

Supplemental Information

Widespread Genetic Heterogeneity in Multiple

Myeloma: Implications for Targeted Therapy

Jens G. Lohr, Petar Stojanov, Scott L. Carter, Peter Cruz-Gordillo, Michael S. Lawrence, Daniel Auclair, Carrie Sougnez, Birgit Knoechel, Joshua Gould, Gordon Saksena, Kristian Cibulskis, Aaron McKenna, Michael A. Chapman, Ravid Straussman, Joan Levy, Louise M. Perkins, Jonathan J. Keats, Steven E. Schumacher, Mara Rosenberg, The Multiple Myeloma Research Consortium, Gad Getz, and Todd R. Golub

SUPPLEMENTAL DATA

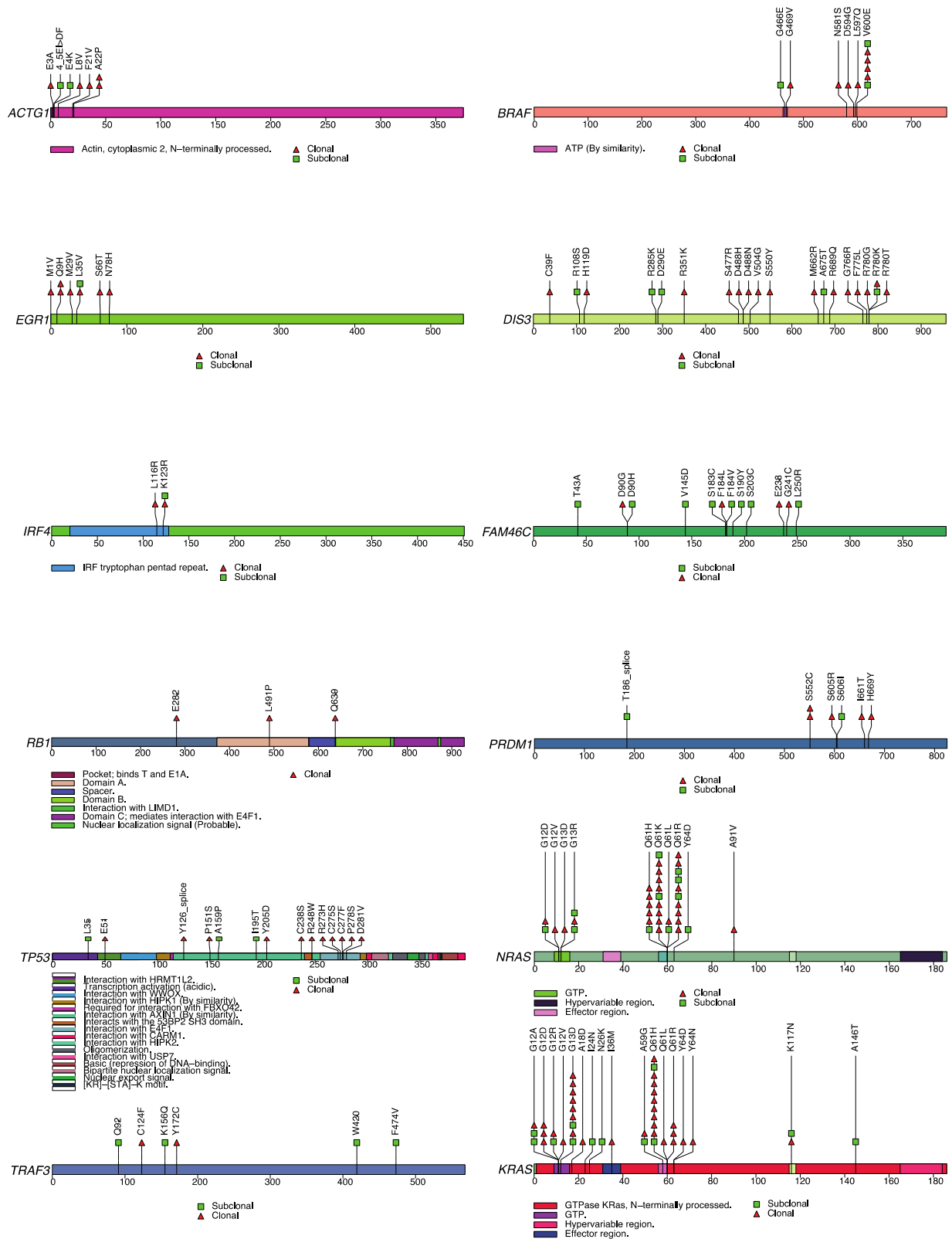


Figure S1, related to Figure 1 and 2. Schematic of the location of somatic mutations in significantly mutated genes. A diagram of the relative positions of somatic mutations is shown for significantly mutated genes in our cohort and mutated genes with potential biological relevance. Whether a mutation occurs in a clonal or subclonal fashion and protein domain information are indicated. Only mutations for which clonality information is available are displayed.

