breast cancer. Some of these studies using animal models and genomic profiling have identified gene expression signatures that were associated with organ specific metastasis. In particular, using an experimental system based on the *in vivo* selection of MDA-MB231- derived breast cancer cell lines with specific organotropism, *Massague* and co-workers have identified genes related to bone, lung and brain metastasis. They have suggested that whereas some genes may determine a breast cancer's overall classification and prognostic signature, some other superimposed tissue-specific metastatic gene expression profile(s) in a subset of tumor cells may affect the cancer's organ specific metastatic behavior [12].

It has been demonstrated that metastases and primary breast cancers show similar gene expression profiles [13-15]. We previously reported gene expression profiling experiments performed on primary breast cancers to identify gene expression profiles for organ-specific metastasis. We have recently presented a novel 15-gene expression signature for bone specific metastasis in breast cancer [16]. This new gene expression signature was found to be associated with the likelihood of bone metastasis development in ER-positive and ER-negative tumors, in the training set as well as in an independent data set including 376 tumors with known clinical metastatic disease. We have shown that 80.5% of the patients with luminal subtype tumors developed bone metastasis as opposed to, respectively, 41.7% and 55.6% of the basal-type and HER2-like tumors

(p 0.001). We have also identified that 70.4% of luminal type tumors, 87.5% of basal type tumors and 77.8% of HER2-like tumors developed visceral organ metastasis (liver, lung and brain). Among basal type tumors 66.7% developed visceral metastasis as first metastasis site and 29.2% of these tumors had only visceral metastasis during the course of disease. Survival analyses revealed that patients who developed visceral metastasis had worse survival outcome, in terms of metastasis specific survival and overall survival and they frequently developed multiple metastasis during the course of the disease [17]. In this study we sought a gene expression profiling identifier to select the subgroup of tumors that are most likely to develop visceral organ metastasis.

Here we present a gene expression signature which is found to be associated with development of visceral organ metastasis in breast carcinomas.

Materials and Methods

Patients and Tumor samples

157 primary breast carcinomas from patients who all developed distant metastases were included in this study. This series of tumors has been described previously [17]. The national ethical guidelines of 'Code for Proper Secondary Use of Human Tissue' developed by Federation of Medical Societies (FMWV) in the Netherlands were followed for this study [18].

Clinical data with detailed information on metastatic behavior, metastasis site and survival outcomes were abstracted from the clinical charts for 151 patients as previously published . Briefly, metastasis site was carefully recorded and classified into ever versus never, first versus not first and only versus not only for each organ site. In addition, data on systemic treatment (chemotherapy, hormonal therapy and targeted therapy) used to treat primary and metastatic disease was also available for a subset of patients (n=142 and n=122, respectively). Tumors were evaluated and histological and immunohistochemical characteristics were assessed as previously published [17].

Gene expression profiling and human breast tumor microarray data sets

RNA extraction, amplification, labeling and hybridization have been described and details are available on Illumina website (http://www.illumina.com) [16]. As described, the arrays were processed in the Central Microarray Facility of the Netherlands Cancer Institute. The data was normalized using robust spline normalization (rsn) and log2 transformed, followed by ComBat (http://www.bu.edu/jlab/wp-assets/ComBat/Abstract.html) to adjust for batch effects. Next to this already published gene expression profiling data set of 157 primary breast tumors a combined data set (GSE2034, GSE12276, GSE2603 and the NKI295 (*microarray-pubs.stanford.edu/wound_NKI/Clinical_Data_Supplement.xls*) captured from public domain was used [19]. A subset of tumors (n=376) with clinically proven metastatic disease was utilized for the analyses [19].

Microarray data analysis

All data were analyzed using the R2 (Microarray Analysis and Visualization Platform) web application, which is publicly available at http://r2.amc.nl. The tumors were also designated to have a “good prognosis” or a “poor prognosis” profile based on the 70-gene prognostic signature as described previously [16].

To validate already published gene expression signatures for lung metastasis [10, 11] and brain metastasis [8], the indicated genes were mapped to Illumina platform via Gene Symbol ID. As previously described [16], respectively a K-means method was used to cluster the patients in 2 groups and a t-test revealed the performance of these signatures in our dataset.

Identification and validation of site-specific metastasis signature

To identify a gene expression signature associated with organ specific metastasis, we used the one-way ANOVA function in R2. Only the genes with an expression level above background level were included in the analysis (total 16051 genes). Samples were split into 2 groups; one group in which the patient developed a metastasis and another group in which a patient never developed a certain organ metastasis. The genes that showed a significant differential expression between these groups (*p* value, 0.001) were included in the signature. The metastatic signature was subsequently validated

in multiple datasets using the K-means and t test function in R2. To further investigate the association between this gene signature and clinical variables multivariate logistic regression tests were applied using SPSS Statistics for Windows (Release version 21.0;IBM Corp.2012, Armond, NY). Overall survival and metastasis free survival were analyzed by the Kaplan-Meier method in the training data set. Due to missing survival data in the publicly available files, additional survival analyses were not conducted in the independent dataset. All statistical tests were two sided and $p < 0.05$ was considered to be statistically significant.

Results

From 157 primary invasive breast cancer from patients who developed distant metastases during follow-up, mRNA expression signatures were assessed using micro array analysis. The patient and tumor characteristics have been described previously [16].

Tumors were subdivided into molecular subtypes with the help of the PAM50 classifier [20]. The distribution of metastatic behavior including site of metastasis, metastasis timeline and survival outcomes among the molecular subtypes have been published previously and are summarized in Table 1. 79.4% Of the patients with Luminal A, 72.5% of Luminal B, 78.6% HER2-like and 87.5% of basal-type tumors received adjuvant systemic therapy. None of the patients received trastuzumab as adjuvant therapy; a subgroup of patients ($n = 10$) received trastuzumab for treatment of metastatic disease. The tumors were also subdivided in “good prognosis signature” and “poor prognosis signature” based on their 70-gene expression profile.

Validation of previously identified gene signature(s) for lung and brain specific metastasis

We have investigated three previously published gene signatures for lung [10, 11] and brain metastasis [8].

The gene expression signature (consisting of 54 genes) predicting lung metastases identified by *Minn* et al. [11] identified 55 primary tumors as having a "lung metastasis" gene expression signature. Out of the 55 tumors which positively tested for this signature, 17 (30.9%) were found to have developed lung metastases. Of negatively tested primary tumors 61 (63.5%) had no lung metastasis (p 0.594, sensitivity: 32.7% and specificity: 61.6%). When separated according to ER status, of 29 positively tested ER-positive tumors 6 (35.3%) and of 26 positively tested ER-negative tumors 11 (42.3%) developed metastatic disease to the lung (p 0.165 and p 0.368, respectively). These results show that the 54 gene lung metastasis signature did not predict the development of lung metastases in our patient series. 70.8% of the basal type tumors, 44.4% of the HER2-like tumors and 27.7% of the luminal type tumors was positive for the lung metastasis associated gene expression signature.

The six-gene expression lung metastasis associated signature of *Landemaine* et al [10] was present in 23 tumors. Nine (39.1%) of these patients with positively tested tumors had lung metastasis, whereas of 128 patients with negatively tested tumors 85 (66.5%) had no metastatic disease to lung (p 0.638, sensitivity: 17.3% and specificity: 85.9%).

87.5% of the basal-type tumors and 0.9% (*n* = 1) of the luminal-type tumors was positive for the signature. None of the HER2-like tumors were positive. The 17-gene expression signature of *Bos* et al. for brain specific metastasis was tested as positive in 56 tumors. Sixteen (28.6%) of the patients with positively tested tumors developed brain metastases and 79 (83.2%) of the patients with negatively tested tumors did not develop brain metastases (p 0.102, sensitivity: 50% and specificity: 66.4%). When the tumors which were positively tested for 17-gene signature [8] grouped according to ER-status, 11 (44%, *n* = 25) of ER-positive tumors and 12 (38.7%, *n* = 31) of ER-negative tumors had brain metastasis (p 0.795 and p 0.74, respectively). All basal-type tumors, 33.3% of the HER2-like tumors and 24.1% of the luminal type tumors were positive for the brain metastasis associated gene expression signature.

Supervised classification of visceral organ metastasis related genes

Subsequently, we have performed supervised classification comparing 112 tumors from patients who developed visceral metastases and 39 tumors from patients who did not develop visceral metastases. Using this approach,14 differentially expressed genes

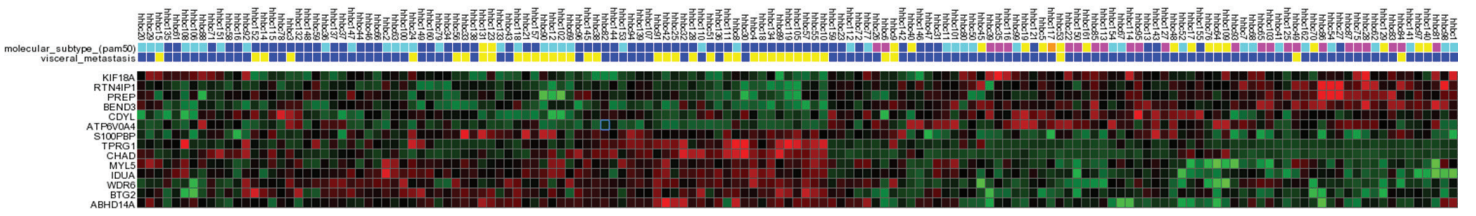


Figure 1. The gene expression pattern of 14 genes of visceral metastasis gene signature.

Heat map displays the gene expression profiling pattern of the 14 differentially expressed genes among 151 tumors. Primary tumors of the patients who developed visceral metastasis are illustrated in blue and the ones without visceral metastatic disease are in yellow. For each primary tumor the expression level of the specific gene is exhibited as red, if up-regulated and green, if down-regulated. Molecular subtypes of primary tumors are also demonstrated as; dark blue for luminal A, light blue for luminal B, yellow for HER2-like, pink for basal type and orange for normal-like tumors.

Table 1. Distribution of metastatic behavior among molecular subtypes

		bone metastasis		visceral metastasis		metastasis timeline*		time to develop metastasis**	time to last event***
		yes	no	yes	no	early	late		
Molecular subtype	Luminal A	54	11	41	24	48	18	45	37
	Luminal B	33	10	35	8	33	10	41	40
	HER2 like	10	8	14	4	15	3	35	21
	Basal type	10	14	21	3	20	4	26	21
	Normal like	1	0	1	0	1	0	10	1
		p .001		p .090		p .749		p .052	p .001

* *cut-off point, 5 years.*
***time from surgery date to the development of first metastasis in months.*
****time from development of first metastasis to the last event (death/last follow-up) in months.*

Table 2. The list of differentially expressed genes in visceral metastatic disease

	Accession number	HUGO	Description	R-value	p-value	Level of expression*
1	ILMN_1669484	WDR6	WD repeat domain 6	-0,319	6,50E-05	<
2	ILMN_1678075	CDYL	Chromodomain protein, Y-like	0,316	7,84E-05	>
3	ILMN_1678186	ATP6V0A4	ATPase, H+ transporting, lysosomal V0 subunit a4	0,318	6,75E-05	>
4	ILMN_1700652	CHAD	Chondroadherin	-0,319	6,42E-05	<
5	ILMN_1703041	IDUA	Iduronidase, alpha-L-	-0,325	4,76E-05	<
6	ILMN_1746948	MYL5	Myosin, light chain 5, regulatory	-0,313	8,97E-05	<
7	ILMN_1751887	PREP	Prolyl endopeptidase	0,357	6,99E-06	>
8	ILMN_1758827	RTN4IP1	Reticulon 4 interacting protein 1	0,314	8,84E-05	>
9	ILMN_1770085	BTG2	BTG family, member 2	-0,389	7,83E-07	<
10	ILMN_1790350	TPRG1	Tumor protein p63 regulated 1	-0,336	2,41E-05	<
11	ILMN_1794213	ABHD14A	Abhydrolase domain containing 14A	-0,321	5,86E-05	<
12	ILMN_2132161	KIF18A	Kinesin family member 18A	0,341	1,81E-05	>
13	ILMN_2294274	S100PBP	S100P binding protein	-0,424	5,87E-08	<
14	ILMN_2375032	BEND3	BEN domain containing 3	0,316	7,91E-05	>

* >, *up-regulated*; <, *down-regulated*

were identified. The 14 genes included in the visceral organ specific gene expression signature were *WDR6*, *CDYL*, *ATP6V0A4*, *CHAD*, *IDUA*, *MYL5*, *PREP*, *RTN4IP1*, *BTG2*, *TPRG1*, *ABHD14A*, *KIF18A*, *S100PBP* and *BEND3* (Table 2). Figure 1 illustrates the heat map with gene expression pattern of these 14 genes in all tumors. Six of these genes, *CDYL*, *ATP6V0A4*, *PREP*, *RTN4IP1*, *BEND3* and *KIF18A* were up-regulated and the other genes were down-regulated. None of these genes overlapped with the genes included in the already published gene expression signatures for lung and brain metastasis [8, 10, 11]. Mapping to the Gene Ontology and Kyoto Encyclopaedia of Genes and Genomes databases revealed that five of these genes (ABDHD14A, IDUA, ATP6V0A4, PREP and KIF18A) were involved in hydrolase activity.

This 14-gene expression signature for visceral metastasis was subsequently validated in the training and the independent datasets. This novel signature was found to be positive in 72 primary tumors of the patients with metastatic breast carcinoma. Out of the 14-gene expression signature positive 72 patients 68 (94%) had visceral organ metastasis. Of 79 patients which were tested as negative for this signature, 35 (44.3%) did not develop visceral metastatic disease (p 2.13e-08, sensitivity: 60.7% and specificity: 89.7%). Among the group of patients which had only visceral metastasis during the disease course (*n* = 18) 88.9% tested positive for the 14-gene expression signature (p 2.0e-04). Among the ones which had visceral organ metastasis as first site of metastasis (*n* = 68) 70.6% tested positive for the signature (p 3.4e-07).

When tested separately in ER-positive and ER-negative tumor groups, 50% of the ER-positive tumors and 60.5% the ER-negative tumors were assessed as visceral metastasis signature positive. Out of 54 ER+/signature+ tumors 94.4% developed metastatic disease in a visceral organ; of 54 ER+ tumors/signature – tumors 53.7% did not develop visceral metastases (p 3.2e-08, sensitivity: 67.1% and specificity: 90.6%). Of 24 ER-/signature+ tumors 91.7% had visceral organ metastasis and of 19 ER-/signature – tumors, 26.3% did not have a visceral organ metastasis (p 0.211, sensitivity: 61.1% and specificity: 71.4%).

Subsequently, the predictive value of 14-gene expression signature was investigated in an independent data set including 376 primary tumors of patients with metastatic breast carcinoma. Of 271 tumors assessed as visceral metastasis signature positive,

170 (62.7%) developed visceral organ metastases. Out of 105 tumors which were tested as negative, 66 (62.9%) had no evidence of metastatic disease to the visceral organs (p 9.68e-06, sensitivity: 81.3% and specificity: 39.5%). This 14-gene expression signature was also assessed separately in ER-positive and ER- negative tumor groups (*n* = 373, ER status was missing in 3 cases). The 14-gene expression signature was found to be positive in 160 of the 245 ER-positive cases. 50% of these ER+/signature+ tumors developed visceral organ metastasis; of 85 ER+/signature – tumors 63.5% did not develop visceral metastases (p 4.50e-02, sensitivity: 72.1% and specificity: 40.3%). There were 128 ER-negative tumors, 104 of which tested as positive for the visceral specific gene expression signature. 76% of these ER-/signature+ tumors developed visceral organ metastases; of 24 ER-negative tumors which were found to be negative for this signature 33.3% had no evidence of visceral organ metastasis (p 4.37e-01, sensitivity: 83.2% and specificity: 24.2%).

Table 3 summarizes the performance of visceral metastasis specific gene expression signature in the data sets described.

Univariate analyses in the training dataset revealed that next to development of visceral organ metastasis (ever, as first site of metastasis and as the only metastasis) the 14-gene expression signature was also found to be significantly correlated to the histologic subtype of the tumor, ER status, PR status and molecular subtype of the primary tumor (p 0.003, $p < 0.001$, $p < 0.001$, $p < 0.001$ and $p =< 0.001$, respectively). The other parameters that were associated with development of visceral organ metastases were tumor size (p 0.002) tumor grade (p 0.009) and tumor type (p 0.008). Multivariate analyses showed that along with tumor type the 14-gene expression signature was remained significantly correlated to visceral organ metastasis (p 0.001, 95% CI 1.43-4.27).

Similarly, univariate analyses in the independent dataset showed that the 14-gene expression signature was significantly correlated with the development of metastatic disease to visceral organs ($p < 0.001$). Molecular subtype of the tumor, ER status, PR status and lymph node status were the other parameters that were statistically related to visceral metastasis in this data set. Multivariate analysis showed that hormone receptor status (ER and HER2) remained significantly correlated to the development of

Table 3. Performance of the gene expression signature

Gene expression signature			Training data set			Independent data set		
						Visceral metastasis		
		Signature	yes	no	p	yes	No	p
14-gene expression signature[a]	All	present	68	4	2.13e-08	170	101	9.68e-06
		absent	44	35		39	66	
	ER-positive	present	51	3	3.2e-08	80	80	4.50e-02
		absent	25	29		31	54	
	ER-negative	present	22	2	0.211	79	25	4.37e-01
		absent	14	5		16	8	

[a] *The 14-gene expression signature developed in this study.*
ER estrogen receptor.

visceral metastases (p 0.047 , 95% CI -3.36-0.003, ER status; p 0.025, 95% CI 0.2-3.2) whereas the 14-gene expression signature was not retained as a significant predictor (p 0.49, 95% CI -0.97-1.9).

Additional survival analyses in the training dataset exhibited that the 14-gene expression signature was associated with survival status of the patients, indicated by metastasis free survival and overall survival (p 0.001 and $p < 0.001$, respectively).

Discussion

Development of visceral organ metastases in breast carcinoma is related to dismal prognosis with poor overall and metastasis free survival rates [17, 21]. The identification of genomic tumor characteristics associated with a higher likelihood of developing visceral organ metastasis will help understanding the mechanisms leading to the development of visceral metastases. In this study, we have compared the gene expression profiles of primary tumors of breast cancer patients who developed visceral organ metastasis to the ones without visceral metastasis using gene-expression microarrays. We have identified a unique group of genes which were differentially expressed in the group of tumors with clinical metastatic disease to the visceral organs. This association between 14-gene expression signature was significant not only with development of visceral organ metastasis (any time, as first site and as only site of metastasis) but also with both overall survival and metastasis free survival.

