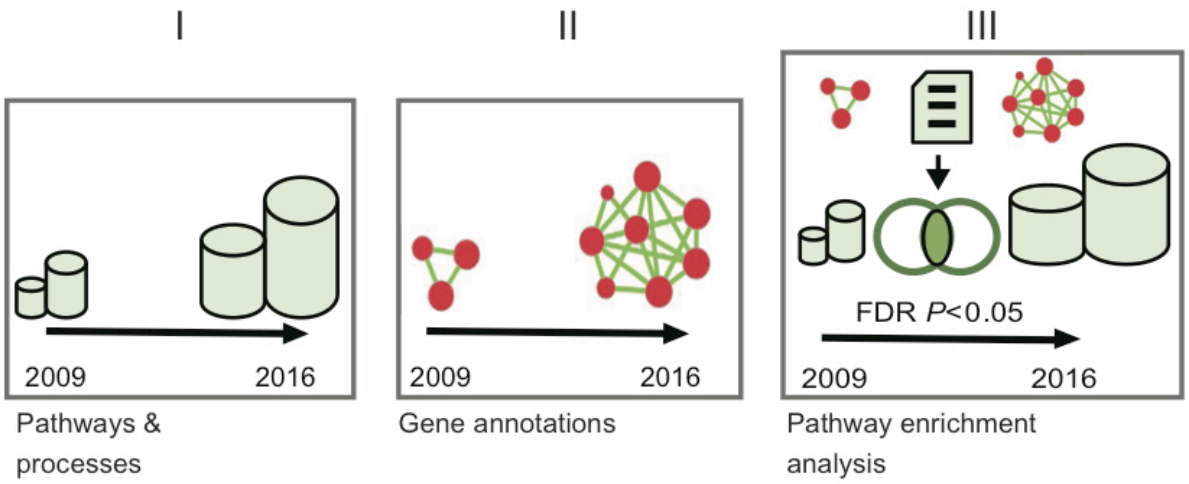
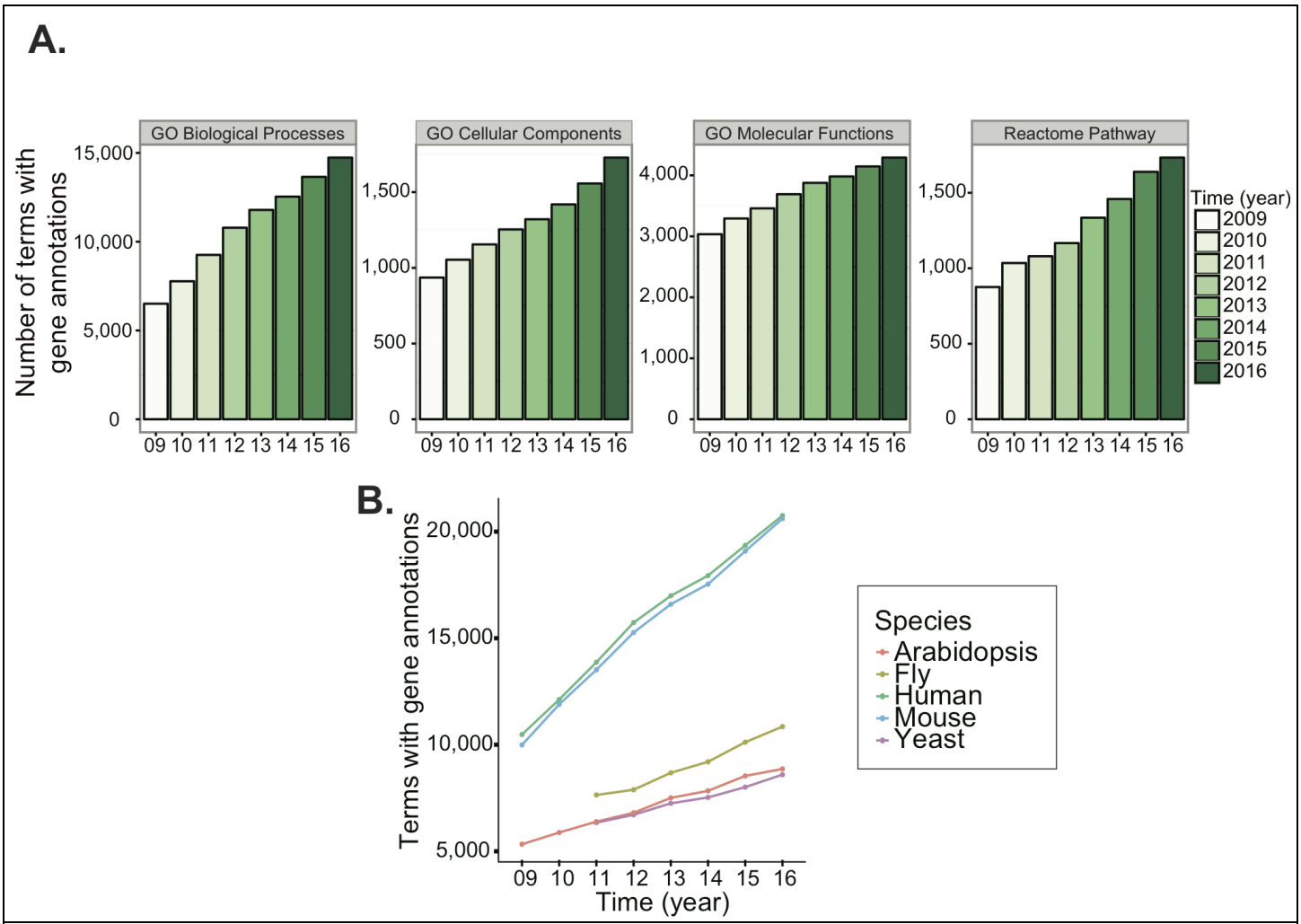


Supplementary Figure 1

Analysis outline.

We analyzed the accumulation of knowledge of gene function during the period 2009-2016 and its impact on practical analysis of gene lists. Our analysis involved three major steps (I-III). First, we studied the evolution of vocabulary of biological processes and pathways from Gene Ontology and the Reactome database (panel I). Second, we studied how gene annotations to these pathways and processes have changed over time (panel II). Third, we evaluated the practical impact of knowledge accumulation by performing pathway enrichment analysis using current and out-dated functional resources on gene lists derived from recent cancer genomics studies (panel III).

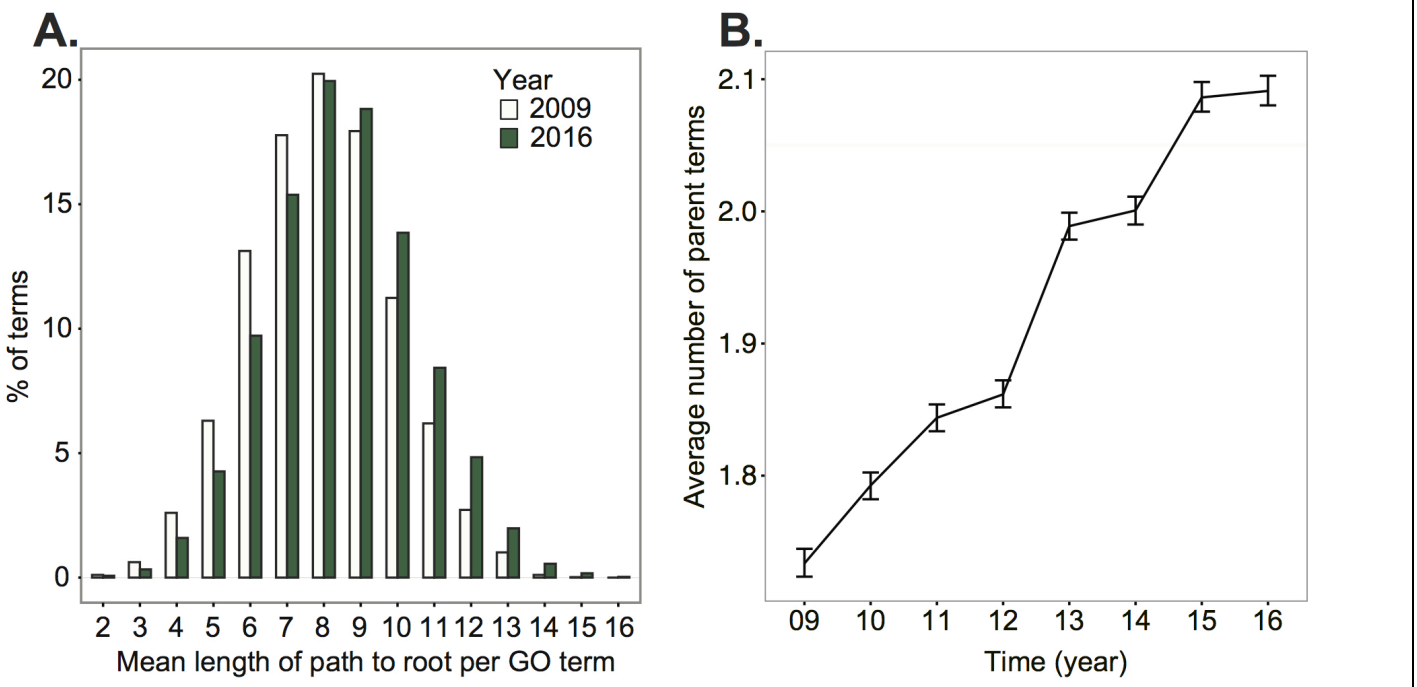
Nature Methods: doi:10.1038/nmeth.3963



Supplementary Figure 2

The vocabulary of biological pathways and processes is growing rapidly.

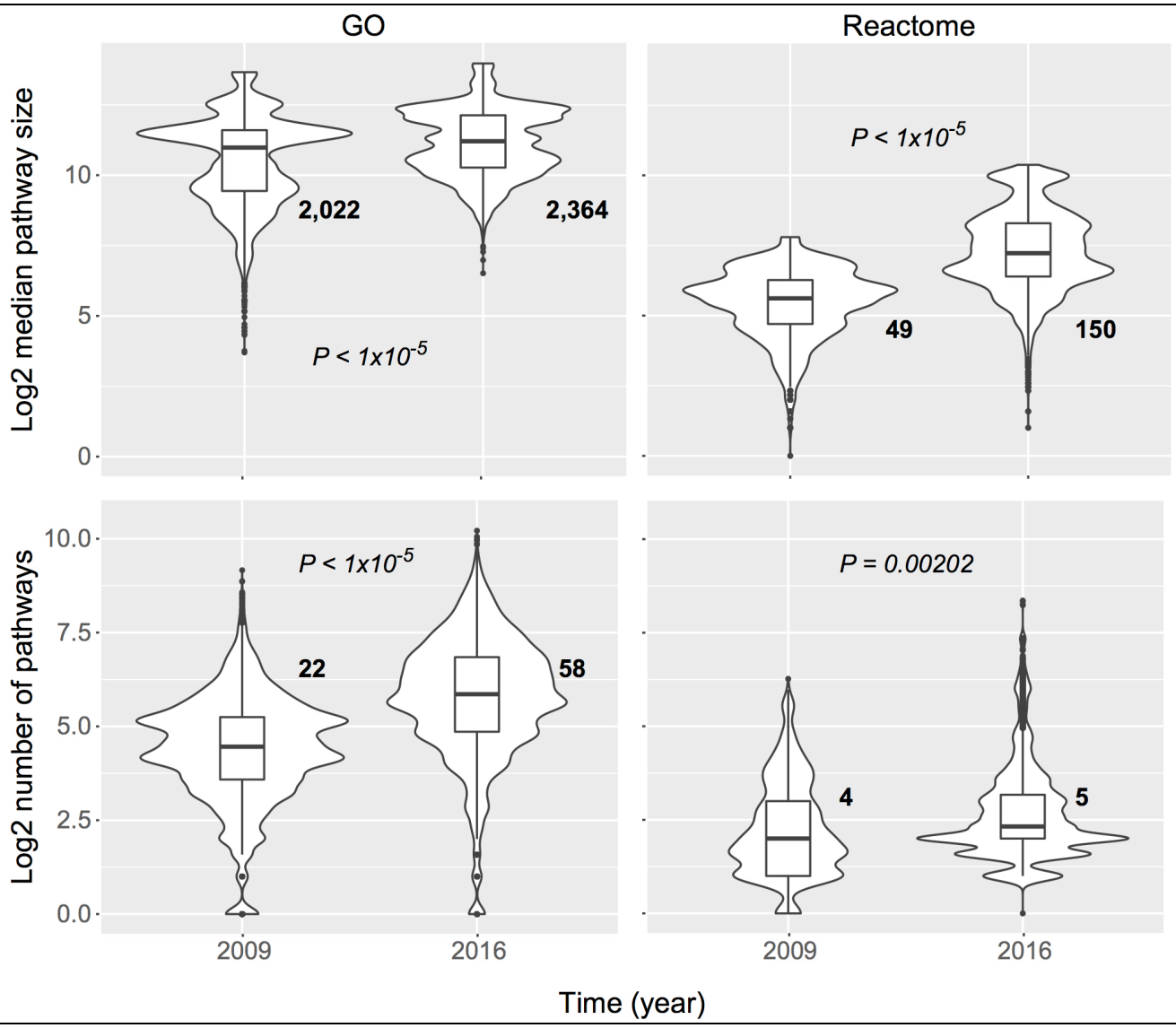
(A) The number of human biological processes and molecular pathways has doubled during 2009-2016. Similar trends are apparent among human cell components and molecular functions. We counted the number of GO terms and Reactome pathways with at least one annotated human gene. (B) The numbers of annotated GO terms have also grown rapidly for model organisms.