Table S1, related to Figure 1. Rank ordered list of the most significantly mutated genes in MM. (provided as an Excel file)

All significantly mutated genes identified in 203 MM patients are shown, rank-ordered by decreasing significance. Significance analysis was also performed on subgroups in which t(11;14) or t(4;14) was detected (individual sheets).

n_nonsilent, n_silent, and n_flank: The numbers of nonsilent, silent, and flanking mutations.

“p_CV” and “q_CV”: p and q values corresponding to the MutSigCV method as published and released by (Lawrence et al., 2013).

“p_homozygous” and “q_homozygous”: p and q values corresponding to prioritization of genes by enrichment of homozygous events.

“p_joint” and “q_joint”: p and q values corresponding to the prioritization of genes that are enriched in mutational hotspots that are significantly conserved (Lohr et al., 2012).

“p_combined” and “q_combined”: p and q values obtained by combining all of the three p values from the three metrics used to rank genes by significance.

NaN = not a number, if data are insufficient to calculate the value.

For more information on the metrics used for ranking genes by significance, and the method of combining p values, please see Supplemental Experimental Procedures.

Table S2, related to Figure 1. Detailed description of SSNVs in MM with analysis of purity, ploidy, clonality and cancer cell fraction. (provided as an Excel file)

Somatic events identified in 203 patients with MM by whole exome and whole genome sequencing are shown in sheet "All_mut". Subclonal coding mutations with LOH in regions with `closest_absolute_copy_number=2`, which occurred on segments with *subclonal* copy number changes are provided as sheet "Mut_LOH_subclonal_SCNA".

The following events were detected by the mutation caller: Missense Mutations, Nonsense Mutations, Splice Site alterations, Frameshift Deletions, Frameshift Insertions, In-Frame Insertions, Nonstop Mutations, In-Frame Deletions, Flanking mutations, Silent Mutations, Intron Mutations, 3' UTR mutations, 5' UTR mutations, Translation start site mutations. The sheet "Validation_targeted_seq" shows individual mutations, which were validated by deep targeted sequencing. Primers were designed for PCR amplification of sites with individual mutations. Individual amplicons were sequenced on an Illumina Miseq. Of 140 individual mutations, 122 were validated, 13 were invalidated, and 5 sites did not have coverage, accounting for a validation rate of 90.4%. NaN = not a number, if data are insufficient to calculate the value.

Table S3, related to Figure 1. Genes enriched with mutations in AID target motifs. (provided as an Excel file)

Significance analysis of WRCY hotspots, as performed previously (Lohr et al., 2012) (also see Experimental procedures). For each gene the following is shown: `num_wrcy`: number of muts in C or G basepairs in WRCY hotspots, `num_total`: number of total mutations, `num_total_C_or_G`: total number mutations in C or G basepairs, `territory_total`: total territory of the gene, `territory_wrcy`: number of C or G territory in WRCY hotspots, `territory_C_or_G_total`: total number of C or G territory in the gene, as well as the p and q values of the statistical analysis.

