

# Gene Expression Profiling of Primary Cutaneous Melanoma and Clinical Outcome

Véronique Winnepenninckx, Vladimir Lazar, Stefan Michiels, Philippe Dessen, Marguerite Stas, Soledad R. Alonso, Marie-Françoise Avril, Pablo L. Ortiz Romero, Thomas Robert, Ovidiu Balacescu, Alexander M. M. Eggermont, Gilbert Lenoir, Alain Sarasin, Thomas Tursz, Joost J. van den Oord, Alain Spatz

On behalf of the Melanoma Group of the European Organization for Research and Treatment of Cancer

**Background:** Gene expression profiling data for human primary cutaneous melanomas are scarce because of the lack of retrospective collections of frozen tumors. To identify differentially expressed genes that may be involved in melanoma progression and prognosis, we investigated the relationship between gene expression profiles and clinical outcome in a cohort of patients with primary melanoma. **Methods:** Labeled complementary RNA (cRNA) from each tissue sample was hybridized to a pangenomic 44K 60-mer oligonucleotide microarray. Class comparison and class prediction analyses were performed to identify genes whose expression in primary melanomas was associated with 4-year distant metastasis-free survival among 58 patients with at least 4 years of follow-up, distant metastasis, or death. Results were validated immunohistochemically at the protein level in 176 independent primary melanomas from patients with a median clinical follow-up of 8.5 years. Survival was analyzed with a Cox multivariable model and stratified log-rank test. All statistical tests were two-sided. **Results:** We identified 254 genes that were associated with distant metastasis-free survival of patients with primary melanoma. These 254 genes include genes involved in activating DNA replication origins, such as minichromosome maintenance genes and geminin. Twenty-three of these genes were studied at the protein level; expression of five (MCM4,  $P = .002$ ; MCM3,  $P = .030$ ; MCM6,  $P = .004$ ; KPNA2,  $P = .021$ ; and geminin,  $P = .004$ ) was statistically significantly associated with overall survival in the validation set. In a multivariable Cox model adjusted for tumor thickness, ulceration, age, and sex, expression of MCM4 (hazard ratio [HR] of death = 4.04, 95% confidence interval [CI] = 1.39 to 11.76;  $P = .010$ ) and MCM6 (HR of death = 7.42, 95% CI = 1.99 to 27.64;  $P = .003$ ) proteins was still statistically significantly associated with overall survival. **Conclusion:** We identified 254 genes whose expression was associated with metastatic dissemination of cutaneous melanomas. These genes may shed light on the molecular mechanisms underlying poor prognosis in melanoma patients. [J Natl Cancer Inst 2006;98:472–82]

Melanoma is the most life-threatening neoplasm of the skin, and its incidence and mortality have been increasing worldwide (1,2). The development of melanoma is a classical example of a neoplasm progressing through discrete stages that have well-known clinical and histologic features. However, the key underlying molecular events have not been clearly elu-

cidated, which may explain why no targeted therapy has been developed and why almost no clinical benefit from new therapies has been clearly demonstrated in patients with melanoma since the late 1970s.

Results of several expression studies designed to investigate the molecular mechanisms associated with melanoma progression have recently been reported (3–13); however, these studies used melanoma cell lines or metastatic tissue samples. No large gene expression profiling study of primary human melanomas from patients with long clinical follow-up has yet been reported because of limitations in the availability of primary human melanoma tissue. In fact, the entire primary tumor is customarily fixed and embedded in paraffin for histologic diagnosis, and so frozen tissue is usually available only from metastatic melanomas. Thus, gene expression profiling data for human primary cutaneous melanomas are scarce, and data with prognostic implication are entirely lacking.

Since the early 1980s, we have systematically frozen representative portions from both primary and metastatic melanomas, as well as from benign melanocytic lesions, and stored them at  $-80^{\circ}\text{C}$ . At the time of the original diagnosis on fixed, paraffin-embedded tissue, we also prepared high-quality sections stained with hematoxylin-eosin from each frozen sample to exclude sampling error or to refine the diagnosis and to assess histologic and/or immunohistochemical prognostic factors. This practice has not introduced difficulties for the diagnosis or microstaging of melanoma in our institutions.

In this study, we used samples of 83 primary melanomas from this collection for genome-wide gene expression profiling. We investigated the relationship between gene expression profiles and clinical outcome in a large cohort of patients with primary

*Affiliations of authors:* Departments of Morphology and Molecular Pathology (VW, JJvdO) and Surgery (MS), University Hospitals, Katholieke Universiteit Leuven, Leuven, Belgium; Divisions of Functional Genomics (VL, PD, TR, OB), Biostatistics and Epidemiology (SM), Dermatology (MFA), Genetics (GL), Cancer Medicine (TT), Pathology (A. Spatz), and Laboratory of Genetic Instability and Cancer (A. Sarasin), Gustave-Roussy Institute, Villejuif, France; Molecular Pathology Program, Centro Nacional de Investigaciones Oncológicas, Madrid, Spain (SRA); Servicio de Dermatología, Hospital 12 de Octubre, Madrid, Spain (PLOR); Surgery Department, Erasmus University Medical Center, Rotterdam, The Netherlands (AMME).

*Correspondence to:* Alan Spatz, MD, Department of Pathology, Gustave-Roussy Institute, 94805 Villejuif Cedex, France (e-mail: spatz@igr.fr).

See “Notes” following “References.”

DOI: 10.1093/jnci/djj103

© The Author 2006. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oxfordjournals.org.

melanoma of the skin, and we applied class comparison and class prediction analyses to identify genes whose expression was associated with 4-year distant metastasis-free survival. Also, we investigated differences in gene expression between paired primary and metastatic melanomas. These results were validated by use of immunohistochemistry in an independent cohort of patients with melanoma who had long-term follow-up.

## PATIENTS AND METHODS

### Patient Populations

Melanoma patients with available fresh frozen tissue were identified from the files of the pathology departments at University Hospitals in Leuven, Belgium (study series;  $n = 83$ ) and at the Gustave-Roussy Institute in Villejuif, France (validation series;  $n = 17$ ). The 17 specimens from the Gustave-Roussy Institute were all the frozen melanomas in the vertical growth phase that were available at this institution. For nine of the 83 patients from Leuven, paired frozen samples of the primary melanoma and its metastasis were available. These nine metastases included five cutaneous and four lymph node metastases. For all cases, representative part(s) of the tumor had been frozen in liquid nitrogen-cooled isopentane at the time of diagnosis and stored at a temperature below  $-80^{\circ}\text{C}$  until use. The rest of the tissue had been fixed in buffered formalin, embedded in paraffin, and used for routine histology. To exclude possible biases in pathologic diagnosis and microstaging caused by freezing only part of the lesion, a high-quality frozen section was made at time of diagnosis. All patients were treated uniformly—i.e., complete excision of the primary melanoma that was followed by reexcision with margins according to the thickness of the primary tumor—and none of the patients had undergone sentinel lymph node biopsy during the pathologic staging procedure. Tissue collections and the specific study protocol were approved by the medical ethical committees and institutional review boards of the University Hospitals at the Katholieke Universiteit Leuven and of the Gustave-Roussy Institute. Each melanoma of our patient cohort was reviewed by three pathologists (V. Winnepenninckx, J. J. van den Oord, and A. Spatz). Clinical data available included age, sex, site of involvement, American Joint Committee on Cancer (AJCC) tumor evaluation (T1A–4B), and AJCC stage at diagnosis (I–IV) (14). The clinical and histopathologic characteristics of the study and validation series are summarized in Table 1.

