Cell Reports, Volume *21*

Supplemental Information

Melanoma Therapeutic Strategies that Select against Resistance by Exploiting MYC-Driven Evolutionary Convergence

Katherine R. Singleton, Lorin Crawford, Elizabeth Tsui, Haley E. Manchester, Ophelia Maertens, Xiaojing Liu, Maria V. Liberti, Anniefer N. Magpusao, Elizabeth M. Stein, Jennifer P. Tingley, Dennie T. Frederick, Genevieve M. Boland, Keith T. Flaherty, Shannon J. McCall, Clemens Krepler, Katrin Sproesser, Meenhard Herlyn, Drew J. Adams, Jason W. Locasale, Karen Cichowski, Sayan Mukherjee, and Kris C. Wood

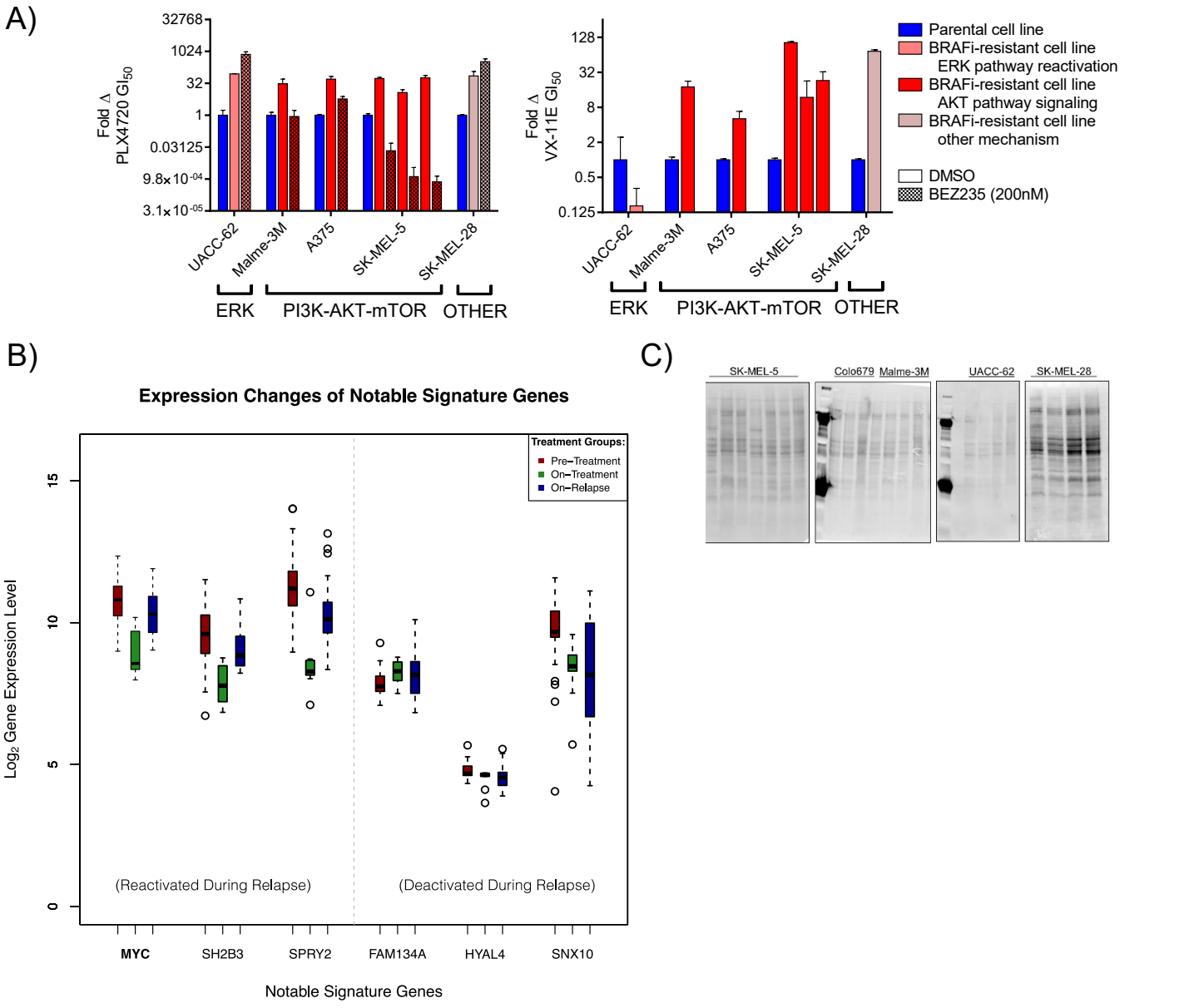


Figure S1. Characterization of MYC activation in melanomas with acquired resistance to BRAF pathway blockade, related to Figure 1. A) Evolved PLX4720 (BRAFi)-resistant cells that could be re-sensitized by the addition of the PI3K/mTOR inhibitor, BEZ235, were classified as having a functional PI3K-AKT-mTOR pathway bypass signaling resistance mechanism. Cells that were resistant to PLX4720 but not also resistant to the ERK inhibitor, VX-11E, were classified as having a functional RAF-ERK signaling re-activation resistance mechanism. Cells that could be fully re-sensitized through stable shRNA-mediated Notch1 suppression were classified as having a Notch1 pathway resistance mechanism and are described elsewhere (Martz, Ottina, Singleton *et al.*, *Sci Signal* 2014, **7**, ra121). One evolved resistant cell line could not be re-sensitized by VX-11E, BEZ235, or shNotch1 and was therefore classified as other/unknown. In all cases, fold change values are relative to parental cells (blue), and data are means (+/- SD) from three experiments. **B)** Expression of indicated genes from the BRAF response signature in cell lines and primary tumors at indicated time points. Error bars indicate 90% confidence intervals and open circles are outliers. **C)** Total protein staining for immunoblots shown in Figure 1F, where for each cell line the left to right ordering of samples is the same as in Figure 1F.

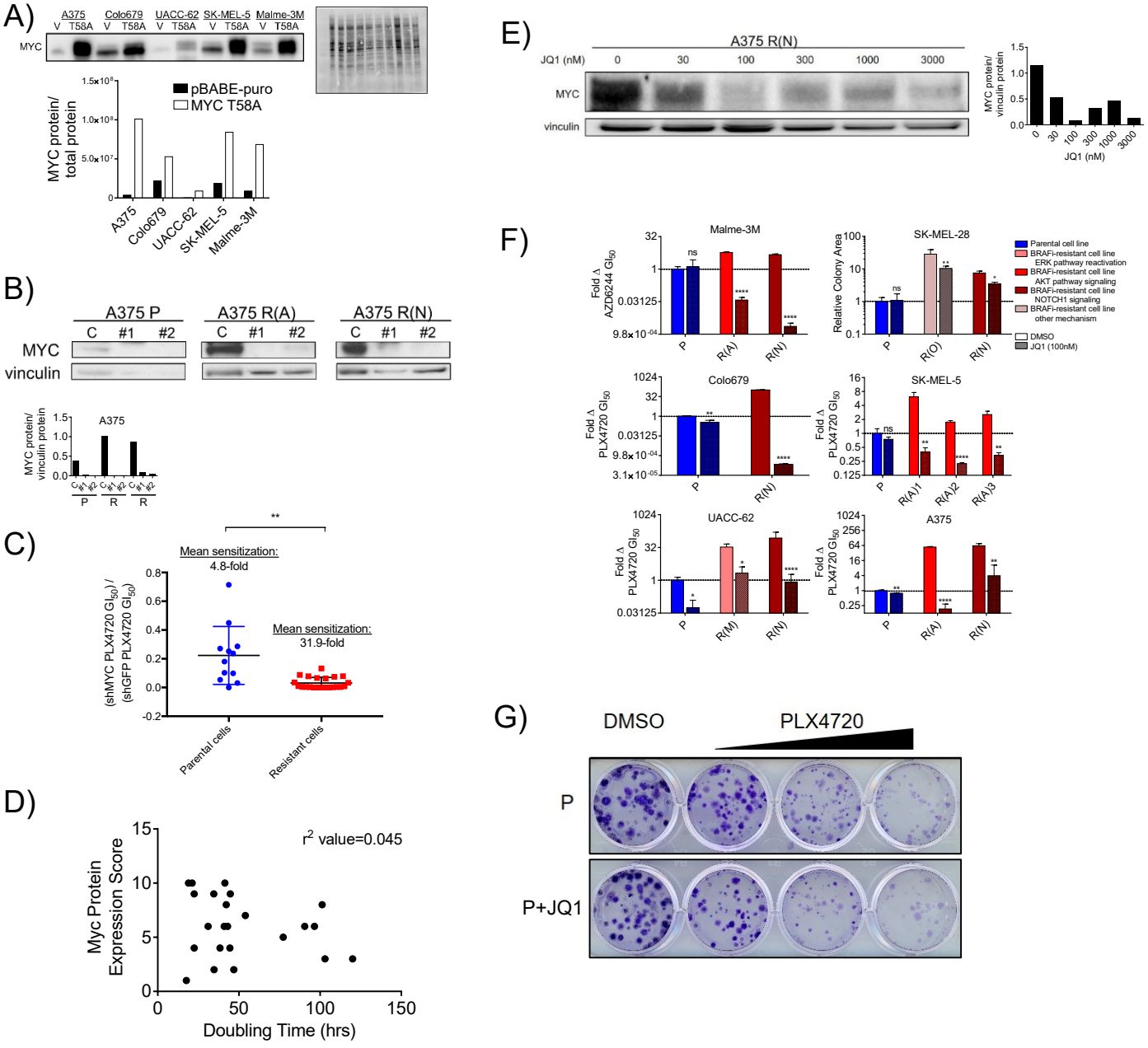


Figure S2. Functional characterization of MYC in BRAFi resistant melanomas , related to Figure 2. A) Ectopic expression of MYC^{T58A} or empty vector control in treatment-naïve, BRAFi-sensitive *BRAF* mutant melanoma cell lines, determined by immunoblotting. Quantification, bottom panel; total protein, right top panel; V, pBABE empty vector. **B)** The indicated cell lines (P, parental and R, resistant), expressing either shGFP (C) or one of two independent shRNAs targeting MYC (#1 or #2), were immunoblotted for expression of MYC and vinculin as a loading control to validate protein knockdown. Quantified, bottom panel. **C)** The ratio of GI50 values for PLX4720 in parental and resistant cell lines expressing shMYC (hairpins 1-2) and shGFP. **D)** MYC protein level was scored on a scale of 1-10, with 10 indicating highest MYC protein expression and correlated to cell line doubling time. **E)** A375 R(N) cells were treated with the indicated doses of JQ1 for 72 hours. Whole cell lysates were immunoblotted for expression of MYC and vinculin as a loading control. Quantified, right panel. **F)** Evolved PLX4720-resistant cell lines could be re-sensitized to PLX4720 or AZD6244 (Malme-3M) by the addition of JQ1 (100 nM) as determined by GI_{50} assay or clonogenic growth assay (in 3 μM PLX4720). JQ1 + BRAFi treated wells are normalized to the values of the relevant JQ1 only treated wells to account for non-specific toxicity. *P* values denote significance between DMSO and JQ1 treatment in each cell line. **G)** Effect on clonogenic growth of A375 cells (P) treated with increasing doses of PLX4720 in the presence or absence of JQ1 (100nM). Data are means (SD) from three experiments. **P* < 0.05; ***P* < 0.01; *****P* < 0.001.

