doi:10.1038/nature12477

Supplementary Note 1. Acknowledgements.

This work was supported by the Wellcome Trust (grant reference 098051); Breakthrough Breast Cancer Research (ICGC 08/09); National Health and Medical Research Council of Australia (NHMRC; 631701, 535903, 1017028); Australian Cancer Research Foundation (ACRF); Queensland Government (NIRAP); The University of Queensland; Brisbane Breast Bank, Cancer Council NSW (SRP06-01; ICGC09-01; SRP11-01); Cancer Institute NSW (06/ECF/1-24, 09/CDF/2-40, 07/CDF/1-03, 10/CRF/1-01, 08/RSA/1-15, 07/CDF/1-28, 10/CDF/2-26,10/FRL/2-03, 06/RSA/1-05, 09/RIG/1-02, 10/TPG/1-04, 11/REG/1-10, 11/CDF/3-26); Garvan Institute of Medical Research; Avner Nahmani Pancreatic Cancer Research Foundation; R.T. Hall Trust; Petre Foundation; Jane Hemstritch in memory of Philip Hemstritch; Spanish Ministry of Economy and Competitiveness through the Instituto de Salud Carlos III (ISCIII) and Red Temática de Investigación del Cáncer (RTICC) del ISCIII (RD12/0036/0036); INCa with the ICGC project; the Ligue Nationale Contre le Cancer (Cartes d'Identité des Tumeurs program); the PAIR-CHC project NoFLIC (funded by INCa and the Association pour la Recherche Contre le Cancer, ARC); the Réseau français des biobanques des tumeurs hépatiques; HEPTROMIC (Framework Programme 7, FP7); BioIntelligence (OSEO); Program for Promotion of Fundamental Studies in Health Sciences from the National Institute of Biomedical Innovation (NIBIO); National Cancer Center Research and Development Funds (23-A-8); Dutch Cancer Society (NKI 2012-5463); Dana-Farber Cancer Institute Women's Executive Council; The Icelandic Centre for Research, RANNIS; The University of Iceland Research Fund; INCa, Synergie-Lyon-Cancer and Centre Léon Bérard; Federal Ministry of Education and Research in Germany (BMBF) within the Program for Medical Genome Research

(01KU1002A to 01KU1002J, ICGC MMML-Seq); PedBrain Tumor Project; German Cancer Aid (109252); German Federal Ministry of Education and Research (BMBF PedBrain #01KU1201A, NGFNplus #01GS0883); NIH grants PO1-155258, RO1-124929 and P50-100007; Cancer Research UK, the Experimental Cancer Medicine Centre and Breakthrough Breast Cancer Unit at Kings College London and by the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's & St Thomas' NHS Foundation Trust and King's College London via the Breast Tissue and Data bank. BASIS project is a European research project funded by the European Community's Seventh Framework Programme (FP7/2010-2014) under the grant agreement number 242006. Individual support: S.B. and L.R.Y. (Wellcome Trust Clinical Research Training Fellow); S.N.Z. (Wellcome Trust Intermediate Clinical Fellow grant reference 100183/Z/12/Z); P.J.C. (Wellcome Trust Senior Clinical Research Fellow grant reference WT088340MA); C.L.-O. is an Investigator of the Botin Foundation; C.D. is supported by the Brussels Region- Impulse Programme Life Sciences; N.B. was supported by a starter grant from the Academy of Medical Sciences. We would like to acknowledge Core Sequencing Facility and IT groups of the Wellcome Trust Sanger Institute, the Breakthrough Breast Cancer Unit, OSBREAC (Oslo breast Cancer consortium, http://ous-research.no/kgjebsen/), the MEDIC Foundation and "Les Amis de l'Institut Bordet", Pathology Queensland; Consortiums members and collaborators: Alexander Claviez, Andreas Rosenwald, Andreas Rosenwald, Arndt Borkhardt, Benedikt Brors, Bernhard Radlwimmer, Chris Lawerenz, Cristina Lopez, David Langenberger, Dennis Karsch, Dido Lenze, Dieter Kube, Ellen Leich, Gesine Richter, Jan Korbel, Jessica Hoell, Jürgen Eils, Kebriah Hezaveh, Lorenz Trümper, Maciej Rosolowski, Marc Weniger, Marius Rohde, Markus Kreuz, Markus

Loeffler, Markus Schilhabel, Martin Dreyling, Martin-Leo Hansmann, Michael Hummel, Monika Szczepanowski, Ole Ammerpohl, Peter F. Stadler, Peter Möller, Ralf Küppers, Siegfried Haas, Sonja Eberth, Stefan Schreiber, Stephan H. Bernhart, Steve Hoffmann, Sylwester Radomski, Ulrike Kostezka, Wolfram Klapper, Christos Sotiriou, Denis Larsimont, Delphine Vincent, Marion Maetens, Odette Mariani, Anieta M. Sieuwerts, John W.M. Martens, Prof Jon G Jonasson, Isabelle Treilleux, Emilie Thomas, Gaëtan Mac Grogan, Cécile Mannina, Laurent Arnould, Laura Burillier, Jean-Louis Merlin, Magali Lefebvre, Frédéric Bibeau, Blandine Massemin, Frédérique Penault-Llorca, Qian Lopez, Marie-Christine Mathieu, INCa-ICGC Breast Cancer Consortium, Per Eystein Lonning, Margrete Schlooz-Vries, Jolien Tol, Hanneke van Laarhoven, Fred Sweep, Peter Bult; The Cancer Genome Atlas (TCGA), and the International Cancer Genome Consortium (ICGC). We would like to thank Gad Getz (and colleagues) for giving us the idea for the easy to read format of Figure 1.

Supplementary Tables 1 and 2

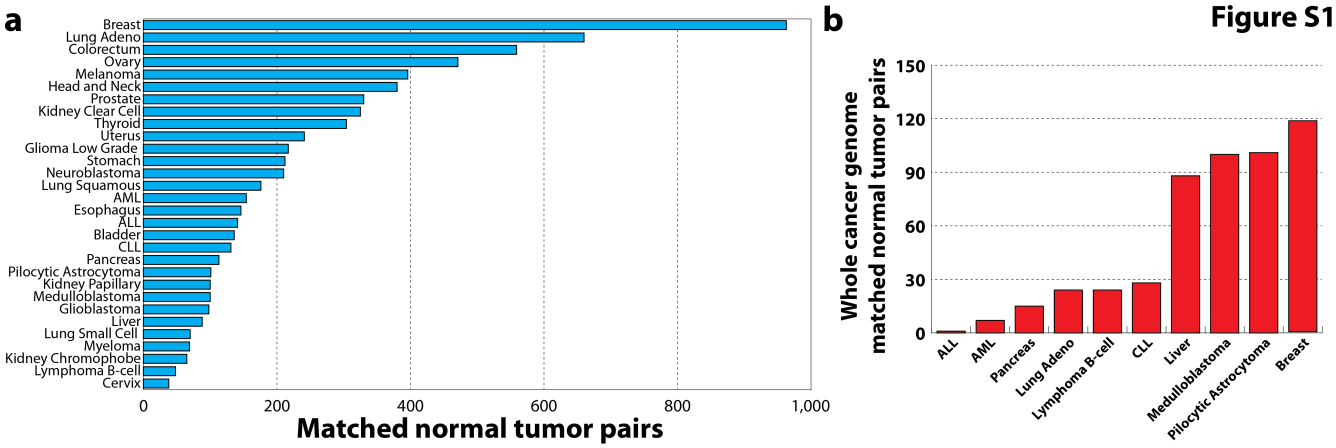
Mutational Signature	Validation Status	Total Mutations in Sample	Total Mutations by Signature	Examined Mutations	Validated Mutations
Signature 1A	V_A	48	40	48	48 (100%)
Signature 1B	V_A	58	55	58	56 (97%)
Signature 2	V_A	76	75	76	72 (95%)
Signature 3	V_A	70	65	70	69 (99%)
Signature 4	V_A	196	192	196	182 (95%)
Signature 5	V_C	332	286	91	75 (82%)
Signature 6	V_A	598	440	598	540 (90%)
Signature 7	V_A	470	432	470	412 (88%)
Signature 8	V_C	290	128	220	192 (87%)
Signature 9	V_B	4,423	2,811	4,423	3,977 (90%)
Signature 10	V_A	12,848	10,558	12,848	9,420 (74%)
Signature 11	V_A	102	100	102	67 (66%)
Signature 12	V_C	2,808	2327	100	93(93%)
Signature 13	V_C	19,797	3,809	100	90(90%)
Signature 14	V_C	12,984	12,984	100	86 (86%)
Signature 15	V_A	784	784	31	30 (97%)
Signature 16	V_A	793	678	73	69 (95%)
Signature 17	V_B	2,627	1,959	2,627	2,476 (94%)
Signature 18	V_A	158	156	158	142 (90%)
Signature 19	V_C	769	769	103	102 (99%)
Signature 20	V_A	885	488	198	198 (100%)
Signature 21	V_C	6,790	4,368	121	103(85%)
Signature U1	N/A	N/A	N/A	N/A	N/A
Signature U2	N/A	N/A	N/A	N/A	N/A
Signature R1	F_C	11,869	7,955	100	2(2%)
Signature R2	F_C	738	738	50	1(2%)
Signature R3	F_C	385	235	83	3(4%)

Supplementary Table 1. Validating mutational signatures. Please refer to online methods for more information about our strategy for validation of mutational signatures. Validation statuses are described in Supplementary Figure 99.

Cancer Type	Samples with Age information	Mutational Signature	P-value (FDR corrected)
ALL	106	Signature 1B	2.13E-04
AML	151	Signature 1B	6.81E-06
Breast	879	Signature 1B	7.23E-04
Colorectum	488	Signature 1B	2.89E-02
Glioma Low Grade	154	Signature 1A	1.50E-07
Head and Neck	299	Signature 1B	4.54E-03
Kidney Chromophobe	21	Signature 1A	3.53E-02
Kidney Clear Cell	294	Signature 1B	7.34E-12
Kidney Papillary	95	Signature 4	3.10E-03
Lymphoma B-cell	24	Signature 1B	1.06E-02
Medulloblastoma	100	Signature 1A	2.83E-10
Melanoma	216	Signature 1B	1.33E-03
Melanoma	216	Signature 7	2.00E-03
Neuroblastoma	192	Signature 1B	2.84E-05
Ovary	425	Signature 1B	7.18E-09
Pilocytic Astrocytoma	63	Signature 1B	4.76E-02
Stomach	148	Signature 1A	3.43E-02
Thyroid	157	Signature 4	2.95E-03

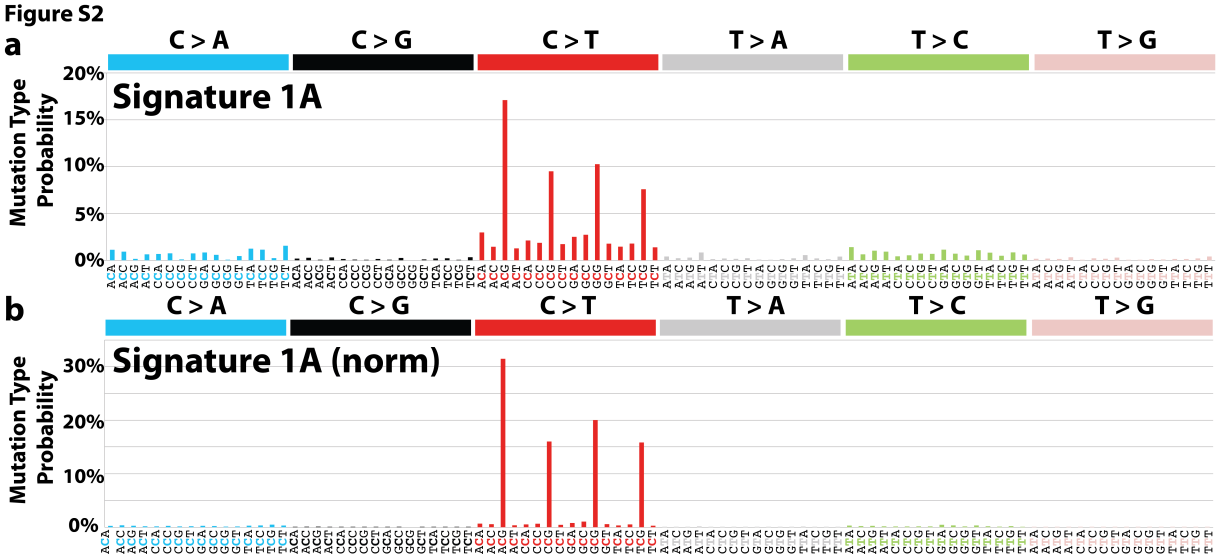
Supplementary Tables 2. Correlation between exposures of validated mutational signatures and age of cancer diagnosis. Please refer to online methods for our approach for associating age of diagnosis and exposures of validated mutational signatures.