### RNA and DNA Extraction and Purification

From all frozen samples, we cut a 5- $\mu\text{m}$  section and used it to assess two previously defined inclusion criteria, i.e., that at least 70% of the cells were tumor cells and that the melanoma was at least 1 mm thick or was in the vertical growth phase (15). Frozen blocks that contained less than 70% melanoma cells (i.e., contained too many adjacent epithelial and stromal components) were trimmed under sterile conditions to remove excess nontumorous components. After trimming, these tumors and tumors with more than 70% melanoma cells were treated in the same way. From each block, we cut 20–40 cryostat sections (each 10  $\mu\text{m}$  thick) under sterile conditions. Sections were collected in Eppendorf tubes containing 350  $\mu\text{L}$  of RLT lysis buffer (Westburg, Leusden, The Netherlands) to which 1% 2-mercaptoethanol was

**Table 1.** Pertinent clinical and histologic characteristics of the microarray study and validation series\*

Characteristic	Study population ( $n = 83$ )	Validation population ( $n = 17$ )
Median age, y (range)	63.5 (1–93)	61 (40–88)
Median Breslow thickness, mm (range)	3.3 (0.4–18.6)	4.9 (1.2–20)
Sex, No.		
Male	27	9
Female	56	8
Site, No.		
Head and neck	10	1
Trunk	24	3
Limbs	44	7
Extremities	5	6
Melanoma type, No.		
Superficial spreading melanoma	49	6
Nodular melanoma	15	7
Acrolentiginous melanoma	7	0
Lentigo maligna melanoma	1	4
Other†	11	0
AJCC stage at diagnosis		
I	27	3
II	38	8
III	15	3
IV	3	3
Ulceration, No.		
Absent	52	0
Present	31	14
NA	0	3
Regression, No.		
Absent	67	12
Present	15	5
NA	1	0
Mitotic rate, No. per $\text{mm}^2$		
0	6	0
1–6	39	10
>6	38	7
Cell type, No.		
Epithelioid	56	10
Epithelioid and spindle	6	5
Spindle	8	0
Other	13	2
TILs‡, No.		
Absent	32	8
Nonbrisk	29	4
Brisk	22	5
Neovascularization, No.		
Absent	72	11
Present	11	6
Elastolysis, No.		
Absent	39	3
Slight or moderate	29	14
Severe (nodular)	9	0
NA	6	0
Median follow-up,§ mo	59	53

\*AJCC = American Joint Committee on Cancer; NA = not available.

†Other = desmoplastic melanoma and small-cell melanoma.

‡TILs = host response by tumor-infiltrating lymphocytes; nonbrisk = focal or discontinuous; brisk = continuous.

§Median follow-up for distant metastasis-free survival was calculated by use of the inverted Kaplan–Meier method.

added immediately before use and vortex mixed for 5 minutes. Eppendorf tubes were then held at room temperature for another hour and frozen. Finally, another section was cut from the rest of the frozen block and stained with hematoxylin–eosin, and the inclusion criteria were evaluated again on this section. Eventually, among 178 frozen primary melanomas in the tissue repository, 32 were rejected at this stage for diagnostic reasons, 22 did not

yield enough RNA, and 41 had degraded RNA. The RNA retrieval ratio for this study set was, therefore, 0.57.

Total RNA was isolated by use of the RNeasy kit from Qiagen after DNase treatment and purified by following the manufacturer's protocol. DNA was recovered from the first column eluate by ethanol precipitation (RNA binds to the RNeasy column, but DNA and proteins do not). DNA was dissolved in TE buffer (10 mM Tris HCl at pH 8.0 and 1 mM EDTA at pH 8.0). DNA was further purified on QIAamp DNA kit from Qiagen. RNA concentrations were adjusted to 100 ng/ $\mu$ L, as measured with a Nanodrop spectrophotometer (NanoDrop Technologies; Wilmington, DE). A pool of RNA, obtained by mixing equal amounts (500 ng) of total RNA from each of the 83 primary tumors in the study set, was used to prepare the reference complementary RNA (cRNA). The quality of RNA preparations was assessed with Lab-on-a-chip Bioanalyzer 2000 technology (Agilent Technologies, Palo Alto, CA), the 28S/18S ribosomal RNA ratio was used as the control. All samples included in this study had a 28S/18S ribosomal RNA ratio of more than 1.5, with an average of 1.8.

### Oligonucleotide Microarray Technology

Agilent oligonucleotide technology uses dual-color analysis, in which probes from tumor tissue and from the reference tissue are differentially labeled by the incorporation of cyanine 3 (Cy3) and cyanine 5 (Cy5) (Agilent Technologies), respectively. Probes were synthesized from 500 ng of total RNA in two steps according to the manufacturer's instructions. In the first step, double-stranded cDNA was synthesized with mouse Moloney murine leukemia virus (Agilent Technologies) reverse transcriptase and an oligo(dT)-T7 RNA polymerase promoter (Agilent Technologies). One microliter of double-stranded cDNA reaction product was diluted in 200  $\mu$ L of nuclease-free water and stored at  $-20^{\circ}\text{C}$  for further real time quantitative polymerase chain reaction (RT-PCR), as a control. In the second step, we synthesized antisense cRNAs that were labeled by the incorporation of CTP-Cy3 or CTP-Cy5 (Perkin Elmer) during in vitro transcription. All reagents were from the Agilent's Fluorescent Linear Amplification kit adapted for use with small amounts of total RNA. Labeled cRNAs were fragmented to an average size of 50–100 nucleotides by heating the samples at  $60^{\circ}\text{C}$  in a fragmentation buffer provided by Agilent. One microgram of purified cRNA from each sample was mixed with the same amount of reference cRNA. Hybridization was performed on whole-human-genome 44K oligonucleotide microarrays (product G4112A; Agilent) with reagents and protocols provided by the manufacturer.

Probe synthesis and hybridization, with a reference design (15), were conducted as single experiments for the analysis of the 83 Cy5-labeled cRNAs from the primary melanomas in the study set with the Cy3-labeled cRNA pool and for the analysis of 17 Cy5-labeled cRNAs from primary tumors in the validation set. A direct comparison was carried out to study differential gene expression between the Cy3-labeled cRNAs from the primary melanoma and the Cy5-labeled cRNA from its metastasis. We also performed dye-swap experiments comparing the Cy5-labeled cRNAs from the primary melanoma with the Cy3-labeled cRNA from its metastasis to increase the precision of these direct comparisons (16).

Feature extraction software provided by Agilent (version 7.2) was used to quantify the intensity of fluorescent images and to

normalize results by subtracting local background fluorescence, according to the manufacturer's instructions. All of the data were imported into Resolver software (Rosetta Biosoftware, Kirkland, WA) for database management, quality control, and analysis.

### RT-PCR

Primer and probes used in this study were as follows (gene symbol/RefSeq identity/assay identity): KLK7/NM\_005046/Hs00192503\_m1; PCNA/NM\_002592/Hs00427214\_g1; KPNA2/NM\_002266/Hs00818252\_g1; MCM6/NM\_005915/Hs00195504\_m1; MCM3/NM\_002388/Hs00172459\_m1; MCM4/NM\_005914/Hs00172459\_m1; PGDS/NM\_014485/Hs00183950\_m1; RFC4/NM\_002916/Hs00427469\_m1; NME1/NM\_000269/Hs00264824\_m1, and 18S ribosomal RNA. Reagents were purchased from Applied Biosystems (Foster City, CA).

Sequences were amplified by PCR with the Taqman Universal Master Mix (Applied Biosystems) and primer and probe reagents (Applied Biosystems), according to the manufacturer's instructions (Applied Biosystems). Reactions were carried out in an ABI 7000 sequence detection system, whose software analyses fluorescent signals and calculates the cycle threshold. Careful monitoring of cycle threshold (Ct)-negative controls for each target showed the absence of carryovers. The expression profile for each specimen was assessed by using the comparative threshold cycle ( $2^{\text{ddCt}}$ ) method, according to the manufacturer's instructions (Applied Biosystems).

### Immunohistochemistry

Tissue microarrays (product MTA1; Beecher Instruments, Sun Prairie, WI) were used for immunohistochemical studies of tumor tissue from 62 patients with adequate material from the original study set and from an additional validation set of 176 patients with primary cutaneous melanoma—i.e., 63 patients diagnosed at Gustave-Roussy Institute, Villejuif, France, and 113 patients diagnosed at the 12 Octubre University Hospital, Madrid, Spain—with a median follow-up of 8.5 years.