The identified gene expression signature included 14 genes, five (CDYL, *ATP6V0A4*, *PREP*, *RTN41P1*, *BEND3* and *Kif18A*) of which were up-regulated in the group of primary tumors of the patients with visceral organ metastasis. Two of these genes (*Kif18a* and *ATP6V0A4*) have been already reported to be associated with human breast carcinogenesis [22-25]. *Kif18a*, kinesin family number 18a, which has function to produce force and movement along microtubules, was previously found to be deregulated in different cancers including breast cancer [25]. It has been shown that overexpression of *Kif18a* is associated with tumor grade, development of metastasis and poor survival. Functional analyses have also shown that ablation of this protein

results in inhibition of proliferative capability of breast cancer cells with inactivation of phosphatidylinositol 3-kinase-Akt signalling pathway [24]. *Zou* et al. have shown that knockdown of kinesin gene family members strongly disrupted the proliferation and induced the apoptosis in both tamoxifen-sensitive and resistant breast cancer cells and they have suggested the potential role of developing novel inhibitors of the kinesins for effective treatment of human cancers including tamoxifen-resistant breast cancer [25]. *ATP6VOA4*, vacuolar ATPase, H+ transporting, lysosomal V0 subunit a4, was another gene that found to be up-regulated in the group of tumors from patients with visceral metastatic disease. Previous studies have suggested an association between vacuolar ATPases (V-ATPase) and tumor invasion [26, 27]. Also, specific V-ATPase inhibitors, such as bafilomycin and concanamycin, have been shown to inhibit invasiveness of MDA-MB231 breast cancer cell lines [22, 23]. Subunit isoforms a3 and a4 are expressed at high levels in highly invasive breast cancer cell lines (MB231). It has been speculated that isoform a4 (*ATP6VOA4*) is involved in targeting V-ATPAse in cell membrane and this V-ATPase plays a role in invasive capability of these cells. This effect may be caused by locally acidifying the extracellular environment which may accelerate tumor invasion via creating an optimal acidic environment for proteases [28]. Further clinical studies investigating the effect of V-ATPase inhibitors on tumor invasiveness and metastasis will be of interest.

Nine out of 14 genes were downregulated in the subgroup of tumors with visceral organ metastasis. One of these downregulated genes is BTG2 (B-cell translocation gene-2), which has antiproliferative activity and has been reported to be altered in breast tumors [7, 29, 30]. It has also been shown that decrease of BTG2 expression in human breast cancer correlates with disease progression [31]. In order to explore the underlying mechanism, *Takahashi* and colleagues have further implemented experimental studies showing that knockdown of BTG2 expression led to increased cell motility. They have also demonstrated that BTG2 suppresses the activation of the HER2 pathway and suggested that HER pathway inhibitors, such as lapatinib, may play a role in controlling the progression of disease among breast cancers with decreased BTG2 expression. Moreover, the same group has demonstrated a modulator role of BTG2 on tamoxifen responsiveness in ER-positive/HER2-negative breast cancer and further validated BTG2 expression as a single predictor of survival following tamoxifen therapy [32]. Likewise, protein expression levels of BTG2 have also been found to be associated with

5-year overall survival in breast cancer patients. A prognostic model combining BTG2 expression, HER2 expression, patient age and Ki67 expression has been proposed with higher prediction accuracy than the currently used prognostic markers [33]. Consistent with the published data, we have shown that lower BTG2 expression was associated with shorter survival time (p 2.6e-04). Luminal type tumors had higher gene expression levels of BTG2 compared to HER2-like and basal type tumors (p 1.6e-10) and our current study has revealed a strong correlation between BTG2 expression and visceral metastasis (p 2.13e-08).

In conclusion, we present a unique 14-gene expression signature for visceral metastasis in breast carcinomas. Further validation of this gene expression signature is warranted in order to test the reproducibility and the robustness of the correlations between the signature and metastatic behavior.

References

1. Cummings MC, Simpson PT, Reid LE, Jayanthan J, Skerman J, Song S, McCart Reed AE, Kutasovic JR, Morey AL, Marquart L *et al*: Metastatic progression of breast cancer: insights from 50 years of autopsies. *J Pathol* 2014, 232(1):23-31.

2. Sant M, Allemani C, Berrino F, Coleman MP, Aareleid T, Chaplain G, Coebergh JW, Colonna M, Crosignani P, Danzon A *et al*: Breast carcinoma survival in Europe and the United States. *Cancer* 2004, 100(4):715-722.

3. Thomssen C, Diel I, Gnant M, Goldmann-Posch U, Goschke C, Haidinger R, Janni W, Lin NU: Follow-up after breast cancer diagnosis. *Breast Care (Basel)* 2013, 8(6):457-460.

4. Paget S: The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev* 1989, 8(2):98-101.

5. Nuyten DS, van de Vijver MJ: Gene expression signatures to predict the development of metastasis in breast cancer. *Breast Dis* 2006, 26:149-156.

6. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T *et al*: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004, 351(27):2817-2826.

7. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ *et al*: A gene-expression signature as a predictor of survival in breast cancer. *NEnglJMed* 2002, 347(25):1999-2009.

8. Bos PD, Zhang XH, Nadal C, Shu W, Gomis RR, Nguyen DX, Minn AJ, van de Vijver MJ, Gerald WL, Foekens JA *et al*: Genes that mediate breast cancer metastasis to the brain. *Nature* 2009, 459(7249):1005-1009.

9. Kang Y, He W, Tulley S, Gupta GP, Serganova I, Chen CR, Manova-Todorova K, Blasberg R, Gerald WL, Massague J: Breast cancer bone metastasis mediated by the Smad tumor suppressor pathway. *Proceedings of the National Academy of Sciences of the United States of America* 2005, 102(39):13909-13914.

10. Landemaine T, Jackson A, Bellahcene A, Rucci N, Sin S, Abad BM, Sierra A, Boudinet A, Guinebretiere JM, Ricevuto E *et al*: A six-gene signature predicting breast cancer lung metastasis. *Cancer Res* 2008, 68(15):6092-6099.

11. Minn AJ, Gupta GP, Padua D, Bos P, Nguyen DX, Nuyten D, Kreike B, Zhang Y, Wang Y, Ishwaran H *et al*: Lung metastasis genes couple breast tumor size and metastatic spread. *Proceedings of the National Academy of Sciences of the United States of America* 2007, 104(16):6740-6745.

12. Gupta GP, Massague J: Cancer metastasis: building a framework. *Cell* 2006, 127(4):679-695.

13. Ramaswamy S, Ross KN, Lander ES, Golub TR: A molecular signature of metastasis in primary solid tumors. *Nat Genet* 2003, 33(1):49-54.

14. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT *et al*: Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002, 415(6871):530-536.

15. Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ: Gene expression profiles of primary breast tumors maintained in distant metastases. *Proceedings of the National Academy of Sciences of the United States of America* 2003, 100(26):15901-15905.

16. Savci-Heijink CD, Halfwerk H, Koster J, van de Vijver MJ: A novel gene expression signature for bone metastasis in breast carcinomas. *Breast cancer research and treatment* 2016, 156(2):249-259.

17. Savci-Heijink CD, Halfwerk H, Hooijer GK, Horlings HM, Wesseling J, van de Vijver MJ: Retrospective analysis of metastatic behaviour of breast cancer subtypes. *Breast cancer research and treatment* 2015, 150(3):547-557.

18. Vermeulen E, Geesink I, Schmidt MK, Steegers C, Verhue D, Brom FW, Aaronson NK, van Leeuwen FE: [Secondary use of human tissue: consent and better information required]. *NedTijdschrGeneeskd* 2009, 153:A948.

19. Harrell JC, Prat A, Parker JS, Fan C, He X, Carey L, Anders C, Ewend M, Perou CM: Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer ResTreat* 2012, 132(2):523-535.

20. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z *et al*: Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009, 27(8):1160-1167.

21. Solomayer EF, Diel IJ, Meyberg GC, Gollan C, Bastert G: Metastatic breast cancer: clinical course, prognosis and therapy related to the first site of metastasis. *Breast cancer research and treatment* 2000, 59(3):271-278.

22. Rojas JD, Sennoune SR, Maiti D, Bakunts K, Reuveni M, Sanka SC, Martinez GM, Seftor EA, Meininger CJ, Wu G *et al*: Vacuolar-type H+-ATPases at the plasma membrane regulate pH and cell migration in microvascular endothelial cells. *Am J Physiol Heart Circ Physiol* 2006, 291(3):H1147-1157.

23. Sennoune SR, Bakunts K, Martinez GM, Chua-Tuan JL, Kebir Y, Attaya MN, Martinez-Zaguilan R: Vacuolar H+-ATPase in human breast cancer cells with distinct metastatic potential: distribution and functional activity. *Am J Physiol Cell Physiol* 2004, 286(6):C1443-1452.

24. Zhang C, Zhu C, Chen H, Li L, Guo L, Jiang W, Lu SH: Kif18A is involved in human breast carcinogenesis. *Carcinogenesis* 2010, 31(9):1676-1684.

25. Zou JX, Duan Z, Wang J, Sokolov A, Xu J, Chen CZ, Li JJ, Chen HW: Kinesin family deregulation coordinated by bromodomain protein ANCCA and histone methyltransferase MLL for breast cancer cell growth, survival, and tamoxifen resistance. *Mol Cancer Res* 2014, 12(4):539-549.

26. Forgac M: Vacuolar ATPases: rotary proton pumps in physiology and pathophysiology. *Nat Rev Mol Cell Biol* 2007, 8(11):917-929.

27. Ohta T, Numata M, Yagishita H, Futagami F, Tsukioka Y, Kitagawa H, Kayahara M, Nagakawa T, Miyazaki I, Yamamoto M *et al*: Expression of 16 kDa proteolipid of vacuolar-type H(+)-ATPase in human pancreatic cancer. *British journal of cancer* 1996, 73(12):1511-1517.

28. Hinton A, Sennoune SR, Bond S, Fang M, Reuveni M, Sahagian GG, Jay D, Martinez-Zaguilan R, Forgac M: Function of a subunit isoforms of the V-ATPase in pH homeostasis and in vitro invasion of MDA-MB231 human breast cancer cells. *J Biol Chem* 2009, 284(24):16400-16408.

29. Duriez C, Falette N, Audoynaud C, Moyret-Lalle C, Bensaad K, Courtois S, Wang Q, Soussi T, Puisieux A: The human BTG2/TIS21/PC3 gene: genomic structure, transcriptional regulation and evaluation as a candidate tumor suppressor gene. *Gene* 2002, 282(1-2):207-214.

30. Mollerstrom E, Delle U, Danielsson A, Parris T, Olsson B, Karlsson P, Helou K: High-resolution genomic profiling to predict 10-year overall survival in node-negative breast cancer. *Cancer Genet Cytogenet* 2010, 198(2):79-89.

31. Takahashi F, Chiba N, Tajima K, Hayashida T, Shimada T, Takahashi M, Moriyama H, Brachtel E, Edelman EJ, Ramaswamy S *et al*: Breast tumor progression induced by loss of BTG2 expression is inhibited by targeted therapy with the ErbB/HER inhibitor lapatinib. *Oncogene* 2011, 30(27):3084-3095.

32. Takahashi M, Hayashida T, Okazaki H, Miyao K, Jinno H, Kitagawa Y: Loss of B-cell translocation gene 2 expression in estrogen receptor-positive breast cancer predicts tamoxifen resistance. *Cancer Sci* 2014, 105(6):675-682.

33. Mollerstrom E, Kovacs A, Lovgren K, Nemes S, Delle U, Danielsson A, Parris T, Brennan DJ, Jirstrom K, Karlsson P *et al*: Up-regulation of cell cycle arrest protein BTG2 correlates with increased overall survival in breast cancer, as detected by immunohistochemistry using tissue microarray. *BMC Cancer* 2010, 10:296.

5

Chapter 6

Association between gene expression profile of the primary tumor and chemotherapy response of metastatic breast cancer

C. Dilara Savci Heijink, Hans Halfwerk, Jan Koster, Marc J. van de Vijver

BMC Cancer. 2017 Nov 13;17(1):755

Abstract

To better predict the likelihood of response to chemotherapy, we have conducted a study comparing the gene expression patterns of primary tumors with their corresponding response to systemic chemotherapy in the metastatic setting. mRNA expression profiles of breast carcinomas of patients that later developed distant metastases were analyzed using supervised and non-supervised classification techniques to identify predictors of response to chemotherapy. The top differentially expressed genes between the responders and non-responders were identified and further explored. An independent dataset which was generated to predict response to neo-adjuvant CT was utilized for the purpose of validation. Response to chemotherapy was also correlated to the clinicopathologic characteristics, molecular subtypes, metastatic behavior and survival outcomes. Anthracycline containing regimens were the most common first line treatment (58.4%), followed by non-anthracycline/non-taxane containing (25.8%) and taxane containing (15.7%) regimens. Response was achieved in 41.6% of the patients to the first line CT and in 21.8% to second line CT. Response was not found to be significantly correlated to tumor type, grade, lymph node status, ER and PR status. Patients with HER2+ tumors showed better response to anthracycline containing therapy (p 0.002). Response to first and second line chemotherapy did not differ among gene expression based molecular subtypes (p 0.236 and p 0.20). Using supervised classification, a 14 gene response classifier was identified. This 14-gene predictor could successfully predict the likelihood of better response to first and second line CT ($p < 0.0001$ and p 0.761, respectively) in the training set. However, the predictive value of this gene set in data of response to neoadjuvant chemotherapy could not be validated. To our knowledge, this is the first study revealing the relation between gene expression profiles of the primary tumors and their chemotherapy responsiveness in the metastatic setting. In contrast to the findings for neoadjuvant chemotherapy treatment, there was no association of molecular subtype with response to chemotherapy in the metastatic setting. Using supervised classification, we identified a classifier of chemotherapy response; however, we could not validate this classifier using neoadjuvant response data.

Introduction

The main aim of treating metastatic breast cancer is to prolong survival of the patients with acceptable toxicity and to palliate the disease-related symptoms. Response to combined chemotherapy agents varies between 50-70% in the metastatic setting [1, 2]. In order to avoid unnecessary chemotherapy treatment it would be of great benefit to be able to distinguish the group of patients which are not likely to respond to chemotherapy in general and to specific chemotherapy regimens. The decision to treat patients with metastatic breast cancer with chemotherapy is usually taken depending on many factors such as patient age and performance status, site of metastasis, hormone receptor status and prior exposure to chemotherapy, [3, 4]. Commonly used first-line therapeutic options in the metastatic setting include anthracycline- and/or taxane-based regimens [5]. In case of disease progression other cytotoxic agents may be applied to maximize the duration of quality time for these patients [6].

The current treatment approaches for metastatic disease consist to a large extent of trial-and-error type models, as predictors of response are lacking. The response rate to the chemotherapy regimens and the median duration of survival differs between breast cancer subtypes [7-11]. Several gene expression profiling studies aimed at the identification of a genomic predictor of chemotherapy response in the neoadjuvant setting have been performed and already provided important insights [12-17]. However a clinically validated gene expression profiling assay to predict the chemotherapy response has not yet been accomplished. Gene expression profiling studies of chemotherapy response in metastatic breast cancer have thus far been lacking.

We have previously investigated the association between the gene expression patterns of primary tumors and metastatic behavior in metastatic breast cancer [18].

In the current study, using the gene expression profiling data of 89 patients, the link between primary tumor and chemotherapy response in the frame of metastatic disease is explored. In order to develop genomic identifiers of chemotherapy responsiveness, gene expression patterns of the primary tumors of the responders and non-responders have been investigated.

Material and methods

Patients and tumor samples

Metastatic breast cancer patients from the *Academic Medical Center* and *Netherlands Cancer Institute (NCI)* were identified (*n* = 263) and a subgroup of patients from whom frozen tumor material from the primary tumor was available, were included in this study. This group constituted of 118 patients whose primary tumors were diagnosed between 1984 and 2000. The study protocol was approved by the Medical Ethical Committee of the *Academic Medical Center* and permission to use the data of the patients from *Netherlands Cancer Institute* was granted by the Core Facility-Molecular Pathology and Biobanking. Relevant clinical data and detailed information on metastatic behavior were abstracted from the clinical charts. Information related to metastatic behavior included data on site of metastasis (ever/never, first/not first and only/not-only for each metastasis site), metastasis pattern (uni/multiple) and metastasis timeline (early/late) has been previously published [11]. Time to develop metastatic disease, time from development of metastatic disease (metastasis-specific survival, MSS) to last event and overall survival (OS) were recorded. Last event date was defined as the most recent follow up date for the patients who were alive and time of death for the others.

Histopathologic examination of the sections from the primary tumors was performed by two pathologists (C.D.S-H and M.J.V.) and as needed immunohistochemical stains and in-situ hybridization were applied in order to determine the hormone receptor and HER2 status as previously described [11].

Chemotherapy data

For each patient, administered systemic therapy was recorded for the adjuvant and metastatic settings separately. The therapy given was grouped as hormonal therapy (HT) and chemotherapy (CT). In addition, the type of the therapeutic agent, the duration and the chronology of the therapy were noted. Due to the heterogeneity of the chemotherapy regimens, we have grouped the chemotherapy regimens into 3 groups as : anthracycline containing, taxane containing and non-anthracycline/non-taxane containing. Response to chemotherapy in metastatic patients was assessed for each line of chemotherapy according to RECIST [19] criteria and classified as complete response

(CR), partial response (PR), stable disease (SD) and progressive disease. For statistical purposes, CR and PR were considered as response and SD and PD were considered as non-response. Response to first and second line CT and each chemotherapeutic group was separately assessed. Response to the given chemotherapy group were scored as response in case of response as first line treatment.

Gene expression profiling

The gene expression profiling experiments have been described previously and detailed information on RNA amplification, labeling and hybridization can be found at Illumina website (http://www.illumuna.com) [18].The gene expression data was normalized utilizing robust spline normalization (rsn) and log2 transformed and followed by ComBat (http://www.bu.edu/jlab/wp-assets/ComBat/Abstract.html). Data analyses were conducted using R2 (Microarray Analysis and Visualization Platform), a publicly available web application (http://r2.amc.nl).

For each tumor, the previously assessed 70-gene prognostic signature [20] was used to categorize tumors as good prognosis or poor prognosis signature. Genes were mapped to the Illumina platform via Gene Symbol ID. 62 Genes were found be present on the Illumina platform corresponding to 65 probes. The probe with the highest variance across the samples was selected in the event of existence of multiple probes for one gene. Tumors were assigned into the good or poor prognostic group based on the Pearson correlation coefficient between the centroids of the original good prognosis template and the gene expression levels of each sample. Classification into molecular subtypes (basal type, HER2 like, luminal A and luminal B type) were done using the genes from the PAM50 classifier [21]. The 21-gene recurrence score for each tumor was calculated as described by King et al. [22].

