doi: 10.1038/nature07423

Supplementary Information

for

Somatic mutations affect key pathways in lung adenocarcinoma

Li Ding, Gad Getz, David A. Wheeler, Elaine R. Mardis, Michael D. McLellan, Kristian Cibulskis, Carrie Sougnez, Heidi Greulich, Donna M. Muzny, Margaret B. Morgan, Lucinda Fulton, Robert S. Fulton, Qunyuan Zhang, Michael C. Wendl, Michael Lawrence, David E. Larson, Ken Chen, David J. Dooling, Aniko Sabo, Alicia C. Hawes, Hua Shen, Shalini N. Jhangiani, Lora R. Lewis, Otis Hall, Yiming Zhu, Tittu Mathew, Yanru Ren, Jiqiang Yao, Steven E. Scherer, Kerstin Clerc, Ginger A. Metcalf, Brian Ng, Aleksandar Milosavljevic, Manuel L. Gonzalez-Garay, John R. Osborne, Rick Meyer, Xiaoqi Shi, Yuzhu Tang, Daniel C. Koboldt, Ling Lin, Rachel Abbott, Tracie L. Miner, Craig Pohl, Ginger Fewell, Carrie Haipek, Heather Schmidt, Brian H. Dunford-Shore, Aldi Kraja, Seth D. Crosby, Christopher S. Sawyer, Tammi Vickery, Sacha Sander, Jody Robinson, Wendy Winckler, Jennifer Baldwin, Lucian R Chirieac, Amit Dutt, Tim Fennell, Megan Hanna, Bruce E. Johnson, Robert C. Onofrio, Roman K. Thomas, Giovanni Tonon, Barbara A. Weir, Xiaojun Zhao, Liuda Ziaugra, Michael C. Zody, Thomas Giordano, Mark B. Orringer, Jack A. Roth, Margaret R. Spitz, Ignacio I. Wistuba, Bradley Ozenberger, Peter Good, Andrew C. Chang, David G. Beer, Mark A. Watson, Marc Ladanyi, Stephen Broderick, Akihiko Yoshizawa, William D. Travis, William Pao, Michael A. Province, George M. Weinstock, Harold E. Varmus, Stacey B. Gabriel, Eric S. Lander, Richard A. Gibbs, Matthew Meyerson, Richard K. Wilson

Excel spreadsheets can be downloaded from http://genome.wustl.edu/supplemental/tsp_nature_2008

Table of Content

A. Supplementary Methods

Section I. Gene list and sample

Gene list origins and tumor sample selection

The TSP gene list consists of 623 genes, many of which have previously been implicated as proto-oncogenes or tumor suppressor genes related to solid tumor development. Genes from selected signaling pathways and additional enzyme families also were selected for this study. In addition, some genes from regions on chromosomes 3 and 22, implicated in lung cancer by copy number analysis, were included. Protein family/domain and gene ontology annotation for the 623 TSP genes are shown in Supplementary Table 1.

To identify statistically significant recurrent mutations, we sequenced these genes in 188 lung adenocarcinomas that were selected from an initial set of 384 snap-frozen lung adenocarcinoma biopsy samples, each containing a minimum of 70% tumor cells according to pathology evaluation. These samples were anonymized, and contributed along with matched normal samples by the Dana-Farber Cancer Institute, MD Anderson Cancer Center, Memorial Sloan-Kettering Cancer Center, University of Michigan, and Washington University in St. Louis. All 384 samples were run on 250K StyI Affymetrix SNP Arrays. Analysis of Affymetrix 250K SNP array data enabled the selection of 188 tumors for re-sequencing, whereby tumors displaying LOH and/or hemizygous deletion for a minimum of 10 Mb on the array were preferentially selected. Here, copy number data were visually inspected to categorize samples based on the following criteria:

- *Pass*: presence of an uninterrupted stretch of loss of heterozygosity (LOH) or hemizygous deletion with a copy number less than 1.4 (indicating >60% tumor purity) for a minimum of 10 Mb.
- *Borderline*: presence of hemizygous deletion with a copy number between 1.4 and 1.6 (indicating 40-60% tumor purity) for a minimum of 10 Mb.
- *Fail*: absence of these qualifying events.

In addition, samples exhibiting any type of technical or protocol failure, such as mismatch between the tumor and normal sample or low quality data, were excluded. These categories, used as a proxy for stromal contamination, were later computationally formalized using a metric defined as the standard deviation of the medians of log2 copy-number ratios in each chromosome arm. The first plate of 94 samples was comprised of samples from the "Pass" and "Borderline" categories.

To test the proposition that copy number data from SNP arrays could be used as a genomic indicator for stromal contamination, as described above, we sequenced the *EGFR*, *KRAS*, and *TP53* genes by both traditional Sanger sequencing and 454 single molecule sequencing. We processed 96 samples chosen randomly from the larger set of 384. The 454 allele frequency generally tracked with both the SNP array category and the ability to detect mutations by Sanger sequencing for each sample. Yet, there were 8 samples that exhibited 454 allele frequencies above 30% (indicating >60% tumor purity) with mutations still detectable by Sanger sequencing, despite the absence of qualifying lesions on SNP arrays (data not shown). In light of these data, the second plate of 94

samples was comprised of the remaining "Pass" samples that were still available, as well as samples chosen at random from the "Borderline" and "Fail" categories (Supplementary Table 17).

Interpro analysis of 623 genes selected for TSP

Peptide sequences for each transcript were created and stored in a fasta file, which was subsequently processed with Interproscan, specifically the HMMPfam and superfamily analyses. The results were loaded into a relational database, enabling queries for protein domain, protein family, gene ontology (GO), etc.

Section II. Sequencing strategy and mutation analysis

Sequencing strategy and mutation detection

We screened 188 primary lung adenocarcinoma tumors for somatic mutations, focusing on the coding exons and splice donor and acceptor sites of 623 human genes. In total, we re-sequenced 247Mb from the full set of tumor genomes utilizing PCR amplification, followed by fluorescent terminator-based capillary sequencing. Coverage analysis demonstrates Phred Q30 coverage for greater than 90% of the target bases. Raw DNA sequence trace data were analyzed using a combination of mutation detection algorithms, including PolyScan[1], PolyPhred[2], PolyDhan (unpublished), PolyTedh (unpublished), and SNPdetector[3] to discover single nucleotide variants and small insertions and deletions (indels). Putative mutations were verified by a second round of PCR-directed sequencing, or were validated by orthogonal methods, including Sequenom genotyping, custom Illumina Golden Gate assays, Applied Biosystems Taqman and/or PCR-based pyrosequencing with the Roche 454 FLX or Biotage platforms (Supplementary Fig. 1). Orthogonal validation included the assay of the matched normal genomic DNA sample (derived from adjacent normal lung tissue or PBLs), to confirm the somatic origin of each variant.

In order to compare results from PCR-directed sequencing and associated variant discovery pipelines across the three sequencing centers (Human Genome Sequencing Center at Baylor College of Medicine, Broad Institute at MIT and Harvard, and The Genome Center at Washington University), 61 genes representing frequently mutated oncogenes and tumor suppressor genes (Supplementary Table 1) were sequenced independently at all three centers and the resulting verified and/or validated mutations were used for concordance analysis. Mutations observed independently by at least two centers were accepted as true positives and had the highest rate of validation by orthogonal methods while mutations discovered by only a single center from a single amplification attempt were likely to be false positives. Therefore, additional supporting data were generated by orthogonal validation methods or independent amplification attempts followed by manual review, to verify all reported mutations in 623 TSP genes.

Description of PCR-based sequencing pipelines

Primers were designed according to target region(s) of interest. The design process included a validation step for both primers and DNA templates to ensure quality. Each DNA template was then amplified with each primer pair for PCR-directed genomic re-sequencing. Dye terminator sequencing was followed by separation and detection on ABI 3730xl sequencers. The resulting data were screened for mutations through a series of automated and manual steps. A supplemental round of "confirmation" sequencing was performed in an attempt to verify putative mutations, which were then further validated utilizing an alternative genotyping and/or sequencing platform. These steps included the matched normal genomic DNA sample to establish the nature of the variant (somatic or germline). All traces were deposited in the NCBI Trace Archive (http://www.ncbi.nlm.nih.gov/Traces/trace.cgi) under randomized trace names to prevent future identification of the individuals.

Validation approaches

Putative single nucleotide variants were validated on the Illumina Golden Gate Genotyping platform, the Sequenom MassARRAY platform or, alternatively using TaqMan, Biotage assays, or 454 re-sequencing. Putative indels verified after second pass sequencing were manually reviewed and annotated, then validated either on the Biotage pyrosequencer or the 454 FLX. Verified and validated mutations were used for downstream analysis.

454 FLX-based indel validation

Putative indels detected in 3730-based resequencing of TSP samples were validated with PCR-directed 454 sequencing. Resulting 454 reads were mapped to the human reference sequence (hg36) using BLAT and CrossMatch (unpublished). Subsequent processing of the mapped 454 reads identified the gap position, if any, and compared it to the position predicted by Sanger sequencing. The list of indels targeted from 3730 data was cross-referenced with the best alignments to determine the overall 454 sequence coverage in both normal and tumor. Target indels were then cross-referenced with 454 indels, matching up indels of the same type, of similar size (within 2 bp), and having similar chromosomal position (within 2 bp). This step was performed separately for normal and tumor. In the event that multiple 454-detected indels matched a target indel, the one with the highest number of supporting reads was retained. To determine validation status, the list of targeted indels was populated with the 454 read coverage, detected indels, and the number of indel-supporting reads. Target indels were deemed as validated when sufficient read coverage was achieved in both samples and the fraction of the reads harboring the indel exceeded a predetermined threshold. The combination of 3730 trace and 454 validation data were used to determine the status of indels.

Biotage-based indel detection

The validation process involves assay design, assay validation, and tumor sample analysis. The process begins with the PSQ Assay Design Software distributed by Biotage which identifies two amplification primers flanking the sequence of interest (one of which is biotinylated) and a sequencing primer. The assay design is validated by both PCR product generation and subsequent Pyrosequencing using human genomic DNA as a template. If the assay works on human genomic DNA, the process is repeated with the designated tumor samples. The data is analyzed by the instrument software and reviewed and assembled by the technician.

Indel annotation

Boundaries of insertion, deletion and complex rearrangements are annotated as follows (see also Reporting Mutations, "MAF file format", below)
Insertions:

* Start Position is the base before the insertion site
* End Position is the base after the insertion site
* The reference sequence is reported as "-"
* The inserted sequence is reported on the positive genomic strand

* When multiple alignments are possible, the position is reported as the 3' most alignment on the annotated gene's strand.

Deletions:
* Start Position is the first base deleted
* End Position is the last base deleted
* The reference sequence is reported as the sequence from the Start Position to the End Position on the positive genomic strand
* The deleted sequence is reported as "-"
* When multiple alignments are possible the position is shifted to the 3' most alignment on the annotated gene's strand.

Complex Indels and Insertions and Deletions with multiple complex alignments:
* Start Position is the first base deleted or first base of the repeat
* End Position is the last base deleted or the last base of the repeat
* The reference sequence is reported as sequence deleted from the Start Position to End Position on the positive genomic strand
* The inserted sequence is reported as the sequence that replaces the reference sequence on the positive genomic strand

Detecting dinucleotide polymorphisms in 454 sequence data

Dinucleotide polymorphisms (DNPs) were detected during the pairwise sequence comparison between read and reference sequences. A DNP event was identified as two substitutions (relative to the reference) at consecutive aligned bases within a single read. All DNP events were then combined to calculate the number of reads supporting each DNP.

Target coverage calculation

Since the position of a SNP might receive a lower quality score due to the presence of double peaks, a sliding window of 15bp was used to evaluate the coverage of the center base of each window. The center base of each window was considered covered when at least 2/3 of the window's bases, typically 10bp out of the 15bp, were covered by Phred quality scores of at least Q30.

In silico predictions of functional consequences of mutations

We used SIFT[4] and PolyPhen[5] to evaluate the potential impact on protein function for 811 missense mutations. Given the amino acid substitution and protein fasta sequence as inputs, these two programs return predictions of possible functional impact for non-synonymous variants. PolyPhen identifies homologues via BLAST search in the NR database, while SIFT searches against the human SWALL database. SIFT predicts functional changes based on sequence homology and physical properties of amino acids while PolyPhen uses sequence, phylogenetic, and structural information to characterize the substitutions. SIFT predicted 430 missense mutations as deleterious and 297 as tolerated. PolyPhen predicted 438 mutations as probably/possibly damaging and 331 as benign. Taken together, these programs identified 579 mutations having potential

functional relevance and these have been prioritized for further functional studies (Supplementary Table 2).

Comparison of mutations with COSMIC and OMIM databases

Detected and annotated mutations were compared to the COSMIC (version 35) and OMIM (1/15/2008 download) databases to assess novelty. For each mutation, all possible matching transcripts were used to determine possible amino acid changes with respect to both residue and position within the associated protein. This change was then checked against the two databases. If both position and residue were identical, the mutation was declared a match. A match was also declared if only the position within the protein matched that of a mutation found within either database. In cases where a mutation affected a splice site, we checked to see if its genomic coordinates were present in COSMIC. (OMIM does not include genomic coordinates.) Mutations not detected by the above methods were considered novel.

Section III. Mutation profiles, background mutation rate, as well as significantly mutated genes, protein domains, and pathways

Molecular signatures of non-synonymous and synonymous mutations

Mutation signatures may reflect histories of mutagen exposure or DNA mismatch repair defects, thereby providing insights into cancer etiology. We have identified 1,013 non-synonymous mutations in 623 genes and another 108 synonymous mutations in a 250 gene subset. This allowed us to analyze the mutation signatures of non-synonymous and synonymous mutations in parallel. Our results show that point mutations are more common at C:G than at A:T. For non-synonymous mutations, 41% are C:G→A:T transversions, while 27% are C:G →T:A transitions. For synonymous mutations the respective numbers are 26% and 61% (Supplementary Fig. 2a). This result indicates that transversion events, leading to non-synonymous mutations, are positively selected for in tumor cells. We then investigated the potential impact of cigarette smoke on the molecular signature by comparing non-synonymous mutations in smokers versus never-smokers. Interestingly, 43% and 13% C:G→A:T transversions are found in smokers and never-smokers respectively, even though the rates of transition and transversion at A:T are comparable for both cohorts (Supplementary Fig. 2b). This result supports the notion that C:G→A:T transversion events in non-synonymous mutations are likely induced by carcinogens in smoke and demonstrates that previous exposure to tobacco-borne carcinogens impacts the mutational signature.

Mutational signatures already have been shown to vary in different cancer types and are characterized by specific sequence contexts. Previous studies also showed that C:G →T:A transitions at CpG sites were targets of activated carcinogens in gastric and colon cancers. We expanded our signature analysis to measure all 12 possible changes of the transcribed strand in various 5' and 3' sequence contexts and showed that a G→T transversion of the transcribed strand with a 5' flanking T and a 3' flanking G is the most commonly observed molecular change in lung cancer. This is followed by a G→T transversion with a 5' G and a 3' T (Supplementary Table 4a). Our results also showed that transition and transversion events at CpGs are not preferentially targeted in lung cancer. The sequence contexts for synonymous mutations (Supplementary Table 4b) are used for background mutation rate estimation.

The relatively large sample size used in this study allowed us to compare the mutational signatures of several frequently mutated genes in lung cancer. *KRAS* and *TP53*, already associated with cigarette smoking, are central to the hypothesis that specific carcinogens link nicotine addiction to lung cancer. Our analysis clearly shows that point mutations at C:G are not only most common in *KRAS* and *TP53* but also in three additional highly mutated genes, *STK11*, *NF1*, and *LRP1B* (Supplementary Table 4d). Furthermore, point mutations in those genes often occur at a G residue on the transcribed strand, establishing a positive relationship between cigarette consumption and C:G→A:T transversions in those genes. Our observations are generally consistent with the fact that most activated carcinogens react predominantly at G and that repair of the resulting adducts would be slower on the non-transcribed strand. A contrasting result is found in the mutational signature of *EGFR* where point mutations are found mainly at A:T and the most frequently observed sequence context is an A:T→C:G transversion flanked by 5' G:C

and 3'C:G (Supplementary Table 4d) . Hence, our mutational signature analysis supports the hypothesis that cigarette smoke-borne carcinogens do not induce *EGFR* mutations.

Statistical analysis of significantly mutated genes

Statistical tests on the observed mutations across samples attempt to identify genes that harbor mutations under selection during tumor development. The general framework is first to estimate a background mutation rate, ρ, based on mutations that were not under selection and then identify genes mutated beyond this rate. It is widely presumed that non-synonymous mutations can be under selection pressure because of their resultant changes to protein sequence and function, while synonymous mutations are not selected against because they are generally considered to be biologically silent. However, they have been implicated to have deleterious effects in some cases [6-9]. This premise provides several convenient bases for estimating ρ, including data-driven methods and methods that use prior knowledge about highly mutable sequence contexts (Supplementary Tables 4&5 and Supplementary Information). The resulting estimates of background mutation rate ranged between 2.98e-6 and 3.30e-6. These estimates do not account for possible non-synonymous mutations that arrest cell growth, which are presumably under negative selection during carcinogenesis. Consequently, the actual rate could be lower, meaning our analysis is conservative.

Statistical treatments and analyses of cancer sequencing results are not yet standardized. There are several important aspects of this problem, each of which can be addressed in a number of possible ways. Here, we describe several procedures for each of the 3 sub-problems: background mutation rate estimation, significance testing, and error analysis. (Of course, we realize additional methods are possible.) We then describe the 3 overall approaches (out of the many possible combinations of the above) that we used for analysis.

(i)Background mutation rate estimation

The number of somatic mutations in each tumor varies widely among the 188 tumors. Across the 623 sequenced genes, 25 tumors had no somatic mutations in the genes we assayed, while others contain exceptionally large numbers of mutations (Fig. 4). We calculated the synonymous mutation rate as the maximum likelihood ratio of the 108 silent mutations in 188 samples encompassing 95,738,393 covered bases over 250 genes. This yields ρ_{SN} = 1.128e-06. The 95% Wilson confidence interval is 0.93e-06 to 1.36e-06. (The Wilson interval is preferable to the traditional Wald interval for low values (see e.g. L.D. Brown, T.T. Cai, A. DasGupta (2001) "Interval Estimation for a Binomial Proportion" Statistical Science, 16(2) 101-133.)

The non-synonymous background rate ρ_{NS} is due to *passenger* non-synonymous mutations, whose number is obviously not known. We can estimate ρ_{NS} in an expected-value context according to $\rho_{NS} = \rho_{SN} \times R$, where R is the ratio of NS to SN changes. R is calculated by examining all possible residue changes for a single mutation based on the codon usage in all genes sequenced in this study. More specifically, each nucleotide of each codon was mutated *in silico* to determine whether a particular change would result in a NS or SN alteration. While this process is straightforward, the choice of how to

weight each hypothetical NS or SN alteration is less clear. An obvious basis for weighting is the sequence-context based mutation spectra for synonymous mutations observed in lung adenocarcinoma (Supplementary Table 4b). However, there are numerous reasonable methods for choosing the specific weights, all of which lead to some amount of variation in the final estimate of ρ_{NS}. We will describe four possible methods here, along with some caveats for each. These methods can be thought of as hierarchical in the sense that they involve increasing degrees of merging context information. (This is quite analogous, for instance, to chi-square testing when there are too few data for the initial number of categories, requiring some categories to be merged.) For example, if we consider the influence of just the two adjacent bases, Supplementary Table 4b represents the extreme case of maximizing the "resolution" of context information. Sampling zeros and data sparseness, along with the possibility that the strictly-synonymous context information itself could bias the result toward synonymous changes, raises the additional possibility of including a carefully chosen number of non-synonymous mutations, as well. (This is something of a "bootstrap" method.)

(1) Direct Weighting: Weight each alteration directly according to its frequency in Supplementary Table 4b (R = 2.41, ρ_{NS} =2.72e-06).The large number of sampling zeros and low integer cell values imply unacceptably large error bars, rendering this method of little practical use. Additional data could remedy this problem.

(2) Uniform Weighting: At the opposite extreme, we could instead give uniform weight to all alterations (R = 3.38, ρ_{NS} = 3.81e-06), but this entirely overlooks differences in context that we know actually exist in these data.

(3) Data-Driven Weighting: Use a data-driven approach to collapse categories in Supplementary Table 4b based on using Fisher's exact test. Specifically, we are doing 2X3 table where the two columns are the two contexts being tested and the three rows represent the three possibilities of any bases could be mutated. The test evaluates whether there is enough information between two different contexts (where a "context" consists of 2 specific flanking nucleotides) of cells to consider them "statistically distinct", whereby they are not collapsed. If there is not enough information to distinguish the contexts, they are collapsed into a single one. For example, if A:A and A:C are the contexts (where the colon is the intervening base position whose mutations are being examined), then the collapsed context would be A:AC. In other words, the downstream flanking base would be either A or C, if there is insufficient information to statistically distinguish A:A from A:C. The collapsing process continues until no further moves can be made within a user-chosen significance threshold. This approach is supposed to provide a non-biased pruning of categories without having to make *ad hoc* assumptions. In addition to the 108 synonymous mutations mentioned above, we included 156 non-synonymous point mutations in this process that we reasonably felt were non-driver occurrences. Specifically, they were picked from genes that had only single non-synonymous events. For a threshold of 5%, we obtain the 18 sequence context categories shown in Supplementary Table 5a. The overall ratio R=2.645 was obtained by concatenating all gene sequences for analysis, while still observing proper context at exon boundaries. This resulted in an overall non-

synonymous rate estimate of ρ_{NS} = 2.98e-06. Gene-specific *R* values and their subsequent non-synonymous rates were obtained by running the same calculation method on individual genes (Supplementary 5b).

(4) Prior Knowledge Weighting: Take prior knowledge and additional biological intuition into account to collapse categories in Supplementary Table 5c. For example, it is well-known that CpG dinucleotides have elevated mutation rates (> 16 fold higher than A's and T's) due to cytosine deamination. Consequently, any C or G appearing in CpG is assigned a particular weight. Similarly, we assign weights to C or G not in CpG and to any A or T. Here, we used the 108 synonymous mutations discussed above, but we again relaxed the restriction on non-synonymous mutations slightly. That is, we included non-synonymous events from those genes not in the list of top 100 mutated genes (sorted by number of mutations per covered bases), which allows inclusion of indels, as there are no silent indel mutations. The resulting 4 categories are summarized in Supplementary Table 5c. The relative mutation rates are then determined directly from these data and are subsequently used as multipliers for their respective raw *in silico* mutation tallies. Finally, the resulting grand totals for non-synonymous and synonymous tallies lead to a ratio of R = 2.93 and a non-synonymous rate of ρ_{NS} = 3.30e-06.

Although we feel the two latter methods described above are an improvement over those adopted in the past, background rate estimation is clearly still something of an art. For example, neither accounts for possible non-synonymous mutations that arrest cell growth. These are presumably under negative selection during carcinogenesis. Consequently, the actual rate could be lower, meaning our analysis is conservative.

(ii) Significance Testing

We used three statistical methods for testing, all based on the nature of mutation, i.e. a base position is either mutated, or it is not mutated. Each is described briefly:

(1) Standard Test: This approach treats all mutations equally. Given a uniform probability for each position in a gene, a one-tailed binomial test was used to assess whether the observed mutation rate was significantly higher than the background mutation rate. Here, P values were obtained directly from the binomial distribution (W. Feller (1968) An Introduction to Probability Theory and Its Applications, John Wiley and Sons). One set of results shown in Supplementary Table 6a use this standard binomial test based on the overall mutation rate of 2.98e-06, as described above.

(2) Gene-Specific Test: We also applied the method of the standard test, except where we used gene-specific background rates instead of a single, global rate. These results are furnished in Supplementary Table 6a, as well.

(3) Category Test: Score mutations according to prevalence. This test is based on the four sequence mutation categories mentioned above (mutated C or G in a CpG dinucleotide, any other mutated C or G, any mutated A or T, and indel). In short,

convolution is used on the category-specific binomials to obtain an overall P-value[10] (see Appendix A). Using the category-based overall rate of 3.30e-06, this test yielded a third list of significantly mutated genes (Supplementary Table 6a).

In all three of these scenarios, we are testing multiple hypotheses (one per gene). Therefore, we control the false discovery rate using the standard Benjamini and Hochberg False Discovery Rate (FDR) procedure. We used an FDR cutoff of 0.1 which bounds the expected fraction of false positives at 10%.

(iii) Error Analysis

Error analysis is important in any problem where only a single, main analysis method is used because it is typically the only avenue for assessing uncertainty. As with the above issues, there are many possibilities for handling this aspect of the problem.

(1) No formal error handling. Use calculated values of background rate exclusively. Usually, this is acceptable only if an alternate method of assessing accuracy is available, for example using several independent testing methods.

(2) Determine confidence intervals, say at 95%, and re-calculate results at low and high values. In this particular problem, this involves estimating not only the intervals on the synonymous rate and the NS/SN ratio, but also subsequently using error propagation analysis to arrive at an overall interval.

(3) Use Monte Carlo methods to sample the distribution of possible background rates and assess error from sampling statistics.

(iv) Main Methods Used for Analysis

It is easy to see from the above that there are many possibilities for an overall statistical method and that these span various levels of sophistication. We chose 3 of these methods for analyzing our results.

(1) Data-driven weighting to determine background mutation rate with the standard binomial test for assessing statistical significance

(2) Data-driven weighting using the gene-specific significance test

(3) Prior-knowledge weighting using the category-specific test

The first two methods are related, while the third is clearly independent of the others. This observation, coupled with the fact that the resulting analyses from the three methods are quite similar, suggested placing somewhat less emphasis on an extremely formalized error analysis. For example background rates from the data-driven and prior-knowledge

approaches, 2.98e-06 and 3.30e-06 respectively, are quite close in value. In fact, when considering statistical uncertainty, their ranges overlap substantially. For example, the 95% confidence interval on the latter value, 2.56e-06 < ρ_{NS} < 4.04e-06 includes the lower estimate by a large margin. This suggested using these two background rates to directly assess rate sensitivity *in lieu* of using confidence intervals on just a single rate.

Distribution of non-synonymous mutations in protein functional domains

We further characterized the distribution of non-synonymous mutations in protein functional domains and found that 645 mutations were within known conserved Pfam domains. An examination of the various functional domains within those driver genes found non-uniform mutation rates, but when we compared the protein domain mutation rates against the estimated background genomic mutation rate, we found that the RAS (71 mutations), P53 DNA binding (54 mutations), protein tyrosine kinase (104 mutations), protein kinase (103 mutations), and protein phosphatase (7 mutations) domains comprised the top mutated driver protein domains. In addition, fibronectin type III repeat domain (fn3, 25 mutations), Furin-like cysteine rich region (involved in receptor aggregation, 8 mutations), and receptor_L_binding domains (16 mutations) also had mutation rates significantly above background (Supplementary Table 6b). This result suggests that mutations within a group of functional domains are positively selected during carcinogenesis and are excellent candidate mutations for therapeutics.

Concurrence and mutual exclusion analysis

Positive correlations between mutations in two genes across tumors may suggest that the mutations (and their associated pathways) function synergistically to promote carcinogenesis, while negative correlation of mutations in two genes may imply that the alteration of either gene (or its associated pathway) is sufficient, wholly or in part, for carcinogenesis. We performed two slightly different permutation tests for mutation correlation between genes (i.e. concurrence or exclusion of mutations). In both tests, we take inter-individual differences of gene mutation prevalence into consideration. FDR corrections have been applied (Supplementary Table 7a-c).

Test 1: we kept the number of samples mutated in a given gene and the number of genes mutated in a given sample the same as observed in the data and randomly permuted the observed mutations across samples and genes. For each permutation, we recorded the number of samples with concurrent and exclusive mutations (Xc and Xe, repectively) for each pair of genes and compared them with the numbers observed in the original data (Nc and Ne). We repeated this process 10,000,000 times and summarize the frequencies of Xc>=Nc and Xe>=Ne, respectively. These frequencies are used as empirical p-values under the null hypothesis (no correlation between genes) (Supplementary Table 7a & 7b).

Test 2: we kept the number of mutations in a given gene and sample the same as observed in the data and randomly permuted the observed mutations across samples and genes. The rest of the analysis is the same as above and we performed 1,000,000 permutations (Supplementary Table 7c).

Analysis of significantly mutated pathways

We examined the distribution of mutations found in the 623 TSP genes within KEGG pathways (release 45). KEGG pathway maps are hierarchically classified, having 5 categories in the top level, 24 subcategories in the second, and 215 individual pathways in the third. We mapped TSP genes to KEGG pathways both by blast (cutoff: P_value ≤ e-10) and by name matching and were able to place 389 genes onto the pathways. NF1 and KRAS were added to the mTOR pathway based on the literature. We used two approaches to test mutation association on all three levels and to determine whether the genes within specific pathways were mutated more often than the genes outside of the pathways. We evaluated the total number of mutations from all genes examined within each pathway, the number of genes mutated, and the total sequenced coding length of genes in all 188 tumors in the pathway.

In one approach, we created a two-way contingency table by counting mutations in all tested genes within and outside of the pathway. Fisher's exact test was then performed to assess significance of pathway mutation enrichment (Supplementary Table 12).

In another approach, we counted both the total number of base pairs within each pathway that were successfully sequenced and the total number of NS mutations observed in each pathway. A one-tailed binomial test was again performed to check for significance of the observed mutation rate in the pathway. The Benjamini and Hochberg False Discovery Rate and Bonferroni correction procedures were applied according to the total number of pathways tested (Supplementary Table 12). Of those pathways deemed to be significant by this process, we further assessed significance in terms of the number of mutated genes. This step disqualifies pathways having just one or two highly-mutated genes.

Section IV. Integrated analyses – mutation, copy number, gene expression, and clinical features

SNP array data pre-processing

The dChip software was used to read Affymetrix 250k Sty array CEL files and to export background-subtracted probe values using the PM-only option. For each probe set (i.e. each SNP), the $\log_2$ value of the average of PM intensities across 12 probes was used as the basic measurement, *S*. For each array, *S* values for 238,304 SNPs were scaled to have a mean of 0 and a variance of 1 by normalization. For each SNP and each sample pair, DNA copy number was estimated based on the difference of *S* between tumor and normal samples (i.e. the $\log_2$ ratio of the tumor and normal intensities).

Gene-based DNA copy number analyses

For genes having greater than 10 SNPs, all SNPs within the gene were used to compute the gene's copy number. For genes with fewer than 10 SNPs, the average intensity of the 10 SNPs closest to the most centrally localized SNP, within or outside of the gene, was used to calculate the copy number. This approach minimizes the risk of higher variability associated with a small number of SNPs. DNA copy numbers are log transformed for plotting.

Tumor RNA extraction and gene expression array and data analysis

For each tumor RNA used for expression analysis, serial 50 um frozen tissue sections were homogenized in Trizol reagent and used for RNA isolation according to the manufacturer's protocol. An H/E stained serial section from each tumor was reviewed to ensure representation of at least 70% neoplastic cellularity. Isolated RNA was quantified by Nanodrop fiberoptic spectrophotometry and RNA integrity was assessed using Agilent Bioanalyzer and RNA Nano assay. One ug of total RNA was convereted to biotinylated cRNA using the Affymetrix single cycle target preparation method, following the manufacturer's protocol and standard reagents. Biotinylated cRNA was fragmented and hybridized to Affymetrix U133Plus2 human GeneChip expression microarrays following the manufacturerer's recommendations. Arrays were washed, stained, and scanned using standard protocols. All target preparation and array hybridization was performed by the Siteman Cancer Center GeneChip facility. Data quality of each array was assessed by reviewling internal poly-A control probe sets and 3':5' probe intensity signal ratios for housekeeping genes. Arrays which passed all QC metrics were transferred to the Siteman Cancer Center Bioinformatics Core at Washington University for data annotation and distribution. The gene expression data were preprocessed using MAS 5.0. Then, the gene expression levels were log transformed and the mean value across all samples was subtracted. The resulting values were divided by standard deviation to obtain normalized gene expression values. The average value was used where multiple probes were available, and the single value was used when only a single probe was available for a given gene.

