BMC Medical Genomics

RESEARCH

Open Access

RNA-seq validation of microRNA expression signatures for precision melanoma diagnosis and prognostic stratification

Christopher G. Love[1*], Lauren Coombs[2] and Ryan Van Laar[1,2]

Abstract

Background New diagnostic tools are needed to improve the diagnosis and risk stratification of cutaneous melanoma. Disease-specific microRNA signatures have been previously described via NanoString profiling of solid biopsy tissue and plasma. This study validated these signatures via next-generation sequencing technology and compared their performance against clinical metrics and other published melanoma signatures.

Methods RNA from 64 plasma and 60 FFPE biopsy samples from individuals with invasive melanoma or related benign/control phenotypes was extracted and enriched for microRNA. RNA sequencing was performed to compute MEL38/MEL12 signature scores. The results were evaluated with published NanoString and RNA sequencing datasets, comprising 548 solid tissue samples and 217 plasma samples, to predict disease status and patient outcome.

Results The MEL38 diagnostic signature classifies patients into discrete diagnostic groups via RNA sequencing in either solid tissue or plasma ($P < 0.001$). In solid tissue, the prognostic MEL12 signature stratifies patients into low-, intermediate- and high-risk groups, independent of clinical covariates. The hazard ratios for 10-year overall survival, based on observed survival intervals, were 2.2 (MEL12 high-risk vs low-risk, $P < 0.001$) and 1.8 (intermediate-risk vs low-risk, $P < 0.001$), outperforming other published prognostic models. MEL12 also exhibited prognostic significance in the plasma of 42 patients with invasive disease.

Conclusions The MEL38 and MEL12 signatures can be assessed in either solid tissue or plasma using RNA-seq and are strong predictors of disease state and patient outcome. Compared with other genomic methods, MEL12 was shown to be the strongest predictor of poor prognosis. MicroRNA expression profiling offers objective, accurate genomic information about a patient's likelihood of invasive melanoma and prognosis.

Keywords RNA-seq, Melanoma, MicroRNA signatures, Cancer, Precision medicine, Gene expression

Background

MicroRNAs play crucial roles in regulating gene expression and have emerged as a potential biomarker for the diagnosis and prognosis of early-stage melanoma [1–4]. Our previous work resulted in two microRNA signatures, MEL38 for diagnosis and MEL12 for prognosis, identified via NanoString nCounter microRNA expression profiling on both solid biopsy tissue and plasma samples representing the spectrum of normal skin, benign naevi, and stage 0 to IV invasive melanoma.

MEL38 mirrors the early molecular changes that occur during the transition from benign to malignant melanocytic lesions. This diagnostic signature comprises 38 miRNAs that are differentially expressed between benign nevi and invasive melanoma. These microRNAs regulate genes involved in cell proliferation, apoptosis, and

*Correspondence:
Christopher G. Love
chris.love@geneseq.com.au
[1] Geneseq Biosciences, Melbourne, VIC, Australia
[2] Australian Clinical Laboratories, Clayton, VIC, Australia

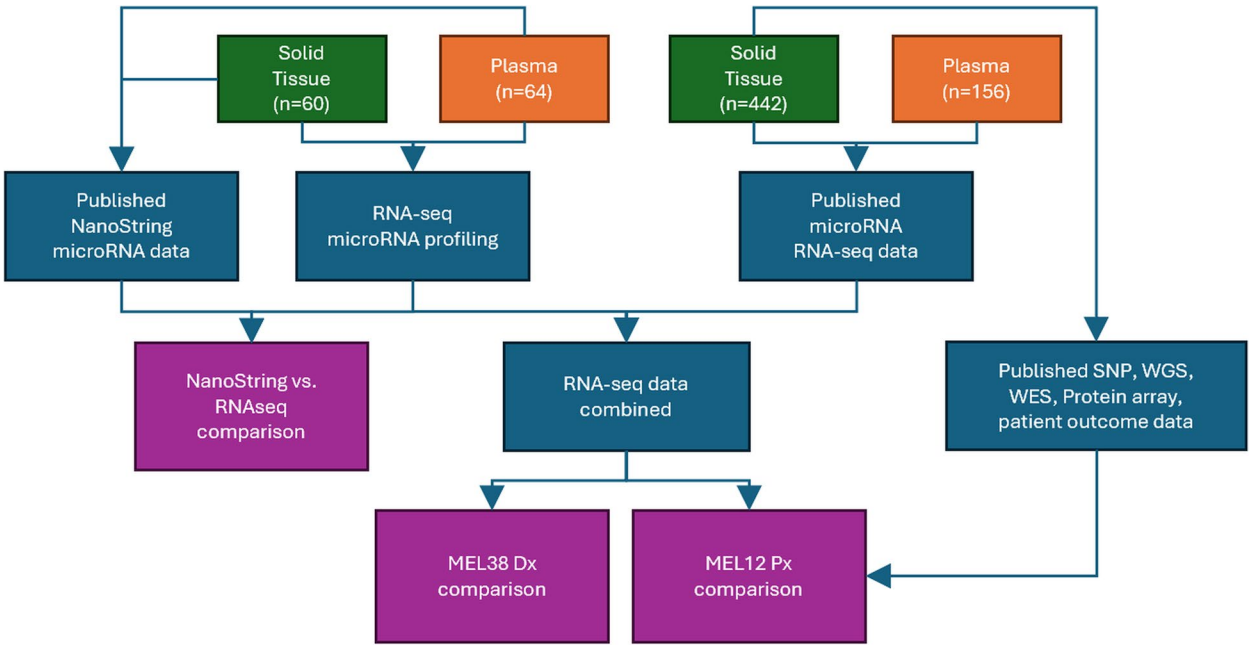

Fig. 1 Schematic diagram of the study design. The diagram demonstrates the samples, datasets and technologies used to validate RNA-seq based microRNA profiling of plasma and solid tissue for melanoma diagnosis and prognosis

migration, capturing the early genetic and epigenetic alterations that signify the onset of malignancy [2, 3]. In contrast, MEL12 represents miRNAs that influence advanced tumour behaviours and patient survival. This prognostic signature consists of 12 miRNAs associated with tumour progression, metastasis, and therapeutic resistance [5].

As a circulating biomarker, MEL38 can detect the presence of invasive melanoma at the systemic level, including stage IA disease, indicating a need for skin examination and additional risk-informed follow-up. When MEL38 is used as a cellular biomarker, microRNA extracted from an excised melanocytic lesion is used to measure the degree of malignancy present and reduce the risk of misdiagnosis associated with melanoma histopathology [6].

The companion MEL12 signature comprises microRNAs whose expression patterns are correlated with the risk of melanoma-specific death. Like MEL38, it was first identified as a circulating biomarker via plasma from patients with stage I-IV invasive disease; therefore, it represents a systemic assessment of a patient and a novel approach to identifying individuals most likely to benefit from invasive diagnostic procedures or clinical trial enrolment [5].