Identification and validation of predictors for chemotherapy response

To identify a gene expression predictor associated to response to chemotherapy, we used the one-way ANOVA function in R2 to select from a set of 15526 genes with an expression level above background. 14 genes had a significant different expression ($p < 0.001$) between the group which the patient had a tumor response (CR and PR) to first line chemotherapy and the group in which the patient had no tumor response

(SD and PD) to chemotherapy were identified. For validation, there are no published datasets for patients with metastatic disease; there are, however, various datasets of gene expression profiles of tumors from patients who underwent neoadjuvant chemotherapy treatment. Therefore, the predictive chemotherapy signature was then validated in an independent data set using the *K*-means and *t* test function in R2. A data set (GSE25066) which includes 488 breast carcinomas with response data in the neoadjuvant setting was used for validation [12] .

To further investigate the association between this 14-gene predictor and clinical variables including response to chemotherapy multivariate logistic regression tests were applied using SPSS Statistics for Windows (Release version 21.0; IBM Corp.2012, Armond, NY). All statistical tests were two sided and $p < 0.05$ was considered to be statistically significant.

Results

Gene expression profiles from primary tumors ($n = 118$) were assessed using microarrays. All patients were known to have developed distant metastasis and underwent (palliative) chemotherapy. The clinicopathologic features of the patients are displayed in Table 1. The mean age at diagnosis was 50.77 years (range 28 to 85 years). Median follow-up time was 63 months (range 9 to 211 months) for all patients and 136.50 months (range 74 to 208 months) for the patients who were alive at last follow-up. In this study group, 17.2% ($n = 21$) previously received neo-adjuvant systemic therapy and 80.4% ($n = 98$) adjuvant systemic therapy as part of the treatment of the primary tumor. Out of 98 patients who were given adjuvant therapy 39.8% ($n = 39$) received only chemotherapy, 15.3% ($n = 15$) received only hormonal therapy and 44.9% ($n = 44$) received chemotherapy and hormonal therapy. Adjuvant chemotherapy consisted of anthracycline containing regimens for 56.3%, non-anthracycline/non-taxane containing regimens for 42.3% and taxane containing drugs for 1.4% of the patients who received adjuvant chemotherapy ($n = 71$). None of the patients received trastuzumab as adjuvant treatment.

In the metastatic setting all patients (*n* = 118) received palliative systemic therapy. Of these patients 49.2% (*n* = 58) received chemotherapy and hormonal therapy , 27.1% (*n* = 32) received only chemotherapy and 23.7% (*n* = 28) only hormonal therapy in the course of metastatic disease. The chemotherapeutic agents given in the metastatic setting were quite heterogeneous. As first line chemotherapy, 58.4% (*n* = 52) received an anthracycline containing regimen, 25.84% (*n* = 23) an non-anthracycline/non-taxane containing regimen and 15.73% (*n* = 14) received a taxane containing regimen (total *n* = 89). Second line CT was given to 63 patients and consisted of an anthracycline containing regimen for 22.2% (*n* = 14), a non-anthracycline/non-taxane containing regimen for 36.5% (*n* = 23) and a taxane containing regimen for 41.3% (*n* = 26) patients. Ten patients received a trastuzumab containing regimen as first line therapy.

The response rate for the first and second line chemotherapy was 41.6% and 21.8%, respectively. Patients who received anthracycline containing therapy showed a response rate of 51.8%, patients who received non-taxane/non-anthracycline containing therapy showed a response rate of 24.3% and the ones who were given taxane containing therapy had a response rate of 30.6%. Table 2 shows the distribution of the administered chemotherapy and response rates among patients. Response to chemotherapy was not found to be significantly correlated with histologic type, tumor grade and lymph node status. Response to first line chemotherapy treatment was better among patients who were younger than 50 years (*p* 0.005).

ER and PR status were not associated with response to chemotherapy treatment, whereas HER2 positive patients showed better response rate to anthracycline containing regimens (*p* 0.002). Out of 13 HER2 positive patients with good response to anthracycline containing regimen, only 3 patients received trastuzumab for the treatment of metastatic disease (23.1%).

When classified into molecular subtypes 95 of the tumors classified as luminal (59, luminal A ; 36, luminal B), 16 tumors as basal, 10 tumor as HER2-like and one tumor as normal like subtype. Out of luminal type tumors 65 and 49; of basal type tumors 16 and 8, of HER2-like tumors 9 and 5 received first and second line chemotherapy respectively. Response to first and second line chemotherapy did not differ among the

Table 1. Clinicopathologic characteristics of the primary tumors

		N	%
Age at diagnosis, years	<50	68	55.7
	>50	54	44.3
Histology	Ductal	105	86.8
	Lobular	12	9.9
	Other	4	3.3
Tumor grade	1	9	7.6
	2	69	58.5
	3	40	33.9
Lymph node status	none	35	30.2
	1-3 positive	35	30.2
	>3 positive	46	39.7
Neoadjuvant chemotherapy	no	100	82.6
	yes	21	17.4
Adjuvant therapy	none	23	19.0
	only CT	39	32.2
	only HT	15	12.4
	CT+HT	44	36.4

CT chemotherapy, HT hormonal therapy.

molecular subtypes (p 0.236 and p 0.20). Molecular subtypes and their corresponding metastatic behavior have already been published [18, 23].

Analyses were further carried out based on specific chemotherapy regimen. Anthracycline containing therapy was given to 41 patients with luminal type tumors, 7 patients with HER2-like tumors and 8 patients with basal-type tumors as first or second line CT in the metastatic setting. Among these patients 51.21% of the patients with luminal type tumors, 71.4% of the patients with HER2-like tumors and 37.5% of the patients with basal type tumors showed response to anthracycline containing therapy (p 0.624). Non-anthracycline/non-taxane containing therapy was given to 28 patients with luminal tumors, 3 patients with HER-2 type tumors and 6 patients with basal type

Table 2. Distribution of the administered CT and response rates among patients

		Response	N	%
First line chemotherapy	A-CT	no	25	48.1
		yes	27	51.9
	NA/NT-CT	no	14	66.7
		yes	7	33.3
	T-CT	no	11	7.,6
		yes	3	21.4
Second line chemotherapy	A-CT	no	10	71,4
		yes	4	28,6
	NA/NT-CT	no	20	87,
		yes	2	8,7
	T-CT	no	17	68,0
		yes	8	32,0

A-CT anthracycline containing chemotherapy, NA/NT-CT non-anthracyline/non-taxane containing chemotherapy, T-CT taxane containing chemotherapy.

tumors. Response rate to non-anthracycline/non-taxane containing regimens was 25%, 33.3% and 16.7% of the patients with luminal, HER2-type and basa- type tumors, respectively (p 0.954).

Taxane containing therapy was administered to 23 patients with luminal type tumors, 3 patients with HER2-type tumors and 10 patients with basal type tumors. Of luminal type tumors 39.1%, of HER2-type tumors 33.3% and of basal-type tumors 10% responded to taxane containing therapy (p 0.033). The association between the molecular subtypes of the tumors and their response status is displayed in Table 3.

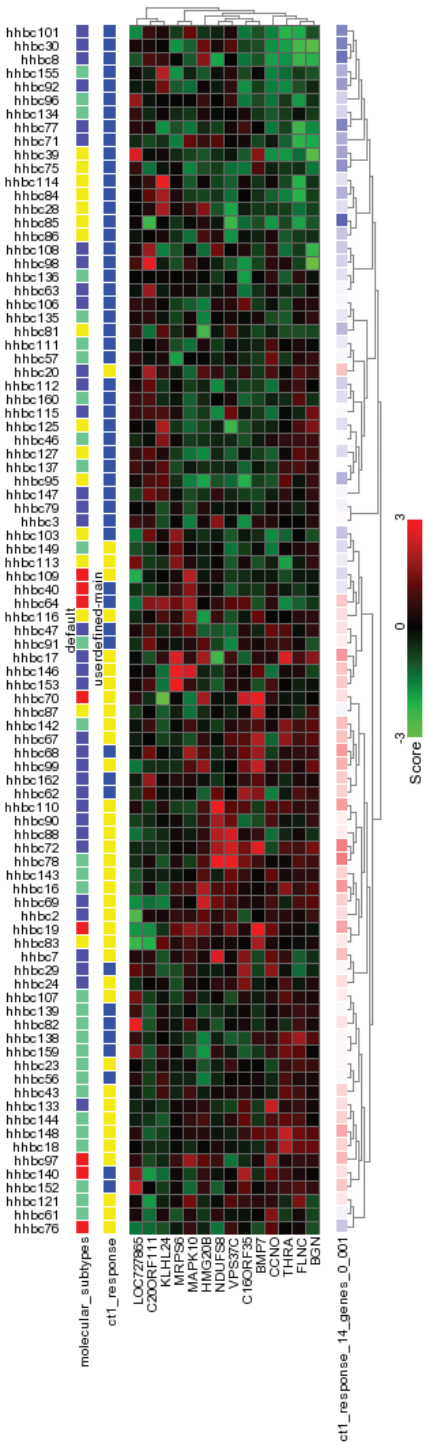


Figure 1. The gene expression pattern of 14-gene predictor for chemotherapy response.

Heat map shows the gene expression profiling pattern of the 14 differentially expressed genes among 89 tumors. Primary tumors of the patients who respond to chemotherapy are illustrated in yellow and the ones without response are in blue. For each primary tumor the expression level of the specific gene is exhibited as red, if up-regulated and green, if down-regulated.

Table 3. The association between the molecular subtypes of the primary tumors and chemotherapy response rates

Response		Molecular subtype							
		Basal		Luminal A		Luminal B		HER2	
		N	%	N	%	N	%	N	%
A-CT	no	5	62,5	12	48,0	8	50,0	2	28,6
	yes	3	37,5	13	52,0	8	50,0	5	71,4
NA/NT-CT	no	5	83,3	12	75,0	9	75,0	2	66,7
	yes	1	16,7	4	25,0	3	25,0	1	33,3
				0		0		0	
T-CT	no	9	90,0	4	36,4	10	83,3	2	66,7
	yes	1	10,0	7	63,6	2	16,7	1	33,3

A-CT anthracycline containing chemotherapy, NA/NT-CT non-anthracyline/non-taxane containing chemotherapy, T-CT taxane containing chemotherapy.

The group of patients who received trastuzumab was composed of 6 with luminal type tumors, 3 with HER2-like tumors an 1 with a basal type tumor. There was no significant association between trastuzumab use and chemotherapy response (p 0.291).

Identification of genomic predictor(s) for chemotherapy response

Using supervised classification the differentially expressed genes between the primary tumors of metastatic breast cancer patients in responders (n = 37) and non-responders for the first line chemotherapy were explored (n = 52). Using supervised classification the top 14 differentially expressed genes between responders and non-responders were selected for further analyses. These 14 differentially expressed genes are listed as *BGN*, *BMP7*, *C16ORF35*, *C20ORF111*, *CCNO*, *FLNC*, *HMG20B*, *KLHL24*, , *LOC727865*, *MAPK10*, *MRPS6*, *NDUFS8*, *THRA* and *VPS37C*. Three of these genes were found to be down-regulated and the rest to be up-regulated in the group of patients with a good response to chemotherapy (Table 4). Figure 1 displays the expression profiling pattern of 14 differentially expressed genes among the patients. This heat map shows that the set of 14 genes separates the responders and non-responders in group of 89 tumors ($p < 0.001$).

Table 4. Performance of the 14-gene predictor for chemotherapy response

			Training data set			Independent data set		
			Chemotherapy response					
		Signature	yes	no	p	yes	no	p
14-gene predictor [a]	All	present	33	10	2.24E-11	47	158	0.254
		absent	4	42		52	231	
	ER-positive	present	23	2	1.02E-10	15	101	0.327
		absent	5	35		15	154	
	ER-negative	present	9	3	3.37E-04	41	63	0.136
		absent	0	12		27	66	

[a] *The 14-gene predictor developed in this study.*
ER estrogen receptor.

The correlation between these 14 differentially expressed genes and chemotherapy response was further explored. In the group of patients who received chemotherapy in the metastatic setting, 43 patients had a tumor with a " chemotherapy responsive " gene expression profile. Out of those 43 patients, 76.7% (*n* = 33) showed good response to first line therapy; whereas out of 46 patients who were predicted to be non-responsive to chemotherapy 91.3% had indeed no response ($p < 0.0001$, sensitivity: 89.2% and specificity: 80.8%) In the case of response to second line CT, 37 were predicted to have good response with the 14-gene predictor and 24.3% (*n* = 9) of these showed good response; out of 27 tumors which were predicted as non-responder 81.5%(*n* = 22) had no response to CT (*p* 0.249, sensitivity: 64.3% and specificity: 44%). However, as this was the set of tumors in which the chemotherapy response signature was identified, validation in an independent dataset is required.

No other gene data set with chemotherapy response data in the metastatic setting was available for the validation of this gene set. Therefore an independent dataset with available chemotherapy response data for neo-adjuvant administered chemotherapy was utilized [12]. This data set included total 488 tumors with available information on chemotherapy response and 205 of these were predicted as responsive with the 14-gene predictor. Of these 205 tumors which were predicted as responsive , 47 (22.9%) showed response to chemotherapy. Out of 283 tumors which were assessed as non-responsive with the predictor, 231 (81.6%) had actually no response to CT (p 0.254, sensitivity: 47.5% and specificity: 59.4%). The validation of this 14-gene predictor is summarized in Table 4.

Other signatures which were developed to predict the response to neoadjuvant chemotherapy were also tested in this study group. DLDA30 signature correctly predicted 76.3% of the responsive and 62% of the non-responsive tumors (p 6,5E-01). In contrast, the genomic grade index (GG1) and genomic predictor of *Hatzis* et al. were not able to distinguish the responsive and non-responsive groups (p 0.317 and p 0.212, respectively). The relationship between the 21-gene recurrence score and CT response in our study set was also further investigated in the subgroup of ER-positive/HER2-negative tumors. Out of 74 tumors 40.5% (*n* = 30) had low-risk, 16.2% (*n* = 12) had intermediate-risk and 43.2% (*n* = 32) had high- risk recurrence scores. A high-risk recurrence scores was found to be correlated with shorter overall survival time and

time to develop metastases (p 0.016 and p 0.033, respectively); but not correlated with survival time after the development of metastatic disease (p 0.117). The recurrence scores were not found to be correlated to chemotherapy response (p 0.854).

Additional analyses to explore the correlation of the 14-gene predictor to the site of metastasis (bone metastasis ever, visceral metastasis ever, bone only metastasis and visceral only metastasis) have not revealed any significant relation (p 0.72, p 0.58, p 0.38 and p 0.80, respectively). Yet it was found that this 14-gene predictor was significantly correlated to time to metastasis (metastasis within 5 year vs later than 5 year), more specifically tumors with present 14-gene expression profile developing metastases at a later time than the others (p 0.021).

Survival analyses revealed no significant association between survival time (overall and metastasis specific) and chemotherapy responsiveness. Survival time also did not differ between patients with a responsive 14-gene predictor and the ones without it.

Discussion

With the purpose of identifying a genomic predictor for response to chemotherapy in metastatic breast cancer, we have compared gene expression profiles of primary breast carcinomas to their response to chemotherapy treatment. We have identified a 14 gene expression profile associated with response to chemotherapy. This gene set was able to successfully predict the group of primary tumors which were more likely to respond to chemotherapy in the training set. We do not have access to a validation cohort of tumors from patients with metastatic breast cancer; therefore, we have studied the predictive value of the 14 gene predictive profile in published series of tumors from patients who underwent neoadjuvant chemotherapy treatment.

Specifically, *Hatzis* et al. have introduced a predictive test for neoadjuvant chemotherapy among patients with HER2-negative tumors. The chemopredictive test algorithm developed by this study was shown to predict the chemosensitivity with positive predictive value of 56% (95% CI, 31%-78%) and absolute risk reduction of 18% (95%

CI, 6%-28%). When compared to the other predictive signatures such as genomic grade index (GG1), PAM50 and DLDA30 [13, 14, 21], the predictive algorithm of *Hatzis* et al. had greater positive predictive value in a validation cohort.

Neoadjuvant chemotherapy is increasingly employed for the treatment of breast cancer and predictors of response to neoadjuvant chemotherapy has been previously studied by several groups. Especially triple negative breast cancer (TNBC), which is characterized by lacking expression of ER, PR and HER-2, has shown to be more sensitive to systemic chemotherapy compared to the non-TNBC group. In particular, pathologic complete remission (pCR) has been reported to be achieved in 21.6-45% of TNBC patients. In contrast, hormone receptor positive tumors have been shown to be associated with very low pCR rates (4.9% -11%) [24-28]. Treatment of patients with HER2-positive tumors with chemotherapy plus HER2 targeted neoadjuvant therapy results in pCR rates of approximately 65% with 37% relative improvement in overall survival and an increase in 10-year overall survival rate from 75.2% to 84% [29-32]. Gene expression based analyses have shown similar results with basal like and HER2-type tumors having better pCR response to neoadjuvant chemotherapy (41.7% - 48.8%), compared to luminal type tumors which have shown to have response rates ranging from 2% to 8.2% [27, 33]. It is also known that, regardless of hormone receptor status and intrinsic subtype of the tumor, patients with residual disease after neoadjuvant chemotherapy have significantly shorter overall and disease free survival than patients who achieve pCR [24-26]. In this study identified chemotherapy response rates in the metastatic setting and their association with molecular subtypes and hormone receptor status differed from the ones in the neoadjuvant setting. Response rates to first line therapy given for metastatic disease was not found to be significantly different between molecular subtypes, i.e. basal like tumors and HER2-type tumors did not show better response rates compared to the luminal type tumors. On the other hand, HER-2 positive tumors were associated with better response which is in agreement with published studies [34, 35].

Recently, the Translational Breast Cancer Research Consortium (TBCRC) has conducted a study to explore the usefulness of the 21-gene recurrence score (RS) in predicting response to therapy among breast cancer patients presenting with Stage IV disease [22]. In the group of 69 patients with ER-positive/HER2-negative tumors, they

have found that both time to first progression (TTP) and 2 year overall survival (OS) time were shorter for the patients with high-risk RS values (≥ 31) and who received first line endocrine therapy. There were no differences by means of TTP and 2-year OS in the group of patients with similarly high-risk RS values who received first-line chemotherapy. Therefore, the 21-gene RS has been suggested as a tool for selection of the patients presenting with stage IV ER-positive/HER2-negative breast cancer who may benefit from first-line chemotherapy. In the current study we have shown that ER-positive/HER2-negative tumors with high-risk recurrence scores had shorter time to develop metastatic disease and shorter overall survival, however we were not able to confirm an association with chemotherapy response.

In this study several limitations have been recognized. As already mentioned, heterogeneity of the given chemotherapeutic agents and non-availability of an independent gene expression data set with CT response information in the metastatic setting are the main limitations to be acknowledged. Nonetheless, the detailed information on response to CT in the setting of metastatic disease in a group of 118 patients is one of the strengths of this study.

