Supplementary Methods

Sample preparation

A portion of pigmented lesions clinically suspicious for melanoma or previously confirmed as melanoma was sampled by punch biopsy at the time of excision. The punch biopsy was bisected; half was sent for permanent section paraffin-embedded histopathological diagnosis, and the other half (along with adjacent normal skin when available) was immediately snap-frozen in liquid nitrogen. The H&E slide was prepared from the snap-frozen tissue embedded in O.C.T. (Sakura Finetek USA, Torrance, CA). Slides were evaluated by a single dermatopathologist [DRF] who identified areas with >70% tumor cellularity, which were then sampled for RNA extraction. Frozen blocks were collected from 42 normal skin samples, 67 primary melanomas, and 20 melanoma metastases. Primary melanomas and melanoma metastases were derived from different patients. Melanoma metastases included lymph nodes (n=15), subcutaneous soft tissue (n=3), spleen (n=1), and small intestine (n=1). Details of 20 paraffin-embedded benign nevi used for immunohistochemical analysis have been previously described (Poynter *et al.*, 2006).

Isolation of RNA, DNA and microarray hybridization

Total RNA from frozen tissues was extracted using TRIzol protocol (Invitrogen, Carlsbad, CA) and Qiagen RNeasy purification kits (Qiagen, Germantown, MA). DNA was extracted from the TRIzol layer containing DNA with 0.1M sodium citrate precipitation followed by ethanol washing and final elution with 8mM NaOH, 1M HEPES, and 0.5M EDTA. Of the 129 samples in the initial (or "discovery") set, 76 samples (46 melanomas, 14 melanoma metastases, and 16 normal skin samples) had adequate quantities of high-quality total RNA determined by Agilent 2100 BioAnalyzer (Agilent Technologies, Palo Alto, CA). cDNA was prepared from 50 ng of total RNA using WT-Ovation™ RNA amplification (NuGEN Technologies, Inc., San Carlos, CA) and was hybridized to the Affymetrix Human Genome U133 PLUS 2.0 Array (Affymetrix, Santa Clara, CA) as per manufacturer's recommendations in the UofM Microarray Core. Arrays were scanned using GeneChip scanner. Expression values were calculated using MAS 5.0 (Affymetrix, Santa Clara, CA).

Melanoma cell lines A375, A2058, CHL-1, and SK-MEL-31 were purchased from ATCC (Manassas, VA). A375, A2058, and CHL-1 were grown in DMEM+10%FBS, and SK-MEL-31 in EMEM+10% FBS. Total RNA was extracted using TRIzol and DNA with Qiagen DNeasy Blood and Tissue kit from 80% confluent cells. RNA was quantified using NanoDrop (Thermo Scientific, Waltham, MA), and its quality was assessed by agarose gel electrophoresis.

Mutational analysis and quantitative RT-PCR

BRAF exon 15 and NRAS exons 1 and 2 were amplified using PCR primers: BRAF-F_5'-TTCATGAAGACCTCACAGTAAAAA-3'; BRAF-R 5'-TCCACAAAATGGATCCAGAC-3'; NRAS1-F_5'-GATGTGGCTCGCCAATTAAC-3'; NRAS1-R_5'-CCGACAAGTGAGAGACAGGA-3'; NRAS2-F_5'-CCCCTTACCCTCCACACC-3'; NRAS2-R_5'-CACAAAGATCATCCTTTCAGAGAA-3'. PCR conditions: 5 min at 95ºC; 40 cycles of 1 min at 95ºC, 1 min at 58ºC, and 1 min at 72ºC; and 10 min at 72ºC. Sequencing was performed in the University of Michigan Sequencing Core. Chromatograms were reviewed using Mutation Surveyor 2.61 and Sequencher 4.6. Quantitative real-time RT-PCR was performed on 69 of 74 arrayed samples because the RNA of four normal samples and one melanoma sample was exhausted. Real-time RT-PCR was performed in triplicates with GAPDH as an endogenous control using High Capacity cDNA Reverse Transcription Kit and SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA) on ABI 7900HT Sequence Detection System. The primers used: HMGA2_RT3_F_5'-TGTGGCCAATGGAACAGTAA-3'; HMGA2_RT3_F_5'-CGACCAACAACAGCAAAGAA-3'. Relative expression of HMGA2 was calculated by ΔCt normalization to the GAPDH expression and compared to the lowest expression level in normal skin samples.