Supplementary Figures 1 to 98

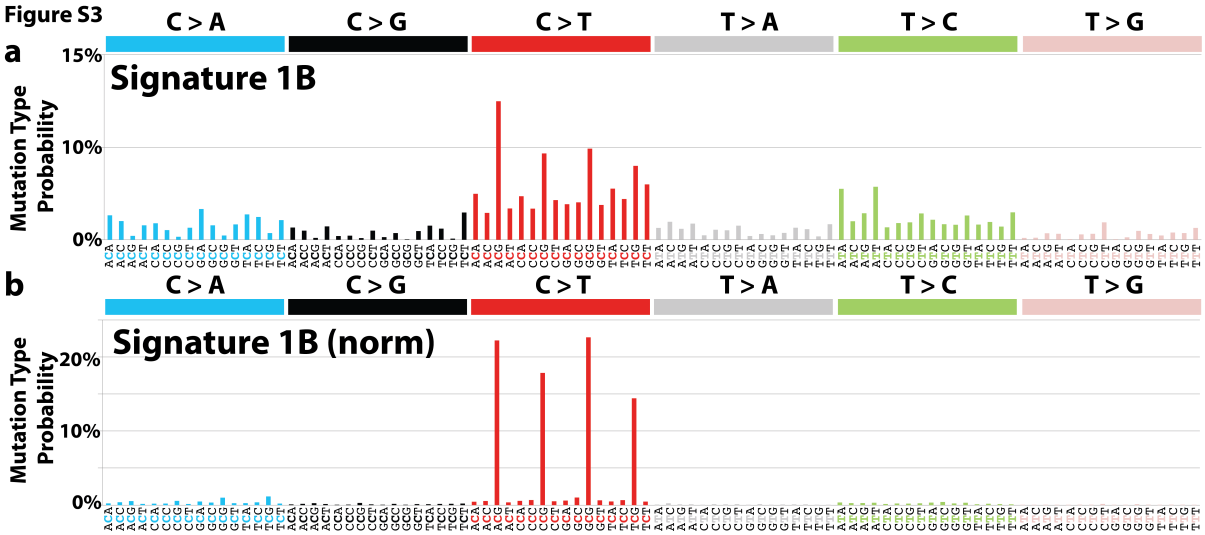


Supplementary Figure 1. Samples used for deciphering signatures of mutational processes. Mutational catalogues of ***(a)*** 7,042 primary cancers derived from 30 different cancer types were examined for mutational signatures, including ***(b)*** 507 whole cancer genome matched normal pairs.

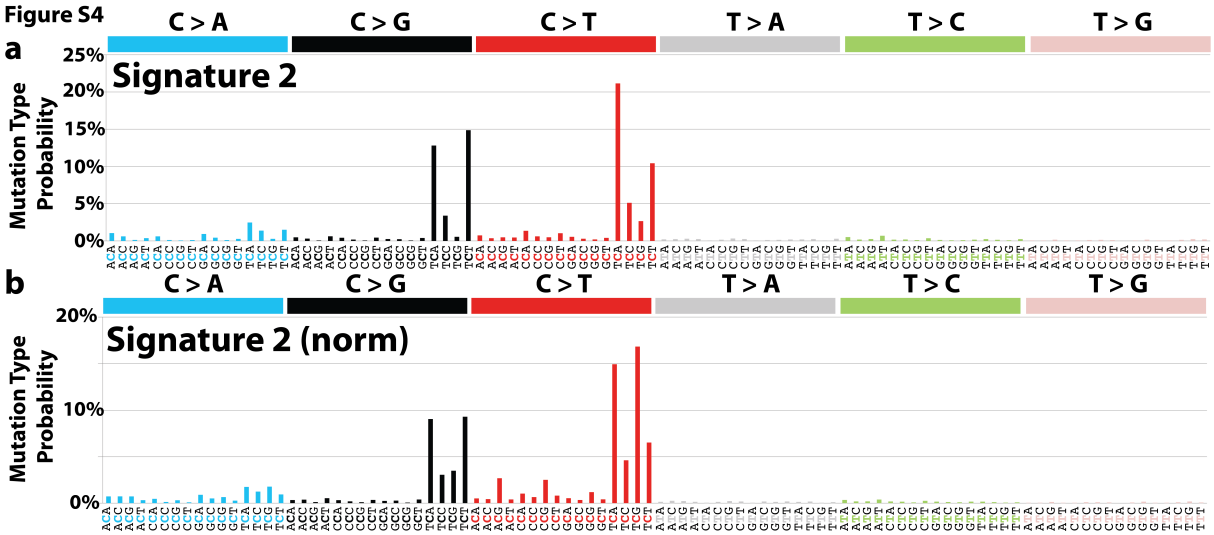
Figure S2



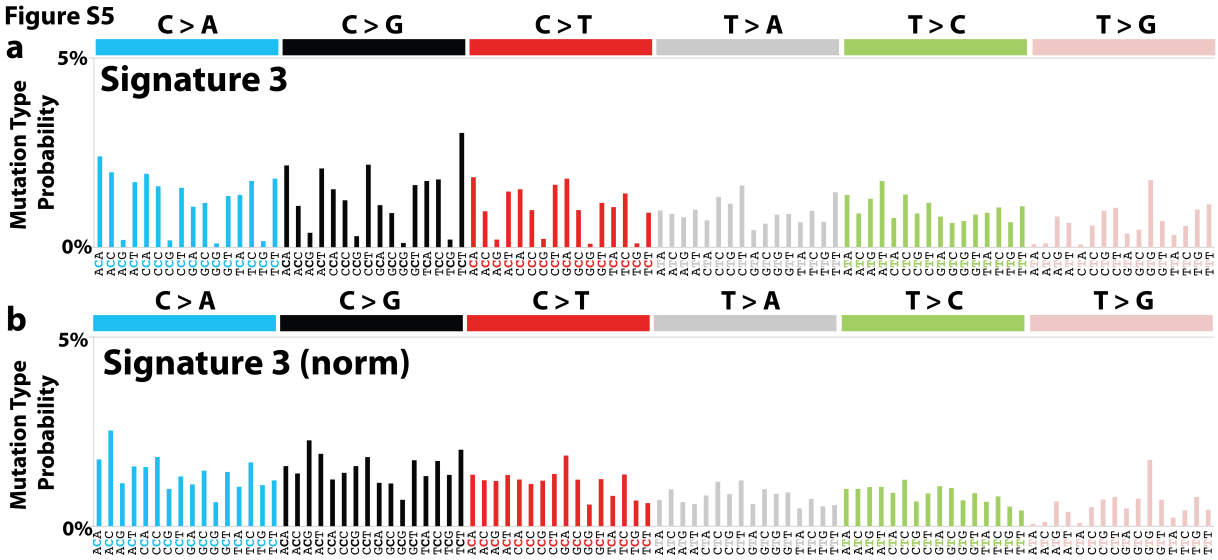
Supplementary Figure 2. Patterns of substitutions for Signature 1A. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



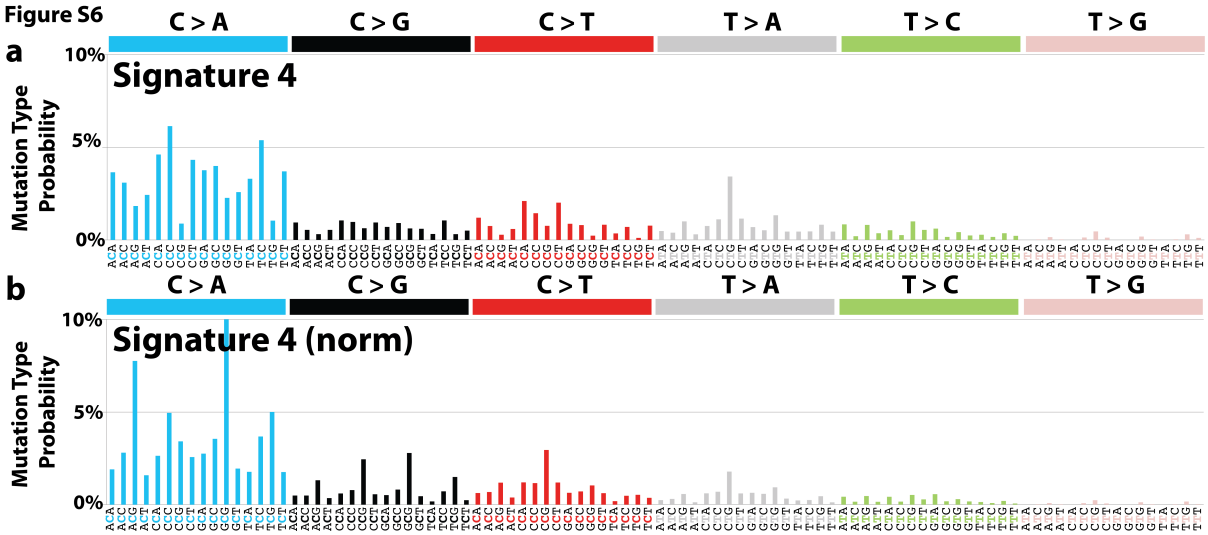
Supplementary Figure 3. Patterns of substitutions for Signature 1B. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



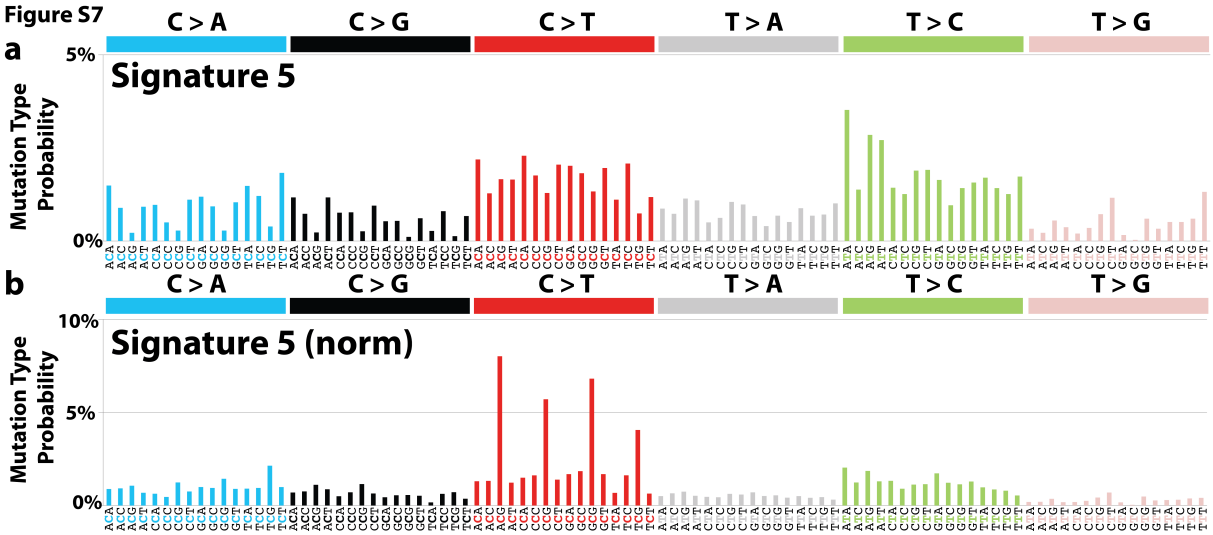
Supplementary Figure 4. Patterns of substitutions for Signature 2. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



