Review and Cross-Validation of Gene Expression Signatures and Melanoma Prognosis

Sarah-Jane Schramm[1], Anna E. Campain[2], Richard A. Scolyer[3,4,5], Yee Hwa Yang[2] and Graham J. Mann[1,3]

In melanoma, there is an urgent need to identify novel biomarkers with prognostic performance superior to traditional clinical and histological parameters. Gene expression-based prognostic signatures offer promise, but studies have been challenged by sample scarcity, cohort heterogeneity, and doubts about the efficacy of such signatures relative to current clinical practices. Motivated by new studies that have begun to address these challenges, we reviewed prognostic signatures derived from gene expression microarray analysis of human melanoma tissue. We used REMARK-based criteria to select the most relevant studies and directly compared their signature gene lists. Through functional ontology enrichment analysis, we observed that these independent data sets converge in part upon immune response processes and the G-protein signaling NRAS-regulation pathway, both important in melanoma development and progression. The signatures correctly predicted patient outcome in independent gene expression data sets with some notably low misclassification rates, particularly among studies involving more advanced-stage tumors. This successful cross-validation indicates that gene expression analysis-based signatures are becoming translationally relevant to care of melanoma patients, as well as improving understanding of the aspects of melanoma biology that determine patient outcome.

Journal of Investigative Dermatology (2012) **132,** 274–283; doi:10.1038/jid.2011.305; published online 29 September 2011

[1]*Sydney Medical School, The University of Sydney at Westmead Millennium Institute, Sydney, New South Wales, Australia;* [2]*School of Mathematics and Statistics, The University of Sydney, Sydney, New South Wales, Australia;* [3]*Melanoma Institute Australia, Sydney, New South Wales, Australia;* [4]*Tissue Pathology and Diagnostic Oncology, Royal Prince Alfred Hospital, Sydney, New South Wales, Australia and* [5]*Discipline of Pathology, Sydney Medical School, The University of Sydney, Sydney, New South Wales, Australia*

Correspondence: Sarah-Jane Schramm, Melanoma Genomics and Genetic Epidemiology Research Group, The University of Sydney at Westmead Millennium Institute, PO Box 412, Westmead, New South Wales 2145, Australia. E-mail: graham.mann@sydney.edu.au

Abbreviations: AJCC, American Joint Committee on Cancer; GEMA, gene expression microarray analysis; TNM, tumor, node, metastasis

Received 14 April 2011; revised 18 July 2011; accepted 1 August 2011; published online 29 September 2011

INTRODUCTION

In the past decade, gene expression microarray (GEMA) technology has shown fluctuating promise of enhancing or even replacing current prognostic strategies in the management of cancer (Hoek, 2007; Shendure, 2008). The prospect of predicting clinical tumor behavior on the basis of its transcriptome remains enticing, but has not yet been realized for most types of cancer (Subramanian and Simon, 2010; Varmus, 2010) including melanoma, because discriminating gene signatures remain elusive. Perhaps, the most advanced use of prognostic GEMAs in a clinical setting is in the case of breast cancer patients. Currently, the US Food and Drug Administration (FDA)-approved MammaPrint 70-gene signature (van't Veer *et al.*, 2002) is undergoing evaluation in a prospective, randomized clinical trial comparing it with traditional clinicopathological criteria for selecting patients for adjuvant chemotherapy (MINDACT Trial: http://www.eortc.be/services/unit/mindact/MINDACT_websiteii.asp). In addition, for estrogen receptor-positive, lymph node-negative breast cancers, the GEMA-derived Oncotype Dx reverse transcriptase-PCR assay (Genomic Health, Redwood City, CA; Paik *et al.*, 2004) is being used to identify patients with high risk of distant recurrence who may benefit from chemotherapy. Skepticism remains, however, that these signatures and others similar to them (e.g., Theros, MapQuant Dx) will enhance clinical decision making and justify their costs (which in some instances are more than US$ 5,000). Questions center primarily around the capacity of GEMAs to outperform current factors in prognostication and, more recently, their potential redundancy in the wake of massively parallel sequencing technologies. Nevertheless, there is evidence that their prognostic power in clinical practice would at least complement traditional factors (Weigelt *et al.*, 2010).

For melanoma, there has hitherto been no such clinical application of GEMAs (Figure 1). This situation is a special problem because disseminated melanoma is one of the most aggressive cancers (Thompson *et al.*, 2009) and, despite the recent advances of BRAF target therapies (Flaherty *et al.*, 2010; Kefford *et al.*, 2010), surgical resection remains the mainstay of treatment. Prognosis is currently assigned almost entirely on the basis of a limited set of clinical and histopathological features (Balch *et al.*, 2009). For patients with clinically localized primary cutaneous melanoma, their prognosis is primarily influenced by tumor thickness, ulcerative state, and dermal mitotic rate. For melanoma patients with metastatic disease, prognosis is influenced by the number, site, and size of metastases, as well as by clinical

