HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2018 May 26.

Published in final edited form as:

Cell Rep. 2015 October 27; 13(4): 840–853. doi:10.1016/j.celrep.2015.09.037.

New Functional Signatures for Understanding Melanoma Biology from Tumor Cell Lineage-Specific Analysis

Florian Rambow[1,2,3,4,9], **Bastien Job**[5,9], **Valérie Petit**[1,2,3,4], **Franck Gesbert**[1,2,3,4], **Véronique Delmas**[1,2,3,4], **Hannah Seberg**[7], **Guillaume Meurice**[5], **Eric Van Otterloo**[7], **Philippe Dessen**[5], **Caroline Robert**[6], **Daniel Gautheret**[5], **Robert A. Cornell**[7], **Alain Sarasin**[8], and **Lionel Larue**[1,2,3,4,*]

[1]Institut Curie, Normal and Pathological Development of Melanocytes, 91405 Orsay, France

[2]Centre National de la Recherche Scientifique (CNRS) UMR3347, 91405 Orsay, France

[3]INSERM U1021, 91405 Orsay, France

[4]Equipe Labellisée – Ligue Nationale contre le Cancer, 91405 Orsay, France

[5]Plateforme de Bioinformatique, UMS AMMICA, Gustave-Roussy, 94805 Villejuif, France

[6]INSERM U981, Gustave-Roussy, 94805 Villejuif, France

[7]Department of Anatomy and Cell Biology, University of Iowa, Iowa City, IA 52242, USA

[8]Centre National de la Recherche Scientifique (CNRS) UMR8200, Gustave-Roussy and University Paris-Sud, 94805 Villejuif, France

SUMMARY

Molecular signatures specific to particular tumor types are required to design treatments for resistant tumors. However, it remains unclear whether tumors and corresponding cell lines used for drug development share such signatures. We developed similarity core analysis (SCA), a universal and unsupervised computational framework for extracting core molecular features common to tumors and cell lines. We applied SCA to mRNA/miRNA expression data from various sources, comparing melanoma cell lines and metastases. The signature obtained was associated with phenotypic characteristics in vitro, and the core genes *CAPN3* and *TRIM63* were implicated in

*Correspondence: lionel.larue@curie.fr.

[9]Co-first author

AUTHOR CONTRIBUTIONS

F.R., B.J., A.S., and L.L. conceived and designed the experiments. F.R., B.J., V.P., F.G., V.D., H.S., G.M., E.V.O., and R.A.C. performed the experiments. F.R., B.J., P.D., R.A.C., and L.L. analyzed the data. C.R., D.G., and L.L. contributed reagents, materials, and analysis tools. F.R., B.J., R.A.C., A.S., and L.L. contributed to the writing of the manuscript. F.R. and B.J. conceived, designed, and performed the experiments; analyzed the data; and contributed to the writing of the manuscript. F.R. was mostly involved in the wet experiments, whereas B.J. was more involved in the computational experiments.

ACCESSION NUMBERS

Cell line data for the expression of genes (23 samples) and miRNAs (21) have been deposited to NCBI GEO and are available under accession number GEO: GSE67638.

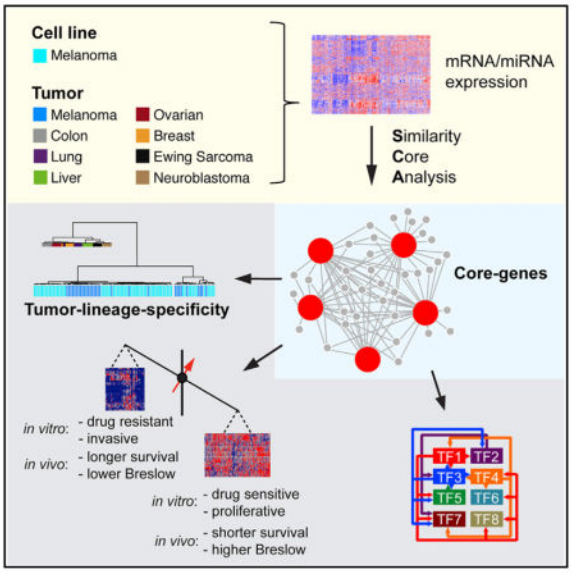
SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.09.037.

melanoma cell migration/invasion. About 90% of the melanoma signature genes belong to an intrinsic network of transcription factors governing neural development (*TFAP2A*, *DLX2*, *ALX1*, *MITF*, *PAX3*, *SOX10*, *LEF1*, and *GAS7*) and miRNAs (*211-5p*, *221-3p*, and *10a-5p*). The SCA signature effectively discriminated between two subpopulations of melanoma patients differing in overall survival, and classified MEKi/BRAFi-resistant and -sensitive melanoma cell lines.

In Brief

Cancer cell lines are at the forefront of drug discovery but are often limited in representing the tumor of origin due to the artificial culture conditions. Rambow et al. develop a computational approach for identifying tumor cell lineage expression cores. These core genes reveal relevant molecular dependencies linking aggressiveness, patient survival, and drug sensitivity.



INTRODUCTION

Genome-wide profiling approaches have provided molecular insight into cancer initiation and progression, promoting the development of targeted drugs for personalized treatment (Gonzalez-Angulo et al., 2010). However, cancer treatments targeting driver mutations are prone to resistance development, due to genetic and epigenetic tumor heterogeneity. Tumor cell-type- or lineage-specific expression characteristics were identified as key predictors of responses to several compounds in a screen of 479 cancer cell lines of 36 different tumor types (Barretina et al., 2012). Cancer-derived cell lines also are used in drug discovery, but may differ from the tumor of origin (Ertel et al., 2006; Gillet et al., 2013; Masters, 2000). In tumors, different cell types interact with each other and the tumor microenvironment, whereas cell lines are essentially clonal. Culture conditions and long-term passaging select the most adapted (potentially artificial) patterns of gene expression, resulting in differences in gene expression and epigenetic status between cell lines and tumors (van Staveren et al., 2009). The unsupervised clustering of whole-transcriptome expression data clearly separates cell lines from tumor samples, even if they contain the same mutations (Domcke et al.,

2013). We therefore need to identify significant similarities in gene expression between cancer-derived cell lines and tumors that might be masked by differences due to biological growth conditions.

Differential gene expression analyses have identified differences, but not core similarities, between cell lines and tumors for particular cancers. Comparisons of cell lines and tumors on this basis are uninformative, as they simply separate in vivo and in vitro samples (Domcke et al., 2013). Supervised gene lists can be used to identify suitable tumor models from cancer cell lines (Dancik et al., 2011; Gillet et al., 2011; Uva et al., 2010), or similarities between cancer cell lines and their tumors of origin can be scored with a tissue similarity index (TSI). This method uses expression data, without identifying the features underlying the similarity (Sandberg and Ernberg, 2005).

We designed a computational framework for the unsupervised extraction of core (= signature) molecular features common to tumors and cell lines. Characterization of this core should provide information about phenotypic in vitro and clinical in vivo characteristics, facilitating lineage-specific approaches. This approach is unique in its use of similarities rather than differences and its retention of weakly expressed genes. Similarity core analysis (SCA) uses a pan-cancer comparative approach to detect tumor cell lineage-specific genes also expressed in vitro, for a given tumor type, corresponding to subtle, biologically meaningful changes in a given lineage. It correctly identified a melanoma-specific molecular signature including genes for migration and invasion proteins, such as TRIM63 and CAPN3, and it underlies melanoma survival.

A similar analysis of microRNA (miRNA) highlighted a regulatory network connecting lineage-specific miRNAs (miR-211, 221, and 10a), transcription factors (e.g., MITF, TFAP2A, SOX10, and DLX2), and target genes potentially involved in phenotype switching. These specific, biologically relevant features demonstrate the utility of SCA for identifying lineage-specific expression cores relevant for tumor biology. SCA is suitable for use with all tumor cell types. Here we applied it to mRNA and miRNA data for melanoma tumors and cell lines.

RESULTS

SCA Identifies a Melanoma-Specific Expression Signature

We applied unsupervised hierarchical clustering to gene expression profiles from tumors and cell lines (Table S1). Melanoma tumors and derived cell lines clustered separately, as is the case for other cancer types (Gillet et al., 2011), whether we used global mRNA levels (17,322 unique genes) (Figure S1A), the 5,000 genes with the most variable expression (Figure S1B), or 244 multidrug resistance (MDR) genes (Figure S1C).

We then developed a computational framework, SCA, to extract cancer type-specific gene expression patterns by comparing the proximity of melanoma cell lines to melanomas and other tumor types. Using expression profiles for 240 tumor samples from eight cancer types and 63 melanoma cell lines (Table S1), we generated 1×10^6 subsets of expression levels for 20 randomly selected genes (randomly selected contexts [RSCs]).

For each RSC, we used principal component analysis (PCA) to capture the characteristic gene expression pattern for each tumor type and to reduce dimensionality, generating a different PCA space for each RSC. We decreased the number of dimensions to the first eight components for each space. We converted the PCA space into a correlation space by projecting the 240 tumors, correlating the expression profile of each tumor with that of each of the eight components. We then reduced the data again by computing the Euclidean centroid for each tumor type from the coordinates of its 30 tumors. The candidate cell line then was projected similarly into the correlation space, and we calculated the Euclidean distances between the cell line and the centroid for each tumor type (Figure S1D).