Supplementary Figure 3

Gene Ontology terms are increasingly specific and interconnected.

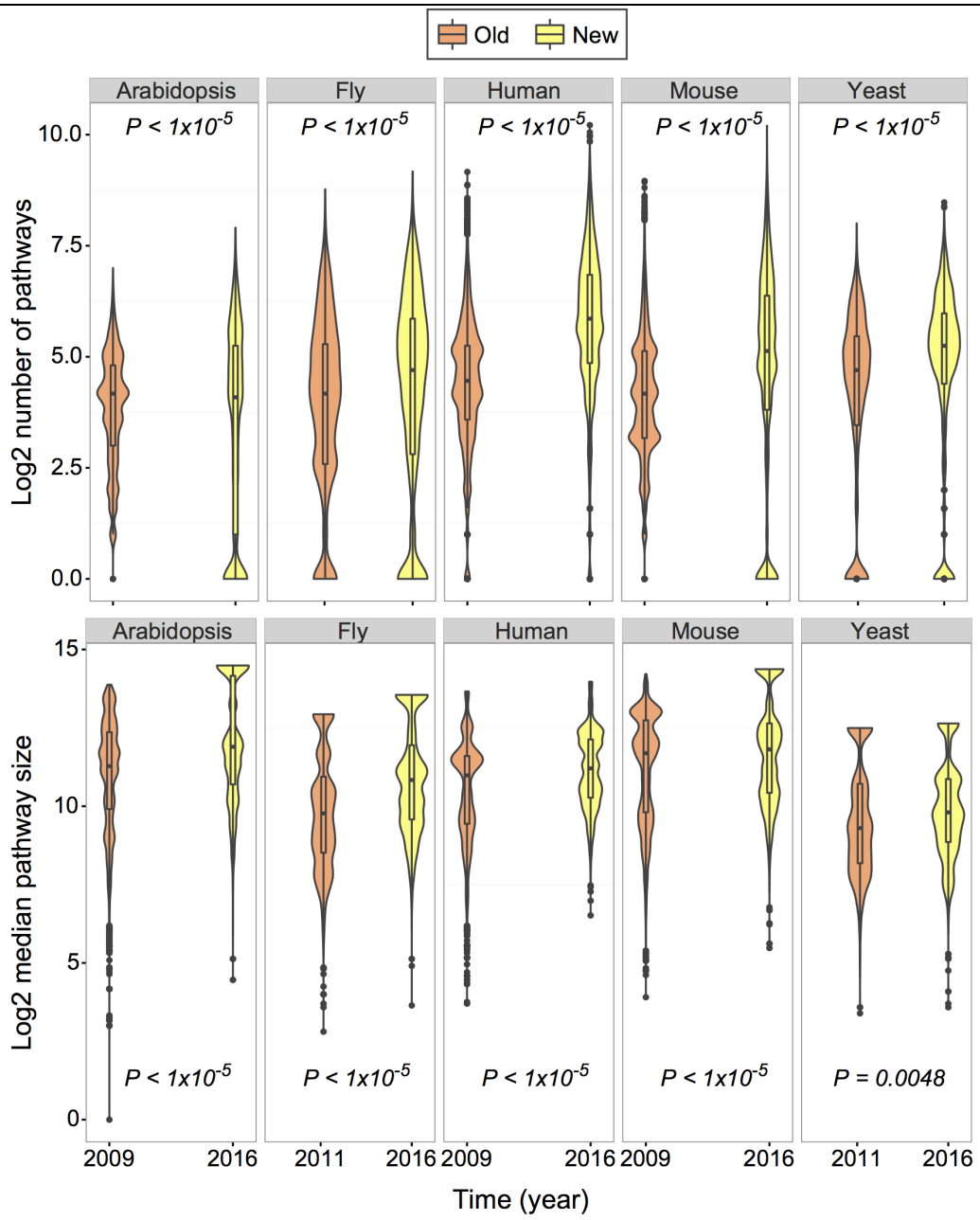
(A) Histogram shows mean length of paths in the Gene Ontology connecting a given term and the root term. Significant increase in the depth of the GO hierarchy between 2009 and 2016 ($P < 10^{-5}$, permutation test) indicates that the biological vocabulary is increasingly detailed and terms are becoming more specific. (B) The average number of parents per GO term has increased over time (2009-2016, 1.73 to 2.09; $P < 10^{-5}$). We used a permutation test (n = 100,000) to compute p-value to evaluate difference of earlier and recent values. Error bars represent 95% confidence intervals from resampling.



Supplementary Figure 4

Human genes and pathways are increasingly annotated.

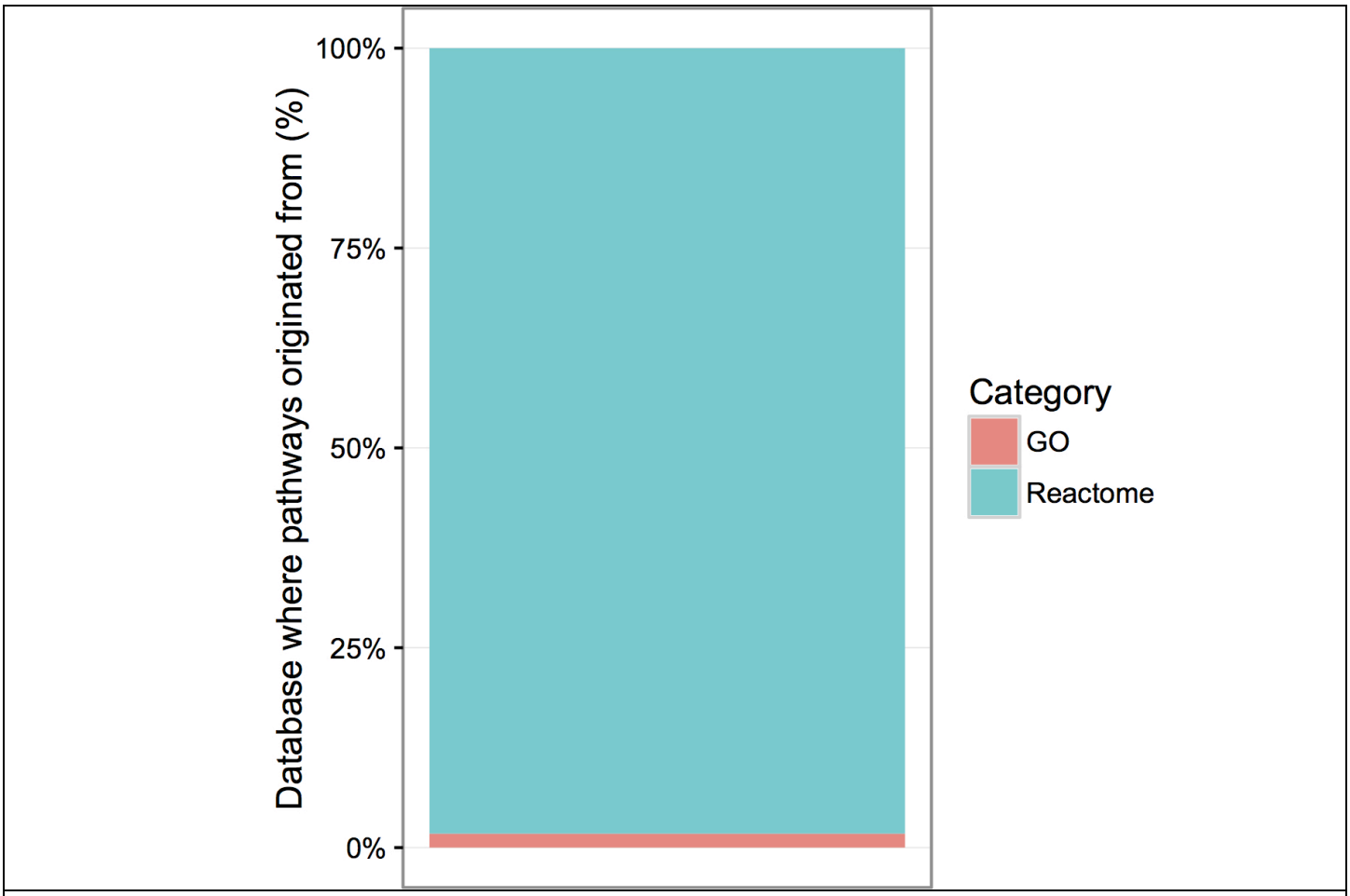
Violin plots show the comparison of pathway size and gene annotation frequency in 2009 and 2016. In the top panels, the median pathway size (total number of genes in pathway) is shown for every gene on log2 scale. In the bottom panels, number of pathways annotated per every gene is shown on log2 scale. P-values were computed using permutation tests (n=100,000). Genes without annotations were excluded. GO biological processes (left) and Reactome pathways (right) are shown separately. The median number of each plot is shown in boldface letters.



Supplementary Figure 5

Genes and pathways of model organisms are increasingly annotated.

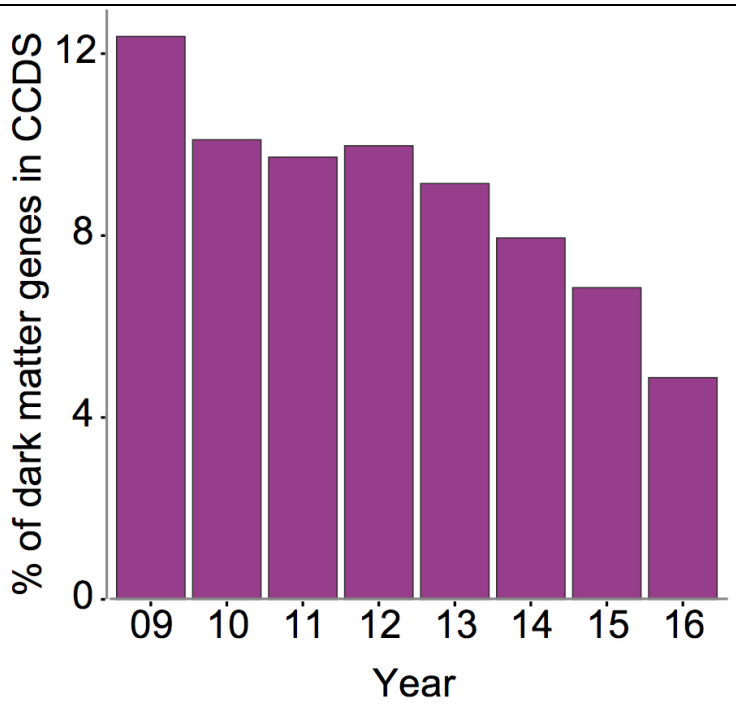
Violin plots show the comparison of pathway size and gene annotation frequency in 2009 and 2016 for human and several model organisms. P-values were computed using permutation tests (n = 100,000). Genes without annotations are excluded.