© 2012 The Society for Investigative Dermatology

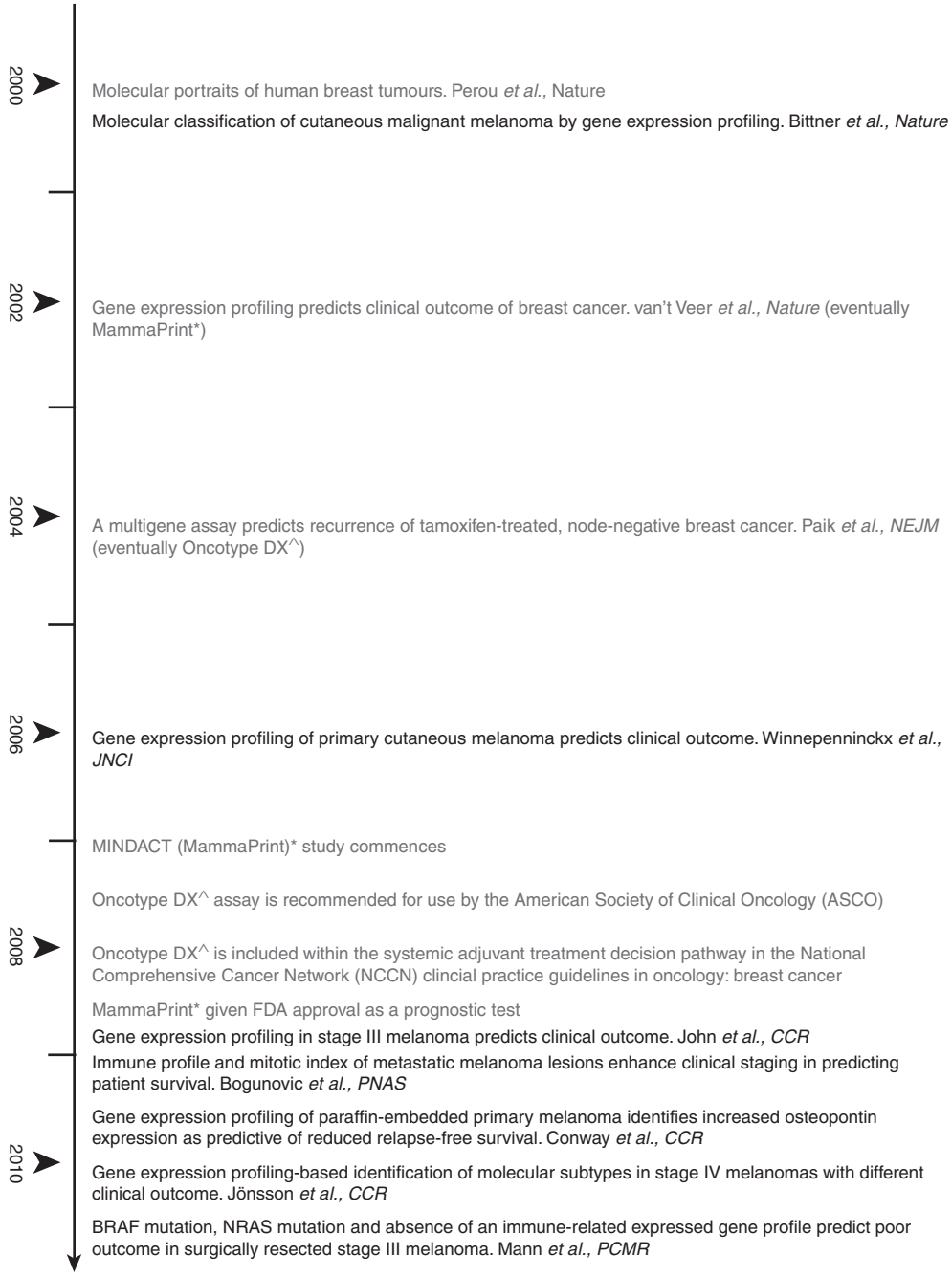

Figure 1. Timeline of selected key events in the evolution of prognostic/predictive signatures in melanoma (in black text) compared with breast cancer (in gray text). *MINDACT (Microarray In Node-negative and 1–3 positive lymph node Disease may Avoid ChemoTherapy) study, a prospective randomized study comparing the MammaPrint 70-gene signature with common clinical–pathological criteria in selecting patients for adjuvant chemotherapy in breast cancer with 0–3 positive nodes. ^Oncotype DX, a 21-gene quantitative RT-PCR (qRT-PCR) assay for formalin-fixed, paraffin-embedded, estrogen receptor (ER)-positive, lymph node–negative breast cancers to predict risk of recurrence.

information such as serum lactate dehydrogenase level among other parameters. Although this approach is useful for assigning broad probabilities of relapse, it has limited predictive power at the level of individual patients and has no direct implications for personalizing therapy. Therefore, the search for novel molecular biomarkers has been intense (reviewed in Gould Rothberg *et al.*, 2009a, b; Schramm and Mann, 2011). For ideal clinical relevance, GEMA-derived molecular subgroups should be readily and reproducibly identifiable, should show clear and independent relationships with specific survival outcomes (prognostic signatures) or therapy options (predictive signatures), and should do so with added or greater sensitivity than the current set of biomarkers.

In the most recent review of melanoma prognostic signatures, Timar *et al.* (2010) compared signature gene lists derived from GEMA assays of human melanoma tissue

(Bittner *et al.*, 2000; Mandruzzato *et al.*, 2006; Winnepenninckx *et al.*, 2006; John *et al.*, 2008). They found that there was little overlap between the signatures, both within and between the four studies that they compared, and attributed this mostly to sample heterogeneity. In the months since that review, melanoma research has experienced a surge in progress, with three special developments underpinning the timeliness of the current review. First, minimum standards for the general methodological approach of GEMA-based prognostic/predictive biomarker discovery (e.g., sample selection and multivariate statistical analysis as discussed in Allison *et al.*, 2006; Subramanian and Simon, 2010; Schramm and Mann, 2011) have been promulgated to enhance their translational reliability. Second, there has been a number of rigorous, independent publications describing outcome-related class prediction in human melanoma samples (see, e.g., Jönsson *et al.*, 2010; Mann *et al.*, 2010; Figure 1). Finally, the encouraging progress of anti-*BRAF* agents in managing advanced-stage melanoma patients has reinforced the pressing need to dissect melanomas into distinct subpopulations for the purpose of personalized therapy (Arkenau *et al.*, 2010). Taken together, these advances warrant an updated review of GEMA-based prognostic/predictive signatures in melanoma.

In this review, we examine the key studies involving GEMA-based work conducted using human melanoma samples in a prognostic/predictive study-design context. To capture all such studies, we first conducted a comprehensive survey of the literature (see Schramm and Mann, 2011 for search parameters). Inclusion in this review required researchers to have a clear hypothesis for supervised classification, and a function or rule used to classify patients where the predicted class was a prognosis-related outcome such as survival or disease-free progression. We excluded studies in which signatures were not validated using independent data sets and/or in which the signature was not assessed for its performance compared with current prognostic factors (multivariate testing). We compared the signatures among eligible studies (Table 1) using two separate methods: we first used MetaCore (from GeneGo, St Joseph, MI) to directly compare the prognostic signature gene lists with each other to identify areas of potential convergence between them that may drive new hypotheses about recurring, aberrant pathways in melanoma. Second, we conducted a formal and systematic cross-validation study in which the capacity of each signature to predict outcome in each of the other selected expression data sets was examined. We observed that carefully selected, independent studies are extracting some similar biological information about gene expression differences between patients with good and poor survival outcomes, although from different biomarkers. Specifically, the consistent observation of altered immune response regulation across GEMA studies suggests that immune-related molecules are strong candidates as potentially valuable biomarkers. Interstudy agreement such as this is broadly useful in highlighting potential directions for future research and can add valuable knowledge to understanding basic melanoma biology *per se*. We also believe that recent studies support prospects for the potential integration of the best of traditional and gene expression-based approaches in clinical practice. Finally, we provide commentary on future directions in the area as work in the field moves forward in the wake of the breast cancer precedent and amid the transforming landscape of genomic medicine.