DNA copy number and Gene expression correlation analysis

To detect the association between DNA copy number and expression level of each gene, Pearson's correlation coefficients and P-values were calculated. Standard FDR and Bonferroni corrections were applied.

DNA copy number and gene expression

Our previous study described the DNA copy number analysis for TSP tumors based on SNP array data[11]. In this study, RNAs from 75 tumor samples were hybridized to the Affymetrix U133Plus2 gene expression array (see Methods). Further pathological analysis suggests that 7 out of 75 tumors are not typical lung adenocarcinomas and likely belong to other types of lung cancers. Statistically significant correlation (FDR < 0.05) between DNA copy number and gene expression level was observed for 2,112 of the 8,046 genes examined. The distribution of Pearson correlation coefficients between DNA copy number and gene expression is shown in supplementary Fig. 3. Further analysis identified 2,096 genes with positive correlation and only 16 genes with negative correlation between DNA copy number and gene expression level. The causes of expression change for many genes in cancer cells are still largely unknown. Our analysis suggests that DNA copy number alterations may be an important factor for driving altered gene expression in lung adenocarcinoma. Notable examples having strong correlation are the adjacent genes *ANKRD10* and *ING1*, two cyclin-dependant kinases *CDK6* and *CDK8*, PIK3 kinase *PIK3CA*, cyclin *CCNE1*, and the zinc finger protein *MYNN* (Supplementary Table 9a). *CDK8* has been shown as a colorectal cancer oncogene that regulates b-catenin activity (Nature, in press). Strong positive correlations also have been found in highly mutated genes such as *NF1*, *KRAS*, *EGFR*, *APC*, and *RB1* (Supplementary Table 9b). Interestingly, we have also identified genes showing discordant changes between DNA copy number and gene expression (data not shown). Notably, several members of the melanoma antigen (*MAGE*) and solute carrier families, having minimal copy number variations, showed drastic gene expression changes across 75 lung adenocarcinoma tumor samples. This is consistent with the notion that regulation of gene expression is a complex process that can be controlled at several steps and by multiple mechanisms in cancer. Besides copy number changes, changes in methylation status, transcription regulation, transcript stability and many other factors can have a significant impact on the gene expression levels in lung cancer cells.

Correlation between mutation and gene expression

To detect the association between gene expression and mutation of each gene, Pearson's correlation coefficient and P-values were calculated. Standard FDR and Bonferroni corrections were applied.

Correlation between mutation and DNA copy number

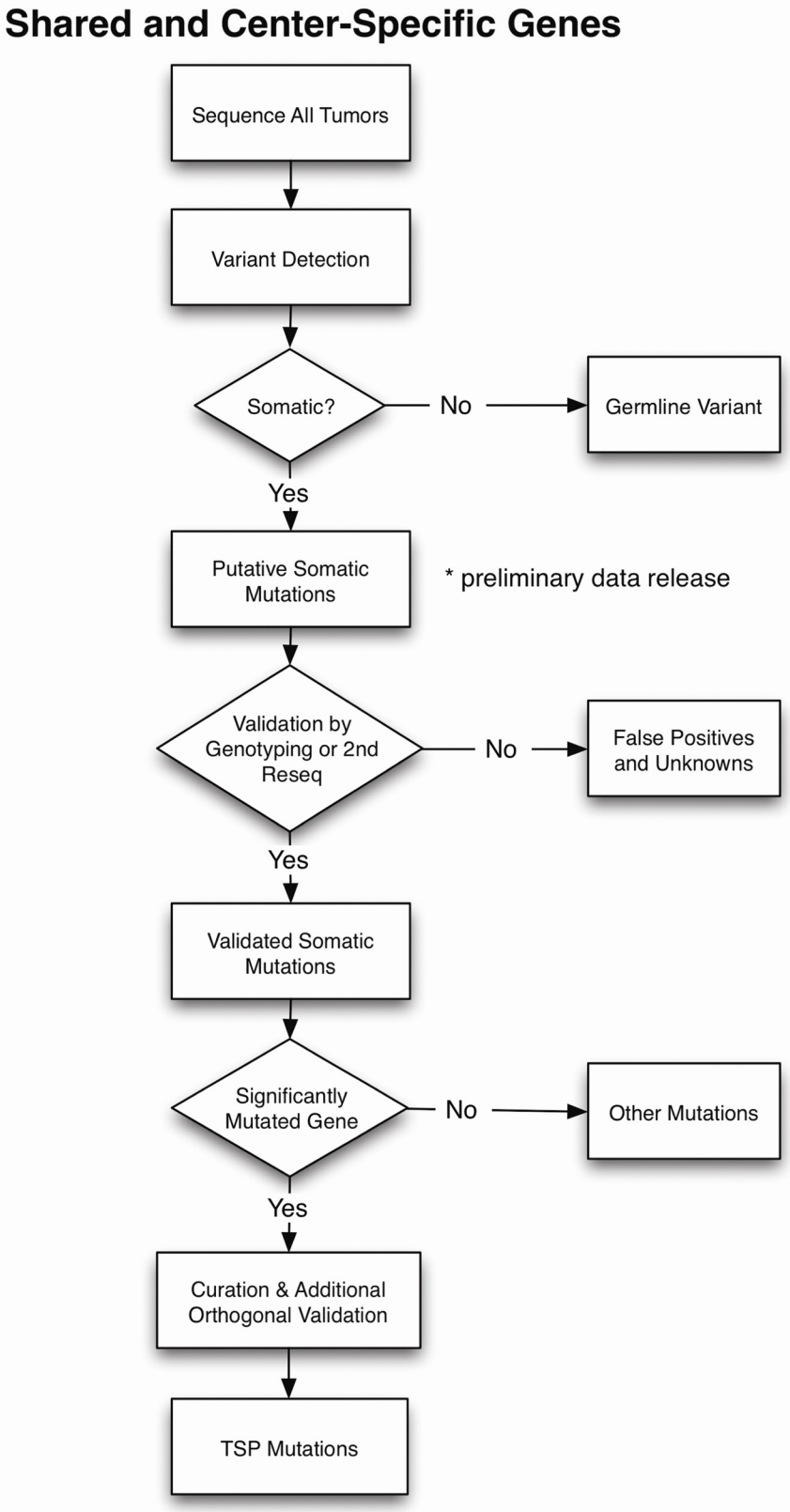
To detect the association between DNA copy number and mutation of each gene, Pearson's correlation coefficient and P-values were calculated. Standard FDR and Bonferroni corrections were applied.

Correlation between mutation and clinical association analysis

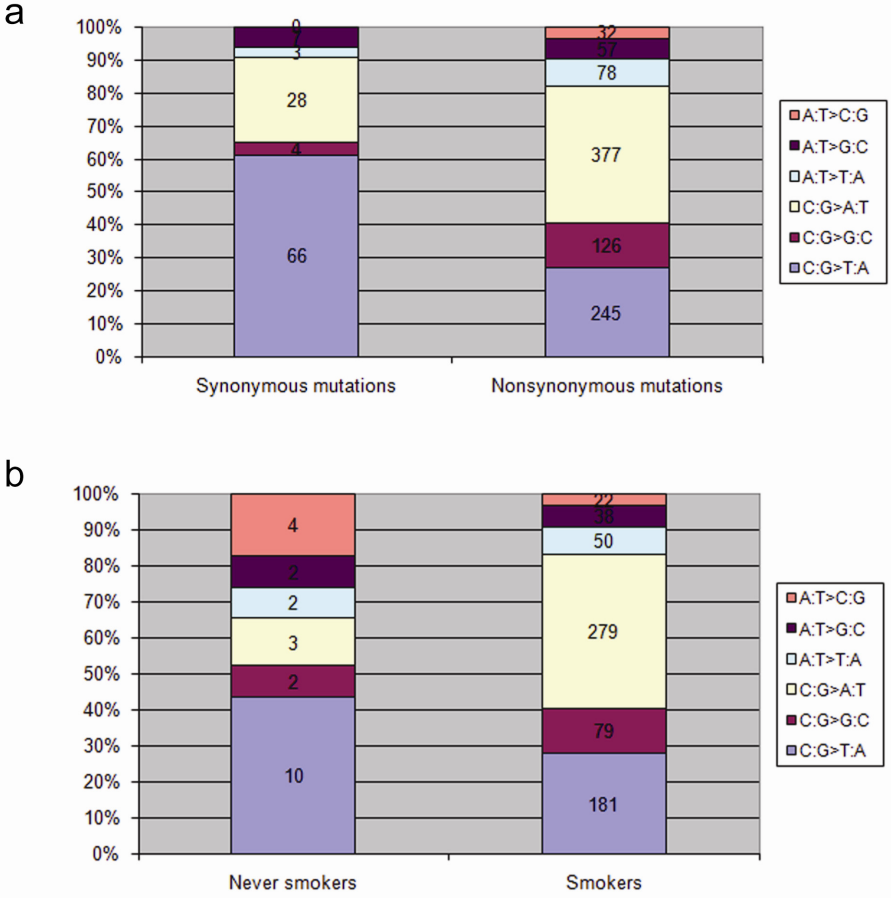
Each gene of each sample was assigned either 1 or 0 according to whether it harbored any mutations. For any two genes, a two-way table was obtained by counting the numbers of 0 and 1 entries over the entire sample set, after which Fisher's exact test was performed to determine their mutation correlation. To establish correlations between mutations and clinical features, Pearson's correlation coefficient and Fisher's exact test were applied (quantitative and qualitative clinical traits, respectively).

B. Supplementary Figures

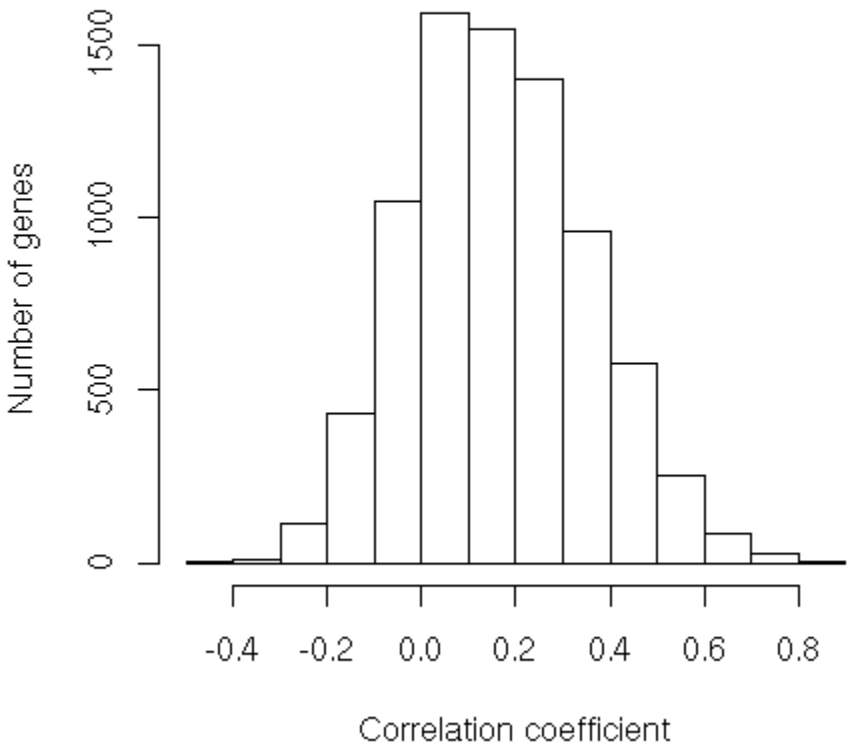
Supplemtary Fig. 1



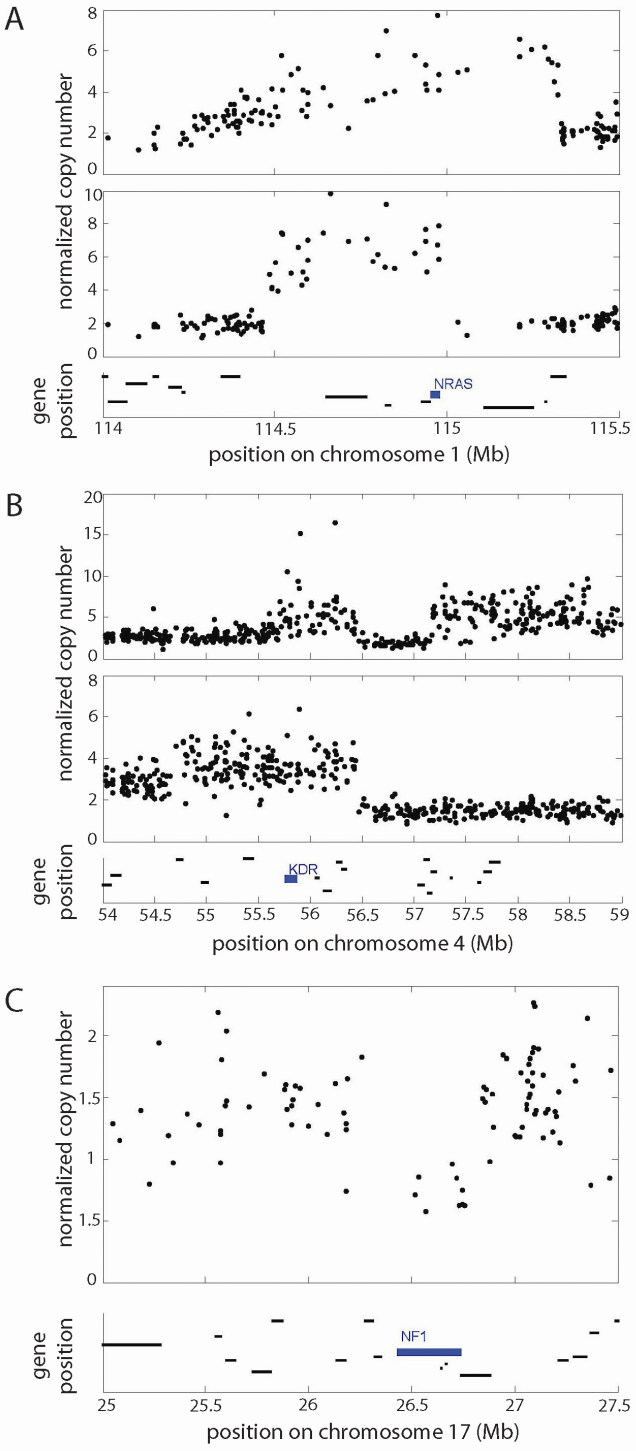
Supplementary Fig. 2



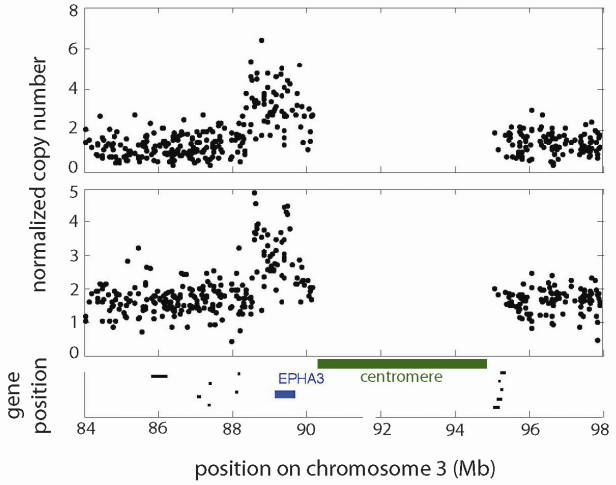
Supplementary Figure 3



Supplementary Fig. 4

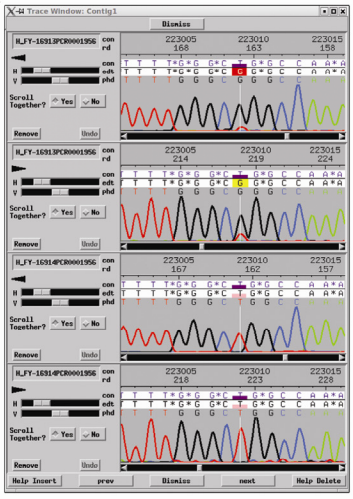


Supplementary Fig.5



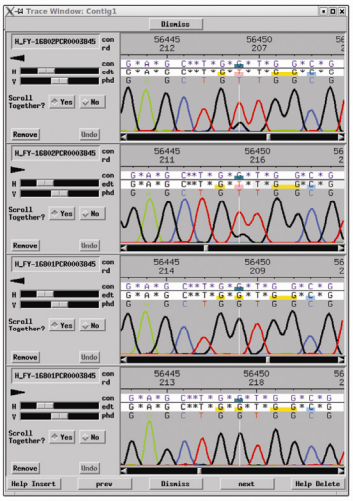
Supplementary Fig. 6

a



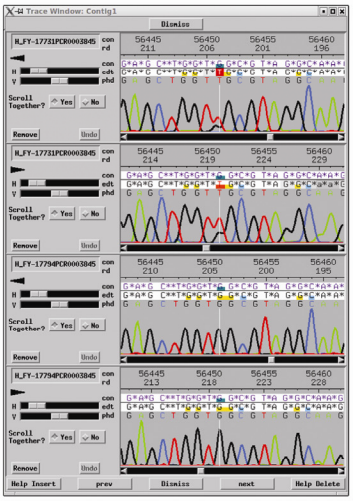
Tumor 16913: amplification of L858R mutant allele
Matched normal 16914

b



Tumor 16802: amplification of G12V mutant allele
Matched normal 16801

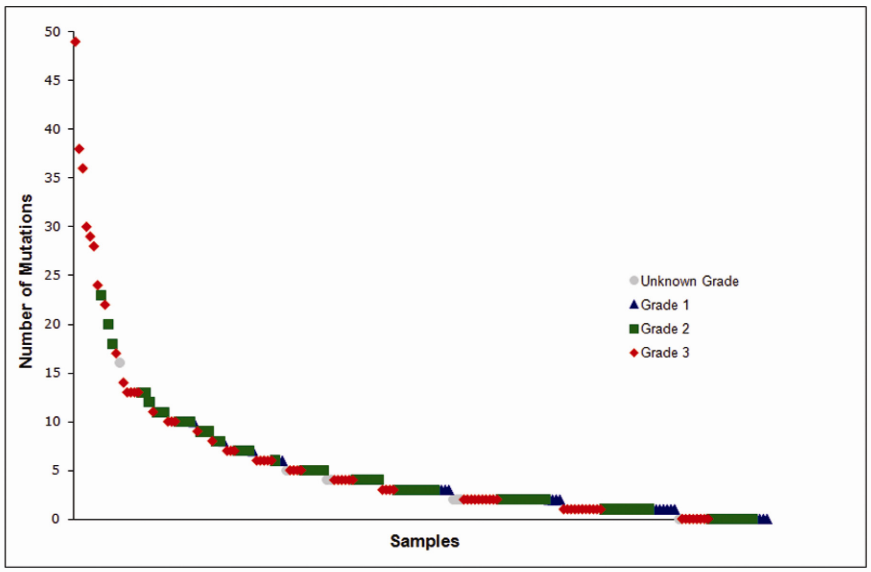
c



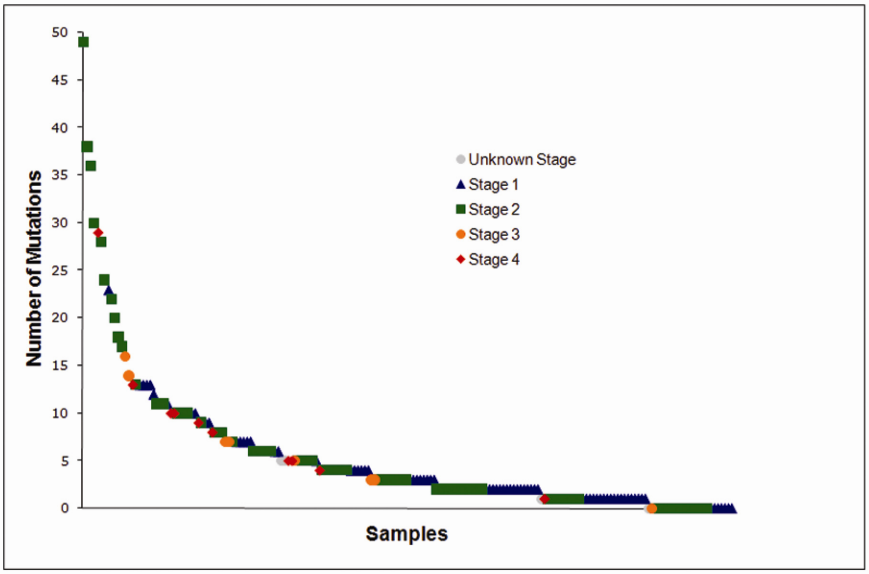
Tumor 17731: amplification of G13C mutant allele
Matched normal 17794

Supplementary Fig. 7

a



b



C. Supplementary Figure Legends

Supplementary Figure 1. Sequencing and mutation detection flow chart. 188 primary lung adencarcinoma samples were screened for mutations in the coding exons as well as splice donor and acceptor sites of 623 human genes. Detected putative somatic mutations were subjected to verification (2^{nd} round of PCR amplification and sequencing), orthogonal validation, and manual review.

Supplementary Figure 2. Mutational signatures in lung adenocarcinoma. a) For non-synonymous mutations, 41% are C:G→A:T transversions, while 27% are C:G →T:A transitions. For synonymous mutations the respective numbers are 26% and 61%. This result suggests that C:G→A:T transversion events, leading to non-synonymous mutations, are positively selected for in tumor cells. b) 43% and 13% C:G→A:T transversions are found in smokers and never smokers respectively, suggesting that C:G→A:T transversion events in non-synonymous mutations are likely induced by carcinogens in smoke.

Suppelemntary Figure 3. Statistical analysis of correlation between DNA copy number alterations and gene expression changes in 75 lung adenocarcinoma samples.

Supplementary Figure 4. Focal copy number changes of significant genes in primary lung adenocarcinoma samples described in Weir et al. A, focal amplifications of *NRAS*. Upper panel, sample Daffy35T. Lower panel, sample Daffy20T. B, focal amplifications of *KDR*. Upper panel, sample 17048. Lower sample, Daffy93T. C, focal deletion in *NF1* in sample Daffy85T. Mutated genes of interest are indicated by blue bars in the figure; locations of neighboring genes indicated by black bars.

Supplementary Figure 5. Focal amplifications of *EPHA3*. Two primary lung adenocarcinoma samples from the set described in Weir et al.[11] show high-level, focal amplification of the *EPHA3* gene, indicated by a blue bar on the figure, as assessed by Affymetrix 250K StyI SNP array data. The positions of neighboring genes are indicated by black bars. The green bar indicates the location of the centromere of chromosome 3. Upper panel, sample 16678. Lower panel, sample 99-43.

Supplementary Figure 6. Lung adenocarcinomas display amplification of mutant alleles of *EGFR* and *KRAS*.

Supplementary Figure 7. Mutation distributions in individual cancer genomes. a) Tumors with higher grade displayed a higher mutation rate than tumors from lower grade. b) Tumors with higher stage displayed a higher mutation rate than tumors from lower stages.

D. Supplementary Tables

Supplementary Table 1. Protein family/domain and GO ontology analysis for the 623 TSP genes.

Gene/Transcript	GO	HMMPfam	Superfamily	Center
AADAC/ENST00000232892		HMMPfam_Abhydrolase_3	superfamily_alpha/beta-Hydrolases	WU
AATK/OTTHUMT00000256055	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	WU
ABL1/ENST00000372348	GO:0007242	HMMPfam_F_actin_bind	superfamily_Protein kinase-like (PK-like)	WU
	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_SH2 domain	
		HMMPfam_SH2	superfamily_SH3-domain	
		HMMPfam_SH3_1		
ABL2/ENST00000367623	GO:0007242	HMMPfam_F_actin_bind	superfamily_Protein kinase-like (PK-like)	Broad
	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_SH2 domain	
		HMMPfam_SH2	superfamily_SH3-domain	
		HMMPfam_SH3_1		
ACTR2/ENST00000377982	GO:0005515	HMMPfam_Actin	superfamily_Actin-like ATPase domain	Broad
ACVR1/ENST00000263640	GO:0004672	HMMPfam_Activin_recp	superfamily_Protein kinase-like (PK-like)	WU
	GO:0004675	HMMPfam_Pkinase	superfamily_Snake toxin-like	
		HMMPfam_TGF_beta_GS		
ACVR1B/ENST00000257963	GO:0004672	HMMPfam_Activin_recp	superfamily_Protein kinase-like (PK-like)	Broad
	GO:0004675	HMMPfam_Pkinase	superfamily_Snake toxin-like	
		HMMPfam_TGF_beta_GS		
ACVR1C/ENST00000243349	GO:0004672	HMMPfam_Activin_recp	superfamily_Protein kinase-like (PK-like)	WU
	GO:0004675	HMMPfam_Pkinase	superfamily_Snake toxin-like	
		HMMPfam_TGF_beta_GS		
ACVR2B/ENST00000352511	GO:0004672	HMMPfam_Activin_recp	superfamily_Protein kinase-like (PK-like)	WU
	GO:0004675	HMMPfam_Pkinase	superfamily_Snake toxin-like	
ACVRL1/OTTHUMT00000260987	GO:0004672	HMMPfam_Activin_recp	superfamily_Protein kinase-like (PK-like)	Broad
	GO:0004675	HMMPfam_Pkinase	superfamily_Snake toxin-like	
		HMMPfam_TGF_beta_GS		
ADRBK1/ENST00000308595	GO:0004672	HMMPfam_PH	superfamily_PH domain-like	Broad
	GO:0004871	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_RGS	superfamily_Regulator of G-protein signaling RGS	
ADRBK2/ENST00000324198	GO:0004672	HMMPfam_PH	superfamily_PH domain-like	Broad
	GO:0004871	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_RGS	superfamily_Regulator of G-protein signaling RGS	

AKT1/ENST00000349310	GO:0004672 GO:0004674	HMMPfam_PH HMMPfam_Pkinase HMMPfam_Pkinase_C	superfamily_PH domain-like superfamily_Protein kinase-like (PK-like)	SHARE
AKT2/ENST00000311278	GO:0004672 GO:0004674	HMMPfam_PH HMMPfam_Pkinase HMMPfam_Pkinase_C	superfamily_PH domain-like superfamily_Protein kinase-like (PK-like)	SHARE
AKT3/ENST00000366539	GO:0004672 GO:0004674	HMMPfam_PH HMMPfam_Pkinase HMMPfam_Pkinase_C	superfamily_PH domain-like superfamily_Protein kinase-like (PK-like)	SHARE
ALK/ENST00000389049	GO:0016020 GO:0004713	HMMPfam_MAM HMMPfam_Pkinase_Tyr	superfamily_EGF/Laminin superfamily_LDL receptor-like module superfamily_Protein kinase-like (PK-like)	Broad
ALS2CR2/ENST00000194530	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
ALS2CR7/ENST00000260967	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
AMHR2/ENST00000257863	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_Snake toxin-like	Broad
ANKK1/ENST00000303941	GO:0004672	HMMPfam_Ank HMMPfam_Pkinase	superfamily_Ankyrin repeat superfamily_Protein kinase-like (PK-like)	Broad
APC/ENST00000257430	GO:0005515 GO:0000074	HMMPfam_APC_15aa HMMPfam_APC_basic HMMPfam_APC_crr HMMPfam_Arm HMMPfam_EB1_binding HMMPfam_SAMP	superfamily_ARM repeat	Broad
ARAF/ENST00000377045	GO:0007242 GO:0004672 GO:0005057	HMMPfam_C1_1 HMMPfam_Pkinase HMMPfam_RBD	superfamily_Cysteine-rich domain superfamily_Protein kinase-like (PK-like) superfamily_Ubiquitin-like	SHARE
ARMET/ENST00000273628			superfamily_Saposin	WU
ATM/ENST00000278616	GO:0016773	HMMPfam_FAT HMMPfam_FATC HMMPfam_PI3_PI4_kinase	superfamily_ARM repeat superfamily_Protein kinase-like (PK-like)	Broad
ATR/ENST00000383103	GO:0016773	HMMPfam_FATC HMMPfam_HEAT	superfamily_ARM repeat superfamily_Protein kinase-like (PK-like)	Broad

		HMMPfam_PI3_PI4_kinase HMMPfam_UME	superfamily_TPR-like	
AURKB/OTTHUMT00000285841	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
AURKC/ENST00000302804	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
AXL/ENST00000301178	GO:0004713	HMMPfam_I-set HMMPfam_Pkinase_Tyr HMMPfam_fn3 HMMPfam_ig	superfamily_Fibronectin type III superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain	WU
BAD/ENST00000309032	GO:0008632 GO:0019221 GO:0005737 GO:0006007 GO:0042593 GO:0008624 GO:0016020 GO:0005741 GO:0005739 GO:0045579 GO:0045582 GO:0005515 GO:0043281			Broad
BAGE3/ENST00000335369		HMMPfam_BAGE		Broad
BAGE4/ENST00000335369		HMMPfam_BAGE		Broad
BAP1/ENST00000296288	GO:0004221	HMMPfam_Peptidase_C12	superfamily_Cysteine proteinases	Broad
BARD1/ENST00000260947	GO:0005622	HMMPfam_Ank HMMPfam_BRCT	superfamily_Ankyrin repeat superfamily_BRCT domain superfamily_RING/U-box	WU
BAX/OTTHUMT00000150522	GO:0042981	HMMPfam_Bcl-2	superfamily_Bcl-2 inhibitors of programmed cell de	WU
BCL2/ENST00000333681	GO:0042981	HMMPfam_BH4 HMMPfam_Bcl-2	superfamily_Bcl-2 inhibitors of programmed cell de	WU
BCL3/OTTHUMT00000150537		HMMPfam_Ank	superfamily_Ankyrin repeat	WU
BCL6/ENST00000232014	GO:0003676	HMMPfam_BTB HMMPfam_zf-C2H2	superfamily_C2H2 and C2HC zinc fingers superfamily_POZ domain	Broad

BLK/ENST00000382364	GO:0007242 GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_SH2 HMMPfam_SH3_1	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain superfamily_SH3-domain	SHARE
BMPR1A/ENST00000372037	GO:0004672 GO:0004675	HMMPfam_Activin_recp HMMPfam_Pkinase HMMPfam_TGF_beta_GS	superfamily_Protein kinase-like (PK-like) superfamily_Snake toxin-like	Broad
BMPR1B/ENST00000264568	GO:0004672 GO:0004675	HMMPfam_Activin_recp HMMPfam_Pkinase HMMPfam_TGF_beta_GS	superfamily_Protein kinase-like (PK-like) superfamily_Snake toxin-like	Broad
BMPR2/ENST00000374580	GO:0004672 GO:0004675	HMMPfam_Activin_recp HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_Snake toxin-like	Broad
BMX/ENST00000380391	GO:0007242 GO:0004713	HMMPfam_BTK HMMPfam_PH HMMPfam_Pkinase_Tyr HMMPfam_SH2	superfamily_PH domain-like superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain	WU
BRAF/ENST00000288602	GO:0007242 GO:0004672 GO:0005057	HMMPfam_C1_1 HMMPfam_Pkinase HMMPfam_RBD	superfamily_Cysteine-rich domain superfamily_Protein kinase-like (PK-like) superfamily_Ubiquitin-like	SHARE
BRAP/ENST00000327551	GO:0005515 GO:0008270	HMMPfam_BRAP2 HMMPfam_zf-C3HC4 HMMPfam_zf-UBP	superfamily_RING/U-box superfamily_RNA-binding domain RBD	Broad
BRCA1/ENST00000357654	GO:0005622 GO:0005515	HMMPfam_BRCT HMMPfam_zf-C3HC4	superfamily_BRCT domain superfamily_RING/U-box	Broad
BRCA2/ENST00000380152		HMMPfam_BRCA-2_OB1 HMMPfam_BRCA-2_OB3 HMMPfam_BRCA-2_helical HMMPfam_BRCA2 HMMPfam_Tower	superfamily_BRCA2 helical domain superfamily_BRCA2 tower domain superfamily_Nucleic acid-binding proteins	Broad
BTK/ENST00000308731	GO:0007242 GO:0004713	HMMPfam_BTK HMMPfam_PH HMMPfam_Pkinase_Tyr HMMPfam_SH2 HMMPfam_SH3_1	superfamily_PH domain-like superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain superfamily_SH3-domain	Broad