Supplementary Figure 6

Contemporary gene annotations include a prominent class of small and specific pathways from the manually curated Reactome resource.

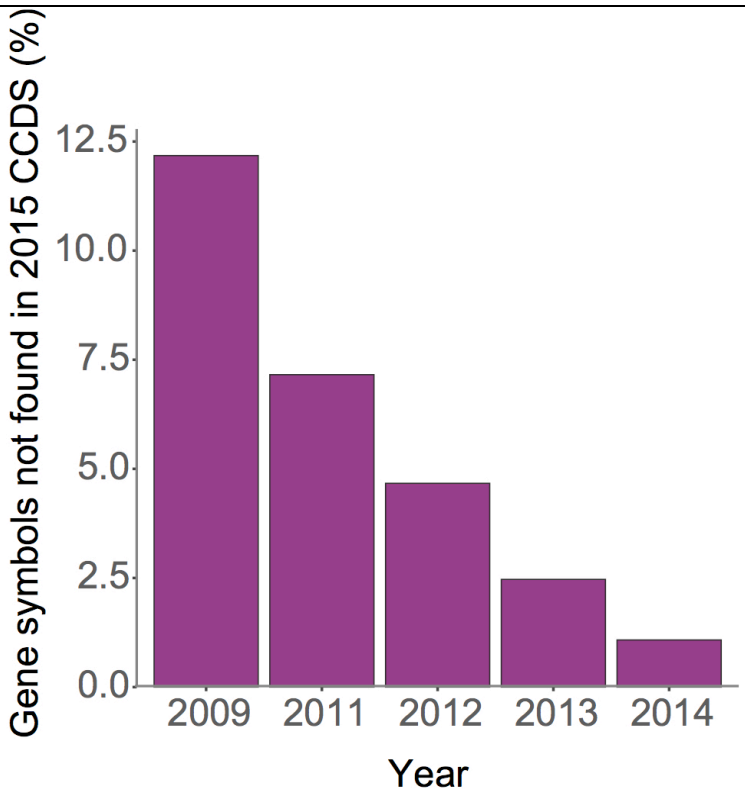
Two-dimensional density plots of median pathway size per gene and numbers of pathway annotations per gene (Fig. 1b) reveal a bimodal distribution of pathways in current annotations from 2016. The group of pathways in the bottom left quadrant of the left panel of Figure 1b primarily represents gene annotations of the Reactome resource (98%). The corresponding genes have relatively few annotations to pathways (below median value) and the pathways themselves contain relatively few genes (also below median value). The group of Reactome pathways is not apparent among annotations of 2009.



Supplementary Figure 7

The fraction of unannotated genes is decreasing consistently.

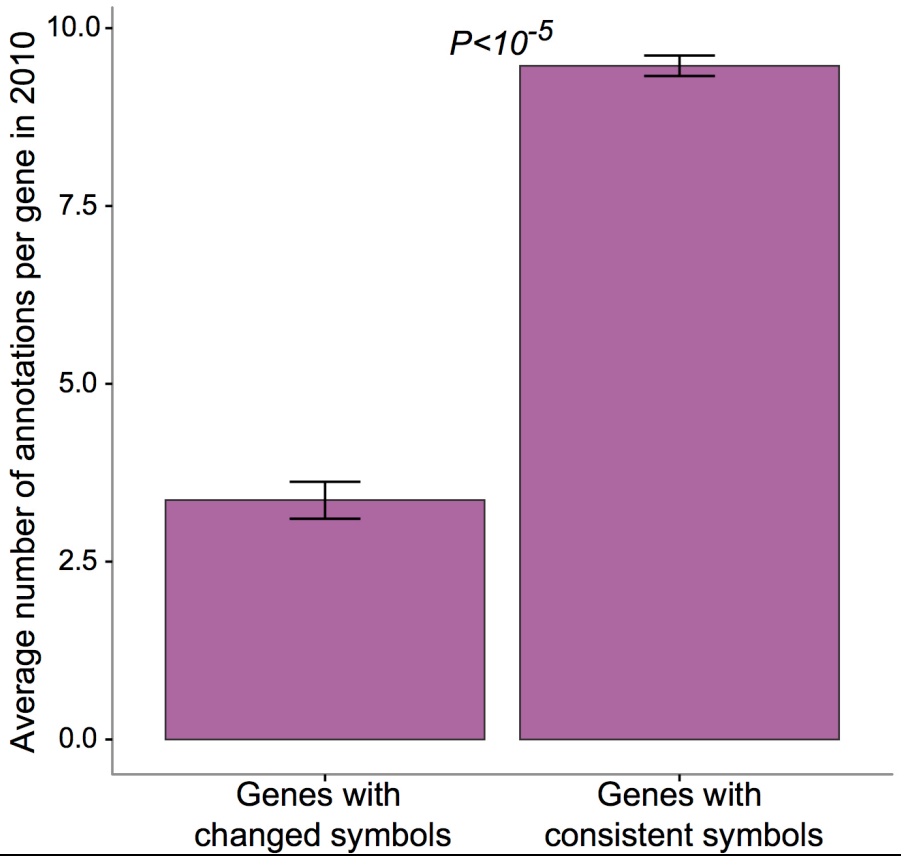
In 2009, one of eight high-confidence protein-coding genes (12.4%) from the CCDS database had no annotations in Gene Ontology or Reactome while this “dark matter” has decreased to 4.9% in 2016. Dark-matter genes included those with no annotations and also genes that only had root-level annotations in GO. We used the closest earlier release of the CCDS database to count annotated genes (e.g., the 2015 release of CCDS for 2016 annotations of GO, as CCDS of 2016 had not been released at the time of the analysis).



Supplementary Figure 8

Changes in gene nomenclature affect functional annotations.

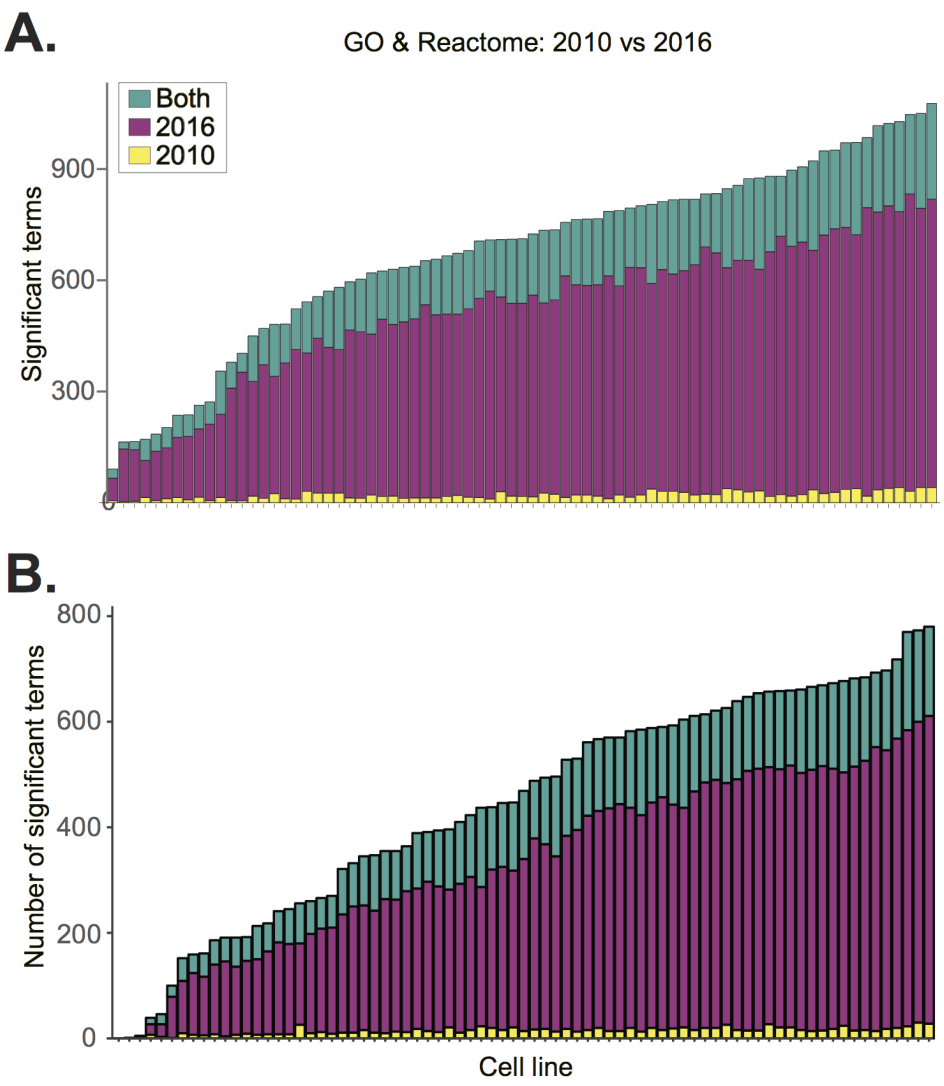
Analysis of human gene lists from current datasets will cause mismatches of gene symbols as standard nomenclature has been updated over the years. We compared the HGNC symbols in the latest CCDS database (2015) to earlier database versions and counted the number of unmatched symbols.



Supplementary Figure 9

Renamed genes have significantly fewer annotations.

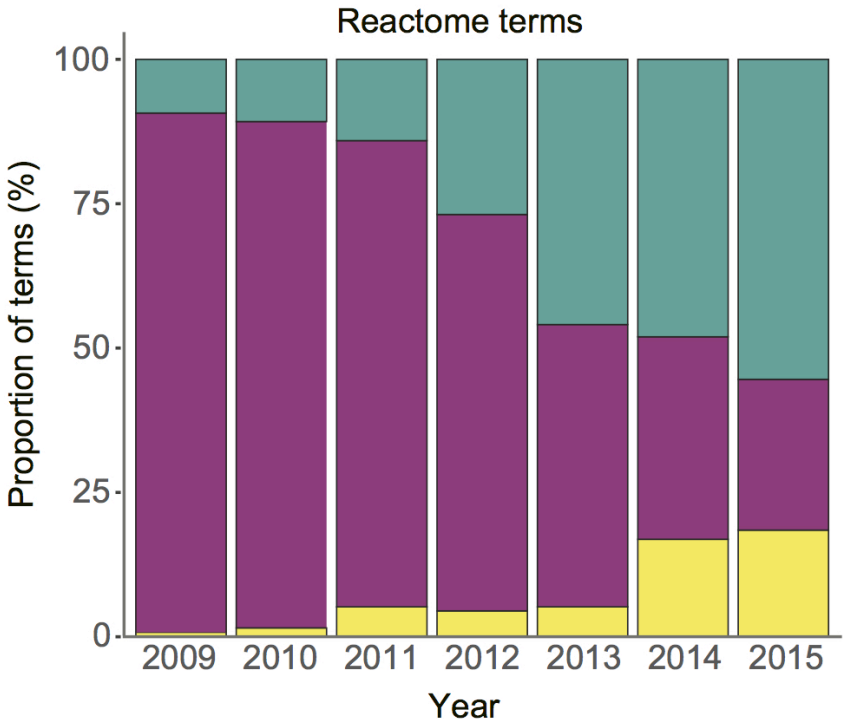
We investigated the number of annotations for genes whose symbols differed in 2010 and 2016 and found that genes with changed symbols have significantly fewer annotations in 2010 than consistently named genes (average number of annotations per gene 3.1 *vs* 9.3, permutation test of n=100,000, $p<10^{-5}$). Error bars represent 95% confidence intervals from resampling.