Conclusions

We present a comprehensive study comparing the gene expression patterns of primary tumors from metastatic breast cancer patients according to their responsiveness of chemotherapy during their treatment of metastatic disease. The 14 differentially expressed genes among these two groups have been further investigated and led to the exploration of couple genes that might play role in the response to CT. In contrast to the findings for neoadjuvant chemotherapy treatment, there was no association of molecular subtype with response to chemotherapy in the metastatic setting. Using supervised classification, we identified a classifier of chemotherapy response; however, we could not validate this classifier using neoadjuvant response data. We believe that the data generated in this study may inspire new studies leading to development of improved and individualized therapy strategies in treatment of metastatic breast cancer.

References

1. El Saghir NS, Tfayli A, Hatoum HA, Nachef Z, Dinh P, Awada A: Treatment of metastatic breast cancer: state-of-the-art, subtypes and perspectives. *Crit Rev Oncol Hematol* 2011, 80(3):433-449.
2. Hortobagyi GN: Treatment of breast cancer. *N Engl J Med* 1998, 339(14):974-984.
3. Johnston SR: The role of chemotherapy and targeted agents in patients with metastatic breast cancer. *Eur J Cancer* 2011, 47 Suppl 3:S38-47.
4. Jones SE: Metastatic breast cancer: the treatment challenge. *Clin Breast Cancer* 2008, 8(3):224-233.
5. Beslija S, Bonneterre J, Burstein H, Cocquyt V, Gnant M, Goodwin P, Heinemann V, Jassem J, Kostler WJ, Krainer M *et al*: Second consensus on medical treatment of metastatic breast cancer. *Ann Oncol* 2007, 18(2):215-225.
6. Lin NU, Thomssen C, Cardoso F, Cameron D, Cufer T, Fallowfield L, Francis PA, Kyriakides S, Pagani O, Senkus E *et al*: International guidelines for management of metastatic breast cancer (MBC) from the European School of Oncology (ESO)-MBC Task Force: Surveillance, staging, and evaluation of patients with early-stage and metastatic breast cancer. *Breast* 2013, 22(3):203-210.
7. Chia SK, Speers CH, D'yachkova Y, Kang A, Malfair-Taylor S, Barnett J, Coldman A, Gelmon KA, O'reilly SE, Olivotto IA: The impact of new chemotherapeutic and hormone agents on survival in a population-based cohort of women with metastatic breast cancer. *Cancer* 2007, 110(5):973-979.
8. Dawood S, Broglio K, Gonzalez-Angulo AM, Buzdar AU, Hortobagyi GN, Giordano SH: Trends in survival over the past two decades among white and black patients with newly diagnosed stage IV breast cancer. *JClinOncol* 2008, 26(30):4891-4898.
9. Giordano SH, Buzdar AU, Smith TL, Kau SW, Yang Y, Hortobagyi GN: Is breast cancer survival improving? *Cancer* 2004, 100(1):44-52.
10. Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, Speers CH, Nielsen TO, Gelmon K: Metastatic behavior of breast cancer subtypes. *JClinOncol* 2010, 28(20):3271-3277.
11. Savci-Heijink CD, Halfwerk H, Hooijer GK, Horlings HM, Wesseling J, van de Vijver MJ: Retrospective analysis of metastatic behaviour of breast cancer subtypes. *Breast cancer research and treatment* 2015, 150(3):547-557.
12. Hatzis C, Pusztai L, Valero V, Booser DJ, Esserman L, Lluch A, Vidaurre T, Holmes F, Souchon E, Wang H *et al*: A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. *JAMA* 2011, 305(18):1873-1881.
13. Hess KR, Anderson K, Symmans WF, Valero V, Ibrahim N, Mejia JA, Booser D, Theriault RL, Buzdar AU, Dempsey PJ *et al*: Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J Clin Oncol* 2006, 24(26):4236-4244.

14. Liedtke C, Hatzis C, Symmans WF, Desmedt C, Haibe-Kains B, Valero V, Kuerer H, Hortobagyi GN, Piccart-Gebhart M, Sotiriou C *et al*: Genomic grade index is associated with response to chemotherapy in patients with breast cancer. *J Clin Oncol* 2009, 27(19):3185-3191.

15. Press MF, Sauter G, Buyse M, Bernstein L, Guzman R, Santiago A, Villalobos IE, Eiermann W, Pienkowski T, Martin M *et al*: Alteration of topoisomerase II-alpha gene in human breast cancer: association with responsiveness to anthracycline-based chemotherapy. *J Clin Oncol* 2011, 29(7):859-867.

16. Pusztai L: Markers predicting clinical benefit in breast cancer from microtubule-targeting agents. *Ann Oncol* 2007, 18 Suppl 12:xii15-20.

17. Rouzier R, Rajan R, Wagner P, Hess KR, Gold DL, Stec J, Ayers M, Ross JS, Zhang P, Buchholz TA *et al*: Microtubule-associated protein tau: a marker of paclitaxel sensitivity in breast cancer. *Proceedings of the National Academy of Sciences of the United States of America* 2005, 102(23):8315-8320.

18. Savci-Heijink CD, Halfwerk H, Koster J, van de Vijver MJ: A novel gene expression signature for bone metastasis in breast carcinomas. *Breast cancer research and treatment* 2016, 156(2):249-259.

19. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M *et al*: New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *EurJCancer* 2009, 45(2):228-247.

20. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ *et al*: A gene-expression signature as a predictor of survival in breast cancer. *NEnglJMed* 2002, 347(25):1999-2009.

21. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z *et al*: Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009, 27(8):1160-1167.

22. King TA, Lyman JP, Gonen M, Voci A, De Brot M, Boafo C, Sing AP, Hwang ES, Alvarado MD, Liu MC *et al*: Prognostic Impact of 21-Gene Recurrence Score in Patients With Stage IV Breast Cancer: TBCRC 013. *J Clin Oncol* 2016, 34(20):2359-2365.

23. Savci-Heijink CD, Kosari F, Aubry MC, Caron BL, Sun Z, Yang P, Vasmatzis G: The role of desmoglein-3 in the diagnosis of squamous cell carcinoma of the lung. *Am J Pathol* 2009, 174(5):1629-1637.

24. Carey LA, Dees EC, Sawyer L, Gatti L, Moore DT, Collichio F, Ollila DW, Sartor CI, Graham ML, Perou CM: The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2007, 13(8):2329-2334.

25. Guarneri V, Lenihan DJ, Valero V, Durand JB, Broglio K, Hess KR, Michaud LB, Gonzalez-Angulo AM, Hortobagyi GN, Esteva FJ: Long-term cardiac tolerability of trastuzumab in metastatic breast cancer: the M.D. Anderson Cancer Center experience. *J Clin Oncol* 2006, 24(25):4107-4115.

26. Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, Symmans WF, Gonzalez-Angulo AM, Hennessy B, Green M *et al*: Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol* 2008, 26(8):1275-1281.

27. Lips EH, Mulder L, de Ronde JJ, Mandjes IA, Koolen BB, Wessels LF, Rodenhuis S, Wesseling J: Breast cancer subtyping by immunohistochemistry and histological grade outperforms breast cancer intrinsic subtypes in predicting neoadjuvant chemotherapy response. *Breast cancer research and treatment* 2013, 140(1):63-71.

28. Ring AE, Smith IE, Ashley S, Fulford LG, Lakhani SR: Oestrogen receptor status, pathological complete response and prognosis in patients receiving neoadjuvant chemotherapy for early breast cancer. *British journal of cancer* 2004, 91(12):2012-2017.

29. Buzdar AU, Ibrahim NK, Francis D, Booser DJ, Thomas ES, Theriault RL, Pusztai L, Green MC, Arun BK, Giordano SH *et al*: Significantly higher pathologic complete remission rate after neoadjuvant therapy with trastuzumab, paclitaxel, and epirubicin chemotherapy: results of a randomized trial in human epidermal growth factor receptor 2-positive operable breast cancer. *J Clin Oncol* 2005, 23(16):3676-3685.

30. Perez EA, Romond EH, Suman VJ, Jeong JH, Sledge G, Geyer CE, Jr., Martino S, Rastogi P, Gralow J, Swain SM *et al*: Trastuzumab plus adjuvant chemotherapy for human epidermal growth factor receptor 2-positive breast cancer: planned joint analysis of overall survival from NSABP B-31 and NCCTG N9831. *J Clin Oncol* 2014, 32(33):3744-3752.

31. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, Gianni L, Baselga J, Bell R, Jackisch C *et al*: Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 2005, 353(16):1659-1672.

32. Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE, Jr., Davidson NE, Tan-Chiu E, Martino S, Paik S, Kaufman PA *et al*: Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med* 2005, 353(16):1673-1684.

33. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, Hess KR, Stec J, Ayers M, Wagner P *et al*: Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2005, 11(16):5678-5685.

34. Andre F, Mazouni C, Liedtke C, Kau SW, Frye D, Green M, Gonzalez-Angulo AM, Symmans WF, Hortobagyi GN, Pusztai L: HER2 expression and efficacy of preoperative paclitaxel/FAC chemotherapy in breast cancer. *Breast cancer research and treatment* 2008, 108(2):183-190.

35. Pritchard KI, Shepherd LE, O'Malley FP, Andrulis IL, Tu D, Bramwell VH, Levine MN, National Cancer Institute of Canada Clinical Trials G: HER2 and responsiveness of breast cancer to adjuvant chemotherapy. *N Engl J Med* 2006, 354(20):2103-2111.

36. Vermeulen E, Geesink I, Schmidt MK, Steegers C, Verhue D, Brom FW, Aaronson NK, van Leeuwen FE: [Secondary use of human tissue: consent and better information required]. *NedTijdschrGeneeskd* 2009, 153:A948.

Chapter 7

Epithelial-mesenchymal transition status of primary breast carcinomas and its correlation with metastatic behavior

C. Dilara Savci Heijink, Hans Halfwerk, Gerrit K.J. Hooijer, Jan Koster, Hugo M. Horlings, Sybren L. Meijer, Marc J. van de Vijver

Accepted for publication in Breast Cancer Res Treat

Abstract

Background: Epithelial-to-mesenchymal transition (EMT) has been implicated as an important step in the development of distant metastases. We therefore wished to study EMT-status of primary breast carcinomas from patients who during follow-up developed distant metastases.

Methods: mRNA expression profiles of primary breast carcinoma samples (n=151) from patients who developed metastatic disease were analyzed and EMT-status was designated using a previously described EMT-core signature. EMT-status of the primary tumor was correlated to clinicopathological characteristics, molecular subtypes, metastasis pattern, chemotherapy response and survival outcomes. In addition, using immunohistochemistry, the expression of several proteins implicated in EMT were studied (*CDH1*, *CDH2*, *NAT1*, *SNAI2*, *TWIST1*, *VIM* and *ZEB1*) compared with the designated EMT-status and survival.

Results: Utilizing the 130-gene-EMT core signature, 66.2% of the primary tumors in the current study was assessed as EMT-activated. In contrast to our expectations, analyses revealed that 84.6% of Luminal A tumors, 65.1% of Luminal B tumors and 55.6% of HER2-like had an activated EMT-status, compared to only 25% of the basal type tumors ($p < 0.001$). EMT-status was not correlated to the pattern of metastatic disease, metastasis specific survival and overall survival. Similarly, there was not a significant association between EMT-status of the primary tumor and chemotherapy response in the metastatic setting. Immunostaining for *NAT1* and *TWIST1* correlated with the EMT-status (p 0.003 and p 0.047, respectively). Multivariate analyses showed that *NAT1* and *TWIST1 staining* was significantly associated with EMT-status regardless of the estrogen receptor status of the tumors (p-values: 0.020 and 0.027, respectively).

Conclusions: The EMT-status of breast cancers, as defined by presence of a core EMT gene expression signature is associated with non-basal type tumors, but not with the pattern of distant metastases. Of several potential immunohistochemical EMT markers, only *NAT1* and *TWIST1* expression were associated with the gene expression based EMT-status.

Introduction

Epithelial-to-mesenchymal transition (EMT) is a complex and dynamic process that involves transdifferentiation of the cells by means of changes in the cell state. This process of epithelial-mesenchymal plasticity plays an established role in embryogenesis and early organ development [1-3]. EMT is initiated with activation of transcription factors such as *Snail*, *Twist*, *Slug* and *Zeb1* and is regulated by modulation of multiple epigenetic regulatory mechanisms [3, 4]. This process results in the loss of epithelial features and acquiring mesenchymal properties such as motility, invasiveness and resistance to apoptosis, eventually leading to colonization and metastasis formation [5]. It is thought that once colonization of the tumor cells at distant sites has occurred, these EMT-derived mesenchymal cells with stem cell like properties go through mesenchymal-to epithelial transition (MET) and re-gain epithelial features and continue to proliferate [6]. Along with its role in cancer metastasis, epithelial–mesenchymal plasticity is also indicated as the origin of systemic therapy resistance in breast cancer stem cells [7-9].

Several transgenic mouse models have provided evidence for the existence of tumor cells with a mesenchymal phenotype in different types of carcinomas [10-13]. Using genetically engineered knock-in reporter mouse lines and fluorescence activated cell sorting, Ye et al have isolated *Slug*$^{+}$ and *Snail*$^{+}$ cells in normal mammary tissue. They have shown that epithelial-to-mesenchymal transition-inducing transcription factors (EMT-TFs) *Snail*, *Twist* and *Zeb1* were expressed in stromal fibroblasts surrounding the mammary ducts, whereas *Slug* was found to be expressed in basal mammary epithelial cells. Adopting a transgenic model of mammary tumor development, as tumors progressed to more undifferentiated phase(s), they have identified that the *Snail*$^{+}$ cancer cells dissociating from epithelium acquired an elongated morphology similar to mesenchymal cells. These cells were found to have lost E-cadherin expression and activated expression of Zeb1. During the process of tumor progression these cells were also shown to gain CK14 expression especially at the invasive edges of the organoids. With these results the authors have demonstrated the potential role of *Snail* and subsequent EMT activation in obtaining basal features usually seen in more aggressive breast tumor types [14].

Despite the increasing interest in this dynamic process, it is still unknown what the exact role of EMT is in the development of distant metastases in human breast cancer. Several authors have suggested that the EMT state of a tumor can range from partial to full as opposed to a static event leading to gain or loss of a function [4, 15-17]. These studies have also identified that the main tumor bulk and the invasive front of the tumor differ, the invasive front being the main area for the EMT program to interact closely with the tumor microenvironment. Individual tumor cells which undergo EMT have been defined at the invasive front of the tumor and have been described as individual cells or small cell groups detaching from the main mass into the adjacent stroma [16, 18-21]. The difficulty to recognize and distinguish these individual cells from the pre-existing stromal cells has contributed to the controversy of existence of clinical evidence of EMT.

Recently, a quantitative EMT scoring system based on gene expression profiling of cell lines was identified. It was shown that each cancer type had its own characteristic EMT spectrum, however EMT-status of the tumors did not correlate to poorer survival or to chemotherapy resistance [22]. A prior EMT-core signature generated by using EMT-induced human mammary epithelial cells was found to be strongly correlated to metaplastic and claudin low breast cancer, but not to other gene expression based subtypes; and lacked to show correlation with poorer survival outcome [23].

To explore the accordance of the concept to reconcile the EMT-ness in clinical practice we have conducted a study utilizing gene expression profiling data from primary breast cancers of a group of patients with known metastatic disease. In the current study, the association between EMT-status of the primary tumors and their pattern of metastatic disease and the possibility of determining this EMT-status with the help of selected routine immunohistochemical stains was investigated.

Material and Methods

Patient and tumor samples

This study was conducted in line with national ethical guidelines of 'Code for Proper Secondary Use of Human Tissue' developed by Federation of Medical Societies (FMWV) in the Netherlands [24]. Metastatic breast cancer patients from the *Academic Medical Center* and the *Netherlands Cancer Institute* with available frozen material from their primary tumors were identified. Relevant detailed clinical information on metastatic disease including the metastasis site, timeline of the metastatic disease and the outcome measures (metastasis specific survival and overall survival) was collected from a group of 151 patients. The clinicopathological features of these tumors and their affiliated metastasis pattern have been reported previously [25]. For each patient, administered chemotherapy and therapy related data including chemotherapy response for the given regimen(s) during the metastatic process was carefully recorded as previously described in a subgroup of the patients (n=142) [26].

The histologic sections from the primary tumors were reviewed and additional routine staining techniques were applied to determine the hormone receptor status of the tumors [25].

Identification and validation of EMT-status

Comparing the first and the last H-E stained sections, the samples with more than 50% tumor cells were used for the gene expression profiling experiments. The details about the RNA isolation and gene expression microarrays (HumanHT-12 v4 Expression BeadChip arrays [Illumina, Inc., > 47,000 probes] have been reported previously [27]. Full information on RNA amplification, labeling and hybridization can be also found on the Illumina website (http://www.illumina.com). Following the robust spline normalization, the data was log2 transformed and processed by ComBat to tailor the batch effects.

The generated data were analyzed with help of R2 (Microarray Analysis and Visualization Platform, http://r2.amc.nl). Molecular subtypes were assessed for each tumor using the Pam50 classifier [28]. Also the percentage of tumor cells on the slides used for gene expression profiling experiments were correlated to the molecular subtypes.

To designate the EMT-status of each tumor, the EMT core gene list of Groger et. al was utilized [29]. The 130 genes of this EMT-core gene list were first mapped to the Illumina platform via Gene Symbol ID. In case of existence of multiples probes for one gene, the one with highest average signal across the samples was selected. Subsequently, a K-means clustering method was applied to separate the tumors into two groups as EMT-activated or not-EMT-activated. By means of K-means clustering method, the EMT core gene list was also applied to an independent set composed of a subset of a combined database including 376 breast cancer samples with distant organ metastases [30]. Additionally, based on gene expression levels, z-scores were calculated for each tumor to define the EMT-status in accordance with distribution of the z-scores. EMT-status identified by K-means method and z-scores of the tumors were subsequently compared to verify the identified EMT-status of the given tumor. Given the significant concordance in assigned EMT-status with K-means method and the z-scores method (*p* 3.92e-20) in our dataset, further analyses and comparisons were carried out on the EMT-status based on K-means method.

To further reconcile the EMT-ness with the help of immunohistochemical stains, a subset of 46 tumors (EMT-activated, n=23 and not-EMT -activated, n=23) were selected. The heat map created by supervised clustering with the EMT-core signature was carefully observed. Based on the current literature information on their established role in EMT, a subgroup of 7 proteins (*CDH1*, *CDH2*, *NAT1*, *SNAI2*, *TWIST1*, *VIM* and *ZEB1*) was selected for further evaluation. To test the representativeness of this subset of genes, respectively a K-means clustering method and a t-test were carried out first to classify the tumors into two groups according to their EMT-status and then to validate the performance of this classification in the same study set (used for immunophenotypic evaluation).

