

RESEARCH ARTICLE

Open Access

Estrogen independent gene expression defines clinically relevant subgroups of estrogen receptor positive breast cancer

Robin M Hallett and John A Hassell*

Abstract

Background: Human breast cancer represents a significantly heterogeneous disease. Global gene expression profiling measurements have been used to classify tumors into multiple molecular subtypes. The capacity to define subtypes of breast tumors provides a framework to enable improved understanding of the mechanisms of breast oncogenesis, as well as to provide opportunities for improved therapeutic intervention in patients.

Methods: We used publicly available gene expression profiling data to identify 'estrogen independent' genes in estrogen receptor alpha (ER+) breast tumors, and subsequently identified 6 subgroups of ER + breast tumors.

Results: Each of the 6 identified subgroups exhibited distinct clinical behaviors and biology. Patients whose tumors comprised subgroups 2,5&6 experienced excellent long-term survival, whereas those patients whose tumors belonged to subgroups 1&4 experienced much poorer survival. Breast tumor cell lines representative of the different subgroups responded to therapeutic compounds in accordance with their subgroup classification.

Conclusions: These data support the existence of 6 distinct subgroups of ER + breast cancer and suggest that knowledge of the ER + subgroup status of patient samples have the potential to guide therapy choice.

Keywords: Breast cancer, Gene expression, Subtypes, Therapies, Estrogen

Background

There is significant molecular and cellular diversity among human breast tumors. Indeed, this heterogeneity is evident from histopathologic features and differences in ER, progesterone receptor (PR) and ERBB2/HER2/NEU status as well as more recent molecular classification schemes based on the expression of large numbers of genes [1-3]. Importantly, these data indicate that breast cancer is an imprecise definition that embodies many molecularly distinct neoplastic disorders that share a common normal breast tissue origin.

The capacity to more accurately define breast cancers and identify tumor subgroups that represent more homogeneous disease entities, provides a framework to increase our understanding of these diseases and provides opportunities to focus treatment options for patients. To this

end investigators have completed relatively large gene expression studies and identified patterns in gene expression that reproducibly stratify breast tumors into each of 5 molecular subtypes. These breast cancer subtypes named basal-like, ERBB2-positive, normal-like, luminal A and luminal B were originally described by Perou *et al.* [1]. The various molecular subtypes possess distinct clinical behaviors thus providing a basis for improved taxonomy for breast cancer. For example, basal-like tumors are highly aggressive, resistant to endocrine therapies but sensitive to conventional chemotherapy, whereas luminal A tumors are more indolent and responsive to endocrine therapies. Importantly, recent and more comprehensive molecular profiling of human breast tumors, including global gene expression, mutation, DNA copy number variation, and protein expression support the original finding that breast cancer falls into major molecular subtypes comprising subsets of genetic and epigenetic abnormalities [4]. Currently, the additional clinical value of molecular classification over traditional histopathological methods is unclear,

* Correspondence: hassell@mcmaster.ca

Department of Biochemistry and Biomedical Sciences, Centre for Functional Genomics, McMaster University, 1200 Main Street West, Hamilton, Ontario L8N 3Z5, Canada

as the molecular subtypes show high correspondence to the expression of ER, PR, and HER2, as well as to tumor grade [3].

It is possible that further refinement of the 'intrinsic' classification scheme of Perou *et al.*, could identify other molecular classes of breast cancer, and provide additional clinical value beyond traditional techniques. For example, ER + tumors generally fall into the luminal A and B molecular subtypes, characterized by expression of the ER as well as cytokeratins typically expressed by luminal epithelial cells [1,3]. However, more recent studies suggest that as many as 12 molecular subgroups of ER + breast cancer exist, demonstrating that the luminal A and B stratification of ER + breast tumors does not fully capture the biological complexity of these tumors [5]. Indeed, further dissection of ER + breast tumors into additional relevant disease subgroups would likely provide further insight into the mechanisms that underlie these tumors, as well as prevent carefully planned studies from being confounded by the heterogeneity found among un-grouped or sub-optimally grouped populations of ER + breast tumors. Notably, the molecular subtypes of breast cancer show subtype specific response to standard chemotherapies as well as experimental compounds, highlighting the value of investigating specific disease subtypes [6]. Hence, the identification and characterization of additional subgroups of ER + breast tumors could focus treatment options for patients with ER + breast tumors, because therapy could be rationally applied based on specific molecular characteristics of the patient's tumor.

We hypothesized that the biology of ER + tumors comprised both estrogen-dependent and -independent components, and furthermore, that investigation and characterization of the estrogen independent component might provide a means to stratify ER + tumors into different distinct disease subgroups. To this end we used publicly available data to identify 'estrogen independent' genes in ER + breast tumors and subsequently identified subgroups of ER + tumors based on molecular differences between tumors identified by these genes. Importantly, we reproducibly identified 6 subgroups of ER + breast tumors that exhibited distinct clinical behavior as well as biology. Moreover, we show that these subgroups have specific responses to therapeutic compounds *in vitro*. Taken together these data support the existence of 6 distinct subgroups of ER + breast cancer, and advance efforts to increase the precision of therapeutic intervention in human breast cancer patients.