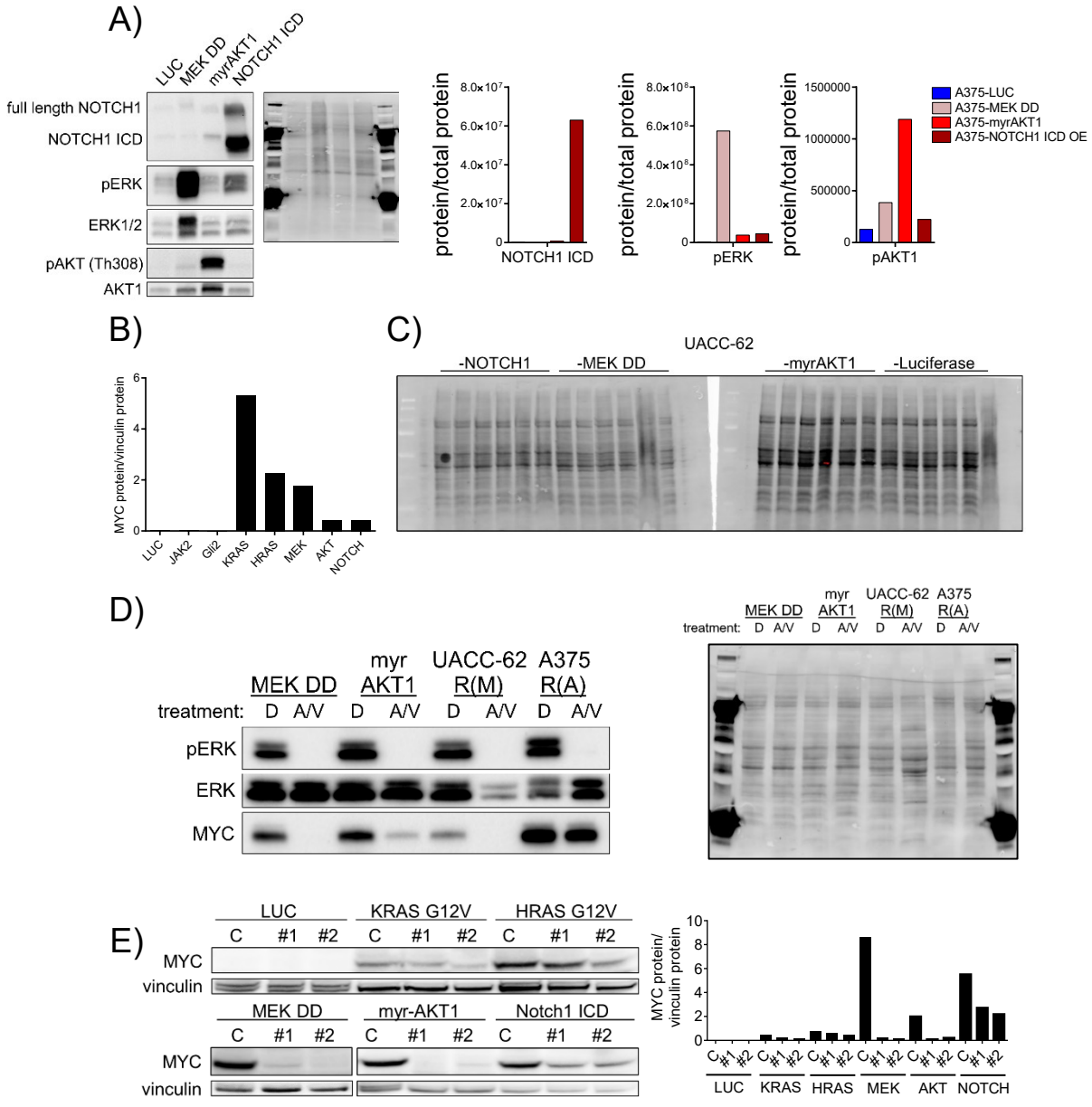


Figure S3. MYC activation and suppression in melanoma cells with engineered activation of major resistance pathways , related to Figure 3. A) A375 cells expressing the pathway activating constructs indicated were immunoblotted for Notch1, phospho-ERK and phospho-AKT1 (Th308) as evidence of activation of each pathway (left panel). Total protein is shown in the center panel and protein expression normalized to total protein is shown in the right panel. **B)** Quantification of immunoblots shown in Figure 3A. **C)** Total protein staining for immunoblots shown in Figure 3C. **D)** Immunoblot of MYC, phospho-ERK and ERK1/2 levels in the indicated cell lines treated with DMSO or 1 μM AZD6244 (“A”, a MEKi) and 1 μM VX-11E (“V”, an ERKi) for 24 hours. At right, total protein staining. **E)** UACC-62 cells expressing luciferase or the indicated pathway activating construct in combination with shGFP (C) or two independent shMYC constructs (#1 or #2) were immunoblotted for expression of MYC and vinculin as a loading control. Blots are quantified in the right panel.

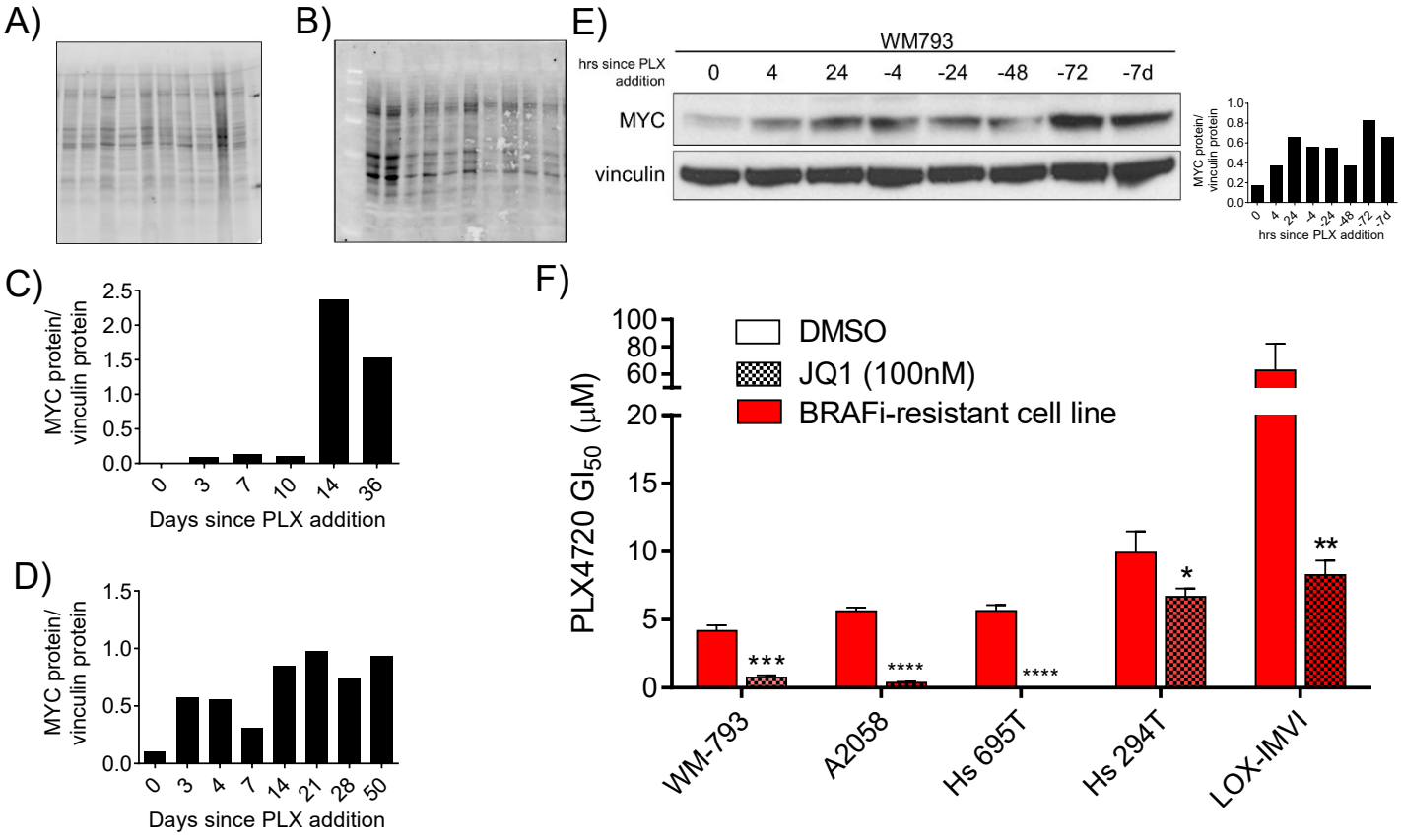


Figure S4. Characterization of the dynamics of MYC expression in intrinsically BRAFi resistant melanoma cells , related to Figure 4. A) Total protein staining for immunoblot shown in Figure 4A. Samples are ordered left to right as in Figure 4A. **B)** Total protein staining for immunoblot shown in Figure 4B. Samples are ordered left to right as in Figure 4B. **C)** Quantification of immunoblot in Figure 4F. **D)** Quantification of immunoblot in Figure 4H. **E)** WM793 cells were treated with 1 μM PLX4720 for 24 hours after which PLX4720 was withdrawn from the culture media and the cells were cultured normally for an additional 7 days. Whole cell lysates were prepared at the indicated time points and immunoblotted for expression of MYC and vinculin as a loading control. Blots are quantified in the right panel. **F)** Intrinsically BRAFi-resistant cell lines could be sensitized to PLX4720 by the addition of the indicated dose of JQ1. JQ1+PLX4720 treated cells are normalized to the viability of cells treated with JQ1 alone to account for nonspecific toxicity. *P* values denote significance between DMSO and JQ1 treatment in each cell line. Data are means (SD) from three experiments. **P* < 0.05; ***P* < 0.01; ****P* < 0.005; *****P* < 0.001.

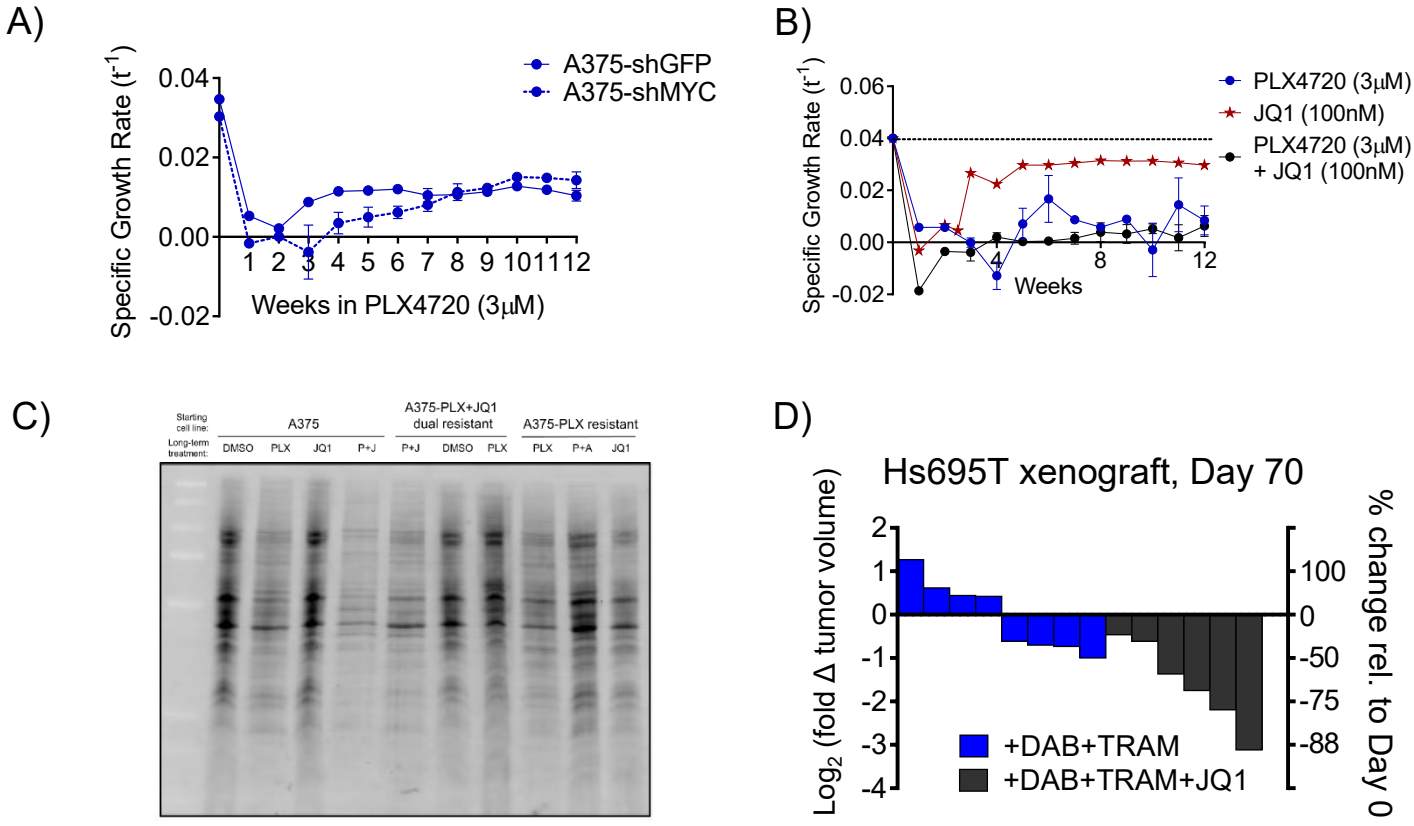
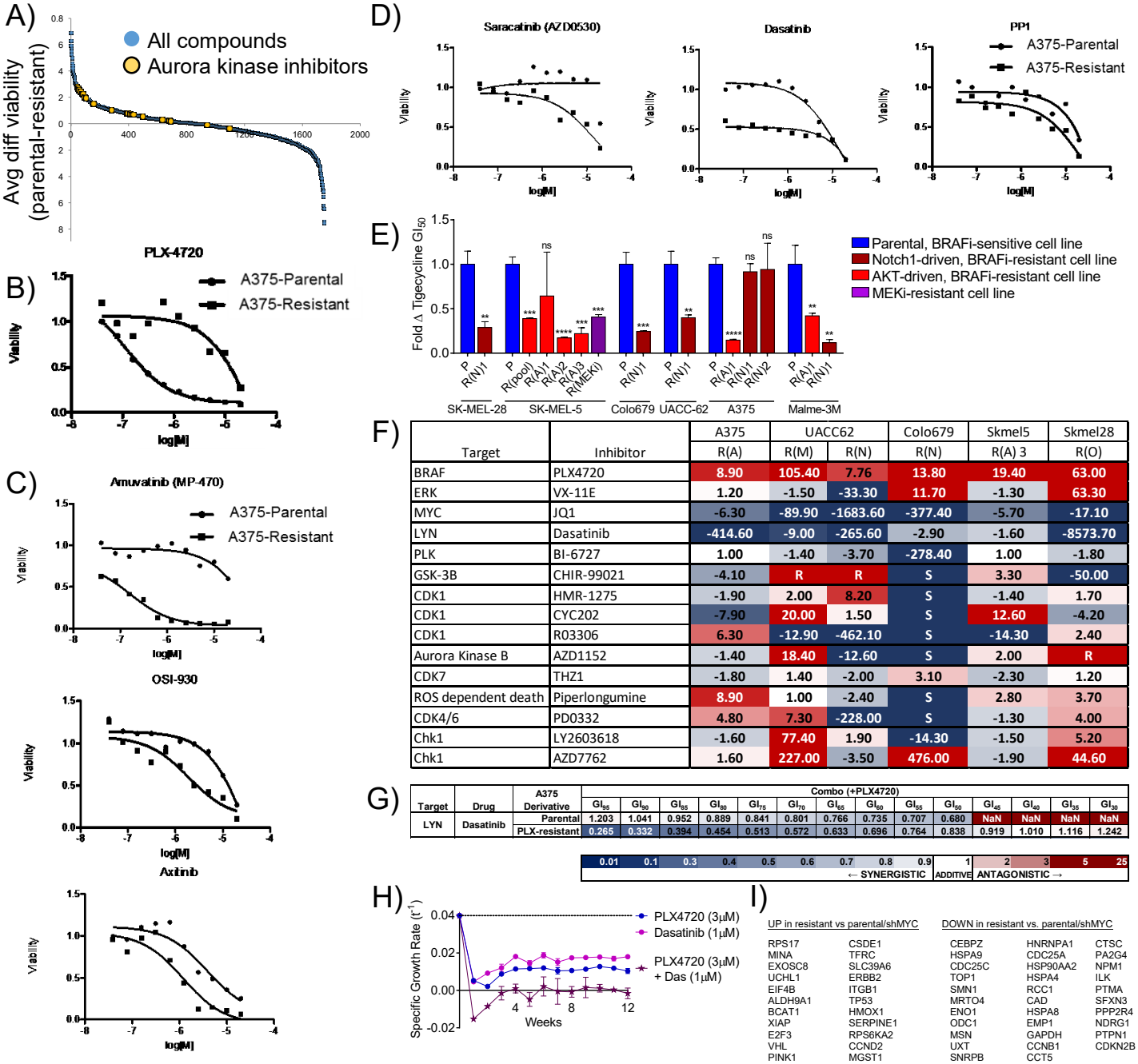