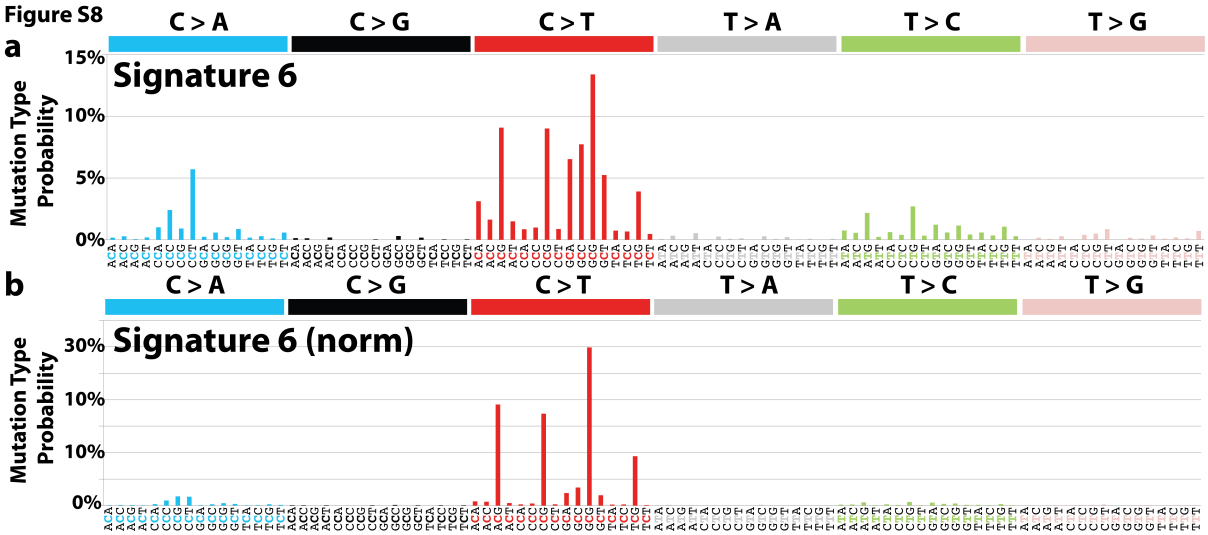
Supplementary Figure 5. Patterns of substitutions for Signature 3. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



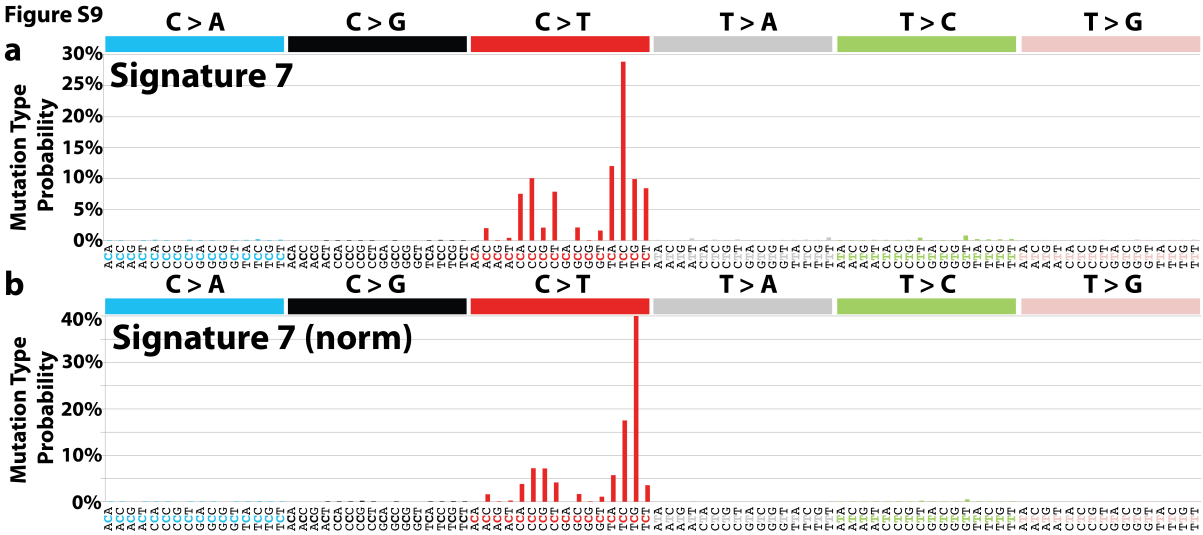
Supplementary Figure 6. Patterns of substitutions for Signature 4. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



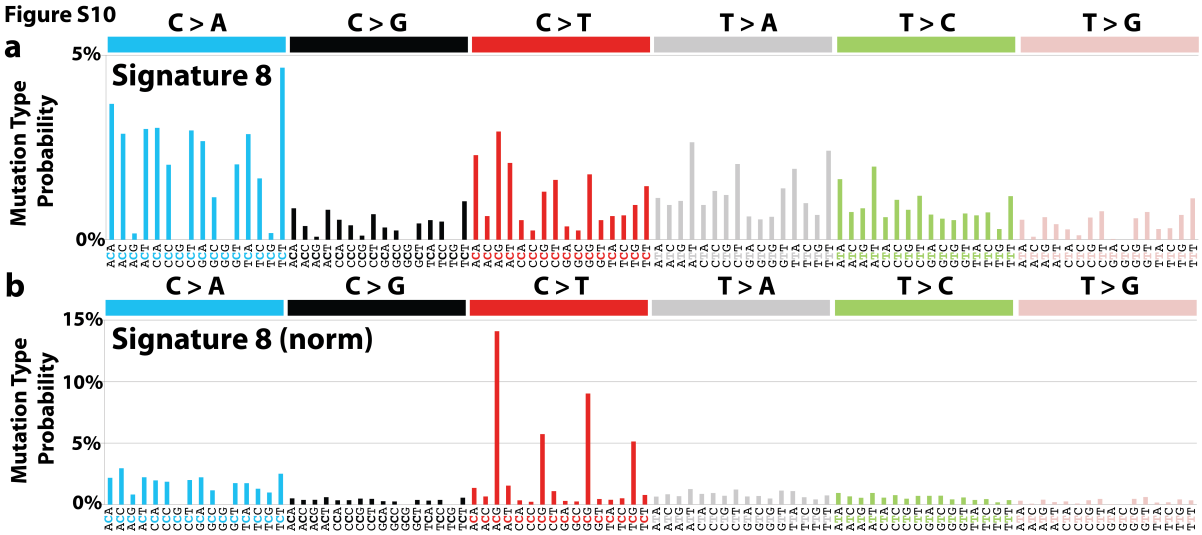
Supplementary Figure 7. Patterns of substitutions for Signature 5. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



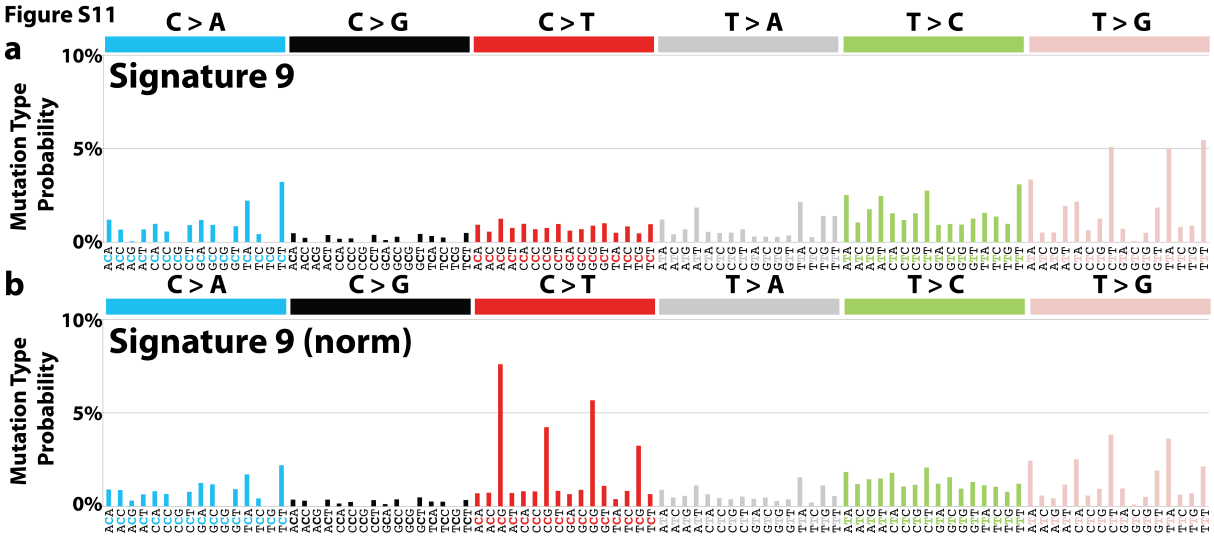
Supplementary Figure 8. Patterns of substitutions for Signature 6. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



Supplementary Figure 9. Patterns of substitutions for Signature 7. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.

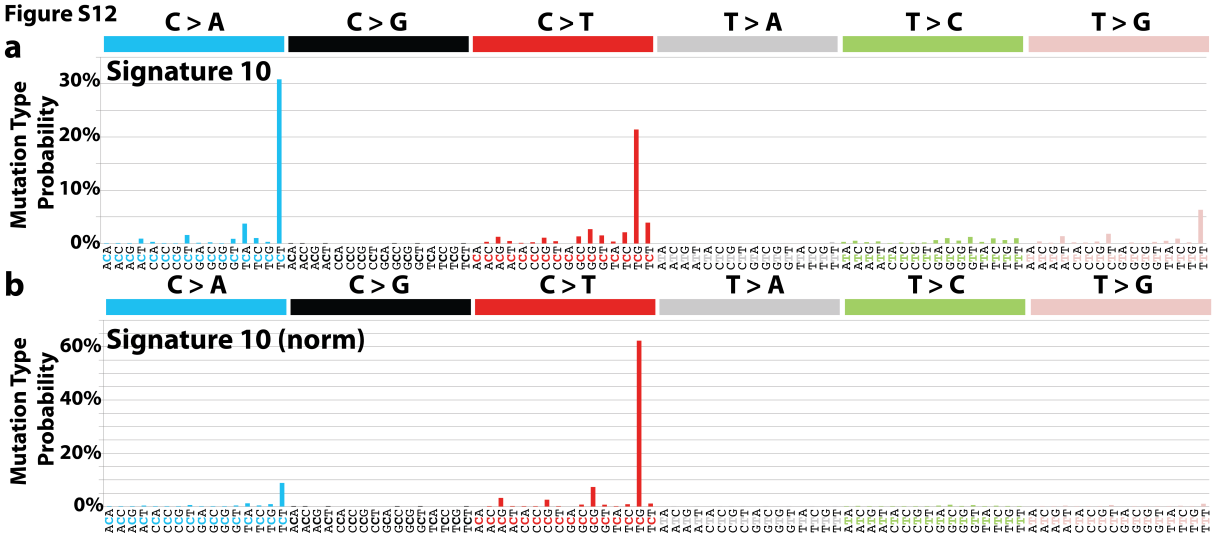


Supplementary Figure 10. Patterns of substitutions for Signature 8. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.

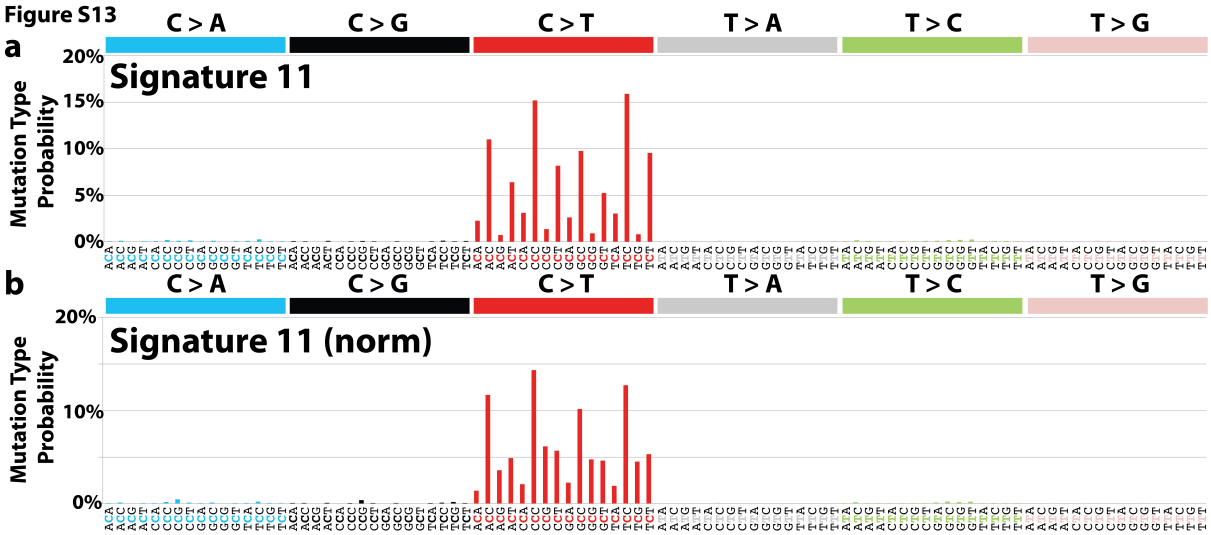


Supplementary Figure 11. Patterns of substitutions for Signature 9. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.