Methods

Human breast tumor data sets

All tumor samples were downloaded from the gene expression omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>). The latter included Letrozole treated tumor samples

(GSE5462) [7], the discovery cohort (GSE6532, 133A array samples, n = 327 [8], the validation cohort (GSE6532 133 Plus 2.0 array samples n = 87 [8], GSE9195 n = 77 [9], GSE17705 n = 298 [10], GSE2034 n = 209 [11], GSE7390 n = 134 [12], Original samples from GSE26971 (n = 136)). Cell line expression profiles were downloaded from ArrayExpress (E-TABM-157) [13]. Raw data files representing the tumor samples were normalised using RMA [14]. TCGA gene expression was obtained from the TCGA research network (<http://cancergenome.nih.gov/>), by downloading level 3 RNAseq data from the TCGA data portal (RSEM normalised) [15]. For GEO cohorts, ER + status was obtained from associated clinical files, which were generally based on histopathological assessment. ER + status for the TCGA cohort was determined using expression cut-offs (250 RSEM normalised transcript counts) for the ESR1 gene. ER + patients were selected from each dataset, and validation cohorts were combined after each probe set/gene was standardized and mean centered.

Cell line drug sensitivity data

We obtained previously reported human breast tumor cell line sensitivity data from Heiser *et al.* [6].

Definition of estrogen independent genes

We calculated within (w, treatment pairs) and between (b, independent primary tumor samples) variation for all tumors. In this fashion probe-sets with greater variation in expression between tumors than between treatment paired samples received high b/w scores, and vice versa.

PAM 50 subtype assignment

Subtype membership assignment was based on the nearest PAM50 centroid (Pearson correlation) [16].

Class discovery

Non-negative matrix factorization was carried out as previously described [17]. Prediction analysis of microarrays (PAM) was carried out as described [18] to discover subgroup specific genes (discovery cohort) and to classify samples (validation cohort, cell lines).

Cell growth assays

All cell lines were obtained from the ATCC and passaged minimally prior to completing these experiments. Cell lines were maintained as suggested by the ATCC. Cell lines were maintained in either RPMI or DMEM supplemented with 10% fetal bovine serum (all from Life Technologies). Cells were seeded at a density 50,000 cells/ml in the wells of a 6-well plate (Corning) in triplicate for each time point. At each time point cells were trypsinized and viable cells were counted with a hemocytometer using Trypan Blue exclusion as a marker of cell viability. Relative

cell growth was calculated as a number of viable of cells relative to control at each time point.

Statistical analysis

Survival analysis and Log-rank tests were used to evaluate survival differences between patient subgroups. We used 10 year distant metastasis free survival (DMFS) or disease free survival (DFS) as the clinical endpoint for these studies, and log-rank tests to detect differences in survival. T-test were used to compare means for 2-group comparisons, whereas ANOVA followed by Dunnett's multiple comparison test was used to compare means for 3 or more groups. Tests were two-sided and a p-value of 0.05 or less was considered statistically significant.

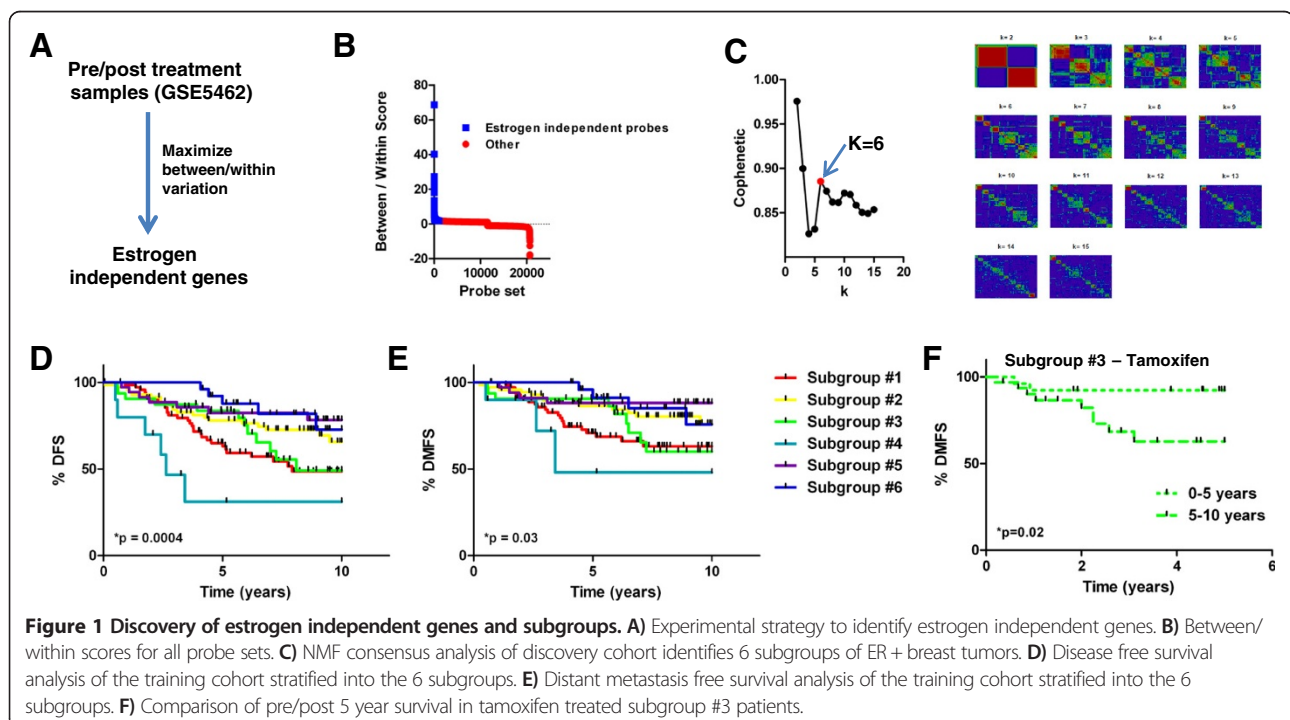
Results

Identification of estrogen independent genes and distinct subgroups of ER + breast cancer