After inactivation of endogenous peroxidase activity in methanol and  $\text{H}_2\text{O}_2$  and heat-induced epitope retreatment in 10 mM Tris HCl at pH 9.0 and 1 mM EDTA, sections were incubated with antibodies against CDC2, CDC6, CENPA, ephrin B6, ECT2, factor X, geminin, HAUSP, interleukin 6, KLK5, KLK7, KLK11, karyopherin- $\alpha$ 2, MCM3, MCM4, MCM6, NEK2, NME1, PCNA, prostaglandin D synthase, RFC4, survivin, or WNT11 proteins. All antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz, CA) except those against CDC2 and CDC6 (Oncogene Science, Cambridge, MA), interleukin 6 (Genzyme, Cambridge, MA), NME1 (Dakopatts, Copenhagen, Denmark), and PCNA (Novocastra, Newcastle-upon-Tyne, United Kingdom). A peroxidase-labeled polymer (Envision; Dakopatts) was added to the reaction mixture, and enzyme activity was detected by use of 3-amino-9-ethylcarbazole as substrate, which produced a brightly red reaction product that contrasted well with the brown melanin in the sections. Sections were independently scored by two pathologists (V. Winnepenninckx and J. J. van den Oord) on a four-level scale (grade 1 = 0; grade 2 = 1%–10%; grade 3 = 11%–50%; and grade 4 = >50%) for the percentage of cells with a stained nucleus (for nuclear immunoreactivity) and/or cytoplasm (for cytoplasmic immunoreactivity). For each lesion, the average cytoplasmic or nuclear expression value obtained with

the immunohistochemical tissue microarrays was used to summarize the expression values of different cores.

### Data Availability

The microarray data analyzed in this paper have been submitted to the Array Express data repository at the European Bioinformatics Institute (<http://www.ebi.ac.uk/arrayexpress/>) under the following accession numbers: E-TABM-1 IGR\_MELANOMA\_STUDY (corresponding to the 83-patient study set), E-TABM-2 IGR\_MELANOMA\_VALID (17-patient validation set), and E-TABM-4 IGR\_MELANOMA\_META (nine-patient set with paired tumors and metastases). The gene expression table used to create Fig. 1 (83 tumors times 11 043 genes) is available under the first accession number.

### Statistical Analysis

Statistical analyses were performed with the SAS System, version 8.2, and BRB-ArrayTools software, version 3.3 (17). Unsupervised hierarchical clustering analysis was performed with the Rosetta Resolver system, version 4, by use of mean centered correlation and average linkage for both sequences and arrays. Statistically significantly regulated sequences were defined as those sequences whose expression in tumors was statistically significantly different from the expression in the pool, as calculated by the Agilent microarray-specific error model (18), in at least one of four experiments (21 of 83 tumors) and for which the intensity of the two channels was higher than 100 in half of the experiments (i.e., at least 42 of 83 experiments). For this analysis, a *P* value of less than .01 was considered statistically significant. After normalization and correction of fluorescence intensities from the scanned images with Agilent Feature Extraction software, 11 043 statistically significantly regulated sequences were retained for analysis. For all analyses on the primary tumors, the 11 043 statistically significantly regulated sequences were used.

The primary endpoint for the study was distant metastasis-free survival, which was defined as the time interval between the diagnosis of the primary cutaneous melanoma and a distant metastasis or death from melanoma. A distant metastasis was defined as a documented melanoma with a location beyond the first regional lymph node basin. Patients alive without distant metastasis or with a death not related to cancer were censored at the date of last follow-up or at the date of death, respectively. Overall survival was defined as the time from diagnosis until death (from any cause). Patients alive at the date of last follow-up were censored at that date.

To identify sequences associated with reduced rates of death from distant metastases, we analyzed data for the 58 of the 83 patients with at least 4 years of follow-up, intercurrent distant metastasis, or death. Eight patients with a reported death related to another cause within 4 years and 17 patients without a reported event but with a follow-up of less than 4 years were excluded from this analysis. Patients were separated into two groups, one group with distant metastasis-free survival of more than 4 years (group M<sup>-</sup>) and one group with distant metastasis-free survival of 4 years or less (group M<sup>+</sup>). We used an exploratory class comparison analysis to determine which sequences were differentially expressed between the two groups. For this analysis, we used a two-sample *t* test whose *P* values were adjusted by the

false discovery rate step-up method of Benjamini and Hochberg (19) to account for multiple testing.

For supervised class prediction analysis of the 58 patients in the study set with at least 4 years of follow-up and the 17 patients in the validation set, we chose the nearest prediction rule, which has been intensively studied on data from several microarray studies [Michiels et al. (20) and references therein]. This prediction rule classifies new patients according to the correlation between the expression of their signature sequences and the average profiles in the two prognostic groups (groups M<sup>+</sup> and M<sup>-</sup>), the predicted category being the one with the highest correlation. The default gene selection procedure from BRB ArrayTools was used, in which the sequences that were differentially expressed at a statistical significance level of *P* = .001 were included in the prediction rule. Leave-one-out cross-validation was used to estimate the prediction accuracy. One sample is left out, and the remaining samples are used to build the prediction rule, which is then used to classify the left-out sample. The entire model-building process was repeated for each leave-one-out training set to provide unbiased estimates of the prediction accuracy (21,22). The statistical significance of the cross-validated misclassification rate was determined by repeating the entire cross-validation procedure on the 58 patients in group M<sup>+</sup> and group M<sup>-</sup> randomly permuted 2000 times (23). The *P* value was obtained from the proportion of times this re-labeling resulted in a lower misclassification rate. All *P* values obtained were adjusted by the false discovery rate step-up method of Benjamini and Hochberg and are referred to as adjusted *P* values.

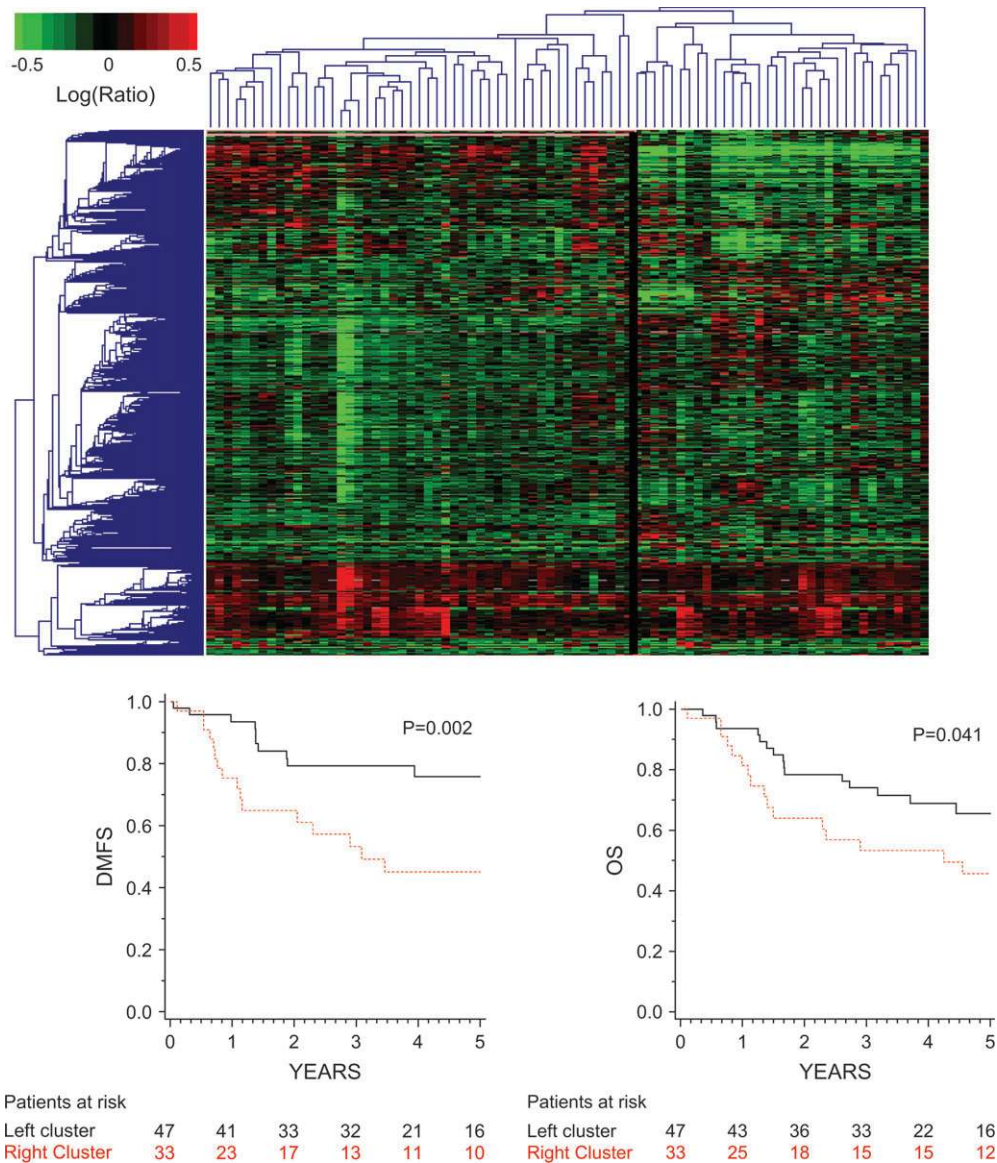
Pearson's correlation values were used to analyze correlations between gene expression and thickness. We used Fisher's exact tests, Wilcoxon tests, and chi-squared tests to assess differences between the clinicopathologic variables of patients whose disease recurred and patients whose disease did not recur.