We evaluated the representation of melanoma specificity by selecting spaces for which the cell line position in the eight-dimensional correlation space was twice as close to the melanoma centroid as to any other centroid. The contribution of genes contribution to melanoma specificity was evaluated by calculating their frequency of occurrence in the RSCs corresponding to the selected spaces, with respect to all the RSCs generated. Fisher's exact test was used to evaluate the significance of enrichment. Finally, genes were ordered by decreasing contribution to melanoma specificity (increasing p value). An SD median (SDM) score was generated to take inter-cell line variability into account and to identify the top 100 genes (Figure 1A).

The top 100 genes (Table S1) of the mRNA signature defined by SCA (Figure S1E; SCA-MEL-mRNA signature or SCA-MEL-mRNA-100) showed expression to be more similar in the 30 melanomas and 63 melanoma cell lines than between melanoma cell lines and other tumor types. The numbers of contributing samples and tumor types were not exhaustive, but SCA can manage larger amounts of data if required.

We used literature vector analysis (LitVAn) to search for terms overrepresented in our signature from publications relating to the SCA-melanoma-mRNA signature (Akavia et al., 2010). Significant enrichment was observed for five terms as follows: melanocyt, melanoma, catenin, pigment, and tyrosinas (Figure 1B). These terms were linked to 36 of the 100 SCA genes, consistent with SCA yielding a cancer type-specific expression signature (Table S1).

The SCA signature was specific to melanoma, separating the 210 tumor samples (30 samples × 7 different cancer types) from the 30 melanoma samples and 63 melanoma cell lines in unsupervised analysis (Figure 1C). Of the SCA-melanoma-mRNA-100 signature, 45 genes were underexpressed (45DN) in melanomas, the other 55 being overexpressed (55UP) (Figure 1C; Table S1). Melanoma cell lines and melanomas clustered together, in two clear subgroups (Figure 1D), in analyses based on the SCA-melanoma-mRNA signature. This suggests that SCA identified a melanoma-specific signature potentially useful for comparisons of cell lines and tumors. The 55 overexpressed SCA genes were significantly enriched in terms associated with melanoma biology, such as melanocyt, melanoma, pigment, tyrosinas, waardenburg, SOX10, neural, and MITF (Table S1), whereas no such enrichment was observed for the other 45 genes (Figure 1E). We therefore focused on the 55UP SCA-melanoma-mRNA signature, which seemed to capture melanoma specificity. The overexpression of these genes in melanoma samples also identifies them as suitable targets for loss-of-function approaches.

Thus, SCA can extract a tumor cell lineage-specific signature from expression data for multiple cancers. Melanoma cell lines with this signature are relevant models of their in vivo counterparts, as illustrated below.

The SCA-Melanoma-mRNA-55UP Signature Is Correlated with Phenotypic Characteristics In Vitro

We investigated the correlation between SCA-melanoma-mRNA-55UP signature expression and in vitro behavior. We obtained expression profiles for 23 independent melanoma cell lines not used in the SCA analysis, in house, with a microarray platform different from that used to generate the previous tumor dataset. Two characteristic groups (groups 1 and 2) were obtained for clustering with the SCA-MEL-mRNA-55UP signature (Figure 2A). Two well-separated groups (groups 1b and 2b) also were obtained (Figure S2) when clustering with the SCA-MEL-mRNA-100 signature. In this second analysis, SK29, Mull, and Gerlach melanoma cell lines shifted from group 1 to group 2, suggesting that these cell lines were less representative of the two groups. This is consistent with the previous finding that 14% of 536 profiled melanoma cell lines were intermediate between proliferative and invasive phenotypes (Widmer et al., 2012).

We analyzed the association between SCA-MEL-mRNA-55UP and in vitro behavior, assessing migratory, invasive, tumorigenic, and proliferative potentials. Four representative melanoma cell lines were chosen from each subgroup for further analysis (Figure 2B). The cell lines of group 1 (G1, T1, 501mel, MNT-1, and SKMel3) generally expressed melanocyte differentiation-associated genes, such as *SOX10*, *MITF*, *TFAP2A*, *TYR*, *TYRP1*, *PAX3*, *RAB38*, *DCT*, *SILV*, *MLANA*, and *EDNRB*, more strongly than those of group 2 (WM1366, WM793, WM852, Lu1205, and A375M). Group 2 cell lines were more aggressive, with a greater wound-closure capacity, a larger capacity to invade Matrigel, and a higher tumorigenic capacity after subcutaneous injection into nude mice. However, they proliferated less than group 1 cell lines (Figure 2C). There are, thus, two groups of melanoma cell lines differing in SCA-MEL-mRNA-55UP signature expression and phenotypic characteristics.

Transcriptional Control of the SCA-MEL-mRNA-55UP Signature through MITF and New Targets Involved in Melanoma Cell Migration and Invasion

The 55UP SCA signature was enriched in MITF target genes (Figure 3A). We investigated the biological roles of *CAPN3* and *TRIM63*, two SCA-MEL-mRNA-55UP signature genes with unknown roles in melanoma biology. Consistent with SCA-based predictions, *CAPN3* and *TRIM63* were overexpressed in group 1 melanoma cell lines (Figures 3B and 3C). Small interfering RNA (siRNA)-mediated MITF knockdown in 501mel melanoma cells decreased CAPN3 and TRIM63 levels, consistent with direct regulation by MITF via the predicted MITF-binding sites (Strub et al., 2011; Figure 3D). *TRIM63* repression by siRNA in 501mel cells increased Matrigel invasion capacity (Figure 3E), whereas *CAPN3* repression enhanced wound closure (Figure 3F), consistent with the greater migration capacity of B16 murine melanoma cells after treatment with calpastatin, a pan-calpain inhibitor (Raimbourg et al., 2013).

The SCA-MEL-mRNA-28SDE Melanoma Signature

We carried out differential gene expression analysis (fold change > 2, adjusted $p < 0.05$) with the SCA-MEL-mRNA-55UP signature in order to identify genes significantly differentially expressed (SDE) between more and less aggressive melanoma cell lines. We identified 28 SDE genes. Clustering of the 23 melanoma cell lines with the SCA-MEL-mRNA-28SDE signature yielded two characteristic groups (groups 1 and 2) identical to those for SCA-MEL-mRNA-100 (Figure S3A). Of the 28 SDE genes, 26 were generally overexpressed in less aggressive (group 1) melanoma cell lines. Only RAGE and the non-coding RNA (uncharacterized LOC100130938) were significantly more strongly expressed in more aggressive (group 2) melanoma cell lines. RAGE overexpression in WM115 primary melanoma cells confers a metastatic phenotype (Meghnani et al., 2014).

SCA Identifies a Melanoma-Specific miRNA Expression Signature

Unsupervised clustering of global (422) miRNA levels (Figure S4A) or for the 100 most variably expressed miRNAs (Figure S4B) clearly separated melanoma cell lines and metastases. We applied SCA to the miRNA profiles of 157 tumor samples (21 samples × 7 cancer types), with 21 melanoma samples and 51 melanoma cell lines (Table S1). We selected the miRNAs making the largest contribution to melanoma specificity, by calculating miRNA enrichment and identifying a core of 51 miRNAs, the SCA-MEL-miR-51 signature (Figure S4C; Table S1). Melanoma samples and cell lines formed a distinct branch in the dendrogram (Figure 4A) and clustered together (Figure 4B) with this signature. Of the SCA-MEL-miR-51 genes, 22 were underexpressed (22DN) and 29 were overexpressed (29UP) in melanoma samples (Figure 4A). LitVAn allows only mRNA gene symbols for input. We therefore predicted mRNA targets for the 22DN and 29UP miRNA signatures by sequence-based approaches only. The predicted miRNA targets had to overlap the SCA-MEL-mRNA signature to qualify for LitVAn analysis (Table S1). The predicted target mRNAs of SCA-MEL-miR-29UP genes were enriched in the melanocyt, waardenburg, melanoma, SOX10, neural, crest, and MITF terms. The predicted mRNA targets of SCA-MEL-miR-22DN were enriched in all these terms except MITF and SOX10 (Figure 4C).

The clustering of 21 independent melanoma cell lines with SCA-MEL-miR-51 yielded two distinct groups (groups 1 and 2) identical to those for SCA-MEL-mRNA-100 and SCA-MEL-mRNA-28SDE (Figure 4D; Figure S2). Thus, SCA extracts meaningful melanoma-associated miRNAs, with expression patterns correlated with the in vitro behavior of melanoma cell lines. We performed a conventional differential miRNA expression analysis between group 1 and group 2 cell lines. Only three miRNAs from SCA-MEL-miR-51 were SDE (adjusted p value < 0.05). The more aggressive melanoma cell lines (group 2) had higher levels of miR-221-3p and miR-10a-5p. miR-221 has been identified as a potential biomarker of melanoma, as high serum miR-221 levels are correlated with tumor thickness (Kanemaru et al., 2011) and higher recurrence risk (Friedman et al., 2012). miR-10a expression has been linked to gene amplification in melanoma (Zhang et al., 2006). Group 1 cell lines overexpressed miR-211-5p (Figure 4E), a melanocyte lineage-specific small non-coding RNA intronic to *TRPM1*, a target gene of MITF. By upregulating *TRPM1* transcription, MITF, which is critical for melanocyte differentiation and survival and for melanoma progression, indirectly drives miR-211 expression (Levy et al., 2010).

