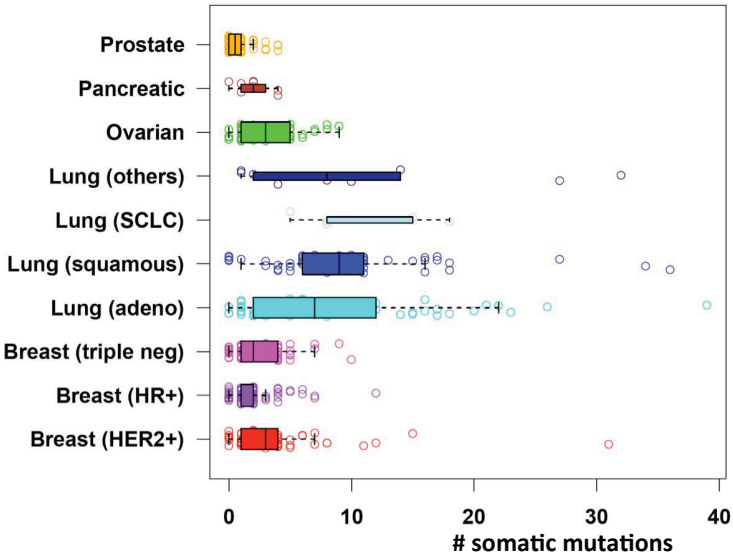
SUPPLEMENTARY INFORMATION

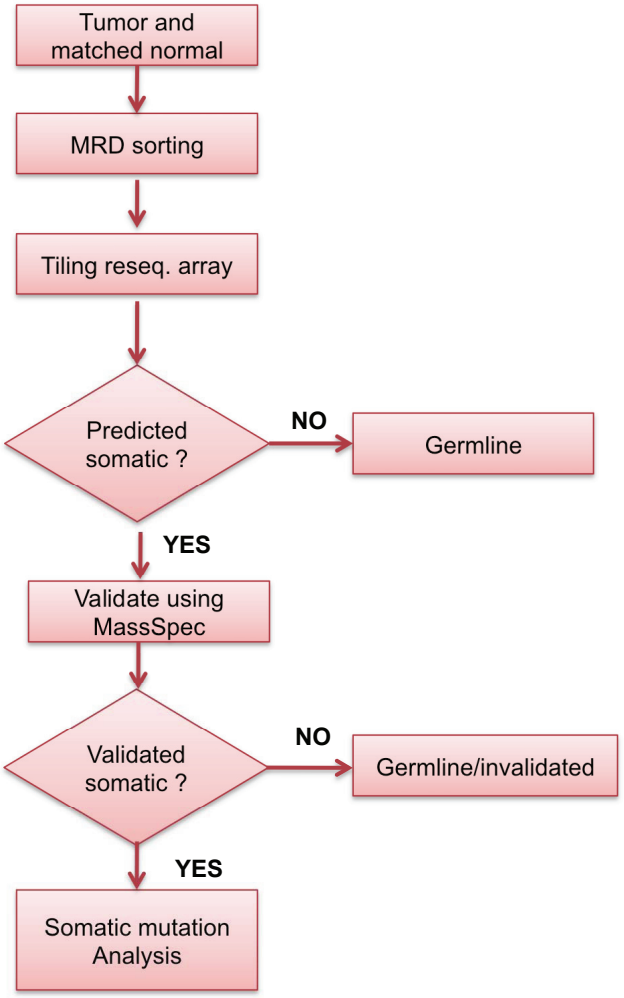
a



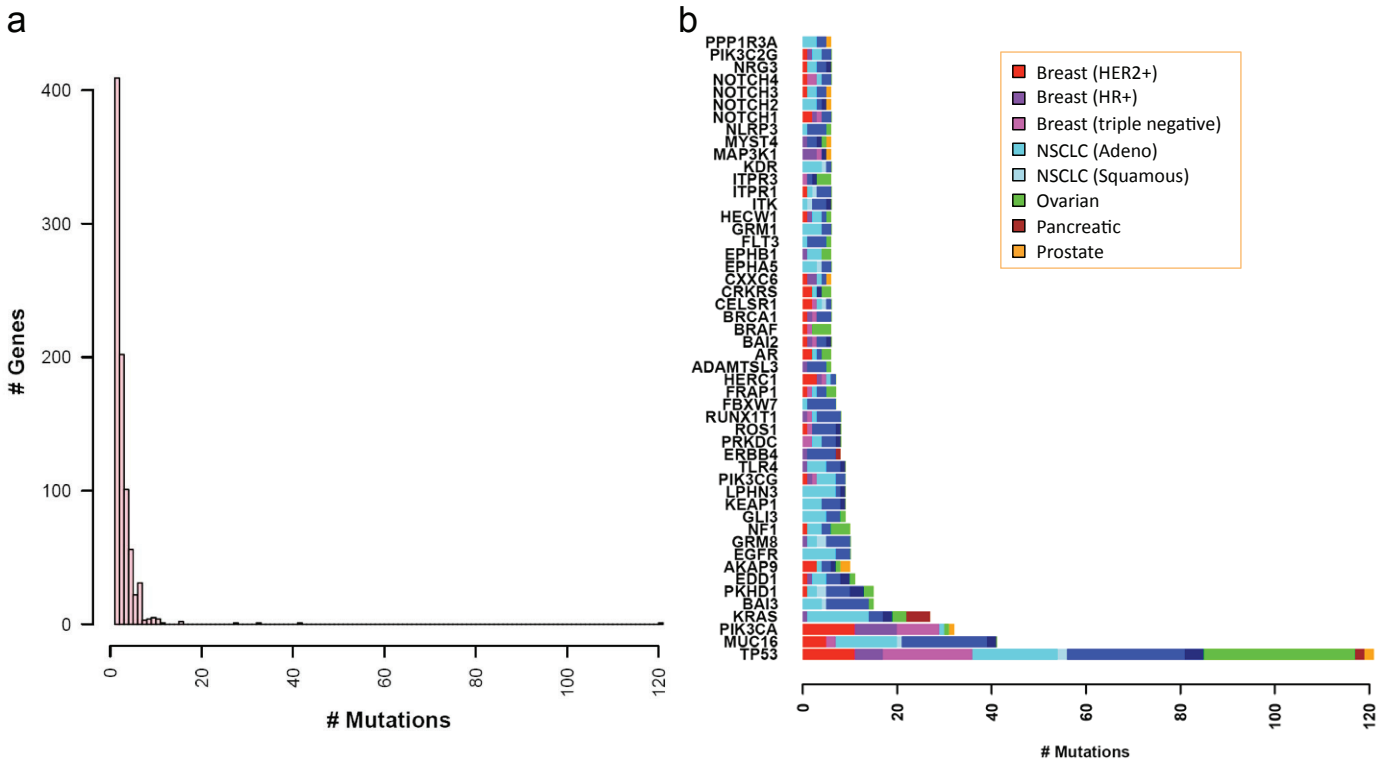
b

Supplementary Fig. 1 | Somatic mutation prevalence across major cancer types.
a, Distribution of somatic mutation prevalence across multiple tumor types. **b**, Mutation rates (# of mutations/Mb) for all changes, protein-altering and synonymous changes across cancer types. Protein-altering mutations include non-synonymous, non-sense and splice-site changes.

Cancer type	# samples	# mutations	# samples mutated	mutation rate (protein-altering)	mutation rate (synonymous)	mutation rate (all changes)
Breast Cancer	183	524	158	1.14	0.32	1.46
Breast Cancer (HER2+)	59	229	54	1.55	0.41	1.96
Breast Cancer (HR+)	65	141	54	0.87	0.26	1.12
Breast Cancer (Triple Neg)	59	154	50	1.04	0.3	1.34
Lung Cancer	134	1276	129	3.8	0.96	4.76
Lung Cancer (Adenocarcinoma)	57	503	55	3.52	0.89	4.41
Lung Cancer (Carcinoma Small Cell)	5	54	5	4.3	0.56	4.86
Lung Cancer (Carcinoma Squamous	63	620	60	3.92	1.02	4.94
Lung Cancer (Others)	9	99	9	4.55	1.24	5.79
Ovarian Cancer	58	174	54	1.19	0.36	1.55
Pancreatic Cancer	8	16	7	0.8	0.25	1.04
Prostate Cancer	58	48	29	0.33	0.08	0.4
All Cancers	441	2038	377	1.84	0.49	2.33