Supplementary Figure 10

Pathway enrichment analysis of essential genes of breast cancer confirms loss of information in outdated annotations.

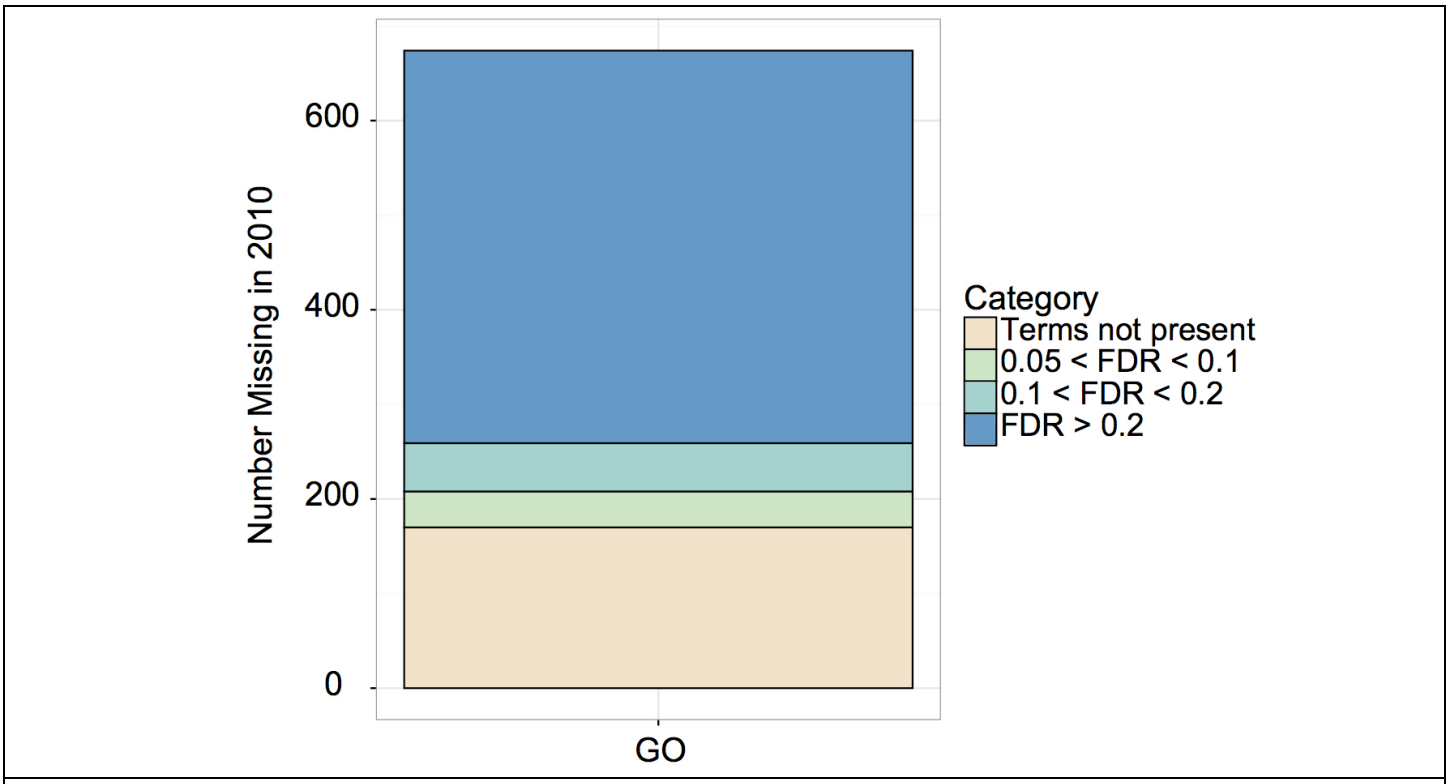
(A) We analysed top-500 essential genes from each of 77 cancer cell lines derived from recent shRNA screens. We studied (i) annotations from 2010 and (ii) annotations from 2016, and quantified enrichments using Fisher's exact test and multiple testing correction (FDR $P < 0.05$). We then compared the resulting enriched terms from both analyses. We found a three-fold increase in detected pathways and processes when data were analyzed with current annotations from 2016 (695 pathways and processes per median cell line) compared to outdated annotations from 2010 (191 per median cell line, 74% missed when accounting for terms only appearing in 2010 annotations). GO biological processes and Reactome pathways were analyzed and respective counts are aggregated in the plot. (B) We repeated our pathway enrichment analysis of breast cancer essential genes by analyzing top-100 essential genes of the same dataset and found a similar difference of the effect of outdated and current pathway annotations (143 v 455 pathways, 71% missed in earlier annotations).



Supplementary Figure 11

Evolution of pathway information affects recently updated and out-of-date software tools.

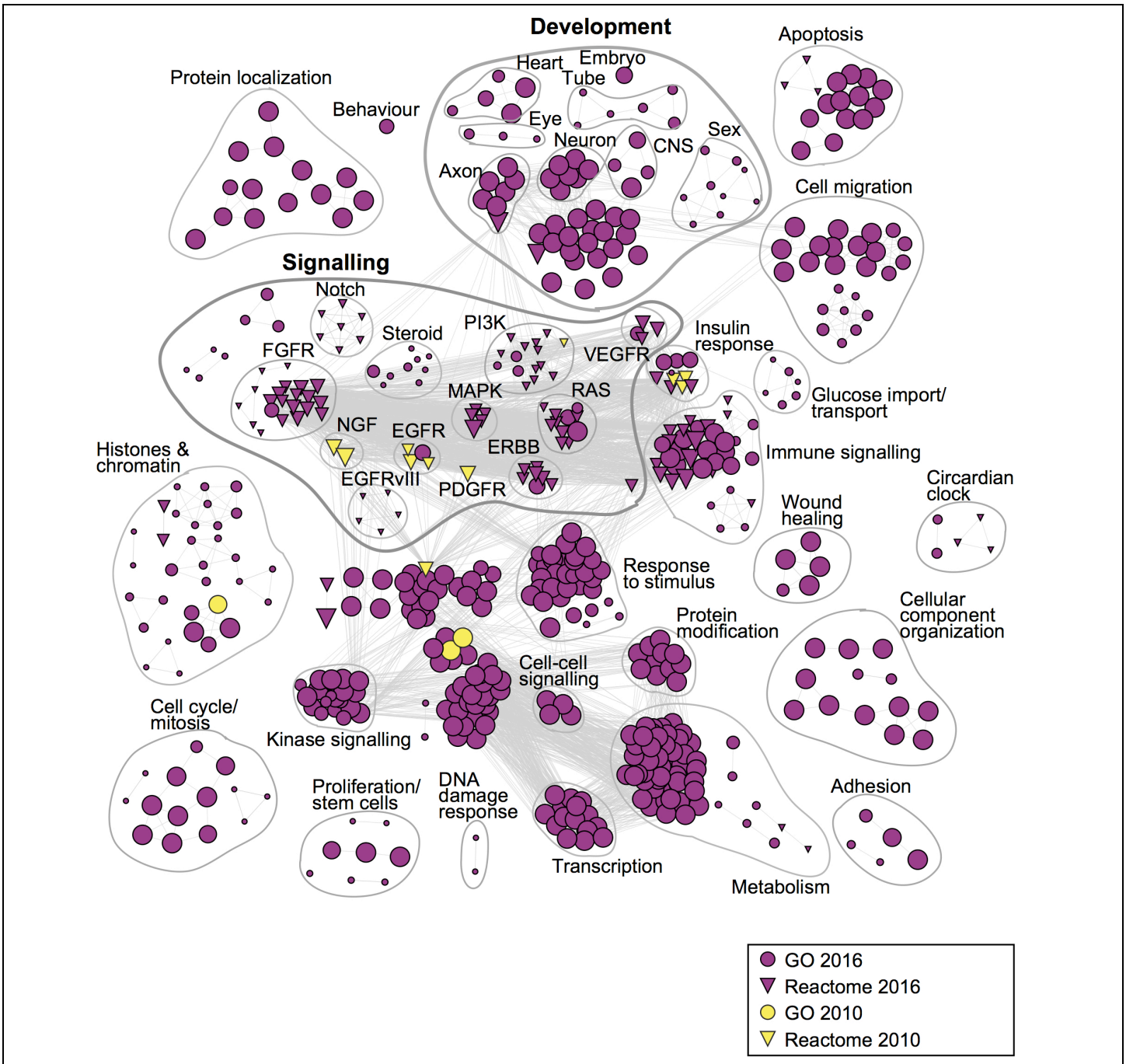
We analyzed significantly mutated driver genes of glioblastoma using gene annotations of 2009-2016 and compared the results of 2016-era analysis with results of each earlier year. Colors indicate the fraction of commonly detected (yellow), 2016-only (purple) and outdated-only (dark blue) pathways from the Reactome resource with statistically significant enrichment (FDR $p<0.05$).



Supplementary Figure 12

Most missed GO terms in 2010-era analysis involve known pathways and processes that do not associate significantly with input genes.

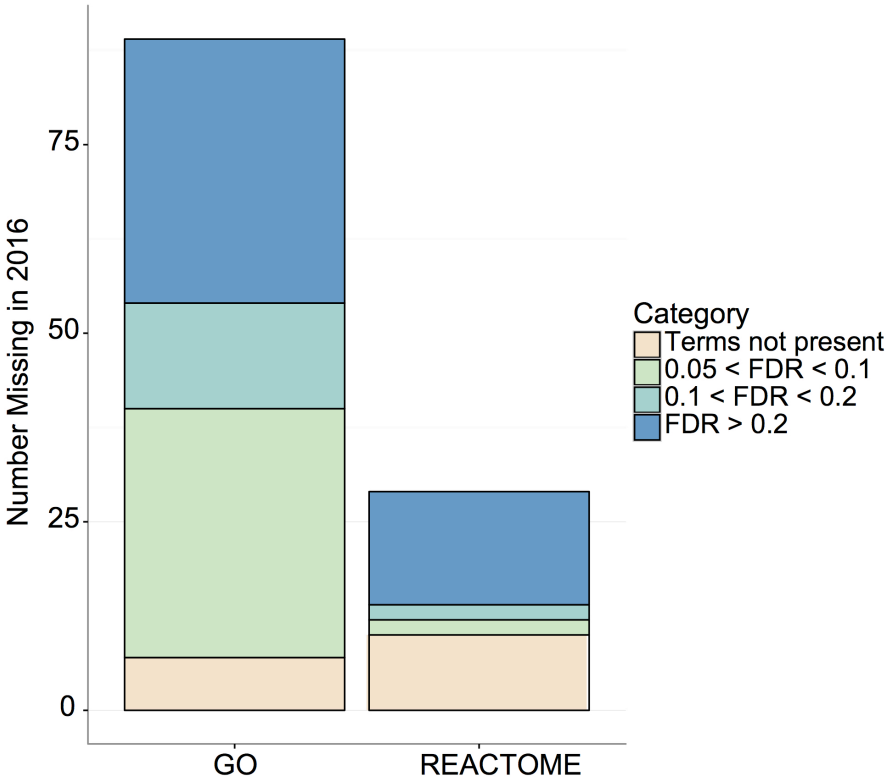
We compared results of pathway enrichment analyses that used annotations from 2010 and 2016. The majority of pathways missed in the outdated annotations (~75%) exist in the 2010 edition of Gene Ontology, however these are not significantly associated to input genes. The remaining 25% represent processes added to GO after 2010.



