SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Whole Exome Sequence (WES) Analysis. Single nucleotide variant (SNV), small insertion-deletion (INDEL) and copy number variation (CNV) calling were performed as in our prior study (Shi et al., 2014). Tumor purities and ploidies, as listed in Supplementary Table S1E, were calculated based on the purity and ploidy calls of Sequenza (Favero et al., 2015) on the WES data using its default parameters. When there were multiple possible (purity, ploidy) values, we manually chose the pair that minimized the ploidy variation across the different tumors from the same patient.

RNASeq Analysis. To ensure robustness in the DGE calls, we used three programs: Cuffdiff (Trapnell et al., 2013), GFOLD (Feng et al., 2012) and RSEM (Li and Dewey, 2011) + EBSeq (Leng et al., 2013). Using Cuffdiff, the paired-end reads were mapped to the UCSC hg19 reference genome using Tophat2 (Kim et al., 2011). DGE was called by with the option *"--frag-bias-correct"* and *"--multi-read-correct"* to improve sensitivity (Roberts et al., 2011). Each sample was run on two different sequencing lanes serving as replicates for computing p-values of differential gene expression (DGE). When there was no lane replicates (for samples Pt #18 baseline, DD-DP and Pt #19 DD-DP), the p-value of the differential expression was called by the *blind* mode. Using *GFOLD*, we supplied the Tophat2 computed BAM alignment files of each lane and computed the adjusted fold change of each genes in the DP or DD-DP samples compared to their patient-matched baseline samples. For a few selected cases where there were no replicates (as noted above), the p-value of an observed fold change was estimated

using a null distribution list of GFOLD fold change values computed from technical replicates of an in-house cell line RNASeq data set. Using RSEM+EBSeq, we first quantified the expression levels in each sample by supplying the unmapped FASTQ files to RSEM's *rsem-calculate-expression* function. Then the expected count output by the previous step was passed to EBSeq using RSEM's *rsem-run-ebseq* function. DGE call for all three programs was defined as a two-fold increase or decrease in the expression of a gene with a FDR adjusted q-value less than or equal to 0.05 ($|\log_2 FC| \geq 1$ with q-value ≤ 0.05). A gene was defined to be differentially expressed when two of the three programs called the gene to be differentially expressed in the same direction. The microarray data of patient #10 was analyzed using the Affymetrix HuGene 2.1 microarray using the Oligo R package (Carvalho and Irizarry, 2010). Microarray RNA expression in the validation data set (Illumina HumanHT-12 V4.0 expression beadchip) was processed by normalization using the *normaliseIllumina* function in the beadarray R package. DGE for microarray-based data between acquired resistant and baseline melanomas was defined by a 1.5-fold up- or down-expression in the acquired BRAFi resistant tumors as compared to its respective baseline. When analyzing immune gene up- and down-expression from RNASeq, we used a 1.5-fold cutoff as the immune compartment was expected to be smaller than the tumor cell compartment. For analysis of the microarray data set, we relaxed the 1.5-fold cutoff to 1.25-fold for immune gene expression analysis.

In addition to DGE, we also tabulated the normalized gene expression levels of each baseline and DP/DD-DP sample. The normalized expression level of genes was expressed in FPKM values as generated by the program Cuffnorm. We applied the options *"--frag-bias-correct"*, *"--multi-read-correct"* and "*--compatible-hits-norm*" in the Cuffnorm run to ensure that both Cuffdiff and Cuffnorm applied the same adjustments and denominator for FPKM computation. Expressed SNVs or INDELs were defined as those found in genes that had (Cuffnorm) FPKM value ≥ 0.1 in the DP or DD-DP samples. When the SNV/INDEL was expected to be a loss-of-function (LOF) event, such as nonsense, frameshift and splice site mutations (expected to result in a truncated and, potentially, degraded mRNA), the expression cutoff of FPKM ≥ 0.1 was applied on the baseline expression of the gene to ensure that the gene was expressed prior to the LOF event. CNV-related DGE events were defined as concurrent copy number gain (called by both ExomeCNV and ExomeDepth) and mRNA up-expression ($\log_2$ FC ≥ 1 with q-value ≤ 0.05) or copy number loss and mRNA down-expression ($\log_2$ FC ≤ -1, q-value ≤ 0.05).

