Research Article

Diagnostic Distinction of Malignant Melanoma and Benign Nevi by a Gene Expression Signature and Correlation to Clinical Outcomes

Jennifer S. Ko[1], Balwir Matharoo-Ball[2], Steven D. Billings[1], Brian J. Thomson[2], Jean Y. Tang[3], Kavita Y. Sarin[3], Emily Cai[3], Jinah Kim[3], Colleen Rock[4], Hillary Z. Kimbrell[4], Darl D. Flake II[4], M. Bryan Warf[4], Jonathan Nelson[4], Thaylon Davis[4], Catherine Miller[4], Kristen Rushton[4], Anne-Renee Hartman[4], Richard J. Wenstrup[4], and Loren E. Clarke[4]

Abstract

Background: Histopathologic examination alone can be inadequate for diagnosis of certain melanocytic neoplasms. Recently, a 23-gene expression signature was clinically validated as an ancillary diagnostic test to differentiate benign nevi from melanoma. The current study assessed the performance of this test in an independent cohort of melanocytic lesions against clinically proven outcomes.

Methods: Archival tissue from primary cutaneous melanomas and melanocytic nevi was obtained from four independent institutions and tested with the gene signature. Cases were selected according to pre-defined clinical outcome measures. Malignant lesions were defined as stage I–III primary cutaneous melanomas that produced distant metastases (metastatic to sites other than proximal sentinel lymph node(s)) following diagnosis of the primary lesion. Melanomas that were metastatic at the time of diagnosis, all re-excisions, and lesions with <10% tumor volume were excluded. Benign lesions were defined as cutaneous melanocytic lesions with no adverse long-term events reported.

Results: Of 239 submitted samples, 182 met inclusion criteria and produced a valid gene expression result. This included 99 primary cutaneous melanomas with proven distant metastases and 83 melanocytic nevi. Median time to melanoma metastasis was 18 months. Median follow-up time for nevi was 74.9 months. The gene expression score differentiated melanoma from nevi with a sensitivity of 93.8% and a specificity of 96.2%.

Conclusions: The results of gene expression testing closely correlate with long-term clinical outcomes of patients with melanocytic neoplasms.

Impact: Collectively, this provides strong evidence that the gene signature adds valuable adjunctive information to aid in the accurate diagnosis of melanoma. *Cancer Epidemiol Biomarkers Prev; 26(7); 1107–13. ©2017 AACR.*

Introduction

Melanoma is now the seventh most common cancer among adults, and approximately 2 million melanocytic lesions are biopsied annually in the United States for clinical suspicion of melanoma (1). Although the majority of these lesions are benign, melanoma will account for an estimated 10,130 deaths in 2016 (2). However, many melanomas are curable if detected early, with a 10-year survival rate of 95% for stage IA melanoma compared to only 10% to 15% for stage IV (3). This marked difference in biologic potential and clinical outcome underscores the value of early and accurate diagnosis.

Many melanocytic lesions can be accurately classified as benign or malignant by a skilled histopathologist using conventional light microscopy; however, numerous studies have demonstrated that pathologists arrive at different diagnoses in 15% to 38% of cases when evaluating the same specimens (1, 4–6). Studies have also shown that the consensus diagnosis of multiple experienced dermatopathologists may not align with ultimate clinical outcomes (7, 8).

To improve diagnostic accuracy, ancillary techniques have been sought to provide adjunctive diagnostic information for pathologists confronting histopathologically ambiguous melanocytic lesions. These methods include array comparative genomic hybridization (aCGH) and fluorescence in-situ hybridization (FISH), both of which rely on the detection of chromosomal copy number aberrations within neoplastic melanocytes (9–11). Recently, a gene expression signature has been developed to evaluate 23 genes that are differentially expressed in malignant melanoma and benign melanocytic nevi using quantitative reverse transcription polymerase chain reaction (qRT-PCR) (12, 13).

Previous validations of the gene signature have demonstrated that this test is capable of differentiating melanoma from benign nevi with greater than 90% accuracy compared with consensus diagnosis of multiple expert dermatopathologists (12, 13). These studies are in line with standard practice to validate new diagnostic methods against the current diagnostic standard or clinical outcomes (14). Collecting long-term clinical follow-up for melanocytic lesions can be a challenge, given the duration of follow-

AACR 1107

up required for benign cases and the high rate of excision of suspected melanoma. In addition, the requirement for metastatic disease to prove malignancy may limit the types of lesions available for analysis. For example, some lesions that fulfill morphologic criteria for melanoma may lack metastatic capability (15). For this reason, aCGH, FISH, and the 23-gene signature have all been previously validated against the current diagnostic standard of histopathology. Although the known limitations of histopathology may impact the reported diagnostic accuracy in such studies, the extent of this impact is unknown. In addition, variability in the diagnostic accuracy of histopathology across different melanocytic subtypes suggests that the magnitude of this effect is likely dependent on the composition and size of the validation cohort.

