Melanoma Prognosis: A REMARK-based systematic review and bioinformatic analysis of immunohistochemical and gene microarray studies

Sarah-Jane Schramm¹, and Graham J. Mann¹

¹University of Sydney at Westmead Millennium Institute and Melanoma Institute Australia, Westmead, NSW, Australia

Running title: Prognostic biomarkers in melanoma

Keywords: melanoma / biomarker / prognosis / REMARK / gene microarray / immunohistochemistry/bioinformatic

Abbreviations: CMM, cutaneous malignant melanoma; REMARK, REporting recommendations for tumor MARKer prognostic studies; NCI-EORTC, National Cancer Institute-European Organisation for Research and Treatment of Cancer; IHC, immunohistochemistry; FDR, false discovery rate; CI, confidence interval; MSM, melanoma-specific mortality; ACM, all-cause mortality; DFS, disease-free survival; DMFS, distant metastasis-free survival; MPD, multimarker prognostic discriminator

Financial support: No external funding was used for this study

Corresponding author / reprint requests: Sarah-Jane Schramm, Melanoma Genomics and Genetic Epidemiology Research Group, University of Sydney at Westmead Millennium Institute, PO Box 412, Westmead, NSW, 2145. e-mail: sj_schramm@wmi.usyd.edu.au. phone: +61 2 9845 9000. fax. +61 2 9845 9100

Disclosure statement: The authors declare no competing financial, personal or other interests.
Abstract

Despite intensive research efforts, within-stage survival rates for melanoma vary widely. Pursuit of molecular biomarkers with improved prognostic significance over clinico-histological measures has produced extensive literature. Reviews have synthesized these data but none have systematically partitioned high-quality studies from the remainder across different molecular methods nor examined system properties of that output. Databases were searched for studies analyzing protein expression by immunohistochemistry (IHC) (n=617, extending the only systematic review to date (1) by 102 studies), or gene expression microarray studies (n=45) in melanoma, in relation to outcome. REMARK-derived criteria were applied to identify high-quality studies. Biomarkers and pathways were functionally assessed using gene ontology software. Most manuscripts did not meet REMARK-based criteria, an ongoing trend that can impede translational research. Across REMARK-compliant literature, 41 proteins were significantly associated with outcome. Multimarker tests consistently emerged among the most promising potential biomarkers, indicating a need to continue assessing candidates in that composite setting. Twenty-one canonical pathways were populated by outcome-related proteins but not by those that failed to show such an association; we propose this set of pathways warrants closer investigation to understand drivers of poor outcome in melanoma. Two gene expression microarray studies met REMARK-based criteria reflecting a genuine paucity of literature in the area. The 254 outcome-related genes were examined for correspondences with the systematically identified protein “signature”. This analysis highlighted PCNA and Survivin as priorities for further examination as biomarkers in melanoma prognosis, and, illustrated ongoing need to integrate alternative approaches to biomarker discovery in melanoma translational research.
Introduction

Melanoma is a malignant proliferation of melanocytes arising from an accumulation of genetic mutations within their onco- and/or tumor suppressor genes. The cancer is a major health problem in several countries due to high incidence (2). Excepting the case of early detection and treatment, cutaneous malignant melanoma (CMM) is an aggressive and therapy-resistant cancer (3) with surgical resection remaining the mainstay of treatment options. Patient outcome after tumor resection can be presumed to depend on a complex sum of patient and tumor characteristics, some of which may be strongly determined by tumor and germline genotype (3). However, prognosis is currently assigned almost entirely on the basis of a limited set of clinical and histological features (4). While it is useful in assigning broad probabilities of relapse, it has limited predictive power at the level of individual patients and has no direct implications for personalising therapy.

Research has focused heavily on assessing protein and, more recently, expressed gene biomarkers in order to classify melanomas into subgroups with distinct molecular features. To have translational relevance, such subgroups should be readily and reproducibly identifiable, show clear and independent relationships with specific survival outcomes and, do so with greater discriminating power than the current system. The hundreds of studies in this area have been generally small and underpowered and have not yet provided a comprehensive picture of CMM. The 2010 implementation of the AJCC Staging System revision (4) therefore includes no novel molecular prognostic factors.