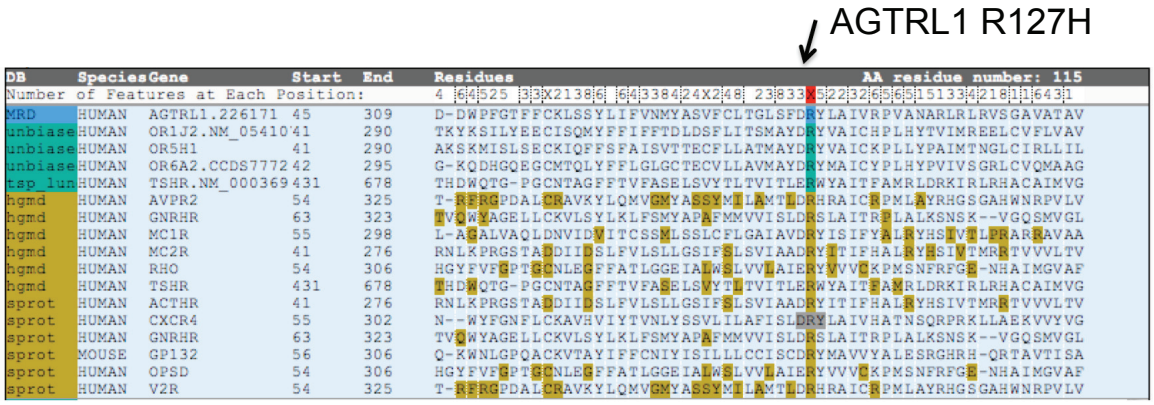


Supplementary Figure 2 | Work flow showing the process used to identify and validate somatic mutations. Tumor DNA and matched normal DNA were subjected to MRD sorting and the sorted amplicons were hybridized to a tiling array and the predicted changes were mapped back to the amplicon reference sequence to identify putative somatic changes as described (Methods). dbSNP database was used to identify germline changes prior to somatic mutation prediction. The putative somatic changes were then validated using Sequenom mass spectrometry. Any changes found in both tumor and normal were classified as new germline changes and the changes found only in the tumor, but not in the normal were classified as somatic.

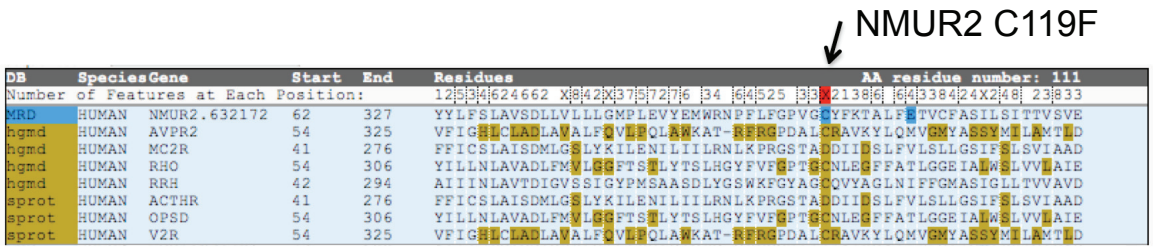


Supplementary Figure 3 | Somatic mutation distribution in genes. (a) Distribution of number of genes grouped by mutation count. The tail end of the graph has genes with the highest number of mutations. (b) The distribution of mutations by tumor type in the top 50 mutated genes.

