ELSEVIER

European Journal of Cancer 40 (2004) 2667–2675

European Journal of Cancer

www.ejconline.com

Review

Molecular portraits of breast cancer: tumour subtypes as distinct disease entities☆

Therese Sørlie

Department of Genetics, Institute for Cancer research, The Norwegian Radium Hospital, Montebello, Oslo, Norway

Received 17 August 2004; accepted 21 August 2004
Available online 30 September 2004

Abstract

This Review describes advances in the characterisation of breast tumour phenotypes using DNA microarrays and the identification of five subtypes of breast cancer with significant clinical implications.

Keywords: Breast cancer; DNA microarrays; Gene expression patterns; Prognosis; Individualised treatment

1. Introduction

The emergence of the field "genomics" and recent advances in global molecular profiling techniques have revolutionised cancer genetics and hold the potential of also revolutionising clinical practice. The diverse entities that collectively are called "cancer" have in common that they arise as a result of accumulation of mutations, chromosomal instabilities and epigenetic changes; together these facilitate an increased rate of cellular evolution and damage that progressively impairs the cell's detailed and complex system of regulation of cell growth and death. Changes in gene activities are further influenced by the microenvironment within and in the vicinity of tumour cells as well as by exogenous factors, such as diet. When one combines all of these factors with inborn genetic variations among individuals, there is every kind of reason to expect tumours to display prodigiously diverse phenotypes. Microarray technologies, applied to the study of DNA, RNA, and protein profiles, as well as to the genome-wide distribution of epigenetic changes, such as DNA methylation, can be used to portray a tumour's detailed phenotype in its unique context. Systematic and detailed characterisation of tumours on a genomic scale can be correlated with clinical information and greatly enhance our understanding of the causes and progression of cancer, ability to discover new molecular markers and possibilities for therapeutic intervention. Eventually, advances in tumour portraiture will naturally lead to improved and individualised treatments.

Despite reduced mortality for breast cancer patients due to earlier diagnosis and implementation of adjuvant chemo- and hormone-therapies, breast cancer is still the most common cause of cancer death in women worldwide [1]. Breast tumours are heterogeneous and consist of several pathological subtypes with different histological appearances of the malignant cells, different clinical presentations and outcomes, and the patients show a diverse range of responses to a given treatment. Furthermore, breast tumour tissue also shows heterogeneity with respect to its microenvironment including specifically the types and numbers of infiltrating lymphocytes, adipocytes, stromal and endothelial cells. The cellular composition of tumours is a central determinant of both the biological and clinical features of an individual's disease.

This Review describes advances in the characterisation of breast tumour phenotypes and the identification

☆ This review is based on the EACR Young Cancer Researcher Award lecture presented at EACR-18 in Innsbruck, July 2004.

E-mail address: tsorlie@labmed.uio.no.

0959-8049/$ - see front matter
doi:10.1016/j.ejca.2004.08.021

of five subtypes of breast cancer with significant clinical implications.

2. Methodological considerations

Most published studies to date have used spotted cDNA arrays that were introduced by Schena and colleagues in 1995 [2], and the data presented herein, are based on such a platform. This method allows the use of two different fluorescent labels to distinguish, on the same spots, the abundance of gene-specific nucleic acid from two different samples. We have routinely used a reference strategy, in which transcripts extracted from the tumour sample were labelled with one fluorescent dye, whereas transcripts from a standard reference [3] were labelled with a different dye. Relative levels of the two transcripts were calculated by log transforming the ratio between the two fluorescent intensities (red to green).

The great value of microarray experiments when applied to the study of cancer is most clear when a large number of samples are analysed similarly and combined, so that variation in gene expression patterns across a large number of tumours can be investigated. For most of our analyses, hierarchical clustering algorithms have been applied to organise both genes and samples into meaningful groups based on similarity in their overall expression patterns [4]. As an alternative to this unsupervised approach, we have apply supervised methods, in which predictive models are built based on existing knowledge of the data [5–8].

3. Global molecular profiling of breast tumours

The phenotypic diversity of tumours is accompanied by a corresponding diversity in gene expression patterns that can be captured by DNA microarrays. Our earlier studies of 65 surgical specimens of human breast tissue from 42 individuals using microarrays representing 8102 genes showed that there was great molecular heterogeneity among the tumours, with multidimensional variation in the patterns of gene expression [9]. To help provide a framework for interpreting the variation in expression patterns observed in the tumour samples, 17 cultured cell lines were also characterised. Clusters of genes with coherent expression patterns could be related to specific features of biological variation among the samples, for example variation in proliferation rates and activation of the interferon-regulated signal transduction pathway. Furthermore, clusters of co-expressed genes were identified whose expression patterns derived from different cell types within the grossly dissected tumours, including the two epithelial cell types, luminal and basal epithelium, as well as stromal cells and lymphocytes. These findings were corroborated by immunohistochemistry.

