**Generation of B508 Resistant Clones**

BMT-819/B508 resistant clones were generated from a parental barcoded H460 cell line (described above). H460 cells were plated onto 10x10cm plates to achieve ~20% confluence at the start of experiment. The following day, B508 was added at 150 nM in complete RPMI and medium was changed every other day with fresh medium containing 150 nM B508. Upon culture to 90% confluence, plates were passaged by diluting 1x10cm plate of H460 cells 1:5 into fresh B508 containing medium for ~5-20 passages, until cells had completely arrested. After ~7 days following arrest, individual clones emerged and were cultured for an additional ~7-14 days in the presence of B508. One individual colony per 10 cm plate was selected by cloning cylinder and transferred to 12-well plates. Each well was then grown to confluence and each clone expanded to 3x10 cm plates for DNA extraction, storage in liquid nitrogen or assay by cell-titer glo in comparison to paclitaxel and 5-Fluorouracil. DNA was purified from 50 clones that were resistant to BMT-819 and B508 but remained sensitive to the antimitotic paclitaxel, by purification with purelink genomic DNA mini kit. Barcode sequences from each clone were amplified from 200 ng of DNA by mixing with 2.5 μL of 10 μM oligo3, 2.5 μL of 10 μM oligo4, 10 μL of 5× Standard Phusion buffer, 1 μL of 10 mM dNTP mix, 0.5 μL Phusion polymerase and 31.5 μL of H2O and amplified using 34 cycles of the following sequence: 98°C 0:15, 98°C 0:10, 59°C 0:15, 72°C 0:30. PCR reactions were cleaned using a DNA clean and concentrator kit and submitted for sanger sequencing using oligo 5. Ten clones that had a unique barcode sequence and acquired resistance to BMT-819/B508 but not paclitaxel, were submitted for whole genome sequencing.

**PCR-Free Whole Genome Sequencing**

Genomic DNA was extracted using purelink genomic DNA kit and 350 ng per sample diluted in 50µL aliquots was acoustically sheared using a Covaris focused-ultrasonicator, targeting 385bp fragments. Following fragmentation, additional size selection was performed using a SPRI cleanup. Library preparation was performed using a commercially available kit provided by KAPA Biosystems (KAPA Hyper Prep without amplification module, product KK8505), and with palindromic forked adapters with unique 8-base index sequences embedded within the adapter (purchased from Roche). Following sample preparation, libraries were quantified using quantitative PCR (kit purchased from KAPA Biosystems), with probes specific to the ends of the adapters using Agilent’s Bravo liquid handling platform. Based on qPCR quantification, libraries were normalized to 2.2 nM and pooled into 24-plexes. Sample pools are combined with NovaSeq Cluster Amp Reagents DPX1, DPX2 and DPX3 and loaded into single lanes of a NovaSeq 6000 S4 flowcell cell using the Hamilton Starlet Liquid Handling system. Cluster amplification and sequencing occur on NovaSeq 6000 Instruments utilizing sequencing-by-synthesis kits to produce 151 bp paired-end reads.

**Analysis of Whole Genome Sequencing**

Genomes sequenced to at least ~30x coverage were assessed with NGS QC Toolkit (v2.3.1), aligned to the reference genome GRCHG38 using BWA-MEM v0.7.1.5r1140, sorted and deduplicated using picard, recalibrated using known sites taken from the latest dbSNP released on 05-14-20 with BaseRecalibrator, and indexed with samtools. Germline variants in reference samples were called from sorted analysis ready BAM files using HaplotypeCaller and somatic mutations arising in resistant clones were called against the original barcoded H460 cell line using a combination of three variant callers: Mutect2, VarScan2, and MuSE according to best practices2. For Mutect2, variants were called with max\_alt\_alleles\_in\_normal\_count raised to 10e6 and max\_alt\_allele\_in\_normal\_fraction set at 0.05. Individual VCF files were hard filtered using the following thresholds: FS>30.0, QD<5.0, DP<10.0, GQ<20.0 with a cluster-window-size of 35. To filter for variants that accumulated during the clonal expansion step, we only considered somatic mutations with a variant allele frequency (VAF) greater than 0.2 or less than 0.7. We removed SNVs and Indels that had minor allele frequency (MAF) of 1% or higher in either 1000Genomes (phase 3) or gnomAD (r2.0.1), and SVs that overlapped DGV and 1000Genomes (phase3). VCFs were then annotated with Ensembl Variant Effect Predictor (VEP), and previously unidentified coding sequence variants were extracted with filter\_VEP using the flags “not Existing\_variation” and “Consequence is coding\_sequence\_variant”. Missense mutations co-occurring across resistant clones within the same gene were extracted by merging VCFs by gene name. Variants that passed all the above-mentioned filters were included in our final somatic call set. Copy number variation between resistant clones and a panel of normals generated from WGS data across five parental H460 cell samples was analyzed by both CNVKit and GATK. For GATK analysis, reads from individual resistant clones were collected using GATK CollectReadCounts. The panel of normals was used to call copy ratios and denoise using GATK DenoiseReadCounts.  Denoised copy-number ratios were plotted per-chromosome. Variants identified by WGS within the OXA1L gene were validated by PCR amplification of the OXA1L gene using oligos OXA1L\_F and OXA1L\_R and sanger sequencing of the purified OXA1L PCR product using oligos OXA1L 1S-10S. For PCR amplification of OXA1L gene, 50 uL PCR reactions using 250 ng of genomic DNA purified from each clone using purelink genomic DNA kit was mixed with 2.5 μL of 10 μM OXA1L\_F, 2.5 μL of 10 μM OXA1L\_R, 10 μL of 5× Standard Phusion buffer, 1 μL of 10 mM dNTP mix, 0.5 μL Phusion polymerase and 31.5 μL of H2O and amplified using 34 cycles with the following sequence: 98°C 0:15, 98°C 0:10, 59°C 0:15, 72°C 2:30, and final extension 72°C 10:00.