The NanoString nCounter system employs a colour-coded molecular barcoding system for the digital quantification of individual miRNAs. Its benefits include specific and sensitive detection of low-abundance microRNAs without the need for amplification. However, its small batch size of twelve samples per run limits its suitability for high-throughput applications [7, 8]. An alternative method, small-RNA next-generation sequencing (NGS or RNA-seq), involves generating cDNA libraries from miRNA molecules followed by high-throughput sequencing. Recent advances in NGS library preparation protocols have eliminated the processing bottleneck of size-separating agarose gels, increasing the feasibility of this method for high-throughput clinical applications [9].

The aim of this study is to validate the use of RNA-seq for evaluating the MEL38 and MEL12 microRNA signatures of melanoma in solid and liquid patient samples, as summarised in Fig. 1. A secondary objective is to further evaluate the performance of these signatures in publicly available RNA-seq datasets, encompassing both solid tissue and plasma profiles from individuals with or without invasive melanoma. The largest of these datasets is from the SKCM (Skin Cutaneous Melanoma) project, which is part of The Cancer Genome Atlas (TCGA), from which other genomic models of melanoma have been developed from or validated, enabling direct comparison with MEL38/12 [10].

The incorporation of standardised microRNA expression profiling into the melanoma care path has the potential to increase diagnostic accuracy by providing personalised and objective information about melanoma development and progression at both the systemic and local level.

Methods

Specimen selection and study design

Sample size calculations were conducted to ensure adequate power for the detection of an area under the curve (AUC) ≥ 0.78 with a small-to-moderate effect size, 80% power and 95% confidence.

The study began with a set of 60 patient samples for FFPE profiling, which were selected from a previously published series of patient samples on the basis of remaining tissue availability [11]. Additionally, 64 previously analysed plasma samples with sufficient remaining sample volume were identified and used for RNA-seq profiling [5]. A summary of the clinicopathological variables for these samples is presented in Table 1.

RNA extraction and quality control

RNA was extracted via the Qiagen miRNeasy FFPE Kit (Qiagen, Germany, Part: 217,504) or the Qiagen miRNeasy Serum/Plasma Advanced Kit (Part: 217,204), as previously described [11, 12]. The extracted RNA from the plasma was further purified via Amicon Ultra 0.5 Centrifugal filter columns (MilliporeSigma, Germany, Part UFC501096) for 80 min at 10,000 g. Purified RNA concentrations were determined with the Invitrogen

Table 1 Clinicopathological characteristics of patient specimens used for protocol optimisation and validation of melanoma microRNA signatures

Solid tissue (FFPE) cohort (*N* = 60)	**Count**	**%**	**Plasma cohort (*N* = 64)**	**Count**	**%**
Age (mean, stdev)	67 (16)		Age (mean, stdev)	60 (17)	
Sex:			Sex:		
Male		34	Male	35	55
Female		26	Female	29	45
Disease status:			Disease status:		
Invasive melanoma (*n* = 28)			Invasive melanoma (*n* = 52)		
Clinical stage			*Clinical stage*		
Stage I	5	18	Stage I	4	8
Stage II	11	39	Stage II	7	13
Stage III	4	14	Stage III	27	52
Stage IV	8	29	Stage IV	14	27
Histological subtype:			*Histological subtype:*		
Epithelioid	1	4	Acral lentiginous	1	2
Nodular	5	18	Epithelioid / Spindle cell	1	2
Superficial spreading	10	36	Lentigo maligna	2	4
Melanoma NOS	12	43	Malignant melanoma, NOS	38	73
			Nodular	5	10
			Superficial spreading	5	10
Other (*n* = 32)					
Naevi	13	41			
Histological subtype:			Other (*n* = 12)		
Compound	5	38	Naevi	8	67
Dysplastic	1	8	*Histological subtype:*		
Intradermal	4	31	Actinic keratosis	2	25
Junctional	2	15	Naevi, NOS	1	13
Solar keratosis	1	8	Seborrheic keratoses	5	63
			Melanoma in-situ	4	33
Melanoma in-situ	19	59			
Histological subtype:			Specimen collection type:		
Melanoma in-situ NOS	6	32	Archival	51	80
Junctional	1	5	Prospective	13	20
Lentigo maligna	10	53			
Superficial spreading	2	11			

microRNA Qubit Assay (Thermo Fisher Scientific, USA). Part Q32880).

Small-RNA library preparation and sequencing

The Revvity NEXTFLEX Small RNA-Seq Kit v4 (Revvity, Inc., USA, Part: NOVA-5132–43), which is optimized for microRNA profiling and allows for multiplexing up to 384 samples, was used for library preparation. Libraries were prepared from 5 ng of small-RNA enriched total RNA, incorporating Unique Dual Indexes (UDIs) for sequencing on an Illumina instrument. The plasma samples included tRNA/YRNA blockers to enrich for microRNA content. Library concentrations were measured via an Agilent 5200 Fragment Analyzer (Agilent, Inc., USA) by Micromon Genomics (Monash University, Victoria, Australia). Libraries were normalised to 2 nM and pooled. Sequencing was performed on an Illumina MiSeq system (Illumina, Inc., USA) at loading concentration of 10 pM using the MiSeq Reagent Kit v3 (Part MS-102–3003) and version 4.1.0.471 of Illumina Local Run Manager software.

Workflow for miRNA identification and quantification via RNA-seq

FASTQ files were generated from Illumina BCL files via the GenerateFASTQ Analysis Module of the the Illumina Local Run Manager software (Illumina Inc., USA). Data from read 1 (R1), which includes full-length mature microRNAs, were used in subsequent steps. Quality control involved trimming bases from the 3' end with a minimum Phred quality score threshold of 20, removing sequencing adapters, and excluding reads under 10 bp. Samples producing fewer than 150,000 total raw reads were excluded from further analysis.

Samples that passed these data QC thresholds were aligned to miRBase (v22) via Bowtie (version 1.2.2) in sRNAbench library mode via full-length alignment and two mismatch parameters [13–15]. Reads aligning to mature miRNAs in the sense orientation were counted and normalized to counts per million (CPM). Additional sequence databases were used to identify other RNA classes present in the samples: the Ensembl genome GRCh38_p13 release 104 and tRNA sequences from GtRNAdb and RNAcentralDB v20.

Public RNA-seq validation datasets

Solid tissue RNA-seq data representing 452 invasive melanoma samples and 37 non melanoma skin samples (1 actinic keratosis, 1 squamous cell carcinoma, 2 psoriatic skin samples and 33 normal skin) were downloaded from the DIANA-Lab MicroRNA Tissue-Expression Database (DIANA-miTED) in CPM format [16]. Data corresponding to samples originally published by the TCGA SKCM consortium were annotated with histopathology stage and observed survival (OBS) interval information from Xiong et al. [16, 17]. The source studies for these samples and the sequencing methods used are provided in the Supplementary Information [Additional File 2].