A unique quality of this study was the availability of tumour pairs; 20 tumours were sampled twice as part of a larger prospective study on locally advanced breast cancer (T3/T4 and/or N2 tumours) [10–12]. Following an open surgical biopsy to obtain the ''before'' sample, each of these patients was treated with doxorubicin for an average of 15 weeks, followed by resection of the remaining tumour, in which case the ''after'' sample was obtained. In addition, primary tumours from two patients were paired with a lymph node metastasis from the same individual. A striking result observed was that most of the tumour pairs (15 of the 20 ''before and after'' pairs and both primary tumour/lymph node metastasis pairs) clustered together on terminal branches in the accompanying dendrogram (Fig. 1). That is, despite the potential confounding effects of an interval of 15 weeks, cytotoxic chemotherapy and different sample preparations, independent samples taken from the same tumour were in most cases recognisably more similar to each other than either was to any of the other samples. This implied that every tumour is unique and has a distinctive gene expression 'signature' or 'portrait'. It also implied that the type and numbers of non-epithelial cells in tumours is a remarkably consistent and enduring feature of each individual tumour. Breast tumours thus appear to be very diverse, but are not internally heterogeneous, when millions of cells are sampled.

To explore the possibilities for refining distinctions among subtypes of breast tumours using microarrays, we took advantage of the paired tumour samples. The specific features of a gene expression pattern that are to be used as the basis for classifying tumours should be similar in any sample taken from the same tumour, and they should vary among different tumours. The paired samples therefore provided a unique opportunity for a deliberate and systematic search for genes whose expression levels reflected such intrinsic characteristics of the tumours. Using well-measured expression data from the paired samples, a subset of genes termed ''intrinsic gene-set'' was selected that consisted of genes with significantly greater variation in expression between different tumours than between paired samples from the same tumour. The rationale behind this list of genes is that it is enriched for those genes whose expression patterns were characteristic of each tumour as opposed to those that varied as a function of tissue sampling, and hence, would be ideally suited for classification.

4. Identification of tumour subtypes

Our analyses have since been extended to include more samples and an intrinsic gene-set of 540 genes se-

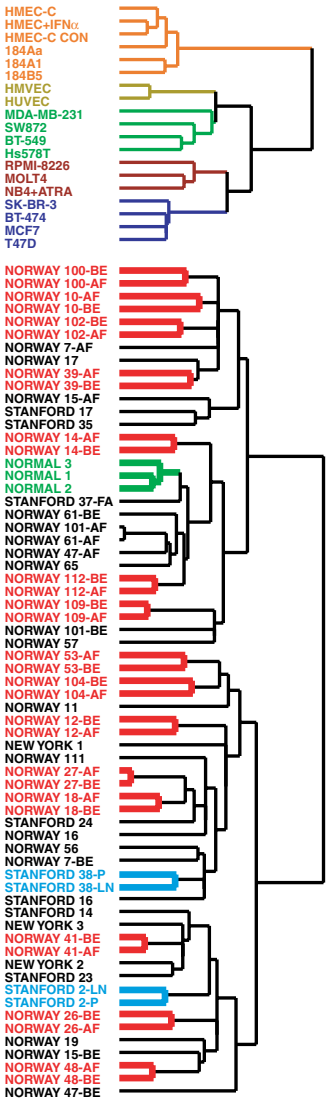

Fig. 1. Dendrogram representing similarities in the expression patterns between experimental samples. In the right part all "before and after" chemotherapy pairs that were clustered on terminal branches are highlighted in red; the two primary tumour/lymph node metastasis pairs are highlighted in light blue; the three clustered normal breast samples are highlighted in light green. In the left part the branches representing the four breast luminal epithelial cell lines are displayed in dark blue; breast basal epithelial cell lines are displayed in orange, the endothelial cell lines in dark yellow, the mesenchymal-like cell lines in dark green, and the lymphocyte-derived cell lines in dark red. Reprinted by permission from Nature 406, 17 August, 747–752, copyright 2000 Macmillan Publishers Ltd.

lected from expression data of 45 tumour pairs (including two primary-lymph node pairs) and the approximately 8000 genes that were common for all experiments [13,14]. Altogether, 122 tissue samples were included in the analysis, 115 carcinomas and seven non-malignant tissues. Most of the tumours were sampled as part of two independent studies evaluating response to chemotherapy of locally advanced breast cancer in an neoadjuvant setting. From the first cohort of patients treated with doxorubicin monotherapy [11], 55 tumour samples were analysed, and from the second cohort of patients treated with 5-fluorouracil and mitomycin C (FUMI) [12], 34 tumour samples were analysed. The remaining 26 samples were primary tumour specimens collected either at Stanford or in Norway.