Supplementary Figure 13

Most pathway enrichments from outdated annotations from 2010 are based on low-quality information.

We repeated the pathway analysis of frequently mutated glioblastoma genes by only analyzing high-quality gene annotations from 2010 and 2016 (IEA annotations from GO were excluded). We found that 96.5% of results from 2016 analysis were missed when 2010 annotations were used, showing that earlier annotations are largely based on low-confidence information.



Supplementary Figure 14

Some enriched pathways and processes are missed in relatively recent gene annotations.

We compared results of pathway enrichment analyses that used annotations from 2015 and 2016. We focused on the terms that were found in the earlier analysis and missed in the most current annotations, including 89/743 (12%) GO terms and 29/116 (25%) Reactome pathways. The majority of missing pathways were part of the pathway database or GO in the up-to-date analysis although not detected at statistically significant levels (blue tones), while a smaller fraction of terms were entirely missing from the analysis, likely because of restructuring of pathways and processes.

Nature Methods: doi:10.1038/nmeth.3963

Supplementary Methods

Collection of pathway tools and citations. We collected web-based pathway enrichment analysis tools that included human gene annotations using literature search and recent review papers (refs 1,2,3). Tools were compared only in the context of gene annotation datasets and not their analytic capacity. The times of the most recent data updates were collected from the websites of selected tools. Developers were contacted if this information was unavailable. We collected primary publications of software and methods from the websites and used Google Scholar to find additional publications. We conducted a PubMed query for each selected publication, determined its PubMed ID and counted its citations in 2015 by defining a custom time range on publication date (01.01.2015-31.12.2015). Citations were summed across multiple publications for the same software tools.

Ontologies and pathways. Functional terminology of biological processes, molecular functions, and cell components was retrieved from the Gene Ontology[4] website (http://geneontology.org/page/download-ontology) and comprised January releases of each year (2009-2016; **Supplementary Table 3**). February release of GO was used for 2012 as the January release was unavailable. Five relationship types were considered (is_a, part_of, regulates, positively_regulates, negatively_regulates). Gene annotations were derived from the Gene Ontology Annotation database[5] (UniProt-GOA, http://www.ebi.ac.uk/GOA/archive). We selected GOA datasets that were released shortly after the corresponding GO ontologies. Genes were annotated to GO terms as well as parent and ancestor terms via all possible paths. Obsolete terms and negative relationships in GO were removed. Molecular pathways from the Reactome[6] database were retrieved from archives and included December releases of previous years. We filtered human genes with non-public status and analyzed protein-coding genes of matched versions of the NCBI Consensus Coding Sequence Database[7] (CCDS, https://www.ncbi.nlm.nih.gov/CCDS). The closest earlier release of CCDS was used for each annotation dataset. Average path lengths and parent counts of GO terms in 2009 and 2016 were evaluated with permutation tests. Terms were permuted uniformly for 100,000 times and simulated mean values derived from permuted terms were compared with actual mean values derived from observed terms. Each p-value was computed as the number of permutations where values from simulated data exceeded values from observed data over the total number of permutations. Pathways, annotations, and enrichments were analysed with custom R scripts available on request.

Analysis of gene annotations. Pathway databases were analyzed for growth in total number of pathway terms separately for the three main ontologies in GO (biological processes, molecular functions, cell components) for each year (2009-2016). The same analysis was repeated for human Reactome pathways and GO annotations for model organisms (mouse, Arabidopsis thaliana, fly, yeast). We counted GO terms and Reactome pathways with at least one annotated gene of the studied species. Path lengths and numbers from terms to roots were computed with custom scripts. Human annotations of GO terms and Reactome pathways contained high-confidence protein-coding genes from the nearest previous release of the CCDS database release (e.g. 2015 release for 2016 annotations) and only included genes with public status. Density of human gene annotations was assessed with two-dimensional density plots. For each gene, number of associated processes and pathways (Y-axis) and median size of corresponding gene sets per gene (Y-axis) were shown. The density plots include non-annotated genes (i.e. “dark matter”) for density estimation but are not shown. Dark matter genes were selected as protein-coding genes of the corresponding CCDS

release that had no annotations in GO or Reactome, or only had top-level GO annotations (one or more of biological_protess, cell_component, molecular_function). GO biological processes per gene were also estimated for model organisms (mouse, A. thaliana, fly, yeast) without filtering of CCDS genes and dark matter. The proportion of missing gene symbols was estimated from earlier CCDS releases relative to the most recent CCDS release of 2015. Quality of human gene annotations was assessed in three mutually exclusive categories – genes with at least one Reactome annotation, genes with at least one non-electronic (non-IEA) annotation in GO, and genes with only IEA (Inferred from Electronic Annotation) annotations in GO. Statistical comparisons of gene annotations across years were conducted with permutation tests similarly to the tests described above.

Pathway Enrichment Analysis. Pathway enrichment analysis was conducted on GO biological processes and Reactome pathways using Fisher's exact tests. GO cell components and molecular functions were excluded prior to analysis. Multiple testing was conducted separately for GO and Reactome terms using the Benjamini-Hochberg False Discovery Rate (FDR) procedure. Terms with FDR $p<0.05$ were considered significant. Enrichment analysis of GO and Reactome terms conservatively comprised separate background gene sets that included all the genes with at least one gene annotation of biological process (GO) and Reactome pathway, respectively. We chose this general enrichment strategy and did not compare outputs of different tools directly, as direct comparison would be confounded by differences in underlying methods, gene symbol mapping, and filtering of annotation data. Two sets of enrichment analyses were conducted on cancer gene lists using gene annotations from 2016 (corresponding to g:Profiler) and 2010 (corresponding to DAVID). First we analyzed essential breast cancer genes from recent shRNA screens of 77 cell lines[8]. We separately analyzed top-100 and top-500 lists of genes according to per-gene zGARP scores provided by the study. We counted shared, outdated-only, and recent-only gene annotations enriched in the analyses (FDR $p<0.05$) and matched these using GO and Reactome term identifiers. The most common terms only found in the up-to-date analysis were visualized with the WordCloud R package. To simulate practical analysis, we did not manually convert outdated gene symbols in breast cancer analysis. We also conducted enrichment analyses of essential breast cancer genes after converting these to earlier HGNC symbols via EntrezGene identifiers in the CCDS database. We observed no major differences of analyses performed with converted and unconverted gene symbols. To confirm this observation, we studied the distributions of annotations of genes and found that the genes whose symbols changed between 2010 and 2016 had significantly fewer annotations than the genes with consistent symbols, and 42% of these corresponded to 'dark matter' genes with no annotations in 2010. A similar pathway enrichment analysis was conducted for 75 frequently mutated glioblastoma genes[9,10] derived from the IntOGen database[11]. In this analysis, we manually mapped outdated gene symbols to create a more conservative scenario. We compared enriched annotations across the years 2009-2015 relative to 2016 and counted common and distinct terms as above. We studied the origin of these differences in the context of changed annotations and changed functional vocabulary. We also visualized pathway enrichments of 2010 and 2016 using the Enrichment Map[12] app of Cytoscape[13]. The enrichment map covered pathways with at least four genes. Our observations were also confirmed when this filter was removed and all pathways were included. Functional themes and signaling pathways in the map were curated manually.

Supplementary Note | Pathway analysis of glioblastoma genes highlights specific processes and druggable pathways missed by earlier annotations

Detailed summary of GBM pathway enrichment results as an Enrichment Map[12] shows biological themes that are missed when using annotations from 2010 (FDR $P < 0.05$, **Figure 1e; Supplementary Tables 4-5**). GBM driver genes are enriched in hallmark cancer processes[14] including cell cycle, apoptosis, cell migration, and signaling. These are apparent when analyzing either out-of-date or current annotations, however only few general terms appear in 2010 data.

Enriched pathways in up-to-date analysis are more specific to neuronal context as expected from brain cancer genes. For example, while apoptosis is found in both analyses, neuronal apoptosis only appears in newer analysis (n=7 genes, FDR P = 0.018). Similarly, current gene annotations emphasize central nervous system (CNS) development (n = 22, FDR $P = 1.63 \times 10^{-8}$) as well as neurogenesis (n = 26, FDR $P = 1.12 \times 10^{-6}$) and gliogenesis (n = 7, FDR P = 0.025), while only CNS development is apparent in older data. Functional themes such as immune response (n = 29, FDR $P = 5.2 \times 10^{-5}$), neurotransmitter signaling (n = 6, FDR P = 0.0013), circadian clock (n = 8, FDR $P = 1.5 \times 10^{-4}$) and glucose signaling (n = 7, FDR P = 0.0016) are only highlighted in new annotations. These processes are expected in the context of current knowledge[15,16], for example enhanced glucose uptake of brain tumor initiating cells helps these overcome nutrient deprivation[17].

Current pathway analysis also highlights specific signaling pathways relevant to GBM biology and therapy development[18,19,20]. For example, Notch (n = 5, FDR P = 0.0019), TGF-β (n = 5, FDR P = 0.027), and fibroblast growth factor (n = 12, FDR $P = 1.13 \times 10^{-6}$) pathways are only enriched among up-to-date gene annotations and reveal translational hypotheses. Notch is targetable with γ-secretase inhibitors (*e.g.* R04929097; Roche) for malignant glioma and for progressive GBM that are currently in phase I and phase II trials, respectively[18]. Drugs of the TGF-β pathway inhibit the ligand (Trabedersen, Antisense Pharma) or the receptor of the signaling cascade (Galunisertib, Eli Lilly)[21]. The current gene annotations also highlight the enrichment of EGFRvIII[22,23] signaling pathway (*EGFR*, *KRAS*, *PIK3CA*, *PIK3R1*, *SOS1*, FDR $P = 1.07 \times 10^{-5}$). Aberrant EGFR signaling is common to many cancer types, however EGFR kinase inhibitors have been unsuccessful in GBM treatment to date. *EGFR* alterations occur in more than 50% of GBMs[20] and the most common alteration causes deletion of exons 2-7 of the gene, known as the EGFRvIII variant that drives tumor progression and correlates with poor prognosis[24]. The recently developed Rindopepimut vaccine targets EGFRvIII and has entered clinical trial[25].

These examples demonstrate the limitations of outdated gene annotations. Glioblastoma has extremely poor outcome as the average patient only survives 15 months after diagnosis regardless of surgery and aggressive chemotherapy. The specific processes and pathways with existing drugs highlight avenues for functional follow-up experiments and candidates for future therapy development. Researchers who use outdated software for analyzing their experimental data will miss out on relevant functional and translational hypotheses.

Supplementary Table 1 | Data update times of web-based pathway enrichment analysis software tools.
The table shows 25 web-based pathway enrichment analysis tools and the dates of their most recent updates of gene annotation databases. The table reflects information from the web sites of corresponding tools and was compiled in February 2016. The list of tools was compiled from review papers and additional literature searches. We only included tools that allowed analysis of human genes.