The ongoing exclusion of molecular markers from prognostication has been explained by inconsistent results between studies due to critical methodological differences (5). To that end, the REporting recommendations for tumor MARKer prognostic studies (REMARK) were published simultaneously in five major journals in 2005 e.g. (5). The development of these guidelines was a major proposition of the National Cancer Institute-European Organisation for Research and Treatment of Cancer (NCI-EORTC) First International
Meeting on Cancer Diagnostics in 2000. So far only one study (1) has used the guidelines to systematically partition high-quality, prognostic biomarker studies from the remainder. In that study of immunohistochemistry (IHC)-based research, Gould Rothberg et. al distinguished just 37 manuscripts from 1797 citations identified in a systematic search up to and including January 15, 2008. From a total of 62 proteins assessed in the REMARK-compliant studies, 30 showed significant associations with at least one measure of prognosis. Based on consistency across outcome measures the most promising prognostic markers were (MCAM)/MUC18, MMP-2, KI67, PCNA, and p16/INK4A.

However work in the field is accelerating. We have analyzed all REMARK-compliant studies in the subsequent two years and report that of a total 80 protein biomarkers tested, 41 have shown a significant relationship with prognosis in primary melanoma. We then applied the REMARK criteria, for the first time, to studies involving gene expression microarray analysis of CMM.

Further, we compared the pool of molecules extracted from IHC-based studies with the eligible gene expression microarray data. Bioinformatic analysis of all eligible data revealed differences and similarities between datasets from candidate-based research (the IHC data) and the unbiased gene expression microarray work. It also aided in identifying biomarkers with the most promising translational potential i.e. gene/protein elements whose relationship with CMM prognosis is strong and clear and supported by multiple, independent lines of evidence. Lastly, we used a systems-based analysis to examine which tumor-sustaining biochemical pathways are being illuminated by this composite of the premier literature in the field.

**Materials and methods**

Identification of high quality studies

IHC-based research
To update the work of Gould Rothberg et al. (1) an analysis of IHC-based literature was conducted using their methods but for the period between their search freeze date (15 January, 2008) and 1 December, 2009. Data were extracted and redacted as described (1). Proteins not already designated into one of the 8 modified Hanahan-Weinberg functional capabilities (as described (6) and modified (1)), were so classified based on the description of the molecule's function(s) given in MetaCore™ (GeneGO, Encinitas, CA).

Gene expression microarray-based research

The above search and review process was repeated for all primary research articles describing the application of gene expression microarray technology in the context of melanoma prognosis, using the Web of Science literature database on 27 November, 2009. The primary query was (melanoma) AND (dna microarray* OR rna microarray* OR cdna microarray* OR gene-expression profil* OR gene expression profil* OR gene-expression signature* OR gene expression signature* OR gene microarray*) AND (prognos* or surviv* or progress* or clinic*) NOT (uea*). Based on work by Allison et al. (7), the REMARK/Gould Rothberg criteria were adapted for application to gene expression microarray studies; manuscripts were included in our bioinformatic analysis if the research:

1. involved assay of primary cutaneous melanoma tissue specimens via gene expression microarrays;
2. was undertaken in the context of a prospective or retrospective cohort design with a clearly defined source population and justifications for all excluded eligible cases;
3. contained clearly described methods about RNA/DNA extraction, the array platform used, and, the choice of data pre-processing methods including image analysis, normalization and data transformation techniques;
4. dealt with multiple hypothesis testing e.g. by implementing false discovery rate (FDR) control methods;
5. involved statistical analysis that provided a gene expression feature and associated prediction rule where the predicted class was a prognosis-related outcome such as survival or disease-free progression;

6. included validation of that outcome-related gene expression feature and classification rule using an independent dataset; and,

7. involved comparison of the utility of that classifier with current prognostic factors (multivariate analysis).