Immunophenotypic evaluation

Whole mount slides of the tumors from a subgroup of patients (total n=46; EMT-activated n=23, not-EMT-activated n=23) were selected for the additional immunophenotypic evaluations.

Immunohistochemical staining for *E-Cadherin* (*CDH1*, clone 24E10, Cell signaling), *N-cadherin* (*CDH2*, clone 32N/Cadherin, BD transduction Laboratories), *NAT1* (Abcam),

SNAI2 (Abcam), *Vimentin* (clone D21H3, Cell Signaling) and *ZEB1* (Sigma Life science) were performed using an automated slide preparation system (Benchmark XT, Ventana Medical Systems, Tucson Arizona, USA). The signal detection for immunohistochemistry was performed with a biotin free ultraview universal DAB detection Kit (Ventana medical systems). Immunohistochemical staining for *TWIST1* (Clone Twist2C1a, Abcam) was performed manually with a Bright DAB detection.

The immunostained slides were scored by two pathologists (C.D. S-H, S.L.M). The invasive edges and the main tumor mass were evaluated separately. Immunostaining patterns for the other antibodies were evaluated semi-quantitatively; the extent and the intensity of the expression were scored for each antibody. For *NAT1*, *SNAI2*, *TWIST* and *ZEB1* a cut-off value of 10% was used to designate tumors as positive or negative. Absence of E-Cadherin expression in > 1% of the tumor cells was considered as loss of expression and presence of *N-cadherin* and *Vimentin* expression in >1% of tumor cells were noted as gain of expression. Immunohistochemical findings were then correlated to the designated EMT-status.

Statistical analysis

The association between the EMT-status of the tumors and clinical variables were further investigated with multivariate logistic regression tests applying SPSS Statistics for Windows (Release version 21.0; IBM Corp. 2012, Armond, NY). Paired t-tests were applied for the comparisons of the immunostaining results to the EMT-status. The statistical tests were two sided and p value being less than 0.05 was considered to be significant.

Results

We studied epithelial-mesenchymal transition status, as assessed by gene expression profiling, for 151 primary invasive breast carcinomas of patients whom all developed metastatic disease. Clinical and pathologic characteristics of the primary breast tumors have been previously described [25] and are shown in Table 1.

Using the 130-gene-EMT core signature, 66.2% (n=100) of the primary tumors in the current study was assigned as EMT-activated and 33.8% (n=51) as not-EMT- activated. The heat map in Figure 1 displays the gene expression profiling pattern of 130 genes of

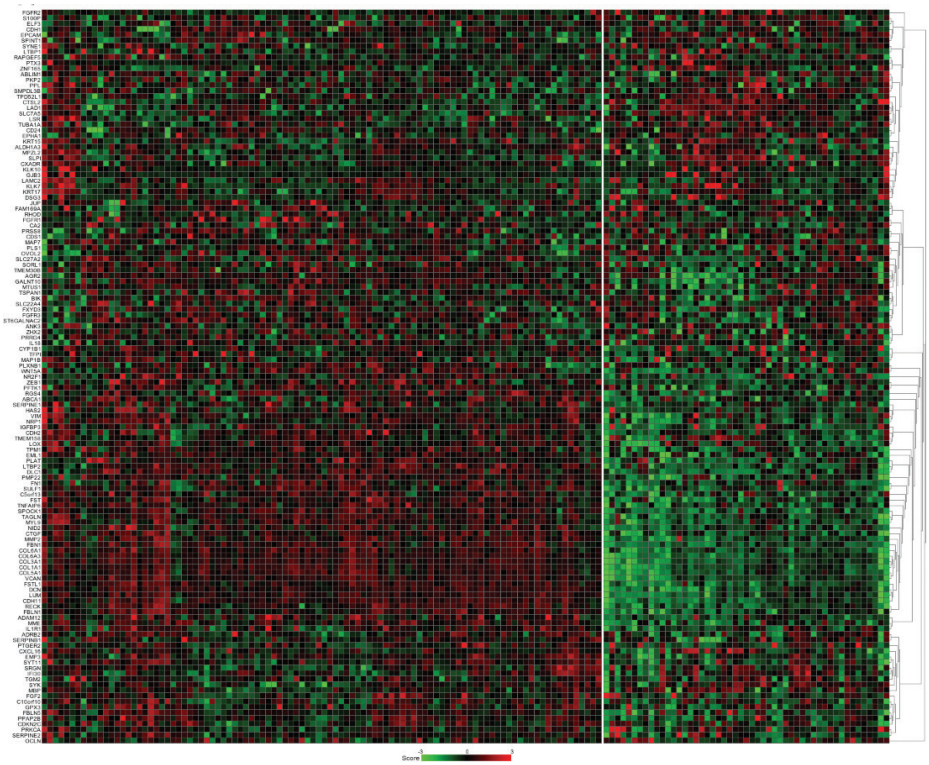


Figure 1. The gene expression pattern of 130-genes of EMT-core signature.

Heat map shows the gene expression profiling pattern of 130-genes of EMT-core signature among 151 patients. For each primary tumor the expression level of the specific gene is exhibited as red, if up-regulated and green, if down-regulated.

Table 1. Clinical and pathological characteristics of metastatic breast cancer patients

		N	%
Age at diagnosis, years	<50	83	52.9
	>50	74	47.1
Surgical procedure	none	4	2.8
	mastectomy	73	51.8
	breast conserving	64	45.4
Adjuvant therapy	none	30	21.1
	only CT	50	35.2
	only HT	17	12.0
	CT+HT	45	31.7
Lymph node status	none	43	29.3
	1-3 positive	48	32.7
	>3 positive	56	38.1
Histology	Ductal	134	86.5
	Lobular	14	9.0
	Other	7	4.5
Tumor grade	1	13	8.6
	2	84	55.3
	3	55	36.2
Time to distant metastasis[a]	early	117	77.0
	late	35	23.0
Metastasis at first presentation	no	141	92.8
	yes	11	7.2
Multiple metastasis sites at first presentation	no	97	64.2
	yes	54	35.8
Multiple metastasis sites during follow up	no	37	24.5
	yes	114	75.5

CT chemotherapy, HT hormonal therapy.
[a] *Cut-off point 5 years.*

the EMT-core signature. In the independent dataset 72,9% (n= 274) of the tumors was identified as EMT-activated and 27,1% (n=102) as not-EMT-activated.

Comparisons of the histopathological characteristics of the tumors from EMT-activated group with the tumors from not-EMT-activated group showed that 76.9% of Grade 1 and 78.5% of grade 2 tumors had an activated EMT-status as opposed to the 45.5% of grade 3 tumors ($p < 0.001$). Histologic type or the size of the tumor were not found to be correlated with the gene expression based EMT-status (p 0.635).

Among ER-positive and PR-positive tumors 77.8% and 81.1% were identified as EMT-activated, respectively ($p < 0.001$). Of Luminal A tumors 84.6%, of Luminal B tumors 65.1% and of HER2-like 55.6% was assessed as having an activated EMT status, whereas only 25% of the basal type tumors were assigned to the EMT-activated group ($p < 0.001$). Similarly, in the independent dataset [30] of total 376 tumors, 100% of the Luminal A tumors , 97,5% of the Luminal B tumors, 92,1% of HER2-like tumors and 8,3% of the basal type tumors were designated as EMT-active ($p < 0.001$). Mean tumor percentage in the basal type tumors was 72.1% (range 66.8 to 77.3). In the non-basal tumor groups, mean percentage was 58.9% in luminal A type, 65.8% in luminal B type and 62.1% in HER2-like tumors ($p < 0.001$).

EMT-status did not differ between the patients who developed metastasis within 5 years' time and the ones who developed metastatic disease later than 5 years (p 0.310). Regarding the metastasis site, out of 108 patients who developed bone metastasis 72.2% had an EMT-activated primary tumor (p 0.021); versus 65.2% of the tumors of the patients with visceral metastasis (p 0.698).

Median overall survival time was 60 months and 37 months for the EMT-activated and the not-EMT-activated group respectively (p 0.162). Metastasis specific survival time was 33 months for the patients with EMT-activated tumors and 19 months for those with non-EMT-activated tumors (p 0.036).

118 patients underwent chemotherapy treatment; 48.1% of the patients with EMT-activated tumors showed a response versus 31.4% of those with not-EMT-activated ones (p 0.130).

We subsequently wished to study the correlation of the gene expression based EMT-status of the tumors with the expression of EMT associated proteins in using immunohistochemistry in a subset of 50 tumors (25 EMT activated; 25 not-EMT activated); the findings are summarized in Table 2. Immunohistochemical evaluation for all stains revealed comparable expression patterns at the invasive edges and at the center of the tumors; therefore analyses were further carried out based on single score.

The expression patterns of staining for *E-cadherin*, *N-cadherin*, *Vimentin*, *SNAI2* and ZEB1 did not differ between the EMT-activated and not-EMT-activated groups (p values: 1.000, 0.699, 0.109, 1.000 and 0.071, respectively).

Table 2. Correlation of immunohistochemical findings and EMT-status

	IHC	EMT status		
		not-activated	activated	p
CDH1	negative	3	2	1.000
	positive	20	21	
CDH2	negative	20	18	0.699
	positive	3	5	
NAT1	negative	16	5	0.003
	positive	7	18	
SNAI2	negative	7	6	1.000
	positive	16	17	
TWIST1	negative	20	13	0.047
	positive	3	10	
VIM	negative	19	23	0.109
	positive	4	0	
ZEB1	negative	17	10	0.071
	positive	6	13	

IHC immunohistochemistry, EMT epithelial-to-mesenchymal transition

NAT1 expression was scored as positive in 25 cases with a range of 30% to 100% positivity in the tumor cells. Out of positively stained cases, 72% was assigned to the EMT-activated group; of negatively stained cases 76.2% was marked as not-EMT-activated (p 0.003). Further evaluations showed that NAT1 expression was not significantly correlated to the overall survival (p 0.223) and metastasis specific survival (p 0.146).

TWIST1 expression was found to be positive in 13 cases with 76.9% of these tumors being EMT-activated and of 60.6% of *TWIST1* negative tumors belonged to the not-EMT-activated group (p 0.047). *TWIST1* expression was not found to be significantly correlated with overall survival and metastatic specific survival (p 0.675, p 0.461, respectively).

To investigate the additional role of *NAT1* and *TWIST1* staining to predict EMT-status, multivariate regression analyses were applied. Multivariate analyses results are displayed in Table 3 and shows that positive *NAT1* and *TWIST1* staining was significantly correlated to EMT-activated status independent of ER-status of the tumor (p values: 0.020 and 0.027, respectively).

Table 3. Multivariate analyses results displaying the correlation between immuno-histochemical findings and EMT-status

	B	Wald x^2	p	Odds ratio	95% C.I.
ER-status	-1.11	0.68	.410	0.33	.02 – 4.63
NAT1-status	3.18	5.37	.020	23.98	1.63 – 352.24
TWIST1-status	2.12	4.92	.027	8.35	1.28 – 54.55

EMT epithelial-to-mesenchymal transition, ER estrogen receptor.

Discussion

In this study, gene expression profiles from primary breast carcinomas of patients with known metastatic disease have been utilized to assess the EMT-status of the primary tumor. Subsequently, the designated EMT-status has been correlated to the metastatic behavior and survival outcomes. In addition, the expression of EMT associated proteins to the EMT-status as assessed by gene expression profiling was studied.

The previously suggested reciprocal link between basal type breast cancer as assessed by immunohistochemistry and EMT [31] was not found in our data set. Unexpectedly, the low grade tumors tended to be more frequently EMT-activated than high grade tumors.

The role of epithelial to mesenchymal transition (EMT) in cancer progression has been demonstrated in several tumor models. Yet, the translation of this concept to clinical breast cancer remains problematic and it has been argued that EMT may not be required for the development of distant metastases [32]. Recent investigations focusing on these debates have led to adoption of a new concept of EMT indicating the flexibility and intermediate hybrid state of this process rather than a rigid state [4, 16, 17, 33, 34]. These recently proposed transitional states and the heterogeneity of EMT may explain the difficulty to visualize the EMT-status. In our study, we were not able to show any significant association between EMT-status and the metastasis time (early versus late metastasis). Overall survival and metastasis specific survival outcomes did not differ significantly between EMT-activated and not-EMT activated group, either. Although these results seem to be opposing to common concept that EMT-active status has bad prognostic implications, they can be due to proposed intermediate hybrid states. Tan et al have already addressed this issue with their study including several types of cancer tissue [22]. Applying a generic EMT signature, they have quantitatively estimated the extent of EMT in human tumor samples and cell lines. In this study, authors have not found a relation between EMT-status and overall and disease free survival. Particularly in the breast carcinoma samples, they have shown that tumors with mesenchymal (*Mes*) profile appeared to have better prognosis than the ones with epithelial (*Epi*) profile. The authors have suggested the role of stromal component and the distribution

of molecular subtypes for the contradictory results. Concordantly, in this study we have demonstrated that the percentage of tumor cells, hence the epithelial component differed among the molecular subtypes and luminal type tumors had relatively more stromal component than the basal type tumors ($p < 0.001$).

Next to its association with cancer progression and metastasis formation, EMT has been linked to chemoresistance in several cancer types [5, 7, 9, 35-38]. Several studies have demonstrated that cells with an EMT profile, rather than directly establishing metastasis, showed more resistance to chemotherapy (CT) and have indicated the potential role of EMT-targeted therapy. In our current study we were not able to demonstrate a link between EMT-status of the primary breast tumors and response to CT in the metastatic setting. The study conducted by the group of Tan has also failed to show a direct translation of EMT status to chemotherapy resistance [22]. These authors have concluded that in addition to acquiring EMT, gaining stem cell-like properties plays an important role in chemoresistance. Several studies have already shown that overexpression of EMT-inducing transcription factors leads to changing luminal lineage cells to a more stem cell-like trait suggesting that these breast cancer stem cells showing an EMT-like profile are more chemotherapy resistant [5, 35, 36, 39]. A generic EMT signature which is developed to assess the EMT-status, may not be the optimal tool to assess the stemness of the cancer cells and their potential response profile [6, 22, 31, 40].

Activation of an EMT program has been suggested as a critical event for cancer progression which grants epithelial cancer cells with more invasive mesenchymal phenotypes [3]. Direct visualization of these cells going through this process and their morphological changes remains an area of interest. To reveal/recognize the cancer cells with EMT-phenotype, we have performed immunostaining for *CDH1*, *CDH2*, *NAT1*, *SNAI2*, *TWIST1*, *VIM* and *ZEB1*. We were not able demonstrate significant difference between EMT-activated and not-EMT-activated group regarding *CDH1*, *CDH2*, *SNAI2*, *VIM* and *ZEB1* expression, in the tumor bulk as well as at the invasive edges of the tumor tissue. Noteworthy, the staining pattern of *TWIST1* and *NAT1* have appeared to be related to the EMT-status of the primary tumor. We have already pointed out this link in a previous gene expression profiling based study [27] and its potential role, particularly as a drug target in cancer development [41-43]. Many investigators have faced

difficulties to detect cancer cells with EMT phenotype. To overcome the main obstacle which is to differentiate the stromal fibroblasts from the cells with EMT phenotype, Yu et al conducted a study using RNA in situ hybridizations on HER2-positive breast tumors in order to distinguish primary tumor cells from the surrounding stromal cells [34]. By using dual-colorimetric RNA-in situ hybridizations, they were able to identify breast cancer cells co-expressing epithelial and mesenchymal markers. Contrary to expectations, these biphenotypic cells were observed mainly in draining lymph nodes but not at the invasive fronts of primary tumors. Alongside the heterogeneous nature of EMT process and possibility of an incomplete EMT state, it has also been suggested that molecular alterations that initiate a signal transduction cascade leading to EMT properties does not necessarily prompt acquirement of a complete mesenchymal phenotype [15].

In conclusion, our results fails to draw a direct line between the gene expression based EMT-status of a primary tumor and its associated metastatic behavior. In this study we have also demonstrated that, immunostaining for *NAT1* and *TWIST1* may be of help to identify the tumor cells with EMT-phenotype. We believe that our study is a valuable addition to the current literature and gives additional perspective on EMT in human metastatic breast carcinomas.

References

1. Lee JM, Dedhar S, Kalluri R, Thompson EW: The epithelial-mesenchymal transition: new insights in signaling, development, and disease. *J Cell Biol* 2006, 172(7):973-981.

2. Nieto MA: Epithelial plasticity: a common theme in embryonic and cancer cells. *Science* 2013, 342(6159):1234850.

3. Thiery JP, Acloque H, Huang RY, Nieto MA: Epithelial-mesenchymal transitions in development and disease. *Cell* 2009, 139(5):871-890.

4. Tam WL, Weinberg RA: The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat Med* 2013, 19(11):1438-1449.

5. Singh A, Settleman J: EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene* 2010, 29(34):4741-4751.

6. Brabletz T: To differentiate or not--routes towards metastasis. *Nat Rev Cancer* 2012, 12(6):425-436.

7. Kajiyama H, Shibata K, Terauchi M, Yamashita M, Ino K, Nawa A, Kikkawa F: Chemoresistance to paclitaxel induces epithelial-mesenchymal transition and enhances metastatic potential for epithelial ovarian carcinoma cells. *Int J Oncol* 2007, 31(2):277-283.

8. Li X, Lewis MT, Huang J, Gutierrez C, Osborne CK, Wu MF, Hilsenbeck SG, Pavlick A, Zhang X, Chamness GC *et al*: Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J Natl Cancer Inst* 2008, 100(9):672-679.

9. Yang AD, Fan F, Camp ER, van Buren G, Liu W, Somcio R, Gray MJ, Cheng H, Hoff PM, Ellis LM: Chronic oxaliplatin resistance induces epithelial-to-mesenchymal transition in colorectal cancer cell lines. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2006, 12(14 Pt 1):4147-4153.

10. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, Rasmussen KE, Jones LP, Assefnia S, Chandrasekharan S *et al*: Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* 2007, 8(5):R76.

11. McCoy EL, Iwanaga R, Jedlicka P, Abbey NS, Chodosh LA, Heichman KA, Welm AL, Ford HL: Six1 expands the mouse mammary epithelial stem/progenitor cell pool and induces mammary tumors that undergo epithelial-mesenchymal transition. *The Journal of clinical investigation* 2009, 119(9):2663-2677.

12. Rhim AD, Mirek ET, Aiello NM, Maitra A, Bailey JM, McAllister F, Reichert M, Beatty GL, Rustgi AK, Vonderheide RH *et al*: EMT and dissemination precede pancreatic tumor formation. *Cell* 2012, 148(1-2):349-361.