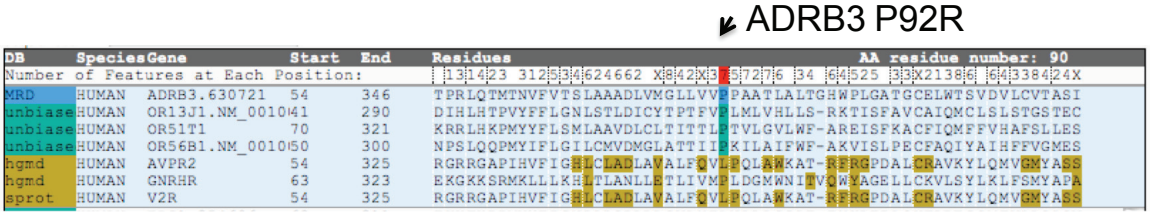
a



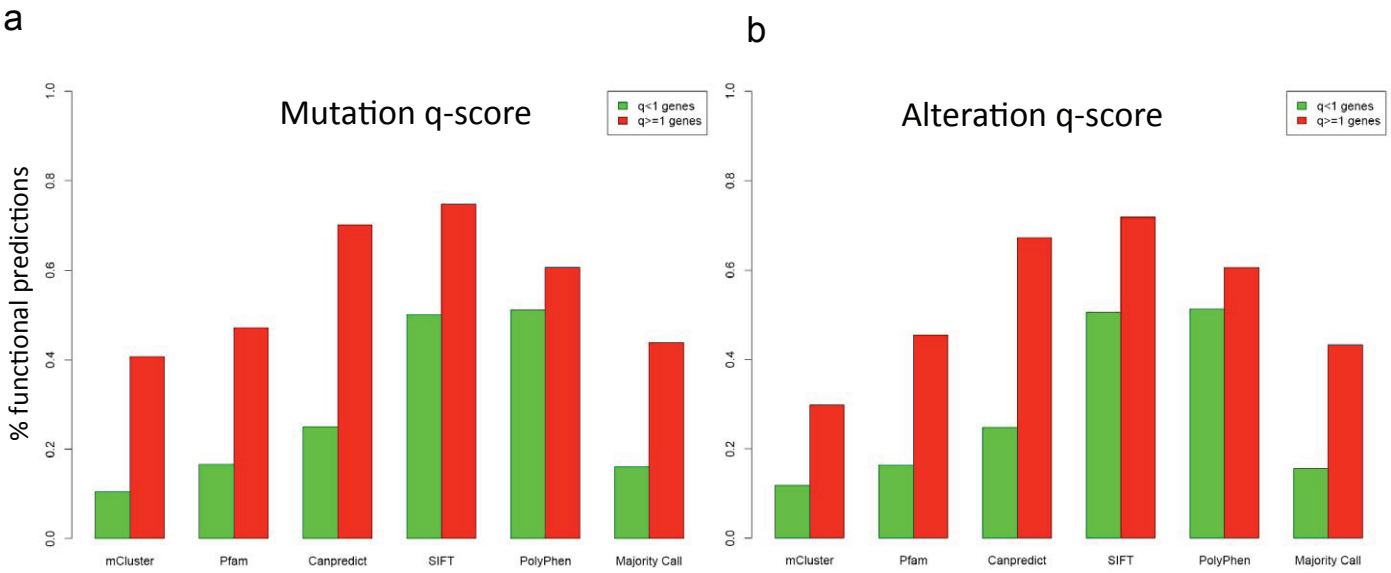
b



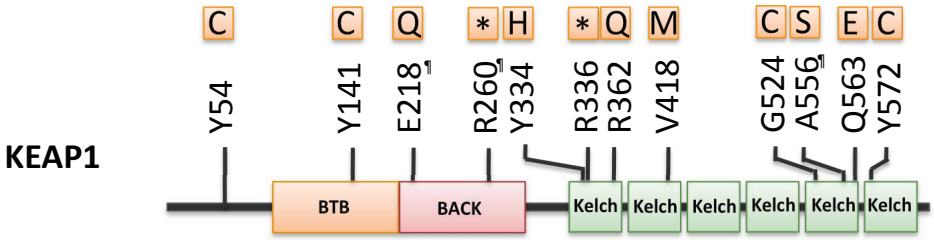
c



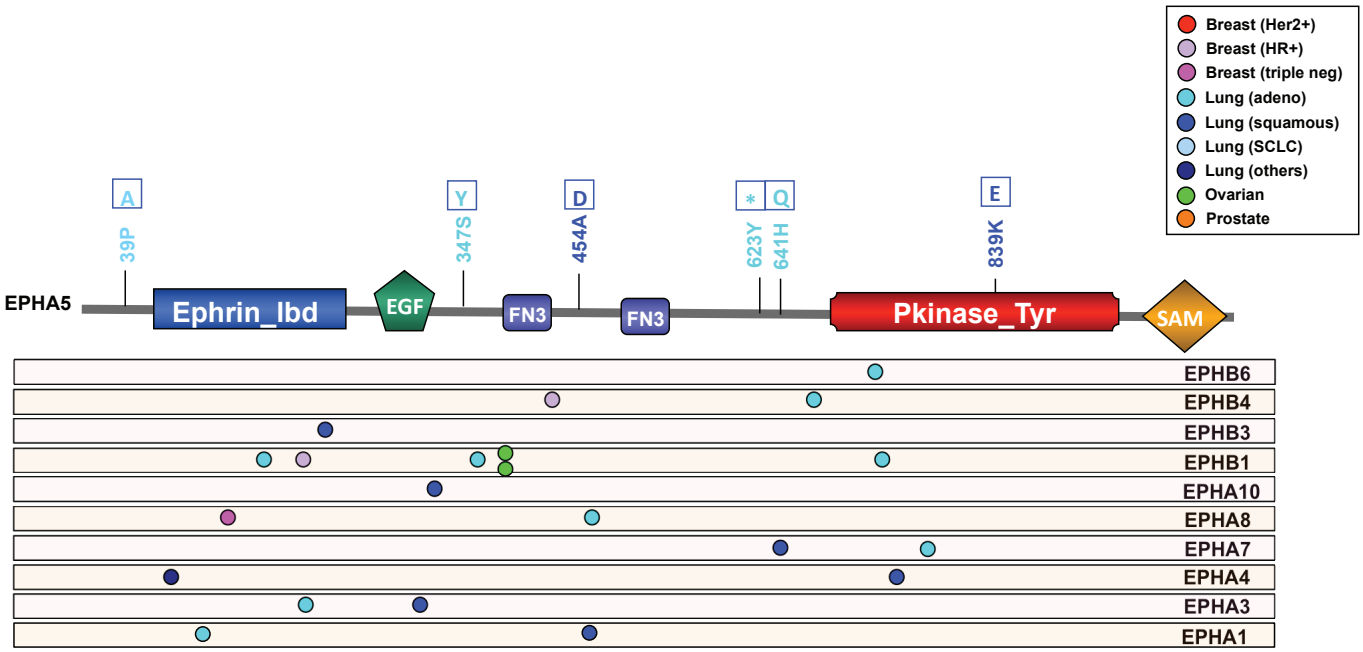
Supplementary Figure 4 | mCluster[40] analysis of AGTRL1 (a), NMUR2 (b) and ADRB3 (c) depicting the occurrence of mutations at the analogous residue in other GPCRs in human disease that correspond to the mutated residue in these proteins.



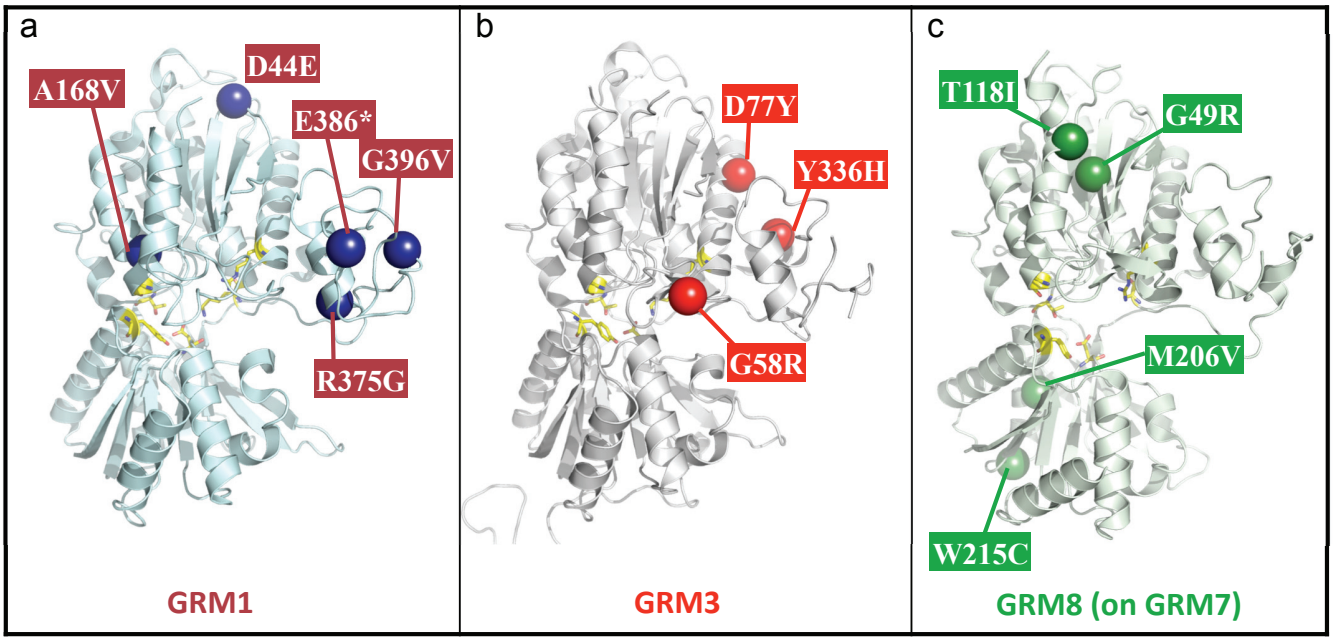
Supplementary Figure 5 | Mutations predicted to have a functional effect are enriched in (a) significantly mutated genes and (b) genes with significant combined prevalence of mutation and copy number alteration. All scored mutations were segregated based on mutation and alteration q-scores into two groups, significant genes (q ≥1) vs. non-significant genes (q<1). The proportions of mutation events in each group receiving a functional prediction based on each of the six criteria are shown. Majority call refers to mutations predicted to have functional effect by at least three out of five methods