As was the case for the IHC-based systematic review, studies were rendered ineligible for inclusion where:

1. they were specific to the assay of acral lentiginous, mucosal, ocular, uveal and/or choroidal melanomas;

2. they predominantly involved non-Caucasian populations in the patient sample;

3. gene expression levels were not from melanoma cells or were limited to tumor-associated stroma or vasculature; and,

4. the same gene microarray data had been used as part of a larger study.

Bioinformatic analysis

To search for prominent biochemical pathways and biological processes enriched by the eligible data, an analysis was performed on each of the gene expression and protein datasets using the MetaCore™ GeneGO™ software Pathway Maps option (default settings). Gene and protein data were then compared with each other using the Compare Experiments algorithm (default settings) which probes the datasets for intersections. Intersections were closely examined to determine whether findings between studies generally agreed in terms of the significance and direction of any outcome-associated expression change. In each case, the sum of evidence for each molecule of interest was deemed either ‘concordant’ (where eligible data from independent studies of that molecule generally agreed with each other), ‘discordant’ (indicating an apparent lack of agreement between independent studies), or, ‘unclear’ (where eligible data from independent studies could not easily be assessed as either concordant or
discordant). Molecules for which gene expression and protein data were not ‘concordant’, in terms of the significance and direction of outcome-associated expression changes, were classified as such. In those situations, the different findings do not necessarily indicate a lack of agreement between the studies but rather, could imply that protein expression may be regulated without a correlated gene transcript expression difference.

Results

Excluded studies

The literature review strategy devised by Gould Rothberg et al. (1) was repeated for the subsequent 23 months. That process identified 284 citations. After evaluating all titles and abstracts, 102 manuscripts were downloaded and systematically assessed in accordance with the REMARK (reported in Supplement 1). Studies were excluded for the following reasons: IHC was not performed (n=27) or IHC methods were incomplete (n=27); primary tumors (n=8) or tumor compartments (n=11) were not the staining target; research involved predominantly non-Caucasians (n=35); study design was case-control (n=27), case-series (n=58) or cross-sectional (n=315); data were redundant (n=6); or only univariate analysis was reported (n=34) (Supplement 2). Where multivariate analysis was undertaken against current prognostic factors, 17 studies did not report an associated HR and/or confidence interval (CI); five of these manuscripts (identified in Supplement 2) were assessed as part of the present study. They involved 8 proteins, of which Akt, Erk, PAR1, and TF were significantly associated ($P \leq 0.05$) with clinical outcome and would have been novel additions to the list of outcome-associated inclusions had the authors provided an HR (CI).

The novel search strategy for gene expression microarray-based literature in the context of melanoma prognosis returned 421 citations (excluding reviews and editorial material). After review of these titles and abstracts, 45 manuscripts were downloaded and assessed in accordance with the REMARK-based criteria described in the methods above (summarised in Supplement 1, reported in Supplement 3). Of these studies, three were excluded as case-series
research where investigators failed to provide details on either the source population of melanocytic tumors or the sampling strategy. Eleven studies were excluded as cross-sectional where they were limited to determining the association between levels of marker expression with melanocytic lesion progression e.g. metastasis signatures. A majority of studies, 17 in total, was excluded because they did not examine primary melanoma. In two of these cases, primary melanoma was examined during a validation phase but was not initially used to generate an outcome-related gene expression signature. One of those studies (8) derived a 150-gene signature from cell lines and used primary tumors as part of their validation set. In this otherwise eligible study, the predictive utility of the gene set could not be significantly associated with clinical outcome in primary tumors ($P > 0.182$), a result attributed to insufficient statistical power and awaiting validation in a larger sample pool. In the other study (9), researchers identified a significant association between high CXCR4 gene expression in colorectal cancer tissue and poor clinical outcome. They then examined this feature in primary melanoma but could not demonstrate an associated significant difference in overall survival for that cohort. Two studies were classified as case-control. These involved a comparison of clinically unaffected skin between patients with melanoma and healthy controls (10), or of deep penetrating nevus and nodular malignant melanoma (11). Five manuscripts were excluded because gene expression microarray analysis of melanoma samples was not undertaken. Three studies were excluded due to data redundancy with previous reports. Two studies were limited to univariate analysis, one in which PLZF was shown to be a significant predictor of long term survival in that context (12). In the other study (13), lower expression of ASK/Dbf4 appeared to be associated with improved relapse-free survival.