The Melanoma-Specific Regulatory Network Is Associated with Melanoma Survival

The SCA-MEL-mRNA-55UP signature includes eight transcription factor genes as follows: *ALX1*, *DLX2*, *GAS7*, *LEF1*, *MITF*, *PAX3*, *SOX10*, and *TFAP2A*. Ingenuity Pathway Analysis, based on published findings, and available chromatin immunoprecipitation sequencing (ChIP-seq) datasets for these transcription factors and their surrogates (Table S1) suggested probable mutual control for these factors (Figure 5A). Moreover, 90% of the SCA-MEL-mRNA-55UP signature was connected to at least one of these transcription factors (Figure 5B); the myelin basic protein gene (*MBP*) was potentially regulated by five of the eight transcription factors (MITF, SOX10, LEF1, PAX3, and TFAP2A). Significant enrichment was observed for TFAP2A, as 40 of the SCA-MEL-mRNA-55UP genes had TFAP2A ChIP-seq peaks in their promoters ($p < 3.77 \times 10^{-14}$). MITF, PAX3, SOX10, LEF1, and TFAP2A have been implicated in melanocyte/melanoma biology (Harris et al., 2010; Koludrovic and Davidson, 2013; Larue and Delmas, 2006; Medic and Ziman, 2009; Melnikova and Bar-Eli, 2008). The roles of ALX1, DLX2, and GAS7 in the melanocyte lineage are unclear, although they have been linked to neural tube development (Achim et al., 2014; Chao et al., 2005; Juriloff and Harris, 2000). Thus, the SCA-MEL-mRNA-55UP signature identified transcriptional programs involving melanoma core genes, driven by melanocyte, neural crest, and neuronal transcription factors.

We identified three miRNAs differentially expressed between melanomas and contributing to a network regulating melanoma phenotype (Figure 5C). Based on the inversely correlated expression levels for these three specific miRNAs, and sequence-based methods (targetscan), we identified 12 potential mRNA targets (including MITF, SOX10, TFAP2A, EDNRB, and MBP). miR-221 overexpression was associated with enhanced invasive, migratory, and tumorigenic behavior of melanoma cell lines in vitro. Moreover, miR-221-3p levels were inversely correlated with expression levels for three potential target genes encoding key transcription factors (*MITF*, *SOX10*, and *TFAP2A*). These findings are consistent with reports that inhibiting miR-221 expression with antagomiRs decreases invasion and migration by melanoma cells (Felicetti et al., 2008). The expression of miR-211 and miR-221-3p was correlated with a weak MITF transcriptional signature, typical of highly invasive melanoma cells (Hoek et al., 2008; Figures S5A and S5B). miR-211 restores adhesion through NUAK1 repression, suggesting that the MITF-miR-211 axis inhibits invasion by blocking adhesion (Bell et al., 2014). miR-221-3p targets MITF, thereby favoring invasion over proliferation. Epithelial-to-mesenchymal transition (EMT) inducers such as FOSL1 (FRA-1) and TGFB1 or TWIST1 (Figure S5C) have been shown to control the expression of miR-10a-5p and miR-221-3p, respectively (Li et al., 2013; Stinson et al., 2011).

Finally, we assessed the clinical relevance of melanoma SCA signatures by survival analyses of independent external data-sets. We downloaded RNA sequencing (RNA-seq) data for 292 melanoma metastases (the Cancer Genome Atlas [TCGA] portal, skin cutaneous melanoma cohort) and clustered them with the SCA-MEL-mRNA-28SDE signature (Figure 5D). Cluster 2 patients (n = 68), generally displaying underexpression of the 28 genes, survived significantly longer overall (p = 0.0191). Studies of two other published datasets (Bogunovic et al., 2009; Jönsson et al., 2010) confirmed that patients overexpressing the 28 SDE core

genes had significantly lower overall (p = 0.0362) and distant metastasis-free (p = 0.0205) survival, respectively (Figures S5D–S5G). By contrast, the melanoma phenotype-specific expression (MPSE) gene (n = 97) signature, derived from melanoma cell lines only, was not informative for survival analysis (Figures S5H and S5I; Widmer et al., 2012). However, a 46-gene signature, consisting mostly of immune-related genes and shown to be informative for patient survival (Mann et al., 2013), also was discriminant for patients' survival in the TCGA dataset (Figures S5J and S5K). Thus, overexpression of the 28 SDE signature, associated with a less aggressive phenotype in vitro, was associated with a worse patient outcome in vivo.

We performed an equivalent survival analysis with the three SDE miRNAs (211-5p, 10a-5p, and 221-3p). Patients from the TCGA cohort with high levels of miR-211-5p and low levels of miR-10a-5p and miR-221-3p had significantly lower overall survival (p = 0.0031) (Figure 5E). High miR-211-5p and low miR-10a-5p and miR-221-3p levels were associated with a less aggressive phenotype in vitro. The previously published miRNA-survival signature, consisting of another 18 miRNAs (Segura et al., 2010), yielded no significant results for analysis of the TCGA melanoma cohort (p = 0.0606) (Figures S5L and S5M).

In summary, SCA of melanoma samples revealed molecular features associated with the cellular and clinical features of melanoma samples, suggesting a melanoma model (Figure 5F).

DISCUSSION

Appropriate in vitro, in cellulo, and in vivo models are required for preclinical pipelines for cancer treatment. For ethical, economic, and biological complexity reasons, cancer cell lines are a good compromise for deciphering molecular and cellular mechanisms before testing therapies. Relevant systems, reproducing in vivo scenarios accurately, are thus required for molecular and cellular studies. Cell lines have long been used for therapeutic drug screening. Murine B16 melanoma cells have been widely used, but with little success. It has become clear that this specific cell line is useful for addressing particular biological questions, but not for the development of new treatments for humans, probably due to differences between this in vitro model and in vivo tumors. Large-scale cell-line panels are increasingly used for drug screening and omics data generation (Kim et al., 2014). Analyses of the correlation of molecular features with sensitivity to small-molecule drugs in many cancer cell lines have highlighted the need to take lineage-specific features into account (Barretina et al., 2012; Basu et al., 2013; Cheung et al., 2011). Our SCA approach could be useful for this. Similarly, tumor samples and cancer cell lines for specific tumors do not cluster in transcriptomics analyses, suggesting clear differences in gene expression that may render cell lines inappropriate for use in treatment development.

SCA identified cancer lineage-specific mRNA/miRNA expression patterns. This computational approach extracts signatures characteristic of both tissue samples and cell lines. The combination of lineage specificity and the relevance of the signature to both in vivo and in vitro conditions make it possible to identify molecular relationships missed by conventional strategies, such as differential gene expression analysis and batch removal

approaches comparing cell lines with tumors (Figure S6). SCA was developed in assuming that cancer cell lines of a specific lineage retain key transcriptional programs reminiscent of the original tumor and in the unsupervised extraction of these programs.

Unlike other methods, SCA identified a molecular mRNA and miRNA signature in melanoma tumors that was common to melanoma cell lines, suggesting that these cell lines were relevant to the corresponding tumor samples. This method can be applied to other tumor types. For melanoma samples, this approach also did the following: (1) identified new mRNA/miRNAs specific to the melanocyte lineage; (2) showed that most of the associated proteins were molecularly linked via eight transcription factors also included in the signature; and (3) showed that, according to SCA-MEL-mRNA-55UP and SCA-MEL-miR signatures, aggressive cell lines corresponded to non-aggressive melanomas and vice versa.

This mRNA/miRNA signature is, by design, specific to the melanocyte lineage. Pharmacological approaches are based on toxicity and specificity. By using the specificity of this signature to target these mRNA/proteins and/or miRNAs, it should be possible to alter their molecular and cellular functions. For instance, CAPN3 (protease) and TRIM63 (E3 ubiquitin ligase) are produced abundantly by aggressive melanomas and regulate in vitro migration and invasion, respectively. The enzymatic activities of these proteins may induce the release of toxic products from prodrugs. This general strategy already has been proposed, with tyrosinase as the enzyme and MITF (the master transcription factor of the melanocyte lineage) as the target (Sáez-Ayala et al., 2013). Such cell lineage-targeted therapy could be coupled with molecularly targeted therapy and/or immunotherapy, yielding additional benefits.

The mRNA components of the SCA-MEL-mRNA signature include mRNAs for key neural crest and melanocyte transcription factors and their targets. *MITF* and *TYR* scored highest in the SCA-MEL-mRNA signature analysis, and our melanoma-specific signature was highly enriched in MITF target genes (30 of 55), including melanocyte differentiation genes, such as *MLANA*, *TYRP1*, *DCT*, and *SILV*, and genes involved in melanosome biogenesis, such as *MFN2* and *RAB38* (Daniele et al., 2014; Loftus et al., 2002). It has been suggested that RAB7 controls melanoma progression through lineage-specific wiring of the endolysosomal pathway (Alonso-Curbelo et al., 2014). Lysosomes and melanosomes have common precursors (Raposo and Marks, 2007). The Cancer Cell Line Encyclopedia (CCLE) cohort was used to explore expression patterns specific to melanoma cell lines by comparison with cell lines for other cancer types. This validates our approach, also using tumor samples and cell lines to identify lineage-specific dependences.