Tool	URL	Last update of pathway database	Reference
GORILLA	http://cbl-gorilla.cs.technion.ac.il/	02-2016	26
g:Profiler	http://biit.cs.ut.ee/gprofiler/	02-2016	27
ToppGene	https://toppgene.cchmc.org/	02-2016	28
PANTHER	http://pantherdb.org/	02-2016	29
InterMine Human Mine	http://intermine.org/ http://www.humanmine.org/	11-2015	30
GoEast	http://omicslab.genetics.ac.cn/GOEAST/	11-2015	31
GeneMerge	http://www.genemerge.net/	09-2015	32
ConsensusPathDB	http://consensuspathdb.org/	09-2015	33
GREAT	http://bejerano.stanford.edu/great/public/html/	02-2015	34
Babelomics	http://babelomics.bioinfo.cipf.es/	12-2014	35
Enrichr	http://amp.pharm.mssm.edu/Enrichr/	11-2014	36
FuncAssociate	http://llama.mshri.on.ca/funcassociate/	06-2014	37
gsGator	http://gsgator.ewha.ac.kr/	01-2014	38
WebGestalt	http://bioinfo.vanderbilt.edu/webgestalt/	01-2013	39
GeneCodis	http://genecodis.cnb.csic.es/	12-2011	40
GoMiner	http://discover.nci.nih.gov/gominer/	01-2011	41
GeneTrail	http://genetrail.bioinf.uni-sb.de/	09-2010	42
EasyGO	http://bioinformatics.cau.edu.cn/easygo/	05-2010	43

GARNet	http://biome.ewha.ac.kr:8080/GSEAWebApp/	05-2010	44
DAVID	https://david.ncifcrf.gov/	01-2010	45
ConceptGen	http://conceptgen.ncibi.org/core/conceptGen/	11-2009	46
GOToolBox	http://genome.crg.es/GOToolBox/	07-2009	47
L2L	http://depts.washington.edu/l2l/	07-2007	48
GoSurfer	http://systemsbio.ucsd.edu/GoSurfer/	03-2007	49
GOstat	http://gostat.wehi.edu.au/	NA	50

Supplementary Table 2 | Citation counts of pathway enrichment analysis software in 2015.

The table shows 25 different pathway tools, the number of citations per publication, the total number of citations of the tool, and the percentage of all citations in 2015. Lists of primary publications of software tools were collected from respective web sites and augmented with additional literature searchers. Citation counts of these papers were derived from PubMed.

Tool	Citation count 2015 per publication	PMID	Total citations in 2015	Total citations in %
Babelomics	2	16845052	**44**	1.13
	5	15980512		
	17	20478823		
	3	18515841		
	2	25897133		
	15	14990455		
ConceptGen	6	20007254	**6**	0.15
ConsensusPathDB	16	18940869	**79**	2.04
	25	21071422		
	38	23143270		
	0	20847220		
DAVID	1417	19131956	**2517**	**64.89**
	261	12734009		
	614	19033363		
	28	14519205		
	72	17576678		
	57	17784955		
	16	17980028		
	21	19728287		
	23	22543366		
	8	18841237		
EasyGO	3	17645808	**3**	0.08
Enrichr	61	23586463	**61**	1.57
FuncAssociate	13	19717575	**21**	0.54

	8	14668247		
g:Profiler	26	17478515	**74**	1.91
	48	21646343		
GARNet	0	21342555	**0**	0
GeneCodis	31	17204154	**108**	2.78
GeneMerge	4	12724301	**4**	0.10
	28	19465387		
	45	22573175		
GeneTrail	12	17526521	**14**	0.36
	2	21592396		
GoEast	32	18487275	**40**	1.03
	8	19615110		
GoMiner	26	12702209	**33**	0.85
	7	15998470		
GORILLA	134	19192299	**134**	3.45
GoSurfer	1	15702958	**1**	0.03
GOstat	24	14962934	**24**	0.62
GOToolBox	3	15575967	**3**	0.08
GREAT	142	20436461	**145**	3.74
gsGator	0	24423189	**0**	0.00
	3	23814184		
InterMine	22	22434830	**50**	1.29
	17	17615057		
	9	22080565		
	1	26092688		
	1	25414324		
L2L	0	16168088	**0**	0
PANTHER	16	19597783	**242**	6.24
	8	17130144		
	132	23193289		
	86	23868073		
ToppGene	72	19465376	**72**	1.86
Webgestalt	89	15980575	**204**	5.26

	113	23703215		
	2	24233776		
	0	26656494		
	20	18511468		
TOTAL			**3879**	100

Supplementary Table 3 | GO, Reactome, and CCDS databases used in the study. The following versions of GO, Reactome, and CCDS databases were used for each year to analyze data described in this manuscript.

Year	GO ontology data version	GO annotations version	Reactome version	CCDS version
2009	NA	70	27	2009-09-02
2010	1.1.939	81	31	2009-09-02*
2011	1.1.1689	93	35	2011-09-07
2012	1.1.2572	106	39	2012-10-25
2013	2012-12-31	117	43	2013-11-29
2014	2013-12-20	129	47	2014-08-07
2015	2014-12-22	140	51	2015-05-12
2016	2015-12-22	152	55	2015-05-12*

*CCDS database from previous year was used as no CCDS is available for that year.

Supplementary Table 4 | Enriched pathways and processes of glioblastoma genes that remain undetected in 2010-era annotations.
We found nine major themes (i.e. groups of related pathways and processes) out of 28 that were only detected as significant when analyzing recent annotations from 2016. Similarly, new sub-pathways (spread across different major themes) were discovered only when using up-to-date gene annotations. Color indicates type of detected pathway: pathways only discovered in 2016-era annotations (pink), and pathways commonly discovered in annotations of 2010 and 2016 (yellow).

Major themes/ pathways	Sub-pathways
Catabolism	Protein catabolism, RNA catabolism
GABAergic synaptic transmission	Synaptic plasticity, neurotransmitter transport
Cognition/ learning/ behaviour	Visual behaviour, associative learning, memory, cognition
Glucose import/ transport	Carbohydrate homeostasis, response to glucose stimulus
Circadian clock	Regulation of circardian rhythm, BMAL1:CLOCK:NPAS2 activates circardian gene expression
Wound healing	Coagulation, platelet activation, homeostasis
Immune signalling	Fc receptor signalling pathway, TCR signalling, cytokine signalling
Homeostasis	Chemical homeostasis, tissue homeostasis
Endocytosis	Vesicle-mediated transport, receptor-mediated endocytosis
Adhesion	Cell-matrix adhesion, cell-substrate adhesion
Response to stimulus	Response to cAMP/ purine-containing compound
	Response to EGF
	Response to UV
	Response to stress
	Response to metal ion/ inorganic substances
	Response to TGFβ
Cellular component organization	RNP complex biogenesis/ assembly
	Negative regulation of organelle / cellular component organization
Histones & chromatin	H3-K9 methylation
	Histone acetylation/ peptidyl-lysine acetylation
	Chromatin (dis)assembly

	DNA alkylation/ (de)methylation
	Histone H3 acetylation
	Histone deacetylation
	Peptidyl-lysine acetylation
	Nucleosome organization
	Histone H3-K9 acetylation
	Chromatin modifying enzymes
DNA replication/ CC checkpoint	Regulation of DNA replication
Cell cycle/ mitosis	mitotic CC phase transition
	Chromosome segregation/ meta-anaphase transition
	Sister chromatid cohesion
	Sister chromatid segregation
Apoptosis/ Neuron death	Neuron death regulation
	Chromosome breakage/ programmed DNA elimination
	Fibroblast apoptotic process
	Neuron apoptotic process
	Cell death signalling via NRAGE/ NRIF/ NADE
	Mesenchymal apoptotic process
	p75 NTR receptor-mediated signalling
	Cell type specific apoptotic process
Protein import/ localization	Protein localization to membrane
	ECM organization
	Potassium ion transmembrane transport
	protein import into nucleus
Metabolism	Regulation of ROS
	Regulation of TNF
Signalling	EGFRvIII
	FGFR
	Notch
	VEGF

	ERK1 and ERK2 cascade
	ERBB
	TGFβ
Development	Tube
	Neuro/ axono/ gliogenesis
	Cartilage
	Head/ body/ face
	Trachea
	Eye
	Sertoli cell
	Liver
	Pattern specification

Supplementary Table 5a | GBM-associated GO terms of the 2016 analysis compared to 2010.
The table shows the top most significant GO terms in 2016 ranked by FDR adjusted p-values compared to FDR p-values detected in 2010.

GO.ID	Description	FDR	Common genes	FDR of GO.ID in 2010
GO:0009893	positive regulation of metabolic process	2.69E-14	AKAP9,ANK3,ARFGEF2,ARHGAP35,ARHGEF6,ARID1A,ATRX,BPTF,BRAF,BRCA1,CASP1,CHD8,CLOCK,CNOT1,CUL1,DIS3,EGFR,EZH2,FN1,HSP90AB1,KALRN,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAP4K3,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TP53,TRIO,WT1,KMT2A	0.005040439
GO:0010604	positive regulation of macromolecule metabolic process	2.61E-13	AKAP9,ANK3,ARID1A,ATRX,BPTF,BRAF,BRCA1,CASP1,CHD8,CLOCK,CNOT1,CUL1,EGFR,EZH2,FN1,HSP90AB1,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAP4K3,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TP53,WT1,KMT2A	0.024811057
GO:0070887	cellular response to chemical stimulus	5.20E-13	ADAM10,AKAP9,ANK3,ARHGEF6,ATRX,BPTF,BRAF,BRCA1,CAD,CASP1,CUL1,EGFR,EZH2,FN1,HDAC9,HSP90AB1,KALRN,KDR,KRAS,LRP6,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,NUP107,PIK3CA,PIK3CB,PIK3R1,PRPF8,PTEN,PTPN11,RB1,SIN3A,SOS1,SOX9,SPTAN1,TGFBR2,TP53,TRIO,WT1,KMT2A	**ND**

GO:0048856	anatomical structure development	2.30E-12	ADAM10,AKAP9,ANK3,ARHGAP35,ARID1A,ATRX,BPTF,BRAF,BRCA1,CAD,CARM1,CHD8,CLOCK,CNOT1,CSDE1,CUL1,EGFR,EZH2,FAT1,FN1,HDAC9,HSP90AB1,IDH1,KALRN,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PBRM1,PCDH18,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,TGFBR2,TJP1,TP53,TRIO,WT1,RPSA,KMT2A	0.002983586
GO:0019222	regulation of metabolic process	2.70E-12	ACAD8,AKAP9,ANK3,ARFGEF2,ARHGAP35,ARHGEF6,ARID1A,ARID2,ATRX,BPTF,BRAF,BRCA1,CARM1,CASP1,CHD8,CLOCK,CLTC,CNOT1,CSDE1,CUL1,DIS3,EGFR,EZH2,FN1,HDAC9,HSP90AB1,IDH1,KALRN,KDM5C,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAP4K3,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,NUP107,PAX5,PBRM1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TP53,TRIO,WT1,ZNF814,KMT2A	0.000413789
GO:0071840	cellular component organization or biogenesis	2.75E-12	ADAM10,AKAP9,ANK3,ARFGEF2,ARHGAP35,ARHGEF6,ARID1A,ARID2,ATRX,BAP1,BPTF,BRAF,BRCA1,CARM1,CHD8,CLOCK,CLTC,CNOT1,DIS3,EGFR,EZH2,FAT1,FN1,HDAC9,HSP90AB1,KALRN,KDM5C,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NUP107,PAX5,PBRM1,PIK3CA,PIK3CB,PIK3R1,PRPF8,PTEN,PTPN11,RB1,RPL5,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TJP1,TP53,TRIO,WT1,RPSA,KMT2A	**ND**