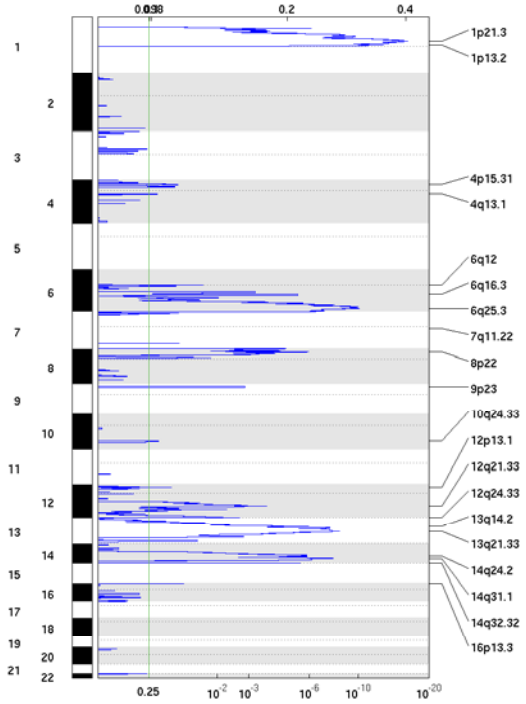
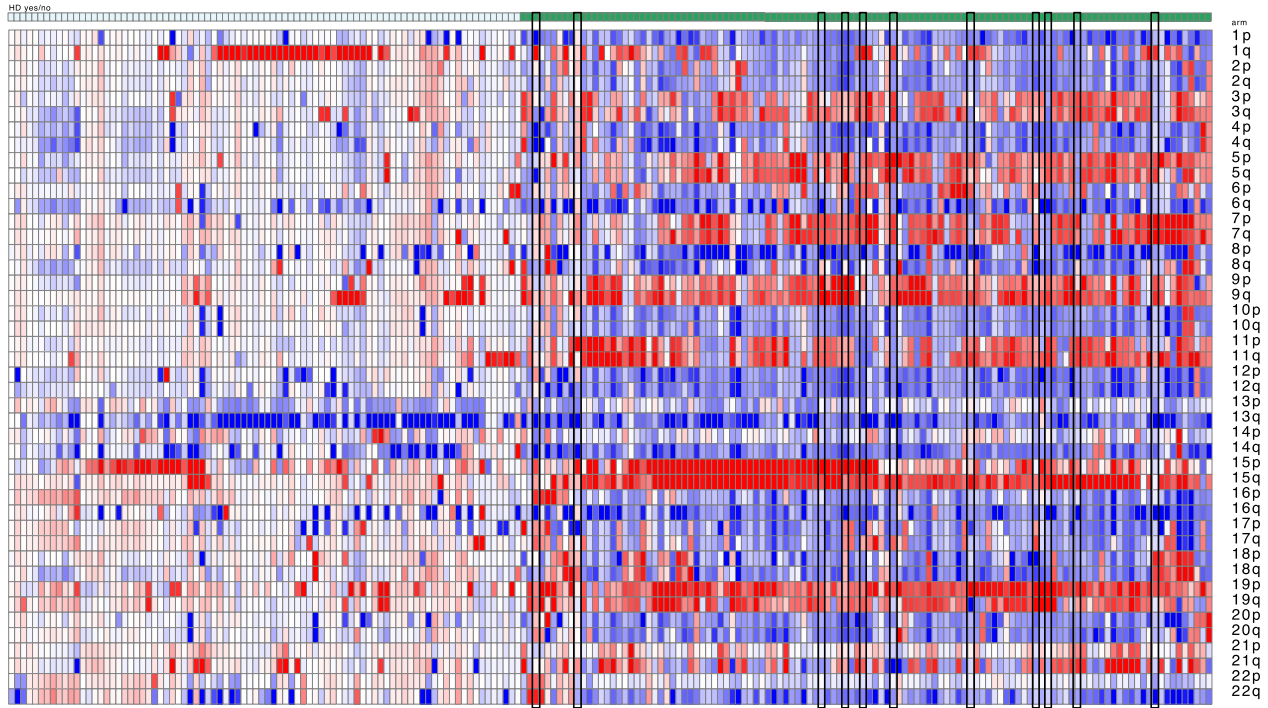


Figure S2, related to Figure 2.

(A) Distinguishing Hyperdiploid from Non-Hyperdiploid MM samples by WES/WGS. We determined the number of arm-level trisomies by WES/WGS, to distinguish hyperdiploid and non-hyperdiploid samples, as outlined in Supplemental Experimental Procedures. The normalized read density ratios for each individual chromosome arm of 202 MM samples are shown. A sample was called hyperdiploid (green symbols in annotation bar) if it had amplified at least 7 chromosome arms. Hyperdiploid samples, in which t(11;14) was identified by FISH, reported to be a rare event, are highlighted (black border). As is true for flow cytometry-based or FISH-based classification, a small number of samples is difficult to classify. Since WES/WGS allows identification of individual underlying chromosomal multiplication events with high resolution, this approach may help to classify challenging cases, such as samples with normal DNA mass in flow cytometric analysis, but chromosome arm amplifications and deletions that balance each other.

(B) GISTIC analysis of copy number variations. Segmented copy number data was obtained by SNP6.0 array from 153 patients. Amplifications (red) and deletions (blue), determined by segmentation analysis of normalized signal intensities were determined by the GISTIC method, as previously described (Beroukhi et al., 2007), and are displayed across the genome. The statistical significance of the aberrations are displayed as FDR q-values to account for multiple hypothesis testing. Chromosome positions are indicated along the y-axis with centromere positions indicated by dotted lines. The individual peaks and the genes they contain, and CNVs of genes located in the 99% confidence interval of these peaks ("Wide Peak Regions") are shown in Table S8.

Table S4, related to Figure 2. Patient and sample characteristics. (provided as an Excel file)

Y=Yes, N=No, NA=data not reported by the contributing institution.

Table S5, related to Figure 2. Prevalence of significantly mutated genes in HD vs. non-HD and pre-treated vs. untreated patients. (provided as an Excel file)

Differences in the frequency of mutations in the most significantly mutated genes between the indicated subgroups (hyperdiploid vs. non-hyperdiploid in sheet "HD_vs_non-HD", previously treated vs. untreated, in sheet "pre-treated_vs_untreated") by Fisher's exact test.