13. Strizzi L, Bianco C, Normanno N, Seno M, Wechselberger C, Wallace-Jones B, Khan NI, Hirota M, Sun Y, Sanicola M *et al*: Epithelial mesenchymal transition is a characteristic of hyperplasias and tumors in mammary gland from MMTV-Cripto-1 transgenic mice. *J Cell Physiol* 2004, 201(2):266-276.

14. Ye X, Tam WL, Shibue T, Kaygusuz Y, Reinhardt F, Ng Eaton E, Weinberg RA: Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. *Nature* 2015, 525(7568):256-260.

15. Christiansen JJ, Rajasekaran AK: Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis. *Cancer Res* 2006, 66(17):8319-8326.

16. Huang RY, Wong MK, Tan TZ, Kuay KT, Ng AH, Chung VY, Chu YS, Matsumura N, Lai HC, Lee YF *et al*: An EMT spectrum defines an anoikis-resistant and spheroidogenic intermediate mesenchymal state that is sensitive to e-cadherin restoration by a src-kinase inhibitor, saracatinib (AZD0530). *Cell Death Dis* 2013, 4:e915.

17. Kong D, Li Y, Wang Z, Sarkar FH: Cancer Stem Cells and Epithelial-to-Mesenchymal Transition (EMT)-Phenotypic Cells: Are They Cousins or Twins? *Cancers (Basel)* 2011, 3(1):716-729.

18. Brabletz T, Jung A, Reu S, Porzner M, Hlubek F, Kunz-Schughart LA, Knuechel R, Kirchner T: Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. *Proceedings of the National Academy of Sciences of the United States of America* 2001, 98(18):10356-10361.

19. Prall F: Tumour budding in colorectal carcinoma. *Histopathology* 2007, 50(1):151-162.

20. Wicki A, Lehembre F, Wick N, Hantusch B, Kerjaschki D, Christofori G: Tumor invasion in the absence of epithelial-mesenchymal transition: podoplanin-mediated remodeling of the actin cytoskeleton. *Cancer Cell* 2006, 9(4):261-272.

21. Wyckoff JB, Wang Y, Lin EY, Li JF, Goswami S, Stanley ER, Segall JE, Pollard JW, Condeelis J: Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. *Cancer Res* 2007, 67(6):2649-2656.

22. Tan TZ, Miow QH, Miki Y, Noda T, Mori S, Huang RY, Thiery JP: Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol Med* 2014, 6(10):1279-1293.

23. Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, Hartwell K, Onder TT, Gupta PB, Evans KW *et al*: Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proceedings of the National Academy of Sciences of the United States of America* 2010, 107(35):15449-15454.

24. Vermeulen E, Geesink I, Schmidt MK, Steegers C, Verhue D, Brom FW, Aaronson NK, van Leeuwen FE: [Secondary use of human tissue: consent and better information required]. *NedTijdschrGeneeskd* 2009, 153:A948.

25. Savci-Heijink CD, Halfwerk H, Hooijer GK, Horlings HM, Wesseling J, van de Vijver MJ: Retrospective analysis of metastatic behaviour of breast cancer subtypes. *Breast cancer research and treatment* 2015, 150(3):547-557.

26. Savci-Heijink CD, Halfwerk H, Koster J, Van de Vijver MJ: Association between gene expression profile of the primary tumor and chemotherapy response of metastatic breast cancer. *BMC Cancer* 2017, 17(1):755.

27. Savci-Heijink CD, Halfwerk H, Koster J, van de Vijver MJ: A novel gene expression signature for bone metastasis in breast carcinomas. *Breast cancer research and treatment* 2016, 156(2):249-259.

28. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z et *al*: Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009, 27(8):1160-1167.

29. Groger CJ, Grubinger M, Waldhor T, Vierlinger K, Mikulits W: Meta-analysis of gene expression signatures defining the epithelial to mesenchymal transition during cancer progression. *PLoS One* 2012, 7(12):e51136.

30. Harrell JC, Prat A, Parker JS, Fan C, He X, Carey L, Anders C, Ewend M, Perou CM: Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer ResTreat* 2012, 132(2):523-535.

31. Sarrio D, Rodriguez-Pinilla SM, Hardisson D, Cano A, Moreno-Bueno G, Palacios J: Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. *Cancer Res* 2008, 68(4):989-997.

32. Tarin D, Thompson EW, Newgreen DF: The fallacy of epithelial mesenchymal transition in neoplasia. *Cancer Res* 2005, 65(14):5996-6000; discussion 6000-5991.

33. Abell AN, Jordan NV, Huang W, Prat A, Midland AA, Johnson NL, Granger DA, Mieczkowski PA, Perou CM, Gomez SM *et al*: MAP3K4/CBP-regulated H2B acetylation controls epithelial-mesenchymal transition in trophoblast stem cells. *Cell Stem Cell* 2011, 8(5):525-537.

34. Yu M, Bardia A, Wittner BS, Stott SL, Smas ME, Ting DT, Isakoff SJ, Ciciliano JC, Wells MN, Shah AM *et al*: Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* 2013, 339(6119):580-584.

35. Creighton CJ, Li X, Landis M, Dixon JM, Neumeister VM, Sjolund A, Rimm DL, Wong H, Rodriguez A, Herschkowitz JI *et al*: Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proceedings of the National Academy of Sciences of the United States of America* 2009, 106(33):13820-13825.

36. Fischer KR, Durrans A, Lee S, Sheng J, Li F, Wong ST, Choi H, El Rayes T, Ryu S, Troeger J *et al*: Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* 2015, 527(7579):472-476.

37. Zhang Y, Toy KA, Kleer CG: Metaplastic breast carcinomas are enriched in markers of tumor-initiating cells and epithelial to mesenchymal transition. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2012, 25(2):178-184.

38. Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, Sugimoto H, Wu CC, LeBleu VS, Kalluri R: Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* 2015, 527(7579):525-530.

39. Guo W, Keckesova Z, Donaher JL, Shibue T, Tischler V, Reinhardt F, Itzkovitz S, Noske A, Zurrer-Hardi U, Bell G *et al*: Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell* 2012, 148(5):1015-1028.

7

Chapter 8

General discussion and concluding remarks

General discussion

Breast cancer is a heterogeneous disease, also with respect to the process of developing distant metastases. There is marked variability in the time interval between the initial presentation of the primary tumor and the manifestation of distant metastases; sites and the sequence of the organs involved; and the survival outcomes including response to systemic therapy for each patient. It has been known for a long time that the metastasis pattern of breast cancer differs by hormone receptor status of the primary tumor. Hormone receptor positive tumors are known to have a tendency to develop bone metastases, while triple negative tumors show increased rates of visceral organ metastases. Human epidermal growth factor receptor 2 (HER2) positive tumors are also reported to have more frequently metastases to the brain compared to HER2-negative tumors [1-8]. Despite the recent improvements in prolonging the survival time of patients with metastatic disease [9-11], patients with triple negative breast tumors continue to have a dismal prognosis following the development of distant metastases [1, 12-15] and have shorter overall survival times compared to the patients with hormone receptor- and/or HER2-positive tumors [9].

The concept of organotropism comprises the propensity of a primary tumor to metastasize to secondary organ-sites in a non-random fashion. This non-random distribution to secondary organ-site involvement is observed in different cancer types as well as within a given type of cancer, which suggests intrinsic heterogeneity between cancer cells [16]. The investigations focusing on this distinct manner of metastasis brought forth several hypothetical explanations for this process. One of the widely accepted metastasis models is Stephen Paget's "*seed and soil*" hypothesis, which advocates a more directed way of metastasizing instead of a haphazard manifestation. According to the "*seed and soil*" hypothesis, the development of metastasis fully relies on the favourable interactions between the subpopulation of tumor cells, "*seed*", and the microenvironment of organ sites, "*soil*", that they preferentially select to grow in [17]. *P*aget' s century old hypothesis is supported by both clinical and experimental research and still forms a basis for organ-specific metastasis related investigation to unravel the underlying mechanisms [18, 19].

The analyses of large numbers of breast carcinomas has revealed that breast cancer shows distinctive and greatly variable genetic alterations [20]. Identification of specific genetic alteration patterns associated with certain clinical behavior may be of clinical value in the clinical management of breast cancer. Genome-wide associated studies using various molecular techniques provide considerable value in studying the underlying biology and the possible links leading to better clinical outcomes in this heterogeneous process. In **chapter 2** of this thesis, a perspective on genomic alterations of breast cancer and their translation into clinical application is presented in conjunction with the article by *Russness* et al [21]. Using array based comparative genomic hybridization (aCGH) data of 4 clinical cohorts, including 569 breast tumors, these authors present two new algorithms to predict the genomic complexity of a given tumor. These two algorithms include a whole-arm aberration index (WAAI); measuring all events involving whole chromosome arms and a complex arm-wise aberration index (CAAI); measuring events including local aberrations to recognize regions with structural complexity. The results reveal that type A tumors, those with whole-arm gain or 1q and/or loss of 16q, are composed mainly of estrogen receptor (ER) positive and luminal A tumors showing high-magnitude WAAI scores. A2 tumors showed more arms with high-magnitude WAAI scores and tended to be more aneuploid compared to A1 tumors. A2 tumors were also histologically higher-grade tumors and associated with poor clinical outcomes. Type B tumors, those with regional loss on 5q and/or gain on 10p, showed more divergent genomic patterns and included mainly basal-type tumors. Most of the HER2-like tumors and normal-like subtype tumors, as well as the 30% of the basal-type tumors, were classified as Type C tumors (tumors showing none of the genetic alterations defined in groups A and B). Importantly, the CAAI score was shown to be independently prognostic with high scores of CAAI indicating poor clinical outcomes compared to low CAAI scores. This study displays the complexity of the analyses needed to overlay the genetic alterations and other clinicopathological parameters. In order to provide individualized therapy with minimal side-effects, genomic tests are increasingly being utilized as a companion diagnostic in the treatment of breast cancer patients. Translation of these genomic applications into the clinic remains challenging because of (*i*) the heterogeneity of the studies and (*ii*) the intricate biology of breast cancer, (*iii*) the complexity of the analyses, (*iv*) small-scale studies with limited sample sizes and therefore lack of independent validation, and (*v*) the relatively uncommon occurrence of most of genetic events in breast cancer [21]. The study by *Russness* et

al, which was discussed in this perspective paper, is an example of integrating genetic alterations with clinicopathological features and the molecular subtype leading to a scoring system with independent prognostic value.

Several comprehensive molecular studies revealed notable differences, regarding clinical behavior, between ER-positive and ER-negative tumors [22-26]. ER-positive tumors tend to develop more often metastases to bone when compared to ER-negative tumors [23, 24, 27, 28]. Regardless of the recent discovery of intrinsic subtypes based on gene expression profiling, immunohistochemistry still is a practical tool to define the subgroups in breast cancer. In the study presented in **chapter 3**, we investigated the presence of organ-specific metastasis and accompanying characteristics of the metastatic breast cancer in a retrospective case series of 263 breast cancer patients, focusing on the immunophenotypic features of the primary tumor. This study demonstrated that subtypes of breast tumors, mainly defined by conventional immunohistochemistry, show significant association with metastatic behavior, with respect to site-specific relapse, metastasis time and survival outcomes. Median overall and metastasis-specific survival times were longer for the patients with hormone receptor positive tumors. Hormone receptor positive tumors showed higher tendency to develop bone metastases, while hormone receptor negative tumors had more propensity to develop visceral organ metastases. HER2 status of the tumor was not found to be correlated to the pattern of metastasis in this data set. Patients who had primary breast tumors, which developed visceral metastases, were also observed to have shorter overall survival times, shorter metastasis specific survival times and more frequent presence of multiple metastases in the course of disease, compared to the tumors without visceral organ metastases. The associations identified in this study are of help for decisions for further follow-up and therapy in individual patients.

Experimental models of metastasis development have identified several distinct gene sets, which are reported to mediate organ-specific metastasis in breast cancer [29-32]. These distinct gene sets discovered in mouse model systems, were subsequently validated in human breast cancer cohorts. Despite the indisputable significance of these well received studies investigating the biology of metastatic breast cancer, little progress has been accomplished since then to identify a clinically applicable gene expression signature for organ-specific metastasis. Moreover, when tested in human

breast cancer samples, these experimentally identified gene sets were not as strongly associated with organ-specific metastasis as in the experimental mouse models. Taking into consideration the fact that primary tumors and the metastases display similar gene expression profiles [33], in **chapters 4** and **5**, we investigated the association between gene expression profiles of primary tumors and their metastasis patterns with the aim of developing predictors for organ-specific distinct metastasis. These chapters present studies based on analysis of gene expression profiling data of 157 primary invasive breast carcinomas of patients with known metastatic disease. Next to the correlation of gene expression profiling to the metastatic behavior of the tumor, bone and visceral organ specific metastases related genes were explored in the **chapters 4** and **5**, respectively.

Chapter 4 introduces a novel gene expression signature identified by using supervised classification. This gene expression signature included 15 genes with three of these genes, namely *NAT1*, *PH-4 and BBS1* being upregulated. These overexpressed genes of the 15-gene bone–specific metastasis signature were linked to protein transport and metabolic and oxidation-reduction processes, corresponding to the earlier studies suggesting the possible role of these genes in modification of the host tissue microenvironment to reach a bone metastasis [34-37]. When tested in an independent data set composed of 376 breast carcinomas with available site-specific metastasis information, our 15-gene signature performed better in terms of correlation with the development of bone metastasis independent of ER-status of the tumors, compared to the previously identified signatures.

In the course of metastatic disease, development of visceral organ metastasis has been linked to worse overall and metastasis-specific survival outcomes and occurrence of more frequent multiple metastases [38]. Therefore, in **chapter 5**, we sought a gene expression signature to identify the group of primary tumors with more likelihood to develop visceral organ metastasis. Utilizing the microarray data generated from 157 primary breast tumors, we were able to identify a set of genes, which were differentially expressed in the primary tumors of the patients who developed visceral organ metastasis. This gene expression signature that was composed of 14 genes was not only significantly associated with development of visceral organ metastasis (as first site, as only site and at any time), but also found to be significantly correlated to

overall and metastasis specific survival outcomes. A robust gene expression predictor for development of site specific metastasis in breast cancer has potential clinical value; gained information on genes being expressed at higher or lower levels in the metastatic setting may initiate the development of novel targeted therapies to prevent development of site specific metastases.

Several gene expression-based studies to identify a genomic predictor of chemotherapy responsiveness in the neoadjuvant setting have been conducted, but have not increased our understanding of chemotherapy responsiveness/resistance [39-48]. Some interesting results of such studies have been obtained; for example, a predictive algorithm for response to neoadjuvant therapy among HER2-negative patients reported to have superior/better positive predictive values in comparison with other predictors to chemosensitivity such as, PAM50, genomic grade index (GGI) and DDLA30 [43]. Additionally, several studies investigating the role of the 21-gene recurrence score in predicting chemotherapy response, have reported encouraging results regarding the use of 21-gene RS as a selective tool, including among patients with Stage IV disease [49-51].

In breast cancer, several prognostic gene expression signatures are used to guide adjuvant systemic therapy decisions. The prognostic value of MammaPrint® in ER-positive/PR-positive breast cancer has been demonstrated in retrospective studies [52-56] and validated by an independent multi-center study [57]. The prospective MINDACT trial (Microarray in Node negative Disease may Avoid ChemoTherapy) has investigated the clinical benefit of the inclusion of the 70-gene assay to standard clinicopathologic criteria to select patients for adjuvant chemotherapy [58]. This study has included 6,693 women with early-stage breast cancer and categorized the patients into risk groups based on their genomic risk (using the 70-gene assay) and their clinical risk (assessed by a modified version of Adjuvant! Online). Among the patient group, which was assigned to have a high clinical risk and low genomic risk, the patients who did not receive adjuvant chemotherapy had a 5-year distant metastasis-free survival rate of 94.4%. Within the same group, those who received chemotherapy had a 5-year survival rate of 95.9%. Furthermore, similar rates of survival outcomes were observed in the group of patients with ER-positive and/or PR-positive/HER2-negative, and either lymph node-negative or lymph node-positive tumors. The results of this study have

indicated that approximately 46% of women with breast cancer, who are identified to have high clinical risk, may avoid chemotherapy [58]. The data presented by the MINDACT study revealed that the addition of MammaPrint® may provide guidance in decision-making on adjuvant chemotherapy for women who have a high clinical risk but low genomic risk. These results have also led to a recent update in recommendations to use biomarkers to guide decisions on adjuvant chemotherapy by American Society of Clinical Oncology clinical practice guideline [59]. Similarly, based on the data reported by preceding retrospective studies [60, 61] showing that the 21-gene recurrence score predicts benefit from adjuvant chemotherapy in ER-positive breast cancer, a prospective clinical trial, Trail Assigning Individualized Options for Treatment (TAILORx) has been performed [62]. The TAILORx study included women with hormone receptor-positive, HER2-negative and lymph node negative breast cancer. Of these patients, 15.9% (*n*= 1626) was assigned to have low-risk profile (recurrence scores 0-10) tumors according to the 21-gene assay (Oncotype DX®). These patients who had low-risk profile tumors underwent endocrine therapy alone, without adjuvant chemotherapy. The analyses have demonstrated that the survival rates at 5-years were 93.8% for disease-free, 99.3% for distant metastases free and 98% for overall survival times for these patients with low-risk profile tumors who received endocrine therapy alone [62]. The 6,711 women with hormone receptor positive/HER-2 negative and lymph node negative breast cancers from the TAILORx study with midrange recurrence scores (11 to 25) assessed by 21-genes assay, were randomized to undergo aduvant chemoendocrine therapy or endocrine therapy. The patients in both arms had similar disease-free (84.3% in the endocrine therapy only group and 83.3% in the chemoendocrine therapy group) and overall (93.8% in the endocrine therapy only group and 93.9% in the chemoendocrine therapy group) survival rates [63].

To this day a clinically validated genomic predictor for chemotherapy response in metastatic breast cancer has not been identified. **Chapter 6** describes a study of the association between gene expression profiles of primary breast tumors and their chemotherapy responsiveness in the metastatic setting. In this study, gene expression profiles of primary breast carcinomas and their corresponding response to given chemotherapy were compared and resulted in a distinct gene set composed of 14-genes. Due to lack of an available data set with chemotherapy response data in the metastatic setting, the performance of this gene set was validated in tumor series from

patients who received neoadjuvant chemotherapy. Next to the newly generated 14-gene predictor, we have also validated the other gene sets that reported to be predictive of response to neoadjuvant chemotherapy. Other than the DLDA30 signature, none of the signatures (GGI, genomic predictor of *Hatzis*, PAM50 and 21-gene RS) tested, were found to predict the response to chemotherapy successfully in the therapy setting of metastatic disease. In this study, we have also identified that chemotherapy response rates and their association with hormone receptor status and molecular subtypes differed from the association reported in the neoadjuvant chemotherapy setting. In contrast to the response rates to neodjuvant chemotherapy, no significant correlation between molecular subtypes and chemotherapy response were determined. Specifically, when compared to luminal-type tumors; basal- type and HER2-like tumors were not found to have better chemotherapy response.