The goal of this study was to enable classification of ER + breast tumors on the basis of genes whose expression is related to the estrogen independent biology of ER + tumors. To this end, we took advantage of the gene expression profiles of 58 ER + breast tumors biopsied from post-menopausal women before and after treatment with Letrozole (n = 116, [7]). Because letrozole treatment induces estrogen deprivation in tumors of post-menopausal women, we hypothesized that genes whose expression showed minimal variation after letrozole treatment could be considered to be expressed independent of estrogen. To identify estrogen-independent genes that might be

useful to identify subtypes of ER + breast tumors, we calculated between/within (b/w) scores for each probe set, which were measurements of probe set variation observed between different primary tumors relative to the variation observed within paired samples pre- and post-treatment. In this fashion, probe sets with high b/w scores showed greater variation between different primary tumors than between treatment paired tumor samples, whereas genes with low b/w scores showed greater variation within treatment paired tumors samples than between different tumors. Therefore, probe sets with high b/w scores were likely not influenced by estrogen deprivation and also showed variable expression among the different tumors prior to treatments, suggesting that they are related to differences in the estrogen independent biology of tumors (Figure 1A). We selected the top (highest b/w scores) 1,000 estrogen independent probe sets (893 genes) for further analysis (Figure 1B, Additional file 1: Table S1).

To investigate whether the expression of the estrogen independent probe sets could capture the phenotypic complexity of ER + breast tumors we completed unsupervised clustering using non-negative matrix factorization (NMF) [17]. NMF is an efficient method to identify molecular patterns that is readily applicable to gene expression data, and therefore can be used as a powerful means for class discovery. In short, NMF identifies metagenes, or distinct gene expression patterns, which are used to determine the optimal value for k, where k represents the number of sample subgroup clusters by calculating a cophenetic co-efficient for each value of k. In short, we applied NMF (for k = 2-10)



to gene expression data representing 262 primary ER + breast tumors (GSE6532, 133A arrays, [19] filtered such that only the 1,000 estrogen independent probe sets were used for class identification. This data set optimally fell into 6 clusters, designated subgroups 1–6 (Figure 1C). Moreover, NMF on an additional independent data set of 298 ER + breast tumors (GSE17705, [10]) using the same 1,000 estrogen independent probe sets also suggested that these patients were also optimally stratified into 6 subgroups (Additional file 2: Figure S1). Hence, we concluded that on the basis of the expression of estrogen independent genes, ER + breast tumors can be categorized optimally into 1 of 6 ER independent subgroups.

To learn whether these groups might encompass disease with different clinical outcomes we compared DFS (Figure 1D, * $p < 0.05$, Log-rank test) and DMFS (Figure 1E, * $p < 0.05$, Log-rank test) among the various subgroups. Interestingly, some subgroups displayed excellent long term outcomes, whereas other groups did not. For example, 10 year DMFS in subgroup 5 patients was 88%, whereas in subgroup 4 patients it was 48%. All patients comprising the various subgroups were uniformly chemotherapy naïve, suggesting that these differences in survival are likely related to the natural progression of their disease, rather than influenced by response to chemotherapy. Interestingly, in tamoxifen treated subgroup #3 ($n = 27$) patients we observed that the majority of DMFS events occurred after 5 years (Figure 1F), the time at which most patients cease tamoxifen treatment, possibly suggesting that these patients would have benefited from tamoxifen treatment beyond 5 years. Unfortunately, this dataset only comprised 8 subgroup #3 patients who did not receive tamoxifen, making the complimentary analysis in tamoxifen naïve patients impractical.