Supplementary Figure 6 | Somatic mutations in KEAP1.
Broad-Complex, Tramtrack and Bric a brac (BTB), BTB And C-terminal Kelch (BACK), Kelch domain. ¶Mutations identified in prevalence screen. *represents nonsense mutations.

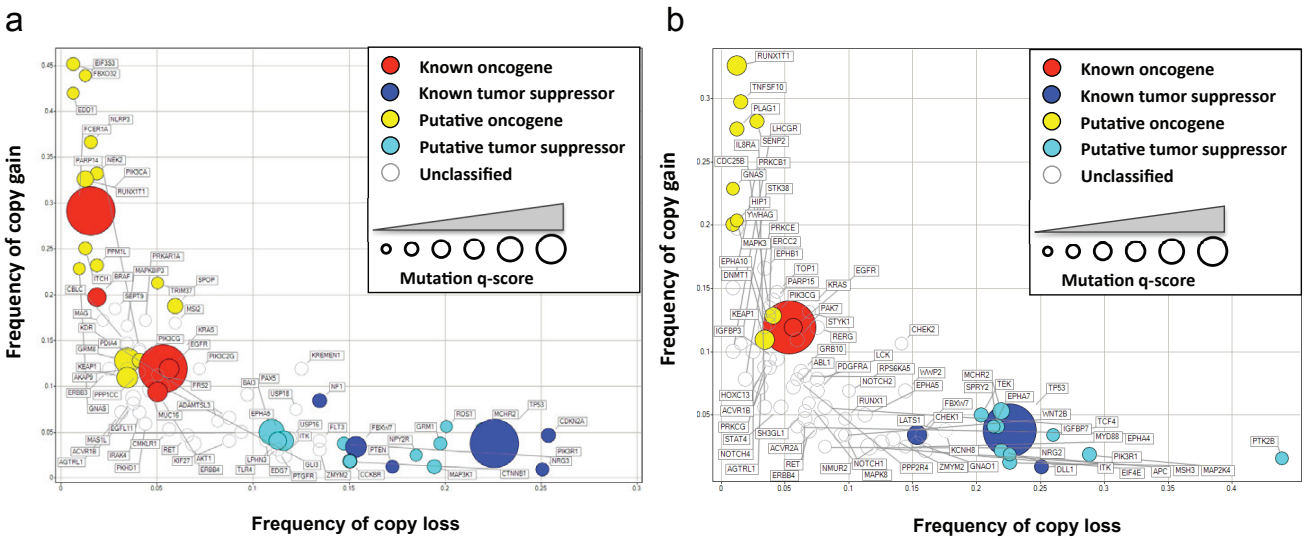


Supplementary Figure 7 | Somatic mutations in Epherin family of genes. Somatic mutation in EPHA5 is depicted on the top. Mutation in other family member mapped on to EPHA5 is shown below. Fibronectin type 3 domain (FN3), low-density lipoprotein-receptor YWTD domain (LY), Transmembrane domain (TM), Protein tyrosine kinase domain (Pkinase_Tyr), Ephrin receptor ligand binding domain(Ephrin_lbd) and Sterile alpha motif (SAM) are shown. * represents nonsense mutation.