For the immunohistochemical validation study, the survival distributions between groups were compared by stratified log-rank tests and stratified Cox models. The multivariable analysis was adjusted for all the clinical and histologic prognostic factors that were selected by a backward selection procedure in a Cox regression model (with a selection probability to stay in the model of .05). The proportional hazards assumption was verified by tests for interactions between time and the covariates; the data confirmed the proportionality assumption. All statistical tests were two-sided.

## RESULTS

### Two Major Melanoma Subsets From Unsupervised Hierarchical Clustering

Using two-channel hybridization to genome-wide oligonucleotide arrays carrying 41 675 probes, we generated gene expression profiles for 83 primary melanomas (study set from University Hospitals in Leuven, Belgium) including 27 stage I, 38 stage II, 15 stage III, and three stage IV tumors, with stages defined according to the AJCC staging system. After normalization and correction of fluorescence intensities from the scanned images with Agilent Feature Extraction software, 11 043 statistically significantly regulated sequences were retained for analysis. Unsupervised hierarchical clustering of the study set, which was based on the similarities in expression for these 11 043



**Fig. 1.** Unsupervised hierarchical clustering of 83 primary melanomas. The unsupervised hierarchical clustering was based on similarities in expression measured over the filtered 11 043 genes. **A)** Two-dimensional representation of transcript logarithmic ratios for 83 primary melanomas. Data were obtained by hierarchical clustering that used the average linkage for both genes and for tumors. For clustering, the distance metric used was 1 minus mean centered correlation. Each row represents one gene and each column represents a tumor. The **color bar** at the top indicates the level of mRNA expression. **Red** = higher level of expression of mRNA in the tumor than the reference level of mRNA (pool of all

primary tumors); **green** = lower level of expression than the reference; **black** = no change from the reference. The **black bar** divides the left cluster of data from the right cluster. **B)** Kaplan–Meier analysis of distant metastasis–free survival (DMFS) and overall survival (OS) among 80 patients according to whether they were found in the left or right cluster defined as the two main clusters cutting the dendrogram at the first split. For 4-year DMFS, left cluster = 76% (95% confidence interval [CI] = 64% to 90%) and right cluster = 46% (95% CI = 30% to 68%). For 4-year OS, left cluster = 69% (95% CI = 57% to 84%) and right cluster 53% (95% CI = 38% to 75%). All statistical tests were two-sided.

sequences, revealed two distinct main clusters—a larger cluster of 49 tumors (left cluster) and a smaller cluster of 33 tumors (right cluster) (Fig. 1). One patient could not be classified into either cluster and therefore was excluded from cluster comparison. A sensitivity analysis that applied three more stringent filtering criteria (and hence identified fewer sequences) identified very similar clusters. When we examined the relationships between the two clusters and the patient and tumor characteristics, we found statistically significant differences in the median thickness of the lesions (1.7 mm in the left cluster versus 5.5 mm in the right cluster;  $P < .001$ ); the right cluster contained the three stage IV and eight of the 15 stage III patients, as well as two-thirds of the patients with ulcerated melanomas. Four-year dis-

tant metastasis–free survival rates were 0.76 (95% confidence interval [CI] = 0.64 to 0.90) for patients in the left cluster and 0.46 (95% CI = 0.30 to 0.68) for patients in the right cluster; overall survival rates were 0.69 (95% CI = 0.57 to 0.84) for patients in the left cluster and 0.53 (95% CI = 0.38 to 0.75) for patients in the right cluster. Comparison analyses showed that distant metastasis–free survival and overall survival were statistically significantly different between the 49 patients included in the left cluster and the 33 patients in the right cluster (for distant metastasis–free survival,  $P = .002$ ; and for overall survival,  $P = .041$ ; log-rank tests). Thus, unsupervised clustering permitted melanomas with increased survival to be distinguished from those with decreased survival.

## Class Comparison: Distant Metastasis-free Survival at 4 Years

To identify gene sequences associated with 4-year distant metastasis-free survival, we performed an exploratory class comparison analysis of gene expression data from 58 patients with at least 4 years of follow-up. These patients were separated into 32 patients with no reported distant metastasis or death within 4 years (group M<sup>-</sup>) and 26 patients with a reported distant metastasis within 4 years (group M<sup>+</sup>). A comparison of the clinical and histologic characteristics of these two groups is presented in Table 2. For each gene examined, the mean fluorescence intensities (logarithm of ratios), as a measure of gene expression, between the two groups were compared by use of a *t* test, in which an adjusted *P* value of less than .05 was statistically significant. A total of 361 sequences were statistically significantly differentially expressed between the two groups (Supplementary Table 1;

available at <http://jncicancerspectrum.oxfordjournals.org/jnci/content/vol98/issue7>).

## Class Prediction: Distant Metastasis-free Survival at 4 Years

We performed a class prediction analysis on the 11 043 filtered sequences by use of the nearest centroid prediction rule. Sequences that were differentially expressed, at a statistical significance level of *P* = .001, were included in the gene classifier. The prediction accuracy of this gene classifier was estimated by a full leave-one-out cross-validation analysis (i.e., repeating the gene selection procedure at each leave-one-out step) on the 58-patient dataset. From this analysis, we obtained an estimated misclassification rate of 29%. Results of a permutation test (i.e., 2000 permutations) indicated that the probability of randomly achieving this misclassification rate by the prediction rule was equal to .04.

We found that the expression of the top 254 sequences (Supplementary Table 1; available at <http://jncicancerspectrum.oxfordjournals.org/jnci/content/vol98/issue7>), which we termed the 254-gene classifier, differed between groups M<sup>+</sup> and M<sup>-</sup> at the statistical significance level of .001. The expression pattern of these 254 sequences in the 58 samples is shown in Fig. 2. Of these 254 sequences, 174 correspond to known genes. Seventy-five of these 174 genes have been previously reported to be involved in the biology of cancer, but only 30 of the 75 genes have been studied in melanoma. We then used the expression of these 254 genes to classify 17 patients from an independent validation set, of whom nine developed distant metastasis within 4 years (group M<sup>+</sup>) and eight remained free from distant metastasis (group M<sup>-</sup>). The nearest centroid prediction rule using the expression of these 254 sequences classified 11 of these 17 patients correctly (65%, 95% CI = 38% to 87%; Fig. 2).

We compared the prediction accuracy of our 254-gene classifier, as estimated by the leave-one-out cross-validation procedure, with the prediction accuracy of the two most statistically significant histologic prognostic parameters in primary melanomas—i.e., tumor thickness and ulceration (14). For this purpose, we constructed a clinical prediction rule that was based on the AJCC melanoma staging system. Patients with primary melanoma that was more than 4 mm thick or patients with an ulcerated melanoma that was more than 2 mm thick were classified as having poor prognosis; all other patients were classified as having favorable prognosis. Using this clinical prediction rule, we obtained a misclassification rate of 28% for our set of 58 patients, which is very close to the misclassification rate of 29% obtained with the 254-gene classifier.