Immunohistochemistry

Immunohistochemical staining was performed on the DAKO Autostainer (DAKO, Carpinteria, CA) using DAKO LSAB+ and 3,3'-diaminobenzidine as the chromogen. 5μm recuts were deparaffinized, labeled with anti-HMGA2 (goat polyclonal antibody, sc-23684; 1:200; Santa Cruz) and counterstained with

hematoxylin. Antibodies were validated on tissue microarrays containing positive (thyroid carcinoma) and negative (normal thyroid gland) controls. Thyroid carcinoma was shown to express high levels of HMGA2 (Belge *et al.*, 2008). Only nuclear immunoreactivity was scored. Intensity of expression was evaluated by a dermatopathologist [DRF] as follows: negative, weak, moderate, or strong. The percentage of HMGA2-stained tumor nuclei was scored as follows: no staining, <25% cells, 25%-50%, 50%-75%, and 75%-100%. For survival analysis, we dichotomized HMGA2 staining data for positive (any percentage of cells with positive intensity) and negative (no staining).

Replication TMA was stained using the same anti-HMGA2 (goat polyclonal antibody, sc-23684; 1:200; Santa Cruz) antibody as used in the original discovery set. Replication melanoma TMA was scored using tissue microarray and automated quantitative assessment of immunofluorescence (TMA-AQUA).

Statistical analysis

Quality was assessed using density plot of log-intensity and RNA degradation plot for each sample. For all 76 samples, signal intensities were quantile normalized using robust multiarray analysis (Irizarry *et al.*, 2003) and log2 transformed. Sample-specific median centering and scaling were applied by standard deviation. Probe sets with no expression and low variability across samples were excluded. Expression values were required to be above the lower quartile of all expression measurements in at least 25% of samples, and the interquartile range across the samples on log2 scale was required to be at least 0.5. After preprocessing and quality assessment, 74 samples and 44,137 probe sets were subjected to analysis.

Two class comparisons were conducted by two-sample t-tests for all comparisons. Multiple testing was addressed by using p-values controlled for overall false discovery rate (FDR) by the Benjamini and Hochberg method (BH-adjusted p-values). Local FDR was evaluated using locfdr (Efron, 2004). Probes' lists were selected using BH-adjusted p-values and ratio of mean expression values across the two groups.

To identify differentially expressed genes in three comparisons (PCM+MM vs. N; PCM vs. N; MM vs. N), we chose probes with BH-adjusted p-values <0.0001 and >1.8-fold differential expression ratio. The same BH-adjusted p-values <0.0001 but different fold change >1.5 was used for MM vs. PCM comparison because of the lower discriminatory power between these two groups based on expression profiles.

Transcriptome was annotated using NetAffx Analysis Center and classified by comprehensive exploration of data using NCBI (http://www.ncbi.nlm.nih.gov), Genecards (http://www.genecards.org), and Gene Ontology (http://amigo.geneontology.org). Gene set enrichment analysis was performed using DAVID (Dennis *et al.*, 2003) on gene lists with BH-adjusted p-value <0.05 and >1.5-fold change. Data have been deposited in NCBI's Gene Expression Omnibus (GSE15605 http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE15605).