Supplementary Figure 8 | (a-c) Mutations mapped on to the extracellular domain of (ECD) GRMs. (a) GRM1 mutations mapped on to GRM1 ECD structure (PDB 1EWK). (b) GRM3 mutations mapped on to GRM3 ECD structure (PDB 2E4U). (c) GRM8 mutations mapped on to GRM7 structure (PDB 2E4Z).

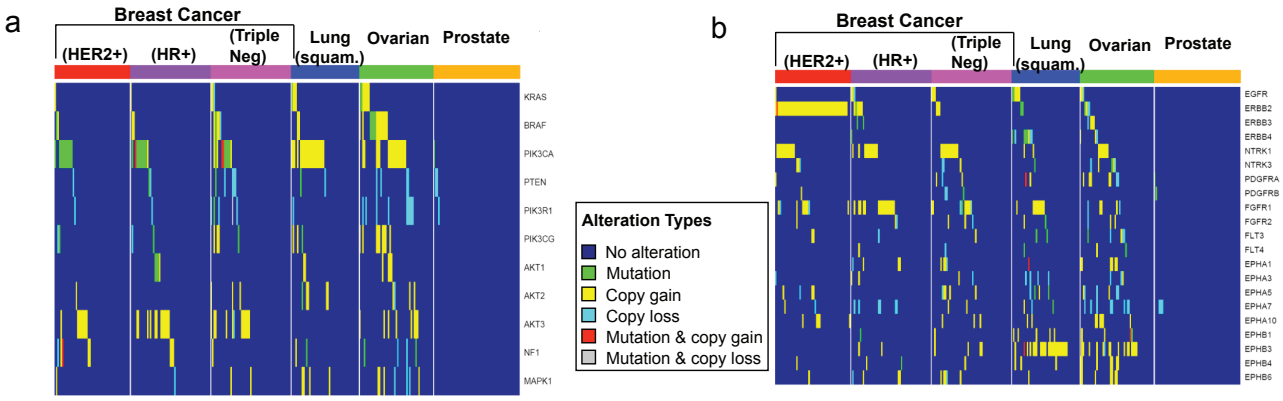
Supplementary Figure 9 | Integrated map of molecular alterations in multiple cancers. Heatmap of mutations and copy number alterations across tumors and genes studied reveals the extent of alternation across tumor types and sub-types. A few known genes and novel ones with alterations are indicated to the left.



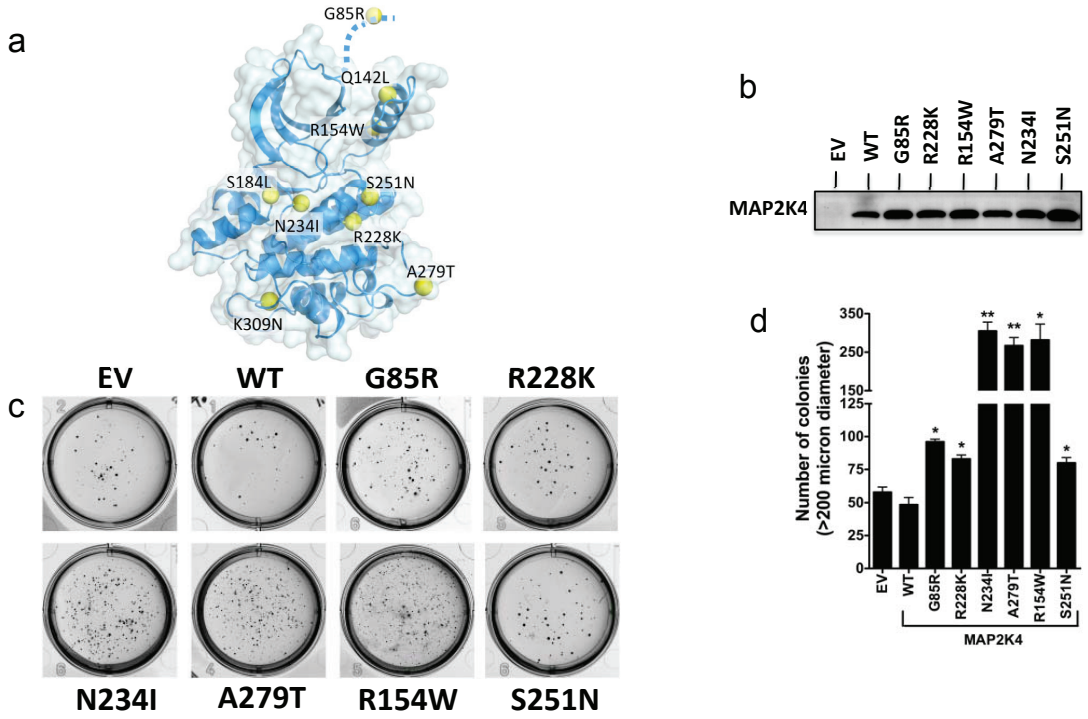
All 112 statistically significant genes with labels