Plasma RNA-seq data from 60 pre-treatment metastatic melanoma patients and 96 control (cancer-free) individuals were obtained from NCBI BioProject PRJNA634142. These data were originally generated via the HTG EdgeSeq miRNA Whole Transcriptome Assay and Illumina sequencing, as published by Bustos, Tran [18]. The raw FASTQ files from this study were processed to CPM values via sRNAbench [15].

MicroRNA signature calculations and statistical analyses

MicroRNA signature weights used to calculate MEL38/MEL12 scores were updated for the dynamic range of RNA-seq data by refitting the two machine learning algorithms as previously described, without feature reselection [3, 5]. Updated MEL38 weights were calculated using a linear kernel SVM with default penalty cost [19]. Updated MEL12 weights were calculated using principal component analysis, using the top two components [20, 21]. Individual sample scores were then calculated using the formula:

$$score = \sum_{i=0}^{n} \omega_i \cdot x_i$$

where:

$$n = number\ of\ microRNAs\ in\ signature\ (MEL12\ or\ MEL38)$$

$$\omega = classification\ weight\ for\ i-th\ microRNA$$

$$x = normalised\ expression\ level\ of\ i-th\ microRNA$$

Statistical analyses were performed using R Statistical Software (v4.1.2; R Core Team 2021), Microsoft Excel (Version 2404, Microsoft, Inc., WA, USA), and MedCalc (version 22.023). Student's t-test was used to compare means across groups for continuous variables. Intra-Class Correlation (ICC) coefficient and their 95% confidence intervals were calculated for average measures [22]. Receiver Operating Characteristic (ROC) analyses were used to assess the association between microRNA scores and disease status, with Area Under the Curve (AUC) calculated as a measure of diagnostic accuracy. Cox proportional hazards regression was used to evaluate the association of MEL12 with patient outcome including available clinical covariates [23]. All *P*-value calculations were two-sided, and values less than 0.05 were considered statistically significant.

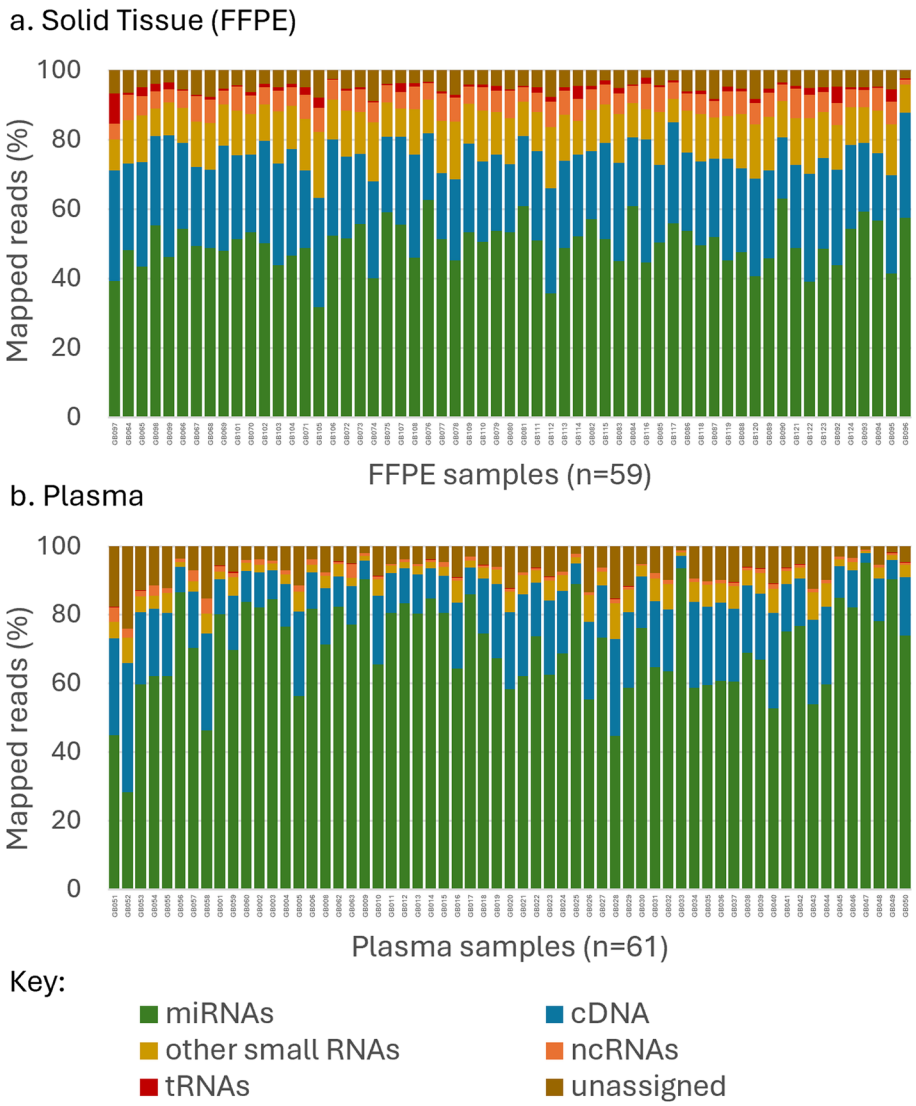

Fig. 2 RNA-seq profiles from extracted total RNA demonstrating enrichment of microRNAs relative to other RNA species. **a** Macro-dissected FFPE tissue sections. **b** Plasma samples

Results

Quality control and microRNA quantification using RNA-seq

To evaluate the performance of the MEL38 and MEL12 signatures using RNA-seq technology, we performed small-RNA extraction and purification on 60 FFPE and 64 plasma samples previously analysed via the NanoString platform. The mean concentration of small-RNA enriched total RNA from the FFPE samples was 40 ng/µl (range 3 to 146) and 18 ng/ul (range 1 to 100) from plasma. The RNA-seq libraries were prepared using the NextFlex Small RNA Kit, and Illumina sequencing resulted in a mean of 521,699 reads per sample, with no significant difference in read numbers between specimen types (*P*-value: 0.36).

After raw data quality control, 59 FFPE and 61 plasma RNA-seq profiles were available for alignment and melanoma signature calculation. Fifty percent of FFPE and 71% of plasma reads were mapped to mature microRNAs in miRbase, as shown in Fig. 2. With a CPM threshold of > 50, the mean number of expressed microRNAs in plasma was 1304, whereas it was 721 for FFPE samples (*P*-value for difference < 0.001). The increased detection of mature microRNA sequences observed in plasma samples can be attributed to the Amicon column concentration step performed during plasma processing or to a

reduced presence of other small RNA types in acellular versus cellular tissues.