To date, evaluation of adjunctive melanoma diagnostic methods against clinical outcomes has been limited to relatively small samples sizes (9, 16–24), and previously published validations of the 23-gene signature have been based on histopathologic diagnosis. However, validation against a large, clinical cohort is necessary to accurately assess the performance of a diagnostic test for clinical use. Therefore, the current study aimed to evaluate the diagnostic accuracy of the 23-gene signature utilizing a large cohort of melanocytic lesions with clinical outcomes. This was done by determining the sensitivity of the gene signature in primary cutaneous melanomas with proven distant metastases and the specificity in melanocytic nevi with long-term event-free follow-up.

Materials and Methods

Sample cohort

The study was conducted with Institutional Review Board oversight and was approved with a waiver for individual patient informed consent (Quorum Review IRB). Archived melanocytic neoplasms not previously assessed by the gene expression signature that met inclusion criteria were independently identified and submitted for testing by the Cleveland Clinic (Cleveland, OH), Nottingham University (Nottingham, United Kingdom), and Stanford University (Stanford, CA). Additional samples were acquired from Avaden Biosciences. Cases were selected based upon pre-specified clinical outcome measures to identify malignant and benign lesions.

Malignant lesions were stage I, II, or III primary cutaneous melanomas that produced distant metastases subsequent to diagnosis of the primary lesion, where distant metastases were defined as melanoma present at anatomic sites not including the proximal sentinel lymph node(s), given the propensity for borderline Spitz lesions to spread exclusively to sentinel lymph nodes. Melanomas known to be metastatic at the time of initial diagnosis were excluded. No minimum length of follow-up was required for cases with documented distant metastases. Benign lesions were required to be diagnosed as benign nevi and have clinical follow-up after the initial biopsy that demonstrated no evidence of local recurrence or in-transit, regional, or distant metastases. A disease-free follow-up time of at least 5 years was recommended for benign sample submission/acquisition from participating institutions. Samples were excluded from patients with a previous history of melanoma, a diagnosis of synchronous or metachronous melanoma or other cancer during the follow-up period, or who received immune therapy or chemotherapy prior to diagnosis of mel-

anoma. Re-excision specimens, metastatic melanomas, and lesions that did not contain at least 10% tumor volume were also excluded in accordance with the technical specifications of the assay (13).

De-identified clinical data for each sample were obtained from the submitting institution and included patient age, gender, anatomic site of the primary lesion, Breslow depth, ulceration status, sentinel lymph node biopsy status, anatomic location of metastases, time to metastasis, and length of follow-up. Eleven melanomas were received without ulceration status indicated. The presence or absence of ulceration was determined by review of hematoxylin and eosin (H&E) stained sections of the lesions by a dermatopathologist (L.E. Clarke) at the testing institution. Histopathologic subtypes (e.g., superficial spreading melanoma, nodular melanoma, etc.) were determined for all cases by a panel of three dermatopathologists (J.S. Ko, S.D. Billings, and L.E. Clarke) who were blinded to the diagnosis of the submitting pathologist(s) and the gene expression score.

Gene expression testing

Gene expression analysis (Myriad Genetic Laboratories, Inc.) was carried out on archival formalin-fixed paraffin-embedded (FFPE) tissue sections from each lesion according to previously described methods (12, 13, 25). Briefly, a pathologist identified representative areas of the lesion using a single H&E-stained section. Lesions with <10% tumor volume were excluded from testing based on previous evidence that these lesions can produce false-negative results due to dilution by nonlesional cells (13). The representative area of the lesion identified for testing was macrodissected from unstained tissue sections and pooled into a single tube for RNA extraction. The qRT-PCR assay measured the differential expression of 23 genes (14 tumor marker genes and 9 housekeeper genes, Table 1).

The gene signature included three categories of tumor marker genes that are differentially expressed in melanoma: *PRAME* (cell differentiation), *S100A9* and 4 related genes (cell signaling response to tissue damage), and 8 immune group genes (tumor immune response signaling). The S100A9 and immune group categories consisted of genes with highly correlated expression and similar biological functions. The differential expression of all genes in each category was consolidated into average component scores. A weighted algorithm was applied to normalize the tumor marker genes to the average differential expression of the housekeeper genes and produce a single numeric score plotted on a scale ranging from -16.7 to $+11.1$ (12). Scores from -16.7 to -2.1 were reported as "likely benign," scores from -2.0 to -0.1 were

Table 1. List of genes included in the qRT-PCR assay

Gene component 1	Gene component 2	Gene component 3	Housekeeping genes
PRAME	*S100A9*	*CCL5*	*CLTC*
	S100A7	*CD38*	*MRFAPI*
	S100A8	*CXCL10*	*PPP2CA*
	S100A12	*CXCL9*	*PSMA1*
	PI3	*IRF1*	*RPL13A*
		LCP2	*RPL8*
		PTPRC	*RPS29*
		SELL	*SLC25A3*
			TXNL1

reported as "indeterminate," and scores from 0.0 to +11.1 were reported as "likely malignant."