**Included studies**

Taken together with the work of Gould Rothberg et al. (1), and in accordance with the REMARK-derived criteria, a total of 51 high-quality, IHC-based melanoma prognostic studies have now been formally distinguished from the remainder with 14 of those studies
identified in the current review. Altogether, these 51 studies have evaluated 80 unique proteins, 18 identified in the present review.

All research groups relied on the use of formalin-fixed, paraffin-embedded tissue (FFPE) for their investigations. Whereas a minority of studies identified in the original review (1) were performed using whole slides (38%), the majority of studies from the last 23 months (64%) employed tissue microarray-based IHC techniques. Sample size for the 14 newly identified studies ranged from 46 to 301, consistent with those previously evaluated (1). The largest study remains that conducted by Weinlich et al. in 2006 (28) involving more than 1270 patients. For the first time some studies (18, 21, 22) evaluated the combined effect of 3 or more markers in a multivariate setting (multimarker prognostic discriminators, MPDs). It is notable that RGS1, β-catenin, p16INK4a, p21WAF1, p53, SPP1/osteopontin and NCOA3 are the only eligible biomarker candidates to have been investigated by more than one research group, with respect to at least two outcomes, and, as part of an MPD set.

Novel factors being used to test for biomarker independence in the more recent eligible studies include: CDK4 expression (20); sun exposure to anatomic site (21); markers of cell adhesion (CD44, αv-integrin) and hypoxia (TNFα and Apaf1) (14); p16 expression and tumor cell proliferation (via KI67 expression) (14, 20); neurotropism (26); and, metastasis during follow up (25). The clinical covariate most frequently accounted for as part of multivariate assessment was Breslow thickness with all recent studies including this factor in their analysis, followed by ulceration, gender, age and Clark level (11, 9, 9 and 6% of recent studies, respectively). Mitotic rate was accounted for much more frequently in the eligible studies conducted in the last 23 months (29%) than in those previously reviewed (1). The median number of covariates assessed alongside protein expression per study has increased from 4 (range = 1 to 8) to 6.5 (range = 1 to 10).
The 33 unique proteins examined in eligible manuscripts in this review are presented in Supplement 4. Of these proteins, 18 (nucleolin, SNF5, α,β-integrin/CD51, LDH5, ING4, GRP78, HER3, α-catenin, annexin1, β-catenin, fibronectin, hairy/enhancer of split-related1, integrin-linked kinase (Ilk), MMP1, E-cadherin, Rho-c, RGS1 and p120-catenin) were not among the 62 eligible species identified in the Gould Rothberg et al. (1) review. Of those 18 molecules, 9 were found to have a significant, independent association with clinical outcome (nucleolin, SNF5, α,β-integrin/CD51, ING4, HER3, α-catenin, β-catenin, fibronectin and RGS1) either individually or as part of an MPD set. Two molecules (BMI1 and p21/WAF1), for which there was insufficient evidence for a relationship with outcome in the original review, now have new evidence supporting significance. This brings the total number of protein biomarkers with REMARK-criteria supported evidence for association with outcome in primary melanoma to 41.