The overall expression patterns of these intrinsic genes showed that the main distinction was between tumours that expressed genes characteristic of luminal epithelial cells, including the oestrogen receptor (ER) and those that were negative for these genes (Fig. 2). Among the luminal epithelial-type tumours, the largest group (termed luminal subtype A, dark blue branches) demonstrated the highest expression of ER, oestrogen-regulated protein LIV-1, the transcription factors hepatocyte nuclear factor 3, alpha (*HNF3A*), X-box binding protein 1 (*XBP1*) and GATA-binding protein 3 (*GATA3*) (Fig. 2G). The second, smaller group of tumours (termed luminal subtype B, light blue branches) showed low to moderate expression of the luminal-specific genes including the ER cluster mentioned above, but was further distinguished from luminal subtype A by the high expression of a novel set of genes such as *GGH*, *LAPTMB4*, *NSEP1* and *CCNE1*, but whose coordinated function is unknown (Fig. 2D). In addition, a basal epithelial-like subtype (red branches) was characterised by high expression of *KRT 5* and *KRT17*, annexin 8, *CX3CL1* and *TRIM29* and which was completely negative for the luminal/ER cluster of genes (Fig. 2E), while the ERBB2+ subtype (pink) was characterised by high expression of several genes in the ERBB2 amplicon at 17q22.24 including *ERBB2*, *GRB7* and *TRAP100* (Fig. 2C). A normal breast tissue-like group (green branches) was identified that showed the highest expression of many genes known to be expressed by adipose tissue and other non-epithelial cell types (Fig. 2F). These tumours also showed strong expression of basal epithelial genes and low expression of luminal epithelial genes. However, it is unclear at this point whether these tumours represent poorly sampled tumour tissue or a distinct, clinically important group.

In conclusion, the inherent properties of the tumours seem to be sustained throughout chemotherapy as well as between a primary tumour and its lymph node metastasis, and could be represented by a relatively small number of genes whose variation in expression formed a platform for classification.

5. Clinical implications

A major goal in the field of oncogenomics is to try to answer the clinically important questions about which tumours will behave aggressively, which tumours will remain dormant, which patients do and do not require systemic therapy and what type of drugs should be used.