All 77 genes containing a mutation receiving at least 4 positive functional predictions

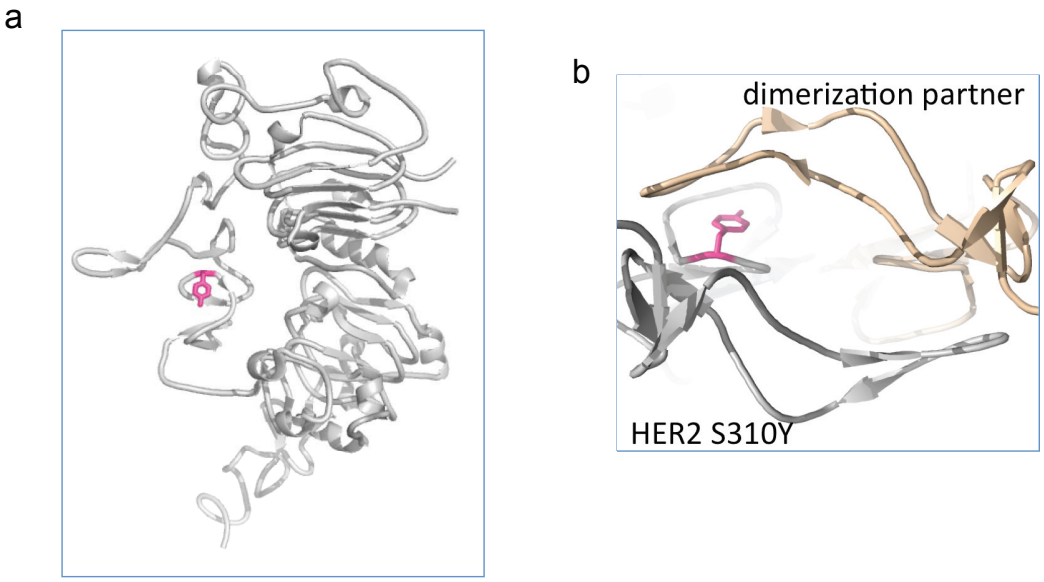
Supplementary Figure 10 | Relative frequency of copy number gain vs. loss for genes with frequency of copy change ≥ 20% and (a) genes with significant prevalence of mutation and copy number alteration, and (b) genes harboring mutations predicted to have a functional consequence by at least three out of five methods. "Frequency" refers to the percentage of all samples analyzed by CGH array that exhibit copy number alteration.



Supplementary Figure 11 |Integrated analysis of RTK/RAS and GPCR signaling pathway alterations in cancers. a-b, Heatmap of integrated molecular alteration in RAS pathway genes (a) and RTKs (b) across cancers.



Supplementary Figure 12 |MAP2K4 mutations promote anchorage independent growth. (a) Model of MAP2K4 structure, generated using SWISS-model sever with the crystal structure of MEK6 (3FME) as the template, depicting somatic mutation sites. (b) Western blot showing the expression of Flag tagged MAP2K4 mutants in HMECs (c-d) A representative image showing anchorage-independent growth (c) and number of colonies (d) formed by HMEC cells stably expressing *MAP2K4* mutants. Data shown (d) are mean ± SEM. (*p<0.05 and **p<0.01). EV= empty vector.



Supplementary Figure 13 |Recurrent HER2 mutation. (a) Mutation 310S>Y (deep pink) located in domain II of HER2 mapped on to the structure of HER2 (PDB 1S78). Domains I-IV of the HER2 extracellular domain is annotated on the structure. (b) HER2 homodimer modeled based on EGFR/TGF-alpha dimeric structure (PDB 1MOX) depicting the location of serine to tyrosine mutation at residue 310 in domain II of HER2 receptor dimerization interface.