In the present analysis, the most certain (lowest \( P \)-value) observed effects on outcome were exhibited by the MPDs: p16/INK4a, together with Survivin and p53 (18); β-catenin combined with fibronectin, ATF2, p16/INK4a, and p21/WAF1 (21); and, RGS1 and SPP1/osteopontin with NCOA3 (22). Strongly outcome-related individual biomarkers included RGS1 (melanoma-specific mortality (MSM), incorporating 7 covariates) (19), nucleolin (all-cause mortality (ACM), 3 dominant clinical covariates) (16), HER3 (ACM, 4 covariates) (25), ING4 (ACM, 6 covariates) (23), and β-catenin (MSM, 7 covariates) (21). The potential utility of SNF5 (ACM) was obscured by inclusion of just one covariate in the multivariate test (24). Tests involving α-catenin showed no significance in a univariate analysis but significant association with outcome in a multivariate setting (21). α,β-integrin/CD51 was reported with borderline significance (MSM, \( P = 0.049 \)) (14).

Of the studies involving gene expression microarray-based examination of primary melanoma, only two (29, 30) met all of the REMARK-derived criteria. In the most recent of these works (29), authors used FFPE primary melanoma tissue to identify osteopontin (SPP1)
expression as predictive of relapse-free survival in the training set. This predictive utility, however, was not maintained in a validation set in a multivariate context accounting for the most important clinical-histological predictive markers. The second eligible study (30) involved analysis of 58 primary cutaneous samples and included class prediction from which a 254-gene expression signature associated with 4-year distant metastasis-free survival (DMFS) was identified. The performance of the 254-gene classifier was validated using the same method in an independent sample set of primary cutaneous melanomas where 11 out of 17 cases were correctly categorized with respect to DMFS. The potential clinical utility of the classifier was then tested against the prediction accuracy of tumor thickness and ulceration which misclassified 28% of patients while the 254-gene classifier misclassified 29% of cases.

In that same study, IHC was used to examine protein expression in relation to outcome for several of the 254-gene classifier elements. Although that validation involved a large independent sample set (n=176) and took place in a multivariate setting, authors used overall survival as an endpoint rather than DMFS making it difficult to directly compare those results with the gene expression data.

Bioinformatic analysis

Protein data

We first analysed the pathway map and functional characteristics of all the proteins tested in the eligible studies (Supplement 5). These were predicted to correspond to mechanisms underlying plausible hypotheses about causation of poor prognosis, namely critical features of melanoma progression, such as proliferation, invasion, adhesion and metastasis. Our data confirm that they had been selected on that basis. Using all eligible proteins identified in the study by Gould Rothberg et al. (1) and the present review (n = 80), 103 canonical pathway maps were found to be significantly enriched by up to 11 of these species (FDR = 0.05). The most heavily populated maps involved the cell cycle with regulation of G1/S transition being the most significant ($P = 7.522e-14$, 11 eligible proteins from map total of 38). Other popularly investigated pathways included those involving transcription, cell adhesion,
development, DNA damage, apoptosis and survival, signal transduction, cytoskeleton remodelling, and immune response.

In terms of the modified Hanahan-Weinberg functional capabilities (6), most of the proteins examined were allocated to tissue invasion and metastasis (23 candidates) or limitless replicative potential (21 candidates). Remaining proteins were distributed among self-sufficiency in growth signals (12 molecules), insensitivity to growth-inhibitory (antigrowth) signals (10 molecules), evasion of programmed cell death (apoptosis) (5 proteins), altered immunocompetence (3 proteins), melanocyte differentiation (3 proteins), and sustained angiogenesis (2 proteins).

All eligible protein species with a significant association with clinical outcome were then identified and examined together. That analysis generated 65 maps that were significantly \( (P < 0.05) \) enriched with between 2 and 6 outcome-related proteins (Supplement 5). The most populated canonical map was the transcription-regulating, P53 signalling pathway \( (P = 6.391\text{e}-11 \text{ with } 6 \text{ out of the } 39 \text{ proteins on that map significantly associated with outcome}) \). Twenty-one pathways were significantly populated by the marker proteins linked to melanoma outcome, but not by those that failed to show such an association (Supplement 5).