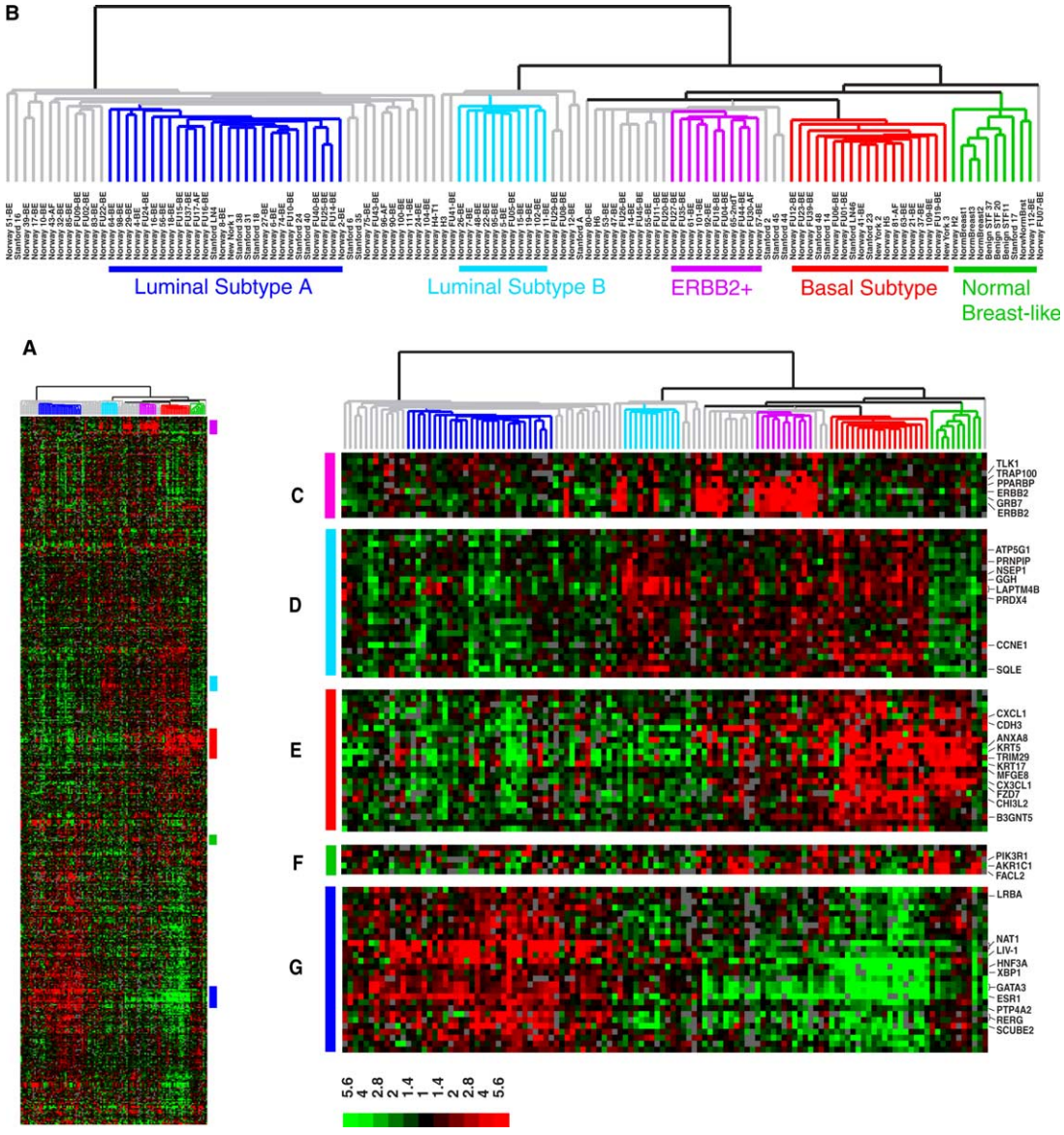

Fig. 2. Hierarchical clustering of 115 tumour tissues and seven non-malignant tissues using the "intrinsic" gene-set. (A) A scaled-down representation of the entire cluster of 540 genes and 122 tissue samples based on similarities in gene expression. (B) Experimental dendrogram showing the clustering of the tumours into five subgroups. Branches corresponding to tumours with low correlation to any subtype are shown in grey. (C) Gene cluster showing the *ERBB2* oncogene and other co-expressed genes. (D) Gene cluster associated with luminal subtype B. (E) Gene cluster associated with the basal subtype. (F) A gene cluster relevant for the normal breast-like group. (G) Cluster of genes including the oestrogen receptor (*ESR1*) highly expressed in luminal subtype A tumours. Scale bar represents fold-change for any given gene relative to the median level of expression across all samples. PNAS, July 8, 2003, vol. 100, no. 14, 8418–8423. Copyright (2003) National Academy of Sciences, USA.

To investigate whether the five different tumour subgroups identified by hierarchical clustering may represent clinically distinct groups of patients, univariate survival analyses comparing the subtypes with respect to overall survival (OS) and relapse-free survival (RFS) were performed. For these analyses, only the patients who were enrolled in the two prospective studies on locally advanced disease were included. The Kaplan–Meier curves based upon four subclasses (excluding the normal-like group) showed a highly significant difference in OS between the patients belonging to the different subclasses (Fig. 3, $P < 0.01$). Specifically, the basal-like and ERBB2+ subtypes were associated with the shortest survival time. Overexpression of the ERBB2 oncoprotein is a well-known prognostic factor associated with poor survival in breast cancer, which was also found for the ERBB2+ group defined in this study.

Perhaps the most intriguing result is the considerable difference in outcome observed between tumours classified as luminal A versus luminal B; both classes contained tumours that expressed the oestrogen receptor to some extent. It should be noted that the prognostic value of the expression-based classes were improved relative to the standard stratification based on ER protein

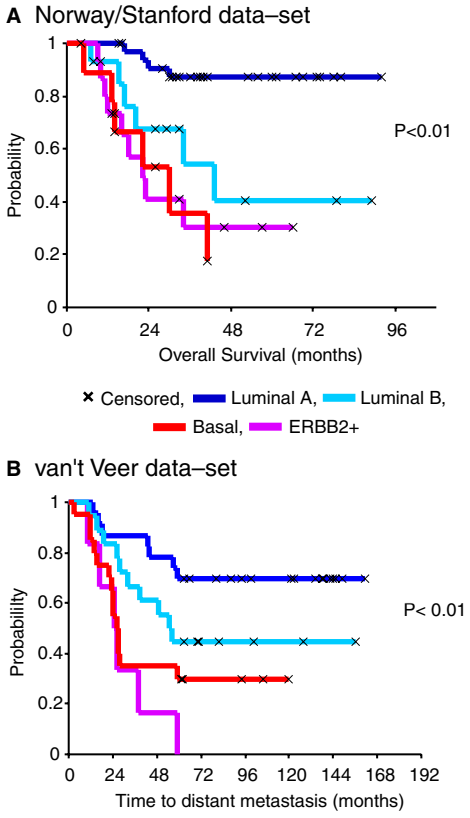

Fig. 3. Kaplan–Meier analysis of disease outcome in two patient cohorts. (A) Overall survival for 72 patients with locally advanced breast cancer in the Norway cohort. The normal-like tumour subgroups were omitted from both data-sets in this analysis. (B) Time to development of distant metastasis in the 97 sporadic cases from van't Veer and colleagues. PNAS, July 8, 2003, vol. 100, no. 14, 8418–8423. Copyright (2003) National Academy of Sciences, USA.

expression (data not shown). The luminal subtype B tumours might represent a clinically distinct group with a different and worse disease course, in particular with respect to relapse. Perhaps this subtype reflects a group of patients who will not benefit from adjuvant tamoxifen despite positive receptor values. (All patients in this study were given adjuvant tamoxifen provided they had positive hormone receptor values.) The potential clinical significance of this molecular subtype is further highlighted by the similarities in expression of some of the luminal B genes with the ER-negative tumours in the basal-like and ERBB2+ subtypes, which suggests that high level of expression of these genes is associated with poor disease outcomes.