Diagnostic accuracy of the gene expression signature

The correlation between the gene expression test result and the clinical outcome (benign or malignant) of the neoplasm was assessed for all eligible cases. Sensitivity (proportion of correctly identified positive/malignant cases) and specificity (proportion of correctly identified negative/benign cases) were calculated after excluding lesions that produced indeterminate results (scores between −2.0 and 0.0).

Results

Cohort description

A total of 293 samples (135 benign, 158 malignant) were submitted by the participating institutions. Of these, 53 did not satisfy the pre-designated inclusion criteria (Fig. 1). An additional 58 samples failed to produce a valid gene expression score (Fig. 1). This failure rate may be due to insufficient mRNA quality, as mRNA extracted from archival samples is more prone to fragmentation compared with recently prepared FFPE lesions (12). In support of this, lesions within the current cohort that were tested within 5 years of the primary biopsy had a much lower failure rate (2/22, 9.1%) compared with lesions tested more than 5 years after biopsy (56/218, 25.7%).

Table 2 summarizes the clinical and histopathologic characteristics of the final cohort of 182 cases, which included 83 benign nevi and 99 malignant melanomas. The average age for patients with malignant and benign lesions was 63 and 42 years, respectively. Sixty-seven percent of melanomas were from males, whereas 33% were from females. Median follow-up time for benign lesions was 74.9 months, with 70% (58/83) of cases having at least 5 years (60 months) of clinical follow-up. Median time to metastasis for the malignant samples was 18 months.

The final cohort included a diverse range of histopathologic variants (Table 3). Overall, 12 distinct melanoma subtypes were tested, including nodular melanoma, superficial spreading melanoma, acral melanoma, and nevoid melanoma. The cohort also contained 18 distinct benign subtypes, including compound nevi, intradermal nevi, compound dysplastic nevi, and Spitz nevi. Median Breslow depth for the melanoma samples was 2.8 mm. The Breslow depth was 0 to 1.0 mm for 16 of the melanomas (including one melanoma *in situ*), 1.01 to 2.0 mm for 20, and >2.0 mm for 62. Ulceration was present in 39% of melanomas (Table 2).
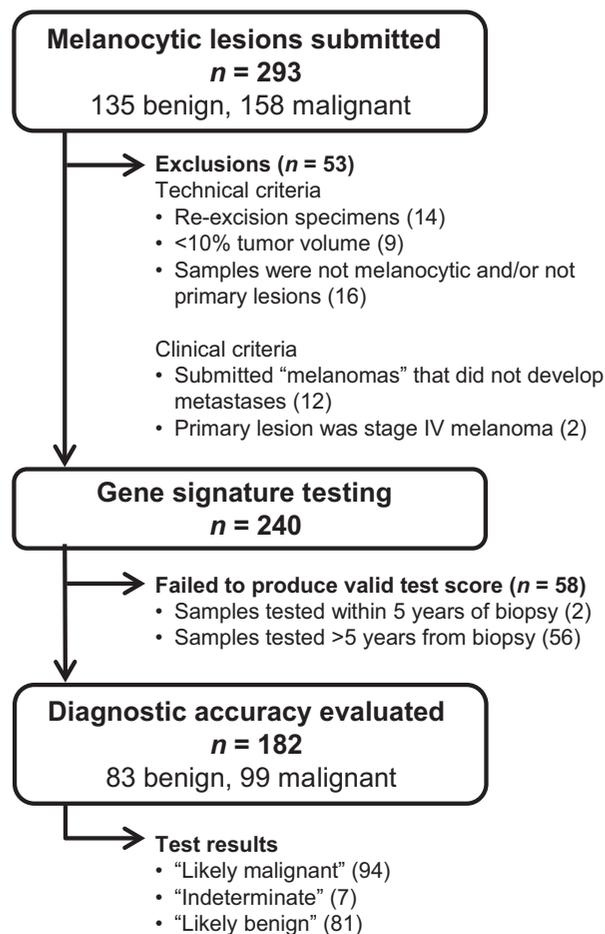
Diagnostic accuracy of the gene expression signature

Of the 99 melanoma cases included in the final cohort, 91 produced positive ("likely malignant") scores, 6 produced negative ("likely benign") scores, and 2 produced indeterminate scores (Table 3, Supplementary Fig. S1). The score accurately classified as malignant all thin (Breslow thickness ≤ 1.0 mm) melanomas ($n = 16$; Supplementary Figs. S2 and S3), all "nevoid" melanomas ($n = 7$; Supplementary Fig. S4), all but one melanoma arising within pre-existing nevi ($n = 6$; Supplementary Fig. S5), and all lentigo maligna melanomas ($n = 2$). Sentinel lymph node biopsy was performed in 14 outcome-based melanoma cases, six of which were positive for melanoma. The primary melanomas from all 14 cases produced positive gene expression scores (Table 4).