Gene data

MetaCore\textsuperscript{TM} software recognised 207 of the 254 gene (element) IDs present in the only gene expression signature deemed eligible for inclusion in this review (30). Approximately half of these genes (110 out of 207) appeared on MetaCore\textsuperscript{TM} canonical pathway maps. Altogether, 20 maps were significantly enriched by the data (Supplement 6). The most significantly populated of these maps was the cell cycle metaphase checkpoint \( (P = 1.799\text{e}-14, 11 \text{ genes from map total of } 36) \). Of the remaining significantly enriched pathways, 11 were related to cell cycle while cellular processes such as apoptosis, DNA damage and immune response were also reported.
Data intersections

The gene and protein data were compared to identify complementary or divergent lines of evidence that might further implicate prognostic relevance of a candidate biomarker and hence warrant further examination. For each molecule, we analysed the extent to which the eligible IHC-based and gene expression microarray-based data were concordant in terms of the significance and/or direction of any outcome-related effect (Figure 1). All proteins significant at the protein level were present on the gene microarray. Most of the genes identified by expression difference have of course not yet been assessed as candidate protein markers.

Just two molecules (PCNA (31) and Survivin (18, 32)) were significantly associated with clinical outcome at both the gene and protein level. In contrast, 6 molecules were assessed as having no significant relationship with outcome at the protein level but appeared on the Winnepenninckx et al. (30) differently expressed gene list (CDK1 (32), Cyclin B (32), GRP78 (26), nm23 (33, 34), Survivin (18, 32) and TOPII (32)).

For 33 molecules (see Figure 1), the evidence sum has been classified as concordant. In all of these cases except one (PCNA (31)), the nature of that agreement was that no significant relationship with expression and clinical outcome was observed. For PCNA, increased gene and protein expression was associated with worse outcome. In three cases (BMI1 (20, 32), KI67 (21, 31, 32, 35, 36) and gp100 (31, 32, 37), the results of independent studies were discordant i.e. the results of IHC-based analysis have disagreed either with respect to the significance of an effect (BMI1 and KI67) or its significance and direction (gp100). In each of those cases, no outcome-related expression change was observed at the gene transcript level in the eligible gene expression dataset (30). For 28 molecules (see Figure 1), where there is only one IHC-based study or for which all protein data generally agree with each other, gene data do not converge upon that same finding in terms of the direction and/or significance of
the effect. In other cases (GRP78 (26), Cyclin B1 (32), CDK1 (32), nm23 (33, 34), Topoisomerase II (32)), the opposite was observed; gene transcript levels differ while no significant, outcome-related effect was observed at the protein level. For the remaining 16 molecules (BCL2 (32, 38), Survivin (18, 32), p16/INK4a (21, 32, 35), p16/INK4a nuclear compartment (18, 21), p21/WAF1 (21, 32), p27/KIP1 (21, 32), Cyclin A (32, 36, 39), p53 (18, 32, 35), AP2α (40, 41), ATF2 (21, 42), Fibronectin (21), tPA (43), α-catenin (16, 21), β-catenin (16, 21), Metallothionein (28, 44) and SPP1/osteopontin (29, 45, 46), concordance could not be assessed due to between-study differences making a direct comparison untenable. In addition, there were 4 molecules for which disparate findings were observed between different measures of clinical outcome. For GP100 (31, 32, 37) and Cyclin A (32, 36, 39), an effect was present for (disease-free survival) DFS but not ACM while for SPP1/osteopontin (45, 46) and tPA (43), the effect was observed for MSM or ACM but not DFS.

Discussion

To help address the ongoing search for useful molecular markers of prognosis in melanoma, we present the most formal and systematic examination of IHC and gene expression microarray-based studies in primary cutaneous melanoma and prognosis to date. For IHC-based literature, the review process employs the stringent REMARK-based method developed by Gould Rothberg and colleagues (1). Our update of their work expands their list of candidate biomarkers from 62 to 81 proteins and highlights promising candidates (e.g. Survivin, PCNA, MPDs) for further validation as biomarkers of prognosis, out of the more than 700 protein species that have been assessed in any way in the literature (47). We have also adapted the REMARK criteria for application to assessment of all published, peer-reviewed gene expression microarray-based research involving primary cutaneous melanoma and prognosis. The gene expression microarray data are not complicated by investigator selection bias, in that the gene list arises from a discovery analysis of associations with
outcome, using a near-full set of expressed gene markers. This has allowed us to systematically compare and synthesize candidate biomarker data derived from alternative selection approaches. We have observed that for 32 molecules the results were concordant between our eligible protein and gene expression studies. We have also attempted to explore system properties emerging from the combination of REMARK-based literature in human primary melanoma to try to better understand the drivers of poor outcome.