Table S6, related to Figure 2. Detection of loci with homozygous deletions. (provided as an Excel file)

Results from the Gistic2.0 (Beroukhim et al., 2007; Mermel et al., 2011) analysis on homozygous deletions. Each cytoband is shown with its respective Gistic2.0 q value, and the genes that belong to each peak.

Table S7, related to Figure 2. Enrichment of mutations in novel gene sets. (provided as an Excel file)

Significance analysis on GSEA canonical genesets as described by Chapman et al (Chapman et al., 2011). N_{genes} = number of genes in geneset, N = total number of bases of genes in geneset, n = number of mutations, n_{pat} = number of patients, n_{site} = number of individual sites, n_{sil} = number of silent mutations, num_homozygous = number of mutations with LOH.

Table S8, related to Figure 2. Focal amplifications and deletions by the GISTIC method.

(provided as an Excel file)

Focal amplifications and deletions across 153 patients with myeloma for whom segmented copy number data was obtained by SNP6.0 array. The most significant focal amplifications (sheet “Sign_GISTIC_amplifications”) and deletions (sheet “Sign_GISTIC_deletions”) were determined by the GISTIC method as outlined in Figure S2B. Raw copy number values of all genes within focally amplified or deleted peaks, located within the 99% confidence interval of the designated peaks are shown (sheet “Sign_CNV_by_gene”).

Table S9, related to Figure 2. Structural variations detected in 3 samples after whole genome sequencing. (provided as an Excel file)

For 3 previously unpublished samples in our current cohort whole genome sequencing was performed, in addition to the 23 samples for which whole genome sequencing and structural variations were reported previously (Chapman et al., 2011).

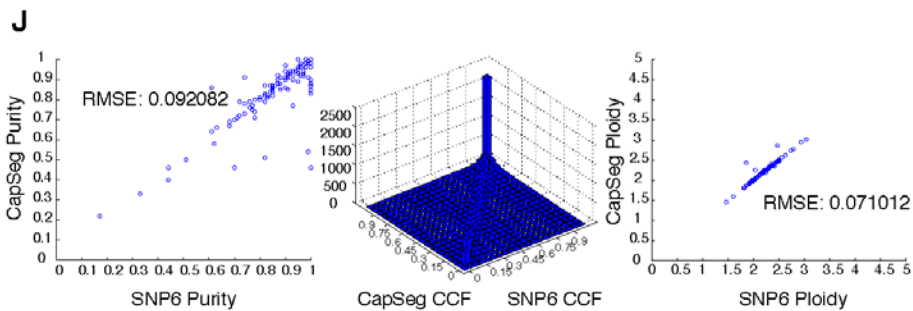
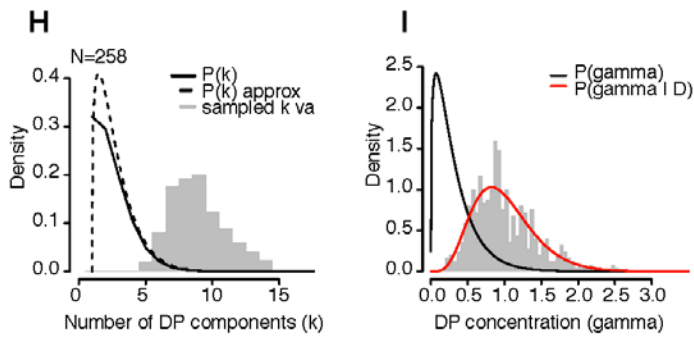
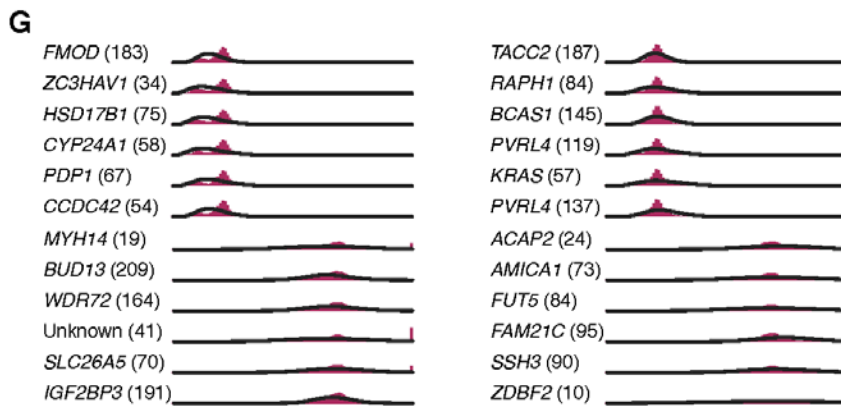
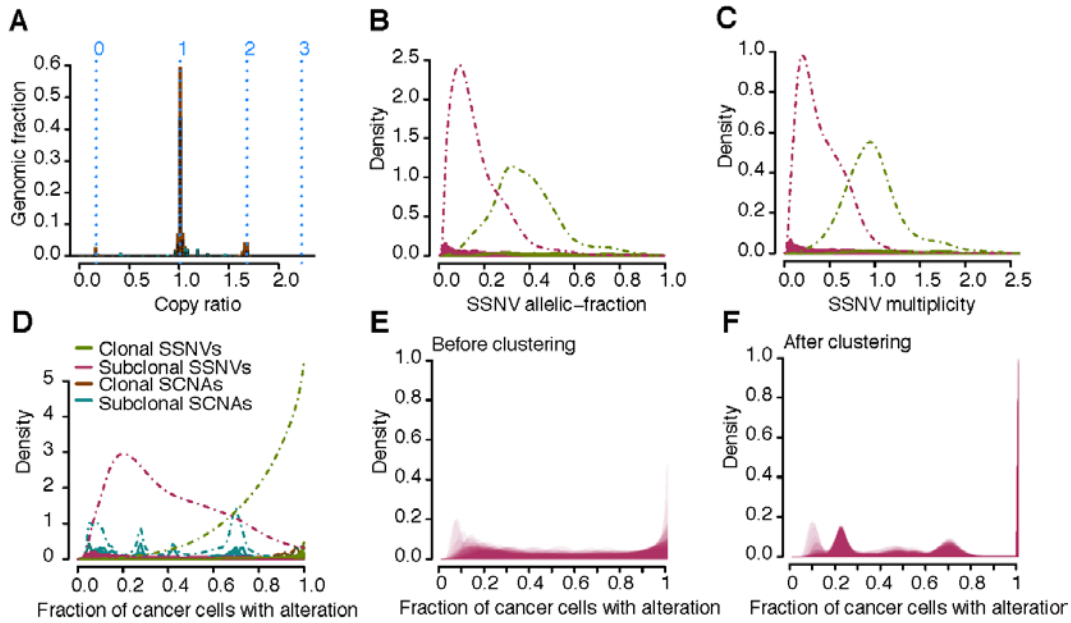