Figure S5. Evolution of resistance in treatment naïve melanoma cells with and without genetic or pharmacological MYC suppression , related to Figure 5. A) A375 cells expressing shGFP or shMYC were cultured in 3 μM PLX4720 for 12 weeks and cells were counted weekly. Growth rates were calculated from the cell counts and plotted. **B)** A375 cells were cultured in 3 μM PLX4720, 100 nM JQ1 or the combination for 12 weeks and cells were counted weekly. Growth rates were calculated from the cell counts and plotted. The growth rate of parental cells treated with DMSO is indicated with a dashed line. Data are means (SD) from three experiments. **C)** Total protein staining for the immunoblot shown in Figure 5C. **D)** Growth of Hs695T xenograft tumors treated with vehicle, JQ1 (45 mg/kg/d), dabrafenib (Dab, 30 mg/kg/d) plus trametinib (Tram, 0.6 mg/kg/d), or the combination over time. Data shown are mean tumor volume +/- SEM, with n=6-8 mice per group. Waterfall plot showing change in tumor sizes in the indicated groups on day 70 of treatment.



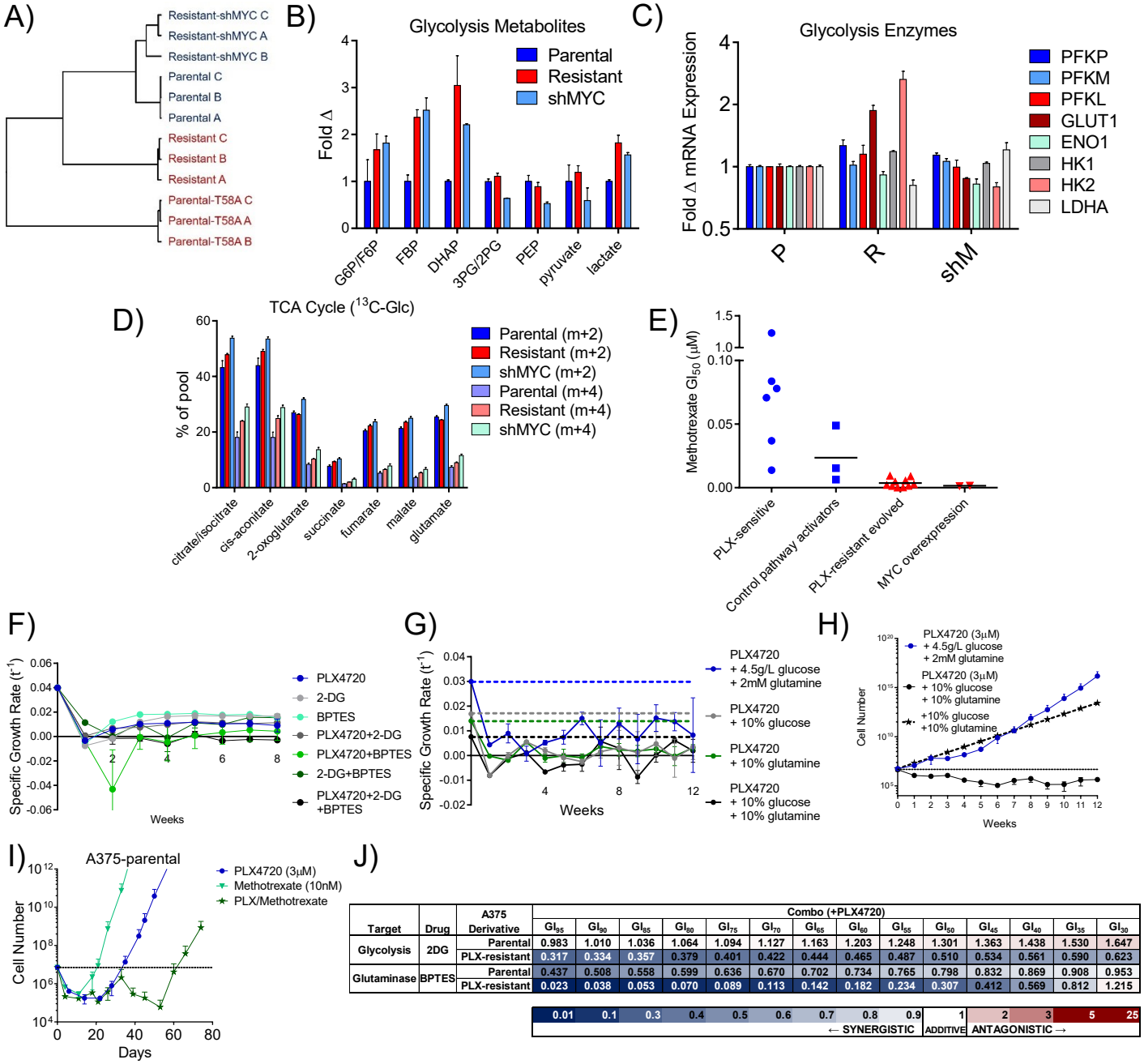
Target	Inhibitor	A375 R(A)	UACC62 R(M)	UACC62 R(N)	Colo679 R(N)	Skmel5 R(A) 3	Skmel28 R(O)
BRAF	PLX4720	8.90	105.40	7.76	13.80	19.40	63.00
ERK	VX-11E	1.20	-1.50	-33.30	11.70	-1.30	63.30
MYC	JQ1	-6.30	-89.90	-1683.60	-377.40	-5.70	-17.10
LYN	Dasatinib	-414.60	-9.00	-265.60	-2.90	-1.60	-8573.70
PLK	BI-6727	1.00	-1.40	-3.70	-278.40	1.00	-1.80
GSK-3B	CHIR-99021	-4.10	R	R	S	3.30	-50.00
CDK1	HMR-1275	-1.90	2.00	8.20	S	-1.40	1.70
CDK1	CYC202	-7.90	20.00	1.50	S	12.60	-4.20
CDK1	R03306	6.30	-12.90	-462.10	S	-14.30	2.40
Aurora Kinase B	AZD1152	-1.40	18.40	-12.60	S	2.00	R
CDK7	THZ1	-1.80	1.40	-2.00	3.10	-2.30	1.20
ROS dependent death	Piperlongumine	8.90	1.00	-2.40	S	2.80	3.70
CDK4/6	PD0332	4.80	7.30	-228.00	S	-1.30	4.00
Chk1	LY2603618	-1.60	77.40	1.90	-14.30	-1.50	5.20
Chk1	AZD7762	1.60	227.00	-3.50	476.00	-1.90	44.60

Target	Drug	A375 Derivative	Combo (+PLX4720) GI95	GI90	GI85	GI80	GI75	GI70	GI65	GI60	GI55	GI50	GI45	GI40	GI35	GI30
LYN	Dasatinib	Parental	1.203	1.041	0.952	0.889	0.841	0.801	0.766	0.735	0.707	0.680	NaN	NaN	NaN	NaN
		PLX-resistant	0.265	0.332	0.394	0.454	0.513	0.572	0.633	0.696	0.764	0.838	0.919	1.010	1.116	1.242

0.01 0.1 0.3 0.4 0.5 0.6 0.7 0.8 0.9 | 1 | 2 3 5 25
← SYNERGISTIC | ADDITIVE | ANTAGONISTIC →

UP in resistant vs parental/shMYC		DOWN in resistant vs. parental/shMYC		
RPS17	CSDE1	CEBPZ	HNRNPA1	CTSC
MINA	TFRC	HSPA9	CDC25A	PA2G4
EXOSC8	SLC39A6	CDC25C	HSP90AA2	NPM1
UCHL1	ERBB2	TOP1	HSPA4	ILK
EIF4B	ITGB1	SMN1	RCC1	PTMA
ALDH9A1	TP53	MRTO4	CAD	SFXN3
BCAT1	HMOX1	ENO1	HSPA8	PPP2R4
XIAP	SERPINE1	ODC1	EMP1	NDRG1
E2F3	RPS6KA2	MSN	GAPDH	PTPN1
VHL	CCND2	UXT	CCNB1	CDKN2B
PINK1	MGST1	SNRPB	CCT5	

Figure S6. Identification of pharmacological strategies to target MYC activated, BRAFi-resistant melanoma cells , related to Figure 6. Analysis of compound screening data. **A)** Distribution of Aurora kinase inhibitors in the primary screen. Secondary GI50 validation assays with **B)** the RAF inhibitor PLX4720, **C)** selected c-KIT inhibitors, and **D)** selected SRC family kinase inhibitors. **E)** Fold change in tigecycline GI_{50} values of the indicated PLX-resistant lines relative to parental lines. *P* values indicate significance between parental and resistant derivatives. **F)** Fold change in GI_{50} of the indicated inhibitors in various PLX4720-resistant evolved cell lines compared to their parental counterparts. Cases where PLX-resistant cells were more resistant to the indicated inhibitor than their parental line are shaded in red and cases that the PLX-resistant lines are sensitized are shaded blue. Cases where GI_{50} was not reached are indicated with 'R' or 'S.' **G)** Combinatorial Index (CI) values for parental A375 or PLX-resistant derivatives at a range of GI_{50} values for the combination of PLX4720 and dasatinib. Synergy is indicated by CI values less than 1. **H)** A375 cells were cultured in 3 μM PLX4720, 1 μM dasatinib, or the combination for 12 weeks and cells were counted weekly. Growth rates were calculated from the cell counts and plotted. The growth rate of parental cells treated with DMSO is indicated with a dashed line. Data are means (SD) of three experiments. **I)** List of MYC target genes with differential expression in resistant compared to parental and shMYC A375 cells.