## Correlation Between Gene Expression and Tumor Thickness

When we used a correlation test to determine the relationship between gene expression and melanoma thicknesses for the 83 melanomas in the study set, we found that 652 sequences were highly correlated with thickness (at an adjusted statistical significance level of *P* < .001; range of absolute correlation coefficients = .50–.71; Supplementary Table 2; available at <http://jncicancerspectrum.oxfordjournals.org/jnci/content/vol98/issue7>). Of these 652 sequences, 145 were included in our 254-gene classifier. From the 652 sequences, 477 sequences were negatively

**Table 2.** Comparison of patient characteristics between the favorable (M<sup>-</sup>) and unfavorable (M<sup>+</sup>) prognosis groups used in the supervised analysis\*

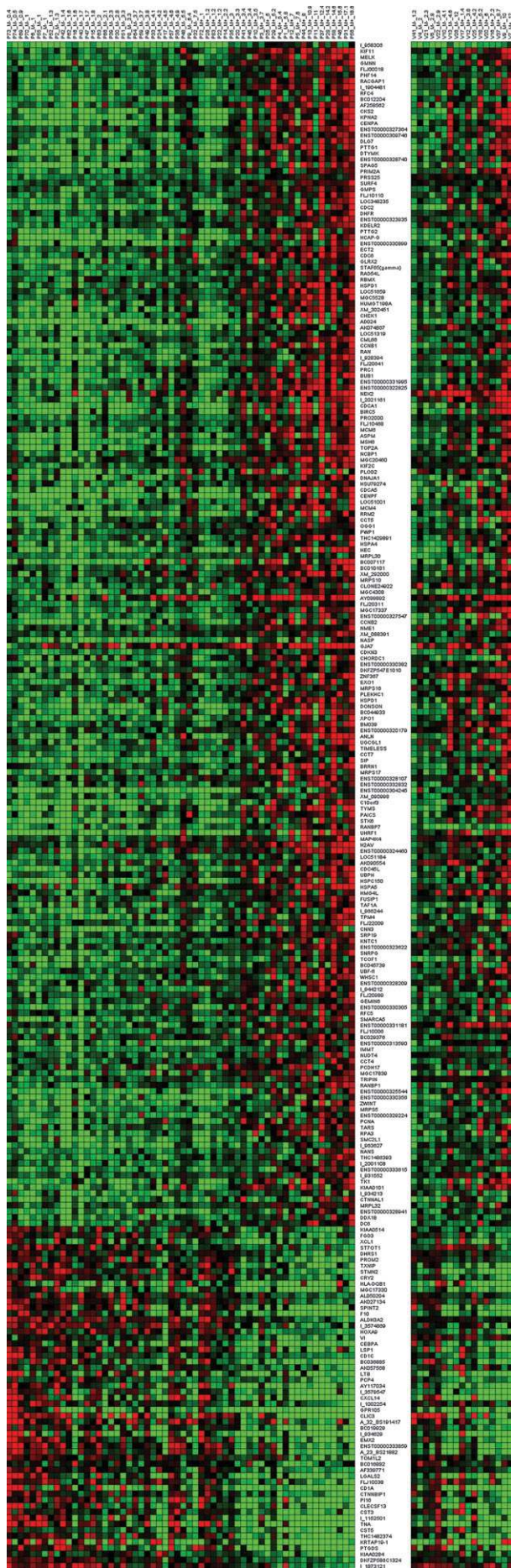
Variable	M <sup>-</sup> group (n = 32)	M <sup>+</sup> group (n = 26)	<i>P</i>
Median age, y (range)†	57 (12–81)	71 (1–89)	.07
Median thickness (Breslow), mm (range)†	2.0 (0.4–7.5)	5.3 (1.0–18.6)	.001
AJCC stage at diagnosis, No.‡			<.001
I	15	2	
II	15	11	
III or IV	2	13	
Sex, No.§			.25
Male	7	10	
Female	25	16	
Site, No.§			.59
Axial	11	11	
Extremities	21	15	
Melanoma type, No.§			.02
SMM	21	11	
Others	6	13	
Ulceration§			.003
Absent	26	11	
Present	6	15	
Regression§			.49
Absent	25	23	
Present	6	3	
Mitotic rate, No.§			.02
<6 mitoses per mm <sup>2</sup>	22	9	
≥6 mitoses per mm <sup>2</sup>	10	17	
Predominant cell type§			1.00
Spindle	9	8	
Others	23	18	
TILs‡			.75
Absent	12	10	
Nonbrisk	10	10	
Brisk	10	6	
Neovascularization or angiotropism, No.§			.27
Absent	29	20	
Present	3	6	
Elastolysis§			.56
Absent to moderate	17	15	
Severe	13	9	

\*AJCC = American Joint Committee on Cancer; SSM = superficial spreading melanoma; TILs = host response by tumor-infiltrating lymphocytes; nonbrisk = focal or discontinuous; brisk = continuous.

†Wilcoxon test, two-sided.

‡chi-squared test, two-sided.

§Fisher's exact test, two-sided.



correlated with thickness (i.e., their expression decreased as the thickness increased) and 175 sequences were positively correlated (i.e., their expression increased as the thickness increased). The expression of 23 of 24 genes involved in DNA repair increased as the thickness of the primary melanoma increased, as did the expression of genes associated with the cell cycle (all eight genes examined), protein folding (all 10 genes examined), chromatin remodeling (all 10 genes examined), and heat shock protein activity (all 11 genes examined). In contrast, expression decreased as the thickness of the primary melanoma increased for genes involved in serine-type endopeptidase inhibitor activity (all 15 genes examined), cell adhesion (all 15 genes examined), cell–cell signaling (all eight genes examined), and transcription factor activity (33 of the 36 genes examined).

### Relationship Between Gene Expression in Primary Melanomas and Its Metastases

Previous studies (24,25) on breast and liver cancers have found only small differences in gene expression between primary cancers and their subsequent metastases. When analyzing the nine pairs of primary melanoma and their metastases from the same patients without preprocessing of the sequences, we found no statistically significant differences in gene expression between primary melanoma tissue and its metastatic tumor tissue (at an adjusted statistical significance level of  $P < .05$  with a paired *t* test).

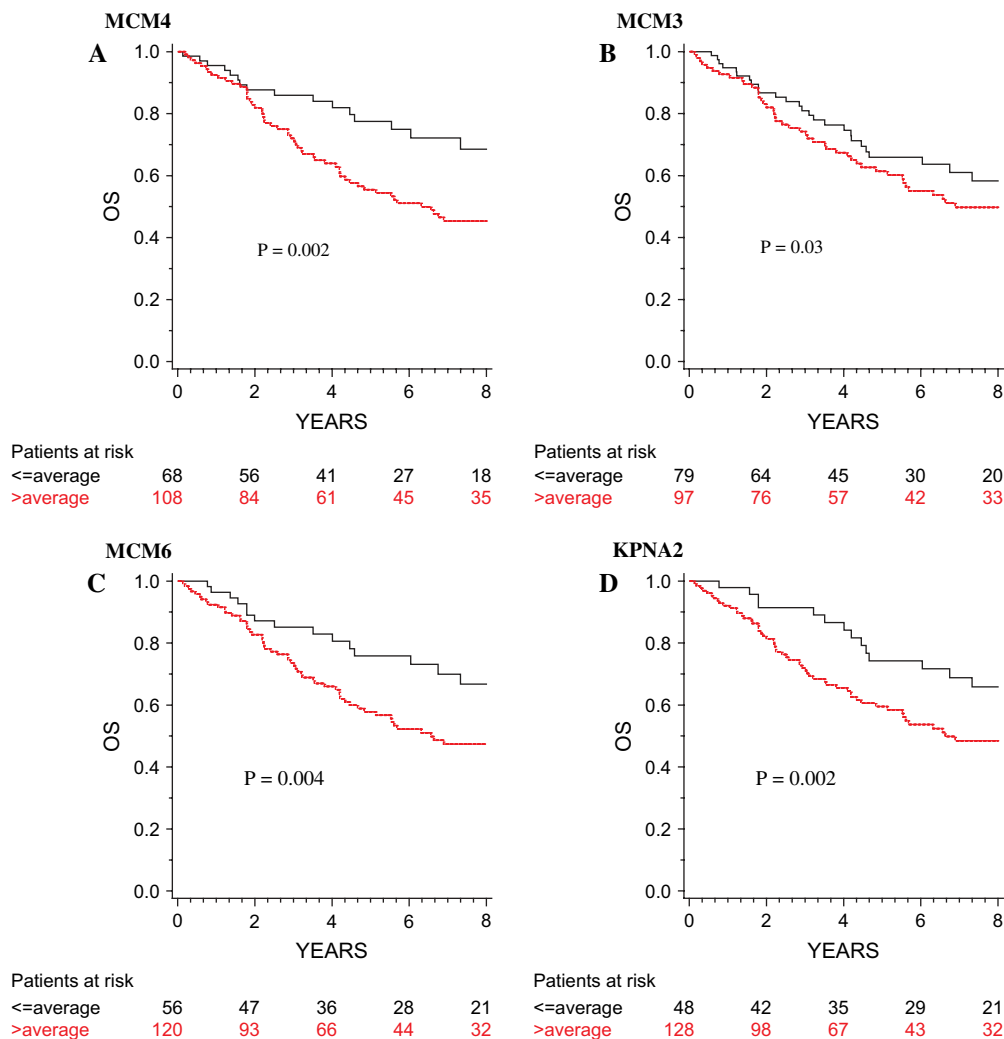
### Confirmation of Gene Expression Results by Immunohistochemistry on Tissue Microarrays