We curated CD8 T-cell exhaustion gene lists (Wherry, 2011) to minimize those likely to be expressed by melanoma cells by excluding genes whose maximum $\log_2$ FPKM was ≥ 1 in our in-house melanoma cell line-derived RNAseq database (n=26 cell lines, data not shown). This resulted in the inclusion of surface receptors *PDCD1* (*PD-1*). *LAG3*, *HAVCR2* (*Tim-3*), *CD160*, and *CD244* as well as the transcription factors *EOMES*, *PRDM1* (*Blimp-1*) and *TBX21* (*T-bet*). We also included *FOXP3* (a regulatory T-cell

transcription factor marker gene) to assess the ratio of FOXP3+ CD4 regulatory T-cells to CD8 T-cells. We excluded *CTLA4* and *NFATC1* (a T-cell exhaustion transcription factor gene like EOMES) since they were expressed in some of our melanoma cell lines.

Gene Set Enrichment Analysis. As there were no available gene-set enrichment analysis specifically catered to patient-matched, paired RNASeq data sets, we devised an algorithm to estimate the enrichment of a gene set based on the fold change of each gene in the set. Specifically, for each sample, we rank ordered the genes based on their fold change in expression from baseline to disease progression. Gene sets with many high-ranked up-expressed genes would have a smaller rank-sum while sets with negative enrichment would have a larger rank-sum. We tested for significance of the rank-sum by using the Wilcoxon rank-sum test and set a p-value cutoff ≤ 0.05. We additionally required that the median of up/down-expression across all genes in the gene set to be at least 10% (i.e., at least half of the samples have up/down-expression larger than or equal to 10%). This analysis was performed using the C6, C7, C2 CGP gene sets from the Molecular Signature Database of the Broad Institute. Additionally, we applied this enrichment analysis on our TCGA Melanoma data set-derived gene signatures for *c-MET*, *LEF1* and on the DENDRITIC CELL MARKERS signature.

For single sample gene set enrichment calculations, we used the GSVA program (Hanzelmann et al., 2013) to derive the absolute enrichment scores (e.g for the *c-MET*

gene signatures in Figure 2B). RNASeq read counts were computed using HTSEQ-COUNT program and were converted to $\log_2$ CPM values using EdgeR (McCarthy et al., 2012) as input for GSVA (run in RNAseq mode). The GSVA scores on the validation dataset were computed based on the RMA normalized expression values as input to the GSVA program run in the microarray mode. GO enrichments analysis was done using the online functional annotation tools DAVID (Huang da et al., 2009b). GO term enrichments were selected from the GO biological process terms in DAVID's Fat database (Huang da et al., 2009a).

Expression-correlated CpG Cluster Analysis. To identify the differentially methylated CpG sites, we analyze the Illumina 450K Methylation array output data using R-package minfi (Aryee et al., 2014). Raw image IDAT files were loaded into the R statistical computing environment using minfi, and the data was normalized using the SWAN function. This output a matrix of methylation indices (β), which were continuous values between 0 and 1 representing the ratio/fraction of the intensity of the methylated-probe signal to the total signal intensity for each probed CpG site. For each site, the methylation change was measured by the percent methylation difference ($\Delta\beta$) from paired baseline to DP/DD-DP samples. Minfi function *dmpFinder* was then applied to calculate the p-value of the logit transformed differential methylation. The p-values were corrected for multiple hypothesis testing with false discovery rates (FDR) q-values. We then used a cutoff of q-value ≤ 0.05 and absolute percent methylation change $|\Delta\beta| \geq 10\%$ to define differential methylation.

To identify CpG clusters with multiple consecutive CpG methylation sites co-regulating their target gene expression (Sofer et al., 2013), we defined a threshold of 3000 base pairs for two CpG sites to be considered in the same CpG cluster. Specifically, we computed the CpG clusters by the following steps:

1) For each gene G, we started with the CpG site C_i and calculated the Pearson's correlation coefficient (r_i) between the percent methylation change β_i of C_i and the G's expression $\log_2$ fold change E_G. We first excluded genes and CpG sites which were altered significantly ($q \leqslant 0.05$) in less than 25% of the DP/DD-DP samples. These genes or CpG sites would be dominated by minor changes which may not produce meaningful correlation values. We also excluded genes whose expressions altered significantly ($q \leqslant 0.05$) only in less than 25% of the DP/DD-DP samples for the same reason. We set the value of percent methylation change β_i and expression change E_G to be zero when the q value was > 0.05. Pearson correlation coefficient was computed only on samples with non-zero values on either β_i or E_G, and the number of such samples had to be at least half of the total samples.
2) If there existed an adjacent site C_j within 3000 base pairs downstream of C_i, we calculated the median of the percent methylation changes of C_i and C_j across the samples. This combined percent methylation change was denoted as β_{i+j} and we computed the correlation r_{i+j} between β_{i+j} and E_G.