Figure S3, related to Figure 3. Example of Bayesian clustering of mutation CCF distributions. Results are shown for the clustering procedure (Experimental Procedures) applied to myeloma sample MM-0618 (purity = 0.85, ploidy = 2.14). **(A-D)**, Outputs of ABSOLUTE. **(A)** Input allelic copy-ratio data (x-axis) and fit absolute copy numbers (blue vertical dotted lines and numbers (top)). **(B)** Distribution of 258 SSNV allelic fractions from whole exome sequencing data. Dotted curves denote the summation of probability densities over individual SSNV Beta distributions. **(C)** Rescaling of SSNV allelic fraction densities to units of multiplicity (average number of mutant alleles per cancer cell). For clonal mutations, these are integer values. **(D)** Probability distribution over CCF values for SSNVs and subclonal SCNAs. Solid curves denote distributions for individual mutations; dashed curves show the sum of density over each mutation class. **(E-I)** Results of the Bayesian clustering procedure applied to SSNV CCF distributions. **(E)** CCF distributions from individual SSNVs **(D)** are shown as curves with transparent fill. **(F)** Posterior CCF distributions (following clustering) for the mutations shown in **(E)**. **(G)** Probability distributions over CCF for a selected subset of individual SSNVs before clustering (black curves), and after clustering (filled red bars). Numbers in parentheses denote the number of sequencing reads covering each SSNV. **(H)** The prior distribution on the number (k) of DP components (clusters) given by the negative binomial density with $r=10$, $\mu=2$. We note that these are the only set parameters of the clustering analysis. The approximation to the prior distribution was generated as described (Experimental Procedures). **(I)** Prior and posterior on DP concentration parameter α (Experimental Procedures). Histogram denotes sampled values. **(J)** Comparison of purity, ploidy and cancer cell fraction, calculated from sequencing data (CapSeg) versus SNP6.0 array is shown. We analyzed 110 samples with ABSOLUTE in two separate runs, using copy number input data from both SNP6.0 and whole exome capture (CapSeg). The left panel shows a scatter plot of the purities for SNP6.0 vs CapSeg. The middle panel shows the two-dimensional distribution of the cancer cell fractions

(CCF) in SNP6.0 vs CapSeg. The right panel shows a scatter plot of the correlation between the ploidies in SNP6.0 vs. CapSeg.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Sample selection and quality assessment of DNA