Additional data processing results are provided in the Supplementary Information [Additional File 1].

Visualisation of MEL38 expression profiles

Count-per-million (CPM) normalised RNA-seq MEL38 expression profiles were extracted from each sample's complete profile, averaged per disease stage, and compared against previously published NanoString results using colour-coded heatmaps, as shown in Fig. 3. Both technologies capture the increasing or decreasing levels of the microRNAs in the signature between varying stages of melanoma progression. ICC coefficients of MEL38 scores for NanoString versus RNA-seq platforms showed good agreement with 0.76 (95% CI: 0.61–0.86) for plasma and 0.97 (95% CI:0.94–0.98) for FFPE.

Calculation of MEL38 scores using RNA-seq data and comparison to NanoString results

To calculate MEL38 scores for samples profiled using RNA-seq, SVM weights were recalculated for RNA-seq scale data, as described above. RNA-seq MEL38 scores were then calculated for each sample and compared to diagnostic classifications using ROC analysis, which resulted in an AUC of 1.0 for FFPE samples and 0.99 for plasma. No significant differences were observed between AUCs of RNA-seq vs NanoString datasets using the Hanley and McNeil method [24] (Fig. 4).

In the FFPE results, variance analysis within the invasive melanoma subset (i.e. Stage I to IV) revealed significant differences in MEL38 scores between early (Stage I-II, mean 4.7) and advanced stages (Stage III-IV, mean 9.7), as shown in Fig. 5a and, which is consistent with previously published results [11].

Validation of microRNA signatures using publicly available melanoma RNA-seq data

FFPE/Solid tissue

The solid tissue RNA-seq profiles generated for this study were combined with publicly available solid tissue RNA-seq profiles, resulting in a dataset containing 35 normal/psoriatic skin samples, 12 nevi, 19 melanoma samples in-situ, 11 locally invasive melanomas and 365 metastatic melanomas (total N = 442, Additional File 2). MEL38 scores were calculated for each sample as described and

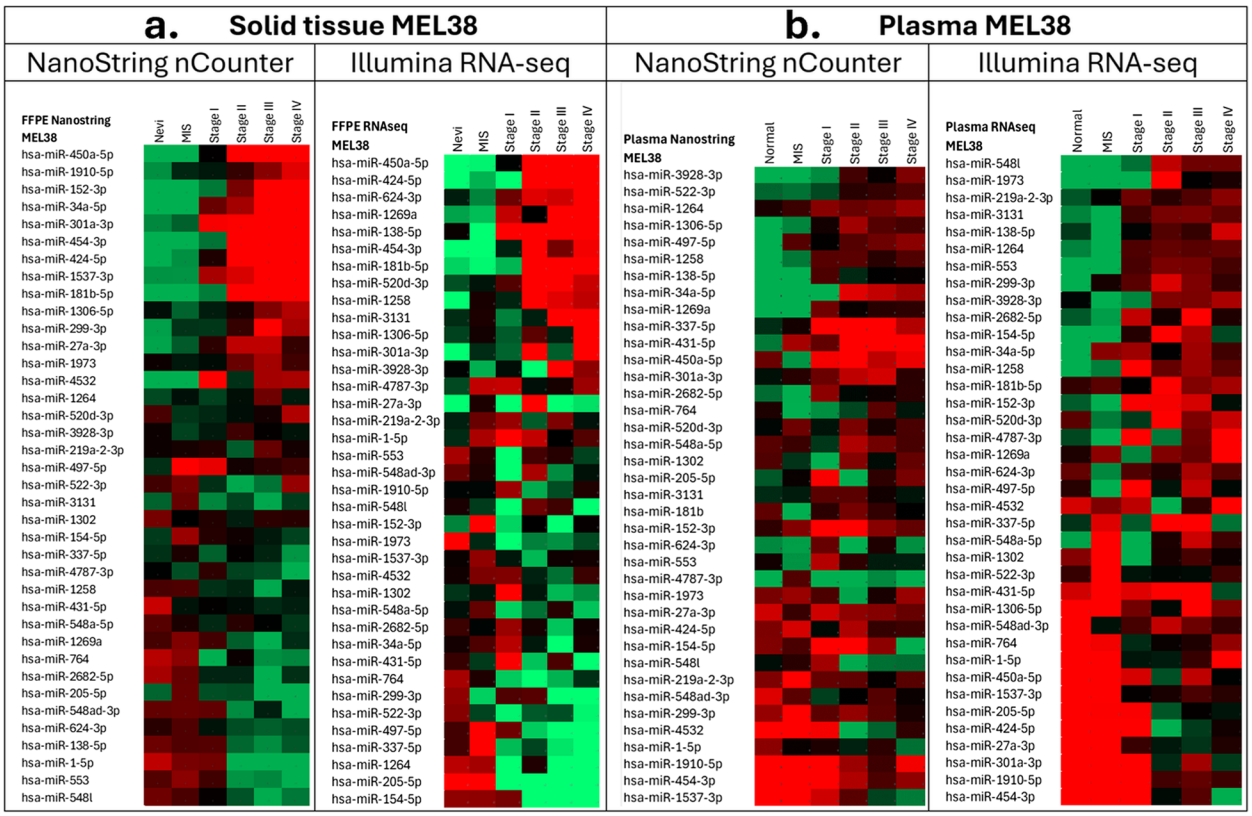

Fig. 3 MEL38 microRNA expression profiles of plasma samples and solid tissue samples grouped by melanoma stage. The relative differences in microRNA expression levels between clinically relevant disease states can be observed in both specimen types and analytical platforms. Dataset includes 59 plasma samples and 61 solid tissue samples. Red: high expression, Green: Low expression