Subgroups are independent of the molecular subtype of breast cancer

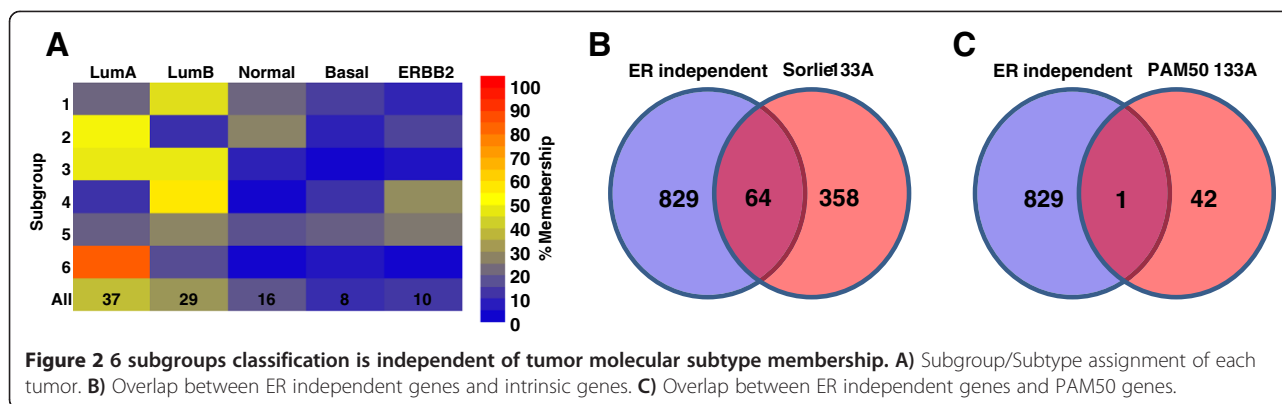
Significant data exists that breast tumors can be stratified into at least 5 molecular subtypes [1,2,16]. Accordingly, we examined whether there was an association between the 6 subgroups we identified and the 5 molecular subtypes of breast cancer. Classification of the 262 ER + tumors used for discovery using the PAM50 genes (43 genes present on 133A arrays), revealed that most of the tumors were classified into either the luminal A (37%) or luminal B (29%) molecular subtypes, whereas of the remainder 16% were normal, 8% were basal, and 10% were ERBB2 (Figure 2A, Additional file 2: Figure S2). Among the 6 ER independent subgroups, only subgroup #6 was strongly associated with any of the 5 molecular subtypes of cancer, and comprised ~82% Lum A, ~14% LumB, and ~4% Basal-like tumors (Figure 2A). Among the 893 ER + independent genes, only 64 overlapped with the Sorlie *et al.* intrinsic genes [2], and only 1 overlapped with the 43 PAM50 genes

[16], present on the 133A array (Figure 2B and C). Hence, we concluded that the classification of ER + breast tumors into the 6 subgroups we identified was relatively independent of their membership in the 5 molecular subtypes of breast cancer.

A framework for ER + breast tumor classification

To confirm that the 6 subgroups we identified were indeed generally representative of ER + breast tumors, we investigated their prevalence in an independent validation dataset comprising 941 ER + chemotherapy-naïve breast cancer patients. Briefly, we identified a 300 probe set classifier (using PAM, top 50 probe sets of each subgroup) to classify tumors into the 6 subgroups (Figure 2A, >80% concordance with NMF classification). Based on the expression of the 300 probe sets, we assigned each tumor comprising the validation data set in the 6 subgroups, using PAM (Figure 2B). Some 84% ($n = 788$) of the tumors in the validation set were assigned with a probability higher than 80% of belonging to one of the 6 subgroups, demonstrating that the classification framework is robust. Notably, the DFS and DMFS characteristics of patients comprising the various 6 groups were found to be highly coincident between the original ($n = 262$) and validation ($n = 941$) cohorts (Figure 3C and D, Additional file 2: Figure S3, Correlation: 0.89, * $p < 0.05$). For instance, 10 year DMFS was lowest in subgroup 4 for both the original and validation datasets. Similar to observations made in our training cohort, we observed that patients with subgroup #3 tumors experienced the majority of DMFS events after 5 years (Additional file 2: Figure S4). To learn whether this phenomenon was related to tamoxifen treatment, we subdivide subgroup #3 patients into tamoxifen treated ($n = 94$) and tamoxifen naïve patients ($n = 32$) and compared pre/post 5 year DMFS survival in each group. Whereas there was no difference between pre/post 5 year DMFS in tamoxifen naïve patients, there was a significant different pre/post 5 year DMFS in tamoxifen treated patients (Figure 3E and F, No tamoxifen, HR: 0.82, $p = 0.8$, Tamoxifen, HR: 0.26, * $p < 0.05$). These results might be interpreted to suggest that in subgroup #3 tumors early relapse is prevented by tamoxifen, albeit relapses resume after the completion of a patient's tamoxifen regimen. Indeed, clinical trials examining the use of tamoxifen for a period greater than 5-years demonstrate that a subset of ER + breast cancer patients benefit from such treatment [20]. Hence, patients with subgroup #3 tumors might represent those who benefit from extended tamoxifen treatment.

As additional validation, we investigated the prevalence of the 6 subgroups in the TCGA breast data set, which comprised 801 ER + breast tumors [4]. Using the PAM classifier described above, the mean probability for classification was 86%, and more than 70% ($n = 580$) of the tumors in the TCGA set were assigned a probability



of 80% or higher of belonging to one of the 6 subgroups (Additional file 2: Figure S5 A&B). Hence, this extra validation data set provides additional evidence for the robustness of the classification framework.

Taken together with our previous data, these results demonstrate that the 6 identified subgroups of ER + breast tumors can be reproducibly identified in independent patient cohorts and provide a clinically relevant means of classifying ER + breast tumors.