6. Validation of subtypes in independent data-sets

Following the identification of these five subtypes, validation in independent data-sets is essential. The robustness of the tumour subtypes was tested by conducting a similar analysis of a data-set published by van't Veer and colleagues [15]. Gene expression data ($\log_{10}$ ratios) were available for 24480 genes in a set of 117 tumours from young breast cancer patients with early stage disease and for whom data on time to development on distant metastases were available. Hierarchical clustering was used, exactly as described for our own data-set, to display the expression patterns of 461 of the "intrinsic" genes in the 97 tumour samples that were obtained from patients diagnosed with sporadic cancer (excluding the 20 familial tumours in this stage of the analysis) [13].

As in the Norway/Stanford data, the clearest discrimination was between tumours that expressed genes in the luminal A/ER cluster at high levels and the tumours that were negative for these genes and exhibited expression profiles characteristic of either the basal or the ERBB2+ subtypes (see Ref [13], Fig. 2). The luminal A/luminal B distinction was also seen, but was less clear in this data-set. The basal subtype was also in this data-set the most homogeneous cluster of tumours, as reflected by high correlation of the samples in this group and the deep branch separating it from the other subtypes.

To investigate whether the subtypes were associated with a significant difference in outcome also in this patient cohort, we performed a univariate Kaplan–Meier analysis with time to development of distant metastasis as a variable. As shown in Fig. 3B, the probability of remaining disease-free was significantly different between the subtypes; patients with luminal A type tumours lived considerably longer before they developed metastatic disease, while the basal and ERBB2+ groups showed much shorter disease-free time intervals. The discrimination in outcome between the luminal A and B subtypes were less distinct, which could be due to differences in tumour stage, patient age or other parameters, but it may reflect the fact that these patients were largely an untreated population, substantiating our hypothesis of an endocrine treatment interaction. While the methodological differences prevent a definitive interpretation, it is notable that the order of severity of clinical outcome associated with the subtypes is similar in the two dissimilar cohorts. What this further indicates, is that the substantial differences in the characteristics of the patients are less important determinants of tumour expression phenotypes than intrinsic biology. Further, we found strong evidence for the universality of a distinction between basal-like and luminal-like subtypes in an additional independent data-set, produced on yet another technology platform [13].

7. Molecular characteristics of the subtypes

7.1. Proliferation cluster

The previously defined proliferation cluster is a group of genes whose levels of expression correlate with

cellular proliferation rates [9,16]. Expression of this cluster of genes varied widely among the tumour samples, and was generally well correlated with the mitotic index. Genes encoding two generally used immunohistochemical markers of cell proliferation, Ki-67 and proliferating cell nuclear antigen (PCNA), were also included in this cluster. More than half of the genes in the proliferation cluster were shown to be cell cycle-regulated when the patterns of expression for these genes were analysed in synchronised HeLa cell cultures [17]. To investigate the expression of these genes in relation to the subtypes, expression data were extracted and while the genes were clustered using average-linkage clustering, the samples were ordered according to the subtypes as presented in [13]. What could be seen from this analysis, was that the basal and the luminal subtype B both highly expressed these proliferation-associated genes, whereas luminal A, the normal-like and to some extent, the ERBB2+ subtype, were mostly negative for the expression of this cluster of genes. This may indicate large differences in the amount of cycling cells among the tumours. The fact that the ERBB2+ subtype tumours were less proliferative, despite showing overexpression of an oncogene and being a poor prognostic group was somewhat surprising. However, this underlines the very strong influence of the ERBB2 amplicon on the expression patterns of these tumours; no other distinct molecular signature protrudes at the transcriptional or at the genomic level (assessed by microarray comparative genomic hybridisation, aCGH).

7.2. TP53

The specific expression patterns and profound differences in the tumour subtypes most likely reflect different alterations in molecular pathways within the tumour cells. P53 plays an important role in directing cellular responses to genotoxic damage and regulates the activation of downstream genes that are involved in apoptosis, cell cycle arrest and DNA repair [18–20]. Previous studies have shown that mutations in the *TP53* gene predict a poor prognosis and are associated with a poor response to systemic therapy [21–23]. Even though *TP53* itself is not differentially expressed across this sample set, it is likely that P53 has a significant role in shaping the gene expression patterns in the various tumour subtypes. The coding region of the *TP53* gene (exons 2–11) was screened for mutations in all but eight tumour samples (not including benign tumours) [14]. The distribution of the frequency of mutations among the different subclasses was significantly different ($P < 0.001$, two-sided). Luminal subtype A contained only 16% mutated tumours, whereas the luminal B, ERBB2+ and basal-like subclasses had 71%, 86% and 75% *TP53*-mutated tumours, respectively. The finding of *TP53* mutations in tumours that simultaneously expressed the *ERBB2* gene at high levels, supports previous observations of an interdependent role for *TP53* and *ERBB2* [12,24].