Bone marrow aspirates and peripheral blood samples were collected at Multiple Myeloma Research Consortium (MMRC) institutions from patients diagnosed with MM, or related diseases and then shipped to the MMRC Tissue Bank for processing as previously described (Ahmann et al., 2008). The studies were approved by the Committee on the Use of Humans as Experimental Subjects of MIT, protocol #0803002647. All patients provided written informed consent under institutional review board approval. Sample processing was slightly modified from previous reports (Salhia et al., 2010). Briefly, bone marrow aspirates were treated with ACK lysis buffer to remove red blood cells. Plasma cells were subsequently isolated using immunomagnetic sorting on a Miltenyi AutoMacs or StemCell Robocept with anti-CD138 antibodies. After immunomagnetic purification, at least 100 nucleated cells were enumerated for κ and λ staining to determine the purity of the CD138 sorts. Mean plasma cell purity (κ and λ) was found to be greater than 90% for all samples. Isolated cells were then pelleted and kept at -80°C for long-term storage. High molecular weight genomic DNA was extracted from frozen cell pellets using the Puregene kit (Qiagen). DNA concentrations were measured using PicoGreen® dsDNA Quantitation (Invitrogen, Carlsbad, CA). DNA sample quality was assessed by gel electrophoresis. The identities of all tumor and normal DNA samples were confirmed by mass spectrometric fingerprint genotyping of 24 common SNPs (Sequenom, San Diego, CA). Data on structural variants, including t(4;14) and t(11;14) were obtained as part of clinical routine testing by FISH at the institution, which contributed the sample. Structural variants were reported for 50 patients.

Whole exome, whole genome sequencing, and detection of copy number variations

Whole-exome capture libraries were constructed from 100ng of tumor and normal DNA following shearing, end repair, phosphorylation and ligation to barcoded sequencing adapters

(Fisher et al., 2011; Gnirke et al., 2009). Ligated DNA was size-selected for lengths between 200-350bp and subjected to exonic hybrid capture using SureSelect v2 Exome bait (Agilent). Samples were multiplexed and sequenced on multiple Illumina HiSeq flowcells (paired end 76bp reads) to average depth of coverage of 89x and 88x for tumor and normals, respectively. For whole-genome sequencing library construction was done with 1-3 micrograms of native DNA from primary tumor and germline samples for each patient. The DNA was sheared to a range of 101-700 bp using the Covaris E210 Instrument, and then phosphorylated and adenylated according to the Illumina protocol. Adapter ligated purification was done by preparatory gel electrophoresis, and size was selected by excision of two bands (500-520 bp and 520-540 bp, respectively) yielding two libraries per sample with average of 380 bp and 400 bp, respectively. The libraries were then sequenced with the Illumina GA-II or Illumina HiSeq sequencer with 76 or 101 bp reads, achieving an average of ~30X coverage depth. The resulting data was analyzed with the current Illumina pipeline, which generates data files (BAM files), which contain the reads and quality parameters. Copy number variations (CNV) of 153 patients of the sequencing cohort were determined by Affymetrix SNP 6.0 array.

Sequence data processing

Massively parallel sequencing data were processed using two consecutive pipelines:

(1) The sequencing data processing pipeline, called “Picard”, developed by the Sequencing Platform at the Broad Institute, starts with the reads and qualities produced by the Illumina software for all lanes and libraries generated for a single sample (either tumor or normal) and produces, at the end of the pipeline, a single BAM file (<http://samtools.sourceforge.net/SAM1.pdf>) representing the sample.

(2) The Broad Cancer Genome Analysis pipeline, also known as “Firehose” (www.broadinstitute.org/cancer/cga/Firehose), starts with the BAM files for each MM sample and matched normal sample from peripheral blood (hg18), and performs various analyses,

including quality control, local realignment, mutation calling, small insertion and deletion identification, rearrangement detection, coverage calculations and others. The details of our sequencing data processing have been described elsewhere (Chapman et al., 2011; Lohr et al., 2012) (see www.broadinstitute.org/cancer/cga).

Calculation of sequence coverage, mutation calling and significance analysis

Somatic single-nucleotide variations (SSNVs) were detected using MuTect (Cibulskis et al., 2013), and we evaluated the fraction of all bases suitable for mutation calling whereby a base is defined as covered if at least 14 and 8 reads overlapped the base in the tumor and in the germline sequencing, respectively. Single nucleotide variants found within coding areas of the genome were annotated for the chromosomal location, the type of the variant, the codon change and the change in the protein sequence using Oncotator (Ramos et al., in preparation). Insertions and deletions in coding areas (both frameshift and in-frame) were detected using the algorithm Indelocator (Sivachenko et al., in preparation; www.broadinstitute.org/cancer/cga/Indelocator).