J)

Target	Drug	A375 Derivative	Combo (+PLX4720)													
			GI_{95}	GI_{90}	GI_{85}	GI_{80}	GI_{75}	GI_{70}	GI_{65}	GI_{60}	GI_{55}	GI_{50}	GI_{45}	GI_{40}	GI_{35}	GI_{30}
Glycolysis	2DG	Parental	0.983	1.010	1.036	1.064	1.094	1.127	1.163	1.203	1.248	1.301	1.363	1.438	1.530	1.647
		PLX-resistant	0.317	0.334	0.357	0.379	0.401	0.422	0.444	0.465	0.487	0.510	0.534	0.561	0.590	0.623
Glutaminase	BPTES	Parental	0.437	0.508	0.558	0.599	0.636	0.670	0.702	0.734	0.765	0.798	0.832	0.869	0.908	0.953
		PLX-resistant	0.023	0.038	0.053	0.070	0.089	0.113	0.142	0.182	0.234	0.307	0.412	0.569	0.812	1.215

0.01	0.1	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	2	3	5	25
← SYNERGISTIC									ADDITIVE	ANTAGONISTIC →			

Figure S7. Metabolic dependencies in MYC activated, BRAFi-resistant melanoma cells , related to Figure 7. A) Hierarchical clustering of gene expression data (RNA-seq) from A375 parental cells (Parental), PLX4720-resistant clonal derivatives (Resistant), resistant A375 cells expressing shMYC (Resistant-shMYC), and parental A375 cells expressing MYC^{T58A} mutant (Parental-T58A). BRAFi-resistant, MYC-activated models are shown in red; BRAFi-sensitive, MYC-inactive models are shown in blue. Changes in **B)** glycolytic intermediates and **C)** expression of glycolytic enzymes among A375 (P), PLX4720-resistant A375 (R), and resistant cells expressing shMYC (shM). **D)** Changes in the incorporation of glucose into TCA cycle from labeled D-glucose. **E)** GI_{50} values of various PLX4720-sensitive (blue) and resistant (red) cell line models treated with methotrexate. **F)** A375 cells were cultured in 3 μM PLX4720, 5 mM 2-DG, 10 μM BPTES or the combinations for 12 weeks and cells were counted weekly. Growth rates were calculated from the cell counts and plotted. **G)** A375 cells were cultured in 3 μM PLX4720 and either DMEM with 4.5 g/L glucose and 2 mM glutamine, DMEM with 0.45 g/L glucose and 2 mM glutamine, DMEM with 4.5 g/L glucose and 0.2 mM glutamine, or DMEM with 0.45 g/L glucose and 0.2 mM glutamine for 12 weeks and cells were counted weekly. Growth rates were calculated from the cell counts and plotted. The growth rate of parental cells in DMEM with 4.5 g/L glucose and 2 mM glutamine is indicated with the blue dashed line. The growth rate of parental cells in DMEM with the various depleted media is indicated with the dashed lines. Growth rates in the depleted media were calculated from daily cell counts over 7 days of culture. **H)** The growth rate of A375 cells in DMEM containing 0.45 g/L glucose and 0.2 mM glutamine was measured daily over 7 days. This growth rate was used to calculate total cell numbers over 12 weeks of growth in the double depleted media and compared to cell counts with PLX4720 treatment. **I)** Projected cell numbers of A375 cultured in 3 μM PLX4720, 10 nM methotrexate or the triple combination. **J)** Combinatorial Index (CI) values for parental A375 or PLX-resistant derivatives at a range of growth inhibition (GI) values for the combination of PLX4720 and 2-DG or the combination of PLX4720 and BPTES. Synergy is indicated by CI values less than 1. Data are means (SD) from three experiments.

EXTENDED EXPERIMENTAL PROCEDURES

1 Statistical Modeling of Genomic Data

Gene Expression Data and Preprocessing Procedures

Microarray gene expression data (Rizos *et al.*, 2014) was collected that included 59 $BRAF^{V600}$-mutant melanoma metastatic samples with 38 progressing and 21 matched pretreatment samples collected from 30 patients. Only two patients, indexed as 5 and 10, were analyzed pre-treatment, on-treatment, and on-relapse (which we term as "progression"). We only utilized the samples in this dataset which had matched pre-treatment and progressed tumors (this included patients 5 and 10). To increase the sample size of the on-treatment tumors used to obtain the BRAFi/MEKi response signature, we added matched pre- and on-treatment microarray gene expression data of melanoma cell lines from two sets of data: M229, M238, and M249 (Nazarian *et al.*, 2010); and SkMel1, SkMel5, SkMel19, SkMel28, and Malme3M (Pratilas *et al.*, 2009). These data are publicly available and can be accessed in the NCBI's Gene Expression Omnibus under the GEO SuperSeries accession numbers: `GSE50509` (Rizos *et al.*, 2014), `GSE24862` (Nazarian *et al.*, 2010), and `GSE10087` (Pratilas *et al.*, 2009).

Collapsing Probes and Data Normalization

Expression data were obtained using the Illumina HumanHT-12 V4.0 Expression Beadchip (Rizos *et al.*, 2014), the Affymetrix Human Gene 1.0 ST Array (Pratilas *et al.*, 2009), and the Affymetrix Human Genome U133A 2.0 Array (Nazarian *et al.*, 2010), respectively. we took the following preprocessing and normalization steps to correct for batch effects and other possible confounders. First, using the annotations of each Affymetrix/Illumina Chip/Platform, each probe of each sample in each experiment was mapped to its corresponding HUGO gene name and symbol. Next, assuming that there is no principled way to account for "missingness", we eliminate the probes that did not match to known genes. For each sample in all three experiments, we follow previous works (Li *et al.*, 2011) and took the probes that mapped to multiple genes and averaged their expression values. Next, we identified the mapped genes that were common in all three chips/platforms and only used these genes in subsequent statistical and pathway analyses. All three datasets were RMA corrected and $\log_2$-transformed. We utilize base 2 because any transformed

2-fold ratio will be converted to an ± 1 scale. This interpretation of fold change helped with interpreting of BAKR regression coefficients and the underlying marginal covariate interactions. Lastly, we perform a final cross-platform quantile normalization (Rudy and Valafar, 2011) on the combined dataset. The goal of the cross-platform normalization was to dilute some of the random noise created by chip-to-chip batch effects and strengthen biological signal in the data. After completing these steps, we were left with a final data set consisting of $n = 68$ samples and $p = 11{,}657$ genes.

Splitting the Data: Phases I and II

The data were split into two phases: Phase I (i.e. generation of the BRAFi/MEKi response signature) corresponded to pre-treatment versus on-treatment samples ($n = 39$); Phase II (i.e. identification of genes that return to pre-treatment status) corresponded to on-treatment versus post-treatment samples ($n = 39$).

Deriving a Gene Signature

In this section, we will use the Phase I dataset to infer a BRAFi/MEKi response signature $\mathcal{G}$ of MAPK pathway addiction. In this study, we use a logistic version of the Bayesian approximate kernel regression (BAKR-logit) model (Crawford *et al.*, 2017) to derive this collection of genes. Note that this methodology is advantageous for statistically analyzing cancer genomic data such as this for a few notable reasons. One, there is no reason to assume a linear relationship between the changes in gene expression and the effect of a given therapeutic strategy. Moreover, crosstalk between signaling pathways, as well as between the genes and nodes within these pathways, has been suggested to be associated with drug response (Bostock, 2005; Bender and Nahta, 2008; Yamaguchi *et al.*, 2014).

Notation

Let $\mathbf{X}$ denote the observed $n \times p$ genotype matrix, where each element x_{ij} is the expression value of the j^{th} gene from the i^{th} sample. Next, let $\mathbf{y} = [y_1, \ldots, y_n]$ represent the following n-dimensional binary response vector with

$$y_i = \begin{cases} 1 & \text{if} \quad \text{MAPK pathway is activated (pre-treatment);} \\ 0 & \text{if} \quad \text{MAPK pathway is deactivated (on-treatment).} \end{cases}$$

Alternatively, we say that $\mathbf{y}$ is drawn as n-independent Bernoulli random variables, such that $y_i \overset{\text{iid}}{\sim} \text{Bern}(\pi_i)$.

Logistic Bayesian Approximate Kernel Regression Model

Consider the following logit link function between the gene expression and binary treatment classes:

$$\text{logit}(\mathbb{E}[y_i \mid \mathbf{x}_i]) = \text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \mathbf{x}_i^\intercal \boldsymbol{\beta},$$

where $\mathbf{x}_i^\intercal$ is the i^{th} row of the $n \times p$ expression matrix $\mathbf{X}$, and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_p]$ is an unknown vector of regression coefficients for each gene $j = 1, \ldots, p$. To facilitate statistical inference and interpretation of this logistic model setup, we define a latent variable $\mathbf{z} = [z_1, \ldots, z_n]$ such that $y_i = 1$ if $z_i > 0$, and $y_i = 0$ otherwise (Kinney and Dunson, 2007). In other words, each z_i is logistically distributed with location parameter $\mathbf{x}_i^\intercal \boldsymbol{\beta}$, with a corresponding probability density function:

$$\mathcal{L}(z_i; \boldsymbol{\beta}) = \frac{\exp\{-(z_i - \mathbf{x}_i^\intercal \boldsymbol{\beta})\}}{[1 + \exp\{-(z_i - \mathbf{x}_i^\intercal \boldsymbol{\beta})\}]^2}. \tag{1}$$

It was shown that this relationship is approximately a non-central t_ν-distribution with location parameter $\mathbf{x}_i^\intercal \boldsymbol{\beta}$ and scale parameter σ_ε^2 (Kinney and Dunson, 2007). This means that we may express (1) as a scale mixture of normals, represented in matrix notation as the following:

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \varepsilon, \quad \varepsilon \sim \text{MVN}(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Omega}^{-1}), \tag{2}$$

where $\boldsymbol{\Omega} = \text{Diag}(\omega_1, \ldots, \omega_n)$ is an $n \times n$ diagonal matrix and $\text{MVN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

As in most gene expression studies, the number of samples is far fewer than the number of genes (i.e. $p \gg n$). For this reason, low rank factorizations of the design matrix $\mathbf{X}$ are used both for numerical stability as well as for statistical efficiency (West, 2003; Liang *et al.*, 2007; Liang *et al.*, 2009; de los Campos *et al.*, 2010; Crossa *et al.*, 2014; Crawford *et al.*, 2017). In addition, there is no reason to assume a linear relationship between the latent variable $\mathbf{z}$ and the changes in gene expression as is assumed in Equation (2). To increase predictive power, we used a nonlinear regression model that can capture higher-order interactions between genes, while also providing effect size estimates for each gene. This model builds on the classical Gaussian process — or the reproducing kernel Hilbert space(RKHS) model — and is formerly

referred to as the Bayesian approximate kernel regression (BAKR) model (Crawford *et al.*, 2017). We summarize the relevant aspects of the BAKR model that we used to infer signatures.

We first specify a RKHS model — a class of models used extensively in machine learning, as well as in genetics, to improve predictive accuracy, and capture higher order nonlinear effects (Gianola *et al.*, 2006; Liang *et al.*, 2007; Liang *et al.*, 2009; de los Campos *et al.*, 2010; Crossa *et al.*, 2014; Howard *et al.*, 2014):

$$\mathbf{z} = \mathbf{K}\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \quad \text{with} \quad \boldsymbol{\alpha} \sim \text{MVN}(\mathbf{0}, \sigma_\alpha^2 \mathbf{K}^{-1}) \quad \text{and} \quad \boldsymbol{\varepsilon} \sim \text{MVN}(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Omega}^{-1}), \tag{3}$$

where $\mathbf{K}$ is an $n \times n$ kernel matrix whose entries are nonlinear functions of pairs of expression profiles, and $\boldsymbol{\alpha}$ is an n-vector of random kernel coefficients. By definition, the kernel matrix $\mathbf{K}$ is symmetric and semi-positive definite:

(1) $k(\mathbf{x}_i, \mathbf{x}_j) = k(\mathbf{x}_j, \mathbf{x}_i) \quad \forall\ i, j,$

(2) $\sum_i \sum_j k(\mathbf{x}_i, \mathbf{x}_j)\alpha_i\alpha_j \geq 0 \quad \forall\ \alpha_i, \alpha_j \in \mathbb{R}.$

The model form in Equation (3) turns a p-dimensional optimization problem into an optimization problem over just n parameters.