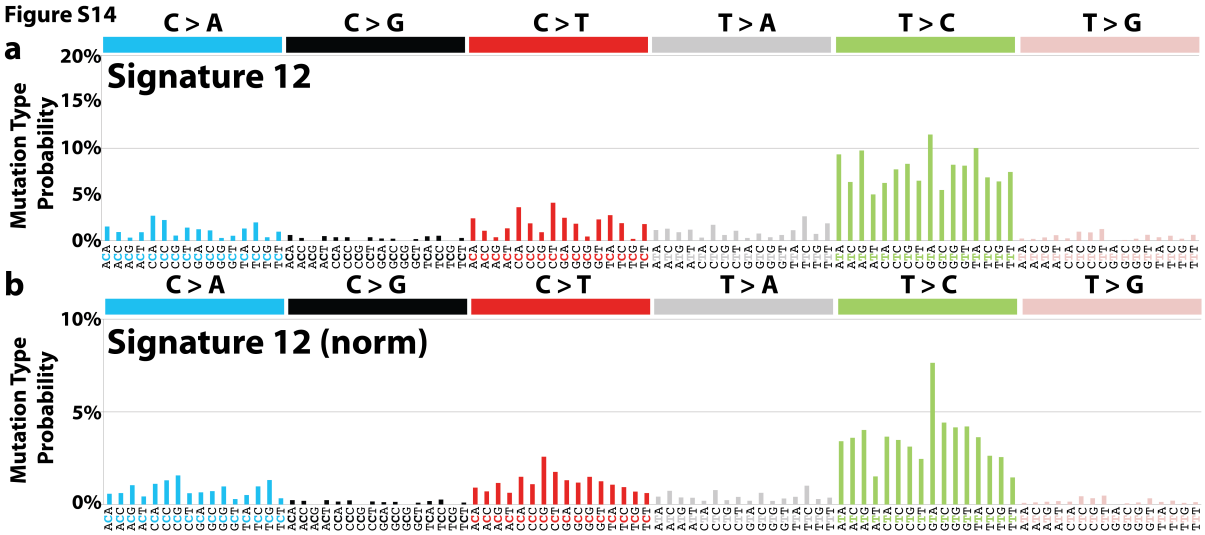
Figure S12



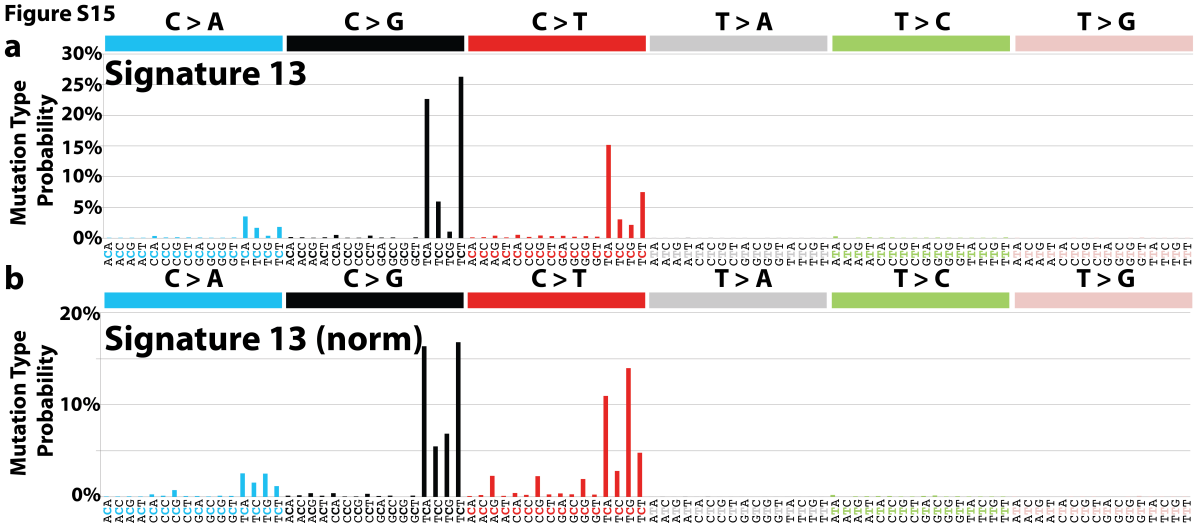
Supplementary Figure 12. Patterns of substitutions for Signature 10. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



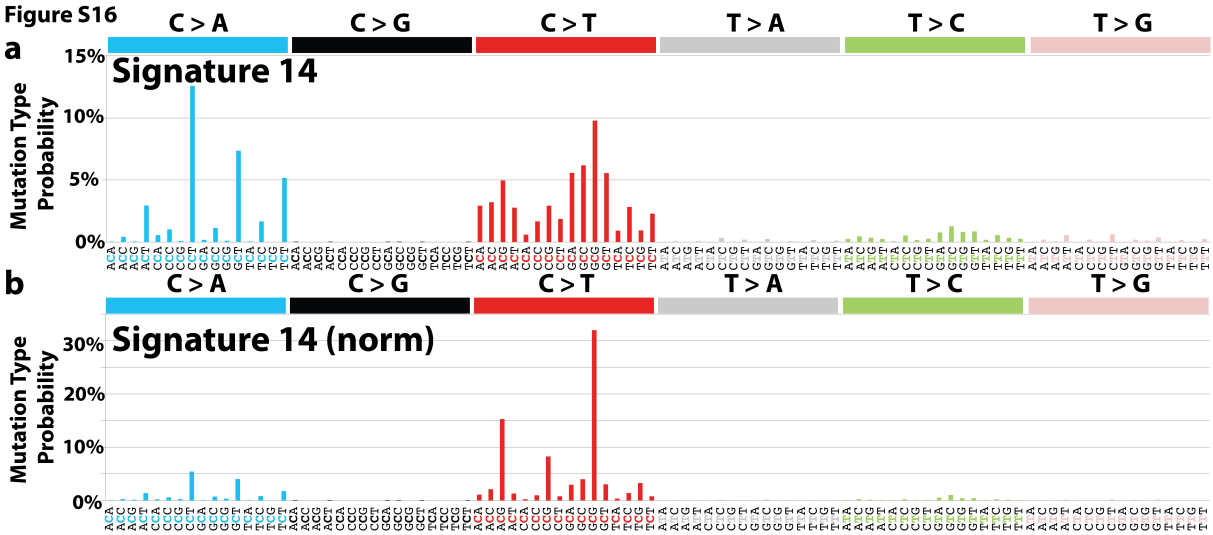
Supplementary Figure 13. Patterns of substitutions for Signature 11. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



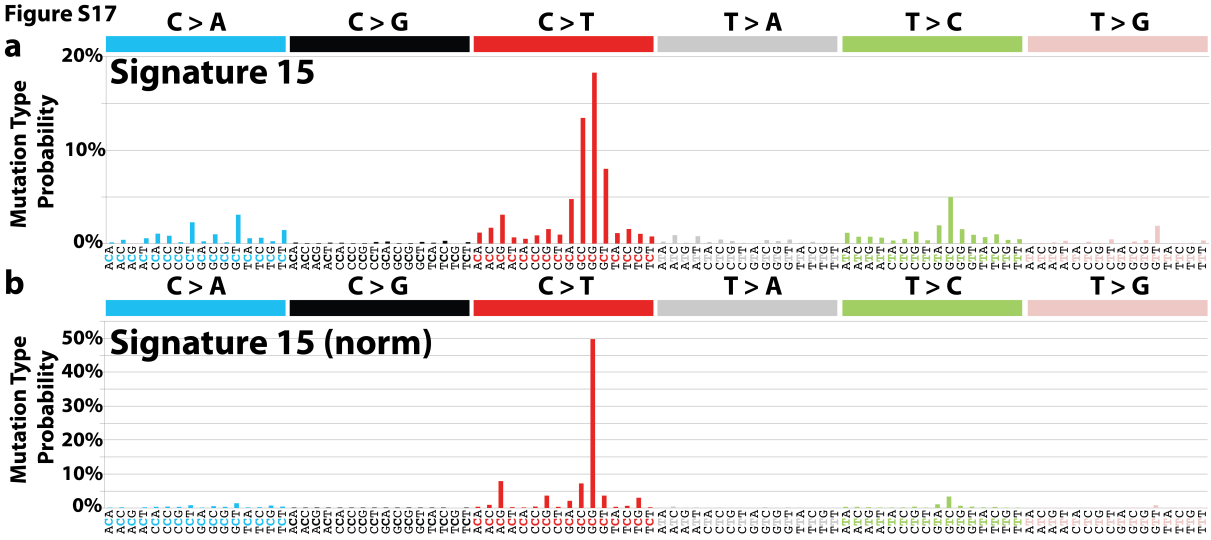
Supplementary Figure 14. Patterns of substitutions for Signature 12. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



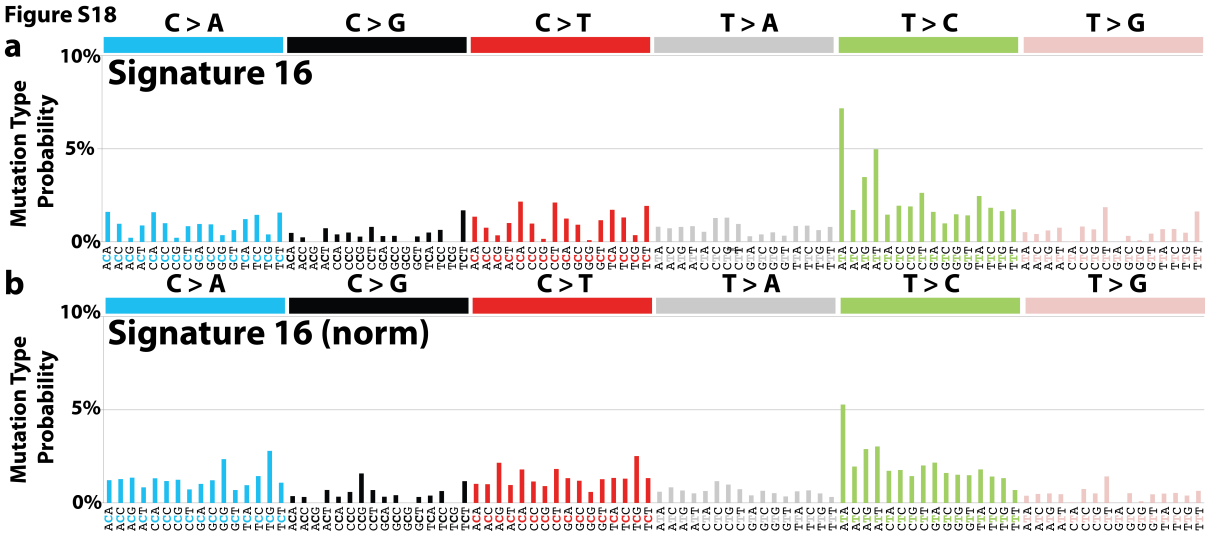
Supplementary Figure 15. Patterns of substitutions for Signature 13. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