To more directly investigate the effect of *TP53* mutations on the expression patterns in these tumours, we searched for genes whose expression was consistently different between *TP53*-mutated and *TP53*-wild-type tumours (two-class SAM – significance analysis of microarrays [8]). A list of 158 genes significantly correlated with *TP53* status was selected that showed a median false discovery rate of less than one percent. The genes were clustered while the samples were ordered according to the subtypes (Fig. 4). As expected, many of the genes that were highly expressed in the basal-like and luminal B type tumours (high frequency of mutations), are cell cycle-regulated genes, such as *PLK1*, *CCNA2*, *STK6* and *MAPK13* [17]. Two of the genes that are co-expressed with the *ERBB2* oncogene, *GRB7* and *STARD3*, were also found, but not the *ERBB2* gene itself. Interestingly, many of the genes whose high expression and specific pattern distinguished luminal subtype B from subtype A, such as *GGH*, *LAPTMB4* and *MYBL2*, were highly expressed in the *TP53*-mutated tumours. Among the genes highly expressed in tumours with a wild-type *TP53* gene, most of the genes in the luminal/ER+ cluster (*ER*, *GATA*-3, *RERG*, etc.) were found, and this recapitulates the low mutation rate of the luminal subtype A. Further studies are needed to determine which of the genes are direct targets of *TP53* and which are only associated with a particular expression phenotype.

7.3. BRCA1

Germ-line mutations in the *BRCA1* gene predispose to early onset breast and ovarian cancer. A previous study reported significantly different gene expression profiles of tumours from *BRCA1*- versus *BRCA2*-carriers [25]. In the data-set produced by van't Veer and colleagues [15], tumours from 18 carriers of *BRCA1* mutations and two carriers of *BRCA2* mutations, were also analysed. When we included these 20 tumours along with the 97 sporadic samples in the clustering analysis, we saw little difference in the overall pattern, except for the striking result that all the tumours from patients carrying *BRCA1* mutations fell within the basal subgroup [13]. This indicates that a mutation in the *BRCA1* gene predisposes for the basal tumour subtype, which is associated with a lack of expression of the oestrogen receptor and poor prognosis. As also reported previously, *BRCA1*-associated breast cancers are usually highly proliferative, *TP53*-mutated and lack expression of *ESR1* and *ERBB2* [26,27]. Recent studies have found that BRCA1 interacts with and regulates the activity of the ER and the androgen receptor [28], thereby providing a link between BRCA1 and hormone-related can-