Of the six metastasis-proven melanomas that produced scores within the benign range of the scale (false-negative results), five had a Breslow thickness >2.0 mm (Table 5). Three of the melanomas that produced false-negative results were desmoplastic melanomas ("pure type"; Supplementary Fig. S6), two were superficial spreading melanoma (one with a substantial vertical growth phase, the other a melanoma arising within a nevus in which much of the targeted tissue appears to have been nevus), and one was a nodular melanoma ("spindle-cell type"). Two melanomas produced scores within the indeterminate range. This included a superficial spreading melanoma as well as a lesion classified as melanoma, not otherwise specified, in which isolated clusters of malignant melanocytes were surrounded by dense aggregates of lymphocytes.

Of the 83 benign cases, 75 produced negative ("likely benign") scores, 3 produced positive ("likely malignant") scores, and 5 produced indeterminate scores (Table 3; Supplementary Fig. S1). Of the three lesions with event-free follow-up that produced malignant scores (apparent false positives), two were compound dysplastic nevi with severe atypia (344 and 84 months of event-free follow-up, respectively), and one was a compound nevus with

Figure 1.
Study flow chart.

Ko et al.

Table 2. Summary of clinical and histologic data by disease status

Variable	Summary statistic	Disease status	
		Malignant (*n* = 99) *N* (%)	Benign (*n* = 83) *N* (%)
Age	Mean (SD)	63 (14)	42 (17)
Gender	Male (%)	66 (67)	29 (35)
	Female (%)	33 (33)	54 (65)
Breslow depth	Median (range)	2.8 (0.4, 17.9)	NA
	0–1.0 mm (%)	16 (16)	—
	1.01–2.0 mm (%)	20 (20)	—
	>2.0 mm (%)	62 (63)	—
	Missing	1	—
Ulceration status	Present (%)	39 (39)	NA
	Absent (%)	60 (61)	—
SLN	Positive	6 (11)	NA
	Negative	8 (15)	—
	Not biopsied	39 (74)	—
	Missing	46	—
Length of follow-up (months)	Median	18	74.9

Abbreviations: NA, not applicable; SLN, sentinel lymph node.

congenital and dysplastic features (180 months of event-free follow-up). One of these lesions was re-excised with a margin of normal skin. Another was not re-excised because the original biopsy site could not be definitively identified due to extensive solar damage and prior treatment with liquid nitrogen for non-melanoma skin cancer (Supplementary Fig. S7). Treatment for the third lesion was unknown. Five lesions without adverse events detected during the follow-up period produced scores within the indeterminate zone of the scale. Four were dysplastic nevi and one was a compound nevus with "special site" features (Table 5).

The sensitivity and specificity of the signature was determined for all eligible samples that produced "likely benign" or "likely malignant" test results (*n* = 175). The resulting sensitivity and specificity were 93.9% and 96.2%, respectively. When benign samples with less than 5 years of clinical follow-up are excluded, the specificity was 94.6%.

Table 3. Performance of the gene signature within individual histopathologic subtypes

	Signature classification		
	Malignant	Benign	Indeterminate
Melanoma subtype			
Nodular melanoma	38	1	0
Nodular melanoma ex nevus	1	0	0
Superficial spreading melanoma	29	1	1
Superficial spreading melanoma ex nevus	4	1	0
Acral melanoma	6	0	0
Acral lentiginous melanoma	2	0	0
Desmoplastic melanoma	0	3	0
Lentigo maligna melanoma	2	0	0
Nevoid melanoma	7	0	0
Spitzoid melanoma	1	0	0
Melanoma, NOS	0	0	1
Melanoma *in situ*	1	0	0
Total	91	6	2
Nevi subtype			
Compound nevi	0	23	0
Compound dysplastic nevi	0	8	0
Compound dysplastic nevi with severe atypia	2	0	3
Compound nevi with congenital and dysplastic features	1	0	0
Compound nevi with congenital features	0	9	0
Compound nevi with special site features	0	0	1
Intradermal nevi	0	21	0
Intradermal nevi with congenital features	0	3	0
Intradermal nevi, mitotically active	0	1	0
Junctional nevi	0	1	0
Junctional dysplastic nevi	0	1	1
Pigmented spindle cell nevus	0	2	0
Spitz nevi	0	1	0
Spitz nevi with desmoplasia	0	1	0
Spitzoid dysplastic nevi	0	1	0
Combined nevi (blue and intradermal)	0	1	0
Combined nevi (desmoplastic and conventional features)	0	1	0
Acral nevi	0	1	0
Total	3	75	5

Abbreviation: NOS, not otherwise specified.