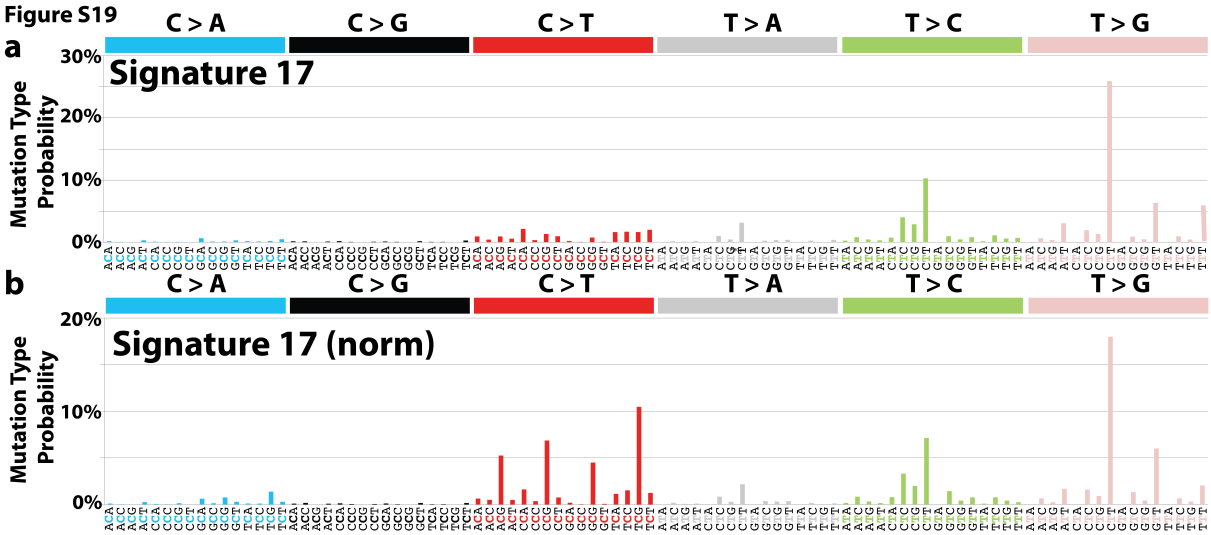
Supplementary Figure 16. Patterns of substitutions for Signature 14. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



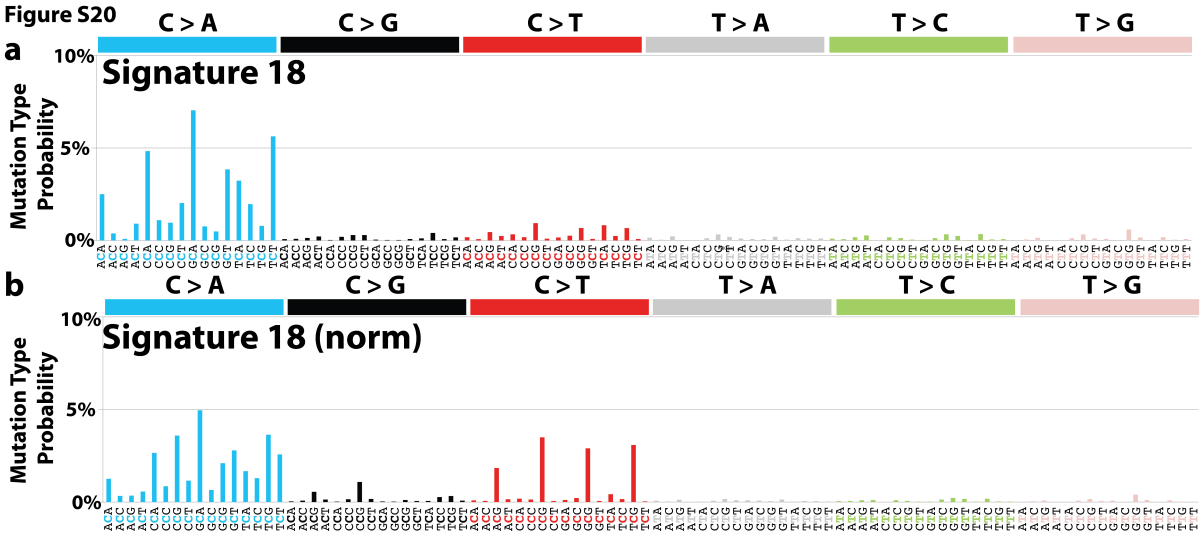
Supplementary Figure 17. Patterns of substitutions for Signature 15. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



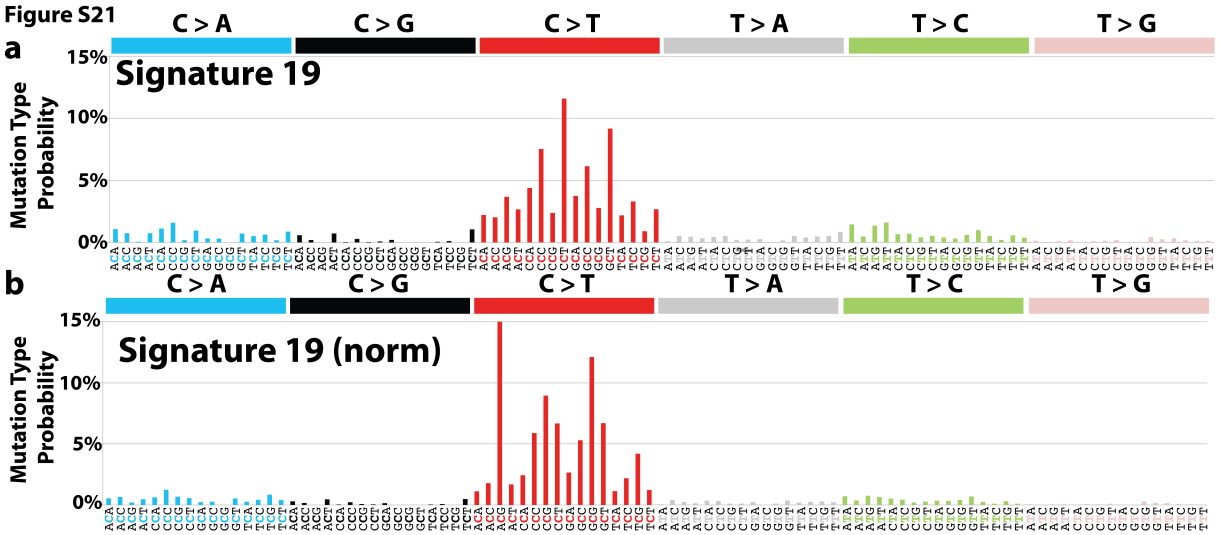
Supplementary Figure 18. Patterns of substitutions for Signature 16. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



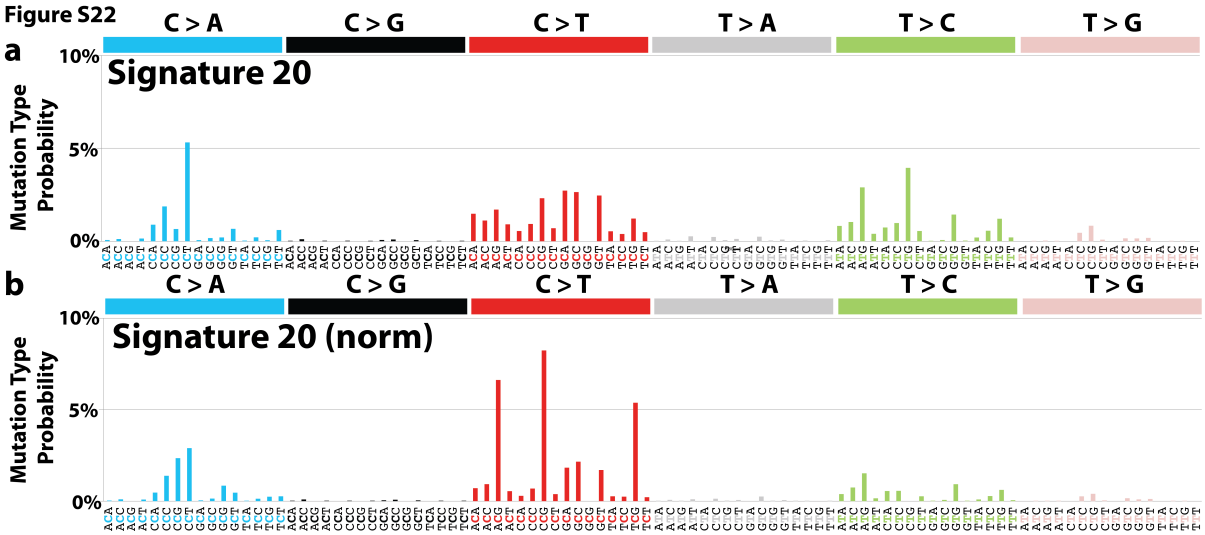
Supplementary Figure 19. Patterns of substitutions for Signature 17. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



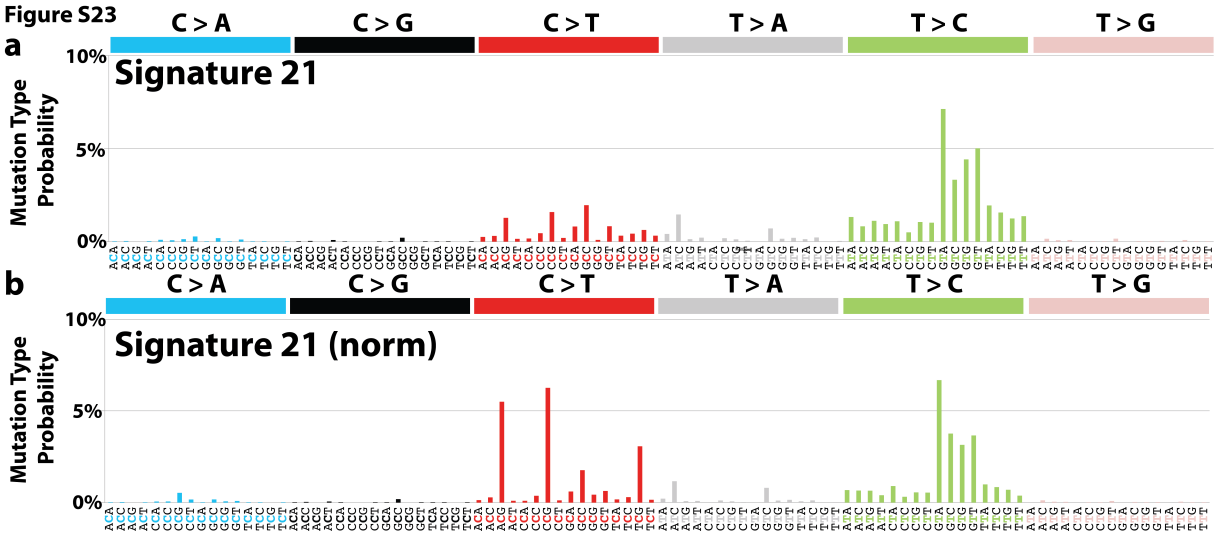
Supplementary Figure 20. Patterns of substitutions for Signature 18. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



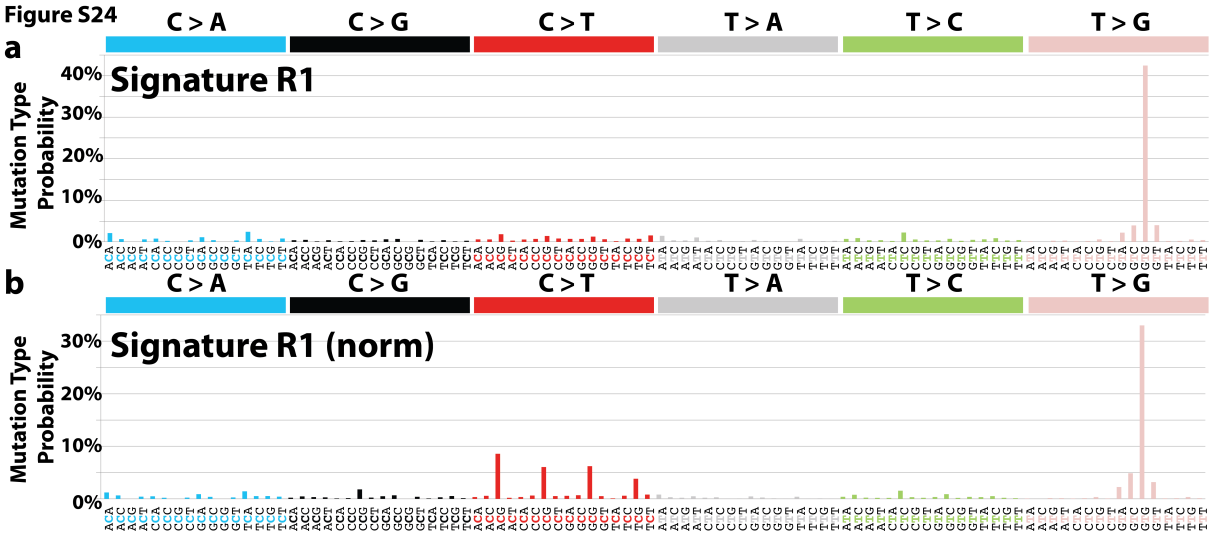
Supplementary Figure 21. Patterns of substitutions for Signature 19. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