A key feature of the BAKR modeling framework is its ability to infer effect sizes for genes while implicitly modeling all possible higher-order interactions between them (Crawford *et al.*, 2017). In high-dimensional settings, this is something that is not feasible for linear models (Yu *et al.*, 2006; Logsdon *et al.*, 2010; Zhang *et al.*, 2010; Guan and Stephens, 2011; Lippert *et al.*, 2011; Zhou and Stephens, 2012; Zhou *et al.*, 2013) or other variable selection testing strategies (Hauck Jr and Donner, 1977; Kleinbaum and Klein, 2010; Li *et al.*, 2011; Wu *et al.*, 2011).

The advantage of BAKR over standard RKHS models is that RKHS models cannot provide estimates of effect sizes for each variable. BAKR overcomes this hurdle by using properties of shift-invariant kernel functions and employing a map between the RKHS model (i.e Equation (3)) and a linear model (i.e. Equation (2)). More specifically, shift-invariant kernel functions can be written as a weighted infinite sum of Fourier bases $\{\psi_i(x)\}_{i=1}^{\infty} = \boldsymbol{\psi}(x)$, such that $k(\|x_i - x_j\|) = \boldsymbol{\psi}(x_i)^\intercal \boldsymbol{\psi}(x_j)$. Though never constructed in practice, under Bochner's Theorem (Bochner, 1934), a p-dimensional approximation of these bases may be constructed using the Fourier Transform of the kernel function such that $k(\|x_i - x_j\|) \approx \sum_{j=1}^{p} \psi_j(x_i)^\intercal \psi_j(x_j) = \widetilde{k}(\|x_i - x_j\|)$ (Rahimi and Recht, 2007; Băzăvan *et al.*, 2012). This allows for a reformulation of Equation

(3) using an approximate kernel matrix of the form:

$$\mathbf{z} = \widetilde{\mathbf{K}}\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \quad \text{with} \quad \boldsymbol{\alpha} \sim \text{MVN}(\mathbf{0}, \sigma_\alpha^2 \widetilde{\mathbf{K}}^{-1}) \quad \text{and} \quad \boldsymbol{\varepsilon} \sim \text{MVN}(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Omega}^{-1}). \tag{4}$$

Going forward, we denote $\boldsymbol{\Psi}$ as the infinite Fourier basis matrix such that $\mathbf{K} = \boldsymbol{\Psi}^\intercal\boldsymbol{\Psi}$, and $\widetilde{\boldsymbol{\Psi}}$ to represent the p-term finite approximation such that $\mathbf{K} \approx \widetilde{\boldsymbol{\Psi}}^\intercal\widetilde{\boldsymbol{\Psi}} = \widetilde{\mathbf{K}}$. These approximate bases can be used in a linear mapping to relate $\boldsymbol{\alpha}$ to $\boldsymbol{\beta}$ in order to extract an effect size for each gene (Crawford *et al.*, 2017). Because of this linear mapping, BAKR can obtain posterior samples of $\boldsymbol{\beta}$ to quantify phenotype association strength and identify associated genes. This is key to defining a robust BRAFi/MEKi response signature. BAKR has the flexibility to take on the approximation of any shift-invariant kernel to model a variety of different genetic architectures; however, we detail our model strictly under the approximation of the Gaussian kernel function (Williams and Seeger, 2001; Vert *et al.*, 2004; Chang *et al.*, 2010):

$$\widetilde{k}(\|\mathbf{x}_i - \mathbf{x}_j\|) \approx k(\|\mathbf{x}_i - \mathbf{x}_j\|) = \exp\left\{-\frac{1}{p}||\mathbf{x}_i - \mathbf{x}_j||^2\right\},$$

where $||\mathbf{x}||^2 = \sum_{i=1}^n \mathrm{x}^2$ is the ℓ_2-norm.

Complete Model Specification

We now state the complete specification of BAKR-logit. Since the approximation of any shift-invariant kernel matrix is also symmetric and semi-positive definite, we took advantage of a low-rank approximation of the nonlinear kernel matrix $\widetilde{\mathbf{K}}$ to dramatically reduce computational cost (West, 2003; Liang *et al.*, 2007; Liang *et al.*, 2009; de los Campos *et al.*, 2010; Crossa *et al.*, 2014; Crawford *et al.*, 2017). This was done by using the spectral decomposition of the approximate Gaussian kernel matrix $\widetilde{\mathbf{K}}$. Specifically:

$$\mathrm{y}_i = \begin{cases} 1 & \text{if} \quad \mathrm{z}_i > 0 \\ 0 & \text{if} \quad \mathrm{z}_i \leq 0 \end{cases} \quad \text{for} \quad i = 1, \ldots, n$$

$$\mathbf{z} = \widetilde{\mathbf{U}}\boldsymbol{\theta} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim \text{MVN}(\mathbf{0}, \sigma_\varepsilon^2 \boldsymbol{\Omega}^{-1}) \tag{5}$$

$$\boldsymbol{\theta} \sim \text{MVN}_q(\mathbf{0}, \sigma_\theta^2 \widetilde{\boldsymbol{\Lambda}}) \tag{6}$$

$$\omega_i, \sigma_\theta^{-2} \sim \Gamma(\nu/2, \nu/2), \quad \sigma_\varepsilon^2 = \frac{\nu - 2}{3\nu}\pi^2 \tag{7}$$

where $\Gamma(a, b)$ is used to represent a Gamma distribution with shape a and rate b. In practice, we follow previous studies (Kinney and Dunson, 2007) and fix $\nu = 7.3$. Note that we utilize the empirical kernel factor representation where $\widetilde{\mathbf{K}} = \widetilde{\mathbf{U}}\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}^{\mathsf{T}}$ and $\boldsymbol{\theta} = \widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}^{\mathsf{T}}\boldsymbol{\alpha}$. For numerical stability and reduction of computational complexity, eigenvectors corresponding to smaller eigenvalues were truncated (West, 2003; Lopes and West, 2004; Liang *et al.*, 2007; Pillai *et al.*, 2007; Liang *et al.*, 2009); so without loss of generality, we implement BAKR by considering $\widetilde{\mathbf{U}}$ to be an $n \times q$ matrix containing the eigenvectors of $\widetilde{\mathbf{K}}$ (i.e. $\widetilde{\mathbf{U}}\widetilde{\mathbf{U}}^{\mathsf{T}} = \mathbf{I}_n$), and $\widetilde{\mathbf{\Lambda}}$ as a $q \times q$ diagonal matrix of the top q eigenvalues of $\widetilde{\mathbf{K}}$. This consideration represented an even greater reduction in required computation since $q \leq n \ll p$. In this study, we chose q to represent the number of eigenvalues that explain 99% of cumulative variance in $\widetilde{\mathbf{K}}$.

Because $\widetilde{\mathbf{K}}$ approximates a matrix that is shift-invariant, we utilized an inverse mapping that allows for inference to be made on the p-dimensional genes $\boldsymbol{\beta}$ in the original genetic space (Crawford *et al.*, 2017):

$$\widetilde{\boldsymbol{\beta}} = \mathbf{X}^{\dagger}\widetilde{\mathbf{\Psi}}^{\mathsf{T}}(\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{U}}^{\mathsf{T}}\widetilde{\mathbf{K}}^{-1}\widetilde{\mathbf{\Psi}}^{\mathsf{T}})^{-1}\widehat{\boldsymbol{\theta}},$$

where $\widehat{\boldsymbol{\theta}}$ is an estimate of the kernel factor coefficients drawn from the posterior distribution $p(\boldsymbol{\theta} \,|\, \mathbf{y})$. Furthermore, since we were strictly concerned with a case in which $p \gg n$, the inverse of $\mathbf{X}\mathbf{X}^{\mathsf{T}}$ exists and hence we let $\mathbf{X}^{\dagger} = \mathbf{X}^{\mathsf{T}}(\mathbf{X}\mathbf{X}^{\mathsf{T}})^{-1}$. By plugging in this quantity, the linear map simplified to

$$\widetilde{\boldsymbol{\beta}} = \mathbf{X}^{\mathsf{T}}(\mathbf{X}\mathbf{X}^{\mathsf{T}})^{-1}\widetilde{\mathbf{U}}\widehat{\boldsymbol{\theta}}. \tag{8}$$

We implemented the mapping from the standard RKHS model to effect sizes as a deterministic step in an MCMC Gibbs sampler resulting in empirical draws $\widetilde{\boldsymbol{\beta}}$ from an implied posterior distribution $p(\boldsymbol{\beta} \,|\, \mathbf{y})$.

Posterior Sampling and Inference

Based on the complete BAKR-logit specification, a standard Gibbs sampler was derived from the following joint posterior distribution:

$$\begin{aligned} p(\mathbf{z}, \boldsymbol{\theta}, \sigma_{\theta}^{-2}, \boldsymbol{\Omega} \,|\, \mathbf{y}) \propto & \left[\prod_{i=1}^{n}\left(\frac{\sigma_{\varepsilon}^2}{\omega_i}\right)^{-\frac{1}{2}} \exp\left\{-\frac{\omega_i}{2\sigma_{\varepsilon}^2}(\mathsf{z}_i - \widetilde{\mathbf{u}}_i^{\mathsf{T}}\boldsymbol{\theta})^2\right\}\mathbb{1}(\mathsf{A}_i)\right] \times \prod_{i=1}^{n}\omega_i^{\frac{\nu}{2}-1}\exp\left\{-\frac{\nu}{2}\omega_i\right\} \\ & \times (\sigma_{\theta}^2)^{-\frac{q}{2}}\exp\left\{-\frac{1}{2\sigma_{\theta}^2}\boldsymbol{\theta}^{\mathsf{T}}\widetilde{\mathbf{\Lambda}}^{-1}\boldsymbol{\theta}\right\} \times (\sigma_{\theta}^{-2})^{\frac{\nu}{2}-1}\exp\left\{-\frac{\nu}{2}\sigma_{\theta}^{-2}\right\} \end{aligned} \tag{9}$$

where $\mathbb{1}(\cdot)$ is an indicator function and each member of the set $\mathbf{A} = [\mathsf{A}_1, \ldots, \mathsf{A}_n]$ is defined as

$$\mathsf{A}_i = \begin{cases} \{\mathsf{z}_i : \mathsf{z}_i > 0\} & \text{if } \mathsf{y}_i = 1, \\ \{\mathsf{z}_i : \mathsf{z}_i \leq 0\} & \text{if } \mathsf{y}_i = 0. \end{cases} \tag{10}$$

After a selection of initial values, samples of parameters and hyper-parameters were drawn sequentially from their respective complete conditional posterior distributions, which we detail below. At each algorithmic step, with all conditioning parameters fixed at their most recent values, we updated the iterates until we created a set of relevant MCMC draws. The complete posterior distributions are given:

(1) For $i = 1, .., n$

$$\mathsf{z}_i^{(t+1)} \,|\, \mathbf{z}^{(t)}, \boldsymbol{\theta}, \sigma_\theta^2, \boldsymbol{\Omega}, \mathbf{y} \sim \begin{cases} \mathsf{N}(\widetilde{\mathbf{u}}_i^\intercal \boldsymbol{\theta}, \sigma_\varepsilon^2 \omega_i^{-1}) \mathbb{1}(\mathsf{z}_i^{(t)} > 0) & \text{if} \quad \mathsf{z}_i^{(t)} > 0, \\ \mathsf{N}(\widetilde{\mathbf{u}}_i^\intercal \boldsymbol{\theta}, \sigma_\varepsilon^2 \omega_i^{-1}) \mathbb{1}(\mathsf{z}_i^{(t)} \leq 0) & \text{if} \quad \mathsf{z}_i^{(t)} \leq 0; \end{cases}$$

(2) $\boldsymbol{\theta} \,|\, \mathbf{z}, \sigma_\theta^2, \boldsymbol{\Omega}, \mathbf{y} \sim \mathsf{MVN}(\mathbf{m}_\theta^*, \mathbf{V}_\theta^*)$ where $\mathbf{V}_\theta^* = \sigma_\varepsilon^2 \sigma_\theta^2 (\sigma_\varepsilon^2 \widetilde{\boldsymbol{\Lambda}}^{-1} + \sigma_\theta^2 \widetilde{\mathbf{U}}^\intercal \boldsymbol{\Omega} \widetilde{\mathbf{U}})^{-1}$ and $\mathbf{m}_\theta^* = \frac{1}{\sigma_\varepsilon^2} \mathbf{V}_\theta^* \widetilde{\mathbf{U}}^\intercal \boldsymbol{\Omega} \mathbf{z}$;

(3) $\widetilde{\boldsymbol{\beta}} = \mathbf{X}^\intercal (\mathbf{X}\mathbf{X}^\intercal)^{-1} \widetilde{\mathbf{U}} \boldsymbol{\theta}$;

(4) $\sigma_\theta^{-2} \,|\, \mathbf{z}, \boldsymbol{\theta}, \boldsymbol{\Omega}, \mathbf{y} \sim \Gamma(a_\theta^*, b_\theta^*)$ where $a_\theta^* = \frac{1}{2}(\nu + q)$ and $b_\theta^* = \frac{1}{2}(\nu + \boldsymbol{\theta}^\intercal \widetilde{\boldsymbol{\Lambda}}^{-1} \boldsymbol{\theta})$;

(5) For $i = 1, .., n$

$\omega_i \,|\, \mathbf{z}, \boldsymbol{\theta}, \sigma_\theta, \mathbf{y} \sim \Gamma(a_\omega^*, b_\omega^*)$ where $a_\omega^* = \frac{1}{2}(\nu + 1)$ and $b_\omega^* = \frac{1}{2\sigma_\varepsilon^2}(\nu \sigma_\varepsilon^2 + \mathsf{e}_i^2)$, with $\mathsf{e}_i = \mathsf{z}_i - \widetilde{\mathbf{u}}_i^\intercal \boldsymbol{\theta}$.