3) If $|r_{i+j}| > |r_i|$, we repeated step 2 to test if combining the next adjacent site downstream of C_j would improve the correlation further.
4) Otherwise, we defined the current best combination starting from C_i as one CpG cluster and repeated step 1 starting from the next CpG site C_j

We defined the correlation score of each CpG cluster to be its Pearson correlation coefficient values. For a gene with multiple CpG clusters, the p-values of the correlation score computed for each CpG cluster were corrected for multiple hypothesis testing with false discovery rates (FDR) q-values across all tested CpG clusters on the gene G. CpG clusters with q-value $\leqslant$ 0.1 were defined as the (expression) correlated CpG clusters of the gene. When a correlated CpG cluster was a subset of another (larger) CpG cluster, we only retained the one with a better correlation score. When there was a tie in the correlation scores, the smaller cluster was retained.

After identifying the correlated CpG clusters of each gene using overall correlation tests across all DP/DD-DP samples, we assessed their association with DGE events in each resistant melanoma sample. For each sample, we selected CpG clusters with significant differential methylation (q-value ≤ 0.05, $|\triangle\beta|$ ≥ 10%) and significant DGE (q-value ≤ 0.05, $|\log_2 FC| \geq 1$). For each of the selected CpG clusters, we assessed whether the direction of the changes of the methylation and RNA expression was concordant with the overall correlation between the CpG cluster and gene expression across all samples (i.e., changes of expression and methylation going in opposite directions when the correlation was negative or changes in the same direction with positive correlation). If

this were true, we nominated the DGE event to be potentially driven by the differential methylation event of the correlated CpG cluster.

GOF and LOF Cancer Gene Events. A list of 855 cancer related genes was compiled in Supplementary Table S1B along with the publication sources of the genes. We defined RNA up-expression (driven by CNV, methylation or other unknown factor) to be a GOF event. SNV and INDEL mutations that have been shown to be activating oncogenes also fell in this category. On the other side, RNA down-expressions, nonsense, frameshift, splice site mutations were defined as LOF events along with missense SNVs and in-frame INDELs that have been shown to cause LOF of tumor suppressor genes. Uncharacterized SNV and INDEL events were considered for calculation in both GOF and LOF lists. Alternative splicing of *BRAF* was specifically detected as per our previous studies (Shi et al., 2014; Shi et al., 2012) and was annotated as a GOF event.

Analysis of the TCGA Melanoma Data. We analyzed The Cancer Genome Atlas (TCGA) Skin Cutaneous Melanoma RNASeq dataset (http://cancergenome.nih.gov/, download date Nov 20th, 2013). For RNA expression analysis, we used the level 3 gene based normalized expression data. To generate *c-MET* and *LEF1* UP and DN signatures, we collected genes that are up- or down-expressed in melanoma samples in the top and bottom quartiles of the expression level of the reference gene (the gene we are interested to derive signature of). Specifically, we chose all genes Y which were

differentially expressed between the samples in the top and bottom quartiles of reference gene X using the two-sided Wilcoxon rank-sum test on the normalized expression values of Y (log_2). To reduce the number of hypothesis testing, we only tested genes Y whose median expressions in the two quartiles differed by at least four fold. We corrected for multiple hypothesis using FDR-adjusted q-values and reported the genes Y with q-values ≤ 0.05 as the signature genes of X. For each gene X, we generated the TCGA_All signature and TCGA_BRAF mutant signature for X where the former was generated using all TCGA samples with RNAseq data (n=356) and the latter was generated based on the subset of the TCGA samples with BRAF mutation (n=182). The signature genes of *c-MET* and *LEF1* were listed in Supplementary Table S2 and S3.