PROGNOSTIC SIGNATURES IN MELANOMA

2006: 254-gene (element) expression signature

The earliest eligible study, referred to here as the Winnepenninckx study (Winnepenninckx *et al.*, 2006), involved analysis of 58 primary cutaneous melanoma samples and identified a 254-gene (element) expression signature associated with a 4-year distant metastasis-free survival. The performance of the classifier was tested in an independent sample set of primary cutaneous melanomas in which 11 out of 17 cases were correctly categorized with respect to their known survival outcome. The potential clinical use of the classifier was then compared with the predictive accuracy of traditional primary tumor-staging factors, tumor thickness, and the presence of ulceration. The researchers first generated a clinical prediction rule whereby patients with nonulcerated primary melanoma >4 mm thick, or patients with an ulcerated melanoma >2 mm thick, were classified as having poor prognosis, whereas all other patients were classified as having favorable prognosis. Considering these two factors, 28% of patients were misclassified, whereas the 254-gene (element) signature misclassified 29% of the cases in which the probability of randomly achieving that misclassification rate was 0.04. Bioinformatics analyses highlighted DNA replication and repair-related genes as being overrepresented within the data set (Kannengiesser *et al.*, 2008; Kauffmann *et al.*, 2008). Although this work initiated the discovery of prognostic signatures in melanoma and led to new insights into biological processes, it has not directly progressed into a clinical application, nor has it led to a change in clinical practice.

2008: 21-gene (element) expression signature

In the next eligible study to emerge (John *et al.*, 2008), researchers used oligonucleotide arrays (30,888 probes) to examine lymph node metastases from 29 patients with stage IIIB and IIIC melanoma. Selecting a cutoff for good prognosis of 24 months, there were 16 patients with poor prognosis and 13 with good prognosis. The mean time to progression for the good prognosis group was 40 months compared with 4 months for the poor prognosis group. Multivariate analysis showed no statistically significant differences in age, gender, American Joint Committee on Cancer (AJCC) stage, the use of adjuvant IFN therapy, or the presence of tumor-infiltrating lymphocytes between the two patient groups. From the list of 2,140 significantly differentially expressed genes, two outcome-related tests were developed, a 21-gene element prognosticator and a 5-gene quantitative PCR prognosticator. In an independent sample of 10 tumors, the quantitative PCR predictor correctly classified 9 of them with respect to outcome. Validation of the 21-gene signature was achieved but with minimal statistical power ($n=4$). We have recently

Table 1. Summary of studies investigating prognostic signatures in human melanoma

Study	Sample size	Cohort	Platform	Classes compared	Signature	Validation	Performance in multivariate setting
Winnepenninckx *et al.* (2006)	58	Primary cutaneous tumor, mixed histology, fresh frozen	Pangenomic 60-mer oligonucleotide microarray (Agilent Technologies) (44,000 probes)	>4-year distant metastasis-free survival versus ⩽4-year distant metastasis-free survival	254 expressed genes (gene elements)	In-house sample (n=17) and IHC (23 proteins)	Signature misclassified 29% of cases[1] whereas standard measures (thickness, ulceration) misclassified 28% of patients
John *et al.* (2008)	29	Node sections (stages IIIB and IIIC), snap-frozen, embedded in OCT compound	Pangenomic oligonucleotide microarray (MWG Biotech) (30,888 probes)	⩾24 months TTP from stage III to stage IV disease versus <24 months TTP stage III to stage IV disease	21 expressed genes (gene elements)	In-house sample (n=4) and expression data set from Winnepenninckx *et al.* (2006)	Signature[2] misclassified 15% of cases for which no statistically significant (P<0.05) between-group differences in median age, sex, AJCC stage, the use of adjuvant IFN, or the presence of TILs were observed
John *et al.* (2008)					5-gene qPCR classifier	In-house sample (n=10)	Signature misclassified 10% of cases for which no statistically significant (P<0.05) between-group differences in median age, sex, AJCC stage, the use of adjuvant IFN, or the presence of TILs were observed
Bogunovic *et al.* (2009)	38	Metastatic tumor (stages III and IV), mixed treatment regimes	Human Genome U133 Plus 2.0 Chips (Affymetrix) (54,675 features)	Prolonged survival (⩾1.5 years) compared with patients with shorter survival	266 expressed genes (gene elements)	Expression data set from John *et al.* (2008)	Signature was a significant, independent predictor of outcome (P=0.03) alongside MI (P=0.0002) where TNM stage was not a independent predictor (P=0.30) in that same test
Conway *et al.* (2009)	254	Primary melanoma, Breslow thickness >0.75 mm, FFPE tissue	DASL Human Cancer Panel (Illumina) (1,536 probes targeting 502 genes)	Relapse-free survival (from diagnosis to date of first relapse at any site), overall survival	Osteopontin (*SPP1*) expression	In-house sample (n=218) quantitative real-time reverse transcription-PCR	Gene expression did not maintain significance in the validation set when adjusted for thickness, ulceration, and mitotic rate (P=0.32)
Jönsson *et al.* (2010)	57	Stage IV melanoma biopsies (subcutaneous and lymph node), fresh frozen	Human WG-6 v2 Expression Beadchip (Illumina) (48,000 probes)	Overall survival between 4 tumor subgroups (high immune response, proliferative, pigmentation, and normal like)	~100, top-ranking significantly differentially expressed genes (gene elements) per subtype	In-house sample (n=20) and expression data sets from Bogunovic *et al.* (2009), John *et al.* (2008), Winnepenninckx *et al.* (2006), Haqq *et al.* (2005), and publicly available cell line data and IHC	Overall survival differed significantly (P=0.01, log-rank test) between the four subtypes with the proliferative subtype associated with the shortest survival (P=0.003) and with no significant (P<0.01) between-subtype differences in age, Breslow thickness, Clark level, or LDH level
Jönsson *et al.* (2010)				Overall survival	*A priori*-defined 30, immune response-related expressed genes		Lower overall survival (P=0.001) in tumors with a low expression of immune response signature
Mann *et al.* (2010)	48	Fresh frozen nodal metastases	Sentrix HumanRef-6 v3.0 Panel (Illumina) (>48,000 probes)	Survival <1 year compared with >4 years after surgery	46-gene (60-probe) expression signature	In-house sample (LOOCV) and expression data sets from Bogunovic *et al.* (2009) and John *et al.* (2008)	Significant, independent predictor of outcome alongside disease stage <III at initial presentation (ORs for survival >4 years, 90% CI): the presence of a nodular component in the primary (6.8, 0.6–76.0), small cell size (11.1, 0.8–100.0), or low pigmentation (3.0, 0.8–100.0) in the nodal metastases, absence of *BRAF* mutation (20.0, 1.0 to >1,000.0), or absence of *NRAS* mutation (16.7, 0.6 to >1,000.0)