In this study, we obtained over 20,000 MCMC samples from the BAKR-logit Gibbs sampler. These samples were selected from a run of a 110,000 iterations, where we keep every 5$^{\text{th}}$ sample, and then follow up with burn-in of 2,000 samples.

Variable Selection: Local False Sign Rate

We are reminded that the purpose of using the BAKR-logit in this study is to discover a set of genes $\mathcal{G}$ (i.e. a gene signature) whose expression levels robustly change following treatment with BRAFi/MEKi in *BRAF*-mutant melanomas. We used the estimate of the original effect sizes $\widetilde{\boldsymbol{\beta}}$ as a metric of the relevance for each gene. Specifically, the metric we used to determine the members of $\mathcal{G}$ is the local false sign rate (lfsr), which is analogous to the local false discovery rate (Efron, 2007). The lfsr provides a measure of

confidence in the sign of the effect rather than confidence of the effect being non-zero (Stephens, 2017). Alternatively, we say that we were more concerned with controlling the minimization of "type S errors" (i.e. the errors of sign), rather than the traditional type I errors (Gelman and Tuerlinckx, 2013). Therefore, we choose to be confident in the directional change of a gene's regulatory pattern (i.e. up-regulation or down-regulation) in the presence of drug. Given posterior samples $\widetilde{\boldsymbol{\beta}}$ for each gene j, we defined the corresponding local false sign rate as (Stephens, 2017)

$$\text{lfsr}_j = \min[p(\beta_j \geq 0 \,|\, \mathbf{y}, \widetilde{\boldsymbol{\beta}}), p(\beta_j \leq 0 \,|\, \mathbf{y}, \widetilde{\boldsymbol{\beta}})]. \tag{11}$$

This selection procedure is a post-hoc deterministic computation applied to the MCMC samples of $\boldsymbol{\beta}$. The BRAFi/MEKi response signature $\boldsymbol{\mathcal{G}}$ is then defined as the set of genes that satisfy $\boldsymbol{\mathcal{G}} = \{j : \text{lfsr}_j \leq c\}$. In this study, we subjectively choose $c = 0.01$. This resulted in a BRAFi/MEKi response signature of $p^* = 68$ genes.

Model Comparisons

To illustrate the utility of the BAKR-logit model, we compared our approach to two linear models which use t-distributed test statistics to identify differentially expressed genes. The first method is a gene-wise association analysis that uses a parametric empirical Bayes approach to borrow strength between genes in order to moderate effect sizes and residual variances (Smyth, 2004). These resulting summary statistics are used to compute Benjamini-Hochberg corrected p-values (BH q-values) for every observed gene, where those with q-values below $q \leq 0.05$ are called significantly differentially expressed. We fit this model using the publicly available `R` package `limma` (Ritchie *et al.*, 2015). The second approach that we considered identifies differentially expressed genes using the classical likelihood ratio test (LRT) (Peng *et al.*, 2002). Briefly, for each gene in turn, we computed a likelihood ratio comparing an alternative model with the gene's expression level to a null model containing just the intercept. Similarly, we then considered those genes with multiplicty corrected p-values below 0.05 as those associated with MAPK inhibition. We fit this score test for nested parametric linear models using the `lrtest` function in the publicly available `R` package `lmtest` (Zeileis and Hothorn, 2002).

Here, we evaluated the ability of the BAKR-logit model, limma, and the LRT to identify genes that fully characterize MAPK inhibition in *BRAF*-mutant melanoma. We are reminded that using the BAKR-logit

resulted in a BRAFi/MEKi response signature of $p^* = 68$ genes. Comparatively, the limma model detected 16 BRAFi/MEKi response signature genes, while the LRT BRAFi/MEKi response signature contained only 8 genes. Notedly, there was substantial overlap between the signature genes identified by the LRT, and those selected by the limma and the BAKR-logit models (see Tables S1 and S7). Again, we want to highlight that the advantage of the BAKR framework is the fact that it also implicitly considers a marginal notion of interaction effects between a given gene and all other genes (Crawford *et al.*, 2017). Hence, BAKR identifies genes as being significant that the other linear modeling approaches fail to detect. For instance, while all three methods select the proto-oncogene *MYC* as being a key downstream component in MAPK inhibition, BAKR is the only methodology to also select genes such as *ID2* and *SKI* as being potential predictors of BRAFi/MEKi response. Briefly, overexpression of the *ID2* gene has been shown to be required for *MYC* signaling (Lasorella *et al.*, 2000). Similarly, oncogenic interactions between *SKI* and *MYC*, coupled with repression of the TGF-β signaling pathway, has been suggested to be consequential to long term treatment strategies in human melanomas (Sun *et al.*, 1999). Therefore, we conclude that the BRAFi/MEKi response signature derived by the BAKR-logit model gives a more complete illustration of MAPK inhibition in *BRAF*-mutant melanoma than both the limma linear model and the LRT.

Lastly, we want to stress the one caveat that this model comparison is strictly empirical in nature. Given that our aim here is to make novel discoveries, we are not afforded any "true" answers like there would be in a typical power simulation study. Nonetheless, BAKR identified the same signature genes as the opposing methods, as well as notable others that the literature suggests to be contextually relevant. A more fundamentally comprehensive review of BAKR, its predictive accuracy, and its power to detect true causal genes and other genetic variants can be found in the original publication (Crawford *et al.*, 2017).

Software and Model Implementation

Software for implementing the BAKR modeling framework is carried out in R and Rcpp code, which is freely available at `https://github.com/lorinanthony/BAKR`.

Identifying Elements of Relapse

Inference of MAPK signaling pathway activity was just the first step in discovering the genes and/or the cellular processes that are important for therapeutic resistance and melanoma reoccurrence. To search for a potential convergent effector of resistance, we began by reasoning that such an effector should follow

two transitional rules: (1) it should be regulated downstream of the driver oncogene, and (2) it should rebound to at least pre-treatment expression or activation states at resistance, independent of the upstream mechanisms driving resistance. We considered signature genes that satisfy this criteria as a signature for patient relapse and we referred to this set as $\boldsymbol{\mathcal{R}}$. Specifically, we used Bayes factors (BF) and marginal likelihoods to determine which of the $p^* = 68$ members of $\boldsymbol{\mathcal{G}}$ belonged in $\boldsymbol{\mathcal{R}}$.

Bayes Factor Computation

In this section, we develop the methodology for identifying an expression signature for melanoma progression (i.e. the members of $\boldsymbol{\mathcal{R}}$). We begin by introducing the Bayes factor, which is the probability of observing the data under one condition relative to another (Kass and Raftery, 1995). In particular, we used the Phase II dataset to obtain $\boldsymbol{\mathcal{R}}$. We redefine the binary phenotype $\mathbf{y}^*$, where now $y_i^* = 1$ denotes a progressed (relapsed) tumor and $y_i^* = 0$ corresponds to a treated sample. Now let $p(\mathbf{y}^* \mid \mathcal{M})$ be the probability of observing the phenotype under some model $\mathcal{M}$, and $p(\boldsymbol{\Theta}^* \mid \mathcal{M})$ be a prior belief about that model's parameters (i.e. gene coefficients in our case). The Bayes factor between two models $\mathcal{M}_1$ and $\mathcal{M}_0$ is then defined as

$$\text{BF}_{10} = \frac{p(\mathbf{y}^* \mid \mathcal{M}_1)}{p(\mathbf{y}^* \mid \mathcal{M}_0)} = \frac{\int p(\mathbf{y}^* \mid \boldsymbol{\Theta}_1^*, \mathcal{M}_1) p(\boldsymbol{\Theta}_1^* \mid \mathcal{M}_1) d\boldsymbol{\Theta}_1^*}{\int p(\mathbf{y}^* \mid \boldsymbol{\Theta}_0^*, \mathcal{M}_0) p(\boldsymbol{\Theta}_0^* \mid \mathcal{M}_0) d\boldsymbol{\Theta}_0^*}. \tag{12}$$

Here, the subscript identifies which models are being compared, while the corresponding order denotes which model is in the numerator and which is in the denominator. As previously shown (Kass and Raftery, 1995; Rouder and Morey, 2012), the Bayes factor is interpretable without recourse to additional criteria or qualification. For example, $\text{BF}_{10} = 5$ means that the data are 5 times more probable under $\mathcal{M}_1$ than under $\mathcal{M}_0$.

We are particularly interested in individual members of the gene signature that track that well with progression — meaning $\boldsymbol{\mathcal{R}} \subseteq \boldsymbol{\mathcal{G}}$. In other words, we want to identify the genes that positively contribute to the explanation of melanoma reoccurrence, conditioned on none of the other genes being present. The reason for this approach is that we are strictly concerned with finding potential drivers of resistance. This is accomplished by computing Bayes factors, while comparing an intercept term (null model, $\mathcal{M}_0$) to the addition of each signature gene independently (alternative models, $\mathcal{M}_j$ for $j = 1, \ldots, p^*$). Typically, Equation (12) can be computationally expensive as one considers all possible combinations of explanatory

variables (Berger and Pericchi, 1996; Barbieri and Berger, 2004). However, given the scope of the problem in question, our logic reduces this search to just $p^* = 68$ Bayes factors that needed to be calculated. Again, we define p^* as the number of genes included in our melanoma signature, (i.e. $|\mathcal{G}| = p^* = 68$).

We used the Laplace method (Guan and Stephens, 2008) to approximate the Bayes factor for the binary responses. Assume that each phenotype is modeled by a logistic regression model,

$$\mathcal{M}_j : \quad \log\left(\frac{\Pr[y_i^* = 1]}{\Pr[y_i^* = 0]}\right) = \mu + x_{ij}^* \beta_j^* = f(\mu, \beta_j^*), \tag{13}$$

where x_{ij}^* is the j^{th} element of the i^{th} row in the $n \times g$ expression matrix $\mathbf{X}^*$ containing the treated and progressed samples from our data and the genes from our signature (i.e. a subset of the Phase II dataset). Hence, β_j^* is defined as the effect parameter for the j^{th} signature gene in $\mathcal{G}$. Lastly, we define μ as a common intercept. Then under the logistic function, Equation (13) can be stated as:

$$\mathcal{M}_j : \quad \pi_i^* = \Pr[y_i^* = 1] = \frac{\exp\{f(\mu, \beta_j^*)\}}{1 + \exp\{f(\mu, \beta_j^*)\}}. \tag{14}$$

By using Equation (14), the log-likelihood of the data is specified as

$$\begin{aligned} \ell(\mathbf{y}^*; \mu, \beta_j^*) &= \sum_i [y_i^* \log \pi_i^* + (1 - y_i^*) \log(1 - \pi_i^*)] \\ &= \sum_i y_i^* f(\mu, \beta_j^*) + \sum_i \log(1 + \exp\{f(\mu, \beta_j^*)\}). \end{aligned} \tag{15}$$

Under the null model, $\mathcal{M}_0$ with $j = 0$, we assume that $\beta_j^* = 0$ and a normal prior on μ

$$p(\mu \,|\, \mathcal{M}_0) \propto \exp\left\{-\frac{1}{2\sigma_\mu^2}\mu^2\right\}.$$