Along its established role in embryogenesis and early organ development, epithelial-to-mesenchymal transition (EMT), is also implicated in cancer progression and therapy resistance in various cancer types [64-69] and breast cancer stem cells [70]. EMT is defined as a complex reversible process of epithelial-mesenchymal plasticity and comprises the transdifferentiation of cell states. In **chapter 7**, we examined the association between gene expression based EMT-status of the primary breast cancers to their metastasis pattern and survival outcomes. In this study, we report distinctive results, which fail to indicate a direct connection between EMT-status of the primary breast cancers and site of metastasis, time to develop metastasis, or overall survival time. We were also not able to show any significant relation between EMT-status and chemotherapy response in the metastatic setting. As opposed to the expectations, only 25% of the basal-type tumors in our study were assessed as EMT-active, compared to the with 84.6% and 65.1% of luminal A and luminal B tumors, respectively. These unexpected findings may be due to a large contribution of the gene expression pattern of stromal cells to the EMT core signature, as there was a strong correlation between tumors positive for the EMT core signature and the percentage of stromal cells. Additional immunohistochemistry applied to reconcile the EMT-ness in the tumor cells did not reveal correlation between expression patterns of immunostaining for Vimentin, E-cadherin, N-cadherin, SNAI2, and ZEB1. Our analyses demonstrated that staining for NAT1 and TWIST1 could be of value to capture EMT-like properties in breast cancer. Notably, NAT1 staining was found to be significantly related to metastasis specific and

overall survival. In addition to the heterogeneity of the EMT process, we explain the discrepancy of our results with newly proposed possible intermediate hybrid states of EMT, which suggests that the EMT-state of a tumor can range from being partial to full-state [68, 71-74].

Concluding remarks

The work described in this thesis, confirms that some traits (i.e. intrinsic subtypes) identified by gene expression profiling of primary breast carcinomas, are correlated with metastasis pattern, but fail to fully conceal the behavior of metastatic breast cancer. Furthermore, our analyses show that correlations between gene expression profiles of the primary breast carcinomas and prediction of chemotherapy response differ in the neoadjuvant and the metastatic setting. To unravel the mechanisms that lead to site-specific metastasis and the mechanisms that underlie the responsiveness/resistance to chemotherapy will require more research to establish clinically validated genomic predictors.

References

1. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P, Narod SA: Triple-negative breast cancer: clinical features and patterns of recurrence. *ClinCancer Res* 2007, 13(15 Pt 1):4429-4434.

2. Fuqua SA: The role of estrogen receptors in breast cancer metastasis. *JMammaryGlandBiolNeoplasia* 2001, 6(4):407-417.

3. Heitz F, Harter P, Lueck HJ, Fissler-Eckhoff A, Lorenz-Salehi F, Scheil-Bertram S, Traut A, du BA: Triple-negative and HER2-overexpressing breast cancers exhibit an elevated risk and an earlier occurrence of cerebral metastases. *EurJCancer* 2009, 45(16):2792-2798.

4. Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, Speers CH, Nielsen TO, Gelmon K: Metastatic behavior of breast cancer subtypes. *JClinOncol* 2010, 28(20):3271-3277.

5. Koo JS, Jung W, Jeong J: Metastatic breast cancer shows different immunohistochemical phenotype according to metastatic site. *Tumori* 2010, 96(3):424-432.

6. Largillier R, Ferrero JM, Doyen J, Barriere J, Namer M, Mari V, Courdi A, Hannoun-Levi JM, Ettore F, Birtwisle-Peyrottes I *et al*: Prognostic factors in 1,038 women with metastatic breast cancer. *AnnOncol* 2008, 19(12):2012-2019.

7. Lin NU, Claus E, Sohl J, Razzak AR, Arnaout A, Winer EP: Sites of distant recurrence and clinical outcomes in patients with metastatic triple-negative breast cancer: high incidence of central nervous system metastases. *Cancer* 2008, 113(10):2638-2645.

8. Slimane K, Andre F, Delaloge S, Dunant A, Perez A, Grenier J, Massard C, Spielmann M: Risk factors for brain relapse in patients with metastatic breast cancer. *AnnOncol* 2004, 15(11):1640-1644.

9. Chia SK, Speers CH, D'yachkova Y, Kang A, Malfair-Taylor S, Barnett J, Coldman A, Gelmon KA, O'reilly SE, Olivotto IA: The impact of new chemotherapeutic and hormone agents on survival in a population-based cohort of women with metastatic breast cancer. *Cancer* 2007, 110(5):973-979.

10. Dawood S, Broglio K, Gonzalez-Angulo AM, Buzdar AU, Hortobagyi GN, Giordano SH: Trends in survival over the past two decades among white and black patients with newly diagnosed stage IV breast cancer. *JClinOncol* 2008, 26(30):4891-4898.

11. Giordano SH, Buzdar AU, Smith TL, Kau SW, Yang Y, Hortobagyi GN: Is breast cancer survival improving? *Cancer* 2004, 100(1):44-52.

12. Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V: Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer* 2007, 109(9):1721-1728.

13. Foulkes WD, Brunet JS, Stefansson IM, Straume O, Chappuis PO, Begin LR, Hamel N, Goffin JR, Wong N, Trudel M *et al*: The prognostic implication of the basal-like (cyclin E high/p27

low/p53+/glomeruloid-microvascular-proliferation+) phenotype of BRCA1-related breast cancer. *Cancer Res* 2004, 64(3):830-835.

14. Haffty BG, Yang Q, Reiss M, Kearney T, Higgins SA, Weidhaas J, Harris L, Hait W, Toppmeyer D: Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. *JClinOncol* 2006, 24(36):5652-5657.

15. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA et *al*: Molecular portraits of human breast tumours. *Nature* 2000, 406(6797):747-752.

16. Lu X, Kang Y: Organotropism of breast cancer metastasis. *JMammaryGlandBiolNeoplasia* 2007, 12(2-3):153-162.

17. Paget S: The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev* 1989, 8(2):98-101.

18. Fidler IJ: Selection of successive tumour lines for metastasis. *NatNew Biol* 1973, 242(118):148-149.

19. Langley RR, Fidler IJ: The seed and soil hypothesis revisited--the role of tumor-stroma interactions in metastasis to different organs. *IntJCancer* 2011, 128(11):2527-2535.

20. Hanahan D, Weinberg RA: The hallmarks of cancer. *Cell* 2000, 100(1):57-70.

21. Russnes HG, Vollan HKM, Lingjaerde OC, Krasnitz A, Lundin P, Naume B, Sorlie T, Borgen E, Rye IH, Langerod A et *al*: Genomic architecture characterizes tumor progression paths and fate in breast cancer patients. *Sci Transl Med* 2010, 2(38):38ra47.

22. Andry G, Suciu S, Pratola D, Sylvester R, Leclercq G, da Costa PM, Legros N, Andry-t'Hooft M, Verhest A, Mattheiem W et *al*: Relation between estrogen receptor concentration and clinical and histological factors: their relative prognostic importance after radical mastectomy for primary breast cancer. *EurJCancer ClinOncol* 1989, 25(2):319-329.

23. Hess KR, Pusztai L, Buzdar AU, Hortobagyi GN: Estrogen receptors and distinct patterns of breast cancer relapse. *Breast Cancer ResTreat* 2003, 78(1):105-118.

24. Koenders PG, Beex LV, Langens R, Kloppenborg PW, Smals AG, Benraad TJ: Steroid hormone receptor activity of primary human breast cancer and pattern of first metastasis. The Breast Cancer Study Group. *Breast Cancer ResTreat* 1991, 18(1):27-32.

25. Raemaekers JM, Beex LV, Koenders AJ, Pieters GF, Smals AG, Benraad TJ, Kloppenborg PW: Disease-free interval and estrogen receptor activity in tumor tissue of patients with primary breast cancer: analysis after long-term follow-up. *Breast Cancer ResTreat* 1985, 6(2):123-130.

26. Winstanley J, Cooke T, Murray GD, Platt-Higgins A, George WD, Holt S, Myskov M, Spedding A, Barraclough BR, Rudland PS: The long term prognostic significance of c-erbB-2 in primary breast cancer. *BrJCancer* 1991, 63(3):447-450.

27. Campbell FC, Blamey RW, Elston CW, Nicholson RI, Griffiths K, Haybittle JL: Oestrogen-receptor status and sites of metastasis in breast cancer. *BrJCancer* 1981, 44(3):456-459.

28. Solomayer EF, Diel IJ, Meyberg GC, Gollan C, Bastert G: Metastatic breast cancer: clinical course, prognosis and therapy related to the first site of metastasis. *Breast cancer research and treatment* 2000, 59(3):271-278.

29. Bos PD, Zhang XH, Nadal C, Shu W, Gomis RR, Nguyen DX, Minn AJ, van de Vijver MJ, Gerald WL, Foekens JA *et al*: Genes that mediate breast cancer metastasis to the brain. *Nature* 2009, 459(7249):1005-1009.

30. Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Cordon-Cardo C, Guise TA, Massague J: A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* 2003, 3(6):537-549.

31. Landemaine T, Jackson A, Bellahcene A, Rucci N, Sin S, Abad BM, Sierra A, Boudinet A, Guinebretiere JM, Ricevuto E *et al*: A six-gene signature predicting breast cancer lung metastasis. *Cancer Res* 2008, 68(15):6092-6099.

32. Minn AJ, Gupta GP, Padua D, Bos P, Nguyen DX, Nuyten D, Kreike B, Zhang Y, Wang Y, Ishwaran H *et al*: Lung metastasis genes couple breast tumor size and metastatic spread. *Proceedings of the National Academy of Sciences of the United States of America* 2007, 104(16):6740-6745.

33. Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ: Gene expression profiles of primary breast tumors maintained in distant metastases. *Proceedings of the National Academy of Sciences of the United States of America* 2003, 100(26):15901-15905.

34. Chi JT, Wang Z, Nuyten DS, Rodriguez EH, Schaner ME, Salim A, Wang Y, Kristensen GB, Helland A, Borresen-Dale AL *et al*: Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. *PLoS Med* 2006, 3(3):e47.

35. Harris AL: Hypoxia--a key regulatory factor in tumour growth. *Nat Rev Cancer* 2002, 2(1):38-47.

36. Pennacchietti S, Michieli P, Galluzzo M, Mazzone M, Giordano S, Comoglio PM: Hypoxia promotes invasive growth by transcriptional activation of the met protooncogene. *Cancer Cell* 2003, 3(4):347-361.

37. Semenza GL: Targeting HIF-1 for cancer therapy. *Nat Rev Cancer* 2003, 3(10):721-732.

38. Savci-Heijink CD, Halfwerk H, Hooijer GK, Horlings HM, Wesseling J, van de Vijver MJ: Retrospective analysis of metastatic behaviour of breast cancer subtypes. *Breast cancer research and treatment* 2015, 150(3):547-557.

39. Ayers M, Symmans WF, Stec J, Damokosh AI, Clark E, Hess K, Lecocke M, Metivier J, Booser D, Ibrahim N *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(12):2284-2293.

40. Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, Mohsin S, Osborne CK, Chamness GC, Allred DC *et al*: Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* 2003, 362(9381):362-369.

41. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, Andre S, Piccart M, Campone M, Brain E et *al*: A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. *Nat Med* 2009, 15(1):68-74.

42. Hannemann J, Oosterkamp HM, Bosch CA, Velds A, Wessels LF, Loo C, Rutgers EJ, Rodenhuis S, van de Vijver MJ: Changes in gene expression associated with response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol* 2005, 23(15):3331-3342.

43. Hatzis C, Pusztai L, Valero V, Booser DJ, Esserman L, Lluch A, Vidaurre T, Holmes F, Souchon E, Wang H et *al*: A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. *JAMA* 2011, 305(18):1873-1881.

44. Hess KR, Anderson K, Symmans WF, Valero V, Ibrahim N, Mejia JA, Booser D, Theriault RL, Buzdar AU, Dempsey PJ et *al*: Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J Clin Oncol* 2006, 24(26):4236-4244.

45. Liedtke C, Hatzis C, Symmans WF, Desmedt C, Haibe-Kains B, Valero V, Kuerer H, Hortobagyi GN, Piccart-Gebhart M, Sotiriou C et *al*: Genomic grade index is associated with response to chemotherapy in patients with breast cancer. *J Clin Oncol* 2009, 27(19):3185-3191.

46. Press MF, Sauter G, Buyse M, Bernstein L, Guzman R, Santiago A, Villalobos IE, Eiermann W, Pienkowski T, Martin M et *al*: Alteration of topoisomerase II-alpha gene in human breast cancer: association with responsiveness to anthracycline-based chemotherapy. *J Clin Oncol* 2011, 29(7):859-867.

47. Pusztai L: Markers predicting clinical benefit in breast cancer from microtubule-targeting agents. *Ann Oncol* 2007, 18 Suppl 12:xii15-20.

48. Rouzier R, Rajan R, Wagner P, Hess KR, Gold DL, Stec J, Ayers M, Ross JS, Zhang P, Buchholz TA et *al*: Microtubule-associated protein tau: a marker of paclitaxel sensitivity in breast cancer. *Proceedings of the National Academy of Sciences of the United States of America* 2005, 102(23):8315-8320.

49. Bear HD, Wan W, Robidoux A, Rubin P, Limentani S, White RL, Jr., Granfortuna J, Hopkins JO, Oldham D, Rodriguez A et *al*: Using the 21-gene assay from core needle biopsies to choose neoadjuvant therapy for breast cancer: A multicenter trial. *J Surg Oncol* 2017, 115(8):917-923.

50. King TA, Lyman JP, Gonen M, Voci A, De Brot M, Boafo C, Sing AP, Hwang ES, Alvarado MD, Liu MC et *al*: Prognostic Impact of 21-Gene Recurrence Score in Patients With Stage IV Breast Cancer: TBCRC 013. *J Clin Oncol* 2016, 34(20):2359-2365.

51. Park KU, Chen YL, Chitale D, Choi S, Ali H, Nathanson SD, Bensenhaver J, Proctor E, Petersen L, Loutfi R et *al*: Utilization of the 21-Gene Recurrence Score in a Diverse Breast Cancer Patient Population: Development of a Clinicopathologic Model to Predict High-Risk Scores and Response to Neoadjuvant Chemotherapy. *Annals of surgical oncology* 2018, 25(7):1921-1927.

52. Bueno-de-Mesquita JM, Linn SC, Keijzer R, Wesseling J, Nuyten DS, van Krimpen C, Meijers C, de Graaf PW, Bos MM, Hart AA *et al*: Validation of 70-gene prognosis signature in node-negative breast cancer. *Breast cancer research and treatment* 2009, 117(3):483-495.

53. Drukker CA, van Tinteren H, Schmidt MK, Rutgers EJ, Bernards R, van de Vijver MJ, Van't Veer LJ: Long-term impact of the 70-gene signature on breast cancer outcome. *Breast cancer research and treatment* 2014, 143(3):587-592.

54. Knauer M, Cardoso F, Wesseling J, Bedard PL, Linn SC, Rutgers EJ, van 't Veer LJ: Identification of a low-risk subgroup of HER-2-positive breast cancer by the 70-gene prognosis signature. *British journal of cancer* 2010, 103(12):1788-1793.

55. Mook S, Schmidt MK, Viale G, Pruneri G, Eekhout I, Floore A, Glas AM, Bogaerts J, Cardoso F, Piccart-Gebhart MJ *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 research and treatment* 2009, 116(2):295-302.

56. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ *et al*: A gene-expression signature as a predictor of survival in breast cancer. *NEnglJMed* 2002, 347(25):1999-2009.

57. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, d'Assignies MS, Bergh J, Lidereau R, Ellis P *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(17):1183-1192.

58. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, Pierga JY, Brain E, Causeret S, DeLorenzi M *et al*: 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med* 2016, 375(8):717-729.

59. Krop I, Ismaila N, Andre F, Bast RC, Barlow W, Collyar DE, Hammond ME, Kuderer NM, Liu MC, Mennel RG *et al*: Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline Focused Update. *J Clin Oncol* 2017, 35(24):2838-2847.

60. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, Ravdin P, Bugarini R, Baehner FL, Davidson NE *et al*: Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal 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.

61. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J *et al*: Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol* 2006, 24(23):3726-3734.

62. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, Geyer CE, Jr., Dees EC, Perez EA, Olson JA, Jr. *et al*: Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med* 2015, 373(21):2005-2014.

63. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, Geyer CE, Jr., Dees EC, Goetz MP, Olson JA, Jr. *et al*: Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med* 2018, 379(2):111-121.

64. Creighton CJ, Li X, Landis M, Dixon JM, Neumeister VM, Sjolund A, Rimm DL, Wong H, Rodriguez A, Herschkowitz JI et *al*: Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proceedings of the National Academy of Sciences of the United States of America* 2009, 106(33):13820-13825.

65. Fischer KR, Durrans A, Lee S, Sheng J, Li F, Wong ST, Choi H, El Rayes T, Ryu S, Troeger J et *al*: Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* 2015, 527(7579):472-476.

66. Kajiyama H, Shibata K, Terauchi M, Yamashita M, Ino K, Nawa A, Kikkawa F: Chemoresistance to paclitaxel induces epithelial-mesenchymal transition and enhances metastatic potential for epithelial ovarian carcinoma cells. *Int J Oncol* 2007, 31(2):277-283.

67. Singh A, Settleman J: EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene* 2010, 29(34):4741-4751.

68. Zhang Y, Toy KA, Kleer CG: Metaplastic breast carcinomas are enriched in markers of tumor-initiating cells and epithelial to mesenchymal transition. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 2012, 25(2):178-184.

69. Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, Sugimoto H, Wu CC, LeBleu VS, Kalluri R: Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* 2015, 527(7579):525-530.

70. Li X, Lewis MT, Huang J, Gutierrez C, Osborne CK, Wu MF, Hilsenbeck SG, Pavlick A, Zhang X, Chamness GC et *al*: Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J Natl Cancer Inst* 2008, 100(9):672-679.

71. Abell AN, Jordan NV, Huang W, Prat A, Midland AA, Johnson NL, Granger DA, Mieczkowski PA, Perou CM, Gomez SM et *al*: MAP3K4/CBP-regulated H2B acetylation controls epithelial-mesenchymal transition in trophoblast stem cells. *Cell Stem Cell* 2011, 8(5):525-537.

72. Huang RY, Wong MK, Tan TZ, Kuay KT, Ng AH, Chung VY, Chu YS, Matsumura N, Lai HC, Lee YF et *al*: An EMT spectrum defines an anoikis-resistant and spheroidogenic intermediate mesenchymal state that is sensitive to e-cadherin restoration by a src-kinase inhibitor, saracatinib (AZD0530). *Cell Death Dis* 2013, 4:e915.