Abbreviations: AJCC, American Joint Committee on bancer; CI, confidence interval; DASL, cDNA-mediated annealing, selection, extension, and ligation; FFPE, formalin-fixed, paraffin-embedded; IHC, immunohistochemistry; LDH, lactate dehydrogenase; LOOCV, leave-one-out cross-validation; MI, mitotic index; OCT, optimal cutting temperature; OR, odds ratio; qPCR, quantitative PCR; TIL, tumor infiltrating lymphocyte; TNM, tumor node metastasis; TTP, time taken to tumor progression.
[1]The probability of randomly achieving this misclassification rate was 0.04.
[2]Of the 21 probes, 13 were assessed in that process.

completed a study (Mann *et al.*, 2010) assessing outcome prediction using stage III melanomas and revalidated this signature (see below). Furthermore, several of the signature genes, including *CHST4*, *MFGE8*, and *KCNIP2*, have been linked to anticancer immune response.

2009: 266-gene (element) expression signature

In an examination of 38 metastatic melanomas, Bogunovic *et al.* (2009), identified 266 genes/gene elements associated with postrecurrence survival. This signature was assessed for its predictive capacity against the 2001 version of the AJCC tumor, node, metastasis (TNM) staging system (Balch *et al.*, 2001), for the presence of tumor-infiltrating lymphocytes, T-cell *CD3* positivity, and mitotic index. The mitotic index was the most significant predictor of outcome (hazard ratio = 2.13, $P = 0.0008$), a result that is now aptly reflected by inclusion of mitotic rate as an important prognostic factor in the current AJCC TNM staging system adopted in 2010 (Balch *et al.*, 2009). The gene signature was, however, also a significant and independent predictor of outcome ($P = 0.03$) in a multivariate context alongside mitotic index ($P = 0.0002$) but not TNM stage ($P = 0.30$). Encouragingly consistent with the work of John *et al.* (2008), several of the signature genes were immune response related (e.g., *ICOS*, *CD3* delta, *ZAP70*, *TRAT1*, *TARP*, *GZMK*, *LCK*, *CD2*, *CXCL13*, *CCL19*, *CCR7*, and *VCAM1*). A cell proliferation theme emerged for genes negatively associated with survival (e.g., *PDE4D*, *CDK2*, *GREF1*, *NUSAP1*, and *SPC24*).

2009: osteopontin expression

Conway *et al.* (2009) conducted the first eligible study to assess formalin-fixed, paraffin-embedded primary melanoma tissue and identified osteopontin expression as predictive of relapse-free survival in the training set. This predictive signal, however, was not maintained in a validation set in a multivariate context when the top-performing histopathological factors, namely, Breslow thickness, mitotic rate, and the presence of ulceration, were taken into account. In spite of this finding, these authors have attempted to address a significant challenge to the field; banks of frozen primary tissue are rare in melanoma. For routine patient clinical care, melanoma specimens are usually fixed in formalin and embedded in paraffin for histopathological diagnosis, for assessment of prognostic parameters, and for determination of surgical margin status. Improved assays have prompted a revisit and analysis of the large biospecimen banks of formalin-fixed, paraffin-embedded tissue. The Conway study supports the view that high-throughput gene expression studies are now possible using archival formalin-fixed, paraffin-embedded tissue. We await upcoming publications in the area to gain further insight into the extent of their translational relevance, including for miRNA analysis (e.g., as in Jukic *et al.*, 2010).

2010: 30-gene (element) expression signature and subtype discriminator

In the second most recent key study, Jönsson *et al.* (2010) examined 57 subcutaneous and lymph node melanoma metastases. Using unsupervised hierarchical clustering and the top 3,000 most variably expressed genes, the researchers identified four tumor subgroups. On the basis of the within-group gene content, subtypes were named as follows: (1) high immune response, (2) proliferative, (3) pigmentation, and (4) normal like. These groups were each observed to be significantly associated with outcome, with patients in the proliferative group having the shortest overall survival relative to those in the other three groups. These researchers also proposed an *a priori*-defined immune response-related gene expression signature of 30 genes and tested its use in predicting outcome. A key strength of the work is that authors used independent in-house data ($n = 20$), as well as data from John *et al.* (2008), Bogunovic *et al.* (2009), Haqq *et al.* (2005), and Winnepenninckx *et al.* (2006), and some publicly available cell line data to test their classifiers. Overall, the results of that validation varied with the prediction rules holding up in some data sets but not in others. In contrast to prior related studies, these researchers also used complementary approaches to further characterize molecular traits of their identified subtypes (including comparative genomic hybridization, reverse transcriptase-PCR, immunohistochemistry, examination of promoter methylation status, multiplex ligation-dependent probe amplification, and mutation screening). It is noteworthy that (1) the proliferative and pigmentation groups were observed to be more genetically heterogeneous than the other groups when tested in this integrated way; (2) *MITF* was highly expressed in the pigmentation group; (3) methylation of *p16INK4a* promoter was associated with the proliferative subtype; (4) six of the eight tumors with *CDKN2A* homozygous deletions were from the proliferative group; and (5) *BRAF* and *NRAS* mutations were found in all groups but were most frequent in the proliferative group.