SCA-MEL-mRNA highlights a regulatory network including eight transcription factors (MITF, TFAP2A, ALX1, DLX2, PAX3, SOX10, LEF1, and GAS7). TFAP2A seems to be a key regulator of this lineage with significant target enrichment ($p < 3.77 \times 10^{-4}$). We identified a new potential regulator of TFAP2A encoded by a gene upstream from the *TFAP2A* promoter: open reading frame 218 on chromosome 6 (C6ORF218) or long intergenic non-protein coding RNA 518 (LINC00518). Given its location and strong coexpression with TFAP2A, LINC00518 may act in *cis*, regulating the expression of its neighbor (Luo et al., 2013). TFAP2A loss contributes to progression from the radial growth

phase to the vertical growth phase in melanoma (Mobley et al., 2012). Our melanoma cell lines with low levels of TFAP2A and MITF expression had a more aggressive phenotype in vitro. TFAP2A controls neural crest differentiation and development and is strongly expressed in neural crest cells migrating from the cranial folds during neural tube closure. Homozygous TFAP2A-knockout mice have neural tube defects and craniofacial and body wall abnormalities (Zhang et al., 1996). ALX1 and DLX2 have similar roles in cranial neural crest development (Merlo et al., 2000; Uz et al., 2010). The other transcription factors from our signature, SOX10, PAX3, LEF1 and GAS7, regulate neural crest and neuronal development (Betancur et al., 2010; Hung et al., 2013; Sauka-Spengler and Bronner-Fraser, 2008), highlighting the lineage specificity of melanoma cells. The transcription factors identified constitute the core of our proposed regulatory network. About 90% of the remaining melanoma-specific signature genes (47 of 55) may be regulated by at least one of these transcription factors.

Variants of the SCA-MEL-mRNA/miR signature in melanoma cell lines predict behavior in vitro and melanoma patient survival. We did not expect tumors to form distinct subclasses associated with tumor aggressiveness. Further in vitro and in vivo studies are now required to gain insight into this serendipitously discovered relationship. Aggressive tumors had signatures corresponding to non-aggressive cell lines and vice versa. This apparent discrepancy probably results from high and modifiable MITF levels. This transcription factor acts as a rheostat, determining subpopulation identity (Carreira et al., 2006). Low MITF levels in melanomas result in invasive cells with stem-like properties, arrested in G1 phase and efficiently inducing tumors (Cheli et al., 2011). By contrast, high MITF levels activate differentiation genes driving melanin production, such as tyrosinase (TYR) and melanosome biogenesis enzyme genes (Cheli et al., 2010). Cell lines consist of largely homogeneous dividing cells. We considered cell lines to be aggressive if they invaded, migrated, and formed tumors in nude mice, even if they divided slowly. Conversely, even rapidly dividing non-aggressive cell lines did not invade, migrate, or form tumors in mice.

We evaluated the prediction by the SCA signature of MEKi and BRAFi drug sensitivity in melanoma cell lines. Strong SCA-MEL-mRNA-28SDE signature expression was associated with sensitivity to MEK and BRAF inhibition (Figures S3B–S3E). Also, we evaluated the SCA-MEL-mRNA-28SDE signature using publicly available primary melanoma samples (n = 297, GEO: GSE57715) and melanoma cell lines (n = 63, GEO: GSE7127) (Figure 6). Interestingly, the SCA-MEL-mRNA-28SDE separated two groups of primary melanomas mixed with melanoma cell lines. After classifying the melanoma cell lines as either proliferative or invasive using Heuristic Online Phenotype Prediction (HOPP, Widmer et al., 2012), we noticed that strong SCA-MEL-mRNA-28SDE expression identified proliferative melanoma cell lines that coclustered with primary melanomas of deeper Breslow depth (Figure 6). This provides further support for our SCA approach, because it is informative for both metastatic and primary melanoma.

Tumors are heterogeneous with some cells proliferating and others invading. Invading cells that cannot proliferate do not induce death rapidly, but the converse may not be true. We analyzed survival in melanoma metastases, proliferative cells that were previously invasive; however, melanoma cells readily switch between these states. There is increasing evidence

for this phenotype switching phenomenon (Hoek and Goding, 2010). We considered the average expression pro-files of melanomas to be the predominant phenotypic state of all melanoma cells within the lesion. A previously defined invasive signature (Widmer et al., 2012) did not overlap with our SCA signature because the invasive program is common to different cancers and, thus, not specific to melanoma (Figure S3F).

SCA may become a standard tool for extracting lineage-specific molecular dependencies for any tumor type. Increasing computational power will make it possible to analyze larger cohorts. SCA can highlight cancer regulatory networks operating in vitro and is particularly valuable for drug development.

EXPERIMENTAL PROCEDURES

SCA

Data Transformation and Context Generation—Expression data for eight tumor types (Table S1) were retrieved from GEO. We aggregated 30 randomly selected samples from each set into a single dataset, which was then quantile-normalized for interdisease scaling. This multi-tumor dataset covered 17,322 genes. A similar miRNA analysis was performed with 21 samples per tumor type, covering 422 miRNAs. One million subsets of 20 randomly selected features (mRNA/miRNA) referred to as RSCs were generated.

PCA Space Computation and Tumor Projection—For each RSC, PCA was performed on the tumor dataset in the same manner as singular value decomposition (SVD) was used by Sandberg and Ernberg (2005), but limited here to the RSC features, whereas Sandberg and Ernberg used a single signature of 303 genes. We limited the dimensions of the resulting PCA space to its first eight, assuming the different types of cancer to be the major source of variance in the tumoral dataset. Tumors used for PCA space construction were then projected into the space by Pearson correlation of their expression to yield a new eight-dimensional correlation space. Euclidean centroids were calculated for each tumor type in this correlation space. Each cell line was then projected similarly into the correlation spaces corresponding to each RSC.

Context Selection Based on Disease Similarity and Specificity—We evaluated the ability of each RSC to capture the features making a cell line uniquely similar to the corresponding disease, with a graphical criterion: its coordinate distance to the melanoma centroid in the correlation space. RSCs were considered to be melanoma-specific if the cell line coordinate in the space was at least twice as close to the melanoma centroid as to any other centroid (Figure S1D). This selection procedure was performed independently for each melanoma cell line.

Feature Enrichment Evaluation and Core Generation—For each cell line, feature enrichment in the selected RSCs versus their occurrence in the one million generated RSCs was evaluated with Fisher’s exact test. Features were then ordered by increasing p values, after false discovery rate correction by the Benjamini-Hochberg method. For core generation, we pooled the feature-ranking results for all independently analyzed cell lines and assigned a score (SDM score) to each feature, by calculating the SD of its rank across

cell lines divided by its median rank. After ordering features in descending order of scores, cores were built by taking the 100 best features for mRNA and all features with a score above 1.0 for miRNA. The SCA code is provided as an R script (https://github.com/aoumess/SCA).

Variability of Gene Expression

See the Supplemental Experimental Procedures.

Sample Classification

See the Supplemental Experimental Procedures.

Data Retrieval, Preprocessing, and Normalization of Public Datasets

See the Supplemental Experimental Procedures.

Literature Vector Analysis

See the Supplemental Experimental Procedures.

Survival Analysis

See the Supplemental Experimental Procedures.

Molecular and Cellular Biology Methods

See the Supplemental Experimental Procedures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Audrey Bechadergue for excellent technical assistance and Richard Marais for helpful comments. We thank Thomas Robert from the Agilent micro-array platform at Gustave-Roussy and the team caring for the animals at Institut Curie, especially J.D. Mam and H. Harmange. This work was supported by the Ligue Nationale Contre le Cancer (Equipe labellisée), Cancéropole Ile-de-France, Institut National du Cancer (INCa), Labex CelTisPhyBio (ANR-11-LBX-0038), and Pair Mélanome.

References

Achim K, Salminen M, Partanen J. Mechanisms regulating GABAergic neuron development. Cell Mol Life Sci. 2014; 71:1395–1415. [PubMed: 24196748]

Akavia UD, Litvin O, Kim J, Sanchez-Garcia F, Kotliar D, Causton HC, Pochanard P, Mozes E, Garraway LA, Pe'er D. An integrated approach to uncover drivers of cancer. Cell. 2010; 143:1005–1017. [PubMed: 21129771]

Alonso-Curbelo D, Riveiro-Falkenbach E, Pérez-Guijarro E, Cifdaloz M, Karras P, Osterloh L, Megías D, Cañón E, Calvo TG, Olmeda D, et al. RAB7 controls melanoma progression by exploiting a lineage-specific wiring of the endolysosomal pathway. Cancer Cell. 2014; 26:61–76. [PubMed: 24981740]

Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehár J, Kryukov GV, Sonkin D, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483:603–607. [PubMed: 22460905]

Basu A, Bodycombe NE, Cheah JH, Price EV, Liu K, Schaefer GI, Ebright RY, Stewart ML, Ito D, Wang S, et al. An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. Cell. 2013; 154:1151–1161. [PubMed: 23993102]

Bell RE, Khaled M, Netanely D, Schubert S, Golan T, Buxbaum A, Janas MM, Postolsky B, Goldberg MS, Shamir R, Levy C. Transcription factor/microRNA axis blocks melanoma invasion program by miR-211 targeting NUAK1. J Invest Dermatol. 2014; 134:441–451. [PubMed: 23934065]

Betancur P, Bronner-Fraser M, Sauka-Spengler T. Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest. Proc Natl Acad Sci USA. 2010; 107:3570–3575. [PubMed: 20139305]