The expression of 23 proteins from the 254-gene classifier that had commercially available antibodies was studied by immunoreactivity. These proteins corresponded to genes selected from the list of genes predictive of distant metastasis-free survival at 4 years, including GMNN, KPNA2, RFC4, CENPA, PTGDS, CDC6, CDC2, F10, EPHB6, NME1, ECT2, NEK2, PCNA, BIRC5, MCM3, MCM4, MCM6, and HAUSP, or associated with ulceration (IL6) or with tumor thickness (WNT11, KLK4, KLK7, and KLK11). For this part, the study material consisted of paraffin-embedded tissue from a total of 238 melanomas, including 62 melanomas from the study set and 176 primary melanomas from an independent set, from patients with a median follow-up of 8.5 years.

In the study set of 62 melanomas, immunoreactivity in the melanoma cores with eight antibodies (i.e., antibodies directed against MCM4, MCM3, MCM6, KPNA2, geminin, NME1, KLK7, and factor X) of the 23 antibodies was statistically significantly associated with distant metastasis-free survival and/or overall survival

**Fig. 2.** Pattern of expression of the 254-gene classifier for 4-year distant metastasis-free survival. **Left)** Pattern of expression of the 254 genes in the 58 patients from the study set. **Right)** Validation set of 17 patients. Each column represents the prognostic profile of the 254 genes for one tumor, and each row represents the relative level of expression of one gene. Genes are rank ordered according to their correlation with the prognosis in the study set. Tumors are ordered according to their true prognosis group (group M<sup>-</sup> to the left; group M<sup>+</sup> to the right) and within each group by increasing thickness values. The **color bar** at the top indicates the level of mRNA expression. **Red** = higher level of expression of mRNA in the tumor than the reference level of mRNA (pool of all primary tumors); **green** = lower level of expression than the reference; **black** = no change from the reference.

**Fig. 3.** Kaplan–Meier analysis of overall survival (OS) among the validation set of 176 patients according to the immunoreactivity of their melanomas in comparison to the average value from the initial study set of patients. **Black** = less than or equal to this average value; **red** = more than this average value. The genes evaluated were MCM4, MCM3, MCM6, and KPNA2. **A)** MCM4. The 4-year OS for patients with a lower or equal expression, compared with the average expression from the study set, was 83% (95% confidence interval [CI] = 73% to 92%); the 4-year OS for patients with higher expression was 66% (95% CI = 56% to 75%). **B)** MCM3. The 4-year OS for patients with a lower or equal expression, compared with the average expression from the study set, was 76% (95% CI = 66% to 86%); the 4-year OS for patients with higher expression was 67% (95% CI = 58% to 77%). **C)** MCM6. The 4-year OS for patients with a lower or equal expression from the study set, was 81% (95% CI = 71% to 92%); the 4-year OS for patients with higher expression was 67% (95% CI = 58% to 75%). **D)** KPNA2. The 4-year OS for patients with a lower or equal expression, compared with the average expression from the study set, was 85% (95% CI = 74% to 95%); the 4-year OS for patients with higher expression was 66% (95% CI = 57% to 75%). All *P* values were from a two-sided stratified log-rank test.



(at a statistical significance level of  $P < .05$ , in accordance with the prognostic significance of the genes selected by the microarray analysis). Furthermore, immunohistochemistry data confirmed the microarray results with respect to the correlation between the expression of MCM4, MCM3, PCNA, and NME1 proteins and the thickness of the melanoma (Pearson's correlation for MCM4,  $r = .35$  and  $P = .007$ ; MCM3,  $r = .47$  and  $P < .001$ ; PCNA,  $r = .33$  and  $P = .01$ ; NME1,  $r = -.39$  and  $P = .002$ ).

The expression of gene products with a statistically significant association with distant metastasis-free survival or with overall survival was further studied by immunohistochemistry in an independent set of 176 primary melanomas. The expression of each gene was separated into two groups by use of the observed average expression value of that gene in the 62-patient study set. From a univariate analysis, a statistically significant association was observed between immunoreactivity and overall survival (at a statistical significance level of  $P < .05$ ) for MCM4 ( $P = .002$ ; Figs. 3, A, and 4, A), MCM3 ( $P = .030$ ; Figs. 3, B, and 4, B), MCM6 ( $P = .004$ ; Figs. 3, C, and 4, C), KPNA2 ( $P = .021$ ; Figs. 3, D, and 4, D), and geminin ( $P = .004$ ; data not shown). Staining was both nuclear and cytoplasmic for MCM4; nuclear for MCM3, MCM6, and geminin; and cytoplasmic for KPNA2. When we applied a backward selection procedure to select the most important clinical and histologic prognostic variables in a multivariable Cox regression model from the variables listed in

Table 1, only four variables were retained—i.e., tumor thickness, ulceration, age, and sex. Immunohistochemically detected expression of MCM4 (hazard ratio [HR] of death = 4.04, 95% CI = 1.39 to 11.76;  $P = .010$ ) and MCM6 (HR of death = 7.42, 95% CI = 1.99 to 27.64;  $P = .003$ ) proteins was still statistically significantly associated with overall survival when adjusted for those four variables in a multivariable Cox model.

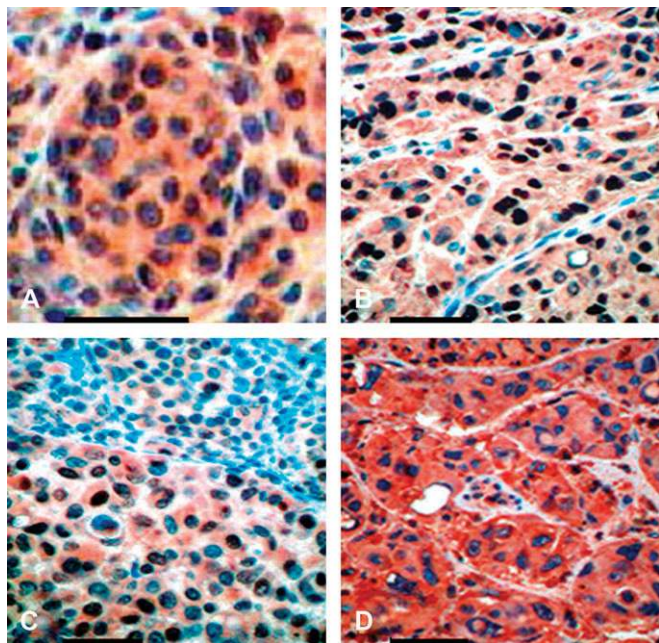
### Confirmation of Expression Measurements by RT-PCR

Nine genes of the 23 genes subjected to protein analysis (i.e., MCM4, MCM3, MCM6, KPNA2, KLK7, PCNA, PTGDS, RFC4, and NME1) were selected for analysis by quantitative RT-PCR, and the correlation between gene expression as measured by microarrays and gene expression as measured by RT-PCR in the study series of 83 primary melanomas was assessed. As shown in Table 3, the two methods exhibited high correlation; all nine correlation coefficients were statistically significant at a level of  $P$  less than .001.