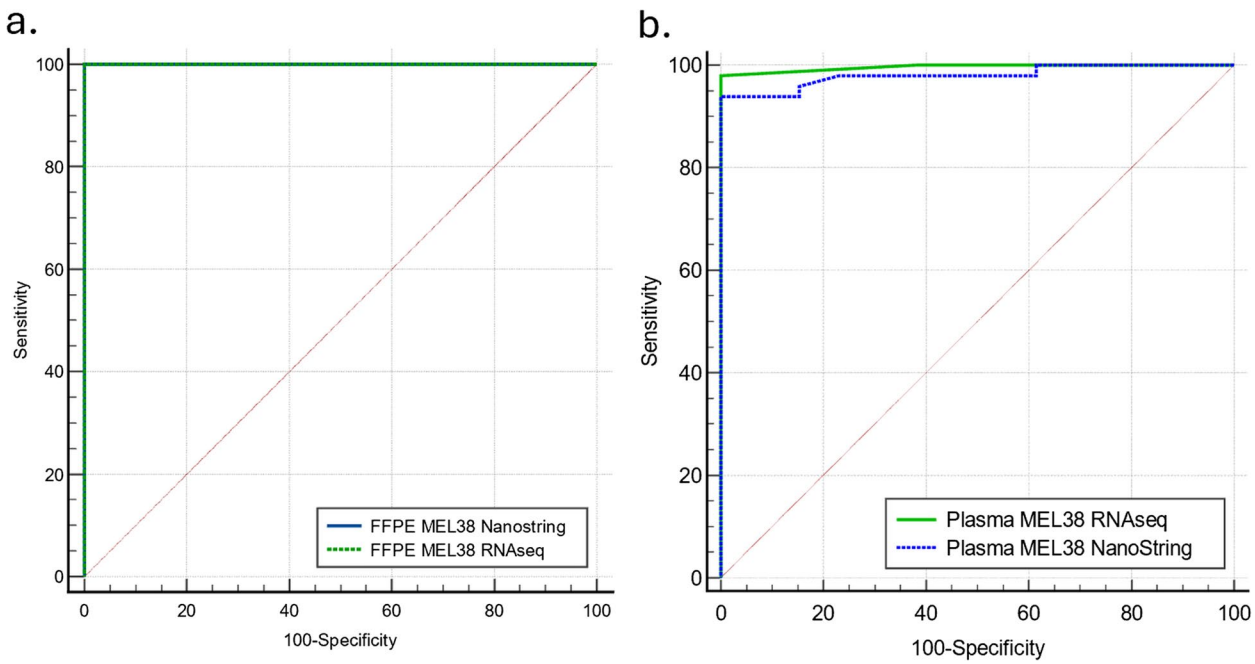

Fig. 4 ROC comparisons between methods and specimen types. **a** ROC curve comparison for 59 FFPE samples analysed with both RNA-seq and NanoString profiling. AUC for both methods: 1.0 ($P < 0.001$). **b** ROC comparison of MEL38 scores for 61 plasma samples profiled using both RNA-seq and NanoString. AUC for NanoString scores: 0.98 ($P < 0.001$), RNA-seq: 0.99 ($P < 0.001$). Pairwise comparison of curves; $P = 0.32$ (not significantly different)

plotted by melanoma stage, revealing complete separation of invasive melanoma and non-invasive melanoma samples (including melanoma in-situ/Stage 0) ($P < 0.001$, Fig. 5a). Within the invasive melanoma subset, MEL38 scores of metastatic melanoma samples (Stage III/IV) were on average threefold higher than locally invasive (I/II) melanoma samples ($P < 0.001$).

In the non-melanoma/other subset, the MEL38 scores of benign nevi and melanoma in-situ samples were on average 1.5 fold higher than normal skin samples ($P = 0.001$), but the difference between nevi and in-situ was not statistically significant. These data provide additional validation of the ability of the MEL38 signature to differentiate between clinically important stages of melanoma development and progression using RNA-seq profiling.

Next, the ability of the MEL12 signature to predict patient outcome using RNA-seq profiling of solid tissue biopsies was evaluated using 256 samples from the TCGA SKCM cohort for which survival data were available. Patients were classified as low, intermediate or high risk on the basis of MEL12 score tertiles, and the difference in rates of death between each group was compared using Kaplan Meier analysis (Fig. 5b). Log rank testing showed the difference in survival between the risk groups to be statistically significant ($P = 0.002$). Hazard ratios indicate a 1.7-fold higher rate of death for MEL12 intermediate- vs low-risk patients (95% CI: 1.2 to 2.5) and a 2.3-fold higher rate for patients classified as MEL12 high risk vs intermediate risk (95% CI: 1.60 to 3.31).

A multivariate Cox proportional hazard (CPH) regression analysis was performed to assess the significance of MEL12 independent to patient age, sex, and AJCC melanoma stage. As shown in Table 2, MEL12 stratification retained its statistical significance, suggesting the signature provides additional prognostic information beyond conventional clinicopathological variables.

Plasma

The 61 plasma RNA-seq profiles generated for this study were combined with 156 profiles from previously published RNA-seq datasets, resulting in a dataset consisting of 217 circulating microRNA profiles from individuals with or without invasive melanoma. Plasma-MEL38 scores were calculated, scaled, and evaluated against the clinical diagnosis of each sample. As depicted in Fig. 5c, the Plasma-MEL38 score perfectly discriminates between individuals with or without invasive melanoma ($P < 0.001$). The mean MEL38 score of normal/MIS samples was 3.8, whereas it was 6.6 for locally invasive melanoma and 7.7 for metastatic melanoma.

We then assessed the prognostic capability of the Plasma-MEL12 signature on the 40 RNA-seq profiles