Bogunovic D, O'Neill DW, Belitskaya-Levy I, Vacic V, Yu YL, Adams S, Darvishian F, Berman R, Shapiro R, Pavlick AC, et al. Immune profile and mitotic index of metastatic melanoma lesions enhance clinical staging in predicting patient survival. Proc Natl Acad Sci USA. 2009; 106:20429–20434. [PubMed: 19915147]

Carreira S, Goodall J, Denat L, Rodriguez M, Nuciforo P, Hoek KS, Testori A, Larue L, Goding CR. Mitf regulation of Dia1 controls melanoma proliferation and invasiveness. Genes Dev. 2006; 20:3426–3439. [PubMed: 17182868]

Chao CC, Chang PY, Lu HH. Human Gas7 isoforms homologous to mouse transcripts differentially induce neurite outgrowth. J Neurosci Res. 2005; 81:153–162. [PubMed: 15948147]

Cheli Y, Ohanna M, Ballotti R, Bertolotto C. Fifteen-year quest for microphthalmia-associated transcription factor target genes. Pigment Cell Melanoma Res. 2010; 23:27–40. [PubMed: 19995375]

Cheli Y, Giuliano S, Botton T, Rocchi S, Hofman V, Hofman P, Bahadoran P, Bertolotto C, Ballotti R. Mitf is the key molecular switch between mouse or human melanoma initiating cells and their differentiated progeny. Oncogene. 2011; 30:2307–2318. [PubMed: 21278797]

Cheung HW, Cowley GS, Weir BA, Boehm JS, Rusin S, Scott JA, East A, Ali LD, Lizotte PH, Wong TC, et al. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. Proc Natl Acad Sci USA. 2011; 108:12372–12377. [PubMed: 21746896]

Dancik GM, Ru Y, Owens CR, Theodorescu D. A framework to select clinically relevant cancer cell lines for investigation by establishing their molecular similarity with primary human cancers. Cancer Res. 2011; 71:7398–7409. [PubMed: 22012889]

Daniele T, Hurbain I, Vago R, Casari G, Raposo G, Tacchetti C, Schiaffino MV. Mitochondria and melanosomes establish physical contacts modulated by Mfn2 and involved in organelle biogenesis. Curr Biol. 2014; 24:393–403. [PubMed: 24485836]

Domcke S, Sinha R, Levine DA, Sander C, Schultz N. Evaluating cell lines as tumour models by comparison of genomic profiles. Nat Commun. 2013; 4:2126. [PubMed: 23839242]

Ertel A, Verghese A, Byers SW, Ochs M, Tozeren A. Pathway-specific differences between tumor cell lines and normal and tumor tissue cells. Mol Cancer. 2006; 5:55. [PubMed: 17081305]

Felicetti F, Errico MC, Bottero L, Segnalini P, Stoppacciaro A, Biffoni M, Felli N, Mattia G, Petrini M, Colombo MP, et al. The promyelocytic leukemia zinc finger-microRNA-221/-222 pathway controls melanoma progression through multiple oncogenic mechanisms. Cancer Res. 2008; 68:2745–2754. [PubMed: 18417445]

Friedman EB, Shang S, de Miera EV, Fog JU, Teilum MW, Ma MW, Berman RS, Shapiro RL, Pavlick AC, Hernando E, et al. Serum microRNAs as biomarkers for recurrence in melanoma. J Transl Med. 2012; 10:155. [PubMed: 22857597]

Gillet JP, Calcagno AM, Varma S, Marino M, Green LJ, Vora MI, Patel C, Orina JN, Eliseeva TA, Singal V, et al. Redefining the relevance of established cancer cell lines to the study of mechanisms of clinical anti-cancer drug resistance. Proc Natl Acad Sci USA. 2011; 108:18708–18713. [PubMed: 22068913]

Gillet JP, Varma S, Gottesman MM. The clinical relevance of cancer cell lines. J Natl Cancer Inst. 2013; 105:452–458. [PubMed: 23434901]

Gonzalez-Angulo AM, Hennessy BT, Mills GB. Future of personalized medicine in oncology: a systems biology approach. J Clin Oncol. 2010; 28:2777–2783. [PubMed: 20406928]

Harris ML, Baxter LL, Loftus SK, Pavan WJ. Sox proteins in melanocyte development and melanoma. Pigment Cell Melanoma Res. 2010; 23:496–513. [PubMed: 20444197]

Hoek KS, Goding CR. Cancer stem cells versus phenotype-switching in melanoma. Pigment Cell Melanoma Res. 2010; 23:746–759. [PubMed: 20726948]

Hoek KS, Schlegel NC, Brafford P, Sucker A, Ugurel S, Kumar R, Weber BL, Nathanson KL, Phillips DJ, Herlyn M, et al. Metastatic potential of melanomas defined by specific gene expression profiles with no BRAF signature. Pigment Cell Res. 2006; 19:290–302. [PubMed: 16827748]

Hoek KS, Eichhoff OM, Schlegel NC, Döbbeling U, Kobert N, Schaerer L, Hemmi S, Dummer R. In vivo switching of human melanoma cells between proliferative and invasive states. Cancer Res. 2008; 68:650–656. [PubMed: 18245463]

Hung FC, Cheng YC, Sun NK, Chao CC. Identification and functional characterization of zebrafish Gas7 gene in early development. J Neurosci Res. 2013; 91:51–61. [PubMed: 23086717]

Jönsson G, Busch C, Knappskog S, Geisler J, Miletic H, Ringnér M, Lillehaug JR, Borg A, Lønning PE. Gene expression profiling-based identification of molecular subtypes in stage IV melanomas with different clinical outcome. Clin Cancer Res. 2010; 16:3356–3367. [PubMed: 20460471]

Juriloff DM, Harris MJ. Mouse models for neural tube closure defects. Hum Mol Genet. 2000; 9:993–1000. [PubMed: 10767323]

Kanemaru H, Fukushima S, Yamashita J, Honda N, Oyama R, Kakimoto A, Masuguchi S, Ishihara T, Inoue Y, Jinnin M, Ihn H. The circulating microRNA-221 level in patients with malignant melanoma as a new tumor marker. J Dermatol Sci. 2011; 61:187–193. [PubMed: 21273047]

Kim N, He N, Yoon S. Cell line modeling for systems medicine in cancers (review). Int J Oncol. 2014; 44:371–376. [PubMed: 24297677]

Koludrovic D, Davidson I. MITF, the Janus transcription factor of melanoma. Future Oncol. 2013; 9:235–244. [PubMed: 23414473]

Larue L, Delmas V. The WNT/Beta-catenin pathway in melanoma. Front Biosci. 2006; 11:733–742. [PubMed: 16146765]

Levy C, Khaled M, Robinson KC, Veguilla RA, Chen PH, Yokoyama S, Makino E, Lu J, Larue L, Beermann F, et al. Lineage-specific transcriptional regulation of DICER by MITF in melanocytes. Cell. 2010; 141:994–1005. [PubMed: 20550935]

Li J, Dong J, Zhang ZH, Zhang DC, You XY, Zhong Y, Chen MS, Liu SM. miR-10a restores human mesenchymal stem cell differentiation by repressing KLF4. J Cell Physiol. 2013; 228:2324–2336. [PubMed: 23696417]

Loftus SK, Larson DM, Baxter LL, Antonellis A, Chen Y, Wu X, Jiang Y, Bittner M, Hammer JA 3rd, Pavan WJ. Mutation of melanosome protein RAB38 in chocolate mice. Proc Natl Acad Sci USA. 2002; 99:4471–4476. [PubMed: 11917121]

Luo H, Sun S, Li P, Bu D, Cao H, Zhao Y. Comprehensive characterization of 10,571 mouse large intergenic noncoding RNAs from whole transcriptome sequencing. PLoS ONE. 2013; 8:e70835. [PubMed: 23951020]

Mann GJ, Pupo GM, Campain AE, Carter CD, Schramm SJ, Pianova S, Gerega SK, De Silva C, Lai K, Wilmott JS, et al. BRAF mutation, NRAS mutation, and the absence of an immune-related expressed gene profile predict poor outcome in patients with stage III melanoma. J Invest Dermatol. 2013; 133:509–517. [PubMed: 22931913]

Masters JR. Human cancer cell lines: fact and fantasy. Nat Rev Mol Cell Biol. 2000; 1:233–236. [PubMed: 11252900]

Medic S, Ziman M. PAX3 across the spectrum: from melanoblast to melanoma. Crit Rev Biochem Mol Biol. 2009; 44:85–97. [PubMed: 19401874]

Meghnani V, Vetter SW, Leclerc E. RAGE overexpression confers a metastatic phenotype to the WM115 human primary melanoma cell line. Biochim Biophys Acta. 2014; 1842:1017–1027. [PubMed: 24613454]

Melnikova VO, Bar-Eli M. Transcriptional control of the melanoma malignant phenotype. Cancer Biol Ther. 2008; 7:997–1003. [PubMed: 18698165]

Merlo GR, Zerega B, Paleari L, Trombino S, Mantero S, Levi G. Multiple functions of Dlx genes. Int J Dev Biol. 2000; 44:619–626. [PubMed: 11061425]

Mobley AK, Braeuer RR, Kamiya T, Shoshan E, Bar-Eli M. Driving transcriptional regulators in melanoma metastasis. Cancer Metastasis Rev. 2012; 31:621–632. [PubMed: 22684365]