The ranking of genes in terms of estimated conferred selective advantage was performed using the mutation statistical analysis algorithm *MutSig*. We used three separate metrics from MutSig to obtain three p values, which we combine using the Truncated Product Method for Combining P-values developed by Zaykin et al (Zaykin et al., 2002).

1. Significance analysis by estimation of gene-specific background mutation rate:

Compared to the version of background mutation frequency calculation used previously (Chapman et al., 2011; Lohr et al., 2012), we have made substantial improvements (Lawrence et al., 2013) in order to calculate a gene specific background mutation rate, using gene characteristics that correlate well with mutation frequency, such as: local relative replication time (Chen et al., 2010), and open vs. closed chromatin status (Lieberman-Aiden et al., 2009).

2. Prioritizing genes that are enriched in mutational hotspots that are significantly conserved:

We performed an additional significance analysis in order to prioritize genes that harbor mutations that are significantly clustered in conserved hotspots, using a metric that was developed and described in Lohr et al (Lohr et al., 2012). In short, for each gene this method permutes the observed number of mutations along the cDNA of the gene, preserving the trinucleotide context in which they occurred. For each permutation as well as the observed mutations, it calculates two metrics:

- A) **The level of clustering of the gene:** To calculate this metric, we take all pairwise distances between the mutations, we calculate their cumulative distribution, and then calculate a Kolmogorov-Smirnov distance between this distribution and a random distribution of pairwise distances.
- B) **Conservation of the gene:** Conservation is obtained from the conservation track of 46 vertebrates provided by the UCSC genome browser. The conservation metric we use is the mean of the conservation values of each mutated site.

We then calculate the joint distribution of the clustering and conservation metrics, and we use this to calculate a score for the observed mutation, as well as for each permutation. We use the permutation scores to form a null distribution, and we project the score of the observed mutations onto this null distribution in order to calculate a p value for the given gene.

3. Prioritizing genes by enrichment in homozygous events:

One of the results of the ABSOLUTE algorithm is a call whether a point mutation is homozygous, based on the multiplicity of the mutation and the local copy number. For each gene, we calculate a fraction f of homozygous mutations for each patient, and we use this fraction to calculate a binomial probability using the Binomial probability density function given by $Binom(h, n, f)$, where h is the number of homozygous mutations, and n is the total number of mutations. We then convolve all the probabilities in order to calculate a final probability, which

represents the probability by chance that the gene harbors k homozygous mutations out of n total mutations, given the respective fractions of homozygous mutations for each patient in which the gene was mutated (in order to control for the fact that some patients have higher homozygous mutation frequency than others, due to large-scale deletions in the genome).

The combination of the three p values from the three metrics described yields a combined p value. Multiple hypothesis correction was performed on these combined p values using the method of Benjamini and Hochberg (Benjamini and Hochberg, 1995) to calculate False Discovery Rate combined q -values.

Classification of Hyperdiploid/Non-Hyperdiploid samples by WES/WGS

In order to designate all samples as either hyperdiploid or non-hyperdiploid, including the ones for which no high density SNP array or ABSOLUTE data from WES were available, we used the WES and WGS data to determine the read density across uniform segments of the exome/genome. We calculated a normalized read density ratio, independently for the long and short arm of each chromosome of every tumor sample. Samples were called hyperdiploid, if trisomy (defined as read density ratio greater than 1.5) was observed in 7 or more chromosomal arms, as shown in Figure S2. As viable cells for FACS analysis or FISH analysis were not available, we validated this approach by comparing the results to the previously published ploidy assessment by ABSOLUTE (Carter et al., 2012). Of note, the ABSOLUTE results are based on orthogonal data, ie. SNP arrays, in the majority of cases (139 samples), and the rest based on exome sequencing (44 samples). The two ploidy estimates significantly correlated ($r_s = 0.71$, $p < 0.000001$, two-tailed Spearman Rank Order test).

Estimation of SSNV cancer cell fraction

For 153 of 203 matched cancer-normal DNA samples, copy number profiles were obtained using the Genome-wide Human SNP Array 6.0 (Affymetrix), according to the manufacturer's

protocol (Genetic Analysis Platform, Broad Institute, Cambridge MA), with allele-specific analysis (Carter et al., 2011). Out of these 153 samples, we were able to obtain good quality ABSOLUTE results for 139. For myeloma samples with no available ABSOLUTE data from SNP arrays, SCNAs were estimated directly from the WES data, by our CAPSEG algorithm based on the ratio of cancer sample read-depth to the average read-depth observed in normal samples for that region. Data obtained from SNP Array and CAPSEG correlated well (**Figure S3J**). We applied the algorithm ABSOLUTE (Carter et al., 2012), to estimate sample purity, ploidy, and absolute somatic copy numbers. These were used to infer the cancer cell fraction (CCF) of point mutations from the WES data (**Figure S3A-D**),