73. Kong D, Li Y, Wang Z, Sarkar FH: Cancer Stem Cells and Epithelial-to-Mesenchymal Transition (EMT)-Phenotypic Cells: Are They Cousins or Twins? *Cancers (Basel)* 2011, 3(1):716-729.

74. Tam WL, Weinberg RA: The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat Med* 2013, 19(11):1438-1449.

Chapter 9

Summary in English and Dutch

SUMMARY

This thesis describes studies aimed at understanding the behavior of metastatic breast carcinoma through the gene expression profiling of primary breast tumors.

Chapter 1 provides a brief introduction to clinical management of metastatic breast cancer and application of genomic research in breast cancer. Furthermore, the rationale and the outline of this thesis are described.

In **chapter 2**, we present a viewpoint on genomic alterations of breast cancer and their translation into clinical application along with an article by *Russness* et al: "Genomic architecture characterizes tumor progression paths and fate in breast cancer patients" [1]. Within the scope of two algorithms developed by the authors to assess the genomic complexity of a tumor, the difficulty of overlaying genetic alterations with clinical and pathological findings is discussed. Translating the genomic alterations into clinical applications remains challenging, as a result of the heterogeneity and complexity of the performed studies, the intricate biology of breast carcinoma and rare occurrence of most of the genetic events. Genomic algorithms, as reported by the authors, can subclassify tumors based on genome-wide DNA copy number gains and losses; and incorporation of clinicopathologic characteristics and genetic alterations carry a promising role in improved patient outcomes.

With the aim of identifying the characteristics of the primary tumor and associated metastatic behavior, including metastatic disease related survival outcomes, in **chapter 3**, we examined the histomorphologic features of the primary tumors, associated metastasis pattern and survival outcomes. The retrospective study presented in this chapter utilized immunohistochemical staining for estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor 2 (HER2), epidermal growth factor receptor (EGFR), cytokeratin 5/6, cytokeratin 14, E-Cadherin, P53, and Ki67 on tissue microarrays that were assembled from 263 primary breast carcinomas of patients who were all known to have developed metastatic disease. Among 263 patients, median time to develop distant metastasis was 30 months (0-15.3 years) with 75.8% of these metastases occurring in the first 5 years following the initial treatment of the primary

tumor. Median overall survival times for the patients with ER-negative/HER2-negative, HER2-positive, ER-positive/HER2-negative/Ki67high, and ER-positive/HER2-negative/Ki67low tumors were 27, 52, 72, 76 and 79 months, respectively. Bone was found to be the most common site of distant metastasis (70.6%), followed by liver (54.5%) and lung (31.4%). The development of visceral metastases was observed in 81% of patients with ER-negative/HER2-negative tumor, 77.4% of with HER-positive, 76.9% of ER-positive/HER2-negative/Ki67low tumor and 75.7% of patients with ER-positive/HER-negative/Ki67high tumor. Of patients with ER+/HER2-/Ki67high tumors 87.8%, with ER+/HER2-/Ki67low tumors 73.1%, with HER2-positive 69.8% and with ER-negative/HER2-negative tumors 55.2% developed bone metastases. In this study, we demonstrate that subtypes of breast carcinoma defined by immunohistiochemical staining for ER, PR and HER2 are strongly correlated to metastatic behavior, in terms of site-specific metastasis, early/late metastasis and survival outcomes. Knowledge of these associations may aid in selection of treatment and follow-up options in metastatic breast cancer.

Subsequently, in **chapters 4** and **5**, we analyzed the gene expression profiles of primary breast carcinomas and their association with metastatic behavior and survival outcomes in search for gene expression signatures for bone- and visceral organ-specific metastases. In these chapters, we analyzed the gene expression profiling data generated from 157 primary breast tumors of patients with distant metastases. The correlation of the generated data with metastatic behavior revealed that 80.5% of luminal-type tumors developed bone metastasis, compared to 55.6 and 41.7% of HER2-like and basal-type tumors. The occurrence of visceral organ metastasis was observed in 87.5% of basal-type tumors, 77.8% of HER2-like tumors and 70.4% of luminal-type tumors. Survival analyses also showed that luminal-type tumors had longer metastasis specific and overall survival times compared to HER2-like and basal type tumors. Furthermore, the analyses led to the identification of differentially expressed genes between tumors with subsequent development of bone/visceral metastases and the ones without bone/visceral metastases. In addition, we report two novel gene expression signatures, respectively a 15-gene and a 14-gene expression signature associated with development of bone and visceral organ metastases. The 15-gene signature for bone-specific metastasis performed better in terms of identifying the tumors that developed bone metastases compared to other bone metastasis-specific

gene signatures from the literature. This novel gene expression signature, which was identified by using a supervised clustering method, was also found to predict the tumor with a higher likelihood of developing bone metastasis in ER-positive as well as ER-negative tumor groups. The 14-gene visceral-metastasis signature was found to be significantly related to the development of visceral metastases in the training and the independent data sets. In addition, the 14-gene signature was also closely related to the survival status of the patients. Considering the differences in the survival outcomes among the tumors with different organ site metastases, the gene expression signatures, such as the ones identified in the abovementioned studies, are potential tools to identify patients with a likelihood of developing bone or visceral metastasis.

Chapter 6 describes a study of the association between gene expression profiling patterns of primary breast cancer and chemotherapy response in the metastatic setting. Our results reveal no significant correlation between the molecular subtypes, tumor grade, lymph node status, and chemotherapy response in the metastatic setting. Response to chemotherapy was noted in 41.6% of the patients receiving first-line chemotherapy and in 21.8% of patients receiving second-line chemotherapy in the metastatic setting. Patients with HER2-positive tumors appeared to show better response to anthracyline-containing regimens compared to other subgroups of tumors. With the help of a supervised classification approach, a classifier composed of 14-genes to predict chemotherapy response in the metastatic setting was identified. This 14-gene classifier predicted the response successfully to first and second-line chemotherapy in the training set. However, we could not further validate this 14-gene predictor for chemotherapy response in an independent data set with available data on chemotherapy response in the neoadjuvant setting. Based on our analyses, we suggest that response to chemotherapy may differ in the neoadjuvant setting and the metastatic setting.

In **Chapter 7,** we studied the link between epithelial-mesenchymal transition (EMT) and metastatic breast cancer. EMT-status was assessed using a previously reported 130-gene-EMT-core signature and compared to the characteristics of metastatic breast cancer. Contrary to the expectations based on current literature, our results failed to indicate a direct association between EMT-status of the primary tumor and the pattern of metastasis and response to chemotherapy. Moreover, we found that

non-basal type tumors were associated with activated EMT-status. Additionally, we used immunohistochemical analyses to capture the EMT-status of a tumor, showing a significant correlation between immunostaining for NAT1 and TWIST1 and EMT-status.

In **chapter 8,** the results that were generated in this thesis are discussed. In conclusion, this thesis reveals that the traits that have been identified through gene expression profiling of primary breast tumors, are associated with metastatic behavior of breast cancer.

References

1. Russnes HG, Vollan HKM, Lingjaerde OC, Krasnitz A, Lundin P, Naume B, Sorlie T, Borgen E, Rye IH, Langerod A *et al*: Genomic architecture characterizes tumor progression paths and fate in breast cancer patients. *Sci Transl Med* 2010, 2(38):38ra47.

NEDERLANDSE SAMENVATTING

In dit proefschrift worden studies beschreven die als doel hebben het (klinische) gedrag van gemetastaseerde mammacarcinomen beter te begrijpen met behulp van genexpressieprofilering van de primaire tumor.

In **hoofdstuk 1** wordt een introductie gegeven over de klinische behandeling van gemetastaseerde mammacarcinomen en over de toepassing van genomisch onderzoek bij borstkanker. Daarnaast wordt onderbouwd waarom de studies in dit proefschrift zijn uitgevoerd met een uiteenzetting van de opzet van het proefschrift.

In **hoofdstuk 2** geven we onze visie op genomische veranderingen in mammacarcinomen en hoe deze veranderingen mogelijk kunnen worden toegepast in de klinische praktijk. Daarbij geldt een publicatie van *Russness* et al, waarin de auteurs twee algoritmes hebben ontwikkeld om de genomische complexiteit van de tumoren in kaart te brengen, als leidraad om de problemen van correlatie van genetische veranderingen aan klinische en pathologische bevindingen te bediscussiëren. Zij gebruiken genomische veranderingen om de tumoren te subclassificeren op basis van genoomwijde DNA-copynumbervariatie. De translatie van genomische veranderingen naar klinische behandeling kent zijn uitdagingen, doordat de meeste studies die zijn uitgevoerd op dit gebied sterk verschillen in opzet en de analyse van de data complex is, maar ook door de ingewikkelde biologie van het mammacarcinoom en de lage frequentie van voorkomen van de meeste genomische veranderingen. Samengevat lijkt de integratie van clinicopathologische karakteristieken met genetische veranderingen een veelbelovende te kunnen krijgen bij het verbeteren van de behandeling van mammacarcinoompatiënten.

In **hoofdstuk 3** worden histomorfologische kenmerken van primaire mammacarcinomen in kaart gebracht met als doel om de latere geassocieerde metastasen, en de metastasevrije overleving, te onderzoeken. In deze retrospectieve studie wordt gebruik gemaakt van immunohistochemische kleuringen voor ER, PR, HER2, EGFR, CK5/6, CK14, E-Cadherine, TP53 en Ki67 op weefsel-microarrays samengesteld uit weefsel van 263 primaire mammacarcinomen waaruit in de loop van de tijd afstandsmetastasen

zijn ontstaan. Bij deze 263 patiënten was de mediane tijd totdat deze metastasen ontstonden 30 maanden (0-15.3 jaar), waarbij 75.8% van deze metastasen is ontstaan binnen 5 jaar na de initiële behandeling van de primaire tumor. De mediane overleving van de patiënten met ER-negatieve/HER2-negatieve, HER2-positieve, ER-positieve/HER2-negatieve/hoge Ki67 en ER-positieve/HER2-negatieve/lage Ki67 tumoren is respectievelijk 27, 52, 72, 76 en 79 maanden. Het skelet was de meest frequent voorkomende locatie van de afstandsmetastasen (70.6%), gevolgd door de lever (54.5%) en longen (31.4%). Viscerale metastasen ontwikkelden zich in 81% van de patiënten met ER-negatieve/HER2-negatieve primaire tumoren, 77.4% van de HER2-positieve tumoren, 76.9% van de ER-positieve/HER2-negatieve/lage Ki67 tumoren en 75.7% van de patiënten met ER-positieve/HER-negatieve/hoge Ki67 tumoren. Daarnaast ontwikkelden 87.8% van de patiënten met ER-positieve/HER2-negatieve/hoge Ki67 tumoren, 73.1% van de patiënten met ER-positieve/HER2-negatieve/lage Ki67 tumoren, 69.8% van de patiënten met HER2-positieve tumoren en 55.2% van de patiënten met ER-negatieve/HER2-negatieve tumoren botmetastasen. Met deze studie tonen we aan dat verschillende subtypes van het mammacarcinoom, gedefinieerd door immunohistochemische kleuringen voor ER, PR en HER2, sterk correleren aan metastatisch gedrag, en meer specifiek aan locatiegebonden metastasen, vroeg/laat optredende metastasen en overleving. Kennis over deze associaties kan behulpzaam zijn bij de behandelkeuze en vervolgstrategieën bij patiënten met een gemetastaseerd mammacarcinoom.

Vervolgens analyseren we in **hoofdstukken 4** en **5** de genexpressieprofielen van primaire mammacarcinomen en de associatie van deze profielen met metastatisch gedrag en overleving, op zoek naar specifieke genexpressieprofielen die correleren met botmetastasen of andere (viscerale) afstandsmetastasen. In deze hoofdstukken worden data van genexpressieprofielen geanalyseerd die zijn gegenereerd uit weefsel van 157 primaire mammacarcinomen van patiënten met afstandsmetastasen. De studie laat zien dat 80.5% van de luminaaltype tumoren botmetastasen ontwikkelt, in vergelijking tot de 55.6 en 41.7% van de HER2-type en basaaltype tumoren. Viscerale afstandsmetastasen worden gezien in 87.5% van de basaaltype tumoren, 77.8% van de HER2-type tumoren en 70.4% van de luminaaltype tumoren. Overlevingsanalyses tonen daarnaast dat de luminaaltype tumoren een langere metastasevrije overleving en algemene overleving hebben dan HER2-type en basaaltype tumoren. Bovendien zijn er

differentieel tot expressie komende genen geïdentificeerd tussen tumoren die uiteindelijk wel afstandsmetastasen genereerden en tumoren zonder afstandsmetastasen in de follow-up. Daarbij rapporteren we twee nieuwe specifieke genexpressieprofielen, één met 15 genen en één met 14 genen, die geassocieerd zijn met het ontwikkelen van respectievelijk botmetastasen en viscerale metastasen, geïdentificeerd middels een gesuperviseerde hiërarchische clustermethode. Het specifieke 15-genenprofiel voor botmetastasen is in het cohort beter voorspellend voor het ontwikkelen van botmetastasen dan andere genexpressieprofielen waarover is gerapporteerd in de literatuur, zowel in de groep van ER-positieve als ER-negatieve tumoren. Het specifieke 14-genenprofiel voor viscerale metastasen is significant gerelateerd aan het ontwikkelen van viscerale metastasen in zowel de trainingsset als de onafhankelijke dataset. Daarnaast is dit 14-genenprofiel ook duidelijk gerelateerd aan de overleving van de patiënten. Aangezien de overleving van patiënten met metastasen naar verschillende organen wisselt, kunnen genexpressieprofielen (zoals de profielen die in bovenbeschreven studies zijn geïdentificeerd) potentieel bijdragen aan het identificeren van patiënten die meer waarschijnlijk botmetastasen of viscerale afstandsmetastasen zullen ontwikkelen.

Hoofdstuk 6 beschrijft een studie naar de associatie tussen genexpressieprofielpatronen van primaire mammacarcinomen en respons op chemotherapie van deze tumoren in het geval van afstandsmetastasen. Onze resultaten laten zien dat er geen significante correlatie is tussen moleculaire subtypes, tumorgraad, lymfklierstatus en respons op chemotherapie in een gemetastaseerde setting. Respons op chemotherapie werd gezien bij 41.6% van de patiënten die eerstelijnschemotherapie kregen en bij 21.8% van de patiënten met tweedelijns chemotherapie. Patiënten met HER2-positieve tumoren lijken een betere respons te tonen op anthracyclinehoudende chemotherapieschema's in vergelijking met de andere tumorsubtypes. Met een benadering via gesuperviseerde classificatie is vervolgens een specifiek profiel van 14 genen geïdentificeerd. Dit specifieke 14-genenprofiel kon de respons op chemotherapie voor zowel de eerste- als tweedelijnstherapie succesvol voorspellen in de trainingsset. We konden dit predictieve 14-genenprofiel echter niet verder valideren bij onderzoek in een onafhankelijke dataset voor respons op neoadjuvante chemotherapie. Op basis van deze resultaten kan dus gesuggereerd worden dat de respons op chemotherapie tussen de neoadjuvante setting en de metastatische setting verschilt.

In **hoofdstuk 7** is de link tussen epitheliale-mesenchymale transitie (EMT) en het gemetastaseerde mammacarcinoom onderzocht. De EMT-status werd geëvalueerd met een specifiek 130-genen-EMT-genexpressieprofiel dat eerder is beschreven in de literatuur. De resultaten zijn vergeleken met de moleculaire subtypes van het mammacarcinoom en met klinische parameters. In tegenstelling tot wat wordt verwacht op basis van de literatuur, laten onze resultaten geen directe associatie zien tussen de EMT-status van de primaire tumor, het patroon van metastasering en de respons op chemotherapie. Wel tonen we aan dat de primaire tumoren die niet van het basale type zijn, zijn geassocieerd met een geactiveerde EMT-status. Als vertaalslag van het genexpressieprofiel passend bij EMT zijn immunohistochemische kleuringen verricht, waarbij een significante correlatie werd gevonden tussen immunohistochemische aankleuring met NAT1 en TWIST1 en EMT-status.

In **hoofdstuk 8** worden de resultaten die beschreven zijn in dit proefschrift bediscussieerd. Concluderend laat dit proefschrift zien dat eigenschappen van primaire mammacarcinomen die met genexpressieprofilering in kaart zijn gebracht behulpzaam kunnen zijn in het voorspellen van het metastatische gedrag van het mammacarcinoom.

Referenties

1. Russnes HG, Vollan HKM, Lingjaerde OC, Krasnitz A, Lundin P, Naume B, Sorlie T, Borgen E, Rye IH, Langerod A *et al*: Genomic architecture characterizes tumor progression paths and fate in breast cancer patients. *Sci Transl Med* 2010, 2(38):38ra47.

Appendix

Nawoord / Acknowledgement
About the author

NAWOORD / ACKNOWLEDGEMENT

Graag zou ik onderstaande personen, die allen in hun rol en op een eigen manier, een fantastische bijdrage aan dit proefschrift hebben geleverd vanuit de grond van mijn hart willen danken:

Alle patiëntes die direct of indirect hebben deelgenomen aan alle besproken onderzoeken

Marc J. van de Vijver

Hans Halfwerk

Hugo M. Horlings

A.D. Bins, N. Bijker, P.J. van Diest, E.J.T. Rutger, V.T.H.B.M. Smit, L.F.A. Wessels

Jan Koster

Gerrit K. Hooijer

Jelle Wesseling

Frederike Dijk

Lianne Koens

Sybren L. Meijer

Marion Zeeman, Agnes Brunsveld

Wim J.V. van Est

Alle medewerkers afdeling pathologie

&

Karin Peeters, Abbe Heijink, Louise van Nieuwenhoven

Annem, Teyzem, Istanbul' daki ailem

Timur ve Banu Savcı

...en mijn liefjes: Andras, Milan en Bente

ABOUT THE AUTHOR

C. Dilara Savcı Heijink was born in 1973 in Adana, Turkey. She received her medical degree in Ankara, Turkey. She obtained her residency training in anatomical pathology at Marmara University in Istanbul, followed by additional training at the Department of Pathology and Laboratory Medicine at Mayo Clinic in Rochester, USA. During her stay at Mayo Clinic, next to the surgical pathology training, she also participated in research activities at the experimental pathology division within the same department, under supervision of G. Vasmatzis and J.C. Cheville. In 2009, she moved to the Netherlands where she was acquainted with Prof. dr. van de Vijver. This prompted Dilara on the track which led to the writing of this thesis. Since September 2013, she has worked as a pathologist at the University of Amsterdam at the Academic Medical Center. Dilara is married to Andras Heijink. They currently live in Amstelveen and have two children, Milan and Bente.