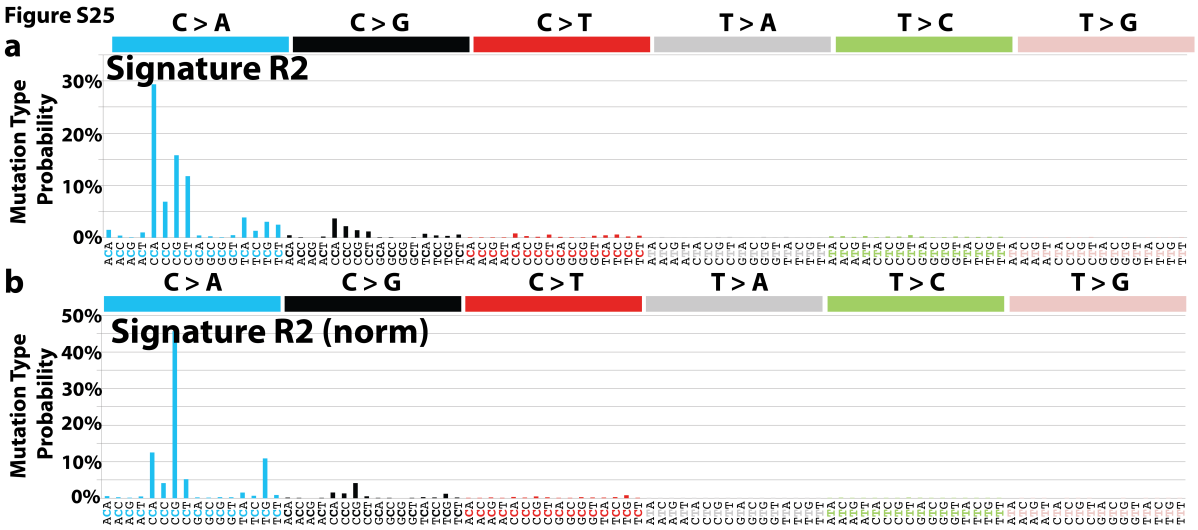
Supplementary Figure 22. Patterns of substitutions for Signature 20. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



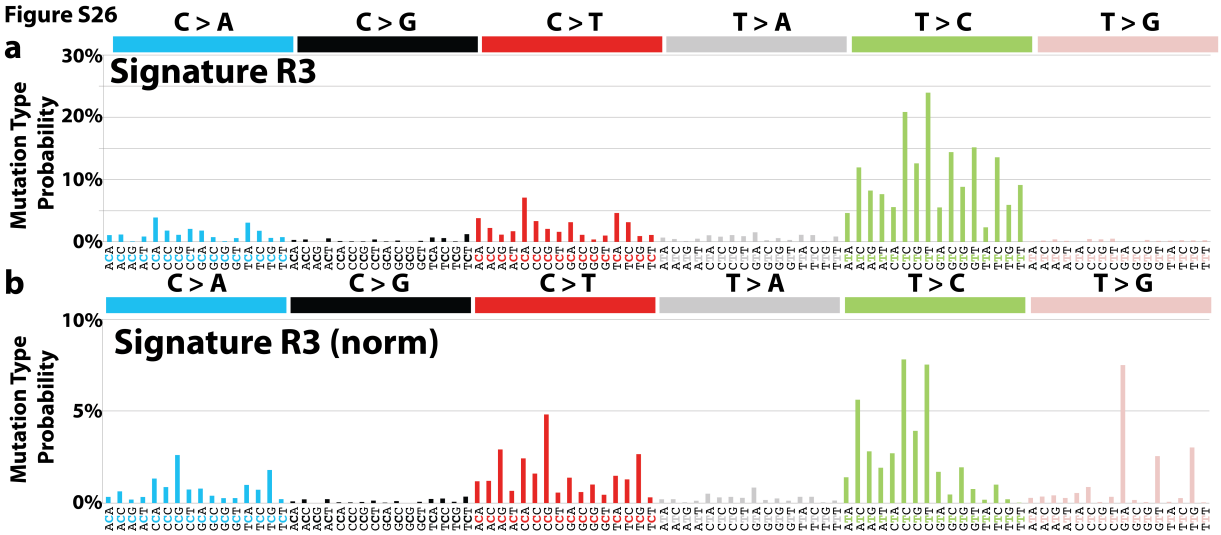
Supplementary Figure 23. Patterns of substitutions for Signature 21. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



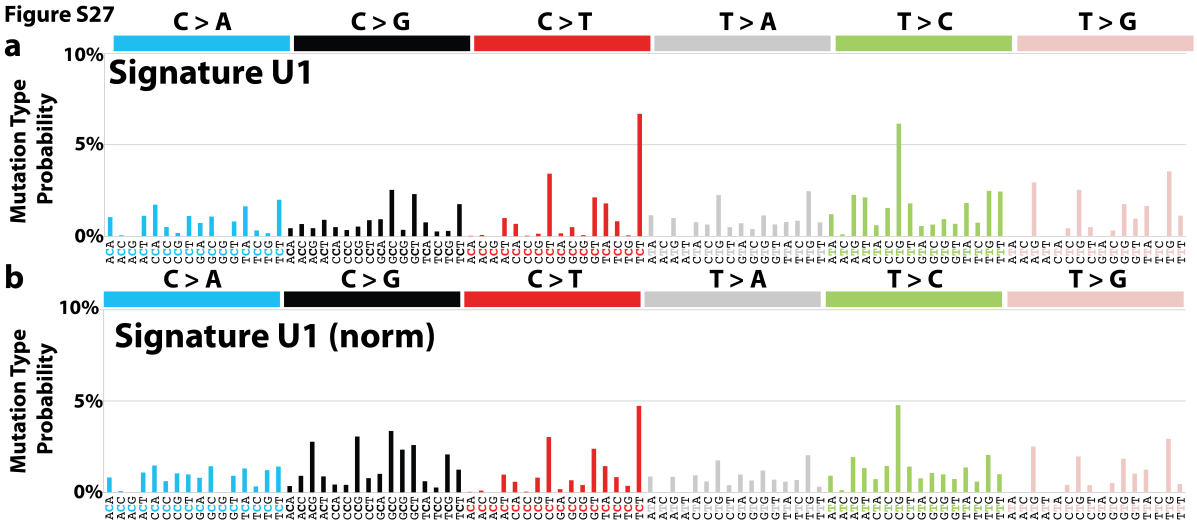
Supplementary Figure 24. Patterns of substitutions for Signature R1. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



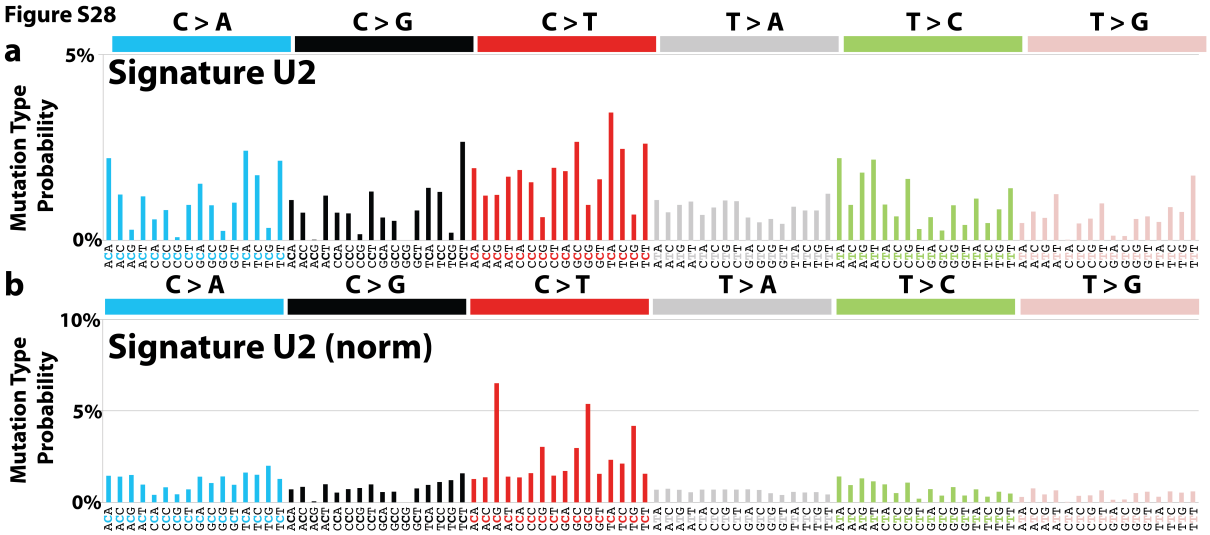
Supplementary Figure 25. Patterns of substitutions for Signature R2. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



Supplementary Figure 26. Patterns of substitutions for Signature R3. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.

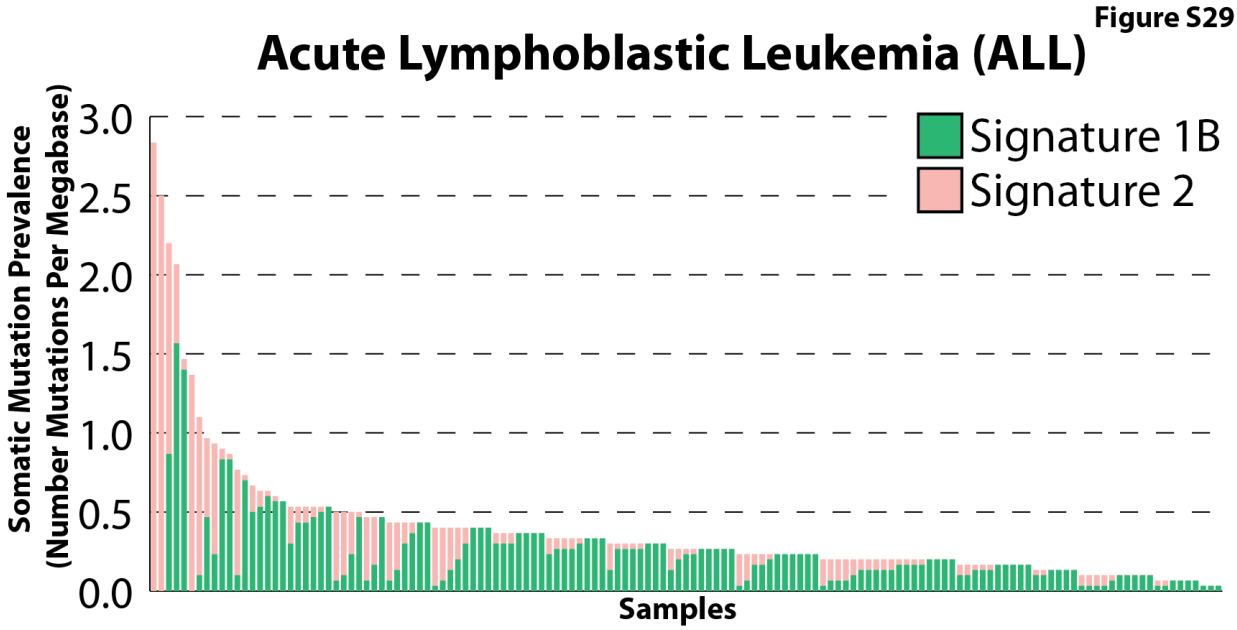


Supplementary Figure 27. Patterns of substitutions for Signature U1. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.