ER + subgroups enable predictive modeling of anti-cancer drug sensitivity

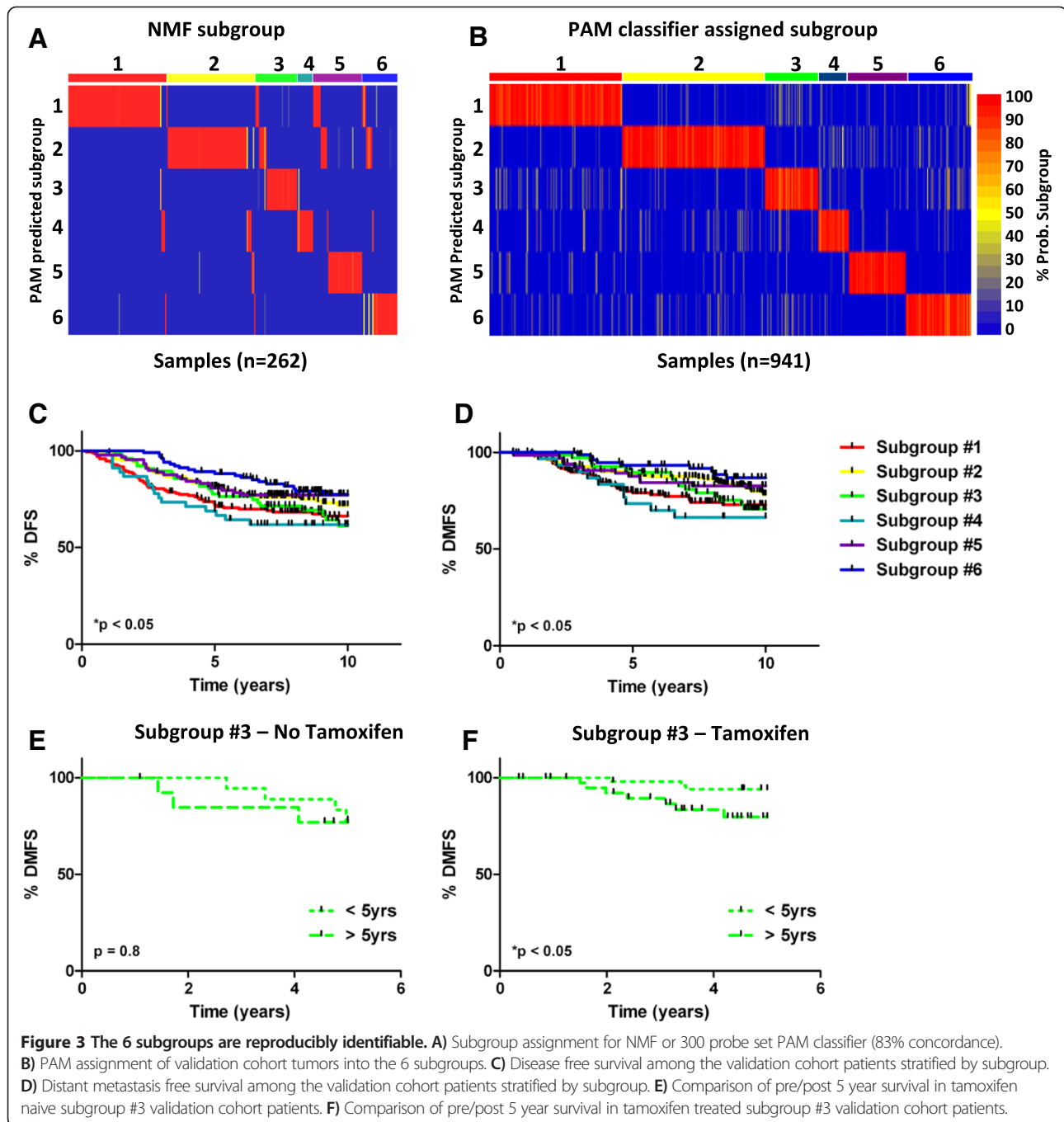
As described above, the established framework allows classification of ER + breast tumors into 1 of 6 subgroups based on patterns in estrogen independent gene expression. We first tested whether this framework could be extended to classify ER + breast tumor cell lines into the same subgroups. We accessed previously described breast tumor cell line gene expression datasets [6,13] and classified ER + breast tumor cell lines into the 6 subgroups. Among 24 ER + breast tumor cell lines, 5 of the 6 subgroups were represented by at least 4 cell lines, thus providing experimental models for these subgroups (Figure 4A, Additional file 1: Table S2). We sought to identify compounds with subgroup specificity for the most aggressive subgroups based on our analyses of DMFS in the patient cohorts. Although patients with subgroup #4 tumors experienced the worst outcome, we failed to identify any subgroup #4 cell lines. Hence we focused our efforts on the second most aggressive subgroup, which was subgroup #1. We observed that subgroup #1 tumors tended to over-express genes involved in the repair of double stranded (ds) DNA breaks, including RAD50 [21] and BARD1 [22], suggesting that subgroup #1 tumors possess a 'dsDNA break' phenotype and may be hypersensitive to agents that induce dsDNA breaks (Figure 4B). We also examined RAD50 and BARD1 expression in cell lines stratified by subgroup, however this analysis was inconclusive likely owing to the fact that most subgroups comprised very few lines (Additional file 2: Figure S6A). Subgroup #1 cell lines were hypersensitive to

etoposide, a potent and specific inducer of dsDNA breaks [23] (Figure 4C). Specifically, we compared the relative growth (to control) of 3 subgroup #1 cell lines and 3 cell lines belonging to other subgroups over 72 hours of treatment with 200nM etoposide. After 72 hours, relative growth was significantly lower in subgroup #1 cell lines (34% of control) compared to the relative growth of non-subgroup #1 cell lines (79% of control)(Figure 4D, *p = 0.008, t-test). To confirm these findings, we obtained breast tumor cell line drug sensitivity data that was previously reported by Heiser *et al.* in 2012 [6], for 18 cell lines that were also present within the cell line gene expression data set. Similar to our previous observations, we observed the subgroup #1 cell lines were generally the most sensitive to etoposide (Figure 4E-i. The mean $-\log_{10}(\text{IC}_{50})$ of subgroup #1 cell lines was significantly lower than cell lines belonging to other subgroups (Figure 4E-ii, *p = 0.02, t-test).