OXA1L\_F: 5′-CGCAAAAGCAAGTCCTCTTC-3′

OXA1L\_R: 5′-GCTAGCTGCTGAAGGCTGAT-3′

Oxa1l FOR\_1: 5′-GAA CCC AGG TTC GAG CTT -3′

Oxa1l FOR\_2: 5′-TGG TGT TTC CTG GTT GGT TT -3′

Oxa1l FOR\_3: 5′-GGC TAA CCT GGA AGA ACA AGA A -3′

Oxa1l FOR\_4: 5′-AAG TAG TTG CAG AAG CTC TGG G -3′

Oxa1l FOR\_5: 5′-GCA CAC CCA TGT GTT ACC TG -3′

Oxa1l FOR\_6: 5′-CAG CCT CTC AAG TAG CTG GG -3′

Oxa1l FOR\_7 : 5′-CAT TTA CTC CCT CAG CCG TC -3′

Oxa1l FOR\_8 : 5′-GCA GCC AGG ATC CAC AAT -3′

Oxa1l FOR\_9 : 5′-TCT ATA AAC CTC TCA TTC TCC CTG -3′

Oxa1l FOR\_10 : 5′-AGT GCC AGG CCT GCA AAG -3′

Oxa1l FOR\_11 : 5′-CTC CAA TTT GTT TTC CCT GG -3′

Oxa1l FOR\_12 : 5′-CAG ACC TTT ACC CAC AAC CC -3′

Oxa1l FOR\_13: 5′-ATA CCC AGC ATT TTG GGA GG -3′

Oxa1l FOR\_14: 5′-ACT TTT GGG AGG CCA AGG -3′

Oxa1l REV\_1: 5′-GGA GGA CAT TGC TAT CAC CG -3′

Oxa1l REV\_2: 5′-TTT TGG AGA GCT GTG TGG AG -3′

Oxa1l REV\_3: 5′-ATT CAG TAG CGA ACA CAA AGC A -3′

Oxa1l REV\_4: 5′-GGC AGA GTG CAT CTA GGT ATC A -3′

Oxa1l REV\_5: 5′-CAG CCA TCA AAG GAA GAA GA -3′

Oxa1l REV\_7: 5′-CAA AAA GCT GAC AGG AGT GGA -3′

Oxa1l REV\_8: 5′-CAC ACT CAA TAT GGT CTC CTG C -3′

Oxa1l REV\_9: 5′-GGA AAT TTC ATG GGT ACT GGA -3′

Oxa1l REV\_10: 5′-CCA TCA ACT CTT TCT CCA GCA -3′

Oxa1l REV\_11: 5′-GTC ATG TAC AAC ACG CTG GG -3′

Oxa1l REV\_12: 5′-TTG ACT TTG GTT TGC TGC TG -3′

Oxa1l REV\_13: 5′-AGC AGA GAT GGC GTT TCC -3′

Oxa1l REV\_15: 5′-GTG TGA GCC ACT TTG CCC -3′

**Sample Preparation and RNA Sequencing**

Total RNA was extracted from 80% confluent 10 cm plates in triplicate across the cell line panel used for this study. A subset of cells was treated with B508 for 6 hours prior to RNA extraction. RNA was extracted using purelink RNA extraction kit according to manufacturer’s protocol. Total RNA samples were prepared into RNAseq libraries using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina sequencers following manufacturer’s recommended protocol. Briefly, for each sample 100 ng total RNA was polyA selected, fragmented, and then converted to double stranded cDNA followed by ligation of sequencing adapters. The library was then PCR amplified 13 cycles using barcoded PCR primers, purified, and pooled before loading onto an Illumina NovaSeq S2 flowcell for 2 x 150 base paired-end sequencing to an average depth of 100x106 reads per sample.

**Data Analysis of RNA Sequencing**

NGS QC Toolkit (v2.3.1) was used to check the sequencing quality, and high-quality reads were aligned to a 151 bp hg38 library using STAR. The htseq-count script distributed with the HTSeq Python package (0.6.1) was employed for counting reads in genes considering the coding strands. For normalization across cell lines and in response to BMT-819/B508 using trimmed mean of m values, the edgeR R script “EdgeR\_MultiVariate.R” was employed. For variant calling, we used GATK package according to suggested RNASeq sample analysis. Briefly, duplicates were removed from aligned and sorted bam files using Picard MarkDuplicates. Cigar strings were split using SplitNCigarReads and indexed using samtools. Germline variants were called using HaplotypeCaller with –don’t-use-soft-clipped-bases and –stand-call-conf 20.0 flags. Variants were then filtered using VariantFiltration with FS>60, MQ<40.0, MQRankSum<-12.5, ReadPosRankSum<-8.0, QD<2.0 and DP<1000. VCF files were annotated using VEP, and coding sequences were extracted for some analysis. For comparison of variants across treated and control groups, Bedtools intersect was used to generate variant sets across replicates, and then plotted per chromosome using the python script “Compare\_Variants\_CHR.py”. For QTL analysis, individual VCFs were combined and converted to .bim, .bed and .fam files based on a per cell line basis.  Plink –assoc was used to generate QTLs using sex as a covariate. Correlation analysis between missense mutations, RNA and protein expression were conducted using the custom python script “Correlation\_Mutations\_RNA.py” and “Correlation\_Mutations\_Protein.py” on merged data generated with the python script “Merge\_Mutations\_RNA\_Protein.py”.