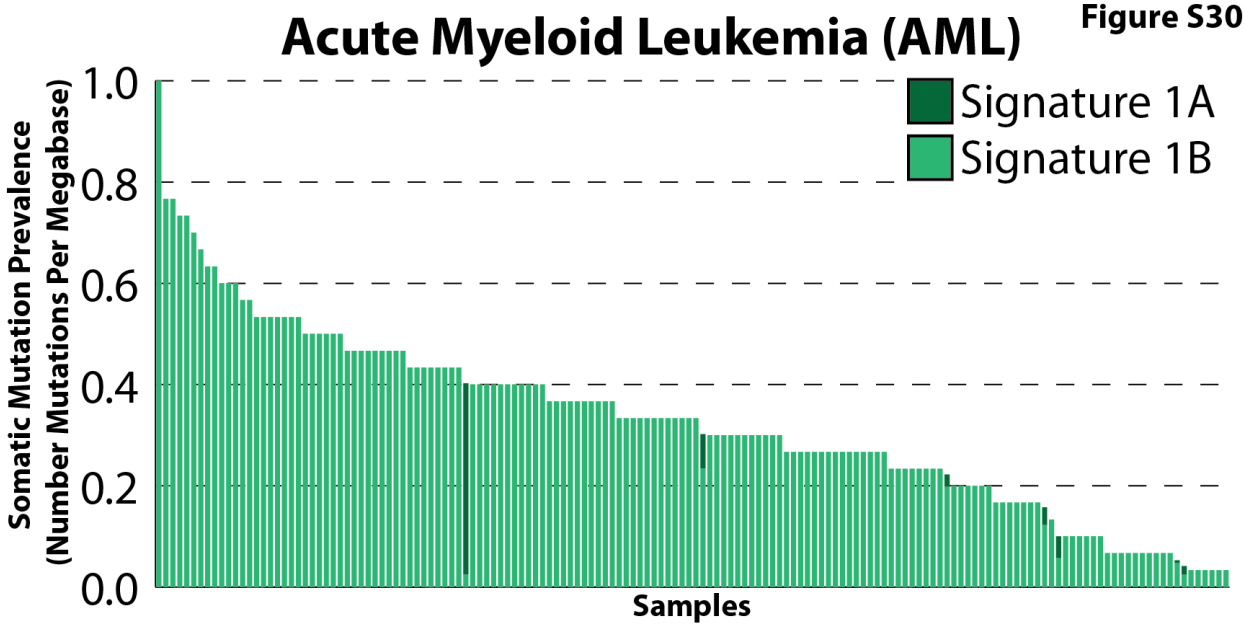
Supplementary Figure 28. Patterns of substitutions for Signature U2. ***(a)*** Mutational signature displayed based on the trinucleotide frequency of the human genome. ***(b)*** Mutational signature displayed based on an equal trinucleotide frequency. In both panels, the mutational signature is displayed using a 96 substitution classification defined by the substitution class and the sequence context immediately 3' and 5' to the mutated base. The probability bars for each of the six types of substitutions as well as the mutated bases are displayed in different colors. The mutation types are displayed on the horizontal axes, while vertical axes depict the percentage of mutations attributed to a specific mutation type. For discussion of the results from the data normalized to equal trinucleotide frequency in the genome see online methods.

Figure S29



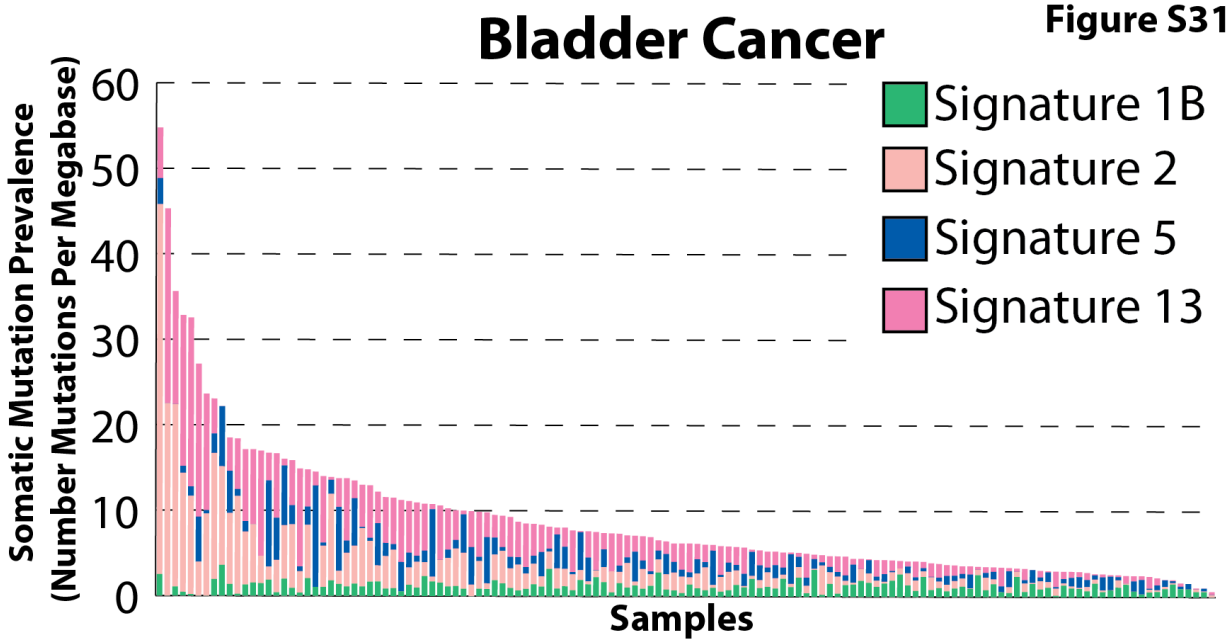
Supplementary Figure 29. Contributions of the signatures of mutational processes operative in ALL. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

Figure S30

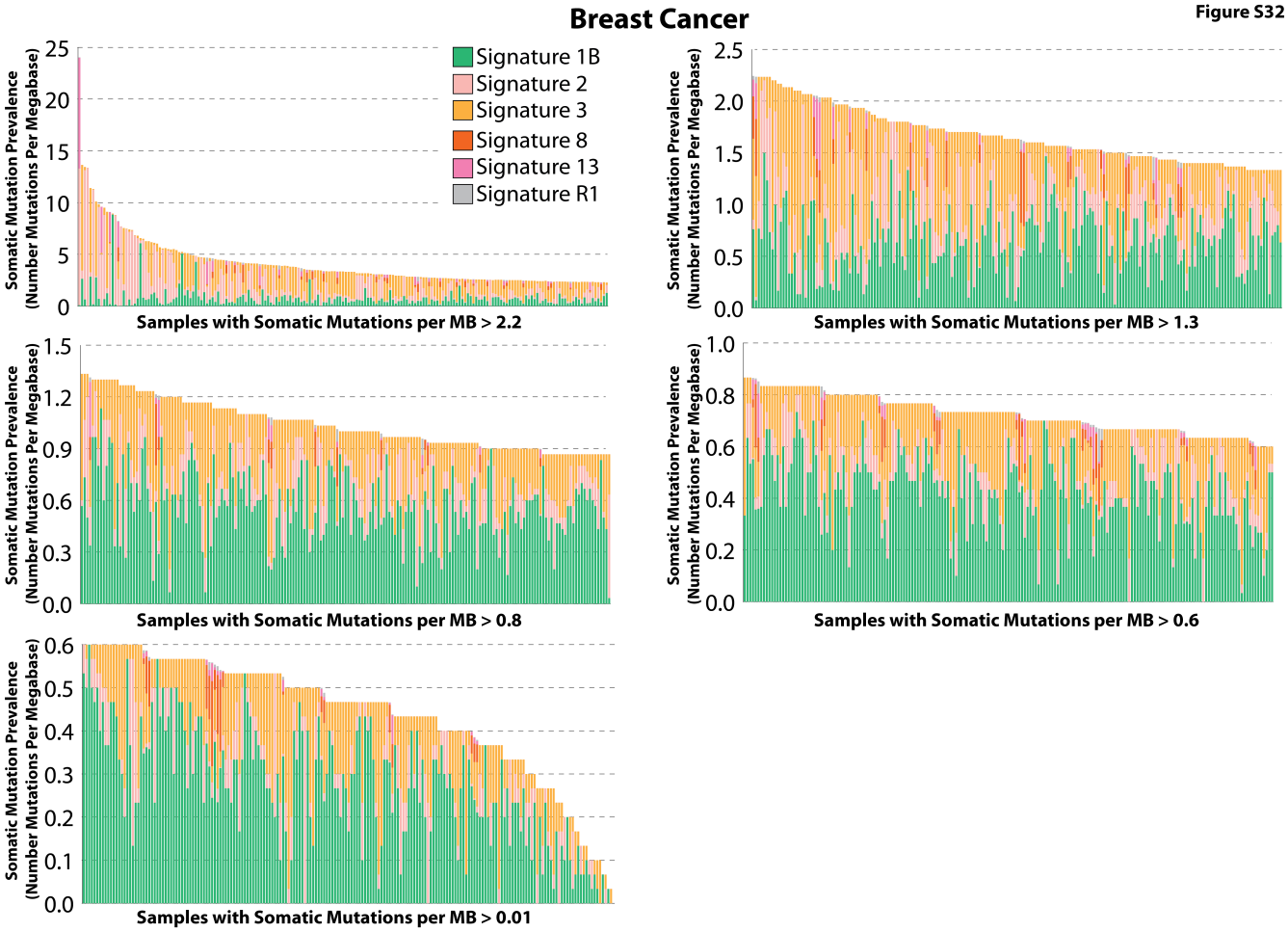


Supplementary Figure 30. Contributions of the signatures of mutational processes operative in AML. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

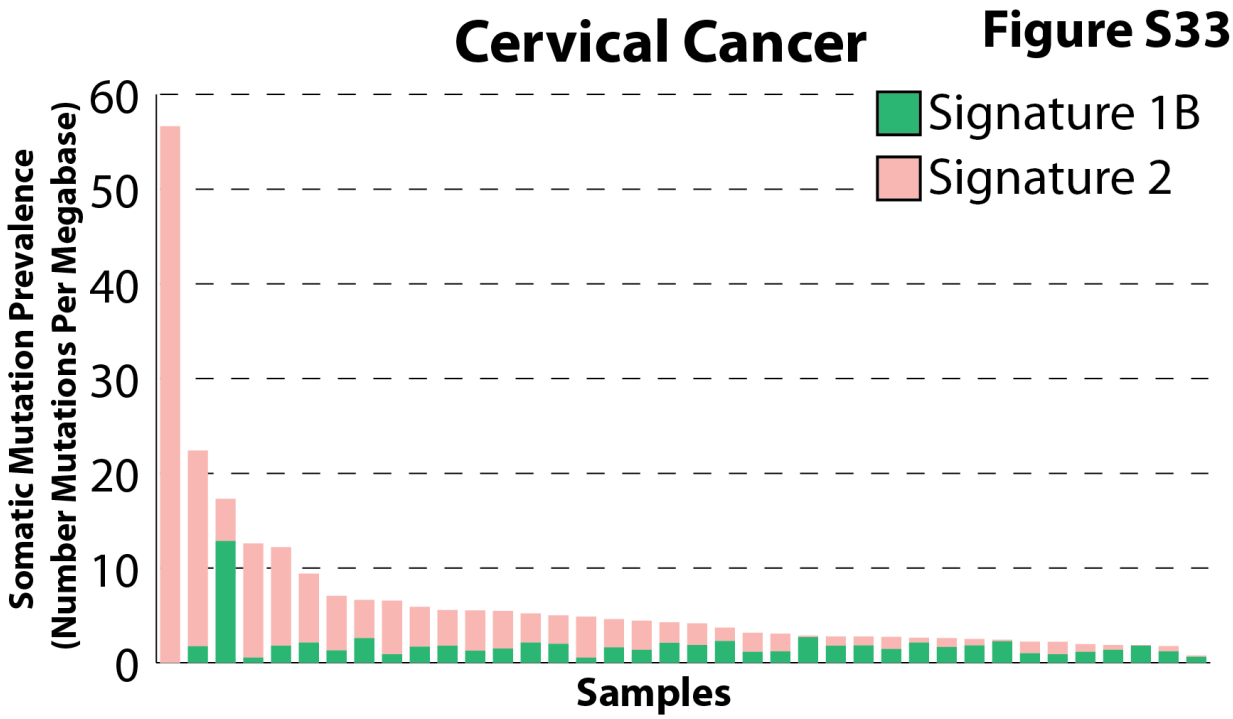
Figure S31



Supplementary Figure 31. Contributions of the signatures of mutational processes operative in bladder cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