BUB1/ENST00000302759	GO:0004672	HMMPfam_Mad3_BUB1_I HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
BUB1B/ENST00000287598		HMMPfam_Mad3_BUB1_I	superfamily_Protein kinase-like (PK-like)	WU
C3orf14/NM_020685				Broad
C3orf18/ENST00000357203	GO:0016021 GO:0016020			WU
C3orf45/ENST00000316436	GO:0016021 GO:0016020			WU
CACNA2D2/ENST00000360963	GO:0016020	HMMPfam_Cache_1 HMMPfam_VGCC_alpha2 HMMPfam_VWA HMMPfam_VWA_N	superfamily_vWA-like	WU
CADPS/ENST00000383710		HMMPfam_C2 HMMPfam_DUF1041 HMMPfam_PH	superfamily_PH domain-like	Broad
CAMKV/ENST00000389606	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
CASK/ENST00000378163	GO:0005515 GO:0004672	HMMPfam_Guanylate_kin HMMPfam_L27 HMMPfam_PDZ HMMPfam_Pkinase HMMPfam_SH3_2	superfamily_L27 domain superfamily_P-loop containing nucleoside triphosph superfamily_PDZ domain-like superfamily_Protein kinase-like (PK-like) superfamily_SH3-domain	Broad
CBL/ENST00000264033	GO:0005515 GO:0004871	HMMPfam_Cbl_N HMMPfam_Cbl_N2 HMMPfam_Cbl_N3 HMMPfam_UBA HMMPfam_zf-C3HC4	superfamily_EF-hand superfamily_N-terminal domain of cbl (N-cbl) superfamily_RING/U-box superfamily_SH2 domain superfamily_UBA-like	Broad
CBLB/ENST00000264122	GO:0005515 GO:0004871	HMMPfam_Cbl_N HMMPfam_Cbl_N2 HMMPfam_Cbl_N3 HMMPfam_UBA HMMPfam_zf-C3HC4	superfamily_EF-hand superfamily_N-terminal domain of cbl (N-cbl) superfamily_RING/U-box superfamily_SH2 domain superfamily_UBA-like	WU
CBLC/ENST00000270279	GO:0005515 GO:0004871	HMMPfam_Cbl_N HMMPfam_Cbl_N2	superfamily_EF-hand superfamily_N-terminal domain of cbl (N-cbl)	WU

		HMMPfam_Cbl_N3 HMMPfam_zf-C3HC4	superfamily_RING/U-box superfamily_SH2 domain	
CCNA1/ENST00000255465	GO:0000074	HMMPfam_Cyclin_C HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCNA2/ENST00000274026	GO:0000074	HMMPfam_Cyclin_C HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCNB3/NM_033031	GO:0000074	HMMPfam_Cyclin_C HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCND1/ENST00000227507	GO:0000074	HMMPfam_Cyclin_C HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCND2/ENST00000261254	GO:0000074	HMMPfam_Cyclin_C HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCND3/ENST00000372991	GO:0000074	HMMPfam_Cyclin_C HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCNE1/ENST00000262643	GO:0000074	HMMPfam_Cyclin_C HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCNE2/ENST00000308108	GO:0000074	HMMPfam_Cyclin_C HMMPfam_Cyclin_N	superfamily_Cyclin-like	WU
CCNG1/ENST00000340828	GO:0000074	HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCNG2/ENST00000354403	GO:0000074	HMMPfam_Cyclin_N	superfamily_Cyclin-like	WU
CCNH/ENST00000380104	GO:0000074	HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCNK/OTTHUMT00000072347	GO:0000074	HMMPfam_Cyclin_N	superfamily_Cyclin-like	BCM
CCNT1/ENST00000261900	GO:0000074	HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCNT2/ENST00000264157	GO:0000074	HMMPfam_Cyclin_N	superfamily_Cyclin-like	Broad
CCRK/ENST00000286878	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
CDC2/ENST00000373811	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
CDC25A/ENST00000302506	GO:0000087	HMMPfam_M-inducer_phosp HMMPfam_Rhodanese	superfamily_Rhodanese/Cell cycle control phosphata	WU
CDC25B/NM_021873	GO:0000087	HMMPfam_M-inducer_phosp HMMPfam_Rhodanese	superfamily_Rhodanese/Cell cycle control phosphata	BCM
CDC25C/ENST00000323760	GO:0000087	HMMPfam_M-inducer_phosp HMMPfam_Rhodanese	superfamily_Rhodanese/Cell cycle control phosphata	WU
CDC2L2/NM_024011	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM

CDC2L5/ENST00000181839	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
CDC2L6/ENST00000368911	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
CDC42BPA/ENST00000366766	GO:0007242	HMMPfam_C1_1	superfamily_Cysteine-rich domain	WU
	GO:0004672	HMMPfam_CNH	superfamily_PH domain-like	
	GO:0004674	HMMPfam_DMPK_coil	superfamily_Protein kinase-like (PK-like)	
	GO:0005083	HMMPfam_PH		
		HMMPfam_Pkinase		
		HMMPfam_Pkinase_C		
CDC42BPB/ENST00000361246	GO:0007242	HMMPfam_C1_1	superfamily_Cysteine-rich domain	BCM
	GO:0004672	HMMPfam_CNH	superfamily_PH domain-like	
	GO:0004674	HMMPfam_DMPK_coil	superfamily_Protein kinase-like (PK-like)	
	GO:0005083	HMMPfam_PBD		
		HMMPfam_PH		
		HMMPfam_Pkinase		
		HMMPfam_Pkinase_C		
CDC42BPG/ENST00000302001	GO:0007242	HMMPfam_C1_1	superfamily_Cysteine-rich domain	BCM
	GO:0004672	HMMPfam_CNH	superfamily_PH domain-like	
	GO:0004674	HMMPfam_DMPK_coil	superfamily_Protein kinase-like (PK-like)	
	GO:0005083	HMMPfam_PBD		
		HMMPfam_PH		
		HMMPfam_Pkinase		
		HMMPfam_Pkinase_C		
CDC7/ENST00000370415	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
CDH1/ENST00000268794	GO:0005509	HMMPfam_Cadherin	superfamily_Cadherin-like	BCM
		HMMPfam_Cadherin_C		
		HMMPfam_Cadherin_pro		
CDH11/ENST00000268603	GO:0005509	HMMPfam_Cadherin	superfamily_Cadherin-like	BCM
		HMMPfam_Cadherin_C		
CDH4/ENST00000360469	GO:0005509	HMMPfam_Cadherin	superfamily_Cadherin-like	BCM
		HMMPfam_Cadherin_C		
		HMMPfam_Cadherin_pro		
CDK10/ENST00000353379	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
CDK2/ENST00000266970	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM

CDK3/ENST00000293215	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
CDK4/ENST00000257904	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
CDK5/ENST00000297518	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
CDK6/ENST00000265734	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
CDK7/ENST00000380832	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
CDK8/ENST00000381527	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
CDK9/ENST00000373265	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
CDKN1A/ENST00000373712	GO:0004861	HMMPfam_CDI		BCM
CDKN1B/ENST00000228872	GO:0004861	HMMPfam_CDI		Broad
CDKN1C/ENST00000313407	GO:0004861	HMMPfam_CDI		BCM
CDKN2A/ENST00000361570		HMMPfam_P19Arf_N		SHARE
CDKN2B/ENST00000276925		HMMPfam_Ank	superfamily_Ankyrin repeat	WU
CDKN2C/ENST00000262662		HMMPfam_Ank	superfamily_Ankyrin repeat	WU
CDKN2D/ENST00000335766		HMMPfam_Ank	superfamily_Ankyrin repeat	BCM
CDKN3/ENST00000335183		HMMPfam_CDKN3	superfamily_(Phosphotyrosine protein) phosphatases	Broad
CEBPA/OTTHUMT00000150648	GO:0003677	HMMPfam_bZIP_2		Broad
CHEK1/ENST00000278916	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
CHEK2/ENST00000382580	GO:0004672	HMMPfam_FHA HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_SMAD/FHA domain	Broad
CISH/OTTHUMT00000264778	GO:0007242	HMMPfam_SH2 HMMPfam_SOCS_box	superfamily_SH2 domain	WU
CNKSR1/ENST00000374253	GO:0005515	HMMPfam_PH HMMPfam_SAM_2	superfamily_PDZ domain-like superfamily_PH domain-like superfamily_SAM/Pointed domain	Broad
CNTN4/OTTHUMT00000263484		HMMPfam_I-set HMMPfam_fn3 HMMPfam_ig	superfamily_Fibronectin type III superfamily_Immunoglobulin	BCM
CNTN6/ENST00000350110		HMMPfam_I-set HMMPfam_fn3 HMMPfam_ig	superfamily_Fibronectin type III superfamily_Immunoglobulin	BCM
CRK/ENST00000382237	GO:0007242	HMMPfam_SH2 HMMPfam_SH3_1	superfamily_SH2 domain superfamily_SH3-domain	SHARE

		HMMPfam_SH3_2		
CRK7/ENST00000300647	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
CRKL/ENST00000354336	GO:0007242	HMMPfam_SH2 HMMPfam_SH3_1 HMMPfam_SH3_2	superfamily_SH2 domain superfamily_SH3-domain	SHARE
CROT/ENST00000331536	GO:0008415	HMMPfam_Carn_acyltransf	superfamily_CoA-dependent acyltransferases	WU
CSF1R/ENST00000286301	GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_V-set HMMPfam_ig	superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	BCM
CSK/ENST00000220003	GO:0007242 GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_SH2 HMMPfam_SH3_1	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain superfamily_SH3-domain	Broad
CTNNB1/ENST00000349496		HMMPfam_Arm HMMPfam_HEAT	superfamily_ARM repeat	WU
CYB561D2/ENST00000232508	GO:0006118	HMMPfam_Cytochrom_B561		WU
CYLD/ENST00000311559		HMMPfam_CAP_GLY	superfamily_Cap-Gly domain superfamily_Cysteine proteinases	WU
CYSLTR2/ENST00000282018	GO:0001584	HMMPfam_7tm_1	superfamily_Family A G protein-coupled receptor-li	Broad
DDR1/NM_013994	GO:0007155 GO:0004713	HMMPfam_F5_F8_type_C HMMPfam_Pkinase_Tyr	superfamily_Galactose-binding domain-like superfamily_Protein kinase-like (PK-like)	BCM
DDR2/ENST00000367922	GO:0007155 GO:0004713	HMMPfam_F5_F8_type_C HMMPfam_Pkinase_Tyr	superfamily_Galactose-binding domain-like superfamily_Protein kinase-like (PK-like)	Broad
DHH/ENST00000266991	GO:0007267 GO:0006508	HMMPfam_HH_signal HMMPfam_Hint	superfamily_Hedgehog/DD-peptidase superfamily_Hedgehog/intein (Hint) domain	BCM
DIRAS3/ENST00000370981		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	BCM
DKFZp434N035/NM_032262				BCM
DMPK/ENST00000342805	GO:0004672 GO:0004674	HMMPfam_DMPK_coil HMMPfam_Pkinase HMMPfam_Pkinase_C	superfamily_Protein kinase-like (PK-like)	WU
DOCK3/ENST00000266037	GO:0005085	HMMPfam_Ded_cyto HMMPfam_SH3_2	superfamily_ARM repeat superfamily_SH3-domain	WU
DUSP6/ENST00000279488	GO:0006470	HMMPfam_DSPc HMMPfam_Rhodanese	superfamily_(Phosphotyrosine protein) phosphatases superfamily_Rhodanese/Cell cycle control phosphata	WU

DUSP24/ENST00000341238	GO:0006470	HMMPfam_DSPc	superfamily_(Phosphotyrosine protein) phosphatases superfamily_Rhodanese/Cell cycle control phosphata	WU
E2F1/ENST00000343380	GO:0000074	HMMPfam_E2F_TDP		Broad
E2F2/ENST00000361729	GO:0000074	HMMPfam_E2F_TDP		Broad
E2F3/ENST00000346618	GO:0000074	HMMPfam_E2F_TDP		BCM
E2F4/ENST00000388769	GO:0000074	HMMPfam_E2F_TDP	superfamily_Chromo domain-like	BCM
E2F5/ENST00000256117	GO:0000074	HMMPfam_E2F_TDP		Broad
E2F6/ENST00000307236	GO:0000074	HMMPfam_E2F_TDP		BCM
EGFR/ENST00000275493	GO:0016020 GO:0004713 GO:0004714	HMMPfam_Furin-like HMMPfam_Pkinase_Tyr HMMPfam_Recep_L_domain	superfamily_Growth factor receptor domain superfamily_L domain-like superfamily_Protein kinase-like (PK-like)	SHARE
EIF4E/ENST00000280892	GO:0003723	HMMPfam_IF4E	superfamily_Translation initiation factor eIF4e	BCM
EIF4EBP1/ENST00000338825	GO:0008190	HMMPfam_eIF_4EBP		WU
EIF4G1/ENST00000382330	GO:0003723 GO:0003743	HMMPfam_MA3 HMMPfam_MIF4G HMMPfam_W2	superfamily_ARM repeat	BCM
EPHA1/ENST00000275815	GO:0005003 GO:0004713	HMMPfam_Ephrin_lbd HMMPfam_Pkinase_Tyr HMMPfam_SAM_1 HMMPfam_fn3	superfamily_Fibronectin type III superfamily_Galactose-binding domain-like superfamily_Growth factor receptor domain superfamily_Protein kinase-like (PK-like) superfamily_SAM/Pointed domain	WU
EPHA2/ENST00000375634	GO:0005003 GO:0004713	HMMPfam_Ephrin_lbd HMMPfam_Pkinase_Tyr HMMPfam_SAM_1 HMMPfam_fn3	superfamily_Fibronectin type III superfamily_Galactose-binding domain-like superfamily_Growth factor receptor domain superfamily_Protein kinase-like (PK-like) superfamily_SAM/Pointed domain	BCM
EPHA3/ENST00000336596	GO:0005003 GO:0004713	HMMPfam_Ephrin_lbd HMMPfam_Pkinase_Tyr HMMPfam_SAM_2 HMMPfam_fn3	superfamily_Fibronectin type III superfamily_Galactose-binding domain-like superfamily_Growth factor receptor domain superfamily_Protein kinase-like (PK-like) superfamily_SAM/Pointed domain	Broad
EPHA4/ENST00000281821	GO:0005003 GO:0004713	HMMPfam_Ephrin_lbd HMMPfam_Pkinase_Tyr	superfamily_Fibronectin type III superfamily_Galactose-binding domain-like	WU

		HMMPfam_SAM_2	superfamily_Growth factor receptor domain	
		HMMPfam_fn3	superfamily_Protein kinase-like (PK-like)	
			superfamily_SAM/Pointed domain	
EPHA5/ENST00000273854	GO:0005003	HMMPfam_BmKX	superfamily_Fibronectin type III	Broad
	GO:0004713	HMMPfam_Ephrin_lbd	superfamily_Galactose-binding domain-like	
		HMMPfam_Pkinase_Tyr	superfamily_Growth factor receptor domain	
		HMMPfam_SAM_1	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_fn3	superfamily_SAM/Pointed domain	
EPHA6/ENST00000389672	GO:0005003	HMMPfam_Ephrin_lbd	superfamily_Fibronectin type III	Broad
	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Galactose-binding domain-like	
		HMMPfam_SAM_1	superfamily_Growth factor receptor domain	
		HMMPfam_fn3	superfamily_Protein kinase-like (PK-like)	
			superfamily_SAM/Pointed domain	
EPHA7/ENST00000369303	GO:0005003	HMMPfam_Ephrin_lbd	superfamily_Fibronectin type III	BCM
	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Galactose-binding domain-like	
		HMMPfam_SAM_1	superfamily_Growth factor receptor domain	
		HMMPfam_fn3	superfamily_Protein kinase-like (PK-like)	
			superfamily_SAM/Pointed domain	
EPHA8/ENST00000166244	GO:0005003	HMMPfam_Ephrin_lbd	superfamily_Fibronectin type III	Broad
	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Galactose-binding domain-like	
		HMMPfam_SAM_1	superfamily_Growth factor receptor domain	
		HMMPfam_fn3	superfamily_Protein kinase-like (PK-like)	
			superfamily_SAM/Pointed domain	
EPHB1/OTTHUMT00000266419	GO:0005003	HMMPfam_Ephrin_lbd	superfamily_Fibronectin type III	Broad
	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Galactose-binding domain-like	
		HMMPfam_SAM_1	superfamily_Growth factor receptor domain	
		HMMPfam_fn3	superfamily_Protein kinase-like (PK-like)	
			superfamily_SAM/Pointed domain	
EPHB2/ENST00000374630	GO:0005003	HMMPfam_Ephrin_lbd	superfamily_Fibronectin type III	Broad
	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Galactose-binding domain-like	
		HMMPfam_SAM_1	superfamily_Growth factor receptor domain	
		HMMPfam_fn3	superfamily_Protein kinase-like (PK-like)	
			superfamily_SAM/Pointed domain	

EPHB3/ENST00000330394	GO:0005003 GO:0004713	HMMPfam_Ephrin_lbd HMMPfam_Pkinase_Tyr HMMPfam_SAM_1 HMMPfam_fn3	superfamily_Fibronectin type III superfamily_Galactose-binding domain-like superfamily_Growth factor receptor domain superfamily_Protein kinase-like (PK-like) superfamily_SAM/Pointed domain	BCM
EPHB4/ENST00000358173	GO:0005003 GO:0004713	HMMPfam_Ephrin_lbd HMMPfam_Pkinase_Tyr HMMPfam_SAM_1 HMMPfam_fn3	superfamily_Fibronectin type III superfamily_Galactose-binding domain-like superfamily_Growth factor receptor domain superfamily_Protein kinase-like (PK-like) superfamily_SAM/Pointed domain	BCM
EPHB6/ENST00000222847	GO:0005003 GO:0004713	HMMPfam_Ephrin_lbd HMMPfam_Pkinase_Tyr HMMPfam_SAM_2 HMMPfam_fn3	superfamily_Fibronectin type III superfamily_Galactose-binding domain-like superfamily_Protein kinase-like (PK-like) superfamily_SAM/Pointed domain superfamily_TNF receptor-like	WU
EPHX1/ENST00000272167	GO:0004301	HMMPfam_Abhydrolase_1 HMMPfam_EHN	superfamily_alpha/beta-Hydrolases	WU
EPM2A/ENST00000367519	GO:0030246 GO:0006470	HMMPfam_DSPc	superfamily_(Phosphotyrosine protein) phosphatases superfamily_Starch-binding domain-like	BCM
ERAS/ENST00000338270		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	SHARE
ERBB2/ENST00000269571	GO:0016020 GO:0004713 GO:0004714	HMMPfam_Furin-like HMMPfam_Pkinase_Tyr HMMPfam_Recep_L_domain HMMPfam_YLP	superfamily_Growth factor receptor domain superfamily_L domain-like superfamily_Protein kinase-like (PK-like)	SHARE
ERBB3/ENST00000267101	GO:0016020 GO:0004713 GO:0004714	HMMPfam_Furin-like HMMPfam_Pkinase_Tyr HMMPfam_Recep_L_domain	superfamily_Growth factor receptor domain superfamily_L domain-like superfamily_Protein kinase-like (PK-like)	SHARE
ERBB4/ENST00000342788	GO:0016020 GO:0004713 GO:0004714	HMMPfam_Furin-like HMMPfam_Pkinase_Tyr HMMPfam_Recep_L_domain HMMPfam_YLP	superfamily_Growth factor receptor domain superfamily_L domain-like superfamily_Protein kinase-like (PK-like)	SHARE
ERCC2/ENST00000221481	GO:0004003 GO:0003677	HMMPfam_DEAD_2 HMMPfam_DUF1227	superfamily_P-loop containing nucleoside triphosph	WU

ERG/ENST00000288319	GO:0005634 GO:0003700	HMMPfam_Ets HMMPfam_SAM_PNT	superfamily_SAM/Pointed domain	WU
ETS1/ENST00000319397	GO:0005634 GO:0003700	HMMPfam_Ets HMMPfam_SAM_PNT	superfamily_SAM/Pointed domain	BCM
ETS2/ENST00000360938	GO:0005634 GO:0003700	HMMPfam_Ets HMMPfam_SAM_PNT	superfamily_SAM/Pointed domain	WU
ETV6/ENST00000266427	GO:0005634 GO:0003700	HMMPfam_Ets HMMPfam_SAM_PNT	superfamily_SAM/Pointed domain	Broad
ETV7/ENST00000340181	GO:0005634 GO:0003700	HMMPfam_Ets HMMPfam_SAM_PNT	superfamily_SAM/Pointed domain	BCM
EVI1/ENST00000264674	GO:0003676	HMMPfam_zf-C2H2	superfamily_C2H2 and C2HC zinc fingers	WU
FAS/ENST00000371875	GO:0005515 GO:0004872	HMMPfam_Death HMMPfam_TNFR_c6	superfamily_DEATH domain superfamily_TNF receptor-like	Broad
FBXW7/ENST00000281708		HMMPfam_F-box HMMPfam_WD40	superfamily_F-box domain superfamily_WD40 repeat-like	BCM
FER/ENST00000379734	GO:0007242 GO:0004713	HMMPfam_FCH HMMPfam_Pkinase_Tyr HMMPfam_SH2	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain	Broad
FES/ENST00000328850	GO:0007242 GO:0004713	HMMPfam_FCH HMMPfam_Pkinase_Tyr HMMPfam_SH2	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain	WU
FEV/ENST00000295727	GO:0003700	HMMPfam_Ets		WU
FGF3/ENST00000334134	GO:0008083	HMMPfam_FGF	superfamily_Cytokine	BCM
FGF4/ENST00000168712	GO:0008083	HMMPfam_FGF	superfamily_Cytokine	BCM
FGF5/ENST00000312465	GO:0008083	HMMPfam_FGF	superfamily_Cytokine	WU
FGFR1/ENST00000310729	GO:0004713	HMMPfam_I-set HMMPfam_Pkinase_Tyr HMMPfam_ig	superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	SHARE
FGFR2/ENST00000351936	GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_ig	superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	SHARE
FGFR3/ENST00000360915	GO:0004713	HMMPfam_I-set HMMPfam_Pkinase_Tyr HMMPfam_ig	superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	SHARE

FGFR4/OTTHUMT00000253412	GO:0005634	HMMPfam_PHD	superfamily_FYVE/PHD zinc finger	SHARE
	GO:0005515	HMMPfam_PWWP	superfamily_SET domain	
	GO:0008270	HMMPfam_SET	superfamily_Tudor/PWWP/MBT	
FGR/ENST00000374005	GO:0007242	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	SHARE
	GO:0004713	HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
FH/ENST00000366560	GO:0003824	HMMPfam_Lyase_1	superfamily_L-aspartase-like	WU
FHIT/ENST00000341848		HMMPfam_HIT	superfamily_HIT-like	BCM
FKBP1A/ENST00000381742	GO:0006457	HMMPfam_FKBP_C	superfamily_FKBP-like	BCM
FLJ25006/ENST00000301037	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
FLJ30473/ENST00000335375	GO:0006118	HMMPfam_Pyr_redox	superfamily_FAD/NAD(P)-binding domain	BCM
	GO:0016491	HMMPfam_Pyr_redox_2	superfamily_FAD/NAD-linked reductases dimerisation	
		HMMPfam_Rieske	superfamily_ISP domain	
FLT1/ENST00000282397	GO:0004713	HMMPfam_I-set	superfamily_Immunoglobulin	SHARE
		HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_ig		
FLT3/ENST00000380982	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Immunoglobulin	BCM
		HMMPfam_ig	superfamily_Protein kinase-like (PK-like)	
FLT4/ENST00000261937	GO:0004713	HMMPfam_I-set	superfamily_Immunoglobulin	SHARE
		HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_ig		
FNDC3A/ENST00000337156		HMMPfam_fn3	superfamily_Fibronectin type III	Broad
FOS/OTTHUMT00000072944	GO:0003677	HMMPfam_bZIP_2	superfamily_A DNA-binding domain in eukaryotic tra	WU
FOSB/ENST00000353609	GO:0003677	HMMPfam_bZIP_2	superfamily_A DNA-binding domain in eukaryotic tra	WU
FOSL2/ENST00000379616	GO:0003677	HMMPfam_bZIP_2	superfamily_A DNA-binding domain in eukaryotic tra	Broad
FOXO1A/ENST00000379561	GO:0003700	HMMPfam_Fork_head		BCM
FOXO3A/ENST00000343882	GO:0003700	HMMPfam_Fork_head		Broad
FOXP1/ENST00000318789	GO:0003700	HMMPfam_Fork_head	superfamily_C2H2 and C2HC zinc fingers	BCM
FRK/ENST00000368626	GO:0007242	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	Broad
	GO:0004713	HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
FYN/ENST00000368678	GO:0007242	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	SHARE
	GO:0004713	HMMPfam_SH2	superfamily_SH2 domain	

		HMMPfam_SH3_1	superfamily_SH3-domain	
GAK/ENST00000314167	GO:0031072 GO:0004672	HMMPfam_DnaJ HMMPfam_Pkinase	superfamily_(Phosphotyrosine protein) phosphatases superfamily_C2 domain (Calcium/lipid-binding domai superfamily_Chaperone J-domain superfamily_Protein kinase-like (PK-like)	Broad
GATA1/ENST00000376670	GO:0003700	HMMPfam_GATA	superfamily_Glucocorticoid receptor-like (DNA-bind	Broad
GBL/NM_022372		HMMPfam_WD40	superfamily_WD40 repeat-like	WU
GCK/ENST00000345378	GO:0004396	HMMPfam_Hexokinase_1 HMMPfam_Hexokinase_2	superfamily_Actin-like ATPase domain	BCM
GGT2/ENST00000249062	GO:0003840	HMMPfam_G_glu_transpept		BCM
GNAI2/ENST00000313601	GO:0004871	HMMPfam_G-alpha	superfamily_P-loop containing nucleoside triphosph superfamily_Transducin (alpha subunit) insertion d	WU
GNAS/ENST00000306120	GO:0004871	HMMPfam_G-alpha	superfamily_P-loop containing nucleoside triphosph superfamily_Transducin (alpha subunit) insertion d	BCM
GNAT1/ENST00000232461	GO:0004871	HMMPfam_G-alpha	superfamily_P-loop containing nucleoside triphosph superfamily_Transducin (alpha subunit) insertion d	WU
GPC3/ENST00000370818	GO:0005578	HMMPfam_Glypican		WU
GRB10/NM_005311	GO:0007242 GO:0007165	HMMPfam_BPS HMMPfam_PH HMMPfam_RA HMMPfam_SH2	superfamily_PH domain-like superfamily_SH2 domain	SHARE
GRB14/ENST00000263915	GO:0007242 GO:0007165	HMMPfam_BPS HMMPfam_PH HMMPfam_RA HMMPfam_SH2	superfamily_PH domain-like superfamily_SH2 domain	SHARE
GRB2/ENST00000316804	GO:0007242	HMMPfam_SH2 HMMPfam_SH3_1	superfamily_SH2 domain superfamily_SH3-domain	SHARE
GRB7/ENST00000309156	GO:0007242 GO:0007165	HMMPfam_BPS HMMPfam_PH HMMPfam_RA HMMPfam_SH2	superfamily_PH domain-like superfamily_SH2 domain	SHARE
GRK1/ENST00000335678	GO:0004672 GO:0004871	HMMPfam_Pkinase HMMPfam_RGS	superfamily_Protein kinase-like (PK-like) superfamily_Regulator of G-protein signaling RGS	Broad

GRK4/ENST00000264763	GO:0004672 GO:0004871	HMMPfam_Pkinase HMMPfam_RGS	superfamily_Protein kinase-like (PK-like) superfamily_Regulator of G-protein signaling RGS	Broad
GRK5/ENST00000369108	GO:0004672 GO:0004871	HMMPfam_Pkinase HMMPfam_RGS	superfamily_Protein kinase-like (PK-like) superfamily_Regulator of G-protein signaling RGS	WU
GRK6/ENST00000355472	GO:0004672 GO:0004871	HMMPfam_Pkinase HMMPfam_RGS	superfamily_Protein kinase-like (PK-like) superfamily_Regulator of G-protein signaling RGS	Broad
GRK7/ENST00000264952	GO:0004672 GO:0004871	HMMPfam_Pkinase HMMPfam_RGS	superfamily_Protein kinase-like (PK-like) superfamily_Regulator of G-protein signaling RGS	Broad
GSK3A/ENST00000222330	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
GSK3B/ENST00000316626	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
HCK/ENST00000262651	GO:0007242 GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_SH2 HMMPfam_SH3_1	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain superfamily_SH3-domain	SHARE
HD/ENST00000355072		HMMPfam_HEAT	superfamily_ARM repeat	Broad
HEMK1/ENST00000232854	GO:0008168	HMMPfam_MTS	superfamily_S-adenosyl-L-methionine-dependent meth	WU
HHIP/ENST00000296575		HMMPfam_EGF_2	superfamily_EGF/Laminin superfamily_Soluble quinoprotein glucose dehydroge	WU
HIC2/ENST00000302897	GO:0003676	HMMPfam_BTB HMMPfam_zf-C2H2	superfamily_C2H2 and C2HC zinc fingers superfamily_POZ domain	BCM
HIF1A/OTTHUMT00000276979		HMMPfam_HIF-1a_CTAD HMMPfam_PAS_3	superfamily_HLH helix-loop-helix DNA-binding domai superfamily_PYP-like sensor domain (PAS domain)	Broad
HRAS/ENST00000388730		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	SHARE
HRPT2/ENST00000367436		HMMPfam_CDC73	superfamily_Cell growth inhibitor/plasmid maintena	BCM
HYAL1/ENST00000266031	GO:0004415	HMMPfam_Glyco_hydro_56	superfamily_(Trans)glycosidases	WU
HYAL2/ENST00000357750	GO:0004415	HMMPfam_Glyco_hydro_56	superfamily_(Trans)glycosidases superfamily_EGF/Laminin	WU
HYAL3/OTTHUMT00000264836		HMMPfam_IFRD HMMPfam_IFRD_C	superfamily_ARM repeat	WU
IFRD2/ENST00000336089		HMMPfam_IFRD HMMPfam_IFRD_C	superfamily_ARM repeat	WU
IGF1R/ENST00000268035	GO:0016020 GO:0004713 GO:0004714	HMMPfam_Furin-like HMMPfam_Pkinase_Tyr HMMPfam_Recep_L_domain	superfamily_Fibronectin type III superfamily_Growth factor receptor domain superfamily_L domain-like	WU

		HMMPfam_fn3	superfamily_Protein kinase-like (PK-like)	
IHH/ENST00000295731	GO:0007267	HMMPfam_HH_signal	superfamily_Hedgehog/DD-peptidase	WU
	GO:0006508	HMMPfam_Hint	superfamily_Hedgehog/intein (Hint) domain	
IKBKB/ENST00000342222	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
	GO:0006464	HMMPfam_ubiquitin	superfamily_Ubiquitin-like	
IKBKE/ENST00000367120	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
ILK/ENST00000299421	GO:0004713	HMMPfam_Ank	superfamily_Ankyrin repeat	BCM
		HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	
INHBA/ENST00000242208	GO:0008083	HMMPfam_TGF_beta	superfamily_Cystine-knot cytokines	BCM
		HMMPfam_TGFb_propeptide		
INSR/ENST00000302850	GO:0016020	HMMPfam_Furin-like	superfamily_Fibronectin type III	SHARE
	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Growth factor receptor domain	
	GO:0004714	HMMPfam_Recep_L_domain	superfamily_L domain-like	
		HMMPfam_fn3	superfamily_Protein kinase-like (PK-like)	
INSRR/ENST00000368195	GO:0016020	HMMPfam_Furin-like	superfamily_Fibronectin type III	SHARE
	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Growth factor receptor domain	
	GO:0004714	HMMPfam_Recep_L_domain	superfamily_L domain-like	
		HMMPfam_fn3	superfamily_Protein kinase-like (PK-like)	
IRAK1/ENST00000369980	GO:0005515	HMMPfam_Death	superfamily_DEATH domain	BCM
	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
IRAK2/ENST00000256458	GO:0005515	HMMPfam_Death	superfamily_DEATH domain	BCM
	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
IRAK3/ENST00000261233	GO:0005515	HMMPfam_Death	superfamily_DEATH domain	Broad
	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
IRAK4/ENST00000356669	GO:0004672	HMMPfam_Pkinase	superfamily_DEATH domain	BCM
			superfamily_Protein kinase-like (PK-like)	
IRS1/ENST00000305123	GO:0005158	HMMPfam_IRS	superfamily_PH domain-like	WU
		HMMPfam_PH		
IRS2/ENST00000375856	GO:0005158	HMMPfam_IRS	superfamily_PH domain-like	Broad
		HMMPfam_PH		
ITK/ENST00000231189	GO:0007242	HMMPfam_BTK	superfamily_PH domain-like	BCM
	GO:0004713	HMMPfam_PH	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_Pkinase_Tyr	superfamily_SH2 domain	