Raimbourg Q, Perez J, Vandermeersch S, Prignon A, Hanouna G, Haymann JP, Baud L, Letavernier E. The calpain/calpastatin system has opposing roles in growth and metastatic dissemination of melanoma. PLoS ONE. 2013; 8:e60469. [PubMed: 23565252]

Raposo G, Marks MS. Melanosomes–dark organelles enlighten endosomal membrane transport. Nat Rev Mol Cell Biol. 2007; 8:786–797. [PubMed: 17878918]

Sáez-Ayala M, Montenegro MF, Sánchez-Del-Campo L, Fernández-Pérez MP, Chazarra S, Freter R, Middleton M, Piñero-Madrona A, Cabezas-Herrera J, Goding CR, Rodríguez-López JN. Directed phenotype switching as an effective antimelanoma strategy. Cancer Cell. 2013; 24:105–119. [PubMed: 23792190]

Sandberg R, Ernberg I. Assessment of tumor characteristic gene expression in cell lines using a tissue similarity index (TSI). Proc Natl Acad Sci USA. 2005; 102:2052–2057. [PubMed: 15671165]

Sauka-Spengler T, Bronner-Fraser M. A gene regulatory network orchestrates neural crest formation. Nat Rev Mol Cell Biol. 2008; 9:557–568. [PubMed: 18523435]

Segura MF, Belitskaya-Levy I, Rose AE, Zakrzewski J, Gaziel A, Hanniford D, Darvishian F, Berman RS, Shapiro RL, Pavlick AC, et al. Melanoma microRNA signature predicts post-recurrence survival. Clin Cancer Res. 2010; 16:1577–1586. [PubMed: 20179230]

Stinson S, Lackner MR, Adai AT, Yu N, Kim HJ, O'Brien C, Spoerke J, Jhunjhunwala S, Boyd Z, Januario T, et al. TRPS1 targeting by miR-221/222 promotes the epithelial-to-mesenchymal transition in breast cancer. Sci Signal. 2011; 4:ra41. [PubMed: 21673316]

Strub T, Giuliano S, Ye T, Bonet C, Keime C, Kobi D, Le Gras S, Cormont M, Ballotti R, Bertolotto C, Davidson I. Essential role of microphthalmia transcription factor for DNA replication, mitosis and genomic stability in melanoma. Oncogene. 2011; 30:2319–2332. [PubMed: 21258399]

Uva P, Lahm A, Sbardellati A, Grigoriadis A, Tutt A, de Rinaldis E. Comparative Membranome expression analysis in primary tumors and derived cell lines. PLoS ONE. 2010; 5:e11742. [PubMed: 20668533]

Uz E, Alanay Y, Aktas D, Vargel I, Gucer S, Tuncbilek G, von Eggeling F, Yilmaz E, Deren O, Posorski N, et al. Disruption of ALX1 causes extreme microphthalmia and severe facial clefting: expanding the spectrum of autosomal-recessive ALX-related frontonasal dysplasia. Am J Hum Genet. 2010; 86:789–796. [PubMed: 20451171]

van Staveren WC, Solís DY, Hébrant A, Detours V, Dumont JE, Maenhaut C. Human cancer cell lines: Experimental models for cancer cells in situ? For cancer stem cells? Biochim Biophys Acta. 2009; 1795:92–103. [PubMed: 19167460]

Widmer DS, Cheng PF, Eichhoff OM, Belloni BC, Zipser MC, Schlegel NC, Javelaud D, Mauviel A, Dummer R, Hoek KS. Systematic classification of melanoma cells by phenotype-specific gene expression mapping. Pigment Cell Melanoma Res. 2012; 25:343–353. [PubMed: 22336146]

Zhang J, Hagopian-Donaldson S, Serbedzija G, Elsemore J, Plehn-Dujowich D, McMahon AP, Flavell RA, Williams T. Neural tube, skeletal and body wall defects in mice lacking transcription factor AP-2. Nature. 1996; 381:238–241. [PubMed: 8622766]

Zhang L, Huang J, Yang N, Greshock J, Megraw MS, Giannakakis A, Liang S, Naylor TL, Barchetti A, Ward MR, et al. microRNAs exhibit high frequency genomic alterations in human cancer. Proc Natl Acad Sci USA. 2006; 103:9136–9141. [PubMed: 16754881]

Highlights

- Similarity core analysis (SCA) is a bioinformatics tool for analyzing expression data
- SCA generates specific transcriptome-miRnome signatures for any tumor type
- SCA clusters aggressive and non-aggressive tumors and cell lines
- Molecular signatures reveal a lineage-specific regulatory network for melanoma

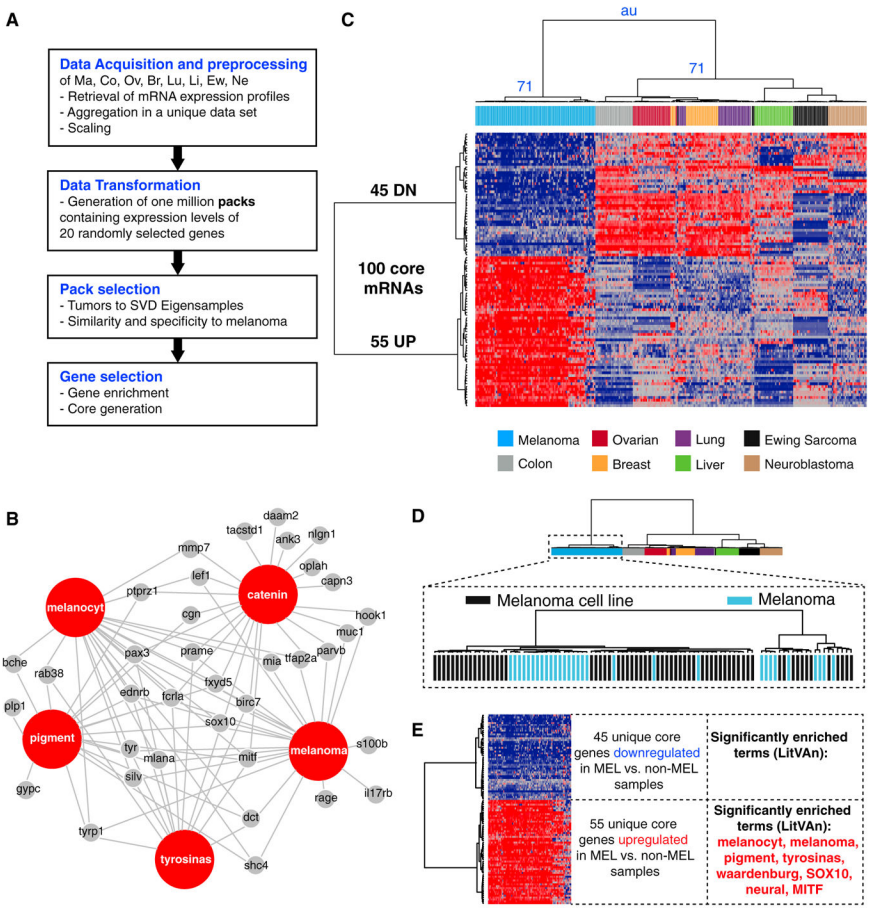


Figure 1. SCA Identifies a Melanoma-Specific Gene Expression Core

(A) SCA flowchart for mRNA. Gene expression data for eight different tumor types (Ma, melanoma; Co, colon; Ov, ovary; Br, breast; Lu, lung; Li, liver; Ew, Ewing sarcoma; Ne, neuroblastoma) were acquired and preprocessed to generate randomly selected contexts (RSCs) or packs of 20 arbitrarily chosen genes. SCA extracts cancer type-specific gene expression patterns for in vitro and in vivo comparisons, by measuring the proximity of cancer cell lines (melanoma) to tumors of the same type (melanoma) and comparing this proximity with that to other tumor types in PCA-based correlation spaces corresponding to one million RSCs. These spaces were screened with a criterion based on the minimal distance between melanoma cell lines and the melanoma centroid. The genes with the largest contribution to melanoma specificity were identified by calculating gene enrichment in the selected spaces and identifying a core of 100 genes (SCA signature).

(B) The SCA-melanoma signature is enriched in melanoma-related terms. Literature vector analysis (LitVAn) showed that 36 of the 100 core genes (gray) were linked to five significantly enriched terms (red) (Table S1). Note that the terms melanocyt and tyrosinas are deliberately truncated to act as wild cards.

(C) Unsupervised clustering of 240 tumor samples from eight different cancer types and 63 melanoma cell lines based on the SCA signature clearly discriminates between melanoma samples and other samples. The SCA-melanoma-mRNA signature consists of 55 overexpressed (55UP) and 45 underexpressed genes (45DN).

(D) The melanoma branch of the dendrogram shows the coclustering of melanoma cell lines (black) and melanomas (blue).

(E) Melanoma specificity is determined by 55UP genes. LitVAn was performed on the 45DN and 55UP core genes separately. The 55 UP core displayed enrichment for eight terms (Table S1).