For the alternative models, where $j = 1, \ldots, p^*$, we put a normal prior on the gene coefficients where

$$p(\mathbf{\Theta}^* \,|\, \mathcal{M}_j) = p(\mu, \beta_j^* \,|\, \mathcal{M}_j) \propto \exp\left\{-\frac{1}{2\sigma_\mu^2}\mu^2 - \frac{1}{2\sigma_{\beta_j}^2}\beta_j^{*2}\right\},$$

with $\mathbf{\Theta}^* = (\mu, \beta^*)$. Hence for each j, Equation (12) then becomes the following:

$$\text{BF}_{j0} = \frac{p(\mathbf{y}^* \,|\, \mathcal{M}_j)}{p(\mathbf{y}^* \,|\, \mathcal{M}_0)} = \frac{\int \ell(\mathbf{\Theta}^*; \mathbf{y}^*, \mathcal{M}_j) p(\mathbf{\Theta}^* \,|\, \mathcal{M}_j) d\mathbf{\Theta}^*}{\int \ell(\mu; \mathbf{y}^*, \mathcal{M}_0) p(\mu \,|\, \mathcal{M}_0) d\mu} \tag{16}$$

Following previous studies (Guan and Stephens, 2008), we may approximate each of the integrals by the Laplace method

$$\int \exp\left\{h(\boldsymbol{\Theta}^*)\right\} d\boldsymbol{\Theta}^* \approx (2\pi)^{\frac{\mathrm{d}}{2}} |H_{\boldsymbol{\Theta}^{**}}|^{-\frac{1}{2}} \exp\left\{h(\boldsymbol{\Theta}^{**})\right\}. \tag{17}$$

where d is the dimension of the integral being approximated, $\boldsymbol{\Theta}^{**}$ is the value at which h is at its maximum, and $|H_{\boldsymbol{\Theta}^{**}}|$ is the absolute value of the determinant of the Hessian matrix of h evaluated at $\boldsymbol{\Theta}^{**}$. Under the null, $\mathcal{M}_0$, and alternative models, $\mathcal{M}_j$, respectively:

$$h(\mu) = \ell(\mu; \mathbf{y}^*) + \log p(\mu);$$
$$h(\boldsymbol{\Theta}^*) = \ell(\boldsymbol{\Theta}^*; \mathbf{y}^*) + \log p(\boldsymbol{\Theta}^*).$$

The derivation of the Hessian matrices in each model case are given in full detail in cited works (Guan and Stephens, 2008).

Bayes Factor Interpretation

We computed the Bayes factor in Equation (16) using the approximation in Equation (17) a total of p^* times — each time comparing one signature gene in $\boldsymbol{\mathcal{G}}$ to the null model containing just the intercept. Taking the log of each ratio j, we examined across results and considered genes with $\log \mathrm{BF}_{j0} > 0$ to track well with progression. Therefore, the set of signature genes satisfying this condition are called *relapse genes* and said to be members of the set $\boldsymbol{\mathcal{R}}$. Alternatively, we say that $\boldsymbol{\mathcal{R}} = \{j : \log \mathrm{BF}_{j0} > 0\}$. To further aid in this interpretation, we also implement a t-test on the relapse genes to show that the Bayes Factor results also mirror that of a standard p-value (see Table S1).

Gene Set Analysis

To investigate the biological relevance of the signature $\boldsymbol{\mathcal{G}}$ and drivers of relapse $\boldsymbol{\mathcal{R}}$, we cross referenced each member of the two sets with gene sets compiled in the molecular signature database (MSigDB) compiled at the Broad Institute (Subramanian *et al.*, 2005). We used gene set enrichment methods to identify classes of genes that are over/underrepresented given a signature, for example $\boldsymbol{\mathcal{G}}$ or $\boldsymbol{\mathcal{R}}$. In this study, we utilized the global test (Goeman *et al.*, 2004) and its corresponding R package `gloabaltest` for gene set analysis.

Briefly, this statistical approach points to whether or not significantly enriched/depleted groups have a higher association with a given phenotype than what is expected by chance. Specifically, it utilizes a logistic regression model on expression measurements to describe a binary phenotype (e.g. $1 \equiv$ treated sample; $0 \equiv$ pre-treated sample). The null hypothesis assumes that all regression coefficients for the members in a particular gene set are zero. Alternatively, we describe this procedure as a test to see if the components within a tested gene set have substantial predictive ability for classifying the targeted phenotypic response (Goeman *et al.*, 2004). This is determined by analyzing the random biological variation between subjects, rather than comparing a gene set with random counterparts. As previously shown (Goeman *et al.*, 2006; Goeman *et al.*, 2011), the global test is designed to have optimal power in the situation where a gene set has many small non-zero regression coefficients. Hence, it is directed to find gene sets for which many genes are associated with the phenotype, even if said relation is minimal. We therefore use the global test model as a way to look for "dysregulation" and differential expression among gene sets. A gene set is considered to be dysregulated if it has a Benjamini-Hochberg corrected p-value (BH q-value) below 0.05.

Determining *MYC* Reactivation on Relapse

Here, we detail the procedure used to determine the probability of *MYC* reactivation in each of the progressed tumor samples. Once again, we defined reactivation as a gene that satisfies the aforementioned *transitional requirements*. For this computation, we specifically looked at the percent change in *MYC* expression between each pre-treated tumor and its matched relapse sample(s). The intuition behind computing this change is that if *MYC* is said to be reactivated when a patient relapses, then its expression levels should be about the same or greater in the corresponding matched progressed tumors. For every patient i, we used the following to compute the percent change in *MYC* expression:

$$\%\text{change}(i) = \frac{\text{pre}(i) - \text{post}(i)}{\text{pre}(i)} \times 100\% \quad \text{for} \quad i = 1, \dots, n.$$

These percent changes were then standardized into z-scores and the probability of reactivation was computed via the transformation $1 - \Phi(z)$, where Φ is the cumulative distribution function of the standard normal distribution. Qualitatively, this value $\Phi(z)$ is similar to a lower one-tailed p-value and is interpreted as the probability of seeing something more extreme than what has already been observed in the data. In our context, this value can be defined as the probability of seeing *MYC* deactivated when a patient re-

lapses. Therefore, we took the complement of a patient relapsing, which we state as $1 - \Phi(z)$, to represent the probability of seeing *MYC* reactivated when a patient relapses. Progressed tumor samples that had probability above 0.5 (i.e. greater than random chance) were said to have *MYC* reactivated.

2 Cell Lines

A375, Colo679, UACC-62, Malme-3M, WM793 and WM1745 cells were grown in RPMI 1640 (Life Technologies Corporation, Carlsbad, CA) supplemented with 10% fetal bovine serum (Sigma-Aldrich Corporation, St. Louis, MO) and 1% penicillin/streptomycin (Life Technologies Corporation). SK-MEL-28, SK-MEL-5, A2058, RPMI-7951, Lox IMVI and Hs294T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Life Technologies Corporation) with 10% fetal bovine serum and 1% penicillin/streptomycin.

3 Chemicals

2-DG was prepared in PBS at 1 M. VX-11E, piperlongumine and PD0332 were prepared as 20 mM stock solutions in DMSO. CHIR-99021, HMR-1275, AZD1152, LY2603618, and AZD7762 were prepared in DMSO at 10 mM. CYC202 was prepared in DMSO at 5 mM. All other inhibitors were prepared as 100 mM stock solutions in DMSO.

4 GI_{50} Assay

To measure the GI_{50} values of specific inhibitors, cells were trypsinized and seeded at 5,000 cells/well in 96-well plates. After a 24-hour incubation, diluent (typically DMSO) or concentrated 10-fold dilutions of the indicated inhibitors (at 1:1000) were added to the cells to yield the highest concentration (see Chemicals). After a 3-day incubation with the treatment, cell viability was assessed with the CellTiter-Glo luminescent viability assay (Promega Corporation, Durham, NC) according to manufacturer's instructions. Growth inhibition was calculated as a percentage of diluent-treated cells and GI_{50} values were determined.

5 Sensitization Assays

In order to quantify the effect of resistance pathway inhibition on the sensitivity of cell lines to ERK pathway inhibition, small molecule sensitization assays were performed as previously described (Martz *et al.*, 2014). The ERK pathway inhibitor GI_{50} values were determined as described in the previous section with the addition of DMSO added to the media at a 1:1000 dilution. The ERK pathway inhibitor GI_{50} values were also determined with the indicated inhibitor in the background at the indicated dose. GI_{50} values were then determined to be the dose of ERK pathway inhibitor that resulted in half-maximal growth inhibition relative to the viability of non-ERK pathway inhibitor-only wells.

6 Immunoblotting

In order to measure protein levels in whole cell lysates, aliquots of cell extracts prepared in lysis buffer (0.5% Triton X-100, 50 mM β-glycerophosphate (pH 7.2), 0.1 mM Na3VO4, 2 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, 0.3 M NaCl, 2 μg/mL leupeptin and 4 μg/mL aprotinin (Sigma-Aldrich)) were submitted to SDS-PAGE. Where indicated, proteins were run on Mini-PROTEAN TGX Stain-Free Precast Gels (Bio-Rad) and total protein was visualized on the ChemiDoc Imaging System (Biorad). After electrophoretic transfer to PDMF, filters were blocked in 5% BSA and probed overnight at 4°C with the following primary antibodies and dilutions: c-MYC (1:200; #764 Santa Cruz Biotechnology, Dallas, TX or 1:10,000 ab32072 Abcam, Cambridge, MA), phospho-ERK (1:1000; #4376 Cell Signaling Technology, Danvers, MA), ERK1/2 (1:1000; #4695 Cell Signaling Technology), Notch1 (1:1000; #3608 Cell Signaling Technology), phospho-AKT1 (1:1000; #13038 Cell Signaling Technology), AKT1 (1:1000; #4691 Cell Signaling Technology), vinculin (1:500; #4650 Cell Signaling Technology), α-tubulin (1:1000, #2125 Cell Signaling Technology), β-actin (1:1000, #4970 Cell Signaling Technology). For quantification of immunoblots, where indicated, densitometry was performed with ImageJ software, background was subtracted and band intensity was normalized to loading control intensity. Alternatively, total protein images and Image Lab software (Bio-Rad) was used to normalize band intensity to total lane protein content.

7 Lentivirus Preparation and DNA Constructs

All expression clones were prepared in lentiviral form as previously described (Martz *et al.*, 2014). In brief, vectors were packaged in 293T cells with an overnight incubation with Fugene (Promega Corporation), pΔVPR and pVSV-G. The virus-containing media was collected after 48 and 72 hours and filtered with a 0.45 μm filter and stored at -80°C until use with 16 μg/mL polybrene (Sigma-Aldrich). The shRNA constructs and other expression vectors are listed in Table S6.

8 *In vitro* Adaption of Inhibitor Resistant Cells

Parental cells were either exposed to escalating doses of inhibitor until logarithmic growth resumed or exposed to a high dose (3 μM) of inhibitor (PLX4720, AZD6244 or VX-11E) and the resultant resistant clones were expanded and cultured. Parental cell lines were cultured concurrently with DMSO. Resistant cell lines were maintained in routine culture with the addition of 3 μM inhibitor. All resistant and DMSO parental control lines were submitted to STR profiling by the Duke University DNA Analysis Facility upon the acquisition of resistance in order to confirm their authenticity.

9 Annexin V Apoptosis Assay

The induction of apoptosis was quantified as described previously (Martz *et al.*, 2014). Briefly, cells were plated in triplicate at 200,000 cells per well in six-well plates. The following day, the growth media was removed and replaced with fresh media containing the indicated dose of drug or diluent (typically DMSO). After a 72-hour incubation in drug, cells were washed in PBS twice and resuspended in a buffer composed of 10 mM HEPES, 140 mM NaCl and 2.5 mM $CaCl_2$ (BD Biosciences, San Jose, CA). Apoptosis was quantified using allophycocyanin-conjugated Annexin V and viability was assessed with 7-Amino-actinomycin D (BD Bioscience). Gating was defined using untreated/unstained cells and treatments were evaluated at 20,000 counts using BD FACSVantage SE.

10 RNA Extraction and Quantitative Real-Time PCR

PCR primers were obtained from Integrated DNA Technologies and are as follows: ACTB forward 5'-CTTCCAGCCTTCCTTCCTGG-3,' reverse 5'-AATGCCAGGGTACATGGTGG-3,' MYC forward 5'-CCACC AGCAGCGACTCTG-3,' reverse 5'-TGTGAGGAGGTTTGCTGTGG-3'. Average cycle threshold (C_t) values were determined for MYC and normalized to the reference gene, β-actin. Relative gene expression was determined using the $\Delta\Delta C_t$ method.