Following the framework previously described (Carter et al., 2012), we computed the posterior probability distribution over CCF c as follows. Consider a somatic mutation observed in a of N sequencing reads on a locus of absolute somatic copy-number q in a sample of purity α . The expected allele-fraction f of a mutation present in one copy in a fraction c of cancer cells is calculated by $f(c) = \alpha c / (2(1 - \alpha) + \alpha q)$, with $c \in [0.01, 1]$. Then $P(c) \propto \text{Binom}(a|N, f(c))$, assuming a uniform prior on c . The distribution over CCF was then obtained by calculating these values over a regular grid of 100 c values and normalizing (**Figure S3D-E**). Further details of this procedure as described recently (Landau et al., 2013).

Clustering analysis of SSNVs

We employed a previously described Bayesian clustering procedure (Escobar and West, 1995). The details of this procedure have been recently described (Landau et al., 2013). Briefly, this approach exploits the assumption that the observed subclonal SSNV CCF values were sampled from a smaller number of subclonal cell populations (subclones). All remaining uncertainty (including the exact number of clusters) was integrated out using a mixture of Dirichlet processes, which was fit using a Gibbs sampling approach, building on a previously described framework (Escobar and West, 1995).

Determining number of predicted subclones

To create the bargraph and error bars in Figure 3a, the results of the Markov chain Monte Carlo (MCMC) simulation were summarized as follows: For each patient we obtained multinomial parameters by calculating the fraction of times each number of subclones (0-10) was represented by the Dirichlet process across the MCMC (after convergence). We simulated 1,500 draws from these distributions and calculated the total number of subclones observed across patients in each draw. Only subclones that have more than one mutation, and have expected CCF > 0.1, with sample purity > 0.7, were included in the analysis. Each bar shows the mean and the standard deviation of the fractions for each total calculated across the draws.

Western blot analysis of MAPK pathway following PLX4720 treatment

Lysate protein concentrations were obtained using the BIO-RAD DC Protein Assay kit, and concentrations were subsequently adjusted to 1 $\mu\text{g}/\mu\text{L}$ final concentrations. Twelve micrograms of protein from each cell lysate was run per well on NuPAGE 4-12% Bis-Tris Midi Gels (Life Technologies WG1403BX10). The gel was blotted onto nitrocellulose membrane paper (Invitrogen LC2001), using the iBlot gel transfer device (Life Technologies IB1001). The membrane was subsequently blocked (LiCor Blocking Buffer 927-40000) for one hour. After blocking, the membrane was incubated overnight at 4°C with primary antibodies against either phosphorylated MEK (Cell-Signaling 9121), total MEK 1 & 2 (Cell Signaling Technology 9122), phosphorylated ERK (Cell Signaling Technology 9101), or total ERK 1 & 2 (Santa Cruz Biotechnology sc-135900) in LiCor Blocking Buffer + 0.1% Tween-20, in addition to GAPDH or beta-actin housekeeping antibodies as controls. Following the overnight primary antibody incubation, the membrane was washed 3 times with TBS-Tween solution. The membrane was then incubated, shaking at room temperature for 1 hour, with the LiCor Odyssey green (Licor Biosciences 926-32211) IRDye or red (Licor Biosciences 926-68020) IRDye secondary

antibody, 0.1% Tween-20, and 0.02% SDS. The membrane was then washed 3 times with TBS-Tween solution again prior to imaging on the LiCor Odyssey machine at a resolution of 85 μm .

SUPPLEMENTAL REFERENCES

Benjamini, Y. H., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* 57, 289-300.

Carter, S. L., Meyerson, M., and Getz, G. (2011). Accurate estimation of homologue-specific DNA concentration-ratios in cancer samples allows long-range haplotyping. *Nature Precedings*, hdl:10101/npre.2011.6494.1, <http://hdl.handle.net/10101/npre.2011.6494.1>.

Chen, C. L., Rappailles, A., Duquenne, L., Huvet, M., Guilbaud, G., Farinelli, L., Audit, B., d'Aubenton-Carafa, Y., Arneodo, A., Hyrien, O., and Thermes, C. (2010). Impact of replication timing on non-CpG and CpG substitution rates in mammalian genomes. *Genome Res* 20, 447-457.

Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E. S., and Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 31, 213-219.

Escobar, M. D., and West, M. (1995). Bayesian Density Estimation and Inference Using Mixtures. *Journal of the American Statistical Association* 90, 577-588.

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., *et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289-293.

Zaykin, D. V., Zhivotovsky, L. A., Westfall, P. H., and Weir, B. S. (2002). Truncated product method for combining P-values. *Genet Epidemiol* 22, 170-185.