References

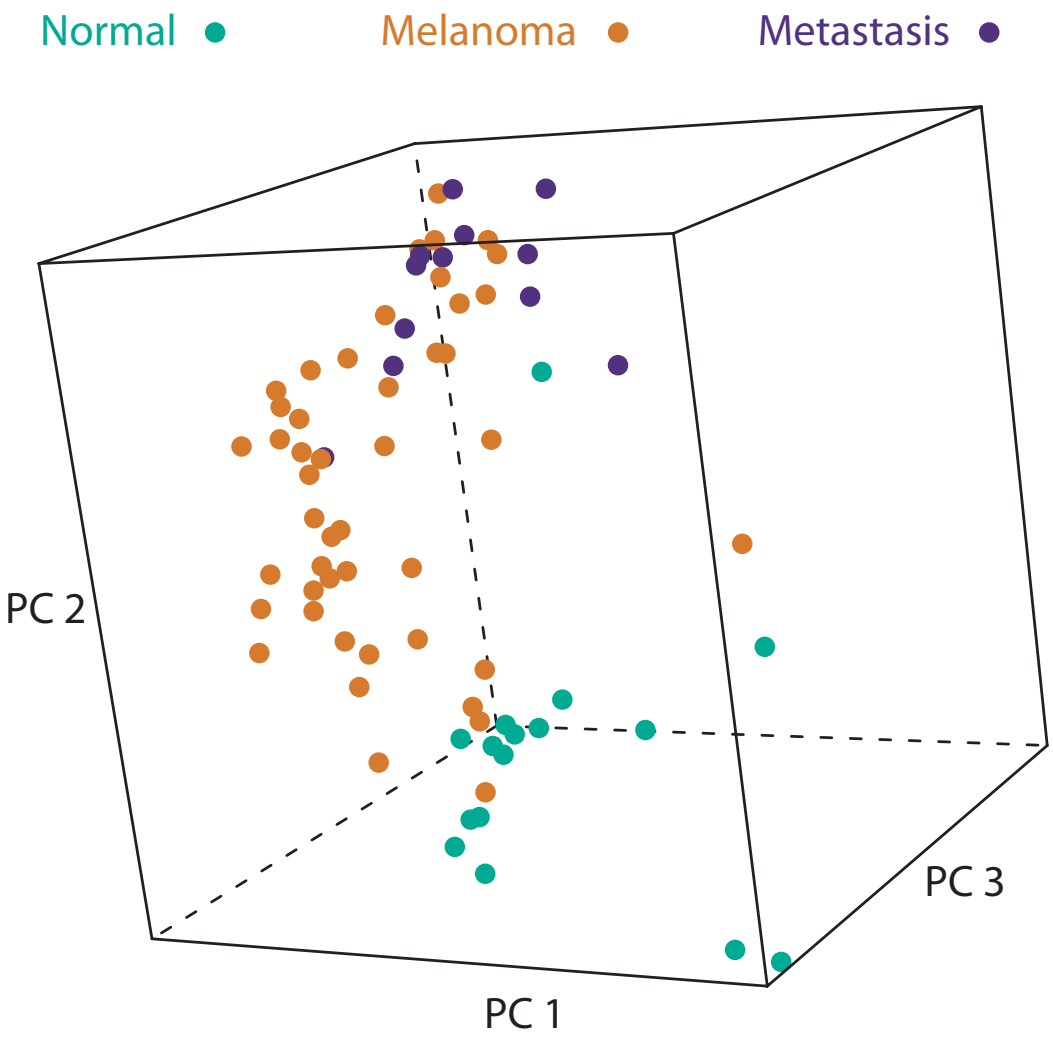
Belge G, Meyer A, Klemke M, Burchardt K, Stern C, Wosniok W*, et al.* (2008) Upregulation of HMGA2 in thyroid carcinomas: a novel molecular marker to distinguish between benign and malignant follicular neoplasias. *Genes Chromosomes Cancer* 47:56-63.

Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC*, et al.* (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome biology* 4:P3.