		HMMPfam_SH2 HMMPfam_SH3_1	superfamily_SH3-domain	
JAG1/ENST00000254958	GO:0007219 GO:0007154	HMMPfam_DSL HMMPfam_EGF HMMPfam_MNNL	superfamily_EGF/Laminin superfamily_PMP inhibitors	BCM
JAG2/OTTHUMT00000074540	GO:0007219 GO:0007154	HMMPfam_DSL HMMPfam_EGF HMMPfam_EGF_CA HMMPfam_MNNL	superfamily_Concanavalin A-like lectins/glucanases superfamily_EGF/Laminin superfamily_PMP inhibitors	BCM
JAK1/ENST00000294423	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain	SHARE
JAK2/ENST00000381652	GO:0007242 GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_SH2	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain	SHARE
JAK3/ENST00000222246	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain	SHARE
JUN/ENST00000371222	GO:0003700	HMMPfam_Jun HMMPfam_bZIP_1	superfamily_A DNA-binding domain in eukaryotic tra	WU
JUNB/ENST00000302754	GO:0003700	HMMPfam_Jun HMMPfam_bZIP_1	superfamily_A DNA-binding domain in eukaryotic tra	BCM
JUND/ENST00000252818	GO:0003700	HMMPfam_Jun HMMPfam_bZIP_1	superfamily_A DNA-binding domain in eukaryotic tra	BCM
JUP/ENST00000310706		HMMPfam_Arm HMMPfam_HEAT	superfamily_ARM repeat	BCM
KDR/ENST00000381338	GO:0004713	HMMPfam_I-set HMMPfam_Pkinase_Tyr HMMPfam_V-set	superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	SHARE
KELCHL/ENST00000328879		HMMPfam_BACK HMMPfam_BTB HMMPfam_Kelch_1	superfamily_Galactose oxidase central domain superfamily_POZ domain	BCM
KIAA1804/NM_032435	GO:0004672	HMMPfam_Pkinase HMMPfam_SH3_1	superfamily_Protein kinase-like (PK-like) superfamily_SH3-domain	WU
KIT/ENST00000381345	GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_ig	superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	SHARE

KLF6/ENST00000380960	GO:0003676	HMMPfam_zf-C2H2	superfamily_C2H2 and C2HC zinc fingers	WU
KRAS/ENST00000256078		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	SHARE
KSR1/ENST00000319524	GO:0007242	HMMPfam_C1_1	superfamily_Cysteine-rich domain	WU
	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	
KSR2/ENST00000379510	GO:0005737	HMMPfam_DUF1170	superfamily_PDZ domain-like	Broad
	GO:0005515	HMMPfam_PDZ	superfamily_PH domain-like	
		HMMPfam_PH	superfamily_SAM/Pointed domain	
		HMMPfam_SAM_1		
LATS1/ENST00000253339	GO:0004672	HMMPfam_Pkinase	superfamily_N-terminal domain of adenylylcyclase a	BCM
	GO:0004674	HMMPfam_Pkinase_C	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_UBA	superfamily_UBA-like	
LATS2/ENST00000382592	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
	GO:0004674	HMMPfam_Pkinase_C	superfamily_UBA-like	
		HMMPfam_UBA		
LCK/ENST00000333070	GO:0007242	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	SHARE
	GO:0004713	HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
LIMD1/ENST00000273317	GO:0008270	HMMPfam_LIM	superfamily_Glucocorticoid receptor-like (DNA-bind	WU
LIMK1/ENST00000336180	GO:0005515	HMMPfam_LIM	superfamily_Glucocorticoid receptor-like (DNA-bind	WU
	GO:0004672	HMMPfam_PDZ	superfamily_PDZ domain-like	
	GO:0008270	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
LIMK2/ENST00000340552	GO:0005515	HMMPfam_LIM	superfamily_Glucocorticoid receptor-like (DNA-bind	Broad
	GO:0004672	HMMPfam_PDZ	superfamily_Myosin phosphatase inhibitor 17kDa pro	
	GO:0008270	HMMPfam_PP1_inhibitor	superfamily_PDZ domain-like	
		HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
LMTK2/ENST00000297293	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	BCM
LMTK3/ENST00000270238	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	WU
LOC150786/ENST00000320501		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	BCM
LOC375133/NM_199283	GO:0016773	HMMPfam_PI3_PI4_kinase	superfamily_Protein kinase-like (PK-like)	BCM
LOC400891/ENST00000342608			superfamily_RNI-like	BCM
LRP1B/ENST00000305375	GO:0016020	HMMPfam_EGF	superfamily_EGF/Laminin	Broad
		HMMPfam_EGF_2	superfamily_LDL receptor-like module	
		HMMPfam_EGF_CA	superfamily_YWTD domain	

		HMMPfam_Ldl_recept_a HMMPfam_Ldl_recept_b HMMPfam_NHL		
LRRK1/ENST00000388948	GO:0005525 GO:0005515 GO:0004672	HMMPfam_Ank HMMPfam_LRR_1 HMMPfam_Miro HMMPfam_Pkinase	superfamily_Ankyrin repeat superfamily_L domain-like superfamily_P-loop containing nucleoside triphosph superfamily_Protein kinase-like (PK-like) superfamily_YVTN repeat-like/Quinoprotein amine de	WU
LRRK2/ENST00000388935	GO:0005525 GO:0005515 GO:0004672	HMMPfam_LRR_1 HMMPfam_LRR_2 HMMPfam_Miro HMMPfam_Pkinase	superfamily_ARM repeat superfamily_Ankyrin repeat superfamily_L domain-like superfamily_P-loop containing nucleoside triphosph superfamily_Protein kinase-like (PK-like) superfamily_WD40 repeat-like	BCM
LTF/ENST00000231751	GO:0005576	HMMPfam_Transferrin	superfamily_Periplasmic binding protein-like II	WU
LTK/ENST00000263800	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_EGF/Laminin superfamily_Protein kinase-like (PK-like)	WU
LYK5/ENST00000336174	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
LYN/ENST00000359619	GO:0007242 GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_SH2 HMMPfam_SH3_1	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain superfamily_SH3-domain	SHARE
LZTR1/ENST00000215739		HMMPfam_BTB HMMPfam_Kelch_1 HMMPfam_Kelch_2	superfamily_Galactose oxidase central domain superfamily_POZ domain	BCM
MAF/ENST00000326043	GO:0003677	HMMPfam_Maf_N HMMPfam_bZIP_Maf	superfamily_A DNA-binding domain in eukaryotic tra	BCM
MAFA/ENST00000333480	GO:0003677	HMMPfam_Maf_N HMMPfam_bZIP_Maf	superfamily_A DNA-binding domain in eukaryotic tra	Broad
MAFB/ENST00000373313	GO:0003677	HMMPfam_Maf_N HMMPfam_bZIP_Maf	superfamily_A DNA-binding domain in eukaryotic tra	WU
MAFF/ENST00000338483	GO:0003677	HMMPfam_bZIP_Maf	superfamily_A DNA-binding domain in eukaryotic tra	Broad
MAFG/ENST00000357736	GO:0003677	HMMPfam_bZIP_Maf	superfamily_A DNA-binding domain in eukaryotic tra	Broad
MAFK/ENST00000343242	GO:0003677	HMMPfam_bZIP_Maf	superfamily_A DNA-binding domain in eukaryotic tra	Broad

MAML1/NM_014757		HMMPfam_MAPEG		Broad
MAML2/NM_032427	GO:0007219 GO:0005634 GO:0045944 GO:0006355 GO:0006350 GO:0003713			WU
MAML3/ENST00000327122	GO:0003677 GO:0007219 GO:0004879 GO:0007399 GO:0005634 GO:0045944 GO:0006355 GO:0006350 GO:0003713 GO:0003700			WU
MAP2K1/ENST00000307102	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MAP2K2/ENST00000262948	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MAP2K4/ENST00000353533	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MAP2K5/ENST00000354498	GO:0004672	HMMPfam_PB1 HMMPfam_Pkinase	superfamily_CAD PB1 domains superfamily_Protein kinase-like (PK-like)	Broad
MAP2K6/ENST00000359094	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MAP2K7/OTTHUMT00000267981	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MAP3K1/OTTHUMT00000132309	GO:0004672 GO:0008270	HMMPfam_Pkinase HMMPfam_SWIM	superfamily_ARM repeat superfamily_Protein kinase-like (PK-like) superfamily_RING/U-box	Broad
MAP3K10/ENST00000253055	GO:0004672	HMMPfam_Pkinase HMMPfam_SH3_1	superfamily_Protein kinase-like (PK-like) superfamily_SH3-domain	WU
MAP3K11/ENST00000309100	GO:0004672	HMMPfam_Pkinase HMMPfam_SH3_1	superfamily_Protein kinase-like (PK-like) superfamily_SH3-domain	BCM
MAP3K12/OTTHUMT00000261165	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MAP3K13/ENST00000265026	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM

MAP3K14/OTTHUMT00000283229	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MAP3K15/ENST00000338883	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MAP3K2/ENST00000344908	GO:0004672	HMMPfam_PB1 HMMPfam_Pkinase	superfamily_CAD PB1 domains superfamily_Protein kinase-like (PK-like)	BCM
MAP3K3/ENST00000361357	GO:0004672	HMMPfam_PB1 HMMPfam_Pkinase	superfamily_CAD PB1 domains superfamily_Protein kinase-like (PK-like)	Broad
MAP3K4/ENST00000265125	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MAP3K5/ENST00000359015	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_SAM/Pointed domain	WU
MAP3K6/ENST00000357582	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_SAM/Pointed domain superfamily_TPR-like	Broad
MAP3K7/ENST00000369329	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MAP3K8/ENST00000375328	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MAP3K9/OTTHUMT00000072328	GO:0004672	HMMPfam_Pkinase HMMPfam_SH3_1	superfamily_Protein kinase-like (PK-like) superfamily_Regulator of G-protein signaling RGS superfamily_SH3-domain	WU
MAP4K1/ENST00000221409	GO:0004672 GO:0005083	HMMPfam_CNH HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MAP4K3/ENST00000378758	GO:0004672 GO:0005083	HMMPfam_CNH HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MAP4K4/ENST00000324219	GO:0004672 GO:0005083	HMMPfam_CNH HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MAP4K5/NM_198794	GO:0004672 GO:0005083	HMMPfam_CNH HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MAPK1/ENST00000215832	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MAPK10/ENST00000359221	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MAPK11/ENST00000330651	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MAPK12/ENST00000215659	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MAPK13/ENST00000358332	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MAPK14/ENST00000229794	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MAPK15/ENST00000322233	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad

MAPK3/ENST00000263025	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MAPK4/NM_002747	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MAPK6/ENST00000261845	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MAPK7/ENST00000299612	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MAPK8/ENST00000374189	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MAPK9/ENST00000347470	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MAPKAPK3/ENST00000357955	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MAST1/ENST00000251472	GO:0005515	HMMPfam_DUF1908	superfamily_PDZ domain-like	BCM
	GO:0004672	HMMPfam_PDZ	superfamily_Protein kinase-like (PK-like)	
	GO:0004674	HMMPfam_Pkinase		
		HMMPfam_Pkinase_C		
MAST2/ENST00000361297	GO:0005515	HMMPfam_DUF1908	superfamily_PDZ domain-like	WU
	GO:0004672	HMMPfam_PDZ	superfamily_Protein kinase-like (PK-like)	
	GO:0004674	HMMPfam_Pkinase		
		HMMPfam_Pkinase_C		
MAST3/ENST00000262811	GO:0005515	HMMPfam_DUF1908	superfamily_PDZ domain-like	BCM
	GO:0004672	HMMPfam_PDZ	superfamily_Protein kinase-like (PK-like)	
	GO:0004674	HMMPfam_Pkinase		
		HMMPfam_Pkinase_C		
MAST4/ENST00000261569	GO:0005515	HMMPfam_DUF1908	superfamily_PDZ domain-like	WU
	GO:0004672	HMMPfam_PDZ	superfamily_Protein kinase-like (PK-like)	
	GO:0004674	HMMPfam_Pkinase		
		HMMPfam_Pkinase_C		
MASTL/ENST00000375940	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
MATK/ENST00000315548	GO:0007242	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	BCM
	GO:0004713	HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
MDM2/OTTHUMT00000261845	GO:0005622	HMMPfam_SWIB	superfamily_RING/U-box	Broad
	GO:0005634	HMMPfam_zf-RanBP	superfamily_SWIB/MDM2 domain	
			superfamily_Znf265 first zinc-finger domain	
MDM4/ENST00000367183	GO:0005634	HMMPfam_SWIB	superfamily_RING/U-box	Broad
			superfamily_SWIB/MDM2 domain	
			superfamily_Znf265 first zinc-finger domain	

MEN1/ENST00000337652	GO:0005634	HMMPfam_Menin		BCM
MERTK/ENST00000295408	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Fibronectin type III	WU
		HMMPfam_fn3	superfamily_Immunoglobulin	
		HMMPfam_ig	superfamily_Protein kinase-like (PK-like)	
			superfamily_Rhodanese/Cell cycle control phosphata	
MET/ENST00000318493	GO:0004713	HMMPfam_PSI	superfamily_E set domains	SHARE
	GO:0004872	HMMPfam_Pkinase_Tyr	superfamily_Plexin repeat	
		HMMPfam_Sema	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_TIG	superfamily_Sema domain	
MGC16703/ENST00000292748		HMMPfam_Tubulin	superfamily_RNI-like	BCM
			superfamily_Tubulin nucleotide-binding domain-like	
MINK1/ENST00000355280	GO:0004672	HMMPfam_CNH	superfamily_Protein kinase-like (PK-like)	Broad
	GO:0005083	HMMPfam_Pkinase		
MKNK1/ENST00000371946	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MKNK2/ENST00000250896	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MLH1/ENST00000231790	GO:0005524	HMMPfam_DNA_mis_repair	superfamily_ATPase domain of HSP90 chaperone/DNA t	WU
		HMMPfam_HATPase_c	superfamily_Ribosomal protein S5 domain 2-like	
MLKL/ENST00000389390	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MLL/ENST00000359313	GO:0003677	HMMPfam_AT_hook	superfamily_Bromodomain	BCM
	GO:0005634	HMMPfam_FYRC	superfamily_FYVE/PHD zinc finger	
	GO:0005515	HMMPfam_FYRN	superfamily_SET domain	
	GO:0008270	HMMPfam_PHD		
		HMMPfam_SET		
		HMMPfam_zf-CXXC		
MON1A/ENST00000296473	GO:0006810	HMMPfam_DUF254	superfamily_SNARE-like	WU
MOS/ENST00000311923	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
MPL/ENST00000372470		HMMPfam_EpoR_lig-bind	superfamily_Fibronectin type III	WU
		HMMPfam_fn3		
MRAS/ENST00000289104		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	Broad
MSH2/ENST00000233146	GO:0005524	HMMPfam_MutS_I	superfamily_DNA repair protein MutS domain II	WU
	GO:0003684	HMMPfam_MutS_II	superfamily_DNA repair protein MutS domain III	
		HMMPfam_MutS_III	superfamily_P-loop containing nucleoside triphosph	
		HMMPfam_MutS_IV		

		HMMPfam_MutS_V		
MSH6/ENST00000234420	GO:0005524	HMMPfam_MutS_I	superfamily_DNA repair protein MutS domain I	WU
	GO:0003684	HMMPfam_MutS_II	superfamily_DNA repair protein MutS domain II	
		HMMPfam_MutS_III	superfamily_DNA repair protein MutS domain III	
		HMMPfam_MutS_IV	superfamily_P-loop containing nucleoside triphosph	
		HMMPfam_MutS_V	superfamily_Tudor/PWWP/MBT	
		HMMPfam_PWWP		
MST1/ENST00000389186	GO:0004252	HMMPfam_Kringle	superfamily_Hairpin loop containing domain-like	WU
		HMMPfam_PAN_1	superfamily_Kringle-like	
		HMMPfam_Trypsin	superfamily_Trypsin-like serine proteases	
MST1R/ENST00000296474	GO:0004713	HMMPfam_PSI	superfamily_E set domains	WU
	GO:0004872	HMMPfam_Pkinase_Tyr	superfamily_Plexin repeat	
		HMMPfam_Sema	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_TIG	superfamily_Sema domain	
MST4/NM_016542	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
MUSK/ENST00000374448	GO:0004713	HMMPfam_Fz	superfamily_Frizzled cysteine-rich domain	Broad
		HMMPfam_I-set	superfamily_Immunoglobulin	
		HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	
MYB/ENST00000341911	GO:0030528	HMMPfam_Cmyb_C	superfamily_Homeodomain-like	WU
		HMMPfam_Myb_DNA-binding		
		HMMPfam_Wos2		
MYBL1/OTTHUMT00000274404	GO:0030528	HMMPfam_Cmyb_C	superfamily_Homeodomain-like	BCM
		HMMPfam_Myb_DNA-binding		
		HMMPfam_Wos2		
MYBL2/ENST00000217026		HMMPfam_Cmyb_C	superfamily_Homeodomain-like	WU
		HMMPfam_Myb_DNA-binding		
MYC/ENST00000377970	GO:0005634	HMMPfam_HLH	superfamily_HLH helix-loop-helix DNA-binding domai	BCM
	GO:0003700	HMMPfam_Myc-LZ		
		HMMPfam_Myc_N		
MYCL1/ENST00000334282	GO:0005634	HMMPfam_HLH	superfamily_HLH helix-loop-helix DNA-binding domai	WU
	GO:0003700	HMMPfam_Myc_N		
MYCN/ENST00000281043	GO:0005634	HMMPfam_HLH	superfamily_HLH helix-loop-helix DNA-binding domai	Broad
	GO:0003700	HMMPfam_Myc_N		

MYO3A/ENST00000265944	GO:0003774 GO:0004672	HMMPfam_IQ HMMPfam_Myosin_head HMMPfam_Pkinase	superfamily_P-loop containing nucleoside triphosph superfamily_Protein kinase-like (PK-like)	Broad
MYO3B/ENST00000263810	GO:0003774 GO:0004672	HMMPfam_IQ HMMPfam_Myosin_head HMMPfam_Pkinase	superfamily_P-loop containing nucleoside triphosph superfamily_Protein kinase-like (PK-like)	WU
MYT1/NM_004535	GO:0003700	HMMPfam_MYT1 HMMPfam_zf-C2HC	superfamily_CCHHC domain	BCM
NCK1/ENST00000288986	GO:0007242	HMMPfam_SH2 HMMPfam_SH3_1	superfamily_SH2 domain superfamily_SH3-domain	Broad
NCK2/ENST00000233154	GO:0007242	HMMPfam_SH2 HMMPfam_SH3_1	superfamily_SH2 domain superfamily_SH3-domain	WU
NCYM/ENST00000307384	GO:0007275			Broad
NF1/ENST00000358273	GO:0005096	HMMPfam_RasGAP	superfamily_ARM repeat superfamily_CRAL/TRIO domain superfamily_GTPase activation domain GAP	WU
NF2/ENST00000334961	GO:0005737 GO:0005856	HMMPfam_Band_41 HMMPfam_ERM	superfamily_Moesin tail domain superfamily_PH domain-like superfamily_Second domain of FERM superfamily_Ubiquitin-like	Broad
NFKB1/ENST00000226574	GO:0005515 GO:0003700	HMMPfam_Ank HMMPfam_Death HMMPfam_RHD HMMPfam_TIG	superfamily_Ankyrin repeat superfamily_DEATH domain superfamily_E set domains superfamily_p53-like transcription factors	BCM
NLK/OTTHUMT00000281110	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
NOTCH1/ENST00000277541	GO:0007219 GO:0016021 GO:0016020	HMMPfam_Ank HMMPfam_EGF HMMPfam_EGF_CA HMMPfam_NOD HMMPfam_NODP HMMPfam_Notch	superfamily_Ankyrin repeat superfamily_Concanavalin A-like lectins/glucanases superfamily_EGF/Laminin superfamily_Notch domain	WU
NOTCH2/ENST00000256646	GO:0007219 GO:0016021	HMMPfam_Ank HMMPfam_EGF	superfamily_Ankyrin repeat superfamily_EGF/Laminin	BCM

	GO:0016020	HMMPfam_EGF_CA HMMPfam_NOD HMMPfam_NODP HMMPfam_Notch	superfamily_Notch domain	
NOTCH2NL/NM_203458		HMMPfam_EGF HMMPfam_EGF_CA	superfamily_EGF/Laminin	Broad
NOTCH3/ENST00000263388	GO:0007219 GO:0016021 GO:0016020	HMMPfam_Ank HMMPfam_EGF HMMPfam_EGF_2 HMMPfam_EGF_CA HMMPfam_NOD HMMPfam_NODP HMMPfam_Notch	superfamily_Ankyrin repeat superfamily_EGF/Laminin superfamily_Notch domain	BCM
NOTCH4/ENST00000375023	GO:0007219 GO:0016020	HMMPfam_Ank HMMPfam_EGF HMMPfam_EGF_CA HMMPfam_NODP HMMPfam_Notch	superfamily_Ankyrin repeat superfamily_Concanavalin A-like lectins/glucanases superfamily_EGF/Laminin superfamily_Notch domain	BCM
NRAS/ENST00000369533	GO:0003677	HMMPfam_CSD	superfamily_Nucleic acid-binding proteins	SHARE
NRK/ENST00000243300	GO:0004672 GO:0005083	HMMPfam_CNH HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_p53-like transcription factors	Broad
NTRK1/ENST00000358660	GO:0005515 GO:0004713	HMMPfam_LRR_1 HMMPfam_Pkinase_Tyr	superfamily_Immunoglobulin superfamily_L domain-like superfamily_Protein kinase-like (PK-like)	SHARE
NTRK2/ENST00000376214	GO:0005515 GO:0004713	HMMPfam_I-set HMMPfam_LRRNT HMMPfam_LRR_1 HMMPfam_Pkinase_Tyr	superfamily_Immunoglobulin superfamily_L domain-like superfamily_Protein kinase-like (PK-like)	SHARE
NTRK3/ENST00000360948	GO:0005515 GO:0004713	HMMPfam_I-set HMMPfam_LRRNT HMMPfam_LRR_1 HMMPfam_Pkinase_Tyr HMMPfam_ig	superfamily_Immunoglobulin superfamily_L domain-like superfamily_Protein kinase-like (PK-like)	SHARE

OSR1/OTTHUMT00000264162	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
P2RXL1/OTTHUMT00000074974	GO:0004872	HMMPfam_P2X_receptor		BCM
P2RY5/ENST00000378434	GO:0001584	HMMPfam_7tm_1	superfamily_Family A G protein-coupled receptor-li	Broad
PAK1/ENST00000278568	GO:0004672	HMMPfam_PBD HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_Wiscott-Aldrich syndrome protein WASP	Broad
PAK2/ENST00000327134	GO:0004672	HMMPfam_PBD HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_Wiscott-Aldrich syndrome protein WASP	Broad
PAK3/ENST00000360648	GO:0004672	HMMPfam_PBD HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
PAK4/ENST00000358301	GO:0004672	HMMPfam_PBD HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
PAK6/ENST00000260404	GO:0004672	HMMPfam_PBD HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
PAK7/ENST00000378429	GO:0004672	HMMPfam_PBD HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
PASK/ENST00000358649	GO:0004672 GO:0006355	HMMPfam_PAS HMMPfam_Pkinase	superfamily_PYP-like sensor domain (PAS domain) superfamily_Protein kinase-like (PK-like)	BCM
PCQAP/ENST00000263205	GO:0005737 GO:0005634 GO:0045735 GO:0006355 GO:0005199 GO:0006350			BCM
PCTK1/ENST00000276052	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
PCTK2/ENST00000261211	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
PCTK3/ENST00000343606	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
PDGFB/ENST00000331163	GO:0008083	HMMPfam_PDGF HMMPfam_PDGF_N	superfamily_Cystine-knot cytokines	Broad
PDGFRA/ENST00000257290	GO:0004713	HMMPfam_I-set HMMPfam_Pkinase_Tyr HMMPfam_ig	superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	SHARE
PDGFRB/ENST00000261799	GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_ig	superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	SHARE

PDK1/ENST00000282077	GO:0005524	HMMPfam_HATPase_c	superfamily_ATPase domain of HSP90 chaperone/DNA t superfamily_alpha-ketoacid dehydrogenase kinase N-	WU
PEBP1/OTTHUMT00000262864		HMMPfam_PBP	superfamily_PEBP-like	Broad
PFTK1/ENST00000306129	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
PHOX2B/ENST00000381741	GO:0003700	HMMPfam_Homeobox	superfamily_Homeodomain-like	BCM
PIK3C2A/ENST00000265970	GO:0004428	HMMPfam_C2	superfamily_ARM repeat	SHARE
	GO:0005942	HMMPfam_PI3K_C2	superfamily_C2 domain (Calcium/lipid-binding domai	
	GO:0016773	HMMPfam_PI3K_rbd	superfamily_PX domain	
	GO:0005515	HMMPfam_PI3Ka	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_PI3_PI4_kinase	superfamily_Ubiquitin-like	
		HMMPfam_PX		
PIK3C2B/ENST00000367187	GO:0004428	HMMPfam_C2	superfamily_ARM repeat	Broad
	GO:0005942	HMMPfam_PI3K_C2	superfamily_C2 domain (Calcium/lipid-binding domai	
	GO:0016773	HMMPfam_PI3K_rbd	superfamily_PX domain	
	GO:0005515	HMMPfam_PI3Ka	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_PI3_PI4_kinase	superfamily_Ubiquitin-like	
		HMMPfam_PX		
PIK3C2G/ENST00000266497	GO:0004428	HMMPfam_C2	superfamily_ARM repeat	WU
	GO:0005942	HMMPfam_PI3K_C2	superfamily_C2 domain (Calcium/lipid-binding domai	
	GO:0016773	HMMPfam_PI3K_rbd	superfamily_PX domain	
	GO:0005515	HMMPfam_PI3Ka	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_PI3_PI4_kinase	superfamily_Ubiquitin-like	
		HMMPfam_PX		
PIK3C3/ENST00000262039	GO:0004428	HMMPfam_PI3K_C2	superfamily_ARM repeat	Broad
	GO:0005942	HMMPfam_PI3Ka	superfamily_C2 domain (Calcium/lipid-binding domai	
	GO:0016773	HMMPfam_PI3_PI4_kinase	superfamily_Protein kinase-like (PK-like)	
PIK3CA/ENST00000263967	GO:0004428	HMMPfam_PI3K_C2	superfamily_ARM repeat	BCM
	GO:0005942	HMMPfam_PI3K_p85B	superfamily_C2 domain (Calcium/lipid-binding domai	
	GO:0016773	HMMPfam_PI3K_rbd	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_PI3Ka	superfamily_Ubiquitin-like	
		HMMPfam_PI3_PI4_kinase		
PIK3CB/ENST00000289153	GO:0004428	HMMPfam_PI3K_C2	superfamily_ARM repeat	Broad
	GO:0005942	HMMPfam_PI3K_p85B	superfamily_C2 domain (Calcium/lipid-binding domai	

	GO:0016773	HMMPfam_PI3K_rbd	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_PI3Ka	superfamily_Ubiquitin-like	
		HMMPfam_PI3_PI4_kinase		
PIK3CD/ENST00000361110	GO:0004428	HMMPfam_PI3K_C2	superfamily_ARM repeat	BCM
	GO:0005942	HMMPfam_PI3K_p85B	superfamily_C2 domain (Calcium/lipid-binding domai	
	GO:0016773	HMMPfam_PI3K_rbd	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_PI3Ka	superfamily_Ubiquitin-like	
		HMMPfam_PI3_PI4_kinase		
PIK3CG/ENST00000359195	GO:0004428	HMMPfam_PI3K_C2	superfamily_ARM repeat	BCM
	GO:0005942	HMMPfam_PI3K_rbd	superfamily_C2 domain (Calcium/lipid-binding domai	
	GO:0016773	HMMPfam_PI3Ka	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_PI3_PI4_kinase	superfamily_Ubiquitin-like	
PIK3R1/ENST00000274335	GO:0003924	HMMPfam_RhoGAP	superfamily_GTPase activation domain GAP	WU
	GO:0005622	HMMPfam_SH2	superfamily_Interferon-induced guanylate-binding p	
	GO:0007242	HMMPfam_SH3_1	superfamily_SH2 domain	
			superfamily_SH3-domain	
PIK3R2/ENST00000222254	GO:0005622	HMMPfam_RhoGAP	superfamily_GTPase activation domain GAP	BCM
	GO:0007242	HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
PIK3R3/ENST00000372006	GO:0007242	HMMPfam_SH2	superfamily_SH2 domain	WU
PIK3R4/ENST00000356763	GO:0004672	HMMPfam_HEAT	superfamily_ARM repeat	Broad
		HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_WD40	superfamily_WD40 repeat-like	
PIK4CA/ENST00000255882	GO:0004428	HMMPfam_PI3Ka	superfamily_ADP-ribosylation	BCM
	GO:0016773	HMMPfam_PI3_PI4_kinase	superfamily_ARM repeat	
			superfamily_Protein kinase-like (PK-like)	
PIM1/ENST00000373507	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
PIM2/ENST00000376509	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
PIM3/ENST00000360612	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
PINX1/ENST00000314787	GO:0003676	HMMPfam_G-patch		Broad
PKN1/ENST00000342216	GO:0005622	HMMPfam_HR1	superfamily_C2 domain (Calcium/lipid-binding domai	BCM
	GO:0004672	HMMPfam_Pkinase	superfamily_HR1 repeat	
	GO:0004674	HMMPfam_Pkinase_C	superfamily_Protein kinase-like (PK-like)	

PKN2/ENST00000370521	GO:0005622	HMMPfam_HR1	superfamily_C2 domain (Calcium/lipid-binding domai	Broad
	GO:0004672	HMMPfam_Pkinase	superfamily_HR1 repeat	
	GO:0004674	HMMPfam_Pkinase_C	superfamily_Protein kinase-like (PK-like)	
PKN3/ENST00000291906	GO:0005622	HMMPfam_HR1	superfamily_C2 domain (Calcium/lipid-binding domai	WU
	GO:0004672	HMMPfam_Pkinase	superfamily_HR1 repeat	
	GO:0004674	HMMPfam_Pkinase_C	superfamily_Protein kinase-like (PK-like)	
PLCG1/ENST00000373272	GO:0007242	HMMPfam_C2	superfamily_C2 domain (Calcium/lipid-binding domai	WU
	GO:0004435	HMMPfam_PH	superfamily_EF-hand	
	GO:0004629	HMMPfam_PI-PLC-X	superfamily_PH domain-like	
		HMMPfam_PI-PLC-Y	superfamily_PLC-like phosphodiesterases	
		HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
PLCG2/ENST00000359376	GO:0007242	HMMPfam_C2	superfamily_C2 domain (Calcium/lipid-binding domai	BCM
	GO:0004435	HMMPfam_PH	superfamily_EF-hand	
	GO:0004629	HMMPfam_PI-PLC-X	superfamily_PH domain-like	
		HMMPfam_PI-PLC-Y	superfamily_PLC-like phosphodiesterases	
		HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
PRKAB1/ENST00000229328	GO:0004553	HMMPfam_AMPKBI	superfamily_E set domains	Broad
		HMMPfam_Isoamylase_N		
PRKACA/ENST00000308677	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
	GO:0004674	HMMPfam_Pkinase_C		
PRKACB/ENST00000370685	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
PRKACG/ENST00000377276	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
PRKAR1A/ENST00000358598	GO:0007165	HMMPfam_RIIa	superfamily_cAMP-binding domain-like	Broad
		HMMPfam_cNMP_binding		
PRKCA/ENST00000284384	GO:0007242	HMMPfam_C1_1	superfamily_C2 domain (Calcium/lipid-binding domai	Broad
	GO:0004672	HMMPfam_C2	superfamily_Cysteine-rich domain	
	GO:0004674	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_Pkinase_C		
PRKCB1/ENST00000303531	GO:0007242	HMMPfam_C1_1	superfamily_C2 domain (Calcium/lipid-binding domai	Broad
	GO:0004672	HMMPfam_C2	superfamily_Cysteine-rich domain	
	GO:0004674	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	