Table 4. Clinical, histologic, and molecular data for cases that received a sentinel lymph node biopsy

Sample ID	Disease status	Histopathologic subtype	Age	Gender	Anatomic site	Breslow depth	Ulceration status	SLN status	Length of follow-up (months)	Gene expression score
9	Malignant	Nodular melanoma	54	M	Left thigh	6	Present	Positive	8	5.2
12	Malignant	Superficial spreading melanoma	60	M	Right shoulder	0.45	Absent	Negative	4	1.5
22	Malignant	Superficial spreading melanoma	76	F	Left lower leg	2.18	Absent	Positive	12	3.8
24	Malignant	Nodular melanoma	80	M	Right scapula	1.93	Present	Negative	37	8.1
25	Malignant	Acral lentiginous melanoma	73	M	Left foot	4.4	Absent	Positive	18	9.4
91	Malignant	Nodular melanoma	56	M	Left leg	6	Present	Positive	144	8.6
94	Malignant	Nevoid melanoma	34	M	Right abdomen	1.25	Present	Negative	42	6.7
96	Malignant	Nodular melanoma	44	F	Right upper back	6.59	Absent	Positive	6	6.7
140	Malignant	Superficial spreading melanoma	53	M	Back	1.2	Absent	Negative	79	6
143	Malignant	Acral lentiginous melanoma	69	M	Left foot	4.1	Absent	Negative	10	5.7
144	Malignant	Superficial spreading melanoma	75	F	Right lower leg	0.67	Absent	Negative	22	2.6
145	Malignant	Nodular melanoma	70	F	Right posterior thigh	1.25	Absent	Negative	66	6.6
159	Malignant	Nevoid melanoma	51	M	Back	8	Absent	Positive	60	3.1
182	Malignant	Nevoid melanoma	71	F	Scalp	1.9	Absent	Negative	13	3.5

Abbreviations: F, female; M, male; SLN, sentinel lymph node.

Discussion

Histopathology is an accurate diagnostic method for many melanocytic lesions and is currently considered the "gold" standard for the diagnosis of melanoma. In line with standard practice, adjunctive molecular tests to aid in the diagnosis of melanoma have largely been developed and evaluated using histopathology as a reference standard (9–12). Previous valida-tion studies of the 23-gene expression signature used here have reported greater than 90% diagnostic accuracy by comparison with consensus diagnosis of multiple expert dermatopathologists (12, 13). However, a separate study of adjunctive diagnostic methods reported a sensitivity of only 55% and a specificity of 88% when the gene signature was compared with consensus histopathologic diagnosis (26). This variability in performance

Table 5. Clinical, histologic, and molecular data for samples producing false-negative, false-positive, and indeterminate test results

Sample ID	Disease status	Histopathologic subtype	Age	Gender	Anatomic site	Breslow depth	Ulceration status	Length of follow-up (months)	Gene expression score
2	Malignant	Melanoma, NOS	58	F	Left chest	2.1	Absent	34	−0.8
15	Malignant	Desmoplastic melanoma	82	M	Right cheek	7.5	Absent	2	−3.3
16	Malignant	Desmoplastic melanoma	58	M	Left shoulder	3.7	Absent	5	−6.5
52	Malignant	Superficial spreading melanoma ex nevus	48	M	Left shoulder	3.2	Absent	13	−3.3
150	Malignant	Nodular melanoma	73	M	Scalp	4	Absent	48	−3.7
154	Malignant	Superficial spreading melanoma	64	M	Right lower extremity	1.67	Absent	33	−1.9
156	Malignant	Superficial spreading melanoma	65	M	Right postauricular	3	Absent	12	−5.7
168	Malignant	Desmoplastic melanoma	73	M	Face	1.2	Absent	18	−2.6
29	Benign	Compound nevus with congenital and dysplastic features (mild atypia)	18	F	Arm	NA	NA	344	0.2
36	Benign	Junctional dysplastic nevus with mild atypia	62	M	Back	NA	NA	84	−1.6
40	Benign	Compound dysplastic nevus with severe atypia	67	F	Right thigh	NA	NA	7	−0.6
79	Benign	Compound dysplastic nevus with severe atypia	48	M	Back	NA	NA	30	−0.8
80	Benign	Compound dysplastic nevus with severe atypia	46	M	Left scapular area	NA	NA	10	−0.1
146	Benign	Compound dysplastic nevus with severe atypia	54	M	Posterior shoulder	NA	NA	84	4.1
166	Benign	Compound dysplastic nevus with severe atypia	57	M	Left arm	NA	NA	180	7.7
172	Benign	Compound nevus with special site features	41	F	Left groin	NA	NA	72	−0.6