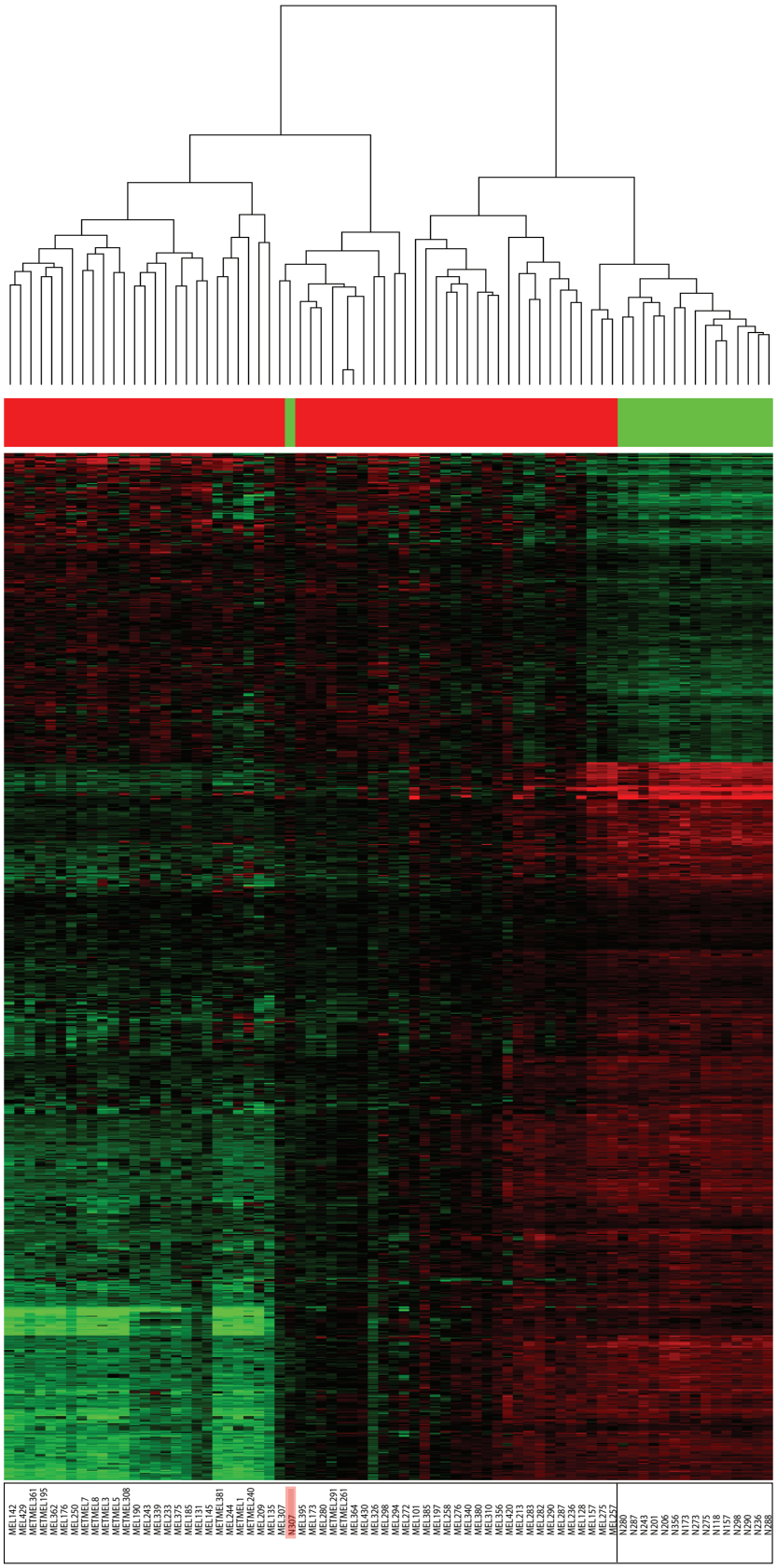
Efron B (2004) Large-scale simultaneous hypothesis testing: the choice of a null hypothesis. *Jour Amer Stat Assoc* 99:96-104.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U*, et al.* (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics (Oxford, England)* 4:249-64.

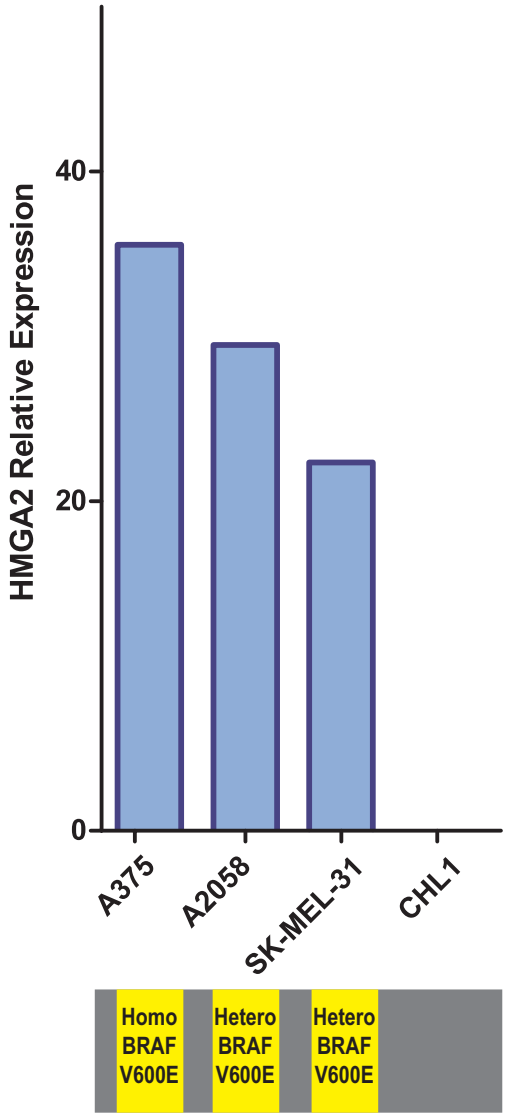
Poynter JN, Elder JT, Fullen DR, Nair RP, Soengas MS, Johnson TM*, et al.* (2006) BRAF and NRAS mutations in melanoma and melanocytic nevi. *Melanoma research* 16:267-73.