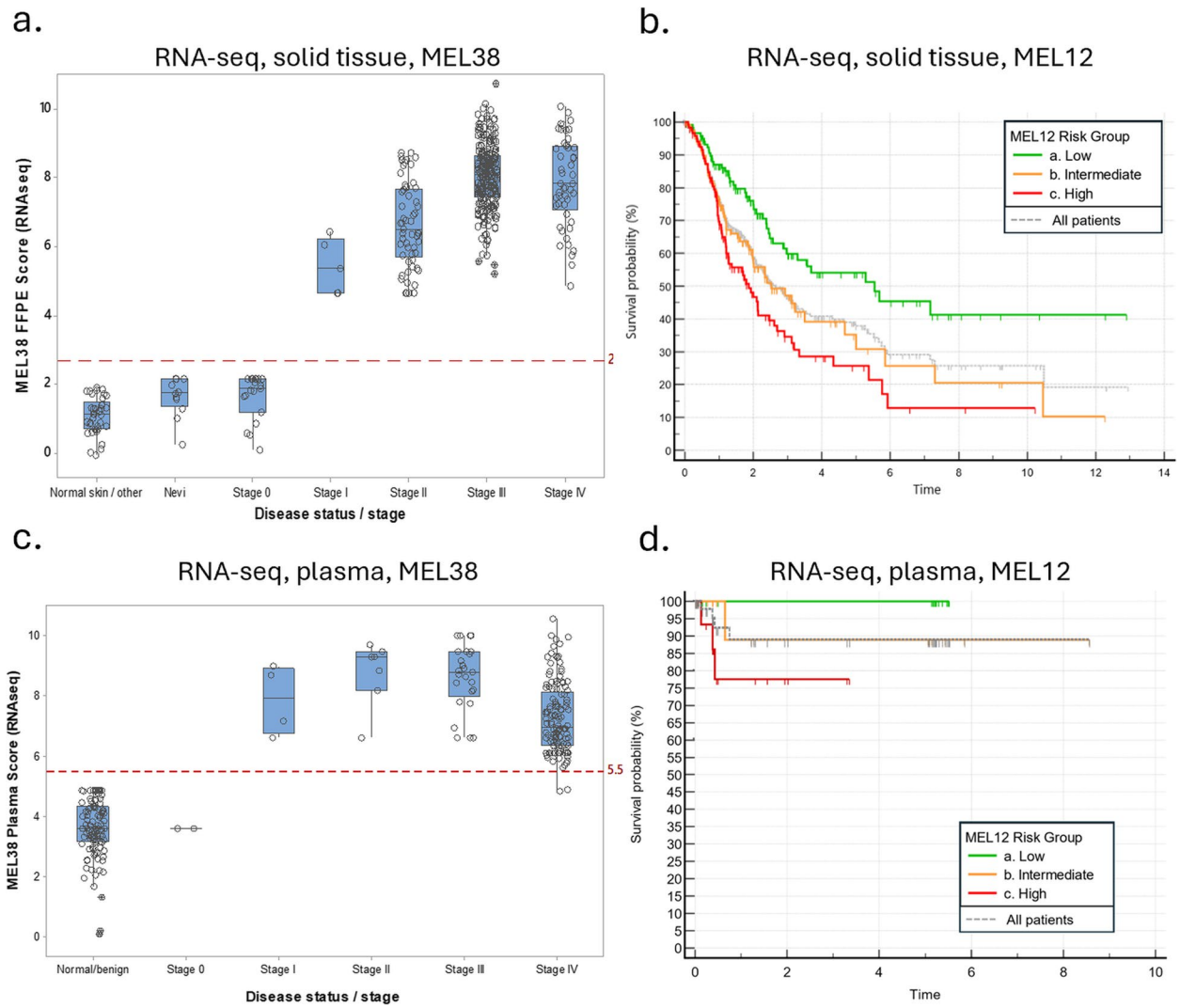

Fig. 5 Signature validation with public datasets. RNA sequencing (RNA-seq) data analysing MEL38 and MEL12 scores in 548 solid tissue and 217 plasma samples. **a** Box plot displaying MEL38 scores from RNA-seq analyses of solid tissue biopsies, categorized by disease status or melanoma stage. The dashed line depicts the MEL38 score threshold between invasive and non-invasive melanoma. **b** Kaplan–Meier survival analysis of melanoma patients grouped by solid tissue MEL12 scores (OBS). Log-rank test ($P=0.002$). **c** Box plot showing MEL38 scores derived from RNA-seq analysis of plasma samples, grouped by disease status. **d** Kaplan–Meier survival analysis of invasive melanoma patients grouped by plasma MEL12 expression levels (MSS). Univariate Cox proportional hazards regression, *P*-value of 0.034. The log-rank test shows a *P*-value of 0.12

Table 2 Cox proportional hazards regression analysis for 256 TCGA-melanoma patients (Stage I-IV) including solid tissue MEL12 risk groups (OBS)

Covariate	*P*	Hazard Ratio	95% CI
Age	0.02	1.01	1.00 to 1.02
Gender = "male"	0.64	0.93	0.67 to 1.28
Melanoma stage (I-IV)	< 0.0001	2.48	1.80 to 3.41
MEL12 = "b. intermediate"	0.004	1.81	1.21 to 2.69
MEL12 = "c. high"	0.0001	2.16	1.45 to 3.20

with available melanoma-specific survival information. MEL12 scores were calculated as described and used to assign patients to low-, intermediate- or high-risk groups based on cohort-specific tertiles. Kaplan–Meier analysis and log rank testing showed differences in survival approaching significance ($P=0.12$, Fig. 5d). Notably, no deaths occurred among patients classified as low risk according to the Plasma-MEL12 signature, highlighting its potential prognostic value.

Benchmarking of MEL12 against other genomic models of invasive melanoma

The TCGA SKCM consortium applied a comprehensive array of genomic technologies to a standardised set of clinically annotated patient samples, including single nucleotide polymorphism (SNP) arrays, DNA whole-genome sequencing (WGS), whole-exome sequencing (WES), protein expression arrays and microRNA sequencing. Methods to stratify patients into groups on the basis of the output of each technology were also described, including the mutation subtypes identified by WGS (BRAF, RAS, NF1, and Triple-WT). Classifications of patients in the SKCM cohort using the novel 121 mRNA signature by Garg et al. ('Cam_121') was also compiled [25].

To compare the prognostic significance of MEL12 against other published algorithms and genomic platforms, CPH models were calculated for each method, using a consistent set of clinical covariates including age, sex, clinical stage, and multiple genomic/transcriptomic classifications, as published by The Cancer Genome Atlas Network [10]. The observed survival data were used as the endpoint in each analysis. Hazard ratios and 95% confidence intervals were compared using a forest plot, as shown in Fig. 6. Individual CPH model results are provided in the Supplementary Information [Additional File 3].

Ranking of classification methods by covariate-adjusted hazard ratios revealed that the high-risk MEL12 group and the 'keratin' RNA-seq group to have largest values (HR: 2.2, significance of variable in model $P<0.001$). Importantly, this figure surpassed the classification based solely on clinical staging (HR: 1.9, $P<0.001$), even after adjusting for age and sex. This indicates that incorporating genomic classification methods into prognostic evaluations is likely to enhance the precision of clinical risk stratification.

In these data, the ability of MEL12 to predict patient outcome was superior to that of the Cam_121 signature (hazard ratio 2.2 vs 1.6), which Garg et al. described as predicting metastasis better than both clinical covariates and other prognostic signatures available at the time of publication [25]. By performing specimen-matched, clinical-covariate adjusted, comparisons of various approaches to patient stratification, the findings show that microRNA expression profiling is equivalent—or superior to—protein, messenger RNA or whole-genome methods for prediction of patient outcome.

Discussion

Melanoma is a multifaceted disease driven by both localised genomic alterations within melanocytes and systemic interactions with the immune system and tumour microenvironment. Our validation of the MEL38 and MEL12 microRNA signatures using RNA-seq technology reinforces the notion that melanoma development and progression are influenced by a complex interplay of local and systemic factors [1, 2].