2010: 60-probe (46-gene) expression signature

In the most recent study to date, we assayed fresh frozen nodal metastases ($n = 48$) coupled with a comprehensive clinical, pathological, and genetic assessment of both the banked nodal melanoma and preceding primary tumor (Mann *et al.*, 2010). The cohort comprised patients selected from the extremes of survival distribution (class A: poor prognosis, survival <1 year after surgical resection, died because of melanoma, cf. class B good prognosis, survival >4 years after surgical resection, with no sign of relapse). The clinical, pathological, and molecular parameters (including results of somatic mutation profiling) were analyzed using multiple imputation and logistic regression for determinants of outcome. The final model included the following independent predictors of good outcome: disease stage $<$III at initial presentation; the presence of a nodular component in the primary tumor; small cell size or lower pigmentation in the nodal metastases; absence of *BRAF* or *NRAS* mutation; and a 46-gene (60-probe) expression signature. This gene signature was validated using two independent later-stage melanoma data sets (John *et al.*, 2008; Bogunovic *et al.*, 2009). Again consistent with prior related studies, bioinformatics analysis demonstrated strong overrepresentation of immune-related

genes in the prognostic gene set. High-dimensional multivariate analysis showed that combining both clinicohistopathological and GEMA data was more effective when compared with reliance on either factor on its own leave-one-out cross-validation error rate = 0.23 for the signature, together with traditional factors cf. 0.25 and 0.27, respectively, for those elements alone).

COMPARATIVE ANALYSES OF PROGNOSTIC SIGNATURES

Direct comparison of eligible prognostic signatures

We performed bioinformatics analysis of each of the aforementioned prognostic signatures in order to identify similar genes between them and examine whether such intersections may potentially illuminate pathways and processes of interest in the broader context of disease drivers and/or correlates. First, the published signature lists from each eligible study were manually extracted. We applied the ''Compare experiments'' algorithm (default settings) in MetaCore (from GeneGo) that performs mapping between objects of each gene list and defines common, similar, and/or unique subsets between them (not accounting for the possibility that intersections may occur by chance) accompanied by gene ontology enrichment analysis. Altogether, we observed 10 pairwise intersections among our eligible data sets (Table 2). Among notable individual genes are those appearing in more than one data set intersection, including *PGHD*, preferentially normally expressed in the brain and a catalyst for conversion of prostaglandin H2 to neuromodulator prostaglandin D2, as well as several cluster of differentiation (CD) molecules. Functional enrichment analysis of the union of all genes present at any intersection between two data sets ($n = 46$, excluding those genes from the 30-gene signature of John *et al.*, 2008) highlighted the immune response as a significant ($P < 0.05$) top-ranking biological process, contributed to by molecules such as *CD2*, *CD3*, *CD79*, *HLA-DQB1*, *ICOS*, *IRF8*, Ikaros, *LTB*, and *VCAM1* (Supplementary Tables S1–3 online). G-protein signaling-related molecules, including several from the *NRAS* regulation pathway (*MHC* class II, *CD3*, *LCK*, and *CALDAG-GEFI*), also feature when the data are analyzed in that way (Supplementary Table S1 online). This finding is not surprising given that *NRAS* and downstream *BRAF* mutations are present in ~60% of melanomas (Lee *et al.*, 2010). Nevertheless, the fact that the *NRAS* pathway signal has emerged across different assay platforms, batches, and statistical methods is evidence of the potential of GEMAs (singularly and when combined with studies as we have performed here) to produce biologically relevant output.

Cross-validation of signatures between data sets

In the second of our study comparisons, we assessed the predictive power of each of the eligible signature gene lists in each of the other eligible gene expression data sets (Supplementary Methods online) (Table 3).

Overall, the signature with the lowest average error rate across all validation expression data was that proposed by Mann *et al.* (2010; 0.28). The John study signature also performed well using other expression data sets (error rates ~0.2–<0.27), except the Winnepenninckx expression data. This observation may be a reflection of stage difference; the Winnepenninckx study involved primary melanomas, whereas the other studies examined more advanced-stage tumors. The Bogunovic study signature also validated with encouragingly low error rates using the John and Jönsson expression data sets (0.17 and 0.20, respectively) but lost its good prognostic capacity using the Mann and Winnepenninckx expression data (0.54 and 0.48, respectively). A close inspection of the cohorts presents a possible explanation for this observation; the Bogunovic sample included stage IV melanoma tissue, as did the Jönsson study. John *et al.* (2008) examined stage IIIB or higher-stage tumors, whereas both the Mann and Winnepenninckx data sets were derived from (slightly) earlier-stage samples (Mann: stage IIIA, B, and C; Winnepenninckx: primary tumor/stages I and II). If GEMAs have discriminated between cohorts on the basis of these arguably subtle differences, then we can be, once again, encouraged by their potential prognostic sensitivity. However, we cannot rule out other plausible explanations for the observed differences, including the effects of immunotherapy treatment of at least some patients in the Bogunovic, John, and Jönsson study cohorts, as well as platform differences and/or varying definitions of the end point being assessed (Table 1). Nevertheless, it is encouraging that the later-stage signatures appear to consistently validate with lower error rates using the later-stage expression data relative to the Winnepenninckx primary tumor stage expression data set (although, again, the possibility that this effect may be due to platform or other differences between studies cannot be ruled out without further testing).

It is noteworthy that the Jönsson *et al.* (2010) *a priori*-defined, immune-related gene list performed most poorly on average (0.45) relative to the other signatures we examined. This occurrence appears to be in spite of a strong immune response-related signal emanating from the combined data sets, as confirmed in our bioinformatics analysis above. In contrast, the Mann signature performed generally well and is *per se* a highly immunological gene set. This anomaly may reflect the improved sensitivity for prognostic gene set detection by algorithm-based methods compared with the more biased hand-picking approach.

DISCUSSION AND COMMENTARY

In the most recent review of melanoma GEMA prognostic signatures, Timar *et al.* (2010) found little in common between them, and they highlighted problems around sample heterogeneity, tissue and survival data availability, lack of independent validation, and platform and statistical differences as the limiting factors. In the present review, we show that work in the field has progressed considerably in the past 18 months. First, rigorous new studies have been published, thereby increasing the number of data sets available for examination and comparison. Second, we have been able to select studies for inclusion in our review based on their fulfillment of a simple set of minimum criteria, developed to identify studies in the area having the greatest translational relevance.