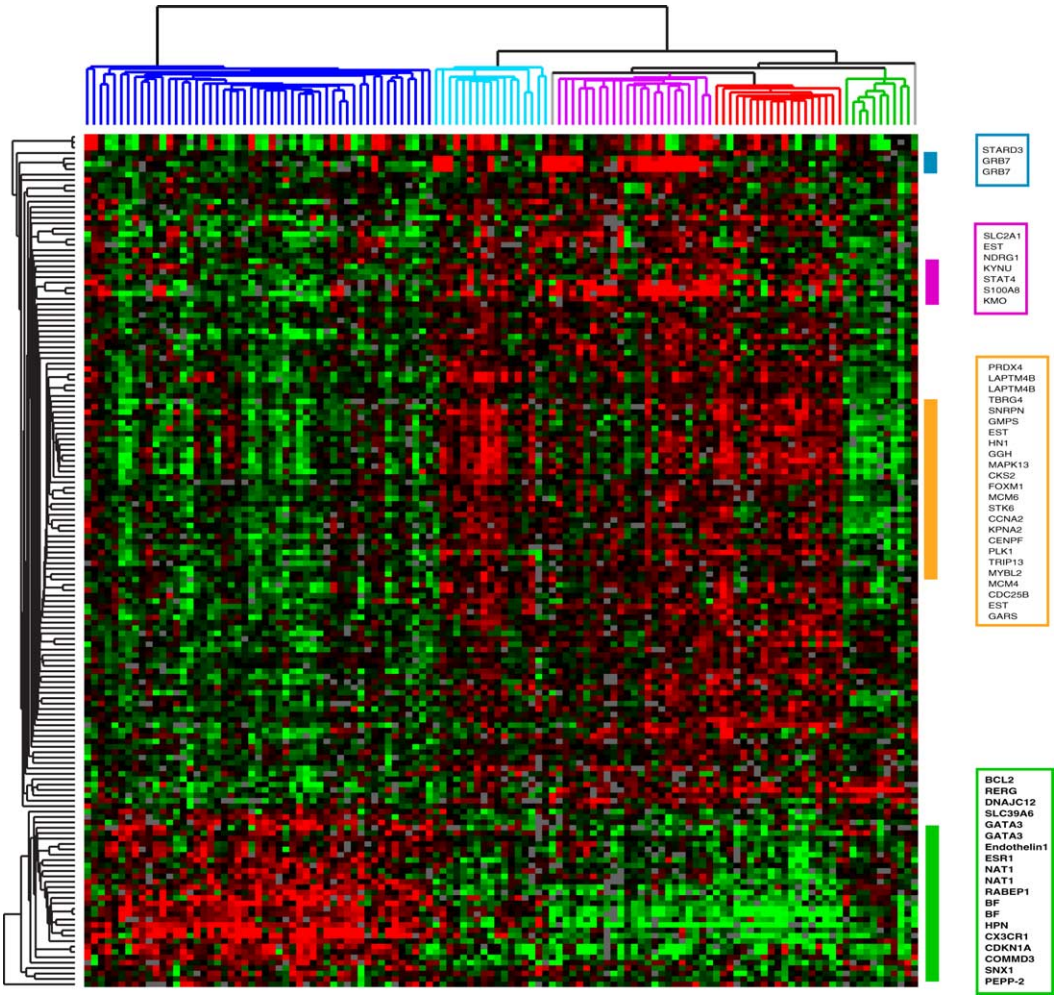

Fig. 4. Cluster of genes associated with TP53 mutational status. Hierarchical clustering of 158 genes whose expression were significantly correlated to the mutation status of the *TP53* gene (SAM, False Discovery Rate <1). Only tumour tissues from the two prospective studies from Norway were included in the cluster analysis. Tumours were ordered according to the subtypes as presented in Fig. 2.

cers. Furthermore, wild-type, but not mutant BRCA1, was able to inhibit oestradiol-induced activation of extracellular signal-related kinase (ERK), as well as the synthesis of cyclins D1 and B1, the activity of cyclin-dependent kinases, Cdk4 and CDK1, and G_1/S and G_2/M cell cycle progression [29].

8. Genome-wide copy number changes

The power of microarrays is also illustrated in their broad range of utility; the same cDNA microarrays can be used to investigate both the structural and the expressed genome. A genome-wide array CGH analysis of most of the tumours described herein, is underway, with a specific emphasis on the distribution of genomic alterations across the subtypes. This parallel measurement of copy number alterations enables us to investigate the global impact of widespread DNA copy number alteration on gene expression in tumour cells. Such studies have already indicated that the overall patterns of gene amplification and elevated gene expression are quite concordant; i.e., a significant fraction of highly amplified genes appear to be correspondingly highly expressed [30,31].

9. Conclusion/future perspectives

We have identified five different subtypes of breast cancer by their variation in gene expression patterns from a set of approximately 500 genes. These genes were selected from several thousands by their robust and steady expression between pairs of samples taken from the same tumour separated by 15 weeks of chemotherapy treatment. A striking conclusion is the stability, homogeneity and uniqueness of the "molecular portraits" provided by the variation in these gene expression patterns. We infer that these portraits faithfully represent the *tumour* itself, and not merely the particular tumour

sample, because we could recognise the distinctive expression pattern of a tumour in independent samples.

Survival analyses showed significantly different outcome for patients belonging to the various subtypes, emphasising the clinical relevance of stratification by such molecular profiling. Differences in mutation frequencies of cancer genes such as *TP53* and *BRCA1* highlights the important roles for these genes and associated pathways in determining the gene expression patterns of the various tumours. In addition, these data allow us to unveil previously unknown genes that may be involved in tumorigenesis. Knowledge about these genes, their products and how they correlate to different aspects of the disease may improve our understanding of the biology of breast cancer and lead to new prognostic and predictive markers. Similar variation in expression of a set of genes across a set of samples indicates similar means of regulation and function, and hence, provides a powerful way of identifying novel biologically important genes that could be used as markers and targets for therapy. The strength of this method lies in its ability to identify clusters of genes that in a unique combination will distinguish subgroups of disease and predict outcome or treatment response. This multi-gene approach will undoubtedly be superior to the standard clinical markers currently in use.

Conflict of interest statement

None declared.