Abbreviations: F, female; M, male.

is consistent with the performance of FISH relative to histopathologic diagnosis, with reported sensitivities ranging from 35% to 87% (10, 18, 24, 26). Important differences in study design, such as the tumor subtypes analyzed and the methods for histopathologic review, likely drive the variability in reported sensitivities of ancillary diagnostic tests. In particular, histopathologic review can vary between studies based on the number of reviewing dermatopathologists, requirement for consensus, and whether reviewers were blinded to other diagnoses made by histopathology or other adjunctive tests. This highlights that although histopathology may be a reasonable reference standard in a carefully designed study, a more robust reference is needed to confirm diagnostic accuracy.

Clinical outcome data offer an objective standard for determining the accuracy of diagnostic methods; however, the challenge of compiling a sufficient number of cases to support robust clinical validation presents significant challenges. This includes the limited proportion of melanoma patients who develop distant metastases, the large number of patients who are lost to follow-up before a benign diagnosis can be definitively confirmed, and the difficulties encountered in acquiring corresponding clinical outcome data and primary biopsy tissue from multiple separate institutions. For these reasons, most reports validating molecular diagnostic tests against clinical outcomes for melanocytic lesions have been small subset analyses of larger studies that use histopathology as the reference standard (9, 16–24).

The current study evaluated the accuracy of a 23-gene expression signature in distinguishing clinical outcome proven melanomas from benign nevi. The cohort included 99 melanomas that metastasized following initial diagnosis and 83 nevi with long-term disease-free follow-up. Excluding indeterminate cases, the signature differentiated these lesions with 93.9% sensitivity and 96.2% specificity.

Many of the melanomas in this cohort were relatively large tumors (63% with a Breslow depth of greater than 2.0 mm). However, a substantial number of early-stage ("thin") melanomas were also included, several of which would have been considered relatively low risk for metastasis by histopathologic staging (i.e., Breslow depth < 1.0 mm, no dermal mitoses, no ulceration; Supplementary Figs. S2 and S3). More than 70% of melanomas in the United States are now <1.0 mm in thickness at the time of diagnosis and account for up to 27% of fatal melanomas (27). The gene expression signature accurately classified all tumors <1.0 mm in thickness included here, suggesting it may prove useful in detecting the subset of thin melanomas with metastatic capacity. Conversely, five of the six false-negative results were produced by "thick" melanomas (Breslow depth greater than 2.0 mm) that may be less likely to pose diagnostic challenges. Although this may be expected given the prevalence of melanomas >2.0 mm thick within the study cohort, it is also possible that molecular differences between thick and thin melanomas may influence the accuracy of the gene expression signature (28–30).

The data presented here show high sensitivity and specificity across the most common clinico-histopathologic melanoma and nevus subtypes. One potential exception appears to be a reduced sensitivity in desmoplastic melanomas (Supplementary Fig. S6). The study cohort contained three desmoplastic melanomas, all of which produced a false-negative result. Other ancillary diagnostic techniques have also shown diminished sensitivity in desmoplastic melanomas relative to other melanoma types, suggesting that desmoplastic melanomas may possess unique molecular characteristics that make them less amenable to detection by existing molecular methods (31–33).

A primary aim of this study was to determine the accuracy of the gene expression signature in detecting lesions with unequivocal proof of malignancy. Therefore, the inclusion criteria of the study defined melanomas as lesions that produced distant metastases subsequent to initial biopsy. Some lesions that satisfy histopathologic criteria for the diagnosis of melanoma but lack metastatic capability would therefore have been excluded. It is also possible that some of the apparent false positives were in fact genuine melanomas that were "cured" by the initial biopsy or excision. Indeed, the histopathologic features of several of the "false-positive" cases were sufficiently concerning that all three reviewing dermatopathologists were unable to exclude melanoma and would have recommended treatment as melanoma (Supplementary Fig. S7). This suggests the benign outcome in these cases might have been due to early surgical intervention. In such cases, the gene expression signature can provide additive information to support excision and close clinical surveillance as is typically recommended for melanocytic neoplasms with "severe" atypia, for example.

Overall, multiple lines of evidence suggest that the gene expression signature differentiates benign and malignant melanocytic lesions with a high degree of accuracy. The data presented here show that the gene signature had high diagnostic accuracy relative to long-term clinical outcomes. In combination with previous validations that used consensus histopathologic diagnosis as the reference standard, this supports the use of the gene signature as an adjunctive diagnostic test to enable the early and accurate diagnosis of melanoma. Additional studies on specific melanoma and nevus subtypes are ongoing and will provide additional insight regarding the performance characteristics of the test.