Table 2. Genes in common, pairwise (identified using MetaCore from GeneGo), between selected proposed prognostic signatures in human melanoma

	Winnepenninckx *et al.* (2006) 254-object signature (*n*=207)	John *et al.* (2008) 21-object signature (*n*=16)	John *et al.* (2008) 5-gene qPCR signature (*n*=5)	Bogunovic *et al.* (2009) 266-object signature (*n*=340)	Conway *et al.* (2009) Osteopontin (*n*=1)	Jönsson *et al.* (2010) 30-gene immune signature (*n*=29)	Jönsson *et al.* (2010) 377-gene molecular subtype discriminator (*n*=349)
Winnepenninckx *et al.* (2006) 254-object signature (*n*=207)							
John *et al.* (2008) 21-object signature (*n*=16)	*MRPS5*						
John *et al.* (2008) 5-gene qPCR signature (*n*=5)	None	NA					
Bogunovic *et al.* (2009) 266-object signature (*n*=340)	*ATAD2* Anillin Exo1 *LTB* *MCM4* ***PGHD*** *TXNIP*	None	None				
Conway *et al.* (2009) Osteopontin (*n*=1)	None	None	None	None			
Jönsson *et al.* (2010) 30-gene immune signature (*n*=29)	None	None	None	*CD79A* *LCK* *SKAP55*	None		
Jönsson *et al.* (2010) 377-gene molecular subtype discriminator (*n*=349)	C12orf24 *CLIC3* *HAI-2* *HLA-DQB1* *ICAT* *MHC* class II *MHC* class II (β-chain) ***PGHD*** *PROM2*	*ITPA*	None	*AADAT* *ARHGAP30* *CALDAG-GEFI* *CDK2* *DLX1* *FLJ11193* G-alpha(s)-specific, prostanoid GCPRs Granzyme K *HLA-DPB1* *ICOS* Ikaros *LCK* *MATP* *MHC* class II *PGE2R2* ***PGHD*** *VNN2*	None	*CD19* *CD3* *CD3* epsilon *LAX1* *LCK*	
Mann *et al.* (2010) 60-probe (46-gene) signature (*n*=53)	None	None	None	*ADAMDEC1* *CD2* *CD3 zeta* *CD3 delta* *CD79* *FAT10* *GBP2* *IRF8* Pleckstrin *VCAM1*	None	*CD79* complex	*GBP4* *Gimap4* *PM17* *TSPAN10*

Abbreviations: NA, not applicable; qPCR, quantitative PCR.
Bracketed *n*-values indicate the number of genes from that particular signature that were positively identified by the MetaCore program. Genes in each box are common between the signatures at that particular intersection. Genes in bold type are common to more than two signatures.
See text for inclusion criteria.

We observed, perhaps unexpectedly, similarities between our carefully selected data sets despite their different platforms and statistical methods. Most of the signatures we examined contained an immune response theme as an overrepresented process. Considering the general body of immunohistochemistry-based examination of melanoma proteins in a prognostic setting (Gould Rothberg *et al.*, 2009b), molecules involved in altered immunocompetence have not been as

Table 3. Leave-one-out cross-validation (LOOCV) error rates for misclassification of patients into outcome-related classes in our cross-validation of gene signatures between independent data sets

Expression data set → **Signature object list ↓**	Bogunovic *et al.* (2009)	John *et al.* (2008)	Jönsson *et al.* (2010)	Mann *et al.* (2010)	Winnepenninckx *et al.* (2006)
Bogunovic *et al.* (2009) (*n*=340)	0.18	0.17	0.20	0.54	0.48
John *et al.* (2008) (*n*=16)	0.24	0.13	0.20	0.27	0.48
Jönsson *et al.* (2010) (*n*=29; *a priori*-defined immune signature)	0.39	0.38	0.17	0.46	0.57
Jönsson *et al.* (2010) (*n*=349; molecular subtype distinguisher)	0.42	0.08	0.17	0.38	0.74
Mann *et al.* (2010) (*n*=46)	0.36	0.13	0.30	0.35	0.35
Winnepenninckx *et al.* (2006) (*n*=207)	0.39	0.50	0.20	0.44	0.70

Bracketed *n*-values indicate the number of objects (e.g., genes, probes, and gene elements) assessed in each particular signature.
See Supplementary Methods online for details.

actively studied (Schramm and Mann, 2011) as proteins from other Hanahan–Weinberg (Hanahan and Weinberg, 2000) cancer capability functional groups (e.g., tissue invasion and metastasis or self sufficiency in growth signals). This finding underscores the importance of a detailed, ongoing examination of the tumor microenvironment for biological correlates with prognosis, as well as potential therapy targets, as has been argued by others (Polak *et al.*, 2009; Bedognetti *et al.*, 2010; Gajewski *et al.*, 2010). It also reiterates the rational platform for identification of immune response-based predictive/prognostic signatures as demonstrated by Jönsson *et al.* (2010) with their *a priori*-defined gene list, although this particular hand-picked signature did not perform as well in formal statistical validation as the other, less-biased gene lists. Is it possible, however, to make a distinction between the immune response *per se* (e.g., CD3 and other markers of infiltrating cells, which could parallel and extend the analysis of tumor-infiltrating lymphocytes, known to be prognostic in both primary and metastatic melanomas; Payette *et al.*, 2009) and genes likely derived from the tumor cells themselves? At least one such attempt has been made by Gould Rothberg *et al.* (2009b) who, in their review of immunohistochemistry-based prognostic biomarkers in melanoma, supplemented the Hanahan–Weinberg cancer capability classification system (Hanahan and Weinberg, 2000) with two additional melanoma-specific categories: ''altered immunocompetence'' and ''melanocyte differentiation.'' Molecules of prognostic interest were then assigned to each of those groups (annexin, MAGE3, and STAT1 to the former and tyrosinase, MART1, and gp100 to the latter) based on extensive literature review of their function.

Some of the signatures we reviewed predicted outcomes in other, equally eligible, expression data sets with notably low rates of misclassification. We propose that this cross-validation illustrates the biological relevance of those particular gene lists in predicting outcome for selected subsets of melanoma patients. Our finding is also consistent with observations made by Weigelt *et al.* (2010) who, in a review of the contribution of gene expression profiling to breast cancer classification, reported that distinct prognostic signatures for specific subgroups of breast cancer may be necessary to accurately define their outcome. One such subset could involve patients with a ''proliferation'' signature. Weigelt *et al.* (2010) and other groups have observed that breast cancer patients with the highest expression of ''proliferation cluster'' genes appear to derive most benefit from chemotherapy. An outcome-related proliferation subtype comprising 100 genes/gene elements was proposed in the Jönsson study and we keenly await the results of further examination in that area. We also look forward to further investigation of outcome or potential therapy-related signatures within the subset of *BRAF*-mutated melanomas, as the distinction of these tumors as a separate molecular subgroup becomes clearer.