To extend these findings, we looked for over-expression of other actionable targets with subgroup selective expression among the 6 subgroups (Additional file 2: Figure S6). IGF2 was over-expressed in subgroup 2 tumors, implicating IGF signaling as a therapeutic target in subgroup#2 tumors. Interestingly, subgroup #3 tumors significantly over-expressed the angiotensin receptor 2 (AGTR2). Whereas AGTR2 hasn't been a target for cancer drug development, it has been a successfully exploited target for the development of hypertension drugs [24]. Other notable targets included over-expression of the anti-apoptotic protein BCL2 in subgroup #3 tumors, and the immune-modulatory target CTLA4 in subgroup #5 tumors. In each case, approved therapeutics exist or are under development that target these highlighted over-expressed genes. These observed patterns could potentially be used to target therapies in ER + breast cancer patients contingent on the subgroup membership of their tumor.

Discussion

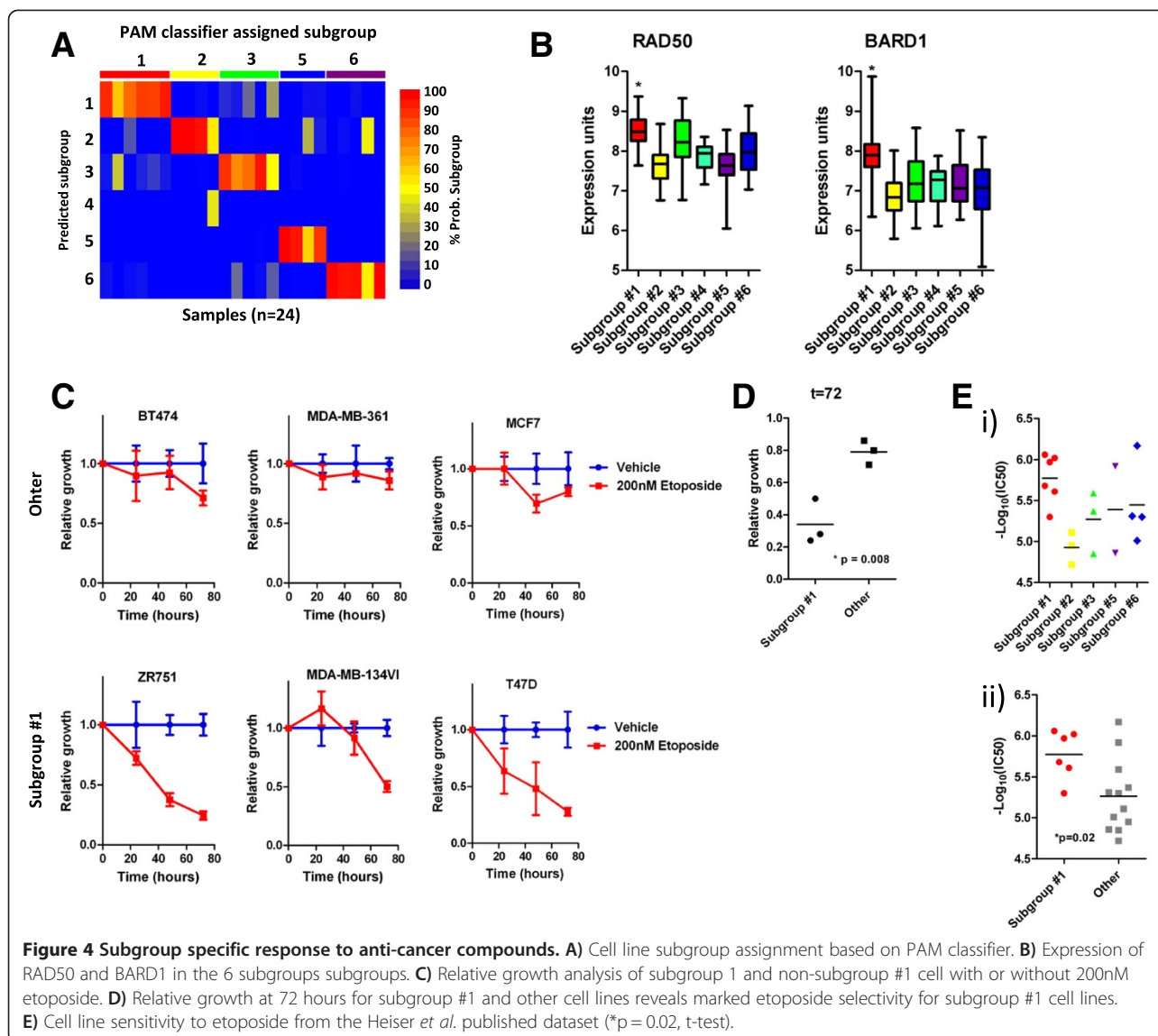
Substantial molecular heterogeneity exists among ER + tumors, which isn't adequately captured by either histopathological variables or more recent molecular subtyping



strategies. Accordingly, we sought to identify novel means of classifying ER+ tumors, and reproducibly identified 6 subgroups of ER+ tumors based on the expression of estrogen independent genes. Notably, we also observed survival and treatment sensitivity differences among the 6 subgroups. Hence, our data suggests that patient subgroup membership may be a useful tool for guiding treatment of ER+ breast cancer patients.