A comparison of features from IHC-based literature between the original review and the present work reveals some interesting changes. Firstly, review by application of the REMARK criteria continues to exclude more than 85% of published, peer-reviewed studies. At least one third of those works have again (1) been omitted due to cross-sectional study design showing that this method continues to dominate as a preferred research approach. Why the more rigorous, quantitative approaches are not more widely used, particularly in a multivariate testing context, is unclear. To bridge the gap between bench side and bed, studies should be fundamentally capable of testing for clear, reproducible and independent relationships between protein expression and clinical outcome. Among the notable examples of research with a translational focus is the work undertaken by Kreizenbeck and colleagues (16). Combining the effects of all six markers examined, these authors undertook hierarchical clustering of composite profiles to define discrete patient subpopulations that could be potentially associated with survival. The analysis yielded four unique clusters with differential overall survival, one of which remained significant in multivariate analysis. In another study (21), researchers examined protein expression for 38 molecules using AQUA and validated these findings in cohort of 246 primary melanomas. Both of these studies, from the same institute, used prior literature to select markers for reassessment in a multimarker and multivariate context and found significant relationships with clinical outcome. These studies demonstrate that the candidate-based approach (discussed in detail in (1)) for prognostic biomarkers can yield quality, clinically-relevant output.
Another positive trend across recent eligible studies is that the commonly used covariates now include mitotic rate. This change is consistent with findings of the final version of the 2009 AJCC staging system (4) in which mitotic rate was, for the first time, identified as one of the three most dominant prognostic factors in patients with localised melanoma. In spite of this shift, many studies continue to exclude dominant clinical covariates from multivariate analysis thereby reducing their translational relevance. Tests for marker independence should include each of the current, stage-specific prognostic factors along with, where possible, variables potentially relevant to the biological function of the molecule of interest.

A novel addition to the eligible literature in recent years is research assessing MPDs (18, 21, 22). When ranked by $P$-value together with single molecule predictors of outcome, these multimarker assays appear to perform better as predictors of outcome compared with individual biomarkers. This finding indicates that the most effective translational approach may well be one in which patient samples are assessed for the combinatorial effects of more than one marker. The contributions of proteome-based work, in its capacity to globally characterise protein expression alterations within tissue, is also expected to improve the rate of discovery for new biomarkers and/or the precision of prognosis (for a recent review see (48)). Gene expression microarray analyses offer the same intrinsic advantages: unbiased coverage and multimarker comparisons.

However application of the REMARK criteria to gene microarray-based studies identified only two eligible studies. This disappointingly low yield was a major limitation of this study in its attempt to unify and compare high quality datasets, but it genuinely reflects a paucity of sufficiently well designed, adequately powered studies, as others have observed (1, 49). A majority of gene microarray studies was excluded because the samples were not primary melanoma: banks of frozen primary tissue are rare in melanoma, and the typical primary tumor is very small compared with many other common cancers. For the much larger tumor
banks of FFPE tissue, improved assays have prompted a revisit and analysis of that tissue (29).