GO:0044260	cellular macromolecule metabolic process	2.75E-12	ACAD8,ADAM10,AKAP9,ANK3,AQR,ARFGEF2,ARHGAP35,ARHGEF6,ARID1A,ARID2,ATRX,BAP1,BPTF,BRAF,BRCA1,CAD,CARM1,CASP1,CHD8,CLOCK,CLTC,CNOT1,CSDE1,CUL1,DIS3,EGFR,EZH2,FN1,HDAC9,HSP90AB1,KALRN,KDM5C,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAP4K3,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,NUP107,PAX5,PBRM1,PIK3CA,PIK3CB,PIK3R1,PRPF8,PTEN,PTPN11,RB1,RPL5,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TP53,TRIO,WT1,ZNF814,RPSA,KMT2A	3.94E-06
GO:0007275	multicellular organismal development	5.30E-12	ADAM10,AKAP9,ANK3,ARHGAP35,ARID1A,ATRX,BPTF,BRAF,BRCA1,CAD,CARM1,CHD8,CLOCK,CNOT1,CSDE1,CUL1,EGFR,EZH2,FN1,HDAC9,HSP90AB1,IDH1,KALRN,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PBRM1,PCDH18,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,TGFBR2,TJP1,TP53,TRIO,WT1,KMT2A	0.000751915
GO:0016043	cellular component organization	5.30E-12	ADAM10,AKAP9,ANK3,ARFGEF2,ARHGAP35,ARHGEF6,ARID1A,ARID2,ATRX,BAP1,BPTF,BRAF,BRCA1,CARM1,CHD8,CLOCK,CLTC,CNOT1,EGFR,EZH2,FAT1,FN1,HDAC9,HSP90AB1,KALRN,KDM5C,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NUP107,PAX5,PBRM1,PIK3CA,PIK3CB,PIK3R1,PRPF8,PTEN,PTPN11,RB1,RPL5,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TJP1,TP53,TRIO,WT1,RPSA,KMT2A	1.6E-04

GO:0048731	system development	6.03E-12	ADAM10,AKAP9,ANK3,ARHGAP35,ARID1A,ATRX,BPTF,BRAF,BRCA1,CAD,CARM1,CHD8,CLOCK,CSDE1,CUL1,EGFR,EZH2,FN1,HDAC9,HSP90AB1,IDH1,KALRN,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PBRM1,PCDH18,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SIN3A,SOS1,SOX9,SPTAN1,TGFBR2,TP53,TRIO,WT1,KMT2A	0.005167947
GO:0032502	developmental process	6.07E-12	ADAM10,AKAP9,ANK3,ARHGAP35,ARID1A,ATRX,BPTF,BRAF,BRCA1,CAD,CARM1,CHD8,CLOCK,CLTC,CNOT1,CSDE1,CUL1,EGFR,EZH2,FAT1,FN1,HDAC9,HSP90AB1,IDH1,KALRN,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PBRM1,PCDH18,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TJP1,TP53,TRIO,WT1,RPSA,KMT2A	0.000751915
GO:0016568	chromatin modification	8.74E-12	ARID1A,ARID2,ATRX,BAP1,BPTF,BRCA1,CARM1,CHD8,CLOCK,EZH2,HDAC9,KDM5C,KDM6A,MEN1,NCOR1,PAX5,PBRM1,RB1,SIN3A,SOX9,TP53,KMT2A	1.06E-07
GO:0048513	organ development	1.03E-11	AKAP9,ARHGAP35,ARID1A,ATRX,BPTF,BRAF,BRCA1,CAD,CARM1,CHD8,CLOCK,CSDE1,CUL1,EGFR,EZH2,HDAC9,HSP90AB1,IDH1,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAX,MEN1,MET,NCOR1,NF1,NFATC4,NR2F2,PAX5,PBRM1,PCDH18,PIK3CA,PIK3R1,PTEN,PTPN11,RB1,SIN3A,SOS1,SOX9,TGFBR2,TP53,WT1,KMT2A	0.000800217

GO:0060255	regulation of macromolecule metabolic process	1.03E-11	ACAD8,AKAP9,ANK3,ARHGAP35,ARID1A,ARID2,ATRX,BPTF,BRAF,BRCA1,CARM1,CASP1,CHD8,CLOCK,CLTC,CNOT1,CSDE1,CUL1,DIS3,EGFR,EZH2,FN1,HDAC9,HSP90AB1,KDM5C,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAP4K3,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,NUP107,PAX5,PBRM1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TP53,WT1,ZNF814,KMT2A	0.002467951
GO:0044767	single-organism developmental process	1.46E-11	ADAM10,AKAP9,ANK3,ARHGAP35,ARID1A,ATRX,BPTF,BRAF,BRCA1,CAD,CARM1,CHD8,CLOCK,CLTC,CNOT1,CSDE1,CUL1,EGFR,EZH2,FN1,HDAC9,HSP90AB1,IDH1,KALRN,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PBRM1,PCDH18,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TJP1,TP53,TRIO,WT1,RPSA,KMT2A	**ND**
GO:0010033	response to organic substance	3.16E-11	ADAM10,AKAP9,ARHGEF6,ATRX,BPTF,BRAF,BRCA1,CAD,CARM1,CASP1,CUL1,EGFR,EZH2,FN1,HDAC9,HSP90AB1,IDH1,KALRN,KDR,KRAS,LRP6,MAP4K3,MAX,MEN1,NCOR1,NEDD4L,NF1,NR2F2,NUP107,PIK3CA,PIK3CB,PIK3R1,PRPF8,PTEN,PTPN11,SIN3A,SOS1,SOX9,SPTAN1,TGFBR2,TP53,TRIO,WT1	0.032600063

GO:0009653	anatomical structure morphogenesis	8.42E-11	ADAM10,AKAP9,ANK3,ARHGAP35,ARID1A,ATRX,BRAF,BRCA1,CARM1,CUL1,EGFR,EZH2,FAT1,FN1,HDAC9,HSP90AB1,KALRN,KDM6A,KDR,KRAS,LRP6,MET,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PIK3CA,PIK3CB,PTEN,PTPN11,RB1,SF3B1,SOS1,SOX9,SPTAN1,TGFBR2,TJP1,TP53,TRIO,WT1	0.00126693
GO:0031325	positive regulation of cellular metabolic process	8.42E-11	AKAP9,ARID1A,ATRX,BPTF,BRAF,BRCA1,CASP1,CHD8,CLOCK,CNOT1,CUL1,EGFR,EZH2,FN1,HSP90AB1,KDR,KRAS,LRP6,MAP3K4,MAP4K3,MEN1,MET,NCOR1,NF1,NFATC4,NR2F2,PAX5,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TP53,WT1,KMT2A	0.003630204
GO:0048608	reproductive structure development	8.42E-11	AKAP9,ARID1A,ATRX,BPTF,CSDE1,EGFR,HSP90AB1,IDH1,KDR,LRP6,MAP3K4,MEN1,MET,NR2F2,PBRM1,PTEN,PTPN11,SOX9,WT1	0.015434574
GO:0071363	cellular response to growth factor stimulus	8.42E-11	AKAP9,ARHGEF6,BPTF,BRAF,CAD,EGFR,FN1,KALRN,KDR,KRAS,MEN1,NCOR1,NEDD4L,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SOX9,SPTAN1,TGFBR2,TP53,TRIO	**ND**
GO:0061458	reproductive system development	8.63E-11	AKAP9,ARID1A,ATRX,BPTF,CSDE1,EGFR,HSP90AB1,IDH1,KDR,LRP6,MAP3K4,MEN1,MET,NR2F2,PBRM1,PTEN,PTPN11,SOX9,WT1	**ND**

GO:0043170	macromolecule metabolic process	9.74E-11	ACAD8,ADAM10,AKAP9,ANK3,AQR,ARFGEF2,ARHGAP35,ARHGEF6,ARID1A,ARID2,ATRX,BAP1,BPTF,BRAF,BRCA1,CAD,CARM1,CASP1,CHD8,CLOCK,CLTC,CNOT1,CSDE1,CUL1,DIS3,EGFR,EZH2,FN1,HDAC9,HSP90AB1,KALRN,KDM5C,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAP4K3,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,NUP107,PAX5,PBRM1,PIK3CA,PIK3CB,PIK3R1,PRPF8,PTEN,PTPN11,RB1,RPL5,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TP53,TRIO,WT1,ZNF814,RPSA,KMT2A	2.36E-05
GO:0007399	nervous system development	1.08E-10	ADAM10,AKAP9,ANK3,ARHGAP35,ARID1A,ATRX,BPTF,BRAF,BRCA1,CARM1,CHD8,EGFR,EZH2,FN1,HDAC9,HSP90AB1,KALRN,KDM6A,KRAS,LRP6,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PCDH18,PTEN,PTPN11,RB1,SOS1,SOX9,SPTAN1,TGFBR2,TP53,TRIO	**ND**
GO:0010467	gene expression	1.23E-10	ACAD8,ADAM10,ANK3,AQR,ARHGAP35,ARID1A,ARID2,ATRX,BPTF,BRAF,BRCA1,CARM1,CASP1,CHD8,CLOCK,CNOT1,CSDE1,DIS3,EGFR,EZH2,FN1,HDAC9,KDM5C,KDM6A,KRAS,LRP6,MAP3K4,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,NUP107,PAX5,PBRM1,PIK3CA,PIK3CB,PIK3R1,PRPF8,PTEN,RB1,RPL5,SF3B1,SIN3A,SOX9,STAG2,TGFBR2,TP53,WT1,ZNF814,RPSA,KMT2A	0.000703641
GO:0070848	response to growth factor	1.23E-10	AKAP9,ARHGEF6,BPTF,BRAF,CAD,EGFR,FN1,KALRN,KDR,KRAS,MEN1,NCOR1,NEDD4L,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SOX9,SPTAN1,TGFBR2,TP53,TRIO	**ND**

GO:0048518	positive regulation of biological process	1.23E-10	ADAM10,AKAP9,ANK3,ARFGEF2,ARHGAP35,ARHGEF6,ARID1A,ATRX,BAP1,BPTF,BRAF,BRCA1,CARM1,CASP1,CHD8,CLOCK,CNOT1,CUL1,DIS3,EGFR,EZH2,FN1,HDAC9,HSP90AB1,KALRN,KDM6A,KDR,KRAS,LRP6,MAP3K4,MAP4K3,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SF3B1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TP53,TRIO,WT1,KMT2A	3.11E-05
GO:0048522	positive regulation of cellular process	1.61E-10	ADAM10,AKAP9,ANK3,ARHGAP35,ARHGEF6,ARID1A,ATRX,BAP1,BPTF,BRAF,BRCA1,CARM1,CASP1,CHD8,CLOCK,CNOT1,CUL1,EGFR,EZH2,FN1,HDAC9,HSP90AB1,KALRN,KDR,KRAS,LRP6,MAP3K4,MAP4K3,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,PAX5,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,RB1,SIN3A,SOS1,SOX9,SPTAN1,STAG2,TGFBR2,TP53,TRIO,WT1,KMT2A	5.84E-06
GO:0051276	chromosome organization	1.76E-10	ARID1A,ARID2,ATRX,BAP1,BPTF,BRCA1,CARM1,CHD8,CLOCK,EZH2,HDAC9,KDM5C,KDM6A,MAP3K4,MEN1,NCOR1,NUP107,PAX5,PBRM1,PTEN,RB1,SIN3A,SOX9,STAG2,TP53,KMT2A	1.34E-06
GO:0006325	chromatin organization	1.88E-10	ARID1A,ARID2,ATRX,BAP1,BPTF,BRCA1,CARM1,CHD8,CLOCK,EZH2,HDAC9,KDM5C,KDM6A,MEN1,NCOR1,PAX5,PBRM1,RB1,SIN3A,SOX9,TP53,KMT2A	1.95E-06

GO:0010468	regulation of gene expression	1.89E-10	ACAD8,ANK3,ARHGAP35,ARID1A,ARID2,ATRX,BPTF,BRAF,BRCA1,CARM1,CHD8,CLOCK,CNOT1,CSDE1,DIS3,EGFR,EZH2,FN1,HDAC9,KDM5C,KDM6A,KRAS,LRP6,MAP3K4,MAX,MEN1,MET,NCOR1,NEDD4L,NF1,NFATC4,NR2F2,NUP107,PAX5,PBRM1,PIK3CA,PIK3CB,PIK3R1,PTEN,RB1,SF3B1,SIN3A,SOX9,STAG2,TGFBR2,TP53,WT1,ZNF814,KMT2A	0.004212933

Supplementary Table 5b | GBM-associated Reactome pathways of the 2016 analysis compared to 2010.
The table shows the top most significant Reactome terms in 2016 ranked by FDR adjusted p-values compared to FDR p-values detected in 2010.