Briefly, the subgroup identification strategy was highly similar to that originally described by Perou *et al.* in 2000

[1]. Whereas Perou *et al.* employed an unsupervised clustering approach with intrinsic genes in unselected breast tumors, we employed unsupervised clustering with estrogen independent genes in breast tumors selected for ER positivity. For this experiment we analysed gene expression profiling data from 58 ER+ tumors biopsied from post-menopausal women before and after treatment with letrozole [7]. We hypothesized that genes whose expression showed minimal variation after letrozole treatment could be considered to be expressed independent of



estrogen, and identified estrogen independent genes based on this assumption. However, many breast cancer patients are pre-menopausal and receive different endocrine therapies for breast cancer treatment, namely tamoxifen. It is unclear whether the definition of estrogen independent genes we propose here would be different in pre-menopausal patients, or patients treated with alternate endocrine agents, and these possibilities represent intriguing avenues for future research. We note however, that subgrouping ER+ tumors based on estrogen independent gene expression was both robust and reproducible in cohorts of tumors that included pre-menopausal patients as well as those treated with tamoxifen, suggesting that our approach is broadly applicable.

There remain several limitations of the work reported herein. All of our conclusions are based on the analysis of retrospective data, which limits its clinical value. We

validated the occurrence, and clinical attributes, of the 6 subgroups in relatively large independent cohorts, however a true estimate of the clinical usefulness of the 6 subgroup classification for ER+ breast cancers would require additional validation in clinical trial samples, as well as completion of a prospective clinical trial examining the capacity of the classification to guide therapy. In addition, it isn't clear if subgroup classification would add meaningful clinical information beyond that obtained from existing prognostic tests designed for ER+ tumors, such as OncotypeDX® [25]. For example, a relevant question that remains is whether the good prognosis subgroups identified here (subgroups 2,5&6) experience similarly excellent survival to the low risk group identified by OncotypeDX®. Additionally, it isn't clear if the relationship between patient outcome and subgroup assignment is a consequence of subgroup association with natural progression of breast

cancer or tumor response to adjuvant endocrine therapy. Many of the patients obtained from publically available sources had incomplete clinical annotations, and they comprise a mixture of patients that received no adjuvant therapy, or adjuvant tamoxifen, likely lasting for 5 years. Based on these data it is difficult to discern how differences in extent or choice of endocrine therapy might influence the relationship between patient outcome and subgroup membership. Hence, although our data suggests the 6 subgroup classification of ER + breast cancer may be useful for guiding therapy in patients, many additional validation experiments are required to confirm our findings.

Conclusion

Ultimately, we propose that the 6 subgroups described here provide a strategy for improved understanding and treatment of ER + breast tumors. We demonstrate that the subgroups are unique and independent of the molecular subtypes of cancer, and provide a clinically relevant means of tumor classification. We anticipate that subgrouping will provide a framework to both guide optimal use of existing therapeutics, as well as gain insight into biological processes that represent relevant targets for development of the next generation of experimental therapies.

Additional files

Additional file 1: Supplemental tables.

Additional file 2: Supplemental figures.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

RH: Conceived, planned, analyzed, performed the experiments in the paper, wrote the manuscript. JAH: Helped write the manuscript, and provided critical feedback for the project. Both authors read and approved the final manuscript.

Acknowledgements

This work was generously supported by grants from the Canadian Breast Cancer Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We wish to acknowledge helpful discussion from Drs. Greg Pond and Anita Bane throughout the course of this work.

Financial support

This work was generously supported by grants from the Canadian Breast Cancer Foundation to JAH.