		HMMPfam_Pkinase_C		
PRKCD/ENST00000330452	GO:0007242	HMMPfam_C1_1	superfamily_C2 domain (Calcium/lipid-binding domai	BCM
	GO:0004672	HMMPfam_Pkinase	superfamily_Cysteine-rich domain	
	GO:0004674	HMMPfam_Pkinase_C	superfamily_Protein kinase-like (PK-like)	
PRKCE/ENST00000306156	GO:0007242	HMMPfam_C1_1	superfamily_C2 domain (Calcium/lipid-binding domai	WU
	GO:0004672	HMMPfam_C2	superfamily_Cysteine-rich domain	
	GO:0004674	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_Pkinase_C		
PRKCG/ENST00000263431	GO:0007242	HMMPfam_C1_1	superfamily_C2 domain (Calcium/lipid-binding domai	BCM
	GO:0004672	HMMPfam_C2	superfamily_Cysteine-rich domain	
	GO:0004674	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_Pkinase_C		
PRKCH/ENST00000332981	GO:0007242	HMMPfam_C1_1	superfamily_C2 domain (Calcium/lipid-binding domai	Broad
	GO:0004672	HMMPfam_C2	superfamily_Cysteine-rich domain	
	GO:0004674	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_Pkinase_C		
PRKCI/OTTHUMT00000267042	GO:0007242	HMMPfam_C1_1	superfamily_CAD PB1 domains	BCM
	GO:0004672	HMMPfam_PB1	superfamily_Cysteine-rich domain	
	GO:0004674	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_Pkinase_C		
PRKCQ/ENST00000263125	GO:0007242	HMMPfam_C1_1	superfamily_C2 domain (Calcium/lipid-binding domai	BCM
	GO:0004672	HMMPfam_Pkinase	superfamily_Cysteine-rich domain	
	GO:0004674	HMMPfam_Pkinase_C	superfamily_Protein kinase-like (PK-like)	
PRKCZ/ENST00000378567	GO:0007242	HMMPfam_C1_1	superfamily_CAD PB1 domains	BCM
	GO:0004672	HMMPfam_PB1	superfamily_Cysteine-rich domain	
	GO:0004674	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
		HMMPfam_Pkinase_C		
PRKD1/ENST00000389985	GO:0007242	HMMPfam_C1_1	superfamily_Cysteine-rich domain	BCM
	GO:0004672	HMMPfam_PH	superfamily_PH domain-like	
		HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
PRKD3/ENST00000379066	GO:0007242	HMMPfam_C1_1	superfamily_Cysteine-rich domain	WU
	GO:0004672	HMMPfam_PH	superfamily_PH domain-like	
		HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	

PRKDC/ENST00000314191	GO:0003677 GO:0016773	HMMPfam_FAT HMMPfam_FATC HMMPfam_HEAT HMMPfam_NUC194 HMMPfam_PI3_PI4_kinase	superfamily_ARM repeat superfamily_Protein kinase-like (PK-like) superfamily_TPR-like	WU
PRKG1/ENST00000373980	GO:0004672	HMMPfam_Pkinase HMMPfam_cNMP_binding	superfamily_Protein kinase-like (PK-like) superfamily_cAMP-binding domain-like	WU
PRKG2/ENST00000264399	GO:0004672	HMMPfam_Pkinase HMMPfam_cNMP_binding	superfamily_Protein kinase-like (PK-like) superfamily_cAMP-binding domain-like	WU
PRKX/ENST00000262848	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
PRKY/ENST00000362758	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
PTCH/ENST00000331920	GO:0008158	HMMPfam_Patched	superfamily_Multidrug efflux transporter AcrB tran	BCM
PTCH2/ENST00000372192	GO:0008158	HMMPfam_Patched	superfamily_Multidrug efflux transporter AcrB tran	WU
PTEN/ENST00000371953	GO:0006470	HMMPfam_DSPc	superfamily_(Phosphotyrosine protein) phosphatases superfamily_C2 domain (Calcium/lipid-binding domai	SHARE
PTK2/OTTHUMT00000275592	GO:0004713	HMMPfam_Focal_AT HMMPfam_Pkinase_Tyr	superfamily_FAT domain of focal adhesion kinase superfamily_Protein kinase-like (PK-like) superfamily_Second domain of FERM	Broad
PTK2B/ENST00000346049	GO:0004713	HMMPfam_Focal_AT HMMPfam_Pkinase_Tyr	superfamily_FAT domain of focal adhesion kinase superfamily_Protein kinase-like (PK-like) superfamily_Second domain of FERM	Broad
PTK7/ENST00000230419	GO:0004713	HMMPfam_I-set HMMPfam_Pkinase_Tyr HMMPfam_ig	superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	Broad
PTPN11/ENST00000351677	GO:0007242 GO:0004725	HMMPfam_SH2 HMMPfam_Y_phosphatase	superfamily_(Phosphotyrosine protein) phosphatases superfamily_SH2 domain	SHARE
PTPRD/ENST00000356435	GO:0004725	HMMPfam_I-set HMMPfam_Y_phosphatase HMMPfam_fn3 HMMPfam_ig	superfamily_(Phosphotyrosine protein) phosphatases superfamily_Fibronectin type III superfamily_Immunoglobulin	WU
PTPRG/ENST00000295874	GO:0004089 GO:0004725	HMMPfam_Carb_anhydrase HMMPfam_Y_phosphatase HMMPfam_fn3	superfamily_(Phosphotyrosine protein) phosphatases superfamily_Carbonic anhydrase superfamily_Fibronectin type III	BCM

RAB43/NM_198490	GO:0000166 GO:0003677 GO:0003924 GO:0004871 GO:0004879 GO:0005515 GO:0005525 GO:0005622 GO:0005634 GO:0005681 GO:0006355 GO:0006397 GO:0006886 GO:0006913 GO:0007165 GO:0007186 GO:0007264 GO:0008380 GO:0015031 GO:0016020 GO:0019001	HMMPfam_RAS		BCM
RAF1/ENST00000251849	GO:0007242 GO:0004672 GO:0005057	HMMPfam_C1_1 HMMPfam_Pkinase HMMPfam_RBD	superfamily_Cysteine-rich domain superfamily_Protein kinase-like (PK-like) superfamily_Ubiquitin-like	SHARE
RAGE/ENST00000361847	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
RALA/ENST00000005257		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	BCM
RALB/ENST00000376224		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	BCM
RAN/OTTHUMT00000263332		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	Broad
RANBP9/ENST00000011619		HMMPfam_LisH HMMPfam_SPRY		WU
RAP1A/ENST00000356415		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	WU
RAP1B/ENST00000250559		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	Broad
RAP2A/ENST00000245304		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	BCM

RAP2B/ENST00000323534		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	WU
RAP2C/ENST00000342983		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	WU
RAPGEF1/ENST00000337036	GO:0005085	HMMPfam_RasGEF HMMPfam_RasGEF_N	superfamily_Ras GEF	WU
RAPTOR/NM_020761		HMMPfam_HEAT HMMPfam_WD40	superfamily_ARM repeat superfamily_WD40 repeat-like	WU
RARB/ENST00000383772	GO:0003700	HMMPfam_Hormone_recep HMMPfam_zf-C4	superfamily_Glucocorticoid receptor-like (DNA-bind superfamily_Nuclear receptor ligand-binding domain	Broad
RASEF/ENST00000330861	GO:0005509	HMMPfam_Ras HMMPfam_efhand	superfamily_EF-hand superfamily_P-loop containing nucleoside triphosph superfamily_alpha-catenin/vinculin	WU
RASSF1/ENST00000357043	GO:0007242 GO:0007165	HMMPfam_C1_1 HMMPfam_RA	superfamily_Cysteine-rich domain	WU
RASSF2/NM_170774	GO:0007165	HMMPfam_RA		BCM
RASSF3/ENST00000336061	GO:0007165	HMMPfam_RA		Broad
RB1/ENST00000267163		HMMPfam_RB_A HMMPfam_RB_B HMMPfam_Rb_C	superfamily_Cyclin-like superfamily_Formin homology 2 domain (FH2 domain)	SHARE
RBL1/ENST00000373664		HMMPfam_RB_A HMMPfam_RB_B	superfamily_Cyclin-like	WU
RBL2/ENST00000262133		HMMPfam_RB_A HMMPfam_RB_B	superfamily_Cyclin-like	WU
RBM5/ENST00000347869	GO:0005622 GO:0003676	HMMPfam_G-patch HMMPfam_RRM_1 HMMPfam_zf-RanBP	superfamily_C2H2 and C2HC zinc fingers superfamily_RNA-binding domain RBD	WU
RBM6/ENST00000266022	GO:0003676	HMMPfam_G-patch	superfamily_C2H2 and C2HC zinc fingers superfamily_NZF domain superfamily_RNA-binding domain RBD	Broad
RCBTB2/ENST00000344532		HMMPfam_BTB HMMPfam_RCC1	superfamily_POZ domain superfamily_RCC1/BLIP-II	Broad
REEP5/ENST00000379639		HMMPfam_TB2_DP1_HVA22		Broad
REL/ENST00000295025	GO:0003700	HMMPfam_RHD HMMPfam_TIG	superfamily_E set domains superfamily_p53-like transcription factors	BCM

RELA/ENST00000308639	GO:0003700	HMMPfam_RHD HMMPfam_TIG	superfamily_E set domains superfamily_p53-like transcription factors	BCM
RELB/ENST00000221452	GO:0003700	HMMPfam_RHD HMMPfam_TIG	superfamily_E set domains superfamily_p53-like transcription factors	WU
RET/ENST00000340058	GO:0005509 GO:0004713	HMMPfam_Cadherin HMMPfam_Pkinase_Tyr	superfamily_Cadherin-like superfamily_Protein kinase-like (PK-like) superfamily_Scorpion toxin-like	SHARE
RHEB/ENST00000262187		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	BCM
RHOA/ENST00000265538		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	WU
RHOB/ENST00000381107		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	Broad
RHOBTB2/OTTHUMT00000273508		HMMPfam_BTB HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph superfamily_POZ domain	Broad
RHOC/ENST00000369633		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	WU
RICTOR/NM_152756			superfamily_ARM repeat superfamily_Cyclin-like	Broad
RIN1/ENST00000311320	GO:0007165	HMMPfam_RA HMMPfam_VPS9	superfamily_SH2 domain superfamily_VPS9 domain (Pfam 02204)	BCM
RIPK1/ENST00000380409	GO:0005515 GO:0004672	HMMPfam_Death HMMPfam_Pkinase	superfamily_DEATH domain superfamily_Protein kinase-like (PK-like)	WU
RIPK2/ENST00000220751	GO:0005515 GO:0004672	HMMPfam_CARD HMMPfam_Pkinase	superfamily_DEATH domain superfamily_Protein kinase-like (PK-like)	WU
RIPK3/ENST00000216274	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
RIPK4/ENST00000352483	GO:0004672	HMMPfam_Ank HMMPfam_Pkinase	superfamily_Ankyrin repeat superfamily_Protein kinase-like (PK-like)	WU
ROBO1/OTTHUMT00000265386		HMMPfam_I-set HMMPfam_fn3	superfamily_Fibronectin type III superfamily_Immunoglobulin	Broad
ROBO2/ENST00000332191		HMMPfam_I-set HMMPfam_fn3	superfamily_Fibronectin type III superfamily_Immunoglobulin	Broad
ROCK1/ENST00000261535	GO:0005622 GO:0007242 GO:0004672 GO:0004674	HMMPfam_C1_1 HMMPfam_HR1 HMMPfam_PH HMMPfam_Pkinase HMMPfam_Pkinase_C	superfamily_Cysteine-rich domain superfamily_PH domain-like superfamily_Prefoldin superfamily_Protein kinase-like (PK-like)	BCM

		HMMPfam_Rho_Binding		
ROCK2/ENST00000315872	GO:0005622	HMMPfam_C1_1	superfamily_Cysteine-rich domain	BCM
	GO:0007242	HMMPfam_HR1	superfamily_PH domain-like	
	GO:0004672	HMMPfam_PH	superfamily_Protein kinase-like (PK-like)	
	GO:0004674	HMMPfam_Pkinase	superfamily_tRNA-binding arm	
		HMMPfam_Pkinase_C		
		HMMPfam_Rho_Binding		
ROR1/ENST00000371079	GO:0004713	HMMPfam_Fz	superfamily_Frizzled cysteine-rich domain	BCM
		HMMPfam_I-set	superfamily_Immunoglobulin	
		HMMPfam_Kringle	superfamily_Kringle-like	
		HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	
ROR2/ENST00000375708	GO:0004713	HMMPfam_Fz	superfamily_Frizzled cysteine-rich domain	BCM
		HMMPfam_I-set	superfamily_Immunoglobulin	
		HMMPfam_Kringle	superfamily_Kringle-like	
		HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	
RPS6KA1/ENST00000374162	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
	GO:0004674	HMMPfam_Pkinase_C		
RPS6KA2/ENST00000366867	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
	GO:0004674	HMMPfam_Pkinase_C		
RPS6KA3/ENST00000379565	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
	GO:0004674	HMMPfam_Pkinase_C		
RPS6KA4/ENST00000334205	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
	GO:0004674	HMMPfam_Pkinase_C		
RPS6KA5/ENST00000261991	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
	GO:0004674	HMMPfam_Pkinase_C		
RPS6KA6/ENST00000262752	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
	GO:0004674	HMMPfam_Pkinase_C		
RPS6KB1/ENST00000225577	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
	GO:0004674	HMMPfam_Pkinase_C		
RPS6KB2/ENST00000312629	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
	GO:0004674	HMMPfam_Pkinase_C		
RPSA/NM_001012321	GO:0003735	HMMPfam_Ribosomal_S2	superfamily_Ribosomal protein S2	WU
RRAS/ENST00000246792		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	WU

RRAS2/ENST00000256196		HMMPfam_Ras	superfamily_P-loop containing nucleoside triphosph	Broad
RUNX1/ENST00000300305		HMMPfam_Runt HMMPfam_RunxI	superfamily_p53-like transcription factors	WU
RUNX2/ENST00000359524		HMMPfam_Runt HMMPfam_RunxI	superfamily_p53-like transcription factors	Broad
RUNX3/ENST00000338888		HMMPfam_Runt HMMPfam_RunxI	superfamily_p53-like transcription factors	Broad
RYK/ENST00000296084	GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_WIF	superfamily_Protein kinase-like (PK-like)	Broad
SCARF2/ENST00000266214		HMMPfam_EGF_2 HMMPfam_Laminin_EGF	superfamily_EGF/Laminin superfamily_Growth factor receptor domain	BCM
SEMA3B/ENST00000316347	GO:0004872	HMMPfam_PSI HMMPfam_Sema HMMPfam_ig	superfamily_Immunoglobulin superfamily_Plexin repeat superfamily_Sema domain	WU
SEMA3F/ENST00000002829		HMMPfam_Sema HMMPfam_ig	superfamily_Immunoglobulin superfamily_Plexin repeat superfamily_Sema domain	WU
SERPIND1/OTTHUMT00000075725	GO:0004867	HMMPfam_Serpin	superfamily_Serpins	BCM
SGK/ENST00000367858	GO:0004672 GO:0004674	HMMPfam_Pkinase HMMPfam_Pkinase_C	superfamily_PX domain superfamily_Protein kinase-like (PK-like)	WU
SGK2/ENST00000341458	GO:0004672 GO:0004674	HMMPfam_Pkinase HMMPfam_Pkinase_C	superfamily_Protein kinase-like (PK-like)	WU
SGKL/ENST00000262211	GO:0005515 GO:0004672 GO:0004674	HMMPfam_PX HMMPfam_Pkinase HMMPfam_Pkinase_C	superfamily_PX domain superfamily_Protein kinase-like (PK-like)	BCM
SHC1/ENST00000368447	GO:0007242	HMMPfam_PID HMMPfam_SH2	superfamily_PH domain-like superfamily_SH2 domain	SHARE
SHC2/ENST00000264554	GO:0007242	HMMPfam_PID HMMPfam_SH2	superfamily_PH domain-like superfamily_SH2 domain	SHARE
SHC3/ENST00000375835	GO:0007242	HMMPfam_PID HMMPfam_SH2	superfamily_PH domain-like superfamily_SH2 domain	SHARE
SHH/ENST00000297261	GO:0007267 GO:0006508	HMMPfam_HH_signal HMMPfam_Hint	superfamily_Hedgehog/DD-peptidase superfamily_Hedgehog/intein (Hint) domain	BCM

SIAH1/ENST00000359996	GO:0005634	HMMPfam_Sina	superfamily_RING/U-box superfamily_TRAF domain-like	WU
SKI/ENST00000378536	GO:0005634	HMMPfam_Ski_Sno HMMPfam_c-SKI_SMAD_bind	superfamily_Putative DNA-binding domain superfamily_SAND domain-like	BCM
SLA/OTTHUMT00000275504	GO:0007242	HMMPfam_SH2 HMMPfam_SH3_1	superfamily_SH2 domain superfamily_SH3-domain	Broad
SLC38A3/ENST00000341160		HMMPfam_Aa_trans		WU
SLC7A4/ENST00000382932	GO:0006810	HMMPfam_AA_permease		BCM
SLK/ENST00000369755	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
SMAD1/ENST00000302085	GO:0005622	HMMPfam_MH1 HMMPfam_MH2	superfamily_SMAD MH1 domain superfamily_SMAD/FHA domain	WU
SMAD2/ENST00000262160	GO:0005622	HMMPfam_MH1 HMMPfam_MH2	superfamily_SMAD MH1 domain superfamily_SMAD/FHA domain	Broad
SMAD3/ENST00000327367	GO:0005622	HMMPfam_MH1 HMMPfam_MH2	superfamily_SMAD MH1 domain superfamily_SMAD/FHA domain	Broad
SMAD4/ENST00000342988	GO:0005622	HMMPfam_MH1 HMMPfam_MH2	superfamily_SMAD MH1 domain superfamily_SMAD/FHA domain	Broad
SMAD5/ENST00000231589	GO:0005622	HMMPfam_MH1 HMMPfam_MH2	superfamily_SMAD MH1 domain superfamily_SMAD/FHA domain	WU
SMAD6/ENST00000288840	GO:0005622	HMMPfam_MH1 HMMPfam_MH2	superfamily_SMAD MH1 domain superfamily_SMAD/FHA domain	Broad
SMAD7/ENST00000262158	GO:0005622	HMMPfam_MH1 HMMPfam_MH2	superfamily_SMAD MH1 domain superfamily_SMAD/FHA domain	Broad
SMAD9/ENST00000379826	GO:0005622	HMMPfam_MH1 HMMPfam_MH2	superfamily_SMAD MH1 domain superfamily_SMAD/FHA domain	BCM
SMG1/NM_015092	GO:0016773	HMMPfam_FATC HMMPfam_HEAT HMMPfam_PI3_PI4_kinase	superfamily_ARM repeat superfamily_Protein kinase-like (PK-like) superfamily_Spectrin repeat	BCM
SMO/ENST00000249373	GO:0004926	HMMPfam_Frizzled HMMPfam_Fz	superfamily_Frizzled cysteine-rich domain	Broad
SNAP29/ENST00000215730		HMMPfam_SNAP-25 HMMPfam_SNARE		BCM

SOS1/ENST00000263879	GO:0003677 GO:0005089 GO:0005085	HMMPfam_Histone HMMPfam_PH HMMPfam_RasGEF HMMPfam_RasGEF_N HMMPfam_RhoGEF	superfamily_DBL homology domain (DH-domain) superfamily_Histone-fold superfamily_PH domain-like superfamily_Ras GEF	WU
SOS2/ENST00000216373	GO:0003677 GO:0005089 GO:0005085	HMMPfam_Histone HMMPfam_PH HMMPfam_RasGEF HMMPfam_RasGEF_N HMMPfam_RhoGEF	superfamily_DBL homology domain (DH-domain) superfamily_Histone-fold superfamily_PH domain-like superfamily_Ras GEF	Broad
SRC/ENST00000373558	GO:0007242 GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_SH2 HMMPfam_SH3_1	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain superfamily_SH3-domain	SHARE
SRMS/ENST00000217188	GO:0007242 GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_SH2 HMMPfam_SH3_2	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain superfamily_SH3-domain	SHARE
STAT1/ENST00000361099	GO:0007242 GO:0003700	HMMPfam_SH2 HMMPfam_STAT_alpha HMMPfam_STAT_bind HMMPfam_STAT_int	superfamily_SH2 domain superfamily_STAT superfamily_Transcription factor STAT-4 N-domain superfamily_p53-like transcription factors	BCM
STAT2/ENST00000314128	GO:0007242 GO:0003700	HMMPfam_SH2 HMMPfam_STAT_alpha HMMPfam_STAT_bind HMMPfam_STAT_int	superfamily_SH2 domain superfamily_STAT superfamily_Transcription factor STAT-4 N-domain superfamily_p53-like transcription factors	BCM
STAT3/ENST00000264657	GO:0007242 GO:0003700	HMMPfam_SH2 HMMPfam_STAT_alpha HMMPfam_STAT_bind HMMPfam_STAT_int	superfamily_SH2 domain superfamily_STAT superfamily_Transcription factor STAT-4 N-domain superfamily_p53-like transcription factors	BCM
STAT4/ENST00000358470	GO:0007242 GO:0003700	HMMPfam_SH2 HMMPfam_STAT_alpha HMMPfam_STAT_bind HMMPfam_STAT_int	superfamily_SH2 domain superfamily_STAT superfamily_Transcription factor STAT-4 N-domain superfamily_p53-like transcription factors	BCM
STAT5A/ENST00000345506	GO:0007242	HMMPfam_SH2	superfamily_SH2 domain	BCM

	GO:0003700	HMMPfam_STAT_alpha	superfamily_STAT	
		HMMPfam_STAT_bind	superfamily_Transcription factor STAT-4 N-domain	
		HMMPfam_STAT_int	superfamily_p53-like transcription factors	
STAT5B/ENST00000293328	GO:0007242	HMMPfam_SH2	superfamily_SH2 domain	BCM
	GO:0003700	HMMPfam_STAT_alpha	superfamily_STAT	
		HMMPfam_STAT_bind	superfamily_Transcription factor STAT-4 N-domain	
		HMMPfam_STAT_int	superfamily_p53-like transcription factors	
STAT6/ENST00000300134	GO:0007242	HMMPfam_SH2	superfamily_SH2 domain	Broad
	GO:0003700	HMMPfam_STAT_alpha	superfamily_STAT	
		HMMPfam_STAT_bind	superfamily_Transcription factor STAT-4 N-domain	
		HMMPfam_STAT_int	superfamily_p53-like transcription factors	
STK10/ENST00000389883	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
STK11/ENST00000326873	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	SHARE
STK23/OTTHUMT00000060999	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
STK24/ENST00000376547	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
STK25/ENST00000316586	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
STK3/ENST00000354930	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
STK32A/ENST00000306304	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
STK32B/ENST00000282908	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
STK32C/ENST00000298630	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
STK36/ENST00000295709	GO:0004672	HMMPfam_HEAT	superfamily_ARM repeat	WU
		HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	
STK38/ENST00000229812	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
	GO:0004674	HMMPfam_Pkinase_C		
STK38L/ENST00000389032	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
	GO:0004674	HMMPfam_Pkinase_C		
STK39/ENST00000355999	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
STK6/ENST00000347343	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
STYK1/ENST00000075503	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	Broad

SUCNR1/ENST00000362032	GO:0001584	HMMPfam_7tm_1	superfamily_Family A G protein-coupled receptor-li	WU
SUFU/ENST00000369902		HMMPfam_SUFU	superfamily_Suppressor of Fused N-terminal domain	Broad

SYK/ENST00000375754	GO:0007242 GO:0004713	HMMPfam_Pkinase_Tyr HMMPfam_SH2	superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain	BCM
TAOK1/ENST00000261716	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_Ribbon-helix-helix	Broad
TAOK2/ENST00000308893	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad
TAOK3/ENST00000359166	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_Ribbon-helix-helix	Broad
TBK1/ENST00000331710	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_Ubiquitin-like	Broad
TCF1/ENST00000257555	GO:0005634 GO:0003700	HMMPfam_HNF-1A_C HMMPfam_HNF-1B_C HMMPfam_HNF-1_N HMMPfam_Homeobox	superfamily_Dimerization cofactor of HNF-1 alpha superfamily_Homeodomain-like superfamily_lambda repressor-like DNA-binding doma	Broad
TEC/ENST00000381501	GO:0007242 GO:0004713	HMMPfam_BTK HMMPfam_PH HMMPfam_Pkinase_Tyr HMMPfam_SH2 HMMPfam_SH3_1	superfamily_PH domain-like superfamily_Protein kinase-like (PK-like) superfamily_SH2 domain superfamily_SH3-domain	BCM
TEK/ENST00000380040	GO:0004713	HMMPfam_EGF_2 HMMPfam_Pkinase_Tyr HMMPfam_fn3	superfamily_EGF/Laminin superfamily_Fibronectin type III superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	BCM
TERT/ENST00000310581			superfamily_DNA/RNA polymerases	SHARE
TESK1/ENST00000336395	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
TESK2/ENST00000372086	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	WU
TFDP1/ENST00000375370	GO:0000074	HMMPfam_DP HMMPfam_E2F_TDP		Broad
TGFB2/ENST00000366929	GO:0008083	HMMPfam_TGF_beta HMMPfam_TGFb_propeptide	superfamily_Cystine-knot cytokines	Broad
TGFBR1/ENST00000374994	GO:0004672 GO:0004675	HMMPfam_Activin_recp HMMPfam_Pkinase HMMPfam_TGF_beta_GS	superfamily_Protein kinase-like (PK-like) superfamily_Snake toxin-like	BCM
TGFBR2/ENST00000359013	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	Broad

		HMMPfam_ecTbetaR2	superfamily_Snake toxin-like	
THAP7/ENST00000215742	GO:0003676	HMMPfam_THAP		BCM
TIE1/ENST00000372476	GO:0004713	HMMPfam_EGF HMMPfam_EGF_2 HMMPfam_Pkinase_Tyr HMMPfam_fn3 HMMPfam_ig	superfamily_EGF/Laminin superfamily_Fibronectin type III superfamily_Immunoglobulin superfamily_Protein kinase-like (PK-like)	WU
TMEM115/ENST00000266025		HMMPfam_DUF1751		WU
TNIK/ENST00000284483	GO:0004672 GO:0005083	HMMPfam_CNH HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like) superfamily_YVTN repeat-like/Quinoprotein amine de	BCM
TNK1/ENST00000311668	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like) superfamily_SAM/Pointed domain superfamily_SH3-domain	Broad
TNK2/ENST00000381916	GO:0004713	HMMPfam_GTPase_binding HMMPfam_Pkinase_Tyr HMMPfam_SH3_1	superfamily_Protein kinase-like (PK-like) superfamily_SAM/Pointed domain superfamily_SH3-domain	Broad
TNNI3K/NM_015978	GO:0004713	HMMPfam_Ank HMMPfam_Pkinase_Tyr	superfamily_Ankyrin repeat superfamily_Protein kinase-like (PK-like)	BCM
TP53/ENST00000269305	GO:0003677	HMMPfam_P53 HMMPfam_P53_TAD HMMPfam_P53_tetramer	superfamily_p53 tetramerization domain superfamily_p53-like transcription factors	SHARE
TP73/ENST00000378295	GO:0003677	HMMPfam_P53 HMMPfam_P53_tetramer HMMPfam_SAM_2	superfamily_Glycosyl hydrolase domain superfamily_SAM/Pointed domain superfamily_p53 tetramerization domain superfamily_p53-like transcription factors	BCM
TP73L/ENST00000264731	GO:0003677	HMMPfam_P53 HMMPfam_P53_tetramer HMMPfam_SAM_2	superfamily_SAM/Pointed domain superfamily_p53 tetramerization domain superfamily_p53-like transcription factors	BCM
TPTE/ENST00000359693			superfamily_(Phosphotyrosine protein) phosphatases superfamily_C2 domain (Calcium/lipid-binding domai superfamily_Voltage-gated potassium channels	BCM
TRIP/ENST00000331456	GO:0005515	HMMPfam_zf-C3HC4	superfamily_Prefoldin superfamily_RING/U-box	WU

TSC1/ENST00000298552		HMMPfam_Hamartin	superfamily_ARM repeat	WU
TSC2/ENST00000219476	GO:0005096	HMMPfam_Rap_GAP	superfamily_ARM repeat	WU
		HMMPfam_Tuberin	superfamily_Rap/Ran-GAP (Pfam 02145)	
TSHR/ENST00000298171	GO:0005515	HMMPfam_7tm_1	superfamily_Family A G protein-coupled receptor-li	Broad
	GO:0001584	HMMPfam_LRR_1	superfamily_L domain-like	
TU3A/ENST00000360997		HMMPfam_DUF1151		BCM
TUSC2/ENST00000232496	GO:0007049			WU
	GO:0008283			
	GO:0007267			
	GO:0045786			
	GO:0005515			
TUSC4/ENST00000232501		HMMPfam_NPR2		WU
TXK/ENST00000381515	GO:0007242	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	BCM
	GO:0004713	HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
TYK2/ENST00000264818	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	BCM
			superfamily_SH2 domain	
TYRO3/ENST00000263798	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Fibronectin type III	WU
		HMMPfam_fn3	superfamily_Immunoglobulin	
		HMMPfam_ig	superfamily_Protein kinase-like (PK-like)	
UBE2L3/ENST00000342192	GO:0006464	HMMPfam_UQ_con	superfamily_UBC-like	BCM
VAV1/ENST00000304076	GO:0005089	HMMPfam_C1_1	superfamily_Calponin-homology domain CH-domain	BCM
	GO:0007242	HMMPfam_CH	superfamily_Cysteine-rich domain	
		HMMPfam_PH	superfamily_DBL homology domain (DH-domain)	
		HMMPfam_RhoGEF	superfamily_PH domain-like	
		HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
VAV2/ENST00000371850	GO:0005089	HMMPfam_C1_1	superfamily_Calponin-homology domain CH-domain	WU
	GO:0007242	HMMPfam_CH	superfamily_Cysteine-rich domain	
		HMMPfam_PH	superfamily_DBL homology domain (DH-domain)	
		HMMPfam_RhoGEF	superfamily_PH domain-like	
		HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	

		HMMPfam_SH3_2		
VAV3/ENST00000370056	GO:0005089	HMMPfam_C1_1	superfamily_Calponin-homology domain CH-domain	WU
	GO:0007242	HMMPfam_CH	superfamily_Cysteine-rich domain	
		HMMPfam_PH	superfamily_DBL homology domain (DH-domain)	
		HMMPfam_RhoGEF	superfamily_PH domain-like	
		HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
		HMMPfam_SH3_2		
VEGFC/ENST00000280193	GO:0008083	HMMPfam_CXCXC	superfamily_Cystine-knot cytokines	Broad
		HMMPfam_PDGF		
VHL/ENST00000256474	GO:0005634	HMMPfam_VHL	superfamily_VHL	SHARE
WEE1/ENST00000299613	GO:0004672	HMMPfam_Pkinase	superfamily_Protein kinase-like (PK-like)	BCM
WT1/ENST00000332351	GO:0003676	HMMPfam_WT1	superfamily_C2H2 and C2HC zinc fingers	WU
	GO:0003700	HMMPfam_zf-C2H2		
XRCC1/ENST00000262887	GO:0005622	HMMPfam_BRCT	superfamily_BRCT domain	WU
	GO:0000012	HMMPfam_XRCC1_N	superfamily_Galactose-binding domain-like	
YES1/ENST00000314574	GO:0007242	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	SHARE
	GO:0004713	HMMPfam_SH2	superfamily_SH2 domain	
		HMMPfam_SH3_1	superfamily_SH3-domain	
ZAK/NM_016653	GO:0004713	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	WU
		HMMPfam_SAM_2	superfamily_SAM/Pointed domain	
ZAP70/ENST00000389517	GO:0007242	HMMPfam_Pkinase_Tyr	superfamily_Protein kinase-like (PK-like)	WU
	GO:0004713	HMMPfam_SH2	superfamily_SH2 domain	
ZMYND10/ENST00000231749	GO:0008270	HMMPfam_zf-MYND		WU
ZNF312/ENST00000283268	GO:0003676	HMMPfam_zf-C2H2	superfamily_C2H2 and C2HC zinc fingers	BCM

Supplementary Table 2. Annotation of somatic mutations found in 188 lung adenocarcinomas. Excel spreadsheet can be assessed at http://genome.wustl.edu/supplemental/tsp_nature_2008.