The MEL38 diagnostic signature robustly classifies patients into discrete diagnostic groups in both solid

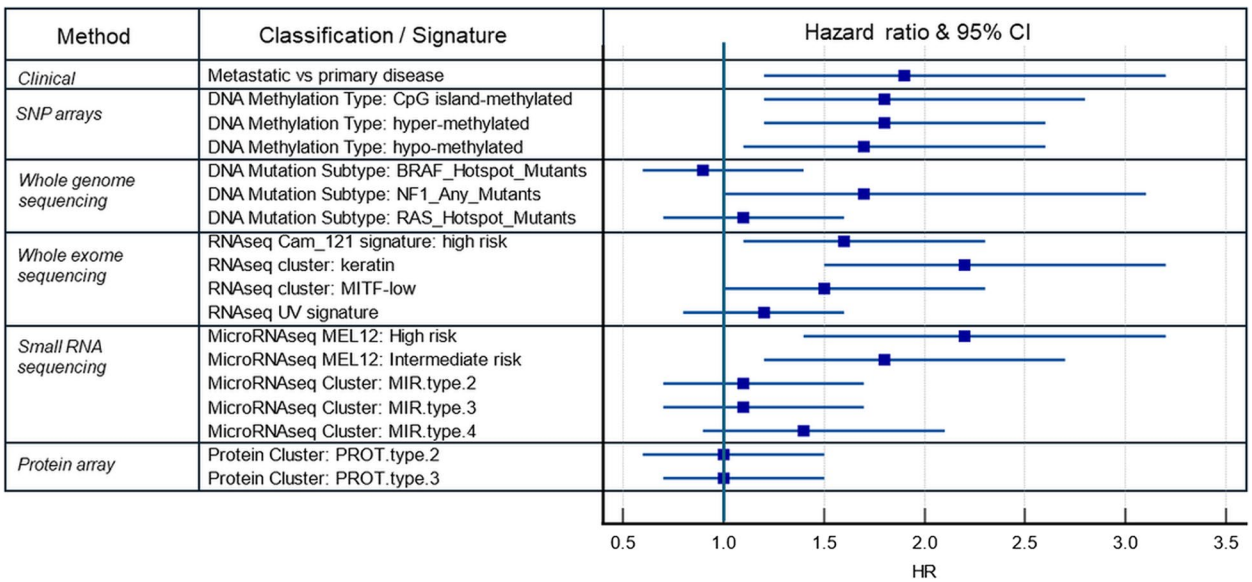

Fig. 6 Comparative prognostic stratification of TCGA SCKM patients. Forest plot of Cox proportional hazard ratio values for the prediction of overall survival risk, using DNA, mRNA, microRNA sequencing or protein array-based classification of patients in the TCGA SKCM cohort ($n=354$). Each CPH model included age, sex, stage and signature classification class. The keratin RNA-seq cluster (vs immune RNA-seq cluster) and the high-risk MEL12 signature (vs the low-risk group) had the highest hazard ratios (2.2, $P<0.001$), reflecting their robust prognostic ability and independence of clinical covariates included in the analysis

tissue and plasma, while the prognostic MEL12 signature stratifies patients into risk categories based on their survival outcomes. These findings align with previous research demonstrating that melanoma progression is influenced by genetic mutations, epigenetic changes, and immune responses. Importantly, our study revealed that these microRNA signatures outperform traditional clinical metrics and other genomic models in predicting disease status and patient outcomes, thereby offering a more comprehensive understanding of melanoma biology and its clinical implications.

Next-generation sequencing methods have traditionally been used for biomarker discovery and validation, while qPCR or microarray-based methods have been better suited for clinical translation owing to their cost and throughput advantages. Advances in library preparation and automation now present an opportunity to maintain the sensitivity of sequencing in a clinical setting, including for small RNA-based assays. The introduction of magnetic bead-based size selection for RNA sequencing libraries has streamlined a previously time-consuming stage of the NGS workflow, while also enhancing sample purity by removing primer dimers and other contaminants. RNA-seq allows the inclusion of up to 384 samples in a single run, offering a level of scalability necessary for high-throughput clinical use.

RNA-seq offers several practical advantages for microRNA profiling in predicting melanoma patient survival. microRNAs are more stable in both tissue and liquid biopsies compared to mRNA, allowing for more reliable detection [26]. They are also measurable with greater sensitivity than protein assays, due to an amplification step [27]. RNA-seq's broad dynamic range further facilitates consistent fold change calculations relative to array platforms [28]. However, RNA-seq for microRNA profiling can have higher initial set-up costs relative to qPCR and require more advanced data handling and bioinformatics pipelines, due to size and complexity of the raw data generated.

However, limitations to this work include the modest size of the plasma MEL12 RNA-seq validation set, due to the limited availability of plasma samples from melanoma patients with corresponding outcome data. Furthermore, the availability of overall survival (OS) rather than melanoma-specific survival (MSS) for patients in the TCGA SKCM cohort prevents censoring or excluding data from patients who died of non-melanoma related causes. Future research should focus on integrating these microRNA profiles into routine clinical practice to enhance diagnostic accuracy and enable risk-based treatment strategies.

This study is unique in that we compared the MEL12 signature to genomic signatures generated using a range of methods and molecule types, including DNA-, mRNA-, and protein-based models, enabling direct and objective assessment of model accuracy. In the multivariate-adjusted prediction of survival, MEL12 was strongly correlated with OBS, notably outperforming the Cam_121 messenger-RNA signature and other genomic models published by the SKCM consortium. The design of the MEL12 signature likely contributes to its prognostic superiority, as the individual microRNAs in the signature were selected for their correlation with length of melanoma-specific survival. This contrasts with Cam_121, which was identified using time to development of distant metastasis, a clinical endpoint with a higher degree of subjectivity and measurement variation than MSS [6].

Conclusions

These data build on previous work demonstrating that the ability of miRNA expression profiling of either solid biopsy tissue or plasma can identify the presence of invasive melanoma at a local or systemic level. Both MEL38 and MEL12 were first identified as circulating biomarkers of disease status, and these data further strengthen the hypothesis that a patient's circulating microRNA profile reflects critical events in melanoma development and progression. Validation of the methods using RNA sequencing techniques is likely to enable greater clinical translation of these methods, improving diagnostic accuracy and facilitating risk-based treatment strategies.

Abbreviations

AUC	Area Under the Curve
cDNA	Complementary DNA
CI	Confidence Interval
CPH	Cox Proportional Hazards
CPM	Counts per Million
FFPE	Formalin-Fixed, Paraffin-Embedded
ICC	Intra-Class Correlation
MEL12	MEL12 melanoma prognostic signature
MEL38	MEL38 melanoma diagnostic signature
miRNA	MicroRNA
MIS	Melanoma In situ
MSS	Melanoma specific survival
NGS	Next-generation sequencing
OBS	Observed Survival Interval
OS	Overall survival
RNA	Ribonucleic acid
RNA-seq	NGS RNA sequencing
ROC	Receiver Operating Characteristic
SKCM	Skin Cutaneous Melanoma
SNP	Single Nucleotide Polymorphism
SVM	Support Vector Machine
TCGA	The Cancer Genome Atlas
UDI	Unique Dual Indexes
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12920-024-02028-w.

Additional file 1.

Additional file 2.

Additional file 3.

Additional file 4.

Acknowledgements
The authors acknowledge use of the services and facilities of Micromon Genomics at Monash University.