### DISCUSSION

This study is, to our knowledge, the first that uses a large retrospective series of frozen samples with long-term follow-up to analyze the genes underlying tumor progression in melanoma.





**Fig. 4.** Expression of gene products in primary melanomas. Immunohistochemical brown staining of primary melanoma using monoclonal antibodies against the following proteins: **A)** MCM4. **B)** MCM3. **C)** MCM6. **D)** KPNA2. Grade 4 (immunostained cells >50% of total) cytoplasmic (**A** and **D**) and nuclear (**B** and **C**) immunoreactivity in the melanoma cells is shown. **Scale bar** = 50  $\mu$ m.

Using a pangenomic oligonucleotide microarray, we found that a 254-gene expression signature could identify patients at risk to develop distant metastasis. We also found several immunohistochemical markers that appear to have prognostic significance. The gene predictor consisted of 254 sequences, of which 174 correspond to known genes (Supplementary Table 3; available at <http://jncicancerspectrum.oxfordjournals.org/jnci/content/vol98/issue7>). From these 174 genes, 141 were underexpressed and 33 were overexpressed in melanomas from patients who remained free of metastasis for 4 years. Seventy-five of these 174 genes have previously been reported to be involved in the biology of cancer, but only 30 of the 75 genes have been studied in melanoma (Supplementary Table 3). These 30 genes include genes involved in the cell cycle (CKS2, CDC2, CCNB1, CENPF, and DHFR), mitosis (HCAP-G and STK6), mitotic spindle checkpoint (BUB1), inhibition (BIRC5) or stimulation (GPR105) of apoptosis, DNA replication (TOP2A, RRM2, TYMS, PCNA, MCM4, and MCM6), stress response (GLRX2, DNAJA1, HSPA4, HSPA5, HSPD1,

**Table 3.** Pearson's correlation coefficients between the logarithmic ratios obtained, respectively, from microarrays and from relative expression by quantitative real-time polymerase chain reaction for the study series of 83 primary melanomas

Gene	Correlation coefficient	No. of observations
KLK7	.82	70
PCNA	.86	81
KPNA2	.86	82
MCM6	.83	82
MCM3	.82	83
MCM4	.76	82
PTGDS	.73	79
RFC4	.77	82
NME1	.74	82

and TXNIP), ubiquitin cycle (SIP), actin and calmodulin binding (CNN3), intracellular signaling (STMN2), negative regulation of Wnt-signaling pathway (CTNNBIP1), inhibition of Mitf expression (EMX2), regulation of proteolysis (TNA), testis cancer (CML66), and metastasis suppression (NME1).

As well as these genes with known activity in melanoma, our study identified genes involved in tumor progression of melanoma that also have been previously shown to be involved in other cancers. Of particular interest are the 33 genes whose expression is higher in melanomas from patients that did not metastasize to distant sites within 4 years after diagnosis than in those that metastasized by 4 years, suggesting an antimetastatic role for these genes (Supplementary Table 3). These 33 genes included PROM2, a gene expressed in neural stem cells but absent from various cancer cell lines (26); genes for several protease inhibitors, such as SPINT2, which encodes the serine peptidase inhibitor Kunitz type 2; and CST3 and CST5, which encode the small molecular weight protein inhibitors of cysteine proteinases, such as cathepsins B, H, and L; PCP4, which encodes the neuron-specific protein PEP-19; CXCL14, which encodes a chemokine with tumor-suppressive functions in prostate cancer (27); and PTGDS, which encodes prostaglandin D synthase. Although the biologic significance of prostaglandins in melanoma progression is unknown, increased expression of prostaglandin D synthase has been correlated with inhibition of cell proliferation in ovarian cancer cells (28).

Among 23 gene products studied, eight immunohistochemical markers were statistically significantly associated with distant metastasis-free survival and/or overall survival. The lack of prognostic significance for the immunoreexpression of the 15 other genes could be explained either by posttranscriptional modifications or by technical constraints, such as those linked to the various fixatives used since the 1980s, when collection of specimens for this study began. Immunoreactivity for these markers was restricted to melanoma cells, thereby excluding the analysis of stromal cells in our study. Because our results were validated in an independent set of 176 primary melanomas, the use of this battery of five antibodies may in the future improve the determination of prognosis and treatment stratification. Karyopherin- $\alpha$ 2, which participates in the nuclear import of proteins with a classical nuclear localization signal, is associated with poor differentiation in several cancers (29), and karyopherin- $\alpha$ 2 overexpression identified melanoma patients with poor outcome.

To ensure that genomic DNA is fully replicated only once during each cell division, it is particularly important that the initiation of DNA replication be tightly controlled in rapidly dividing cells and that the site of the origin where DNA replication begins is also tightly regulated. This regulation is ensured by the geminin-mediated inhibition of Cdt1, which controls further recruitment of specific minichromosome maintenance proteins (MCM2–7) to the replication origins. The list of 361 sequences associated with clinical outcome, at an adjusted statistical significance level of *P* value of less than .05, contained almost all MCM proteins that unwind DNA at replication origins. These findings are particularly striking because expression of MCM proteins is absolutely necessary for the active DNA replication and cell division that is required for tumor growth. This pathway is also regulated by geminin, which is regarded as a tumor suppressor gene that regulates genomic stability during cell cycle (30,31). Geminin regulates genomic stability in cycling cells by preventing firing (or activation) of new replication origins before completion of a mitotic cycle, to ensure that DNA is replicated

only once per cell cycle. Geminin prevents the interaction between Cdt1, an essential component of the replication licensing system, and MCM2–7, the minichromosome maintenance helicase complex (30). Overexpression of the MCMs and geminin has been found in other types of tumors (32), but to our knowledge this is the first demonstration of such overexpression in primary melanomas associated with high risk of metastasis and death. Geminin protein was detected by immunohistochemistry exclusively in the cytoplasm of melanoma cells but was detected predominantly in the nucleus of uninvolved overlying epidermal cells. The lack of geminin in the nuclei of melanoma cells may indicate that a posttranslational modification is interfering with its nuclear import and thereby contributing to tumor progression by enhancing genomic instability.

Finally, in tumors that were associated with poor prognosis, it appears that the global increase in cell replication (as also indicated by overexpression of PCNA) is also associated with dysregulation of the activation of replication origins, which could result in both enhanced proliferation and destabilization of the genome. In this study, we have identified genes whose expression appears to be associated with *in vivo* metastatic dissemination of primary cutaneous melanomas.

To obtain enough mRNA for our analyses, this study was limited to melanomas in the vertical growth phase and/or with a thickness of more than 1 mm. Consequently, the genes identified in this study may be relevant for the transition the vertical growth phase to metastasis but do not provide information on earlier steps in tumor progression. Moreover, as this study was retrospective in nature, sentinel node involvement was not assessed in our study population. The supervised analysis to predict distant metastasis-free survival at 4 years showed a misclassification rate of 29% for gene classifier as estimated by leave-one-out cross-validation. The permutation test provided a *P* value of .04, indicating that the observed misclassification rate was statistically significantly better than what could be expected by chance. The prognostic accuracy of our 254-gene classifier was similar to that predicted by the well characterized prognostic factors tumor thickness and ulceration.