Supplementary Table 3a. Top genes with higher frequencies of nonsense, splice site, and frame-shift mutations in 188 lung adenocarcinomas according to Fisher's exact test .

Gene	Num. of Missense/In_Frame_Indels	Num. of Nonsense/Nonstop/Splice_site/Frame_Shift_Indels	P_value	FDR	Bonferroni
STK11	9	25	6.49E-12	2.31E-09	2.31E-09
RB1	0	7	6.63E-06	0.001181	0.002362
NF1	6	10	0.000112	0.013331	0.039992
TP53	43	23	0.001166	0.103745	0.414979
APC	6	7	0.003985	0.283723	1
ACVR1B	0	2	0.032425	1	1
BAP1	0	2	0.032425	1	1
PIK3R2	0	2	0.032425	1	1
LTK	3	3	0.076083	1	1
FYN	1	2	0.085518	1	1
IRAK2	1	2	0.085518	1	1
NOTCH4	1	2	0.085518	1	1
PRKDC	4	3	0.115594	1	1
ERBB2	2	2	0.150735	1	1
ABL2	0	1	0.179664	1	1
ARAF	0	1	0.179664	1	1
BAX	0	1	0.179664	1	1
CACNA2D2	0	1	0.179664	1	1
EPHA2	0	1	0.179664	1	1
EVI1	0	1	0.179664	1	1
HD	0	1	0.179664	1	1
IKBKB	0	1	0.179664	1	1
JAG1	0	1	0.179664	1	1
MAP2K6	0	1	0.179664	1	1
MAPK8	0	1	0.179664	1	1
MST1R	0	1	0.179664	1	1
MUSK	0	1	0.179664	1	1
PIM2	0	1	0.179664	1	1
PRKAR1A	0	1	0.179664	1	1

RAB43	0	1	0.179664	1	1
RANBP9	0	1	0.179664	1	1
RIPK1	0	1	0.179664	1	1
RPS6KA4	0	1	0.179664	1	1
SRC	0	1	0.179664	1	1
TGFBR1	0	1	0.179664	1	1

Total number of Missense/In_Frame_Indels in all genes tested: 831
Total number of Nonsense/Splice_site/Frame_Shift_Indels in all genes tested: 182

Supplementary Table 3b. Recurrent somatic mutations at 28 sites across seven genes.

Gene	# of samples with mutation	Chromosome	Start position	End position	Amino acid change	Variant type
KRAS	27	12	25289552	25289552	p.G12C,R,S	Missense
KRAS	21	12	25289551	25289551	p.G12A,D,V	Missense
EGFR	11	7	55227009	55227009	p.L858R	Missense
KRAS	7	12	25289549	25289549	p.G13C	Missense
TP53	4	17	7517867	7517867	p.G266*	Nonsense
EGFR	4	7	55209960	55209974	p.E746_A750del	In Frame Del
TP53	3	17	7517845	7517845	p.R273L	Missense
TP53	3	17	7517819	7517819	p.R282W	Missense
STK11	3	19	1158092	1158092	p.Y60fs	Frame Shift Del
NRAS	3	1	1.15E+08	1.15E+08	p.Q61L	Missense
KRAS	3	12	25289551	25289552	p.G12F	Missense
EGFR	3	7	55209959	55209973	p.E746	A750del
TP53	2	17	7519192	7519192	p.T155P	Missense
TP53	2	17	7519186	7519186	p.V157F	Missense
TP53	2	17	7519174	7519174	p.A161T	Missense
TP53	2	17	7519167	7519167	p.Y163C	Missense

TP53	2	17	7518901	7518901	e5+1	Splice Site
TP53	2	17	7518273	7518273	p.G245C	Missense
TP53	2	17	7517839	7517839	p.C275S	Missense
STK11	2	19	1173005	1173006	c.920_e7+1	Splice Site
STK11	2	19	1158021	1158021	p.Q37*	Nonsense
PAK7	2	20	9494832	9494832	p.T397K	Missense
KRAS	2	12	25289548	25289548	p.G13D	Missense
FLT4	2	5	1.8E+08	1.8E+08	p.Y1333H	Missense
EGFR	2	7	55227018	55227018	p.L861Q	Missense
EGFR	2	7	55226933	55226933	p.L833V	Missense
EGFR	2	7	55216507	55216508	p.771insSVD	In Frame Ins
EGFR	2	7	55216506	55216507	p.771insGL	In Frame Ins

Supplementary Table 4a. Signature analysis of all 12 possible changes of the transcribed strand in various 5' (before colon) and 3' (after colon) sequence contexts for nonsynonymous mutations found in 623 TSP genes in lung adenocarcinomas.

	A:A	A:C	A:G	A:T	C:A	C:C	C:G	C:T	G:A	G:C	G:G	G:T	T:A	T:C	T:G	T:T
A>C	0	1	0	1	1	2	1	1	0	0	1	0	0	0	0	1
A>G	0	1	1	3	1	3	5	8	0	0	0	2	0	5	1	5
A>T	3	3	0	1	8	2	17	7	3	3	3	4	1	0	1	0
C>A	3	8	8	3	2	9	4	2	5	7	5	11	3	6	2	5
C>G	5	0	3	1	2	2	5	6	1	2	3	1	8	6	1	9
C>T	3	5	6	1	11	13	17	3	4	8	6	0	14	8	5	7
G>A	8	4	9	4	10	7	10	6	12	12	7	14	9	1	6	15
G>C	9	3	3	1	8	3	4	2	6	5	3	8	6	2	4	4
G>T	24	14	18	13	15	20	19	15	24	14	27	20	6	14	43	8
T>A	0	0	5	3	2	0	4	0	2	0	5	0	0	1	0	0
T>C	0	0	1	2	1	1	6	2	4	0	1	0	2	1	1	0
T>G	0	0	2	0	0	0	12	2	1	0	1	1	0	2	1	1

Supplementary Table 4b. Signature analysis of all 12 possible changes of the transcribed strand in various 5' (before colon) and 3' (after colon) sequence contexts for synonymous mutations in 250 genes in lung adenocarcinomas.

	A:A	A:C	A:G	A:T	C:A	C:C	C:G	C:T	G:A	G:C	G:G	G:T	T:A	T:C	T:G	T:T
A>C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A>G	0	0	0	0	0	1	1	0	0	1	0	0	0	0	1	1
A>T	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
C>A	0	0	1	0	3	2	2	2	0	0	0	0	0	2	1	0
C>G	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
C>T	4	3	2	0	2	4	2	2	0	2	4	1	5	3	3	1
G>A	1	0	6	3	4	0	1	0	0	1	3	0	2	2	4	1
G>C	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0
G>T	0	0	0	0	0	1	4	0	1	2	1	0	2	3	0	1
T>A	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
T>C	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
T>G	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Supplementary Table 4c. Signature analysis of all 12 possible changes of the transcribed strand in various 5' (before colon) and 3' (after colon) sequence contexts for 108 synonymous mutations and 156 non-synonymous point mutations from genes with only one mutation.

	A:A	A:C	A:G	A:T	C:A	C:C	C:G	C:T	G:A	G:C	G:G	G:T	T:A	T:C	T:G	T:T
A>C	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
A>G	0	0	1	2	0	3	2	2	0	1	0	1	0	1	1	3
A>T	1	2	0	0	2	0	4	3	0	1	0	0	0	0	0	0
C>A	0	1	2	0	3	3	3	3	1	0	0	0	0	4	2	1
C>G	0	0	0	0	0	0	0	1	0	0	1	0	2	2	0	3
C>T	5	4	3	1	4	8	5	2	0	4	6	1	8	3	5	4
G>A	3	0	8	3	9	0	3	1	6	2	3	0	5	3	5	5
G>C	1	1	1	0	2	2	1	0	2	1	1	1	2	0	2	0
G>T	2	2	3	1	3	6	7	2	5	7	3	3	2	4	2	1
T>A	0	0	0	1	1	0	1	0	1	0	1	0	0	0	0	0
T>C	1	0	0	0	0	1	1	0	0	0	0	1	0	1	0	0
T>G	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0

Supplementary Table 4d. Signature analysis of nonsynonymous mutations in individual genes.

Gene	A>C	A>G	A>T	C>A	C>G	C>T	G>A	G>C	G>T	T>A	T>C	T>G
RPS6KA6	0	0	0	0	0	1	0	0	1	0	0	0
HCK	0	0	1	1	0	1	0	0	1	0	0	0
PRKCE	0	0	0	0	0	0	2	1	0	0	0	0
ERBB4	1	2	2	1	1	0	0	0	2	0	0	0
AURKB	0	0	0	0	0	0	0	1	0	0	0	0
RAB6C	0	0	0	0	0	0	1	0	0	0	0	0
STK24	0	0	0	0	0	1	0	0	0	0	0	0
MAP4K1	0	0	0	0	0	1	0	1	1	0	0	0
JAG2	0	0	0	0	0	0	1	0	2	0	0	0
FH	0	0	0	0	0	0	1	0	0	0	0	0
ROCK1	0	1	0	0	0	0	0	0	0	0	0	0
ZAP70	0	0	0	0	0	1	0	0	0	0	0	0

STK10	0	0	0	0	0	0	0	1	0	0	0	0
TPTE	0	1	0	0	0	0	0	0	0	0	0	0
ACVR2B	0	0	0	0	0	0	0	0	1	0	0	0
MST1R	0	0	0	0	0	1	0	0	0	0	0	0
RYK	0	1	0	0	0	0	0	0	0	0	0	0
WT1	0	0	0	0	0	0	0	1	0	0	0	0
IRAK1	0	0	0	0	0	1	0	0	1	0	0	0
STK38L	0	0	0	0	0	0	1	0	0	0	0	0
FES	0	0	0	0	0	1	1	0	0	0	0	0
ACVR1B	0	0	0	0	0	0	1	0	0	0	0	0
KDR	0	0	0	1	0	1	1	1	3	0	0	0
MATK	1	0	0	0	0	0	0	0	0	0	0	0
KSR1	0	0	0	0	0	1	0	0	0	0	1	0
STAT2	0	0	0	0	0	0	0	1	0	0	0	0
CBLB	0	0	0	0	1	0	0	0	1	0	0	0
TP53	2	5	3	0	1	6	5	2	28	1	1	1
NTRK1	0	0	0	0	1	1	1	0	3	0	0	1
PTEN	0	0	0	0	0	2	1	0	0	0	1	0
ROBO2	0	0	0	1	0	0	1	0	1	0	1	0
MAST2	0	0	0	0	1	2	0	0	0	0	0	0
PDGFRB	0	0	0	0	1	2	0	0	0	0	0	0
STYK1	0	1	0	0	0	0	0	0	0	0	0	0
LZTR1	0	0	1	0	0	1	0	0	0	0	0	0
EPHA3	0	0	0	3	0	0	1	1	4	1	0	1
PRKCG	0	0	0	1	1	0	0	1	0	1	0	0
JUP	0	0	0	0	0	2	0	0	0	0	0	0
STK11	0	1	3	1	1	6	2	0	7	0	0	0
CRKRS	0	0	0	0	0	0	0	1	1	0	0	0
FLT1	0	0	0	2	0	0	0	1	0	0	0	0
INSR	0	0	0	0	0	0	1	0	2	0	0	0
TGFBR2	0	0	1	0	0	0	0	0	0	0	0	0

RAPGEF1	0	0	0	1	0	0	0	0	0	0	0	0
LCK	0	0	0	0	0	1	0	0	0	0	0	0
MEN1	0	0	0	0	0	0	0	0	1	0	0	0
STK32C	0	0	0	0	0	0	1	0	0	0	0	0
CDC2L5	0	0	0	0	2	0	0	0	0	0	0	0
RAF1	0	0	0	0	0	0	0	0	1	0	0	0
MST4	0	0	0	0	0	0	0	0	1	0	0	0
PDK1	0	0	0	0	0	0	1	0	0	0	0	0
ERG	0	0	0	1	0	0	0	0	1	0	0	0
LMTK3	0	0	0	0	0	0	0	2	2	0	0	0
SRMS	0	0	0	0	0	0	1	0	0	0	0	0
LATS2	0	0	0	0	0	0	1	0	0	0	0	0
PIK4CA	0	0	0	0	0	0	1	0	0	0	0	0
MKNK2	0	0	1	1	0	0	0	0	0	0	0	0
SMAD2	0	0	0	0	1	0	0	0	0	0	0	0
SLA	0	0	0	0	0	0	1	0	0	0	0	0
P2RXL1	0	0	0	0	0	0	0	0	1	0	0	0
FYN	0	0	0	0	0	0	1	0	1	0	0	0
FGFR1	0	0	0	0	1	0	1	0	0	0	0	0
MET	0	0	0	0	0	0	0	0	1	1	0	0
MRAS	0	0	0	0	0	0	1	0	0	0	0	0
HD	0	0	0	0	0	0	0	0	1	0	0	0
MAP3K3	0	0	0	0	0	1	0	0	1	0	0	0
AXL	0	0	0	0	0	0	0	0	0	1	0	0
CDK5	0	0	0	0	0	1	0	0	0	0	0	0
ERAS	0	0	0	1	0	0	0	0	1	0	0	0
VEGFC	1	0	0	0	0	0	0	0	1	0	0	0
RPS6KA1	0	0	0	0	0	0	0	0	1	0	0	0
PTPRD	0	1	0	2	2	0	0	0	2	2	0	1
BRCA2	0	0	0	0	0	0	0	1	1	0	0	0
MUSK	0	0	0	0	1	0	0	0	0	0	0	0

MYO3B	1	1	0	1	0	0	0	0	3	1	0	0
CDK9	0	0	0	0	0	0	1	0	0	0	0	0
PRKCQ	0	0	0	0	0	0	0	0	1	0	0	0
PLCG2	0	0	0	0	0	0	0	0	2	0	0	0
PIK3C2B	0	0	0	0	0	0	1	0	0	0	0	0
CDC42BPG	0	0	0	0	0	0	1	0	1	0	0	0
JAK3	0	0	1	1	0	0	0	0	1	0	0	0
PAK4	0	0	0	0	1	0	0	0	2	0	0	0
JAK1	0	0	0	0	0	0	1	1	0	0	0	0
CAMKV	0	0	0	0	0	0	1	0	1	0	0	0
FGFR2	0	1	0	1	0	0	1	0	0	0	0	0
SHC1	0	0	0	0	0	0	0	0	0	0	1	0
TAOK3	0	0	0	0	0	0	0	0	1	0	0	0
KIT	0	0	0	0	0	1	0	0	0	0	0	0
PRKAR1A	0	0	1	0	0	0	0	0	0	0	0	0
BRCA1	0	0	0	0	0	0	0	0	1	0	0	0
STK36	0	0	0	0	0	0	0	1	1	0	0	0
EPHA7	0	0	0	1	1	1	0	0	2	0	0	0
PKN3	0	0	0	0	0	0	0	1	1	0	0	0
ACVRL1	0	0	0	0	0	1	0	0	0	0	0	0
INHBA	0	0	0	1	0	0	0	0	3	0	0	0
IRAK2	0	0	0	0	0	0	1	0	1	0	0	0
MYO3A	0	0	1	0	0	0	0	0	2	0	0	0
AKT1	0	0	0	0	0	0	1	0	0	0	1	0
PRKCH	0	0	0	0	0	0	0	0	1	0	0	0
NTRK2	0	1	0	2	0	0	1	1	1	0	0	0
JAK2	0	1	0	0	0	1	0	1	0	0	1	0
GPC3	0	0	0	0	0	0	1	0	0	0	0	0
CDKN2C	0	0	0	0	0	0	1	0	0	0	0	0
MAPK6	0	0	0	0	1	1	0	0	1	0	0	0
SMO	0	0	0	1	0	0	0	0	0	0	0	0

PDGFRA	0	0	0	1	1	1	2	1	0	1	0	0
BMPR1A	0	0	0	0	0	1	0	0	0	0	0	0
MAP2K5	0	0	0	0	0	0	0	0	2	0	0	0
LOC400891	0	0	0	1	0	0	0	0	0	0	0	0
STAT3	0	0	0	0	0	0	0	0	1	0	0	0
STK32A	0	0	1	0	0	0	0	0	0	0	0	0
LRP1B	0	0	1	0	1	4	1	1	8	1	0	0
EPHA2	0	0	0	0	0	0	0	0	1	0	0	0
PAK7	0	0	0	3	1	0	0	0	0	0	0	0
ROBO1	0	0	0	0	1	0	0	2	0	0	0	0
MAML2	0	1	0	0	0	0	0	0	0	0	0	0
KIAA1804	0	1	0	0	1	0	0	0	3	0	0	0
EPHB6	0	0	0	2	0	0	1	0	2	0	0	0
TEK	0	0	0	1	0	0	0	0	1	0	0	0
LRRK1	0	0	0	1	0	0	1	0	0	0	0	0
TSC1	0	0	0	0	0	1	0	0	1	0	0	0
BARD1	0	0	0	0	0	0	1	0	1	0	0	0
PRKACA	0	0	0	0	0	0	1	0	0	0	0	0
HEMK1	0	0	1	0	0	0	0	0	0	0	0	0
RIPK3	0	0	0	0	0	1	0	0	0	0	0	0
PRKG1	0	0	0	0	0	0	1	0	0	0	0	0
TNK2	0	0	0	0	1	0	0	0	0	0	0	0
PTPN11	0	0	0	1	0	1	0	0	1	0	0	0
PRKX	0	0	1	0	0	0	0	0	0	0	0	0
NF1	0	1	1	0	0	1	0	3	7	0	0	1
TUSC4	1	0	0	0	0	0	0	0	0	0	0	0
SLC7A4	0	0	0	1	0	0	0	0	0	0	0	0
PIK3CG	0	0	1	1	1	0	0	0	2	0	0	0
PLCG1	0	0	0	0	0	0	1	0	0	0	0	0
NLK	0	1	0	0	0	0	0	0	0	0	0	0
MDM2	0	0	0	0	0	0	0	0	1	0	0	1

GNAS	0	0	1	1	1	1	1	0	1	0	0	0
KSR2	0	0	1	0	0	0	1	0	0	0	0	0
CYB561D2	0	0	0	1	0	0	0	0	0	0	0	0
MLL	0	0	1	0	0	0	1	0	0	0	0	0
MAP4K3	0	0	0	0	0	0	0	0	2	0	0	0
PFTK1	0	0	0	0	0	0	1	1	1	0	0	0
CTNNB1	0	0	0	0	0	1	1	0	0	0	0	0
AIFM3	0	0	1	0	0	0	1	0	0	0	0	0
MKNK1	0	0	2	0	0	0	0	0	0	0	0	0
MYCL1	0	0	0	1	0	0	0	1	0	0	0	0
EGFR	0	0	0	1	0	0	3	0	1	1	0	14
PAK3	0	0	1	2	0	0	0	1	1	0	0	0
IKBKB	0	0	0	0	0	0	0	0	1	0	0	0
STK3	0	0	0	0	0	0	0	1	0	0	0	0
MAP3K15	0	0	0	0	0	1	0	0	1	0	0	0
MAML1	0	0	0	0	0	1	0	0	2	0	0	0
BCL3	0	0	0	0	1	0	0	0	0	0	0	0
TYK2	0	0	0	0	1	1	0	0	0	0	0	0
SEMA3F	0	0	0	0	0	1	0	0	0	0	0	0
TFDP1	0	0	1	0	0	0	1	1	1	0	0	0
RIPK4	0	0	0	0	1	0	0	0	0	0	0	0
PCTK2	0	0	0	0	1	0	0	0	0	1	0	0
BMPR1B	0	0	0	0	1	0	0	0	1	0	0	0
KIAA1303	0	0	0	0	0	0	1	0	1	0	0	0
CCNE1	0	0	0	0	0	1	0	0	0	0	0	0
ACVR1C	0	0	0	0	0	0	0	0	0	1	0	0
CNTN4	0	0	0	0	0	0	0	0	0	1	0	0
ROR2	0	0	0	1	1	1	0	1	1	0	0	0
SGK2	0	0	0	0	0	0	1	0	0	0	0	0
PIK3CA	0	0	0	0	0	1	0	0	0	0	0	0
PIK3C2G	0	0	0	0	1	0	1	0	1	0	1	0

NOTCH2	0	0	0	0	0	0	0	0	1	0	0	0
BMX	0	0	0	1	0	0	0	0	0	0	0	0
PRKDC	0	0	1	0	0	0	2	1	2	0	0	0
MAP3K2	0	0	0	0	0	1	0	0	0	0	0	0
ALK	0	0	0	0	1	1	0	0	1	1	0	0
MINK1	0	0	0	0	0	0	0	0	1	0	0	0
MSH6	0	0	0	0	0	1	1	1	0	0	1	0
GRB7	0	0	0	0	0	0	0	1	1	0	0	0
ATM	1	0	0	1	0	1	1	2	3	1	2	0
ERBB3	0	0	0	0	0	1	2	0	0	0	0	0
ERBB2	0	0	0	0	0	1	2	0	0	0	0	0
TRAIP	0	0	0	0	0	0	0	0	1	0	0	0
SYK	0	0	0	0	0	0	0	1	0	0	0	0
GSK3A	0	1	0	0	0	1	0	0	0	0	0	0
MAPK7	0	0	0	0	0	0	1	0	0	0	0	0
INSRR	0	0	1	2	1	0	0	0	0	0	1	1
PRKCI	0	1	0	0	0	0	0	0	0	0	0	0
DDR2	0	1	0	0	0	0	0	0	0	0	0	0
PTK2	0	0	1	0	0	0	0	0	0	0	0	1
CYSLTR2	0	0	0	0	0	0	1	0	1	0	0	0
FOS	0	0	0	0	0	0	0	0	0	0	1	0
GAK	0	0	0	0	0	0	0	0	1	0	0	0
MAP3K10	0	0	0	0	0	0	0	1	1	0	0	0
PAK6	0	0	0	0	0	0	1	0	0	0	0	0
TP73L	0	0	0	1	0	0	0	0	3	0	0	0
JUNB	0	0	0	0	0	0	0	0	1	0	0	0
DDR1	0	0	0	0	0	0	0	0	1	1	0	0
CHEK1	0	0	0	0	0	1	0	0	0	0	0	0
SLC38A3	0	0	2	0	0	1	0	0	0	0	0	0
RPS6KA4	0	0	0	0	0	0	0	1	0	0	0	0
MAP4K4	0	0	0	0	0	1	0	0	0	0	0	0

EVI1	0	0	1	0	0	0	0	0	0	0	0	0
DOCK3	0	0	0	1	1	0	0	2	0	1	1	0
LATS1	0	0	0	1	0	1	0	0	1	0	0	0
RHOB	0	0	0	0	0	0	0	0	1	0	0	0
RALA	0	0	0	0	0	0	0	0	1	0	0	0
GRK7	0	0	1	0	0	0	0	0	1	0	0	0
LTK	0	0	0	2	0	1	3	0	0	0	0	0
RBL2	0	0	0	0	0	1	1	0	0	0	0	0
ITK	0	0	0	1	0	0	0	0	2	1	0	0
ALS2CR2	0	0	1	0	0	0	0	0	0	0	0	0
CCNG1	0	0	0	0	0	1	0	0	0	0	0	0
PFTK2	0	0	0	1	0	0	0	0	0	0	0	0
HRAS	0	0	0	0	0	0	0	0	0	0	1	0
PRKACB	0	0	0	0	0	0	0	0	3	0	0	0
STAT5B	0	0	0	0	0	1	0	0	1	0	0	0
ARAF	0	0	0	0	0	0	1	0	0	0	0	0
EPHA6	0	0	1	0	0	0	0	0	0	0	0	0
FOXO3	0	0	1	0	0	0	0	0	1	0	0	0
SKI	0	0	0	0	0	0	0	1	0	0	0	0
APC	0	0	1	1	1	4	1	0	1	0	0	0
SMAD4	0	0	0	2	0	0	0	1	0	0	0	0
BCL6	0	0	0	0	0	1	0	0	1	0	0	0
CDKN2B	0	0	0	0	0	1	0	0	0	0	0	0
FNDC3A	0	0	0	0	1	0	1	0	0	0	0	0
MAP3K6	0	0	0	0	2	0	0	0	0	0	0	0
FGFR4	0	0	0	1	2	0	1	0	1	0	0	0
IFRD2	0	0	0	1	0	1	0	0	0	0	0	0
XRCC1	0	0	1	0	0	0	0	0	0	0	0	0
STAT6	0	1	0	0	0	0	0	0	0	0	0	0
PIK3C3	0	0	0	0	0	0	0	1	4	1	0	0
SEMA3B	0	0	0	0	0	1	0	0	1	0	0	0

GCK	0	0	0	1	0	0	0	0	0	0	0	0
CDH11	0	1	1	1	0	1	0	0	2	0	0	0
EPHX1	0	0	0	0	0	0	0	0	1	0	0	0
PIK3CD	0	0	0	1	0	0	1	0	0	0	0	0
RASSF2	0	0	0	0	0	0	0	0	2	0	0	0
CDKN2A	0	0	1	0	0	1	2	2	1	0	0	0
CDC42BPA	0	0	2	0	1	0	1	1	0	0	0	0
VAV2	0	0	0	0	0	0	0	0	2	0	0	0
MAPK8	0	0	0	0	0	1	0	0	0	0	0	0
PTCH1	0	1	0	1	0	0	0	1	1	0	0	0
IGF1R	0	0	0	0	1	0	0	0	1	0	0	0
KLHL22	0	0	0	0	0	1	0	0	0	0	0	0
RIPK1	0	0	0	0	0	1	0	0	0	0	0	0
MAST4	0	0	0	0	0	0	0	1	1	0	0	0
PRKD1	0	0	0	0	0	0	0	0	1	0	0	0
IRS1	0	0	0	0	0	0	0	1	2	0	0	0
TSHR	0	0	0	1	0	0	1	0	1	0	0	0
MYC	0	0	0	0	0	0	0	0	1	0	0	0
SRPK3	0	0	0	1	0	0	0	0	0	0	0	0
ATR	0	0	0	0	0	0	1	0	0	0	0	0
CSF1R	0	0	0	1	0	0	1	0	0	0	0	0
NRK	0	0	0	0	0	1	1	0	0	0	0	0
GRB14	0	0	0	0	0	0	0	0	1	0	0	0
NOTCH3	0	0	0	0	0	2	0	0	0	0	0	0
NTRK3	0	0	1	4	0	0	1	0	0	1	0	0
CCNT2	0	0	0	0	1	0	0	0	0	0	0	0
ABL1	0	0	0	0	0	1	0	0	1	0	0	0
PTPRG	0	0	0	0	0	0	1	0	0	0	0	0
MST1	0	0	0	0	0	0	0	0	0	0	1	0
NF2	0	0	1	0	0	0	0	0	0	0	0	0
LRRK2	0	0	0	1	0	0	1	0	0	0	0	0

PRKD3	0	0	0	0	0	2	0	0	0	0	0	0
TYRO3	0	0	0	1	0	0	0	0	1	0	0	0
TSC2	0	0	0	1	0	0	1	0	1	0	0	0
TNNI3K	0	0	0	0	0	0	0	1	0	0	0	0
MAPK4	0	0	0	1	0	0	1	0	0	0	0	0
FLT4	0	0	0	0	1	0	0	1	1	0	2	0
CDH4	0	0	0	0	0	0	1	0	1	0	0	0
LTF	0	0	0	0	0	0	0	0	0	1	0	0
EPHA5	0	0	0	1	1	0	2	0	3	0	0	0
MAP3K12	0	0	0	0	0	1	1	0	0	0	0	0
BUB1	0	0	0	0	0	0	0	0	1	0	0	0
YES1	0	0	0	0	0	0	0	0	2	0	0	0
ETS1	0	0	0	0	0	0	1	0	0	0	0	0
CASK	0	0	0	0	0	0	0	1	0	0	0	0
KRAS	0	0	1	0	0	0	11	8	38	0	0	0
SMG1	0	1	0	0	0	1	0	0	1	0	0	0
MYCN	0	0	0	0	0	1	0	0	0	0	0	0
BRAF	0	0	0	0	1	0	0	0	2	0	0	0
RBM5	0	0	0	0	0	0	1	0	0	0	0	0
TEC	0	0	0	0	1	0	0	0	2	0	0	0
FER	0	0	0	0	0	0	0	0	1	0	0	0
CCNB3	0	0	0	0	0	0	0	0	1	0	0	0
NOTCH1	0	0	0	0	0	0	1	0	1	0	0	0
RALB	0	0	0	0	0	0	0	0	1	0	0	0
GSK3B	0	0	1	0	0	0	1	0	0	0	0	0
CDK7	0	0	0	0	0	0	0	1	0	0	0	0
BAX	0	0	0	0	0	1	0	0	0	0	0	0
ANKK1	0	0	0	0	0	0	0	0	1	0	0	0
MERTK	0	1	1	0	0	2	0	0	0	0	0	0
AKT3	0	0	1	0	0	0	0	0	0	0	0	0
VAV3	0	0	0	1	1	0	0	0	0	0	0	0

DMPK	0	0	0	0	0	1	1	0	0	0	0	0
FRK	0	0	0	0	0	0	1	0	0	0	0	0
PINX1	0	0	0	0	0	0	0	0	1	0	0	0
EPHA4	0	0	0	0	0	0	0	0	1	0	0	0
TERT	0	0	0	0	0	1	2	1	1	0	0	0
IKBKE	0	0	0	0	0	0	0	1	0	0	0	0
RELA	0	0	0	0	0	0	1	0	0	0	0	0
PASK	0	0	0	0	0	0	0	0	1	0	0	0
TP73	0	0	0	0	0	0	0	0	1	0	0	0
PRKCD	0	0	0	0	0	0	0	0	1	0	0	0
RANBP9	0	0	0	0	0	0	0	1	0	0	0	0
FEZF2	0	0	0	1	0	0	0	0	0	0	0	0
EPHB4	0	0	0	0	0	0	1	0	1	0	0	0
MYBL2	0	0	0	1	0	0	0	0	1	0	0	0
CNTN6	0	0	0	0	0	0	0	0	1	0	0	0
RAP2B	0	0	0	0	0	0	0	0	1	0	0	0
RPSA	0	0	0	0	0	1	0	0	0	0	0	0
RET	0	0	1	1	1	0	0	0	0	0	0	0
MASTL	0	0	0	0	0	0	1	0	0	0	0	0
ZAK	0	0	0	1	1	0	0	0	1	0	0	0
ZMYND10	0	1	0	0	0	1	1	0	0	0	1	0
EPHB1	0	0	1	0	1	0	0	1	2	0	0	0
VAV1	0	0	0	0	0	0	1	0	0	0	1	0
RB1	0	0	0	0	0	1	2	0	2	0	1	0
RASEF	0	0	0	0	0	1	0	0	0	0	0	0
STAT5A	0	1	0	0	0	0	0	0	0	0	0	0
RBL1	0	0	0	0	0	0	1	0	1	0	1	0
NRAS	0	0	3	0	0	0	0	0	0	0	0	0
AURKA	0	0	0	0	0	1	0	0	0	0	0	0
CROT	0	1	0	0	0	0	0	0	0	0	0	0
LMTK2	1	0	0	0	0	0	0	0	3	0	0	0

VHL	0	0	0	0	0	0	1	0	0	0	0	0
BAP1	0	0	0	0	0	1	0	0	1	0	0	0
MAP3K9	0	0	0	0	0	0	0	1	0	0	0	0
RCBTB2	0	0	0	0	0	0	0	0	1	0	0	0
PRKCB1	0	0	0	0	0	0	1	0	1	0	0	0
MAP3K4	0	0	0	0	0	0	0	0	3	0	0	0
MAST1	0	0	0	1	0	0	1	0	0	0	0	0
EPHA1	0	0	0	0	0	0	1	0	0	0	0	0
EIF4G1	0	0	0	0	0	1	0	0	0	0	0	0
SRC	0	0	0	0	0	0	0	0	1	0	0	0
FBXW7	0	1	0	0	0	0	0	0	1	0	0	0
GATA1	0	0	0	1	1	0	0	0	0	0	0	0
ERCC2	0	0	0	0	0	0	1	0	0	0	0	0
PTCH2	0	0	0	0	0	0	0	0	2	0	0	0
HYAL1	0	0	0	0	0	0	1	0	0	0	0	0
PHOX2B	0	0	0	0	0	0	0	0	0	0	0	1
CDC2L2	0	0	0	0	0	0	0	1	0	0	0	0
ADRBK2	0	0	0	0	0	1	0	0	0	0	0	0
CACNA2D2	0	0	0	0	0	0	0	0	1	0	0	0
KLF6	0	1	0	0	0	0	0	0	0	0	0	0
NOTCH4	0	0	0	0	0	0	0	0	1	0	0	0
PIK3C2A	0	1	0	0	0	1	0	0	0	0	0	0
MYB	0	0	0	0	0	0	1	0	1	0	0	0
CDC25A	0	0	0	0	0	0	1	0	1	0	0	0
RIN1	0	0	0	0	0	1	0	0	0	0	0	0
PKN2	0	0	0	0	1	0	0	0	0	0	0	0
total	9	35	56	83	55	111	134	71	294	22	22	23

Supplementary Table 5a. Collapsed categories shown in Supplementary Table 4b based on using Fisher's exact test ($P < 0.05$). A total of 18 categories were considered to be statistically distinct.