Disclosure of Potential Conflicts of Interest

Authors' Contributions

Conception and design: J.S. Ko, C. Rock, D.D. Flake II, M.B. Warf, A.R. Hartman, R.J. Wenstrup, L.E. Clarke
Development of methodology: C. Rock, M.B. Warf, K. Rushton, A.R. Hartman, L.E. Clarke
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J.S. Ko, B. Matharoo-Ball, S.D. Billings, B.J. Thomson, J.Y. Tang, K.Y. Sarin, E. Cai, J. Kim, C. Rock, H.Z. Kimbrell, C. Miller, K. Rushton, R.J. Wenstrup, L.E. Clarke
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): J.S. Ko, J.Y. Tang, J. Kim, D.D. Flake II, M.B. Warf
Writing, review, and/or revision of the manuscript: J.S. Ko, B.J. Thomson, J.Y. Tang, J. Kim, C. Rock, H.Z. Kimbrell, D.D. Flake II, M.B. Warf, C. Miller, K. Rushton, A.R. Hartman, R.J. Wenstrup, L.E. Clarke

References

1. Shoo BA, Sagebiel RW, Kashani-Sabet M. Discordance in the histopathologic diagnosis of melanoma at a melanoma referral center. J Am Acad Dermatol 2010;62:751–6.
2. NCI. SEER Cancer Statistics Factsheets: Melanoma of the Skin. Rockville, MD: National Cancer Institute. Available from: http://seer.cancer.gov/statfacts/html/melan.html.
3. American Cancer Society. 2016 Survival rates for melanoma skin cancer, by stage. Atlanta, GA: American Cancer Society. Available from: http://www.cancer.org/cancer/skincancer-melanoma/detailedguide/melanoma-skin-cancer-survival-rates-by-stage.
4. Farmer ER, Gonin R, Hanna MP. Discordance in the histopathologic diagnosis of melanoma and melanocytic nevi between expert pathologists. Hum Pathol 1996;27:528–31.
5. McGinnis KS, Lessin SR, Elder DE, Guerry Dt, Schuchter L, Ming M, et al. Pathology review of cases presenting to a multidisciplinary pigmented lesion clinic. Arch Dermatol 2002;138:617–21.
6. Veenhuizen KC, De Wit PE, Mooi WJ, Scheffer E, Verbeek AL, Ruiter DJ. Quality assessment by expert opinion in melanoma pathology: experience of the pathology panel of the Dutch Melanoma Working Party. J Pathol 1997;182:266–72.
7. Cerroni L, Barnhill R, Elder D, Gottlieb G, Heenan P, Kutzner H, et al. Melanocytic tumors of uncertain malignant potential: results of a tutorial held at the XXIX Symposium of the International Society of Dermatopathology in Graz, October 2008. Am J Surg Pathol 2010; 34:314–26.
8. Gerami P, Busam K, Cochran A, Cook MG, Duncan LM, Elder DE, et al. Histomorphologic assessment and interobserver diagnostic reproducibility of atypical spitzoid melanocytic neoplasms with long-term follow-up. Am J Surg Pathol 2014;38:934–40.
9. Bastian BC, Olshen AB, LeBoit PE, Pinkel D. Classifying melanocytic tumors based on DNA copy number changes. Am J Pathol 2003;163:1765–70.
10. Gerami P, Jewell SS, Morrison LE, Blondin B, Schulz J, Ruffalo T, et al. Fluorescence *in situ* hybridization (FISH) as an ancillary diagnostic tool in the diagnosis of melanoma. Am J Surg Pathol 2009;33:1146–56.
11. Gerami P, Li G, Pouryazdanparast P, Blondin B, Beilfuss B, Slenk C, et al. A highly specific and discriminatory FISH assay for distinguishing between benign and malignant melanocytic neoplasms. Am J Surg Pathol 2012; 36:808–17.
12. Clarke LE, Warf BM, Flake DD2nd, Hartman AR, Tahan S, Shea CR, et al. Clinical validation of a gene expression signature that differentiates benign nevi from malignant melanoma. J Cutan Pathol 2015;42:244–52.
13. Clarke L, II DDF, Busam K, Cockerell C, Helm K, McNiff J, et al. An independent validation of a gene expression signature to differentiate malignant melanoma from benign melanocytic nevi. Cancer 2017;123: 617–28.
14. U.S. Food and Drug Administration. Guidance for Industry and FDA staff: statistical guidance on reporting results from studies evaluating diagnostic tests. Silver Spring, MD: U.S. Food and Drug Administration.
15. Glusac EJ. The melanoma 'epidemic', a dermatopathologist's perspective. J Cutan Pathol 2011;38:264–7.
16. Egnatios GL, Ferringer TC. Clinical follow-up of atypical spitzoid tumors analyzed by fluorescence in situ hybridization. Am J Dermatopathol 2016;38:289–96.
17. Fang Y, Dusza S, Jhanwar S, Busam KJ. Fluorescence in situ hybridization (FISH) analysis of melanocytic nevi and melanomas: sensitivity, specificity, and lack of association with sentinel node status. Int J Surg Pathol 2012;20:434–40.