Some of our results, especially those related to dysfunction of the activation or firing or replication origins, provide new diagnostic tools for the accurate diagnosis of melanoma and shed new light on the molecular mechanisms underlying poor prognosis in melanoma patients. Several of the genes listed in the 254-gene signature are targets of established or experimental therapies in other solid cancers, including kinesins (33), cyclin-dependent kinases [such as CDK-2 and -4 (34)], primase (35), and DHFR (36). Also, with regard to the list of proteins whose expression was correlated with distant metastasis-free survival or overall survival, peptides have been developed that bind to karyopherin- $\alpha$ 2 (37) and prevent its interaction with nuclear factor  $\kappa$ B and with geminin (38), resulting in suppression of cell proliferation. These results emphasize the importance of some interactions in melanoma progression, such as the control of geminin-CDT1 interaction or of MCMs binding, that may shed light on the molecular mechanism underlying poor prognosis in melanoma patients.

## REFERENCES

(1) Geller AC, Miller DR, Annas GD, Demierre MF, Gilchrist BA, Koh HK. Melanoma incidence and mortality among US whites, 1969-1999. *JAMA* 2002;288:1719-20.

- (2) MacKie RM, Bray CA, Hole DJ, Morris A, Nicolson M, Evans A, et al. Incidence of and survival from malignant melanoma in Scotland: an epidemiological study. *Lancet* 2002;360:587-91.
- (3) Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, et al. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 2000;406:536-40.
- (4) Carr KM, Bittner M, Trent JM. Gene-expression profiling in human cutaneous melanoma. *Oncogene* 2003;22:3076-80.
- (5) Clark EA, Golub TR, Lander ES, Hynes RO. Genomic analysis of metastasis reveals an essential role for RhoC. *Nature* 2000;406:532-5.
- (6) Hoek K, Rimm DL, Williams KR, Zhao H, Ariyan S, Lin A, et al. Expression profiling reveals novel pathways in the transformation of melanocytes to melanomas. *Cancer Res* 2004;64:5270-82.
- (7) McDonald SL, Edington HD, Kirkwood JM, Becker D. Expression analysis of genes identified by molecular profiling of VGP melanomas and MGP melanoma-positive lymph nodes. *Cancer Biol Ther* 2004;3:110-20.
- (8) Pavay S, Johansson P, Packer L, Taylor J, Stark M, Pollock PM, et al. Microarray expression profiling in melanoma reveals a BRAF mutation signature. *Oncogene* 2004;23:4060-7.
- (9) Wang E, Miller LD, Ohnmacht GA, Mocellin S, Perez-Diez A, Petersen D, et al. Prospective molecular profiling of melanoma metastases suggests classifiers of immune responsiveness. *Cancer Res* 2002;62:3581-6.
- (10) de Wit NJ, Rijntjes J, Diepstra JH, van Kuppevelt TH, Weidle UH, Ruiter DJ, et al. Analysis of differential gene expression in human melanocytic tumour lesions by custom made oligonucleotide arrays. *Br J Cancer* 2005;92:2249-61.
- (11) Busam KJ, Zhao H, Coit DG, Kucukgol D, Jungbluth AA, Nobrega J, et al. Distinction of desmoplastic melanoma from non-desmoplastic melanoma by gene expression profiling. *J Invest Dermatol* 2005;124:412-8.
- (12) Vogl A, Sartorius U, Vogt T, Roesch A, Landthaler M, Stolz W, et al. Gene expression profile changes between melanoma metastases and their daughter cell lines: implication for vaccination protocols. *J Invest Dermatol* 2005;124:401-4.
- (13) Haqq C, Nosrati M, Sudilovsky D, Crothers J, Khodabakhsh D, Pulliam BL, et al. The gene expression signatures of melanoma progression. *Proc Natl Acad Sci U S A* 2005;102:6092-7.
- (14) Balch CM, Buzaid AC, Soong SJ, Atkins MB, Cascinelli N, Coit DG, et al. Final version of the American Joint Committee on Cancer staging system for cutaneous melanoma. *J Clin Oncol* 2001;19:3635-48.
- (15) Elder D. Tumor progression, early diagnosis and prognosis of melanoma. *Acta Oncol* 1999;38:535-47.
- (16) Dobbin K, Shih JH, Simon R. Questions and answers on design of dual-label microarrays for identifying differentially expressed genes. *J Natl Cancer Inst* 2003;95:1362-9.
- (17) Simon R, Peng A. BRB-ArrayTools. Bethesda, MD; 2003. Available at: <http://linus.nci.nih.gov/BRB-ArrayTools.html>.
- (18) Hughes TR, Mao M, Jones AR, Burchard J, Marton MJ, Shannon KW, et al. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat Biotechnol* 2001;19:342-7.
- (19) Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* 2003;19:368-75.
- (20) Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005;365:488-92.
- (21) Ntzani EE, Ioannidis JP. Predictive ability of DNA microarrays for cancer outcomes and correlates: an empirical assessment. *Lancet* 2003;362:1439-44.
- (22) Simon R, Radmacher MD, Dobbin K, McShane LM. Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. *J Natl Cancer Inst* 2003;95:14-8.
- (23) Radmacher MD, McShane LM, Simon R. A paradigm for class prediction using gene expression profiles. *J Comput Biol* 2002;9:505-11.
- (24) Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. *Nat Genet* 2003;33:49-54.
- (25) Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ. Gene expression profiles of primary breast tumors maintained in distant metastases. *Proc Natl Acad Sci U S A* 2003;100:15901-5.
- (26) Fargeas CA, Flore M, Huttner WB, Corbeil D. Characterization of prominin-2, a new member of the prominin family of pentaspan membrane glycoproteins. *J Biol Chem* 2003;278:8586-96.
- (27) Schwarze SR, Luo J, Isaacs WB, Jarrard DF. Modulation of CXCL14 (BRAK) expression in prostate cancer. *Prostate* 2005;64:67-74.

- (28) Su B, Guan M, Xia J, Lu Y. Stimulation of lipocalin-type prostaglandin D synthase by retinoic acid coincides with inhibition of cell proliferation in human 3AO ovarian cancer cells. *Cell Biol Int* 2003;27:587–92.
- (29) Blagosklonny MV, An WG, Melillo G, Nguyen P, Trepel JB, Neckers LM. Regulation of BRCA1 by protein degradation. *Oncogene* 1999;18:6460–8.
- (30) Lee C, Hong B, Choi JM, Kim Y, Watanabe S, Ishimi Y, et al. Structural basis for inhibition of the replication licensing factor Cdt1 by geminin. *Nature* 2004;430:913–7.
- (31) Okorokov AL, Orlova EV, Kingsbury SR, Bagneris C, Gohlke U, Williams GH, et al. Molecular structure of human geminin. *Nat Struct Mol Biol* 2004;11:1021–2.
- (32) Gonzalez MA, Tachibana KE, Laskey RA, Coleman N. Control of DNA replication and its potential clinical exploitation. *Nat Rev Cancer* 2005;5:135–41.
- (33) Wood KW, Cornwell WD, Jackson JR. Past and future of the mitotic spindle as an oncology target. *Curr Opin Pharmacol* 2001;1:370–7.
- (34) Senderowicz AM. Novel direct and indirect cyclin-dependent kinase modulators for the prevention and treatment of human neoplasms. *Cancer Chemother Pharmacol* 2003;52 Suppl 1:S61–73.
- (35) Blasco F, Penuelas S, Cascallo M, Hernandez JL, Alemany C, Masa M, et al. Expression profiles of a human pancreatic cancer cell line upon induction of apoptosis search for modulators in cancer therapy. *Oncology* 2004;67:277–90.
- (36) Hanauske AR. Translational research with pemetrexed in breast cancer. *Oncology (Huntingt)* 2004;18(13 Suppl 8):66–9.
- (37) Cunningham MD, Cleaveland J, Nadler SG. An intracellular targeted NLS peptide inhibitor of karyopherin alpha:NF-kappa B interactions. *Biochem Biophys Res Commun* 2003;300:403–7.
- (38) Yoshida K, Inoue I. Peptide binding to Geminin and inhibitory for DNA replication. *Biochem Biophys Res Commun* 2004;317:218–22.

## NOTES

V. Winnepenninckx, V. Lazar, and S. Michiels contributed equally to this work.

We thank Mrs. Virginie Marty for preparing the tissue microarrays and Bernadette Smets for immunohistochemical assistance.

Manuscript received July 25, 2005; revised January 20, 2006; accepted February 21, 2006.