AC:ACGT	A:T>C:G:0	A:T>G:C:12	A:T>T:A:15
GT:ACGT	A:T>C:G:3	A:T>G:C:10	A:T>T:A:3
CG:ACGT	C:G>A:T:47	C:G>G:C:11	C:G>T:A:54
T:ACGT	C:G>A:T:19	C:G>G:C:14	C:G>T:A:43
A:AT	C:G>A:T:2	C:G>G:C:0	C:G>T:A:14
A:CG	C:G>A:T:8	C:G>G:C:1	C:G>T:A:8

Supplementary Table 5b. Gene-based background mutation rates.

Gene	Background mutation rate
AADAC	3.82E-06
AATK	2.41E-06
ABL1	2.81E-06
ABL2	3.45E-06
ACTR2	3.77E-06
ACVR1	3.31E-06
ACVR1B	2.92E-06
ACVR1C	3.48E-06
ACVR2B	2.64E-06
ACVRL1	2.61E-06
ADRBK1	2.55E-06
ADRBK2	3.39E-06
AIFM3	2.78E-06
AKT1	2.78E-06
AKT2	2.76E-06
AKT3	4.04E-06
ALK	2.81E-06

ALS2CR2	3.72E-06
AMHR2	3.03E-06
ANKK1	2.72E-06
APC	4.11E-06
ARAF	2.81E-06
ARMET	2.86E-06
ATM	3.94E-06
ATR	3.85E-06
AURKA	3.63E-06
AURKB	2.88E-06
AURKC	2.94E-06
AXL	2.90E-06
BAD	2.83E-06
BAGE3	2.83E-06
BAGE4	2.83E-06
BAP1	2.89E-06
BARD1	3.68E-06
BAX	2.65E-06
BCL2	2.43E-06
BCL3	2.46E-06
BCL6	3.07E-06
BLK	2.73E-06
BMPR1A	3.63E-06
BMPR1B	3.69E-06
BMPR2	3.77E-06
BMX	3.47E-06
BRAF	3.26E-06
BRAP	3.12E-06
BRCA1	3.67E-06
BRCA2	4.05E-06
BTK	3.48E-06
BUB1	3.81E-06
BUB1B	3.65E-06

C3orf14	4.56E-06
C3orf18	2.79E-06
CACNA2D2	2.69E-06
CADPS	3.07E-06
CAMKV	3.10E-06
CASK	3.56E-06
CBL	3.15E-06
CBLB	3.76E-06
CBLC	2.63E-06
CCNA1	3.26E-06
CCNA2	3.25E-06
CCNB3	3.09E-06
CCND1	2.42E-06
CCND2	2.70E-06
CCND3	2.60E-06
CCNE1	3.03E-06
CCNE2	3.99E-06
CCNG1	3.42E-06
CCNG2	3.49E-06
CCNH	3.52E-06
CCNK	3.28E-06
CCNT1	3.68E-06
CCNT2	3.69E-06
CCRK	2.63E-06
CDC2	3.93E-06
CDC25A	3.08E-06
CDC25B	2.78E-06
CDC25C	3.30E-06
CDC2L2	2.65E-06
CDC2L5	3.18E-06
CDC2L6	3.27E-06
CDC42BPA	3.77E-06
CDC42BPB	2.84E-06

CDC42BPG	2.60E-06
CDC7	3.92E-06
CDH1	3.10E-06
CDH11	3.12E-06
CDH4	2.63E-06
CDK10	2.52E-06
CDK2	3.14E-06
CDK3	2.81E-06
CDK4	3.24E-06
CDK5	2.64E-06
CDK6	2.99E-06
CDK7	3.63E-06
CDK8	3.55E-06
CDK9	2.80E-06
CDKN1A	3.05E-06
CDKN1B	2.76E-06
CDKN1C	2.16E-06
CDKN2A	2.88E-06
CDKN2B	2.74E-06
CDKN2C	2.94E-06
CDKN2D	2.69E-06
CDKN3	3.49E-06
CEBPA	2.11E-06
CHEK1	3.97E-06
CHEK2	3.44E-06
CISH	2.94E-06
CNKSR1	2.80E-06
CNTN4	3.65E-06
CNTN6	3.84E-06
CRK	2.70E-06
CRKL	2.96E-06
CRKRS	3.42E-06
CROT	3.98E-06

CSF1R	2.85E-06
CSK	2.77E-06
CTNNB1	3.89E-06
CYB561D2	2.58E-06
CYLD	3.62E-06
CYSLTR2	3.09E-06
DDR1	2.77E-06
DDR2	3.37E-06
DHH	2.48E-06
DIRAS3	2.62E-06
DKFZP434N035	3.70E-06
DMPK	2.51E-06
DOCK3	3.08E-06
DUSP24	3.20E-06
DUSP6	2.44E-06
E2F1	2.40E-06
E2F2	2.55E-06
E2F3	2.65E-06
E2F4	3.01E-06
E2F5	2.95E-06
E2F6	2.84E-06
EGFR	2.88E-06
EIF4E	3.49E-06
EIF4EBP1	2.75E-06
EIF4G1	3.17E-06
EPHA1	2.87E-06
EPHA2	2.70E-06
EPHA3	3.53E-06
EPHA4	3.34E-06
EPHA5	3.54E-06
EPHA6	3.44E-06
EPHA7	3.86E-06
EPHA8	2.61E-06

EPHB1	3.02E-06
EPHB2	2.74E-06
EPHB3	2.75E-06
EPHB4	2.68E-06
EPHB6	2.86E-06
EPHX1	2.75E-06
EPM2A	2.77E-06
ERAS	2.67E-06
ERBB2	2.74E-06
ERBB3	2.89E-06
ERBB4	3.57E-06
ERCC2	2.61E-06
ERG	3.05E-06
ETS1	3.16E-06
ETS2	2.94E-06
ETV6	2.94E-06
ETV7	2.85E-06
EVI1	3.36E-06
FAS	3.46E-06
FBXW7	3.95E-06
FER	3.86E-06
FES	2.62E-06
FEV	2.18E-06
FEZF2	2.29E-06
FGF3	2.49E-06
FGF4	2.21E-06
FGF5	3.23E-06
FGFR1	2.69E-06
FGFR2	3.10E-06
FGFR3	2.37E-06
FGFR4	2.56E-06
FGR	2.65E-06
FH	3.61E-06

FHIT	3.14E-06
FKBP1A	3.18E-06
FLJ25006	3.12E-06
FLJ38608	2.62E-06
FLT1	3.38E-06
FLT3	3.23E-06
FLT4	2.61E-06
FNDC3A	4.51E-06
FOS	2.60E-06
FOSB	2.47E-06
FOSL2	2.57E-06
FOXO1A	2.67E-06
FOXO3	2.60E-06
FOXP1	3.25E-06
FRK	3.46E-06
FYN	3.35E-06
GAK	2.53E-06
GATA1	2.95E-06
GBL	2.59E-06
GCK	2.65E-06
GGT2	2.52E-06
GNAI2	2.64E-06
GNAS	3.14E-06
GNAT1	2.55E-06
GPC3	3.20E-06
GRB10	2.90E-06
GRB14	3.19E-06
GRB2	3.13E-06
GRB7	2.90E-06
GRK1	2.38E-06
GRK4	3.29E-06
GRK5	2.80E-06
GRK6	2.66E-06

GRK7	2.96E-06
GSK3A	2.54E-06
GSK3B	3.36E-06
HCK	2.88E-06
HD	2.95E-06
HEMK1	3.00E-06
HHIP	3.40E-06
HIC2	2.69E-06
HIF1A	4.05E-06
HRAS	2.55E-06
HRPT2	4.01E-06
HYAL1	3.04E-06
HYAL2	3.17E-06
HYAL3	2.90E-06
IFRD2	2.90E-06
IGF1R	2.84E-06
IHH	2.67E-06
IKBKB	2.89E-06
IKBKE	2.68E-06
ILK	3.33E-06
INHBA	2.76E-06
INSR	2.71E-06
INSRR	2.83E-06
IRAK1	2.53E-06
IRAK2	2.85E-06
IRAK3	3.37E-06
IRAK4	4.25E-06
IRS1	2.92E-06
IRS2	2.18E-06
ITK	3.30E-06
JAG1	2.88E-06
JAG2	2.52E-06
JAK1	3.21E-06

JAK2	3.99E-06
JAK3	2.72E-06
JUN	2.30E-06
JUNB	2.38E-06
JUND	2.12E-06
JUP	2.62E-06
KDR	3.42E-06
KIAA1303	2.65E-06
KIAA1804	2.75E-06
KIT	3.35E-06
KLF6	2.74E-06
KLHL22	2.74E-06
KRAS	3.88E-06
KSR1	2.73E-06
KSR2	2.71E-06
LATS1	4.28E-06
LATS2	2.65E-06
LCK	2.74E-06
LIMD1	3.00E-06
LIMK1	2.70E-06
LIMK2	2.84E-06
LMTK2	2.80E-06
LMTK3	2.36E-06
LOC375133	2.60E-06
LOC400891	2.85E-06
LRP1B	3.81E-06
LRRK1	2.67E-06
LRRK2	3.63E-06
LTF	2.96E-06
LTK	2.62E-06
LYK5	2.96E-06
LYN	3.40E-06
LZTR1	2.61E-06

MAF	2.12E-06
MAFA	2.12E-06
MAFB	2.26E-06
MAFF	2.30E-06
MAFG	2.42E-06
MAFK	2.56E-06
MAML1	2.12E-06
MAML2	3.31E-06
MAML3	3.29E-06
MAP2K1	2.98E-06
MAP2K2	2.43E-06
MAP2K4	3.40E-06
MAP2K5	3.27E-06
MAP2K6	3.36E-06
MAP2K7	2.52E-06
MAP3K1	3.43E-06
MAP3K10	2.50E-06
MAP3K11	2.72E-06
MAP3K12	3.07E-06
MAP3K13	3.41E-06
MAP3K14	2.89E-06
MAP3K15	2.99E-06
MAP3K2	3.86E-06
MAP3K3	3.09E-06
MAP3K4	3.52E-06
MAP3K5	3.32E-06
MAP3K6	2.60E-06
MAP3K7	3.75E-06
MAP3K8	3.32E-06
MAP3K9	2.85E-06
MAP4K1	2.74E-06
MAP4K3	3.91E-06
MAP4K4	3.39E-06

MAP4K5	3.80E-06
MAPK1	3.03E-06
MAPK10	3.25E-06
MAPK11	2.53E-06
MAPK12	2.51E-06
MAPK13	2.57E-06
MAPK14	3.19E-06
MAPK15	2.60E-06
MAPK3	2.44E-06
MAPK4	2.41E-06
MAPK6	3.78E-06
MAPK7	3.01E-06
MAPK8	3.74E-06
MAPK9	3.76E-06
MAPKAPK3	2.92E-06
MAST1	2.57E-06
MAST2	3.09E-06
MAST3	2.63E-06
MAST4	2.29E-06
MASTL	3.61E-06
MATK	2.62E-06
MDM2	3.54E-06
MDM4	3.65E-06
MEN1	2.62E-06
MERTK	3.17E-06
MET	3.86E-06
MGC16703	2.33E-06
MINK1	2.88E-06
MKNK1	3.04E-06
MKNK2	2.66E-06
MLH1	3.31E-06
MLKL	3.20E-06
MLL	3.43E-06

MON1A	2.75E-06
MOS	2.59E-06
MPL	2.98E-06
MRAS	2.99E-06
MSH2	3.59E-06
MSH6	3.66E-06
MST1	2.85E-06
MST1R	2.91E-06
MST4	3.55E-06
MUSK	3.59E-06
MYB	3.43E-06
MYBL1	4.06E-06
MYBL2	2.80E-06
MYC	2.54E-06
MYCL1	2.57E-06
MYCN	2.47E-06
MYO3A	3.74E-06
MYO3B	3.49E-06
MYT1	2.86E-06
NCK1	4.29E-06
NCK2	2.51E-06
NCYM	3.19E-06
NF1	3.60E-06
NF2	3.00E-06
NFKB1	3.44E-06
NLK	3.42E-06
NOTCH1	2.46E-06
NOTCH2	3.27E-06
NOTCH2NL	3.39E-06
NOTCH3	2.69E-06
NOTCH4	2.87E-06
NRAS	3.59E-06
NRK	3.69E-06

NTRK1	2.66E-06
NTRK2	3.12E-06
NTRK3	3.11E-06
OSR1	2.65E-06
P2RXL1	2.76E-06
P2RY5	3.63E-06
PAK1	3.32E-06
PAK2	3.62E-06
PAK3	3.88E-06
PAK4	2.48E-06
PAK6	2.72E-06
PAK7	2.84E-06
PASK	2.76E-06
PCQAP	2.76E-06
PCTK1	2.88E-06
PCTK2	3.68E-06
PCTK3	2.68E-06
PDGFB	2.76E-06
PDGFRA	3.26E-06
PDGFRB	2.76E-06
PDK1	3.24E-06
PEBP1	2.67E-06
PFTK1	3.42E-06
PFTK2	3.37E-06
PHOX2B	2.51E-06
PIK3C2A	3.97E-06
PIK3C2B	2.80E-06
PIK3C2G	3.71E-06
PIK3C3	3.77E-06
PIK3CA	3.94E-06
PIK3CB	3.70E-06
PIK3CD	2.54E-06
PIK3CG	3.03E-06

PIK3R1	3.55E-06
PIK3R2	2.48E-06
PIK3R3	3.65E-06
PIK3R4	3.83E-06
PIK4CA	2.62E-06
PIM1	2.53E-06
PIM2	2.93E-06
PIM3	2.33E-06
PINX1	3.32E-06
PKN1	2.52E-06
PKN2	3.99E-06
PKN3	2.70E-06
PL6	2.66E-06
PLCG1	2.86E-06
PLCG2	2.88E-06
PRKAB1	2.79E-06
PRKACA	2.75E-06
PRKACB	3.86E-06
PRKACG	2.62E-06
PRKAR1A	3.71E-06
PRKCA	3.10E-06
PRKCB1	3.04E-06
PRKCD	2.84E-06
PRKCE	2.95E-06
PRKCG	2.77E-06
PRKCH	3.06E-06
PRKCI	3.92E-06
PRKCQ	3.07E-06
PRKCZ	2.64E-06
PRKD1	3.25E-06
PRKD3	4.09E-06
PRKDC	3.29E-06
PRKG1	3.31E-06

PRKG2	3.40E-06
PRKX	2.55E-06
PRKY	2.50E-06
PTCH1	2.85E-06
PTCH2	2.68E-06
PTEN	3.69E-06
PTK2	3.55E-06
PTK2B	2.79E-06
PTK7	2.91E-06
PTPN11	3.57E-06
PTPRD	3.64E-06
PTPRG	3.23E-06
RAB43	2.48E-06
RAB6C	3.30E-06
RAF1	3.33E-06
RAGE	2.83E-06
RALA	3.58E-06
RALB	3.17E-06
RAN	3.39E-06
RANBP9	2.96E-06
RAP1A	3.26E-06
RAP1B	3.89E-06
RAP2A	2.57E-06
RAP2B	2.23E-06
RAP2C	3.36E-06
RAPGEF1	2.70E-06
RARB	3.51E-06
RASEF	2.48E-06
RASSF1	2.79E-06
RASSF2	2.98E-06
RASSF3	2.88E-06
RB1	3.69E-06
RBL1	3.77E-06

RBL2	3.94E-06
RBM5	3.42E-06
RBM6	3.73E-06
RCBTB2	3.28E-06
REEP5	2.92E-06
REL	3.93E-06
RELA	2.89E-06
RELB	2.41E-06
RET	2.56E-06
RHEB	3.34E-06
RHOA	3.42E-06
RHOB	2.49E-06
RHOBTB2	2.98E-06
RHOC	2.92E-06
RICTOR	3.74E-06
RIN1	2.50E-06
RIPK1	3.33E-06
RIPK2	3.73E-06
RIPK3	2.96E-06
RIPK4	2.45E-06
ROBO1	3.76E-06
ROBO2	3.73E-06
ROCK1	3.81E-06
ROCK2	3.67E-06
ROR1	3.21E-06
ROR2	2.59E-06
RPS6KA1	2.78E-06
RPS6KA2	2.67E-06
RPS6KA3	3.93E-06
RPS6KA4	2.45E-06
RPS6KA5	3.84E-06
RPS6KA6	3.98E-06
RPS6KB1	3.79E-06

RPS6KB2	2.81E-06
RPSA	3.78E-06
RRAS	2.51E-06
RRAS2	3.15E-06
RUNX1	2.46E-06
RUNX2	2.84E-06
RUNX3	2.55E-06
RYK	3.15E-06
SCARF2	2.34E-06
SEMA3B	3.00E-06
SEMA3F	2.82E-06
SERPIND1	3.03E-06
SGK	3.04E-06
SGK2	2.90E-06
SGKL	3.92E-06
SHC1	3.16E-06
SHC2	2.32E-06
SHC3	2.61E-06
SHH	2.28E-06
SIAH1	3.47E-06
SKI	2.91E-06
SLA	2.86E-06
SLC38A3	2.69E-06
SLC7A4	2.61E-06
SLK	3.98E-06
SMAD1	3.51E-06
SMAD2	3.79E-06
SMAD3	2.74E-06
SMAD4	3.87E-06
SMAD5	3.85E-06
SMAD6	2.36E-06
SMAD7	2.41E-06
SMAD9	2.92E-06

SMG1	3.60E-06
SMO	2.69E-06
SNAP29	2.96E-06
SOS1	3.89E-06
SOS2	3.84E-06
SRC	2.52E-06
SRMS	2.45E-06
SRPK3	2.59E-06
STAT1	3.24E-06
STAT2	2.93E-06
STAT3	2.99E-06
STAT4	3.50E-06
STAT5A	2.68E-06
STAT5B	2.83E-06
STAT6	2.87E-06
STK10	2.78E-06
STK11	2.52E-06
STK24	2.74E-06
STK25	2.58E-06
STK3	3.78E-06
STK32A	3.36E-06
STK32B	2.81E-06
STK32C	2.41E-06
STK36	3.00E-06
STK38	3.85E-06
STK38L	3.93E-06
STK39	3.17E-06
STYK1	3.45E-06
SUCNR1	3.38E-06
SUFU	2.86E-06
SYK	2.98E-06
TAOK1	3.86E-06
TAOK2	2.69E-06

TAOK3	3.29E-06
TBK1	4.10E-06
TCF1	2.59E-06
TEC	3.88E-06
TEK	3.41E-06
TERT	2.58E-06
TESK1	2.73E-06
TESK2	3.35E-06
TFDP1	2.80E-06
TGFB2	3.16E-06
TGFBR1	3.44E-06
TGFBR2	2.99E-06
THAP7	2.72E-06
TIE	2.81E-06
TNIK	3.36E-06
TNK1	2.74E-06
TNK2	2.59E-06
TNNI3K	3.67E-06
TP53	3.15E-06
TP73	2.75E-06
TP73L	3.14E-06
TPTE	3.55E-06
TRAIP	2.96E-06
TSC1	3.22E-06
TSC2	2.50E-06
TSHR	2.97E-06
TU3A	2.71E-06
TUSC2	2.77E-06
TUSC4	2.80E-06
TXK	3.67E-06
TYK2	2.72E-06
TYRO3	2.91E-06
UBE2L3	3.31E-06

VAV1	2.83E-06
VAV2	2.67E-06
VAV3	3.57E-06
VEGFC	3.05E-06
VHL	2.60E-06
WEE1	2.90E-06
WT1	2.47E-06
XRCC1	3.08E-06
YES1	4.11E-06
ZAK	3.58E-06
ZAP70	2.61E-06
ZMYND10	2.80E-06

Supplementary Table 5c. Collapsed categories shown in Supplementary Table 4b based on CpG context. Relative mutation rates measured for each category are shown as well.

Mutations occurring at	Mutation rate relative to total
C or G in CpG	3.80
C or G	1.38
A or T	0.31

Supplementary Table 6a. Significantly mutated genes (FDR < 0.1) identified using three different approaches. Significantly mutated genes in lung adenocarcinomas ranked by P_value. Fixed background mutation rate (BMR) of 2.98E-06, gene-based BMRs, and category-based BMR of 3.30E-06, were used for the calculation.

			Fixed rate		Gene-based rate		Category-based rate	
Gene	Number of mutations	Number of base pairs covered	P_value	FDR	P_value	FDR	P_value	FDR
TP53	66	231907	2.56E-104	7.96E-102	9.62E-103	5.99E-100	<1E-11	<1E-8
KRAS	61	132847	4.00E-109	2.49E-106	3.29E-102	1.03E-99	<1E-11	<1E-8
STK11	34	240919	2.24E-44	4.65E-42	7.52E-47	1.56E-44	<1E-11	<1E-8
EGFR	34	749156	2.91E-28	4.53E-26	9.88E-29	1.54E-26	<1E-11	<1E-8
LRP1B	17	2635000	0.003135	0.092203	0.028039	0.415906	0.00856	0.197
NF1	16	1533738	2.47E-05	0.0022	0.000209	0.018594	0.000114	0.00887
ATM	14	1758433	0.001095	0.052459	0.011731	0.259916	0.00015	0.00921
APC	13	1613273	0.001456	0.060476	0.01857	0.332869	8.64E-05	0.00769
EPHA3	11	570668	1.85E-06	0.000192	8.89E-06	0.000923	1.01E-05	0.00105
PTPRD	10	1054147	0.001555	0.060565	0.006227	0.175736	0.000163	0.00921
CDKN2A	9	148848	1.24E-09	1.55E-07	9.14E-10	1.14E-07	1.40E-08	1.74E-06
ERBB4	9	754156	0.00055	0.042844	0.001866	0.089403	0.000138	0.00921
KDR	8	785965	0.002908	0.090575	0.006488	0.175736	0.00663	0.172
FGFR4	7	472000	0.000644	0.044568	0.000261	0.020329	0.000784	0.0349
NTRK1	7	484439	0.000748	0.046618	0.000379	0.02624	0.0101	0.218
RB1	7	514299	0.001054	0.052459	0.003408	0.151634	0.00146	0.057
NTRK3	7	528299	0.001227	0.05462	0.001543	0.080118	0.00319	0.105
EPHA5	7	581785	0.002104	0.077118	0.005304	0.159716	0.00593	0.161
PDGFRA	7	630666	0.00327	0.092203	0.00528	0.159716	0.0143	0.259
GNAS	7	635372	0.003404	0.092203	0.004496	0.159716	0.0192	0.299
LTK	6	446513	0.002515	0.087038	0.001313	0.074357	0.0312	0.442
INHBA	5	242179	0.000904	0.05118	0.000635	0.039581	0.00147	0.0574
PAK3	5	317212	0.002903	0.090575	0.008551	0.213096	0.00723	0.17
ZMYND10	4	240810	0.006294	0.130701	0.005058	0.159716	0.000435	0.0208
NRAS	3	109611	0.004571	0.109518	0.007599	0.197264	0.000216	0.0112
SLC38A3	3	173817	0.015837	0.281891	0.012038	0.259916	0.00271	0.0992

Supplementary Table 6b. Pfam domains have significantly higher mutation rates above background (2.98e-06)

Pfam domain	# of mutations	Covered length	P_value	FDR	Bonferroni
P53	54	332196	9.46E-73	2.21E-70	2.21E-70
Ras	71	1107696	3.02E-67	3.53E-65	7.06E-65
Pkinase_Tyr	104	12298020	8.27E-20	6.45E-18	1.94E-17
Pkinase	103	18050256	1.65E-09	9.65E-08	3.86E-07
Recep_L_domain	16	1668876	6.49E-05	0.00276	0.015189
P53_tetramer	4	71064	7.08E-05	0.00276	0.016562
fn3	25	3825048	0.000333	0.01113	0.077909
Furin-like	8	617016	0.000644	0.018828	0.150626
NESP55	4	138180	0.000864	0.022456	0.202104
Ldl_recept_b	16	2162376	0.001063	0.024866	0.248662
MAM	3	88548	0.002516	0.053514	0.588659
ig	15	2224980	0.003539	0.069002	0.828019
Ldl_recept_a	14	2125716	0.005741	0.103346	1
Y_phosphatase	7	737148	0.007404	0.116746	1
TGFb_propeptide	3	131412	0.007484	0.116746	1
Ephrin_lbd	8	1005048	0.011798	0.17254	1
CXCXC	2	60348	0.014356	0.197613	1

Supplementary Table 7a. Permutation analysis of concurrence among mutations in lung adenocarcinoma (P < 0.1) by keeping the number of mutated samples in a given gene and the number of genes mutated in a given samples as constant.

Gene 1	Gene 2	Number of samples with mutation in Gene 1	Number of samples with mutation in gene 2	Number of samples with mutation in both genes	Number of samples with mutation in only one gene	Expected number of samples with mutation in both genes by permutation	Expected number of samples with mutation in only one gene by permutation	P_value	FDR
PIK3C3	PTPRD	5	10	4	7	0.708041	13.58392	0.000967	0.196301
NTRK2	PDGFRA	6	6	3	6	0.523618	10.95276	0.006901	0.84158
FGFR4	NTRK2	7	6	3	7	0.605003	11.78999	0.011104	0.880429
FGFR4	PDGFRA	7	6	3	7	0.60568	11.78864	0.01137	0.880429
PDGFRA	PTPRD	6	10	3	10	0.842755	14.31449	0.032151	1
EPHA3	LTK	10	6	3	10	0.84256	14.31488	0.032326	1
EPHA3	TP53	10	64	9	56	5.963705	62.07259	0.037207	1
LRP1B	TP53	16	64	13	54	9.327048	61.3459	0.037567	1
APC	LRP1B	11	16	5	17	2.265757	22.46849	0.042354	1
DOCK3	LTK	5	6	2	7	0.440355	10.11929	0.057283	1
NF1	PTPRD	13	10	4	15	1.727687	19.54463	0.063166	1
EPHA7	NF1	6	13	3	13	1.068754	16.86249	0.06482	1
NF1	PDGFRA	13	6	3	13	1.069703	16.86059	0.065295	1
GNAS	PIK3C3	7	5	2	8	0.509754	10.98049	0.077146	1
NTRK3	STK11	7	34	5	31	2.748182	35.50364	0.079204	1
INSRR	PDGFRA	6	6	2	8	0.523021	10.95396	0.08151	1
DOCK3	KDR	5	8	2	9	0.576516	11.84697	0.097414	1
ATM	KRAS	13	60	10	53	7.362754	58.27449	0.099267	1

Supplementary Table 7b. Permutation analysis of mutual exclusion among mutations in lung adenocarcinoma ($P < 0.1$) by keeping the number of mutated samples in a given gene and the number of genes mutated in a given samples as constant.

Gene 1	Gene 2	Number of samples with mutation in Gene 1	Number of samples with mutation in gene 2	Number of samples with mutation in both genes	Number of samples with mutation in only one gene	Expected number of samples with mutation in both genes by permutation	Expected number of samples with mutation in only one gene by permutation	P_value	FDR
EGFR	KRAS	30	60	0	90	15.88841	58.22318	0	0
EGFR	STK11	30	34	1	62	10.33993	43.32014	7.00E-06	0.002842
ATM	TP53	13	64	1	75	7.669006	61.66199	9.50E-05	0.025713
EGFR	LRP1B	30	16	1	44	5.403172	35.19366	0.007052	0.84158
STK11	TP53	34	64	12	74	18.55174	60.89652	0.007255	0.84158
EGFR	EPHA3	30	10	0	40	3.502502	32.995	0.009515	0.880429
NF1	STK11	13	34	1	45	4.926834	37.14633	0.011927	0.880429
EGFR	TP53	30	64	11	72	16.61589	60.76822	0.015579	0.990524
KRAS	MYO3B	60	7	1	65	4.061141	58.87772	0.020671	0.990524
GNAS	KRAS	7	60	1	65	4.057897	58.88421	0.020786	0.990524
EGFR	NF1	30	13	1	41	4.469908	34.06018	0.022752	0.990524
EGFR	KDR	30	8	0	38	2.834468	32.33106	0.024322	0.990524
EGFR	ERBB4	30	8	0	38	2.834851	32.3303	0.024455	0.990524
MYO3B	STK11	7	34	0	41	2.747733	35.50453	0.025325	0.990524
GNAS	STK11	7	34	0	41	2.74798	35.50404	0.025368	0.990524
FGFR4	STK11	7	34	0	41	2.746662	35.50668	0.025542	0.990524
NTRK1	STK11	7	34	0	41	2.746753	35.50649	0.025617	0.990524
KRAS	TP53	60	64	24	76	29.86544	64.26912	0.032204	1
EGFR	EPHA5	30	7	0	37	2.498488	32.00302	0.03856	1
EGFR	NTRK3	30	7	0	37	2.497404	32.00519	0.038661	1
INSRR	STK11	6	34	0	40	2.368251	35.2635	0.042782	1
KRAS	STK11	60	34	13	68	17.72879	58.54242	0.042897	1

PDGFRA	STK11	6	34	0	40	2.367546	35.26491	0.043273	1
KRAS	NF1	60	13	4	65	7.359486	58.28103	0.044014	1
EGFR	LTK	30	6	0	36	2.154057	31.69189	0.061911	1
CDKN2A	STK11	9	34	1	41	3.49261	36.01478	0.069095	1
EGFR	PTPRD	30	10	1	38	3.499732	33.00054	0.072	1
LRP1B	STK11	16	34	3	44	5.958087	38.08383	0.076116	1
APC	TP53	11	64	4	67	6.53659	61.92682	0.094491	1
EGFR	PIK3C3	30	5	0	35	1.80596	31.38808	0.098305	1

Supplementary Table 7c. Permutation analysis of concurrence and mutual exclusion among mutations in lung adenocarcinoma (P < 0.1) by keeping the number of mutations in a given gene and sample as constant.