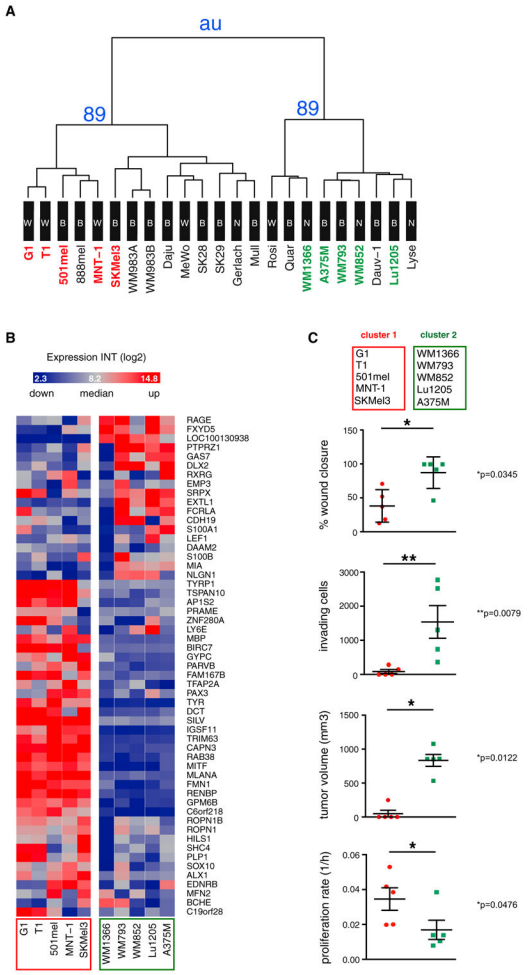


Figure 2. In Vitro Phenotypic Characteristics of Melanoma Cell Lines Are Correlated with SCA-Melanoma-mRNA Signature

(A) Of the independent melanoma cell lines, 23 not used for SCA were profiled and used as the validation set. These melanoma cell lines were clustered using the 55UP gene core. Two major clusters of cell lines were obtained. Confidence in clusters is indicated with bootstrap values (approximate unbiased p value, AU). The major classes are independent ($p = 0.4639$) of mutations of $BRAF^{V600E}$ (B), $NRAS^{Q61K,L,R}$ (N), or wild-type status for BRAF/NRAS (W). Five cell lines of cluster 1 (G1, T1, 501mel, MNT-1, and SKMel3) and five cell lines of cluster 2 (WM1366, WM793, WM852, Lu1205, and A375M) were chosen for further phenotypic analyses.

(B) 55UP expression heatmap shows the ten chosen melanoma cell lines ($2.3 < \log2$ expression intensity < 14.8).

(C) Phenotypic characterization of cluster 1 and 2 melanoma cell lines in vitro. Wound scratch analysis was performed as a surrogate for migratory capacity, Matrigel invasion assay for invasive capacity, subcutaneous tumor growth for tumorigenic potential, and doubling time for proliferation rate. In general, cluster 2 cell lines had greater migratory and invasive capacities and were more tumorigenic, but less proliferative than cluster 1 cell lines (Mann-Whitney test, $*p < 0.05$).

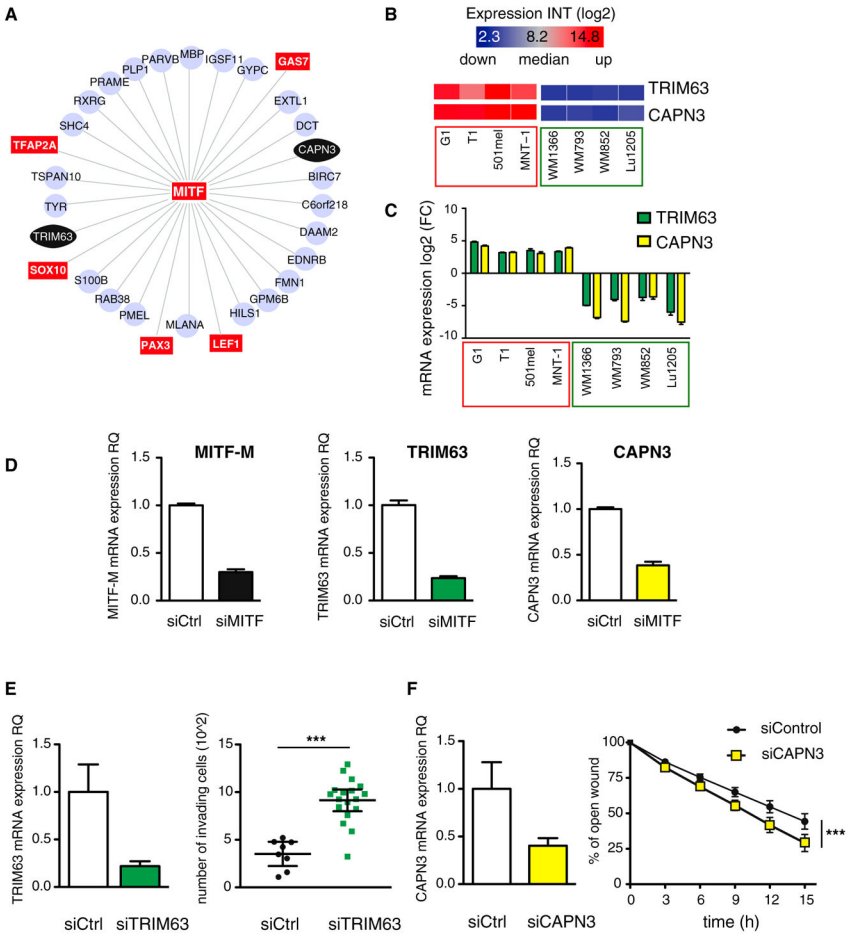


Figure 3. SCA-Melanoma-mRNA Genes *TRIM63* and *CAPN3* Are Potential MITF Targets Involved in Melanoma Cell Migration and Invasion

(A) The 55UP SCA signature was enriched in MITF target genes (30 such genes included; $p < 2.7E–23$, hypergeometric distribution test [Hoek et al., 2006]). Gene symbols are as follows: red, transcription factor; blue, other; and black, gene of further interest.

(B and C) *TRIM63* and *CAPN3* expression levels are higher in cluster 1 than in cluster 2 cell lines, according to (B) microarray and (C) qRT-PCR analysis (representative experiment shown, the error bar represents the SDV of technical triplicates).

(D) MITF downregulation by an siRNA approach decreases the levels of TRIM63 and CAPN3 mRNA.

(E) The repression of TRIM63 by an siRNA approach increases the Matrigel invasion capacity of 501mel melanoma cells after 48 hr. Unpaired t test with Welch correction, ***$p < 10^{-4}$.

(F) The repression of *CAPN3* in 501mel cells, by an siRNA approach, increases wound closure. Unpaired t test with Welch correction, ***$p < 10^{-3}$.

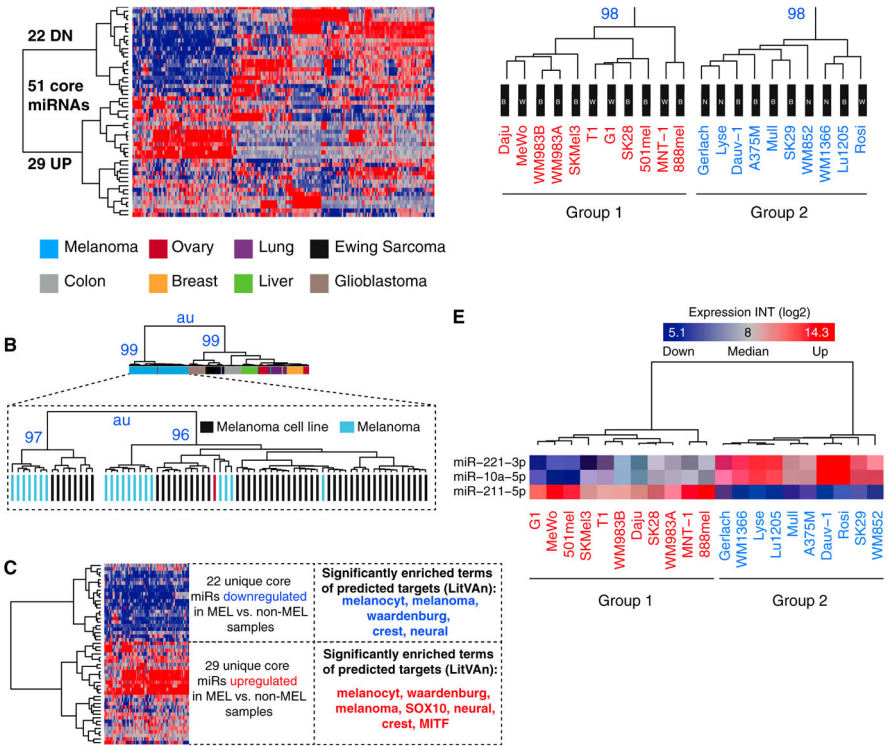


Figure 4. SCA Identifies a Melanoma-Specific miRNA Expression Core

(A) Unsupervised clustering of 168 tumor samples from eight different cancer types (Ma, melanoma; Co, colon; Ov, ovary; Br, breast; Lu, lung; Li, liver; Ew, Ewing sarcoma; Gl, glioblastoma) and 51 melanoma cell lines based on the SCA-melanoma-miRNA signature. Samples are color coded according to their origin. Melanoma samples (blue) cluster together and are well discriminated on the basis of expression of the 51 core miRNAs. The 51 core miRNAs can be separated into 22 miRNAs generally less strongly expressed in melanoma than in other cancers (22DN) and 29 miRNAs generally more strongly expressed in melanoma than in other cancer samples (29UP).