Authors' contributions
LC, RVL were involved in study design. LC, CGL, RVL all contributed to data acquisition, data interpretation and data analysis. All authors contributed to writing the original draft and approved the final manuscript.

Funding
This research received no external funding.

Data availability
The RNA-seq dataset generated during the current study are available in the NCBI Gene Expression Omnibus repository (GEO) under accession GSE266859, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE266859. The NanoString datasets analysed during the current study are available from the corresponding author on reasonable request. Other datasets analysed during the study are available from the microRNA Tissue Expression Database, DIANA miTED (https://dianalab.e-ce.uth.gr/mited).

Declarations

Ethics approval and consent to participate
Use of archival plasma and FFPE tissue was approved by Australian Clinical Labs internal medical advisory board and satisfies criteria for use of human tissue by diagnostic pathology companies as outlined by the Australian Government's National Health and Medical Research Council [29]. Plasma specimen collection and use were approved by the Bellberry Human Research and Ethics Committee (TGA code EC00450, Application 2019–03-282-A-2), with informed consent obtained for prospectively collected materials.

Consent for publication
Not applicable.

Competing interests
CGL and RVL are employed and have equity in Geneseq Biosciences. LC is an employee of Australian Clinical Labs.

Received: 30 July 2024 Accepted: 11 October 2024
Published online: 25 October 2024

References

1. Fattore L, Costantini S, Malpicci D, Ruggiero CF, Ascierto PA, Croce CM, et al. MicroRNAs in melanoma development and resistance to target therapy. Oncotarget. 2017;8(13):22262–78.
2. Van Laar R, Lincoln M, Fereday S. Characterisation and validation of Mel38; A multi-tissue microRNA signature of cutaneous melanoma. PLoS One. 2019;14(2): e0211504.
3. Van Laar R, Lincoln M, Van Laar B. Development and validation of a plasma-based melanoma biomarker suitable for clinical use. Br J Cancer. 2018;118(6):857–66.
4. Stark MS, Tyagi S, Nancarrow DJ, Boyle GM, Cook AL, Whiteman DC, et al. Characterization of the Melanoma miRNAome by Deep Sequencing. PLoS One. 2010;5(3): e9685.
5. Van Laar R, Latif B, King S, Love C, Taubenheim N, Kalansooriya E, et al. Validation of a microRNA liquid biopsy assay for diagnosis and risk stratification of invasive cutaneous melanoma. Br J Dermatol. 2023;189(3):292–301.
6. Elmore JG, Barnhill RL, Elder DE, Longton GM, Pepe MS, Reisch LM, et al. Pathologists' diagnosis of invasive melanoma and melanocytic proliferations: observer accuracy and reproducibility study. BMJ. 2017;357: j2813.
7. Kolbert CP, Feddersen RM, Rakhshan F, Grill DE, Simon G, Middha S, et al. Multi-platform analysis of microRNA expression measurements in RNA from fresh frozen and FFPE tissues. PLoS One. 2013;8(1): e52517.
8. Chatterjee A, Leichter AL, Fan V, Tsai P, Purcell RV, Sullivan MJ, et al. A cross comparison of technologies for the detection of microRNAs in clinical FFPE samples of hepatoblastoma patients. Sci Rep. 2015;5: 10438.
9. Yeri A, Courtright A, Danielson K, Hutchins E, Alsop E, Carlson E, et al. Evaluation of commercially available small RNASeq library preparation kits using low input RNA. BMC Genomics. 2018;19(1):331.
10. The Cancer Genome Atlas Network. Genomic classification of cutaneous melanoma. Cell. 2015;161(7):1681–96.
11. Van Laar R, King S, McCoy R, Saad M, Fereday S, Winship I, et al. Translation of a circulating microRNA signature of melanoma into a solid tissue assay to improve diagnostic accuracy and precision. Biomark Med. 2020;15(13):1111–22.
12. van Laar RK, Lincoln MT, van Laar BJ. A plasma microRNA biomarker of melanoma as a personalised assessment of treatment response. Melanoma Res. 2019;29(1):19–22.
13. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 2014;42(Database issue):D68–73.
14. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3): R25.
15. Aparicio-Puerta E, Gomez-Martin C, Giannoukakos S, Medina JM, Scheepbouwer C, Garcia-Moreno A, et al. sRNAbench and sRNAtoolbox 2022 update: accurate miRNA and sncRNA profiling for model and non-model organisms. Nucleic Acids Res. 2022;50(W1):W710–7.
16. Kavakiotis I, Alexiou A, Tastsoglou S, Vlachos IS, Hatzigeorgiou AG. DIANA-miTED: a microRNA tissue expression database. Nucleic Acids Res. 2022;50(D1):D1055–61.
17. Xiong J, Bing Z, Guo S. Observed Survival Interval: A Supplement to TCGA Pan-Cancer Clinical Data Resource. Cancers (Basel). 2019;11(3):280.
18. Bustos MA, Tran KD, Rahimzadeh N, Gross R, Lin SY, Shoji Y, et al. Integrated assessment of circulating cell-free MicroRNA signatures in plasma of patients with melanoma brain metastasis. Cancers (Basel). 2020;12(6):1692.
19. Chang C-C, Lin C-J. LIBSVM: A library for support vector machines. ACM Trans Intell Syst Technol. 2011;2(3):Article 27.
20. Simon R, Lam A, Li MC, Ngan M, Menenzes S, Zhao Y. Analysis of gene expression data using BRB-ArrayTools. Cancer Inform. 2007;3:11–7.
21. Bair E, Tibshirani R. Semi-supervised methods to predict patient survival from gene expression data. PLoS Biol. 2004;2(4): E108.
22. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86(2):420–8.
23. Schoonjans F, Zalata A, Depuydt CE, Comhaire FH. MedCalc: a new computer program for medical statistics. Comput Methods Programs Biomed. 1995;48(3):257–62.
24. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology. 1983;148(3):839–43.
25. Garg M, Couturier D-L, Nsengimana J, Fonseca NA, Wongchenko M, Yan Y, et al. Tumour gene expression signature in primary melanoma predicts long-term outcomes. Nat Commun. 2021;12(1):1137.
26. Neagu M, Constantin C, Cretoiu SM, Zurac S. miRNAs in the diagnosis and prognosis of skin cancer. Front Cell Dev Biol. 2020;8: 71.

27. Pritchard CC, Cheng HH, Tewari M. MicroRNA profiling: approaches and considerations. Nat Rev Genet. 2012;13(5):358–69.
28. Tam S, de Borja R, Tsao MS, McPherson JD. Robust global microRNA expression profiling using next-generation sequencing technologies. Lab Invest. 2014;94(3):350–8.
29. Sainsbury P. Development and oversight of ethical health promotion quality assurance and evaluation activities involving human participants. Health Promot J Austr. 2015;26(3):176–81.

Publisher's Note