Differential patient survival was computed based on the 10-year survival data of the TCGA melanoma series of patients (available for 247 patients; *BRAF* mutant only, 118 patients) with high expression (top quartile) vs. low expression (bottom quartile) of the genes of interest (i.e., the top 30 GOF and LOF genes along with reported MAPKi resistance-related genes). Significance was tested using log-rank test p-value. We examined possible confounding factors such as age, tumor ulceration and AJCC stage by including them in a Cox proportional hazard model. Specifically, AJCC tumor stage II, III, and IV, regardless of their subcategories A, B and C, were grouped under nominal class ‘Stage II’, ‘Stage III’, and ‘Stage IV’, respectively. Stage 0, Stage I and Stage II NOS were grouped under ‘Stage I’. Age (in days) was examined as a continuous

variable. Ulceration was included as a binary 'Y' and 'N' nominal variable. A GOF event gene has to have significant hazard ratio > 1 ($p \leq 0.05$) for the top quartile patients and has differential survival log-rank p-value ≤ 0.05 to be defined to associate with worse survival. Similarly, an LOF gene event must have hazard ratio < 1 ($p \leq 0.05$) for the bottom quartile patients and have differential survival log-rank p-value ≤ 0.05.

To validate the expression-correlated CpG clusters we identified, we interrogated the TCGA Skin Cutaneous Melanoma data (http://cancergenome.nih.gov/), which includes both RNASeq expression and methylation data (for 335 tumor samples). For each CpG cluster, we first tested if its β value varies by at least 10% across all the samples (CpG clusters with multiple CpG sites were represented by the median of the β values of the CpG sites). If this criterion were met, we then selected the samples in top and bottom quartiles of the methylation level of the CpG cluster of interest and compared the expression of the CpG's annotated target gene in the two quartiles by wilcoxon rank sum test. Significance was defined by a difference of at least two-fold in the median FPKM expression levels between the two quartiles and one-sided wilcoxon rank-sum test p-value ≤ 0.05 (i.e., the direction of the test being defined by the overall correlation value of the CpG cluster in our data set; e.g., negative correlation would require higher expression level in the bottom quartile and vice versa). Similar analysis was also performed based on the top and a bottom quartile expressions of *c-MET*'s transcription factors.

Western Blot Antibody Reagents. p-ERK1/2 (T202/Y204), p-MEK1/2 (S217/221), p-AKT (T308/S473), p-c-MET (S338), total ERK1/2, MEK1/2, AKT, c-MET, p-β-catenin (S33/37 and T41), LEF1, p-YAP1 (S127), PARP1, cleaved PARP1 (D214), MEK1, BIM, GAPDH (Cell Signaling Technology), β-catenin (BD Bioscience), YAP1 (Santa Cruz), and TUBULIN (Sigma).

SUPPLEMENTAL REFERENCES

Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D., and Irizarry, R.A. (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics *30*, 1363-1369.

Carvalho, B.S., and Irizarry, R.A. (2010). A framework for oligonucleotide microarray preprocessing. Bioinformatics *26*, 2363-2367.

Favero, F., Joshi, T., Marquard, A.M., Birkbak, N.J., Krzystanek, M., Li, Q., Szallasi, Z., and Eklund, A.C. (2015). Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. Ann Oncol *26*, 64-70.

Feng, J., Meyer, C.A., Wang, Q., Liu, J.S., Shirley Liu, X., and Zhang, Y. (2012). GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. Bioinformatics *28*, 2782-2788.

Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics *14*, 7.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res *37*, 1-13.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2011). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Leng, N., Dawson, J.A., Thomson, J.A., Ruotti, V., Rissman, A.I., Smits, B.M., Haag, J.D., Gould, M.N., Stewart, R.M., and Kendziorski, C. (2013). EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics *29*, 1035-1043.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323.

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res *40*, 4288-4297.

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L., and Pachter, L. (2011). Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biol *12*, R22.

Shi, H., Hugo, W., Kong, X., Hong, A., Koya, R.C., Moriceau, G., Chodon, T., Guo, R., Johnson, D.B., Dahlman, K.B.*, et al.* (2014). Acquired Resistance and Clonal Evolution in Melanoma during BRAF Inhibitor Therapy. Cancer Discov *4*, 80-93.

Shi, H., Moriceau, G., Kong, X., Lee, M.K., Lee, H., Koya, R.C., Ng, C., Chodon, T., Scolyer, R.A., Dahlman, K.B.*, et al.* (2012). Melanoma whole-exome sequencing identifies (V600E)B-RAF amplification-mediated acquired B-RAF inhibitor resistance. Nat Commun *3*, 724.

Sofer, T., Schifano, E.D., Hoppin, J.A., Hou, L., and Baccarelli, A.A. (2013). A-clustering: a novel method for the detection of co-regulated methylation regions, and regions associated with exposure. Bioinformatics *29*, 2884-2891.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol *31*, 46-53.

Wherry, E.J. (2011). T cell exhaustion. Nat Immunol *12*, 492-499.