11 Gene Expression Analysis

A375 cells (parental and evolved PLX4720-resistant clonal derivatives) expressing shGFP (parental and resistant) or shMYC (resistant only) were grown to ~80% confluency in triplicate in normal growth media and submitted to gene expression analysis. RNA-Seq libraries were generated using I-L-070 Kapa stranded mRNA-seq kit (Kappa Biosystems (Wilmington, MA) and the final libraries checked for quality control on a Qubit Fluorometer (Thermo Fisher) and Agilent 2200 Tapestation (Agilent Technologies), which was followed by sequencing on the Illumina HiSeq2000/2500 V4 (Illumina). RNA-seq data was processed by the Duke University Genome Analysis and Bioinformatics Core Facility using the TrimGalore toolkit which employs Cutadapt to trim low quality bases and Illumina sequencing adapters from the 3' end of the reads. Only reads that were 20nt or longer after trimming were kept for further analysis. Reads were mapped to the GRCm37r75 version of the human genome and transcriptome (Kersey *et al.*, 2012) using the STAR RNA-seq alignment tool (Dobin *et al.*, 2013). Reads were kept for subsequent analysis if they mapped to a single genomic location. Gene counts were compiled using the HTSeq tool. Only genes that had at least 10 reads in any given library were used in subsequent analysis. Normalization and differential expression was carried out using the DESeq2 (Love *et al.*, 2014) Bioconductor (Huber *et al.*, 2015) package with the R statistical programming environment. The false discovery rate was calculated to control for multiple hypothesis testing. Gene set enrichment analysis (Mootha *et al.*, 2003) was performed to identify differentially regulated pathways and gene ontology terms for each of the comparisons performed.

References

Barbieri, M. M. and Berger, J. O. (2004). Optimal predictive model selection. *Annals of Statistics 32*, 870–897.

Băzăvan, E. G., Li, F. and Sminchisescu, C. (2012). *Fourier Kernel Learning* pp. 459–473. Berlin, Heidelberg: Springer Berlin Heidelberg.

Bender, L. M. and Nahta, R. (2008). HER2 Cross Talk and Therapeutic Resistance in Breast Cancer. *Frontiers in bioscience : a journal and virtual library* , *13*, 3906–3912.

Berger, J. O. and Pericchi, L. R. (1996). The Intrinsic Bayes Factor for Model Selection and Prediction. *Journal of the American Statistical Association* , *91*, 109–122.

Bochner, S. (1934). A Theorem on Fourier-Stieltjes Integrals. *Bulletin of the American Mathematical Society* , *40*, 271–276.

Bostock, R. M. (2005). Signal Crosstalk and Induced Resistance: Straddling the Line Between Cost and Benefit. *Annual Review of Phytopathology* , *43*, 545–580.

Chang, Y.-W., Hsieh, C.-J., Chang, K.-W., Ringgaard, M. and Lin, C.-J. (2010). Training and testing low-degree polynomial data mappings via linear SVM. *J. Machine Learning Research* , *11*, 1471–1490.

Crawford, L., Wood, K. C., Zhou, X. and Mukherjee, S. (2017). Bayesian approximate kernel regression with variable selection. *Journal of the American Statistical Association*.

Crossa, J., Pérez, P., Hickey, J., Burgueño, J., Ornella, L., Cerón-Rojas, J., Zhang, X., Dreisigacker, S., Babu, R., Li, Y., Bonnett, D. and Mathews, K. (2014). Genomic Prediction in CIMMYT Maize and Wheat Breeding Programs. *Heredity* , *112*, 48–60.

de los Campos, G., Gianola, D., Rosa, G. J. M., Weigel, K. A. and Crossa, J. (2010). Semi-Parametric Genomic-Enabled Prediction of Genetic Values using Reproducing Kernel Hilbert Spaces Methods. *Genetics Research (Cambridge)* , *92*, 295–308.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* , *29*, 15–21.

Efron, B. (2007). Size, power and false discovery rates. *Annals of Statistics* , *35*, 1351–1377.

Gelman, A. and Tuerlinckx, F. (2013). Type S error rates for classical and Bayesian single and multiple comparison procedures. *Computational Statistics* , *15*, 373–390.

Gianola, D., Fernando, R. L. and Stella, A. (2006). Genomic-Assisted Prediction of Genetic Value With Semiparametric Procedures. *Genetics* , *173*, 1761–1776.

Goeman, J. J., van de Geer, S. A., de Kort, F. and van Houwelingen, H. C. (2004). A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics* , *20*, 93–99.

Goeman, J. J., van de Geer, S. A. and van Houwelingen, H. C. (2006). Testing against a High Dimensional Alternative. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* , *68*, 477–493.

Goeman, J. J., van Houwelingen, H. C. and Finos, L. (2011). Testing against a high-dimensional alternative in the generalized linear model: asymptotic type I error control. *Biometrika* , *98*, 381–390.

Guan, Y. and Stephens, M. (2008). Practical issues in imputation-based association mapping. *PLoS Genetics* , *4*, e1000279.

Guan, Y. and Stephens, M. (2011). Bayesian variable selection regression for genome-wide association studies and other large-scale problems. *Annals of Applied Statistics* , *5*, 1780–1815.

Hauck Jr, W. W. and Donner, A. (1977). Wald's test as applied to hypotheses in logit analysis. *Journal of the american statistical association* , *72*, 851–853.

Howard, R., Carriquiry, A. L. and Beavis, W. D. (2014). Parametric and Nonparametric Statistical Methods for Genomic Selection of Traits with Additive and Epistatic Genetic Architectures. *G3 (Bethesda)* , *4*, 1027–1046.

Huber, W., Carey, V. J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B. S., Bravo, H. C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K. D., Irizarry, R. A., Lawrence, M., Love, M. I., MacDonald, J., Obenchain, V., Oles, A. K., Pages, H., Reyes, A., Shannon, P., Smyth, G. K., Tenenbaum, D., Waldron, L. and Morgan, M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Meth* , *12*, 115–121.

Kass, R. E. and Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association* , *90*, 773–795.

Kersey, P. J., Staines, D. M., Lawson, D., Kulesha, E., Derwent, P., Humphrey, J. C., Hughes, D. S. T., Keenan, S., Kerhornou, A., Koscielny, G., Langridge, N., McDowall, M. D., Megy, K., Maheswari, U., Nuhn, M., Paulini, M., Pedro, H., Toneva, I., Wilson, D., Yates, A. and Birney, E. (2012). Ensembl Genomes: an integrative resource for genome-scale data from non-vertebrate species. *Nucleic Acids Research* , *40*, D91–D97.

Kinney, S. K. and Dunson, D. B. (2007). Fixed and Random Effects Selection in Linear and Logistic Models. *Biometrics* , *63*, 690–698.

Kleinbaum, D. G. and Klein, M. (2010). *Maximum Likelihood Techniques: An Overview*, pp. 103–127. Statistics for Biology and Health. New York, NY: Springer New York.

Lasorella, A., Noseda, M., Beyna, M. and Iavarone, A. (2000). Id2 is a retinoblastoma protein target and mediates signalling by Myc oncoproteins. *Nature* , *407*, 592–598.

Li, Q., Birkbak, N. J., Gyorffy, B., Szallasi, Z. and Eklund, A. C. (2011). Jetset: selecting the optimal microarray probe set to represent a gene. *BMC Bioinformatics* , *12*, 474.

Liang, F., Mao, K., Mukherjee, S. and West, M. (2009). Nonparametric Bayesian kernel models. Department of Statistical Science, Duke University, Discussion Paper.

Liang, F., Mukherjee, S. and West, M. (2007). The Use of Unlabeled Data in Predictive Modeling. *Statistical Science* , *22*, 189–205.

Lippert, C., Listgarten, J., Liu, Y., Kadie, C. M., Davidson, R. I. and Heckerman, D. (2011). FaST linear mixed models for genome-wide association studies. *Nat Meth* , *8*, 833–835.

Logsdon, B. A., Hoffman, G. E. and Mezey, J. G. (2010). A variational Bayes algorithm for fast and accurate multiple locus genome-wide association analysis. *BMC Bioinformatics* , *11*, 1–13.

Lopes, H. F. and West, M. (2004). Bayesian Model Assessment in Factor Analysis. *Statistica Sinica* , *14*, 41–67.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* , *15*, 1–21.

Martz, C. A., Ottina, K. A., Singleton, K. R., Jasper, J. S., Wardell, S. E., Peraza-Penton, A., Anderson, G. R., Winter, P. S., Wang, T., Alley, H. M., Kwong, L. N., Cooper, Z. A., Tetzlaff, M., Chen, P.-L., Rathmell, J. C., Flaherty, K. T., Wargo, J. A., McDonnell, D. P., Sabatini, D. M. and Wood, K. C. (2014). Systematic identification of signaling pathways with potential to confer anticancer drug resistance. *Science Signaling* , *7*, ra121–ra121.

Mootha, V. K., Lindgren, C. M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., Houstis, N., Daly, M. J., Patterson, N., Mesirov, J. P., Golub, T. R., Tamayo, P., Spiegelman, B., Lander, E. S., Hirschhorn, J. N., Altshuler, D. and Groop, L. C. (2003). PGC-1[alpha]-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* , *34*, 267–273.

Nazarian, R., Shi, H., Wang, Q., Kong, X., Koya, R. C., Lee, H., Chen, Z., Lee, M.-K., Attar, N., Sazegar, H., Chodon, T., Nelson, S. F., McArthur, G., Sosman, J. A., Ribas, A. and Lo, R. S. (2010). Melanomas acquire resistance to BRAF(V600E) inhibition by RTK or N-RAS upregulation. *Nature* , *468*, 973–977.

Peng, C.-Y. J., Lee, K. L. and Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The journal of educational research* , *96*, 3–14.

Pillai, N. S., Wu, Q., Liang, F., Mukherjee, S. and Wolpert, R. (2007). Characterizing the Function Space for Bayesian Kernel Models. *Journal of Machine Learning Research* , *8*, 1769–1797.

Pratilas, C. A., Taylor, B. S., Ye, Q., Viale, A., Sander, C., Solit, D. B. and Rosen, N. (2009). V600EBRAF is associated with disabled feedback inhibition of RAF–MEK signaling and elevated transcriptional output of the pathway. *Proceedings of the National Academy of Sciences* , *106*, 4519–4524.

Rahimi, A. and Recht, B. (2007). Random Features for Large-Scale Kernel Machines. *NIPS* , *3*, 5.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* , *43*, e47–e47.

Rizos, H., Menzies, A. M., Pupo, G. M., Carlino, M. S., Fung, C., Hyman, J., Haydu, L. E., Mijatov, B., Becker, T. M., Boyd, S. C., Howle, J., Robyn, S., Thompson, J. F., Kefford, R. F., Scolyer, R. A. and Long, G. V. (2014). BRAF inhibitor resistance mechanisms in metastatic melanoma; spectrum and clinical impact. *Clinical Cancer Research* , *20*.

Rouder, J. N. and Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research* , *47*, 877–903.

Rudy, J. and Valafar, F. (2011). Empirical comparison of cross-platform normalization methods for gene expression data. *BMC Bioinformatics* , *12*, 1–22.

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* , *3*, Article3.

Stephens, M. (2017). False discovery rates: a new deal. *Biostatistics* , *18*, 275–294.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. and Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* , *102*, 15545–15550.

Sun, Y., Liu, X., Eaton, E. N., Lane, W. S., Lodish, H. F. and Weinberg, R. A. (1999). Interaction of the Ski Oncoprotein with Smad3 Regulates TGF- Signaling. *Molecular Cell* , *4*, 499–509.

Vert, J.-P., Tsuda, K. and Schölkopf, B. (2004). A primer on kernel methods. *Kernel Methods in Computational Biology* , *47*, 35–70.

West, M. (2003). Bayesian Factor Regression Models in the "Large p, Small n" Paradigm. *Bayesian Statistics* , *7*.

Williams, C. K. I. and Seeger, M. (2001). Using the Nyström Method to Speed Up Kernel Machines. In *Advances in Neural Information Processing Systems 13*, (Leen, T. K., Dietterich, T. G. and Tresp, V., eds)„ pp. 682–688. MIT Press.

Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M. and Lin, X. (2011). Rare-Variant Association Testing

for Sequencing Data with the Sequence Kernel Association Test. *The American Journal of Human Genetics* , *89*, 82–93.

Yamaguchi, H., Chang, S.-S., Hsu, J. and Hung, M.-C. (2014). Signaling cross-talk in the resistance to HER family receptor targeted therapy. *Oncogene* , *33*, 1073–1081.

Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S. and Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* , *38*, 203–208.

Zeileis, A. and Hothorn, T. (2002). Diagnostic Checking in Regression Relationships. *R News* , *2*, 7–10.

Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., Bradbury, P. J., Yu, J., Arnett, D. K., Ordovas, J. M. and Buckler, E. S. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat Genet* , *42*, 355–360.

Zhou, X., Carbonetto, P. and Stephens, M. (2013). Polygenic Modeling with Bayesian Sparse Linear Mixed Models. *PLoS Genet* , *9*, e1003264.

Zhou, X. and Stephens, M. (2012). Genome-wide Efficient Mixed Model Analysis for Association Studies. *Nat Genet* , *44*, 821–824.