The bioinformatic analysis presented in this review focuses on the convergence or otherwise of evidence for molecular markers of prognosis within and between high-quality datasets. While cross-study comparisons of this nature are not a substitute for validation, they do have the potential to illuminate the contrasting outputs in terms of biomarker identification both at the level of individual molecules and gene ontologies. Most notably across all studies, only two candidates (PCNA and Survivin) had evidence at the gene and protein level for a relationship with clinical outcome. To date, neither of these proteins has been included in an MPD panel and therefore further investigation of their potential to contribute in that area is a high priority. The remaining 31 molecules for which studies agreed in their findings had insufficient evidence for an association with outcome and may not be worth pursuing further as candidates for prognostication. For the 30 molecules for which gene and protein data were not concordant, 24 of those molecules had altered protein but not gene expression. Whether these differences are biological in origin or perhaps an artefact of the cut point selection criteria used in the Winnepenninckx study (30) is not clear and awaits the results of further examination by both methods. For molecules whose relationship with outcome is characterised by unclear or discordant findings, several factors may be at play. Again, those differences may be the result of study design effects such as sample size, cut point selection, and the occasionally fickle nature of antibody binding specificity (as in (20)). Alternatively, phenotype switching (such as from proliferative to invasive states) may be a biological explanation for these differences (50).

The pathway analysis of IHC-based data reported a broad range of biological processes (DNA damage, apoptosis, signal transduction, cytoskeleton remodeling, development, immune response, regulation of lipid metabolism, neurophysiological process, cell adhesion and muscle contraction) so a pathway signal specific to the proteins actually associated with
outcome was hard to discern. In contrast, a cell cycle theme emerged from gene ontology enrichment analysis of the gene expression microarray data. The inclusion of a single gene expression microarray dataset in our analysis limits discussion around comparison of the gene ontology enrichment analyses for each of our eligible gene and protein datasets. It is also disappointing but perhaps not unexpected that the IHC work has not pointed with greater accuracy toward enriched pathways and processes, despite the evident success of the candidate-based approach in identifying individual molecules of interest (21). Nonetheless, we look forward to ongoing integration of these alternative approaches as a useful method of highlighting key molecules and pathways for further investigation in relation to melanoma prognosis, with a focus on translational research, as has been demonstrated here.
References


penetrating Nevi and metastatic vertical growth phase melanoma. Cancer Epidemiology Biomarkers 


Identifies Increased Osteopontin Expression as Predictive of Reduced Relapse-Free Survival. 


Figure Legends

Figure 1: Results of tests for association of protein expression (as evaluated using IHC) or gene expression (as assessed by gene expression microarray) with melanoma patient outcome were compared with each other to determine which findings were concordant across REMARK-compliant studies. As per the graphic key in the centre of the image, the sum of evidence for each molecule examined across the eligible literature was classified as ‘concordant’ (where eligible, independent studies of that molecule generally agreed with each other in terms of significance and direction of an effect), ‘discordant’ (indicating apparent lack of such agreement), or ‘unclear’ (where the studies could not easily be assessed as either agreeing or disagreeing). Where IHC-data generally agreed with each other but differed overall (in terms of the significance and/or direction of an effect) with gene expression data, molecules were classified as ‘gene and protein data not concordant’ (since protein expression may be regulated without a correlated gene transcript expression difference). Molecules are grouped by broad functional class as adapted from the Hanahan-Weinberg ‘Hallmarks of Cancer’ (6) by (1). The candidate biomarkers significantly ($P < 0.05$) associated with outcome in at least one eligible study appear in black font. Those molecules appearing in black font and in the ‘concordant’ quadrant (PCNA, Survivin) may be considered as high priority candidates for further validation, especially relative to those molecules appearing in gray (for which there is insufficient evidence that their expression levels are associated with outcome in melanoma).
Molecular Cancer Therapeutics

Melanoma Prognosis: A REMARK-based systematic review and bioinformatic analysis of immunohistochemical and gene microarray studies.

Sarah-Jane Schramm and Graham J Mann

Mol Cancer Ther Published OnlineFirst June 9, 2011.

Updated version
Access the most recent version of this article at:
doi:10.1158/1535-7163.MCT-10-0901

Supplementary Material
Access the most recent supplemental material at:
http://mct.aacrjournals.org/content/suppl/2011/06/09/1535-7163.MCT-10-0901.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.