A prevailing question around GEMAs in clinical practice is whether they can outperform current clinicopathological markers of outcome. For example, in the Winnepenninckx study, the prognostic capacity of the gene expression data was marginally inferior to tumor thickness and to the presence of ulceration. In contrast, Mann *et al.* (2010). have more recently presented an outcome prediction model that is optimized when both clinicopathological factors and gene expression data are used in combination (error rate = 0.23) rather than alone (error rates = 0.27 and 0.25, respectively). It has been argued that the primary purpose of a new marker is to augment currently available knowledge (Dalton and Kattan, 2010). Therefore, novel gene signatures should continue to be assessed in a multivariable setting alongside existing factors to evaluate any additive prognostic contribution of the tests combined. Efforts to begin building databases that present gene expression data alongside corresponding patient clinical data are in early stages; e.g., the integrated clinical omics database (iCOD; Shimokawa *et al.*, 2010). As is the case for breast cancers, the value of signatures in clinical decision making remains to be elucidated using much larger data sets and rigorous clinical trial experimental designs (discussed in more detail elsewhere; Pepe *et al.*, 2008).

One of the most pressing questions moving forward is whether expression-based predictive/prognostic signatures will continue to add value in clinical practice in a new era of high-throughput genomic analyses. We agree with others

(e.g., Tinker *et al*., 2006; Madabhushi *et al*., 2011) that the most clinically valid and useful prognostic models will probably continue to be derived from a combination of techniques as a function of their costs and effectiveness. For example, the demonstrated relevance of protein signatures by immunohistochemistry profiling (such as in Kreizenbeck *et al*., 2008; Piras *et al*., 2008; Gould Rothberg *et al*., 2009a) and the promise of computer-assisted image analysis (Alexe *et al*., 2009) all suggest that no single technique will supplant all of the others. Many of the lessons learned through the process of integrating useful gene expression signatures with traditional prognostic factors are also likely to be applicable to subsequent technologies.

Among the exciting developments for better understanding melanoma is work being undertaken in the burgeoning field of cancer systems biology (Hawkins *et al*., 2010). Questions such as whether we can exploit our observations of local and global melanoma gene or protein network properties to benefit patients are pertinent (Faratian *et al*., 2009). For example, using breast cancer gene expression microarray data, Taylor *et al*. (2009) have proposed a method that reveals that altered modularity of the human interactome may be useful as a prognostic indicator. In melanoma research, the output of transcriptomic, proteomic, and miRNA signatures is beginning to surge (Glud *et al*., 2009; Han *et al*., 2010; Jukic *et al*., 2010; Philippidou *et al*., 2010). Coupled with ongoing developments in gene/protein network construction algorithms (Djebbari and Quackenbush, 2008; Han, 2008; Kreeger and Lauffenburg, 2010), we are increasingly well placed to better understand the complex interplay of separate biological components via the network analysis approach.

In their commentary on breast cancer GEMA, Weigelt *et al*. (2010) report that, ''In hindsight, the pitfalls of the approaches chosen to define signatures predicting response, both methodological and conceptual, have become clear to many.'' In the present review, we show that the benefit of this hindsight has not been ignored by the melanoma research community. Rather, we have reaffirmed the ongoing relevance of carefully designed GEMAs to improving disease management in the era of personalized medicine, particularly for advanced-stage disease.

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

We are grateful to all our colleagues in the melanoma program for their support and intellectual engagement. We are also grateful to two anonymous reviewers for their valuable observations, some of which we have included verbatim. This study was supported by a University of Sydney Postgraduate Award and a Sydney Medical School Research Top-up Scholarship (to S-JS), an Australian Postgraduate Award (to AEC), an ARC Future Fellowship (to YHY), a Cancer Institute of New South Wales Clinical Research Fellowship (to RAS), and program grants of the National Health and Medical Research Council of Australia, and Cancer Institute of New South Wales (to RAS and GJM).

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at http://www.nature.com/jid

REFERENCES

Alexe G, Monaco J, Doyle S *et al*. (2009) Towards improved cancer diagnosis and prognosis using analysis of gene expression data and computer aided imaging. *Exp Biol Med* 234:860–79

Allison DB, Cui XQ, Page GP *et al*. (2006) Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 7:55–65

Arkenau HT, Kefford R, Long GV (2010) Targeting BRAF for patients with melanoma. *Br J Cancer* 104:392–8

Balch CM, Buzaid AC, Soong SJ *et al*. (2001) Final version of the American Joint Committee on Cancer staging system for cutaneous melanoma. *J Clin Oncol* 19:3635–48

Balch CM, Gershenwald JE, Soong SJ *et al*. (2009) Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol* 27:6199–206

Bedognetti D, Wang E, Sertoli MR *et al*. (2010) Gene-expression profiling in vaccine therapy and immunotherapy for cancer. *Expert Rev Vaccines* 9:555–65

Bittner M, Meitzer P, Chen Y *et al*. (2000) Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 406:536–40

Bogunovic D, O'Neilla DW, Belitskaya-Levya I *et al*. (2009) Immune profile and mitotic index of metastatic melanoma lesions enhance clinical staging in predicting patient survival. *Proc Natl Acad Sci USA* 106: 20429–34

Conway C, Mitra A, Jewell R *et al*. (2009) Gene expression profiling of paraffin-embedded primary melanoma using the DASL assay identifies increased osteopontin expression as predictive of reduced relapse-free survival. *Clin Cancer Res* 15:6939–46

Dalton JE, Kattan MW (2010) Recent advances in evaluating the prognostic value of a marker. *Scand J Clin Lab Invest* 70:59–62

Djebbari A, Quackenbush J (2008) Seeded Bayesian Networks: constructing genetic networks from microarray data. *BMC Syst Biol* 2:1–13

Faratian D, Clyde RG, Crawford JW *et al*. (2009) Systems pathology–taking molecular pathology into a new dimension. *Nat Rev Clin Oncol* 6: 455–64

Flaherty KT, Puzanov I, Kim KB *et al*. (2010) Inhibition of mutated, activated BRAF in metastatic melanoma. *N Engl J Med* 363:809–19

Gajewski TF, Louahed J, Brichard VG (2010) Gene signature in melanoma associated with clinical activity a potential clue to unlock cancer immunotherapy. *Cancer J* 16:399–403

Glud M, Klausen M, Gniadecki R *et al*. (2009) MicroRNA expression in melanocytic nevi: the usefulness of formalin-fixed, paraffin-embedded material for miRNA microarray profiling. *J Invest Dermatol* 129:1219–24

