Supplementary Figure S1. Functional modules among 384 genes selected for the NMF algorithm. Significantly overexpressed genes in each of the three clusters were selected on the basis of adjusted P value ≤ 0.05 for each of the pair-wise comparisons. Functional gene sets were identified using Ingenuity Pathway Analysis (IPA) Software or curated based on published literature.
A.

Chemo-naïve (n=176)

Platinum-refractory (n=41)
Supplementary Figure S2. Co-occurring somatic mutations in KRAS-mutant LUAC subgroups. (A) Non-random patterns of co-occurring somatic mutations in the three KRAS-mutant LUAC clusters. Genes with somatic mutations present in ≥10/68 tumors (≥14.7%) in the TCGA dataset were included in this analysis. At a FDR of 0.05, 11 genes were identified as differentially represented in the three clusters using Fisher’s exact test. Columns in this plot correspond to individual tumor samples, ordered on the basis of expression cluster membership. Cases with identified somatic mutations are depicted in black. (B) Evolution of KRAS co-mutations with prior cytotoxic chemotherapy. Left panel: In a cohort of 176 chemotherapy-naïve KRAS-mutant tumors compiled by combining data from Imielinski et al (n=49), Chitale et al (n=48), Ding et al (n=60) and PROSPECT (n=19), only 4% of KRAS-mutant tumors harbor concurrent mutations in TP53 and STK11. Resistance to cytotoxic, platinum-based, chemotherapy is associated with increased prevalence of somatic mutations in TP53 in the BATTLE-2 cohort, giving rise to more frequent triple mutant (KPL$^m$) LUACs. Note that in this pie-chart groups are defined on the basis of somatic mutations rather than mRNA expression and are therefore indicated with a “m” superscript.
Supplementary Figure S3. Distribution of bi-allelic deletion/ amplification events in the three KRAS-mutant LUAC subsets. Only genes with bi-allelic deletion or amplification (-2/+2 CN change) detected in at least three separate tumors were included in the analysis. The top 100 genes based on Fisher’s exact test P value are presented, ordered according to cytoband location. At a FDR of 0.3 this analysis identified bi-allelic deletions in CDKN2A, CDKN2B and C9orf53, all located in 9p21.3, as differentially represented in the three clusters.
Supplementary Figure S4. Alterations in genes involved in G1/S transition in the three KRAS-mutant LUAC subgroups.
Supplementary Figure S5. Clonal fractions for somatically mutated genes in the 68 KRAS-mutant LUACs (TCGA). Overall, 67.64% of identified somatic mutations were clonal in all tumors where the mutation was detected. Mutations in KRAS, STK11, TP53, KEAP1 and ATM were clonal in 98.53%, 100%, 95.24%, 100% and 81.82% of mutant tumors respectively.
Supplementary Figure S6. Co-mutation plot for the combined PROSPECT/CHITALE validation dataset of chemotherapy-naïve KRAS-mutant LUACs. Expression data were available for 47/48 tumors originally reported by Chitale et al. Complete exome sequence data were available for 19/41 KRAS-mutant tumors from PROSPECT – tumors lacking somatic mutation data for STK11/LKB1 or TP53 are color-coded ochre and were excluded from the co-mutation analysis.
Supplementary Figure S7. Co-occurring genetic events in PI3K-AKT-MTOR pathway genes in the three KRAS-mutant LUAC subsets.
Supplementary Figure S8. LKB1-loss-dependent up-regulation of adaptive branches of the unfolded protein response in KRAS-mutant LUAC cell lines. (A) Quantitative RT-PCR for ATF4 and spliced XBP1 (sXBP1), key transcription factors downstream of the PERK and IRE1 apical UPR sensors in the A549 and H460 isogenic systems. Restoration of LKB1 results in significantly reduced baseline levels of ATF4 transcripts in H460 cells, but has no effect on sXBP1 transcript levels. Conversely, significant down-regulation of sXBP1 upon LKB1 add-back is observed only in the A549 cell line. The graph shows Ct values normalized to the Ct values of ACTB and expressed as fold change compared to the mean vector control value. Error bars represent SD from the mean of duplicate wells. Un-paired t test is used for statistical comparison. Asterisk denotes statistical significance at the P≤0.05 level. (B) Western blot analysis of key UPR proteins in the A549 and H460 isogenic systems. Reduced phosho-S6 and phospho-p70S6K levels upon LKB1 add-back support wild-type functionality of the transgene. In-keeping with results from qRT-PCR, re-introduction of LKB1 impacts on the IRE1-sXBP1 branch of the UPR in A549 and on the PERK-phospho-eIF2a-ATF4 branch in H460 cells.
Supplementary Figure S9. Enumeration of distinct immune cell populations in the KRAS LUAC sub-groups from PROSPECT. The percentage of cells that stain positive for the specified marker is determined by digital image analysis in five random square regions (1mm$^2$) of the tumor core. Statistical comparison is based on Wilcoxon rank-sum test, with $p \leq 0.05$ considered significant.
Supplementary Figure S10. Analysis of overall survival (OS) in KRAS-mutant LUAC subsets. A merged dataset of 164 KRAS-mutant LUACs with available mRNA expression data and OS data are included in this analysis.
Supplementary Figure S11. Integrated analyses for a higher pre-determined number of KRAS-mutant LUAC subsets. Top panel: k=4, bottom panel k=5. Note that tumors with co-mutations in STK11/LKB1 and TP53 still cluster separately. The majority of tumors with co-mutations in either TP53 or ATM express high levels of TTF1, most evident from the five-subgroup partitioning. TCGA expression clusters (proximal proliferative, proximal inflammatory, terminal respiratory unit) and integrated clusters (iCluster) are also indicated.