React.ID	Description	FDR	Common genes	FDR of Reactome.ID in 2010
R-HSA-1266738	Developmental Biology	3.44E-06	ADAM10,AKAP9,ANK3,ARHGAP35,BRAF,CARM1,CLTC,EGFR,EZH2,FN1,HSP90AB1,KALRN,KDM6A,KDR,KRAS,MET,NCOR1,NF1,NR2F2,PTPN11,SOS1,SPTAN1,TRIO	ND
R-HSA-210993	Tie2 Signaling	3.44E-06	KRAS,PIK3CA,PIK3CB,PIK3R1,PTPN11,SOS1	1.26E-05
R-HSA-1236394	Signaling by ERBB4	3.54E-06	AKAP9,BRAF,CUL1,EGFR,FN1,KRAS,NCOR1,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-422475	Axon guidance	9.12E-06	ADAM10,AKAP9,ANK3,ARHGAP35,BRAF,CLTC,EGFR,FN1,HSP90AB1,KALRN,KDR,KRAS,MET,NF1,PTPN11,SOS1,SPTAN1,TRIO	ND
R-HSA-5637810	Constitutive Signaling by EGFRvIII	1.58E-05	EGFR,KRAS,PIK3CA,PIK3R1,SOS1	ND
R-HSA-5637812	Signaling by EGFRvIII in Cancer	1.58E-05	EGFR,KRAS,PIK3CA,PIK3R1,SOS1	ND
R-HSA-5663202	Diseases of signal transduction	1.67E-05	ADAM10,CUL1,EGFR,HDAC9,KRAS,LRP6,NCOR1,PIK3CA,PIK3CB,PIK3R1,PTPN11,SOS1,TGFBR2	ND
R-HSA-166520	Signalling by NGF	1.67E-05	AKAP9,ARHGEF6,BRAF,EGFR,FN1,KALRN,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1,TRIO	0.00410468
R-HSA-449147	Signaling by Interleukins	1.76E-05	AKAP9,BRAF,CASP1,CUL1,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTPN11,SOS1,SPTAN1	ND
R-HSA-177929	Signaling by EGFR	2.31E-05	ADAM10,AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	1.26E-05

R-HSA-1236382	Constitutive Signaling by Ligand-Responsive EGFR Cancer Variants	2.39E-05	EGFR,KRAS,PIK3CA,PIK3R1,SOS1	ND
R-HSA-1643685	Disease	2.39E-05	ADAM10,CUL1,EGFR,HDAC9,HSP90AB1,IDH1,KRAS,LRP6,NCOR1,NEDD4L,NUP107,PIK3CA,PIK3CB,PIK3R1,PTPN11,RPL5,SOS1,TGFBR2,RPSA	ND
R-HSA-2454202	Fc epsilon receptor (FCERI) signaling	2.39E-05	AKAP9,BRAF,CUL1,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-1643713	Signaling by EGFR in Cancer	2.39E-05	EGFR,KRAS,PIK3CA,PIK3R1,SOS1	ND
R-HSA-5637815	Signaling by Ligand-Responsive EGFR Variants in Cancer	2.39E-05	EGFR,KRAS,PIK3CA,PIK3R1,SOS1	ND
R-HSA-5654687	Downstream signaling of activated FGFR1	2.97E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-5654696	Downstream signaling of activated FGFR2	2.97E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-5654708	Downstream signaling of activated FGFR3	2.97E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-5654716	Downstream signaling of activated FGFR4	2.97E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-512988	Interleukin-3, 5 and GM-CSF signaling	2.97E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTPN11,SOS1,SPTAN1	ND
R-HSA-5654736	Signaling by FGFR1	2.97E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND

R-HSA-5654741	Signaling by FGFR3	2.97E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-5654743	Signaling by FGFR4	2.97E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-1433557	Signaling by SCF-KIT	2.97E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-4420097	VEGFA-VEGFR2 Pathway	2.97E-05	AKAP9,BRAF,EGFR,FN1,KDR,KRAS,NCF2,NF1,PIK3CA,PIK3CB,PIK3R1,SOS1,SPTAN1	ND
R-HSA-190236	Signaling by FGFR	3.07E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-5654738	Signaling by FGFR2	3.07E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-1227986	Signaling by ERBB2	3.18E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	ND
R-HSA-194138	Signaling by VEGF	3.18E-05	AKAP9,BRAF,EGFR,FN1,KDR,KRAS,NCF2,NF1,PIK3CA,PIK3CB,PIK3R1,SOS1,SPTAN1	ND
R-HSA-186763	Downstream signal transduction	3.30E-05	AKAP9,BRAF,EGFR,FN1,KRAS,NF1,PIK3CA,PIK3CB,PIK3R1,PTEN,PTPN11,SOS1,SPTAN1	0.000443176

Supplementary References

1 Creixell, P. *et al.* Pathway and network analysis of cancer genomes. *Nat Methods* **12**, 615-621 (2015).

2 Khatri, P., Sirota, M. & Butte, A. J. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Computational Biology* **8**, e1002375 (2012).

3 Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research* 37, 1-13 (2009).

4 Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics* **25**, 25-29 (2000).

5 Huntley, R. P. *et al.* The GOA database: gene Ontology annotation updates for 2015. *Nucleic Acids Research* **43**, D1057-1063 (2015).

6 Croft, D. *et al.* The Reactome pathway knowledgebase. *Nucleic Acids Research* **42**, D472-477 (2014).

7 Pruitt, K. D. *et al.* The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. *Genome Research* **19**, 1316-1323 (2009).

8 Marcotte, R. *et al.* Functional Genomic Landscape of Human Breast Cancer Drivers, Vulnerabilities, and Resistance. *Cell* **164**, 293-309 (2016).

9 Tamborero, D. *et al.* Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Scientific Reports* **3**, 2650 (2013).

10 Rubio-Perez, C. *et al.* In silico prescription of anticancer drugs to cohorts of 28 tumor types reveals targeting opportunities. *Cancer Cell* **27**, 382-396 (2015).

11 Gonzalez-Perez, A. *et al.* IntOGen-mutations identifies cancer drivers across tumor types. *Nat Methods* **10**, 1081-1082 (2013).

12 Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G. D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLOS ONE* **5**, e13984 (2010).

13 Cline, M. S. *et al.* Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols* **2**, 2366-2382 (2007)

14 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).

15 Li, A. *et al.* Circadian gene Clock contributes to cell proliferation and migration of glioma and is directly regulated by tumor-suppressive miR-124. *FEBS letters* **587**, 2455-2460 (2013).

16 D'Urso, P. I. *et al.* miR-155 is up-regulated in primary and secondary glioblastoma and promotes tumour growth by inhibiting GABA receptors. *International Journal of Oncology* **41**, 228-234 (2012).

17 Flavahan, W. A. *et al.* Brain tumor initiating cells adapt to restricted nutrition through preferential glucose uptake. *Nature Neuroscience* **16**, 1373-1382 (2013).

18 Takebe, N. *et al.* Targeting Notch, Hedgehog, and Wnt pathways in cancer stem cells: clinical update. *Nature Reviews Clinical Oncology* **12**, 445-464 (2015).

19 Seoane, J., Le, H. V., Shen, L., Anderson, S. A. & Massague, J. Integration of Smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation. *Cell* **117**, 211-223 (2004).

20 Brennan, C. W. *et al.* The somatic genomic landscape of glioblastoma. *Cell* **155**, 462-477 (2013).

21 Neuzillet, C. *et al.* Targeting the TGFbeta pathway for cancer therapy. *Pharmacology & Therapeutics* **147**, 22-31 (2015).

22 Huang, P. H. *et al.* Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 12867-12872 (2007).
23 Inda, M. M. *et al.* Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma. *Genes & Development* **24**, 1731-1745 (2010).
24 Shinojima, N. *et al.* Prognostic value of epidermal growth factor receptor in patients with glioblastoma multiforme. *Cancer Research* **63**, 6962-6970 (2003).
25 Padfield, E., Ellis, H. P. & Kurian, K. M. Current Therapeutic Advances Targeting EGFR and EGFRvIII in Glioblastoma. *Frontiers in Oncology* **5**, 5 (2015).
26 Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).
27 Reimand, J., Kull, M., Peterson, H., Hansen, J. & Vilo, J. g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Research* **35**, W193-200 (2007).
28 Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Research* **37**, W305-311 (2009).
29 Mi, H., Muruganujan, A. & Thomas, P. D. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research* **41**, D377-386 (2013).
30 Lyne, R. *et al.* Cross-organism analysis using InterMine. *Genesis* **53**, 547-560 (2015).
31 Zheng, Q. & Wang, X. J. GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. *Nucleic Acids Research* **36**, W358-363 (2008).
32 Castillo-Davis, C. I. & Hartl, D. L. GeneMerge--post-genomic analysis, data mining, and hypothesis testing. *Bioinformatics* **19**, 891-892 (2003).
33 Kamburov, A., Wierling, C., Lehrach, H. & Herwig, R. ConsensusPathDB--a database for integrating human functional interaction networks. *Nucleic acids research* **37**, D623-628 (2009).
34 McLean, C. Y. *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology* **28**, 495-501 (2010).
35 Medina, I. *et al.* Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. *Nucleic Acids Research* **38**, W210-213 (2010).
36 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research* (2016).
37 Berriz, G. F., King, O. D., Bryant, B., Sander, C. & Roth, F. P. Characterizing gene sets with FuncAssociate. *Bioinformatics* **19**, 2502-2504 (2003).
38 Kang, H. *et al.* gsGator: an integrated web platform for cross-species gene set analysis. *BMC bioinformatics* **15**, 13 (2014).
39 Zhang, B., Kirov, S. & Snoddy, J. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Research* **33**, W741-748 (2005).
40 Carmona-Saez, P., Chagoyen, M., Tirado, F., Carazo, J. M. & Pascual-Montano, A. GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. *Genome Biology* **8**, R3 (2007).
41 Zeeberg, B. R. *et al.* GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biology* **4**, R28 (2003).

42 Backes, C. *et al.* GeneTrail--advanced gene set enrichment analysis. *Nucleic Acids Research* **35**, W186-192 (2007).
43 Zhou, X. & Su, Z. EasyGO: Gene Ontology-based annotation and functional enrichment analysis tool for agronomical species. *BMC Genomics* **8**, 246 (2007).
44 Rho, K. *et al.* GARNET--gene set analysis with exploration of annotation relations. *BMC Bioinformatics* **12 Suppl 1**, S25 (2011).
45 Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**, 44-57 (2009).
46 Sartor, M. A. *et al.* ConceptGen: a gene set enrichment and gene set relation mapping tool. *Bioinformatics* **26**, 456-463 (2010).
47 Martin, D. *et al.* GOToolBox: functional analysis of gene datasets based on Gene Ontology. *Genome Biology* **5**, R101 (2004).
48 Newman, J. C. & Weiner, A. M. L2L: a simple tool for discovering the hidden significance in microarray expression data. *Genome Biology* **6**, R81 (2005).
49 Zhong, S. *et al.* GoSurfer: a graphical interactive tool for comparative analysis of large gene sets in Gene Ontology space. *Applied Bioinformatics* **3**, 261-264 (2004).
50 Beissbarth, T. & Speed, T. P. GOstat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* **20**, 1464-1465 (2004).