Gould Rothberg BE, Berger AJ, Molinaro AM *et al*. (2009a) Melanoma prognostic model using tissue microarrays and genetic algorithms. *J Clin Oncol* 27:5772–80

Gould Rothberg BE, Bracken MB, Rimm DL (2009b) Tissue biomarkers for prognosis in cutaneous melanoma: a systematic review and meta-analysis. *J Natl Cancer Inst* 101:452–74

Han J-DJ (2008) Understanding biological functions through molecular networks. *Cell Res* 18:224–37

Han M-J, Wang H, Beer LA *et al*. (2010) A systems biology analysis of metastatic melanoma using in-depth three-dimensional protein profiling. *Proteomics* 10:4450–62

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

Haqq C, Nosrati M, Sudilovsky D *et al*. (2005) The gene expression signatures of melanoma progression. *Proc Natl Acad Sci USA* 102:6092–7

Hawkins RD, Hon GC, Ren B (2010) Next-generation genomics: an integrative approach. *Nat Rev Genet* 11:476–86

Hoek KS (2007) DNA microarray analyses of melanoma gene expression: a decade in the mines. *Pigm Cell Res* 20:466–84

John T, Black MA, Toro TT *et al*. (2008) Predicting clinical outcome through molecular profiling in stage III melanoma. *Clin Cancer Res* 14: 5173–80

Jönsson G, Busch C, Knappskog S *et al*. (2010) Gene expression profiling-based identification of molecular subtypes in stage IV melanomas with different clinical outcome. *Clin Cancer Res* 16:3356–67

Jukic DM, Rao UNM, Kelly L *et al.* (2010) Microrna profiling analysis of differences between the melanoma of young adults and older adults. *J Transl Med* 8:23

Kannengiesser C, Spatz A, Michiels S *et al.* (2008) Gene expression signature associated with BRAF mutations in human primary cutaneous melanomas. *Mol Oncol* 1:425–30

Kauffmann A, Rosselli F, Lazar V *et al.* (2008) High expression of DNA repair pathways is associated with metastasis in melanoma patients. *Oncogene* 27:565–73

Kefford R, Arkenau H, Brown MP *et al.* (2010) Phase I/II study of GSK2118436, a selective inhibitor of oncogenic mutant BRAF kinase, in patients with metastatic melanoma and other solid tumors. *J Clin Oncol* 28:15s

Kreeger PK, Lauffenburg DA (2010) Cancer systems biology: a network modeling perspective. *Carcinogenesis* 31:2–8

Kreizenbeck GM, Berger AJ, Subtil A *et al.* (2008) Prognostic significance of cadherin-based adhesion molecules in cutaneous malignant melanoma. *Cancer Epidemiol Biomarkers Prev* 17:949–58

Lee JH, Choi JW, Kim YS (2010) Frequencies of BRAF and NRAS mutations are different in histologic types and sites of origin of cutaneous melanoma: a meta-analysis. *Br J Dermatol* 164:776–84

Madabhushi A, Lucas D, Rodriguez-Canales J *et al.* (2011) Computer aided diagnostic tools aim to empower rather than replace pathologists: lessons learned from computational chess. *J Pathol Inform* 2:25–7

Mandruzzato S, Callegaro A, Turcatel G *et al.* (2006) A gene expression signature associated with survival in metastatic melanoma. *J Transl Med* 4:1–11

Mann GJ, Pupo GM, Campain AE *et al.* (2010) *BRAF* mutation, *NRAS* mutation and absence of an immune-related expressed gene profile predict poor outcome in surgically resected stage III melanoma. *Pigment Cell Melanoma Res* 23:6

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

Payette MJ, Katz M, Grant-Kels JM (2009) Melanoma prognostic factors found in the dermatopathology report. *Clin Dermatol* 27:53–74

Pepe MS, Feng Z, Janes H *et al.* (2008) Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *J Natl Cancer Inst* 100:1432–8

Perou CM, Sorlie T, Eisen MB *et al.* (2000) Molecular portraits of human breast tumours. *Nature* 406:747–52

Philippidou D, Schmitt M, Moser D *et al.* (2010) Signatures of microRNAs and selected microRNA target genes in human melanoma. *Cancer Res* 70:4163–73

Piras F, Perra MT, Murtas D *et al.* (2008) Combinations of apoptosis and cell-cycle control biomarkers predict the outcome of human melanoma. *Oncol Rep* 20:271–7

Polak ME, Borthwick NJ, Jager MJ *et al.* (2009) Melanoma vaccines-the problems of local immunosuppression. *Hum Immunol* 70:331–9

Schramm S-J, Mann GJ (2011) Melanoma prognosis: a REMARK-based systematic review and bioinformatic analysis of immunohistochemical and gene microarray studies. *Mol Cancer Ther* 10:1520–8

Shendure J (2008) The beginning of the end for microarrays? *Nat Methods* 5:585–7

Shimokawa K, Mogushi K, Shoji S *et al.* (2010) iCOD: an integrated clinical omics database based on the systems-pathology view of disease. *BMC Genomics* 11:S19

Subramanian J, Simon R (2010) Gene expression-based prognostic signatures in lung cancer: ready for clinical use? *J Natl Cancer Inst* 102: 464–74

Taylor IW, Linding R, Warde-Farley D *et al.* (2009) Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat Biotechnol* 27:199–204

Thompson JF, Scolyer RA, Kefford RF (2009) Cutaneous melanoma in the era of molecular profiling. *Lancet* 374:362–5

Timar J, Gyorffy B, Raso E (2010) Gene signature of the metastatic potential of cutaneous melanoma: too much for too little? *Clin Exp Metastasis* 27:371–87

Tinker AV, Boussioutas A, Bowtell DDL (2006) The challenges of gene expression microarrays for the study of human cancer. *Cancer Cell* 9: 333–9

van't Veer LJ, Dai HY, van de Vijver MJ *et al.* (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415:530–6

Varmus H (2010) Ten years on–the human genome and medicine. *N Engl J Med* 362:2028–9

Weigelt B, Baehner FL, Reis JS (2010) The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. *J Pathol* 220:263–80

Winnepenninckx VR, Lazar V, Michiels S *et al.* (2006) Gene expression profiling of primary cutaneous melanoma and clinical outcome. *J Natl Cancer Inst* 98:472–82