References

1. Sasco AJ. Breast cancer and the environment. *Horm Res* 2003, **60**(Suppl. 3), 50.
2. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray [see comments]. *Science* 1995, **270**, 467–470.
3. Novoradovskaya N, Whitfield ML, Basehore LS, Novoradovsky R, Pesich R, Usary J, *et al*. Universal reference RNA as a standard for microarray experiments. *BMC Genom* 2004, **5**, 20.
4. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998, **95**, 14863–14868.
5. Brown MP, Grundy WN, Lin D, Cristianini N, Sugnet CW, Furey TS, *et al*. Knowledge-based analysis of microarray gene expression data by using support vector machines. *Proc Natl Acad Sci USA* 2000, **97**, 262–267.
6. Khan J, Wei JS, Ringner M, Saal LH, Ladanyi M, Westermann F, *et al*. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat Med* 2001, **7**, 673–679.
7. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci USA* 2002, **99**, 6567–6572.
8. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 2001, **98**, 5116–5121.
9. Perou CM, Sorlie T, Eisen MB, Van de RM, Jeffrey SS, Rees CA, *et al*. Molecular portraits of human breast tumours. *Nature* 2000, **406**, 747–752.
10. Aas T, Borresen AL, Geisler S, Smith-Sorensen B, Johnsen H, Varhaug JE, *et al*. Specific P53 mutations are associated with de novo resistance to doxorubicin in breast cancer patients. *Nat Med* 1996, **2**, 811–814.
11. Geisler S, Lønning PE, Aas T, Johnsen H, Fluge Ø, *et al*. Influence of TP53 gene alterations and c-erbB2 expression on the response to treatment with doxorubicin in locally advanced breast cancer. *Cancer Res*.
12. Geisler S, Borresen-Dale AL, Johnsen H, Aas T, Geisler J, Akslen LA, *et al*. TP53 gene mutations predict the response to neoadjuvant treatment with 5-fluorouracil and mitomycin in locally advanced breast cancer. *Clin Cancer Res* 2003, **9**, 5582–5588.
13. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, *et al*. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 2003, **100**, 8418–8423.
14. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, *et al*. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001, **98**, 10869–10874.
15. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, *et al*. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002, **415**, 530–536.
16. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, *et al*. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* 2000, **24**, 227–235.
17. Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE, *et al*. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol Biol Cell* 2002, **13**, 1977–2000.
18. Balint EE, Vousden KH. Activation and activities of the p53 tumour suppressor protein. *Br J Cancer* 2001, **85**, 1813–1823.
19. Guimaraes DP, Hainaut P. TP53: a key gene in human cancer. *Biochimie* 2002, **84**, 83–93.
20. Vogelstein B, Lane D, Levine AJ. Surfing the p53 network. *Nature* 2000, **408**, 307–310.
21. Bergh J, Norberg T, Sjogren S, Lindgren A, Holmberg L. Complete sequencing of the p53 gene provides prognostic information in breast cancer patients, particularly in relation to adjuvant systemic therapy and radiotherapy. *Nat Med* 1995, **1**, 1029–1034.
22. Berns EM, Foekens JA, Vossen R, Look MP, Devilee P, Henzen-Logmans SC, *et al*. Complete sequencing of TP53 predicts poor response to systemic therapy of advanced breast cancer. *Cancer Res* 2000, **60**, 2155–2162.
23. Borresen AL, Andersen TI, Eyfjord JE, Cornelis RS, Thorlacius A, Borg A, *et al*. TP53 mutations and breast cancer prognosis: particularly poor survival rates for cases with mutations in the zinc-binding domains genes chromosomes. *Cancer* 1995, **14**, 71–75.
24. Nakopoulou LL, Alexiadou A, Theodoropoulos GE, Lazaris AC, Tzonou A, Keramopoulos A. Prognostic significance of the co-expression of p53 and c-erbB-2 proteins in breast cancer. *J Pathol* 1996, **179**, 31–38.
25. Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, *et al*. Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 2001, **344**, 539–548.
26. Armes JE, Trute L, White D, Southey MC, Hammet F, Tesoriero A, *et al*. Distinct molecular pathogeneses of early-onset breast cancers in BRCA1 and BRCA2 mutation carriers: a population-based study. *Cancer Res* 1999, **59**, 2011–2017.
27. Grushko TA, Blackwood MA, Schumm PL, Hagos FG, Adeyanju MD, Feldman MD, *et al*. Molecular-cytogenetic analysis of

HER-2/neu gene in BRCA1-associated breast cancers. *Cancer Res* 2002, **62**, 1481–1488.
28. Fan S, Wang J, Yuan R, Ma Y, Meng Q, Erdos MR, *et al*. BRCA1 inhibition of estrogen receptor signaling in transfected cells. *Science* 1999, **284**, 1354–1356.
29. Razandi M, Pedram A, Rosen EM, Levin ER. BRCA1 inhibits membrane estrogen and growth factor receptor signaling to cell proliferation in breast cancer. *Mol Cell Biol* 2004, **24**, 5900–5913.
30. Hyman E, Kauraniemi P, Hautaniemi S, Wolf M, Mousses S, Rozenblum E, *et al*. Impact of DNA amplification on gene expression patterns in breast cancer. *Cancer Res* 2002, **62**, 6240–6245.
31. Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE, *et al*. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* 2002, **99**, 12963–12968.