18. Gaiser T, Kutzner H, Palmedo G, Siegelin MD, Wiesner T, Bruckner T, et al. Classifying ambiguous melanocytic lesions with FISH and correlation with clinical long-term follow up. Mod Pathol 2010;23:413–9.
19. Gerami P, Cooper C, Bajaj S, Wagner A, Fullen D, Busam K, et al. Outcomes of atypical spitz tumors with chromosomal copy number aberrations and conventional melanomas in children. Am J Surg Pathol 2013;37:1387–94.
20. Gerami P, Scolyer RA, Xu X, Elder DE, Abraham RM, Fullen D, et al. Risk assessment for atypical spitzoid melanocytic neoplasms using FISH to identify chromosomal copy number aberrations. Am J Surg Pathol 2013;37:676–84.
21. Minca E, Clarke L, Flake D, Ko J, Billings S. Correlation of myPath melanoma gene expression score with clinical outcome on a series of melanocytic lesions. In: Proceedings of the American Society of Dermatopathology 52nd Annual Meeting; 2015 Oct 8–11;. San Francisco, CA. Oakbrook Terrace, IL: ASDP; 2015.
22. North JP, Vetto JT, Murali R, White KP, White CR, Bastian BC. Assessment of copy number status of chromosomes 6 and 11 by FISH provides independent prognostic information in primary melanoma. Am J Surg Pathol 2011;35:1146–50.
23. Raskin L, Ludgate M, Iyer RK, Ackley TE, Bradford CR, Johnson TM, et al. Copy number variations and clinical outcome in atypical spitz tumors. Am J Surg Pathol 2011;35:243–52.
24. Vergier B, Prochazkova-Carlotti M, de la Fouchardiere A, Cerroni L, Massi D, De Giorgi V, et al. Fluorescence *in situ* hybridization, a diagnostic aid in ambiguous melanocytic tumors: European study of 113 cases. Mod Pathol 2011;24:613–23.
25. Warf MB, Flake DD, Adams D, Gutin A, Kolquist KA, Wenstrup RJ, et al. Analytical validation of a melanoma diagnostic gene signature using formalin-fixed paraffin-embedded melanocytic lesions. Biomark Med 2015;27:1–10.
26. Minca EC, Al-Rohil RN, Wang M, Harms PW, Ko JS, Collie AM, et al. Comparison between melanoma gene expression score and fluorescence in situ hybridization for the classification of melanocytic lesions. Mod Pathol 2016;29:832–43.
27. Criscione VD, Weinstock MA. Melanoma thickness trends in the United States, 1988–2006. J Invest Dermatol 2010;130:793–7.
28. 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 USA 2005;102:6092–7.
29. Jaeger J, Koczan D, Thiesen HJ, Ibrahim SM, Gross G, Spang R, et al. Gene expression signatures for tumor progression, tumor subtype, and tumor thickness in laser-microdissected melanoma tissues. Clin Cancer Res 2007;13:806–15.
30. Smith AP, Hoek K, Becker D. Whole-genome expression profiling of the melanoma progression pathway reveals marked molecular differences between nevi/melanoma in situ and advanced-stage melanomas. Cancer Biol Ther 2005;4:1018–29.
31. Blokhin E, Pulitzer M, Busam KJ. Immunohistochemical expression of p16 in desmoplastic melanoma. J Cutan Pathol 2013;40:796–800.
32. Gerami P, Beilfuss B, Haghighat Z, Fang Y, Jhanwar S, Busam KJ. Fluorescence in situ hybridization as an ancillary method for the distinction of desmoplastic melanomas from sclerosing melanocytic nevi. J Cutan Pathol 2011;38:329–34.
33. Wiesner T, Kiuru M, Scott SN, Arcila M, Halpern AC, Hollmann T, et al. NF1 mutations are common in desmoplastic melanoma. Am J Surg Pathol 2015;39:1357–62.

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

Diagnostic Distinction of Malignant Melanoma and Benign Nevi by a Gene Expression Signature and Correlation to Clinical Outcomes

Jennifer S. Ko, Balwir Matharoo-Ball, Steven D. Billings, et al.

Updated version	Access the most recent version of this article at: doi:10.1158/1055-9965.EPI-16-0958
Supplementary Material	Access the most recent supplemental material at: http://cebp.aacrjournals.org/content/suppl/2017/04/04/1055-9965.EPI-16-0958.DC1

Cited articles	This article cites 29 articles, 2 of which you can access for free at: http://cebp.aacrjournals.org/content/26/7/1107.full#ref-list-1

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, use this link http://cebp.aacrjournals.org/content/26/7/1107. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.