Supplementary Figure S1: Three-dimensional plot of the principal components (PC) from the expression data corresponding to 74 samples and 44,137 probes.



Supplementary Figure S2: Unsupervised agglomerative hierarchical clustering of gene expression profiles of the top 1000 differentially expressed genes.



Supplementary Figure S3: Quantitative RT-PCR analysis of HMGA2 expression in melanoma cell lines. A375 is homozygous, and A2058 and SK-MEL-31 are heterozygous for *BRAF* V600E mutation. CHL1 does not have mutations in codon 600 in *BRAF*.

Supplementary Table S1: Number of probes and genes differentially expressed in each two-class comparison.

Comparison	Number of probe sets								Number of genes	
	Unadjusted P<0.05		FC>1.8*		BH-p<0.0001		FC>1.8* BH-p<0.0001		FC>1.8* BH-p<0.0001	
	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
PCM+MM vs. N (58 vs. 16)	9205	8213	279	1002	1046	1654	238	800	164	571
PCM vs. N (46 vs. 16)	8354	7629	274	812	784	1175	216	575	150	411
MM vs. N (12 vs. 16)	9686	9096	597	1744	449	1620	193	1126	140	771
MM vs. PCM (12 vs. 46)	5877	6452	252	690	96	56	43	19	33	17

* Fold change threshold is 1.5 for MetM vs. PCM comparison. Fold change of 1.8 for upregulated probes corresponds to 1/1.8=0.55556 for downregulated probes. Fold change of 1.5 for upregulated probes corresponds to 1/1.5=0.66667 for downregulated probes.
PCM – primary cutaneous melanoma; MetM – metastatic melanoma; N – normal skin
BH-p – Benjamin and Hochberg adjusted p-value; FC – fold change.

Supplementary Table S2: Signal ratio between mutant and WT measured by Sanger sequencing.

Sample	BRAF	NRAS mutant/WT signal ratio	NRAS	NRAS mutant/WT signal ratio
MEL101	T/T		WT	
MEL128	T/T		WT	
MEL131	T/T		Q61R	1
MEL135	T/T		WT	
MEL142	V600R	1	WT	
MEL145	T/T		WT	
MEL157	T/T		WT	
MEL173	V600E	1	WT	
MEL176	V600E	1.33	WT	
MEL185	T/T		WT	
MEL190	V600E	1.17	WT	
METMEL195	V600E	1.86	WT	
MEL197	T/T		Q61R	0.33
MEL209	V600E	0.88	WT	
MEL213	V600E	1.88	WT	
MEL233	T/T		Q61K	0.53
MEL236	L597S	1	WT	
METMEL240	V600E	6	WT	
MEL243	V600E	1	WT	
MEL244	V600E	2	WT	
MEL250	T/T		WT	
MEL257	T/T		WT	
MEL258	T/T		WT	
METMEL261	T/T		WT	
MEL272	V600E	1	WT	
MEL275	T/T		WT	
MEL276	T/T		WT	

MEL280	V600E	2	WT	
MEL282	T/T		WT	
MEL283	T/T		WT	
MEL287	T/T		WT	
MEL290	V600E	0.8	WT	
METMEL291	T/T		Q61K	0.77
MEL294	T/T		WT	
MEL298	V600K	0.7	WT	
MEL307	V600K	0.7	WT	
METMEL308	V600E	0.8	WT	
MEL310	T/T		Q61R	1
MEL326	V600E	1	WT	
MEL339	V600E	0.8	WT	
MEL340	T/T		WT	
MEL356	V600E	6	WT	
METMEL361	T/T		WT	
MEL362	V600E	1.43	WT	
MEL364	T/T		WT	
MEL375	T/T		Q61L	0.5
MEL380	V600E	0.8	WT	
METMEL381	V600E	3	WT	
MEL385	T/T		WT	
MEL395	T/T		WT	
MEL420	T/T		WT	
MEL429	T/T		Q61L	2
MEL430	L597R	0.7	WT	
METMEL2007-01	T/T		Q61L	2.3
METMEL2007-03	V600E	6	WT	
METMEL2007-05	V600E	0.7	WT	
METMEL2007-07	V600E	6	WT	
METMEL2007-08	V600K	2	WT	

Mean mutant/WT signal ratio		1.91		1.05