Received: 9 June 2014 Accepted: 4 November 2014

Published: 24 November 2014

References

- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale AL, Brown PO, Botstein D: **Molecular portraits of human breast tumours.** *Nature* 2000, **406**(6797):747–752.
- Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lønning PE, Brown PO, Børresen-Dale AL, Botstein D: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc Natl Acad Sci U S A* 2003, **100**(14):8418–8423.
- Sotiriou C, Pusztai L: **Gene-expression signatures in breast cancer.** *N Engl J Med* 2009, **360**(8):790–800.
- TCGA: **Comprehensive molecular portraits of human breast tumours.** *Nature* 2012, **490**(7418):61–70.
- Gatza ML, Lucas JE, Barry WT, Kim JW, Wang Q, Crawford MD, Datto MB, Kelley M, Mathey-Prevot B, Potti A, Nevins JR: **A pathway-based classification of human breast cancer.** *Proc Natl Acad Sci U S A* 2010, **107**(15):6994–6999.
- Heiser LM, Sadanandam A, Kuo WL, Benz SC, Goldstein TC, Ng S, Gibb WJ, Wang NJ, Ziyad S, Tong F, Bayani N, Hu Z, Billig JI, Dueregger A, Lewis S, Jakkula L, Korkola JE, Durinck S, Pepin F, Guan Y, Purdom E, Neuvial P, Bengtsson H, Wood KW, Smith PG, Vassilev LT, Hennessy BT, Greshock J, Bachman KE, Hardwicke MA, et al: **Subtype and pathway specific responses to anticancer compounds in breast cancer.** *Proc Natl Acad Sci U S A* 2012, **109**(8):2724–2729.
- Miller WR, Larionov AA, Renshaw L, Anderson TJ, White S, Murray J, Murray E, Hampton G, Walker JR, Ho S, Krause A, Evans DB, Dixon JM: **Changes in breast cancer transcriptional profiles after treatment with the aromatase inhibitor, letrozole.** *Pharmacogenet Genomics* 2007, **17**(10):813–826.
- Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, Ellis P, Harris A, Bergh J, Foekens JA, Klijn JG, Larsimont D, Buysse M, Bontempi G, Delorenzi M, Piccart MJ, Sotiriou C: **Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade.** *J Clin Oncol* 2007, **25**(10):1239–1246.
- Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemand F, Tutt AM, Gillet C, Ellis P, Ryder K, Reid JF, Daidone MG, Pierotti MA, Berns EM, Jansen MP, Foekens JA, Delorenzi M, Bontempi G, Piccart MJ, Sotiriou C: **Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen.** *BMC Genomics* 2008, **9**:239.
- Symmans WF, Hatzis C, Sotiriou C, Andre F, Peintinger F, Regitnig P, Daxenbichler G, Desmedt C, Domont J, Marth C, Delaloge S, Bauernhofer T, Valero V, Booser DJ, Hortobagyi GN, Pusztai L: **Genomic index of sensitivity to endocrine therapy for breast cancer.** *J Clin Oncol* 2010, **28**(27):4111–4119.
- Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J, Jatkoa T, Berns EM, Atkins D, Foekens JA: **Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer.** *Lancet* 2005, **365**(9460):671–679.
- Desmedt C, Piette F, Loi S, Wang Y, Lallemand F, Haibe-Kains B, Viale G, Delorenzi M, Zhang Y, d'Assignies MS, Bergh J, Lidereau R, Ellis P, Harris AL, Klijn JG, Foekens JA, Cardoso F, Piccart MJ, Buysse M, Sotiriou C, TRANSBIG Consortium: **Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series.** *Clin Cancer Res* 2007, **13**(11):3207–3214.
- Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, Speed T, Spellman PT, DeVries S, Lapuk A, Wang NJ, Kuo WL, Stilwell JL, Pinkel D, Albertson DG, Waldman FM, McCormick F, Dickson RB, Johnson MD, Lippman M, Ethier S, Gazdar A, Gray JW: **A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes.** *Cancer Cell* 2006, **10**(6):515–527.
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data.** *Biostatistics* 2003, **4**(2):249–264.
- Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome.** *BMC Bioinformatics* 2011, **12**:323.
- Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM, Bernard PS: **Supervised risk predictor of breast cancer based on intrinsic subtypes.** *J Clin Oncol* 2009, **27**(8):1160–1167.
- Brunet JP, Tamayo P, Golub TR, Mesirov JP: **Metagenes and molecular pattern discovery using matrix factorization.** *Proc Natl Acad Sci U S A* 2004, **101**(12):4164–4169.
- Tibshirani R, Hastie T, Narasimhan B, Chu G: **Diagnosis of multiple cancer types by shrunken centroids of gene expression.** *Proc Natl Acad Sci U S A* 2002, **99**(10):6567–6572.
- Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, Desmedt C, Larsimont D, Cardoso F, Peterse H, Nuyten D, Buysse M, Van de Vijver MJ, Bergh J, Piccart M, Delorenzi M: **Gene**

expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 2006, **98**(4):262–272.

20. Davies C, Pan H, Godwin J, Gray R, Arriagada R, Raina V, Abraham M, Medeiros Alencar VH, Badran A, Bonfill X, Bradbury J, Clarke M, Collins R, Davis SR, Delmestri A, Forbes JF, Haddad P, Hou MF, Inbar M, Khaled H, Kielanowska J, Kwan WH, Mathew BS, Mittra I, Müller B, Nicolucci A, Peralta O, Pernas F, Petruzella L, Pienkowski T, *et al*: Long-term effects of continuing adjuvant tamoxifen to 10 years versus stopping at 5 years after diagnosis of oestrogen receptor-positive breast cancer: ATLAS, a randomised trial. *Lancet* 2013, **381**(9869):805–816.
21. Johzuka K, Ogawa H: Interaction of Mre11 and Rad50: two proteins required for DNA repair and meiosis-specific double-strand break formation in *Saccharomyces cerevisiae*. *Genetics* 1995, **139**(4):1521–1532.
22. Meza JE, Brzovic PS, King MC, Klevit RE: Mapping the functional domains of BRCA1. Interaction of the ring finger domains of BRCA1 and BARD1. *J Biol Chem* 1999, **274**(9):5659–5665.
23. Hande KR: Clinical applications of anticancer drugs targeted to topoisomerase II. *Biochim Biophys Acta* 1998, **1400**(1–3):173–184.
24. Timmermans PB, Wong PC, Chiu AT, Herblin WF, Benfield P, Carini DJ, Lee RJ, Wexler RR, Saye JA, Smith RD: Angiotensin II receptors and angiotensin II receptor antagonists. *Pharmacol Rev* 1993, **45**(2):205–251.
25. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004, **351**(27):2817–2826.

doi:10.1186/1471-2407-14-871

Cite this article as: Hallett and Hassell: Estrogen independent gene expression defines clinically relevant subgroups of estrogen receptor positive breast cancer. *BMC Cancer* 2014 **14**:871.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