(B) The melanoma branch of the dendrogram shows the coclustering of melanoma cell lines (black) and melanomas (blue). Note the clustering of one ovary tumor (red) with the melanoma samples. The melanoma cell lines and tumors separate into two major groups, suggesting that the cell lines of group 1 are more representative of group 1 melanomas than group 2 melanomas, and vice versa.

(C) LitVAn was performed for the sequence-based predicted targets of the 22DN and 29UP miRNAs separately. The 29UP miRNAs were predicted to target 39 of the 100 core mRNAs and these miRNAs were enriched in the terms melanocyt, waardenburg, melanoma, SOX10, neural, crest, and MITF (Table S1). The 22DN miRNAs were predicted to target 30 of the 100 core mRNAs; these 30 mRNAs were analyzed by LitVAn and were found to be enriched in the terms melanocyt, melanoma, waardenburg, crest, and neural (Table S1).

(D) Of the independent melanoma cell lines not used for SCA, 21 were analyzed with the 29UP miRNA core. Two major clusters of cell lines were obtained. The major classes are independent ($p = 0.43$) of mutations of *BRAF*^{V600E} (B), *NRAS*^{Q61K,L,R} (N), or wild-type status for BRAF/NRAS (W).

(E) Differential miRNA analysis was performed between group 1 and group 2 melanoma cell lines with the 29UP miRNAs. Three miRNAs (221-3p, 10a-5p, and 211-5p) were found to be differentially expressed; miR-221-3p and miR-10a-5p were more strongly expressed in group 1 than in group 2 melanoma cell lines, whereas miR-211-5p was less strongly expressed in group 1 than in group 2 melanoma cell lines.

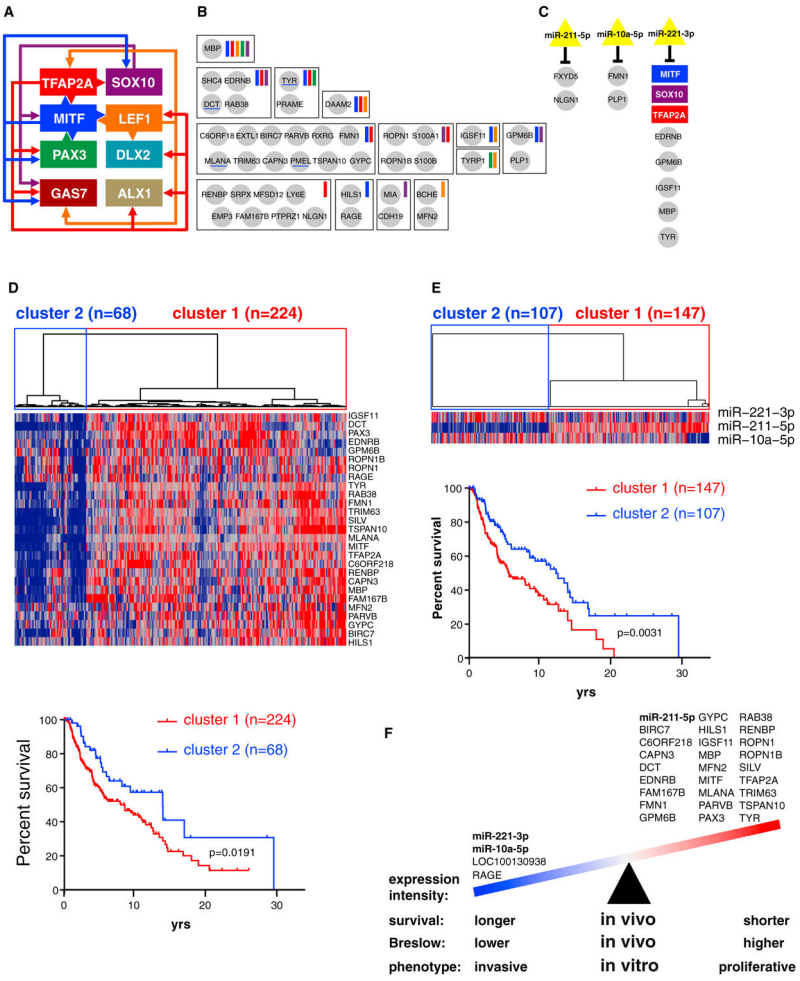


Figure 5. The Melanoma-Characteristic Gene Network Is Associated with Patient Survival

Interactions between genes/proteins of the melanoma-specific signature (55UP) were inferred from Ingenuity Pathway Analysis (IPA, literature-based) and publicly available ChIP-seq data (Table S1). The ALX1- and DLX2-related ChIP-seq data were not available at the time of publication.

(A) A core of eight transcription factors (TFs) characteristically overexpressed in melanoma and identified by SCA. Directed arrows of a specific color indicate that a specific TF regulates or potentially regulates other TFs. Potential regulation is evaluated according to the integration of ChIP-seq data, expression, and known interactions by IPA (see Table S1 and GEO: GSE67638). For instance, MITF (blue) regulates TFAP2A, SOX10, LEF1, PAX3, and probably also GAS7.

(B) Over 90% of the melanoma signature genes have been reported or predicted to be targeted by the core TFs. The color code corresponds to the eight TFs.

(C) Sequence-based prediction (mirDIP) and correlation of expression approaches used to integrate the melanoma-specific miRNA and mRNA signatures. Yellow triangles represent miR. Each potentially binds to sequences located in the 3′ region of the indicated mRNA, decreasing the levels of the corresponding mRNA.

(D) Melanoma metastases (SKCM cohort, TCGA portal) were clustered on the basis of the expression of the 28 core genes (27 of 28 genes were exploitable, LOC100130938 was not recognized) generally overexpressed in melanoma and SDE between aggressive and non-aggressive melanoma cell lines. The patients of the two principal clusters differed significantly in terms of overall survival ($p = 0.0191$, Mantel-Cox test).
(E) Melanoma metastases (TCGA portal) were clustered on the basis of the expression of the three core miRNAs generally overexpressed in melanoma and SDE between aggressive and non-aggressive melanoma cell lines. The patients of the two principal clusters differed significantly in terms of overall survival ($p = 0.0031$, Mantel-Cox test).
(F) Summary of the model linking molecular in vitro and in vivo features for melanoma based on SCA is shown.

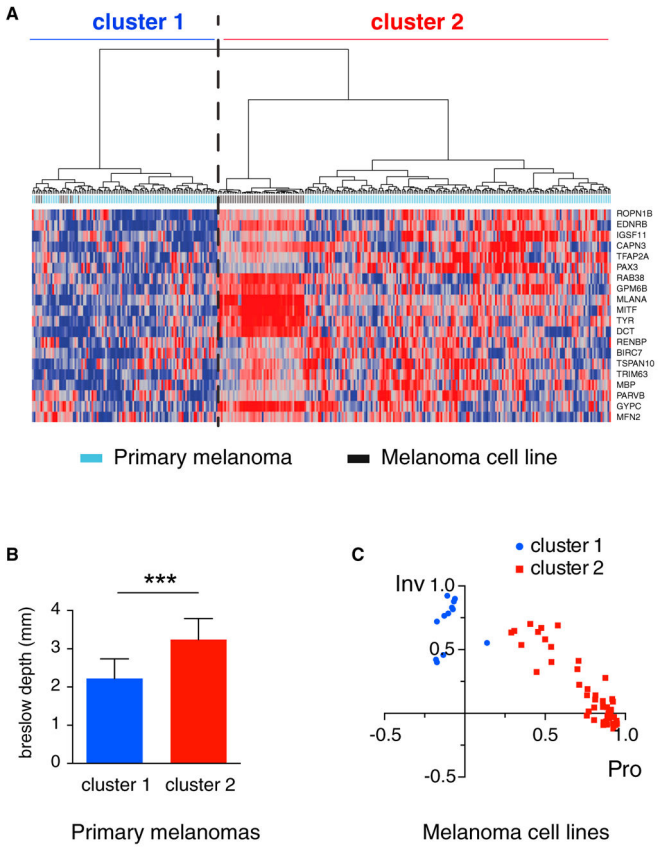


Figure 6. SCA Identifies Melanoma Cell Lines Being Representative of Primary Melanomas that Differ in Breslow Thickness

(A) Primary melanomas (n = 297, GEO: GSE57715) and melanoma cell lines (n = 63, GEO: GSE7127) were clustered using the SCA-MEL-mRNA-28SDE signature (20 of 28 genes were exploitable in this dataset). The dendrogram shows two distinct clusters containing each primary melanomas and melanoma cell lines.

(B) Primary melanomas of cluster 1 are generally of lower Breslow depth compared to cluster 2 (***$p < 0.0001$, Mann-Whitney test)

(C) Scoring of cluster 1 and cluster 2 melanoma cell lines using melanoma phenotype-specific expression (MPSE) correlation plot. In this plot, Pearson correlation coefficients against the proliferative (x axis) and invasive (y axis) standard signatures are plotted (Widmer et al., 2012). Perfect correlation of a melanoma cell line's expression signature with the standard invasive signature would result in a correlation coefficient (r) = 1 (y axis), and perfect correlation with the standard proliferative signature in r = 1 (x axis). According to the MPSE correlation plot, melanoma cell lines of cluster 1, which cocluster with primary melanomas of lower Breslow thickness, are most likely to exhibit an invasive phenotype.