Gene 1	Gene 2	P_value	Correlation[positive(1) and negative(-1)]	FDR
EGFR	KRAS	0	-1	0
STK11	EGFR	0.000056	-1	0.022736
PIK3C3	PTPRD	0.000575	1	0.136213
ATM	TP53	0.000671	-1	0.136213
NTRK2	PDGFRA	0.001877	1	0.304825
PDGFRA	FGFR4	0.003435	1	0.46487
NTRK2	FGFR4	0.008477	1	0.687477
LTK	EPHA3	0.010365	1	0.687477
APC	LRP1B	0.011811	1	0.687477
EGFR	TP53	0.011997	-1	0.687477
NF1	STK11	0.01236	-1	0.687477
NF1	EGFR	0.012527	-1	0.687477
PDGFRA	PTPRD	0.013042	1	0.687477
LRP1B	EGFR	0.013149	-1	0.687477
EPHA3	EGFR	0.013699	-1	0.687477
DOCK3	CDH11	0.014314	1	0.687477
LTK	DOCK3	0.014393	1	0.687477
STK11	TP53	0.017367	-1	0.783445
NF1	KRAS	0.020248	-1	0.865336

INSRR	PDGFRA	0.021521	1	0.873753
PDGFRA	NF1	0.028267	1	1
ERBB4	EGFR	0.029278	-1	1
PRKDC	PDGFRA	0.032299	1	1
EPHA7	NF1	0.036185	1	1
KDR	EGFR	0.043148	-1	1
EPHA7	LRP1B	0.04346	1	1
PTPRD	NF1	0.047795	1	1
GNAS	KRAS	0.049786	-1	1
MYO3B	KRAS	0.049879	-1	1
DOCK3	PTPRD	0.052729	1	1
KRAS	TP53	0.057688	-1	1
EPHA5	EGFR	0.062234	-1	1
NTRK3	EGFR	0.062797	-1	1
LTK	CDH11	0.062879	1	1
NTRK1	STK11	0.062956	-1	1
FGFR4	STK11	0.063005	-1	1
GNAS	STK11	0.063109	-1	1
PDGFRA	STK11	0.063112	-1	1
MYO3B	STK11	0.063657	-1	1
CDH11	EPHA3	0.065685	1	1
STK11	KRAS	0.066042	-1	1
APC	TP53	0.072656	-1	1
PTPRD	EGFR	0.082674	-1	1
CDH11	PRKDC	0.082786	1	1
INSRR	NTRK1	0.082881	1	1
INSRR	GNAS	0.082897	1	1
PIK3C3	GNAS	0.082947	1	1
NTRK2	PRKDC	0.083003	1	1
PIK3C3	PDGFRA	0.083038	1	1
LRP1B	STK11	0.088567	-1	1
PDGFRA	EPHA3	0.092409	1	1
INSRR	STK11	0.09263	-1	1

LTK	EGFR	0.092744	-1	1
PIK3C3	EGFR	0.092902	-1	1
CDH11	EGFR	0.092987	-1	1
PRKDC	EPHA3	0.093062	1	1
EPHA7	APC	0.094953	1	1
NTRK3	STK11	0.095363	1	1
APC	EGFR	0.095808	-1	1
INSRR	LRP1B	0.096179	1	1
PIK3C3	NTRK2	0.097889	1	1
PIK3C3	INSRR	0.098443	1	1

Supplementary Table 8. Mutations in a set of genes show positive correlation with the overall mutation rate.

Gene	Num of mutations	Num. of mutations in samples with mutation	Covered base (bp) for samples with mutation	Mutation rate in samples wih mutation	Num. of mutations in samples without mutation	Covered base (bp) for samples without mutation	Mutation rate in samples wihout mutation	P_value	FDR	Bonferroni
TP53	66	584	58444432	9.99E-06	363	1.14E+08	3.18E-06	4.68E-68	1.67E-65	1.67E-65
PRKDC	7	163	5935361	2.75E-05	843	1.64E+08	5.13E-06	3.52E-59	6.27E-57	1.25E-56
MYB	2	85	1638543	5.19E-05	926	1.71E+08	5.43E-06	5.56E-51	6.60E-49	1.98E-48
PDGFRA	7	140	5274829	2.65E-05	866	1.67E+08	5.19E-06	2.09E-49	1.86E-47	7.45E-47
DDR1	2	75	1609421	4.66E-05	936	1.71E+08	5.49E-06	6.01E-42	4.28E-40	2.14E-39
FGFR4	7	140	6183893	2.26E-05	866	1.66E+08	5.22E-06	8.10E-42	4.81E-40	2.88E-39
NTRK2	6	128	5290566	2.42E-05	879	1.67E+08	5.27E-06	5.35E-41	2.72E-39	1.90E-38
IRS1	3	90	2748187	3.27E-05	920	1.69E+08	5.44E-06	2.50E-38	1.11E-36	8.90E-36
INSRR	6	125	5420552	2.31E-05	882	1.67E+08	5.30E-06	6.11E-38	2.41E-36	2.17E-35
LRP1B	17	216	14443031	1.50E-05	780	1.56E+08	5.01E-06	2.90E-37	1.03E-35	1.03E-34
ROBO2	5	113	4576600	2.47E-05	895	1.67E+08	5.35E-06	7.58E-37	2.40E-35	2.70E-34
PIK3C2G	4	98	3439595	2.85E-05	911	1.68E+08	5.41E-06	8.10E-37	2.40E-35	2.88E-34
PTPRD	10	162	9130117	1.77E-05	841	1.63E+08	5.17E-06	8.25E-36	2.26E-34	2.94E-33
SMG1	3	83	2516901	3.30E-05	927	1.68E+08	5.51E-06	2.09E-35	5.31E-34	7.43E-33
MAST4	2	71	1837120	3.86E-05	940	1.7E+08	5.51E-06	9.59E-35	2.15E-33	3.41E-32

TYK2	2	71	1834392	3.87E-05	940	1.7E+08	5.52E-06	9.66E-35	2.15E-33	3.44E-32
VAV1	2	66	1627946	4.05E-05	945	1.71E+08	5.54E-06	1.55E-33	3.24E-32	5.51E-31
ACVRL1	2	69	1833040	3.76E-05	942	1.71E+08	5.52E-06	3.91E-33	7.66E-32	1.39E-30
LMTK3	4	69	1834426	3.76E-05	940	1.7E+08	5.52E-06	4.09E-33	7.66E-32	1.45E-30
PTPRG	2	65	1636558	3.97E-05	946	1.7E+08	5.56E-06	1.69E-32	2.87E-31	6.00E-30

Supplementary Table 9a. Top 100 genes showing strong correlation between DNA copy number and gene expression. 75 samples have been used for this analysis.

Gene	Correlation Coefficient	P-value	FDR
ING1	0.787589	0	0
ATP11B	0.790337	0	0
ARMC1	0.797561	0	0
MYNN	0.79396	0	0
C19orf12	0.802711	0	0
IPO8	0.778825	2.22E-16	2.98E-13
CDK8	0.773543	4.44E-16	3.97E-13
TAF2	0.773286	4.44E-16	3.97E-13
BRMS1L	0.772436	4.44E-16	3.97E-13
PIK3CA	0.768906	8.88E-16	7.15E-13
LSM1	0.751424	7.99E-15	5.85E-12
MFN1	0.746529	1.49E-14	9.19E-12
TOMM70A	0.7463	1.53E-14	9.19E-12
RFC4	0.745962	1.60E-14	9.19E-12
PAN3	0.745238	1.75E-14	9.41E-12
C19orf2	0.741453	2.78E-14	1.40E-11
TXNL1	0.740337	3.20E-14	1.51E-11
WHSC1L1	0.730942	9.70E-14	4.34E-11
ANKRD10	0.725764	1.75E-13	7.43E-11
GOLGA7	0.725289	1.85E-13	7.45E-11
ARMC8	0.718214	4.07E-13	1.46E-10
MIPEP	0.718096	4.12E-13	1.46E-10
NAPG	0.717964	4.18E-13	1.46E-10
SIAH2	0.705852	1.52E-12	5.11E-10
GATAD1	0.70454	1.75E-12	5.62E-10
CDC42SE2	0.703671	1.91E-12	5.91E-10
METTL4	0.702787	2.09E-12	6.23E-10
RAE1	0.699145	3.03E-12	8.71E-10
MRPS28	0.698281	3.31E-12	9.18E-10

FXR1	0.693729	5.21E-12	1.40E-09
SENP2	0.692081	6.14E-12	1.59E-09
KPNA4	0.690294	7.31E-12	1.84E-09
PNPT1	0.689265	8.08E-12	1.91E-09
H2AFY	0.689256	8.09E-12	1.91E-09
SEC23IP	0.687162	9.90E-12	2.28E-09
RB1CC1	0.686496	1.06E-11	2.36E-09
NEK3	0.684686	1.26E-11	2.73E-09
PAK2	0.68375	1.37E-11	2.91E-09
THUMPD1	0.682872	1.49E-11	3.08E-09
PDCD6	0.680715	1.83E-11	3.68E-09
MRPL47	0.679662	2.02E-11	3.96E-09
ASH2L	0.67872	2.21E-11	4.23E-09
KIAA0196	0.678123	2.33E-11	4.36E-09
MAPK8	0.675327	3.02E-11	5.52E-09
ZNF302	0.67314	3.69E-11	6.60E-09
AKAP9	0.671911	4.13E-11	7.08E-09
CEBPG	0.671885	4.14E-11	7.08E-09
SNAP29	0.669106	5.32E-11	8.91E-09
KRIT1	0.668002	5.87E-11	9.64E-09
THOC1	0.662984	9.16E-11	1.47E-08
NDUFB5	0.657272	1.50E-10	2.34E-08
MTMR6	0.657214	1.51E-10	2.34E-08
COG3	0.654448	1.91E-10	2.91E-08
ZNF336	0.652457	2.27E-10	3.38E-08
RASA1	0.649977	2.79E-10	4.01E-08
FBXW11	0.649966	2.79E-10	4.01E-08
RYK	0.649618	2.88E-10	4.06E-08
SIP1	0.649243	2.97E-10	4.12E-08
DBR1	0.648354	3.20E-10	4.36E-08
PHC3	0.647082	3.55E-10	4.76E-08
CTTN	0.646429	3.75E-10	4.94E-08
PPP1R12A	0.643372	4.82E-10	6.25E-08

TBCA	0.642533	5.16E-10	6.59E-08
PPFIA1	0.641975	5.40E-10	6.78E-08
FAM91A1	0.640789	5.94E-10	7.35E-08
CMAS	0.640512	6.07E-10	7.41E-08
EIF3S3	0.639411	6.64E-10	7.97E-08
ANKRD46	0.637646	7.65E-10	9.05E-08
GTF3A	0.635617	8.99E-10	1.05E-07
BAG4	0.635452	9.11E-10	1.05E-07
BRWD1	0.634212	1.00E-09	1.14E-07
DCTD	0.631675	1.23E-09	1.36E-07
CHD7	0.631594	1.23E-09	1.36E-07
CBLL1	0.630978	1.29E-09	1.40E-07
WDSOF1	0.630884	1.30E-09	1.40E-07
RAB20	0.629063	1.50E-09	1.59E-07
E2F3	0.628737	1.54E-09	1.60E-07
TPD52	0.628628	1.55E-09	1.60E-07
NUP153	0.628439	1.58E-09	1.61E-07
PEPD	0.627701	1.67E-09	1.67E-07
PPM1A	0.627613	1.68E-09	1.67E-07
TRIM35	0.625971	1.91E-09	1.86E-07
FBXO25	0.625912	1.92E-09	1.86E-07
KBTBD11	0.624305	2.17E-09	2.07E-07
PRKCI	0.623971	2.22E-09	2.10E-07
KIAA0804	0.621695	2.64E-09	2.44E-07
ZRF1	0.62167	2.64E-09	2.44E-07
TERF1	0.621577	2.66E-09	2.44E-07
FBXO3	0.62047	2.89E-09	2.62E-07
MAP3K1	0.619208	3.18E-09	2.84E-07
COG5	0.618188	3.43E-09	3.04E-07
CRI2	0.617574	3.59E-09	3.13E-07
FBXO33	0.617479	3.62E-09	3.13E-07
GABPA	0.616849	3.79E-09	3.25E-07
VPS41	0.616634	3.85E-09	3.26E-07

KIAA0274	0.616026	4.03E-09	3.38E-07
NIPBL	0.614466	4.52E-09	3.75E-07
PPP4R2	0.613194	4.97E-09	4.08E-07
MRPS11	0.612392	5.27E-09	4.26E-07
SDHC	0.612334	5.29E-09	4.26E-07

Supplementary Table 9b. Highly mutated genes showing strong correlation between DNA copy number and gene expression. 75 samples have been used for this analysis.

Gene	Correlation Coefficient	P-value	FDR
EGFR	0.549427	3.31E-07	1.24E-05
KRAS	0.523062	1.48E-06	4.08E-05
NF1	0.5175	1.99E-06	5.22E-05
APC	0.50529	3.77E-06	8.72E-05
RB1	0.46645	2.47E-05	0.000374
PRKDC	0.36225	0.001405	0.008871
EGFR	0.549427	3.31E-07	1.24E-05
KRAS	0.523062	1.48E-06	4.08E-05
NF1	0.5175	1.99E-06	5.22E-05
APC	0.50529	3.77E-06	8.72E-05
RB1	0.46645	2.47E-05	0.000374
PRKDC	0.36225	0.001405	0.008871
EGFR	0.549427	3.31E-07	1.24E-05

Supplementary Table 10a. Correlation between mutation and copy number change in 188 lung adenocarcinomas.

Gene	Correlation Coefficient	P_value	FDR
PTEN	-0.3789	8.24E-08	4.37E-06
EGFR	0.331369	3.40E-06	9.01E-05
HCK	0.262651	0.000271	0.00479
APC	-0.24861	0.000581	0.007702
KRAS	0.224593	0.001944	0.017873
EPHB1	0.223757	0.002023	0.017873
TP53	-0.17321	0.017451	0.132128
DOCK3	0.158571	0.029744	0.197051
NTRK1	0.152884	0.036208	0.213224
MERTK	0.14518	0.046825	0.235745
CDC42BPA	0.141571	0.052629	0.235745
ITK	-0.14113	0.053376	0.235745

*Genes with ≥4 mutations used for the analysis

* *STK11, LTK, PTCH1, ZMYND10, PRKCG, LMTK3,* and *ERBB2* were not included in this analysis because there are no SNPs in those genes on affy 250k sty array.

Supplementary Table 10b. Correlation between mutation and gene expression in 41 lung adenocarcinomas.

Gene	Correlation Coefficient	P_value	FDR
KRAS	0.628025	1.10E-05	0.000364
EGFR	0.589485	5.01E-05	0.000826
STK11	-0.36258	0.019819	0.218013
EPHB1	0.314724	0.045055	0.32049
CDC42BPA	0.309995	0.048559	0.32049
RB1	0.290554	0.065344	0.35939
PIK3C2G	-0.26772	0.090589	0.427062
PRKDC	-0.24786	0.118166	0.487435

DOCK3	0.231416	0.145447	0.506924
ATM	-0.22129	0.1644	0.506924
ROBO1	-0.21156	0.184241	0.506924
APC	-0.20903	0.189652	0.506924
PDGFRA	-0.20448	0.199697	0.506924

*Genes with ≥4 mutations used for the analysis.

⋆*PTCH1* and *KIAA1804* were not included in the analysis because there are no expression probes for those genes.

Supplementary Table 11. Copy number variations and mutations found in genes from MAPK signaling, Wnt signaling, P53 signaling, cell cycle, and mammalian target of rapamycin (mTOR) pathways. Excel spreadsheet can be assessed at http://genome.wustl.edu/supplemental/tsp_nature_2008.

Supplementary Table 12. Pathways significantly mutated in lung adenocarcinoma (mTOR pathway has been modified based on the literature and cancer type-specific pathways are not included). Fisher's exact test and binomial test were used. Background mutation rate of 2.98E-06 was used. See supplementary for details. Some cancer type-specific pathways are not shown.

KEGG level	Pathway	# of genes in pathway	# of genes outside of pathway	# of mutations in pathway	Covered bases in pathway	# of mutation outside of pathway	Covered bases outside of pathway	P_value (Fisher's exact test)	FDR (Fisher's exact test)	P_value (binomial test)	FDR (binomial test)
level_3	04540 Gap junction [PATH:ko04540]	34	355	137	10482548	614	147881919	3.62E-37	2.53E-36	4.13E-44	2.89E-43
level_3	04320 Dorso-ventral axis formation [PATH:ko04320]	18	371	106	8118994	645	150245473	1.48E-28	8.41E-28	1.65E-34	8.84E-34
level_3	04730 Long-term depression [PATH:ko04730]	22	367	92	6880300	659	151484167	5.77E-26	3.09E-25	6.25E-31	2.59E-30
level_3	04115 p53 signaling pathway [PATH:ko04115]	21	368	103	8337017	648	150027450	7.96E-25	4.03E-24	1.10E-31	4.76E-31
level_3	04530 Tight junction [PATH:ko04530]	29	360	97	7804914	654	150559553	1.15E-23	5.49E-23	3.90E-30	1.54E-29
level_3	04720 Long-term potentiation [PATH:ko04720]	28	361	90	6991890	661	151372577	1.40E-23	6.38E-23	3.84E-29	1.46E-28
level_3	04912 GnRH signaling pathway [PATH:ko04912]	37	352	126	11592020	625	146772447	5.04E-23	2.18E-22	3.94E-33	1.99E-32
level_3	04010 MAPK signaling pathway [PATH:ko04010]	109	280	289	37995369	462	120369098	2.13E-20	8.79E-20	2.06E-43	1.34E-42
level_3	04310 Wnt signaling pathway [PATH:ko04310]	28	361	104	9273619	647	149090848	2.23E-20	8.83E-20	1.07E-28	3.90E-28
level_3	04150 mTOR signaling pathway [PATH:ko04150]	36	353	142	14499992	609	143864475	3.47E-20	1.32E-19	1.63E-32	7.80E-32
level_3	04210 Apoptosis [PATH:ko04210]	32	357	115	11649739	636	146714728	1.21E-16	4.39E-16	5.91E-27	2.07E-26
level_3	04662 B cell receptor signaling pathway [PATH:ko04662]	25	364	92	8742434	659	149622033	1.29E-15	4.51E-15	9.79E-24	2.97E-23
level_3	04810 Regulation of actin cytoskeleton [PATH:ko04810]	60	329	188	23311155	563	135053312	2.27E-15	7.65E-15	7.13E-32	3.24E-31
level_3	04360 Axon guidance [PATH:ko04360]	44	345	149	18678583	602	139685884	1.21E-11	3.79E-11	3.42E-25	1.15E-24
level_3	04370 VEGF signaling pathway [PATH:ko04370]	33	356	101	11321588	650	147042879	3.36E-11	1.02E-10	7.71E-21	2.26E-20
level_3	04020 Calcium signaling pathway [PATH:ko04020]	20	369	84	9165249	667	149199218	3.93E-10	1.16E-09	2.75E-18	7.36E-18
level_3	04012 ErbB signaling pathway [PATH:ko04012]	68	321	177	24438022	574	133926445	9.23E-10	2.62E-09	4.78E-25	1.55E-24
level_3	04110 Cell cycle [PATH:ko04110]	46	343	133	17688973	618	140675494	1.78E-08	4.92E-08	1.46E-20	4.16E-20

Supplementary Table 13a. Some highly mutated genes display copy number variations. CN > 2.5 is used for assigning amplifications and CN < 1.5 is used for assigning deletions.

Gene	# of samples with mutation	# of samples with amplification	# of samples with deletion	# of samples with both mutation and amplification	# of samples with both mutation and deletion
KRAS	60	11	4	7	1
TP53	64	13	9	2	5
EGFR	30	16	0	6	0
CDKN2A	9	1	21	0	1
EPHA3	10	7	3	1	0
NF1	13	7	1	0	0
ERBB4	8	8	1	0	0
FGFR4	7	24	12	2	0
NTRK1**	7	43	1	4	0
INHBA	5	18	0	0	0
ATM	13	2	2	0	0
RB1	7	2	1	0	0
NTRK3	7	0	0	0	0
APC	11	4	4	1	1
PTPRD	10	0	8	0	1
EPHA5	7	6	4	0	0
LRP1B	16	5	2	0	0
KDR	8	3	0	1	0
PAK3	5	13	5	0	0
PDGFRA	6	1	0	0	0
GNAS	7	16	1	1	0
CDH11	4	1	1	1	0
NTRK2	6	0	0	0	0
NRAS	3	5	3	1	0

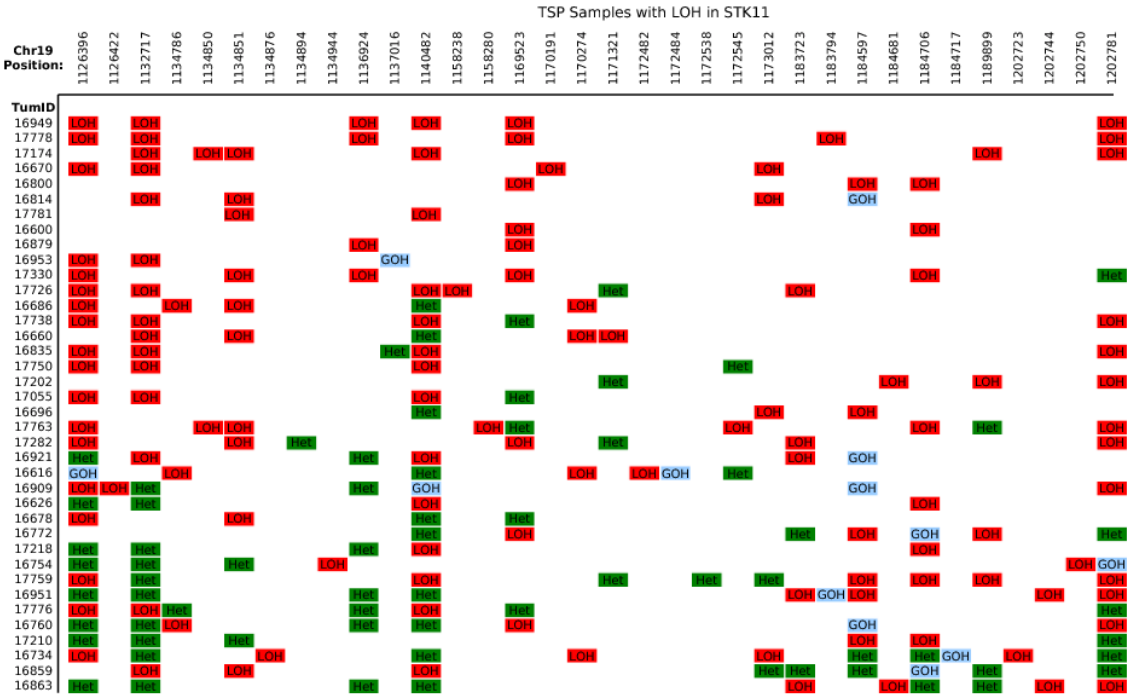
** NTRK1 is not from focal amplification region

Supplementary Table 13b. Some highly mutated genes display copy number variations. CN>3 is used for assigning amplifications and CN < 1.5 is used for assigning deletions.

Gene	# of samples with mutation	# of samples with amplification	# of samples with deletion	# of samples with both mutation and amplification	# of samples with both mutation and deletion
KRAS	60	3	4	2	1
TP53	64	1	9	1	5
EGFR	30	5	0	2	0
CDKN2A	9	0	21	0	1
EPHA3	10	1	3	0	0
NF1	13	0	1	0	0
ERBB4	8	0	1	0	0
FGFR4	7	5	12	1	0
NTRK1**	7	11	1	1	0
INHBA	5	3	0	0	0
ATM	13	0	2	0	0
RB1	7	0	1	0	0
NTRK3	7	0	0	0	0
APC	11	0	4	0	1
PTPRD	10	0	8	0	1
EPHA5	7	0	4	0	0
LRP1B	16	0	2	0	0
KDR	8	1	0	0	0
PAK3	5	2	5	0	0
PDGFRA	6	1	0	0	0
GNAS	7	3	1	0	0
CDH11	4	0	1	0	0
NTRK2	6	0	0	0	0
NRAS	3	0	3	0	0

** NTRK1 is not from focal amplification region.

Supplementary Table 14. Samples with at least two consecutive LOH events in STK11.



Supplementary Table 15a. Clinical information for samples used for SNP array, gene expression array, and DNA re-sequencing experiments and analyses.

	SNP array	Gene expression array	Re-sequencing
Total	383	75	188
Smoker (*C, *F) Never-smoker	245 (82,163) 39	45 (19,26) 8	124 (39,85) 20
Male Female	165 215	37 38	72 115
Grade 1, 2,3	34, 138,173	1, 20,27	24, 85,71
T-Stage 1,2,3,4	154,164,26, 18	21,32,11, 6	75,91,8,10

Note: Some samples with incomplete clinical information. *C: current smoker; *F: former smoker.

Supplementary Table 15b. Clinical information for 188 lung adenocarcinomas sequenced (including major subtype). Missing missing values for major subtype are coded as 99. Excel spreadsheet can be assessed at http://genome.wustl.edu/supplemental/tsp_nature_2008.

Supplementary Table 16. Tumors with no somatic mutations but with copy number variations (Amp: CN>2.5; Del: CN < 1.5) in the focal regions reported in Weir et al.[11]

Sample	chr	start	End	mean_cn
16746	8	1.21E+08	1.24E+08	2.513812
16804	7	69693285	69813285	2.626914
	17	34800000	35180000	2.67426
	22	19065446	20135446	2.575507
	19	40270000	40430000	2.6881
	17	41560000	41710000	2.586792
16810	5	8880000	10510000	2.573645
	5	19720000	23090000	2.61864
16917	19	34790000	35420000	2.578989
	17	34800000	35180000	2.542202
	22	19065446	20135446	2.661321
	19	40270000	40430000	2.995505
16997	5	750000	1620000	2.501844
17054	9	9410000	10400000	1.393052
	5	58400000	59060000	1.425969
	17	34800000	35180000	2.653528
	11	68520000	69360000	2.713188
	5	750000	1620000	2.639555
	22	19065446	20135446	2.576534
	6	43760000	44120000	2.56516
	6	30240000	30530000	2.661515
	7	91573285	92883285	1.407908
	19	40270000	40430000	2.650569

17144	5	8880000	10510000	2.864439
	5	19720000	23090000	3.174002
17252	5	750000	1620000	2.761063
	5	19720000	23090000	2.630595
17737	18	21540000	21900000	2.865068

Supplementary Table 17. copy number data were visually inspected to categorize samples into three categories: pass, borderline, and fail.

Sample_id	SNP array category	#_of_mutations
16997	Borderline	0
17240	Borderline	0
16530	Borderline	0
16674	Borderline	0
16766	Borderline	0
17054	Borderline	0
17252	Fail	0
16614	Fail	0
16746	Fail	0
16804	Fail	0
17144	Fail	0
17248	Fail	0
17737	Fail	0
16612	Fail	0
16688	Fail	0
16720	Fail	0
16722	Fail	0
16758	Fail	0
16810	Fail	0
16889	Fail	0
16917	Fail	0
17208	Fail	0
17148	Pass	0
16768	Pass	0
16780	Pass	0
16630	Borderline	1
16696	Borderline	1
16770	Borderline	1

16841	Borderline	1
16963	Borderline	1
16594	Borderline	1
16712	Borderline	1
16742	Borderline	1
16754	Borderline	1
16616	Fail	1
17150	Fail	1
17154	Fail	1
16684	Fail	1
16825	Fail	1
17146	Fail	1
17206	Fail	1
17238	Fail	1
17292	Fail	1
17294	Fail	1
17745	Fail	1
16636	Fail	1
16730	Fail	1
16792	Fail	1
16845	Fail	1
17228	Fail	1
17264	Fail	1
16606	Pass	1
16774	Pass	1
16784	Pass	1
17769	Pass	1
16786	Pass	1
16979	Borderline	2
16925	Borderline	2
17272	Borderline	2
16740	Borderline	2
16750	Borderline	2

16760	Borderline	2
16772	Borderline	2
17194	Borderline	2
16839	Fail	2
16927	Fail	2
17158	Fail	2
16821	Fail	2
16919	Fail	2
16937	Fail	2
17234	Fail	2
17250	Fail	2
17258	Fail	2
17260	Fail	2
17328	Fail	2
16664	Fail	2
16907	Fail	2
16710	Pass	2
16993	Pass	2
17182	Pass	2
17326	Pass	2
17754	Pass	2
17778	Pass	2
16965	Pass	2
17306	Pass	2
16646	Pass	2
17727	Borderline	3
17747	Borderline	3
16975	Borderline	3
17184	Borderline	3
17741	Borderline	3
17781	Borderline	3
16947	Borderline	3
17176	Fail	3

17731	Fail	3
17216	Fail	3
17728	Fail	3
16638	Pass	3
17172	Pass	3
17320	Pass	3
16973	Pass	3
17302	Pass	3
16596	Pass	3
16698	Pass	3
16748	Pass	3
17222	Borderline	4
17304	Borderline	4
17316	Borderline	4
17324	Borderline	4
16883	Borderline	4
16800	Borderline	4
16823	Fail	4
16915	Fail	4
17170	Fail	4
17286	Fail	4
16901	Pass	4
17298	Pass	4
17738	Pass	4
17776	Pass	4
17152	Pass	4
17739	Borderline	5
17760	Borderline	5
17186	Fail	5
17280	Fail	5
17743	Fail	5
17784	Fail	5
16861	Pass	5

16875	Pass	5
17202	Pass	5
17777	Pass	5
17750	Pass	5
16626	Borderline	6
16863	Borderline	6
17288	Borderline	6
17766	Borderline	6
16913	Fail	6
17730	Fail	6
16670	Pass	6
17735	Pass	6
16857	Fail	7
16881	Fail	7
16905	Fail	7
17190	Fail	7
16640	Fail	7
16706	Fail	7
16827	Pass	7
17055	Pass	7
16929	Fail	8
17308	Fail	8
16921	Pass	8
17733	Pass	8
17226	Fail	9
17330	Pass	9
16648	Pass	9
16859	Pass	9
17042	Borderline	10
17242	Borderline	10
17156	Fail	10
16909	Pass	10
17282	Pass	10

17726	Pass	10
17763	Pass	10
16949	Pass	10
16600	Borderline	11
17060	Borderline	11
16724	Borderline	11
16879	Pass	11
17268	Borderline	12
16608	Borderline	13
16734	Borderline	13
16802	Pass	13
17174	Pass	13
16814	Pass	13
17262	Pass	13
16951	Pass	14
16953	Pass	16
17218	Pass	17
17746	Pass	18
16632	Borderline	20
16628	Borderline	22
17290	Borderline	23
17759	Pass	24
16678	Pass	28
16686	Fail	29
16835	Pass	30
16660	Fail	36
16668	Pass	38
17210	Pass	49

E. Supplementary References

1. Chen, K. et al. PolyScan: an automatic indel and SNP detection approach to the analysis of human resequencing data. *Genome Res* **17**, 659-66 (2007).
2. Nickerson, D.A., Tobe, V.O. & Taylor, S.L. PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res* **25**, 2745-51 (1997).
3. Zhang, J. et al. SNPdetector: a software tool for sensitive and accurate SNP detection. *PLoS Comput Biol* **1**, e53 (2005).
4. Ng, P.C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* **31**, 3812-4 (2003).
5. Sunyaev, S. et al. Prediction of deleterious human alleles. *Hum Mol Genet* **10**, 591-7 (2001).
6. Ars, E. et al. Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Hum Mol Genet* **9**, 237-47 (2000).
7. Greenman, C., Wooster, R., Futreal, P.A., Stratton, M.R. & Easton, D.F. Statistical analysis of pathogenicity of somatic mutations in cancer. *Genetics* **173**, 2187-98 (2006).
8. Yang, Z., Nielsen, R., Goldman, N. & Pedersen, A.M. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**, 431-49 (2000).
9. Pagani, F., Raponi, M. & Baralle, F.E. Synonymous mutations in CFTR exon 12 affect splicing and are not neutral in evolution. *Proc Natl Acad Sci U S A* **102**, 6368-72 (2005).
10. Getz, G. et al. Comment on "The consensus coding sequences of human breast and colorectal cancers". *Science* **317**, 1500 (2007).
11. Weir, B.A. et al. Characterizing the cancer genome in lung adenocarcinoma. *Nature* **450**, 893-8 (2007).