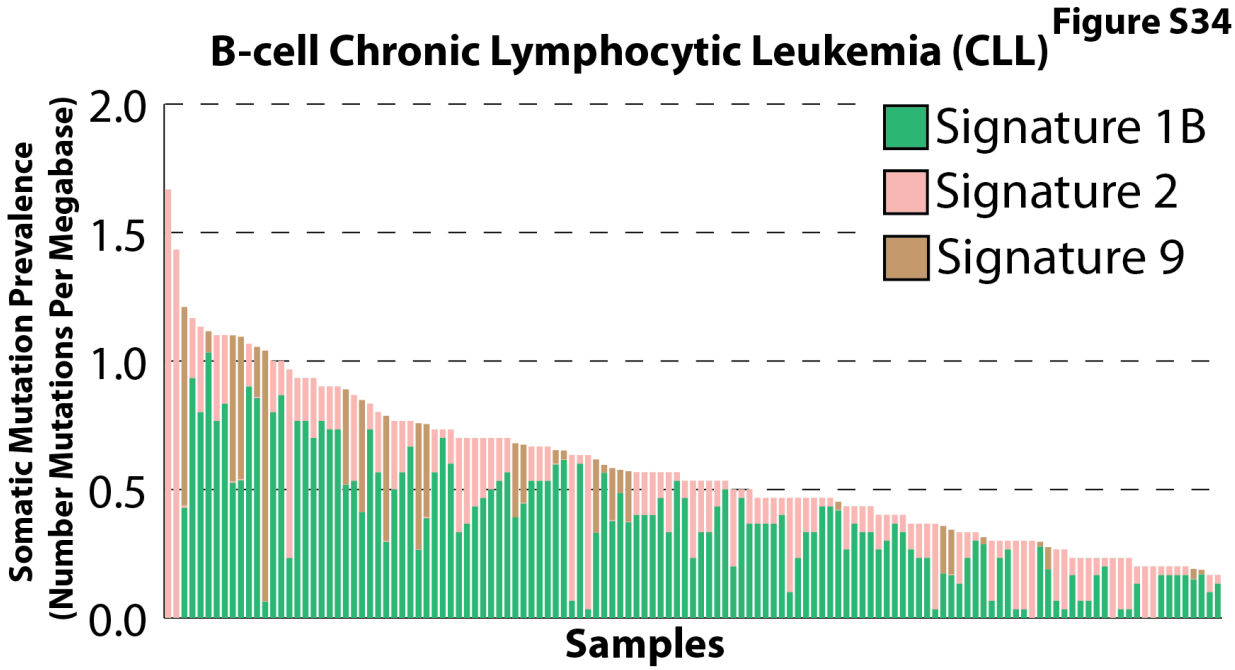


Supplementary Figure 32. Contributions of the signatures of mutational processes operative in breast cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.



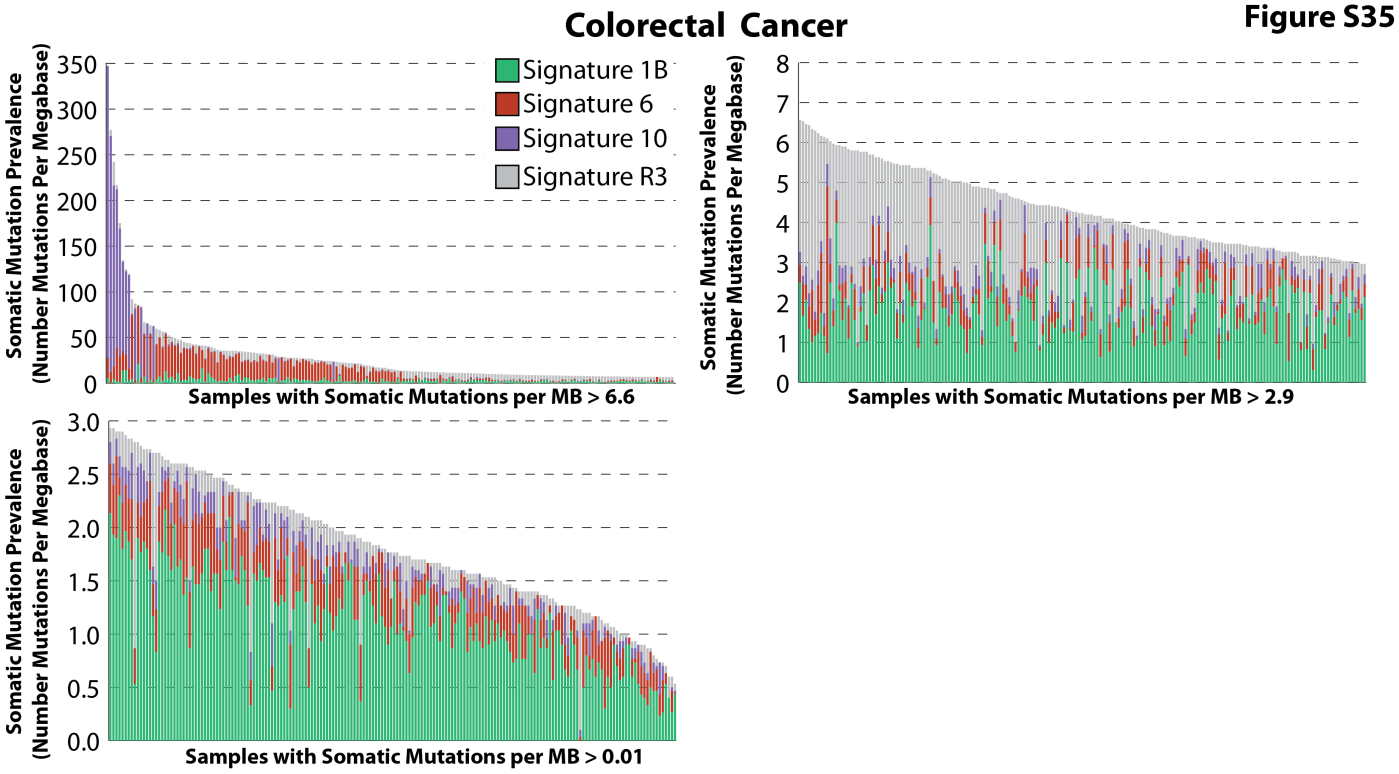
Supplementary Figure 33. Contributions of the signatures of mutational processes operative in cervical cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

Figure S34

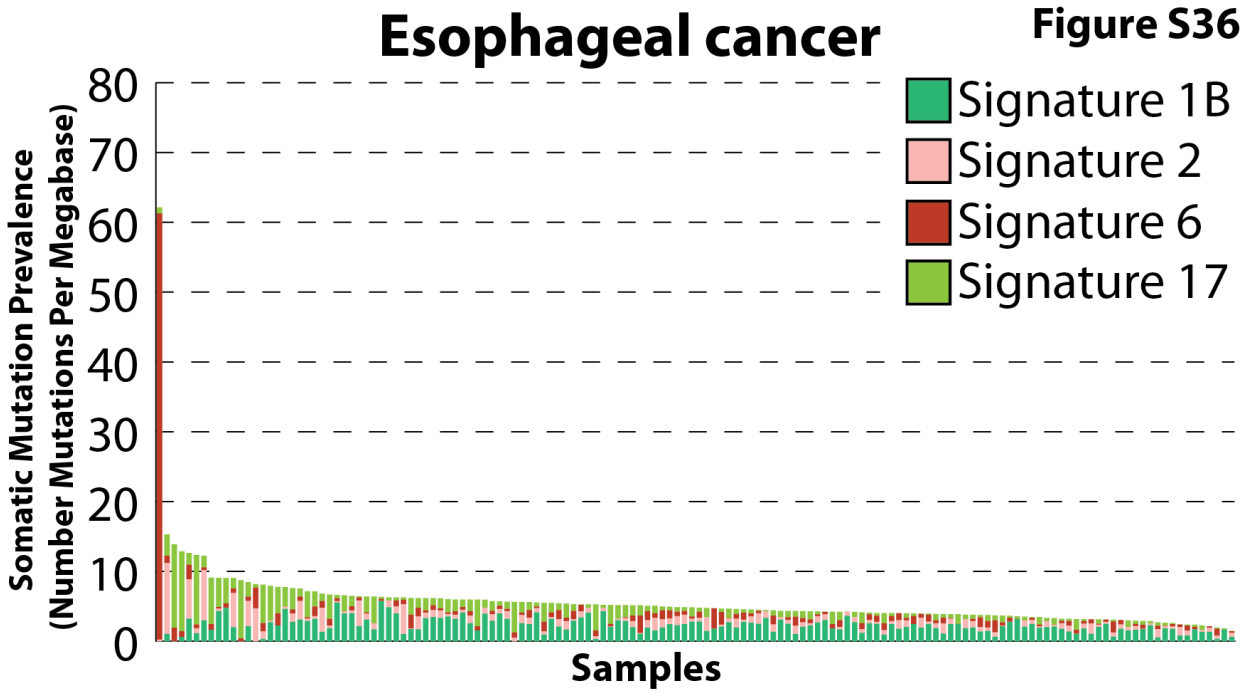


Supplementary Figure 34. Contributions of the signatures of mutational processes operative in CLL. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

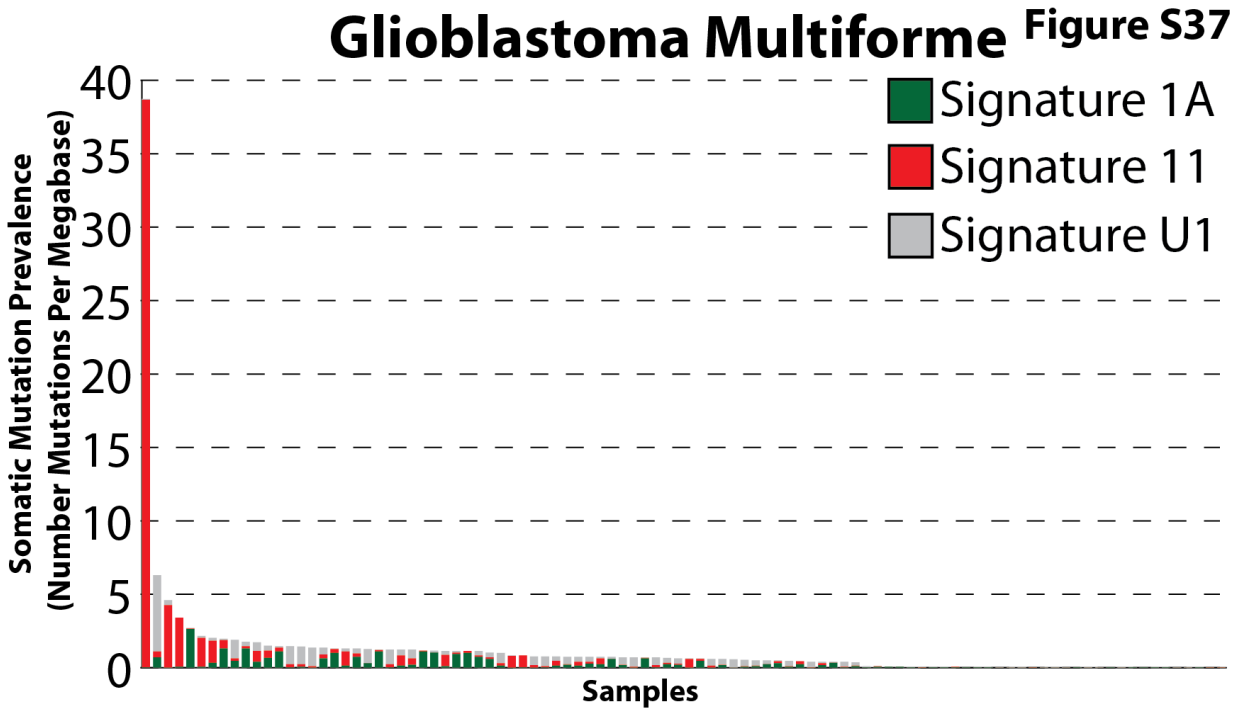
Figure S35



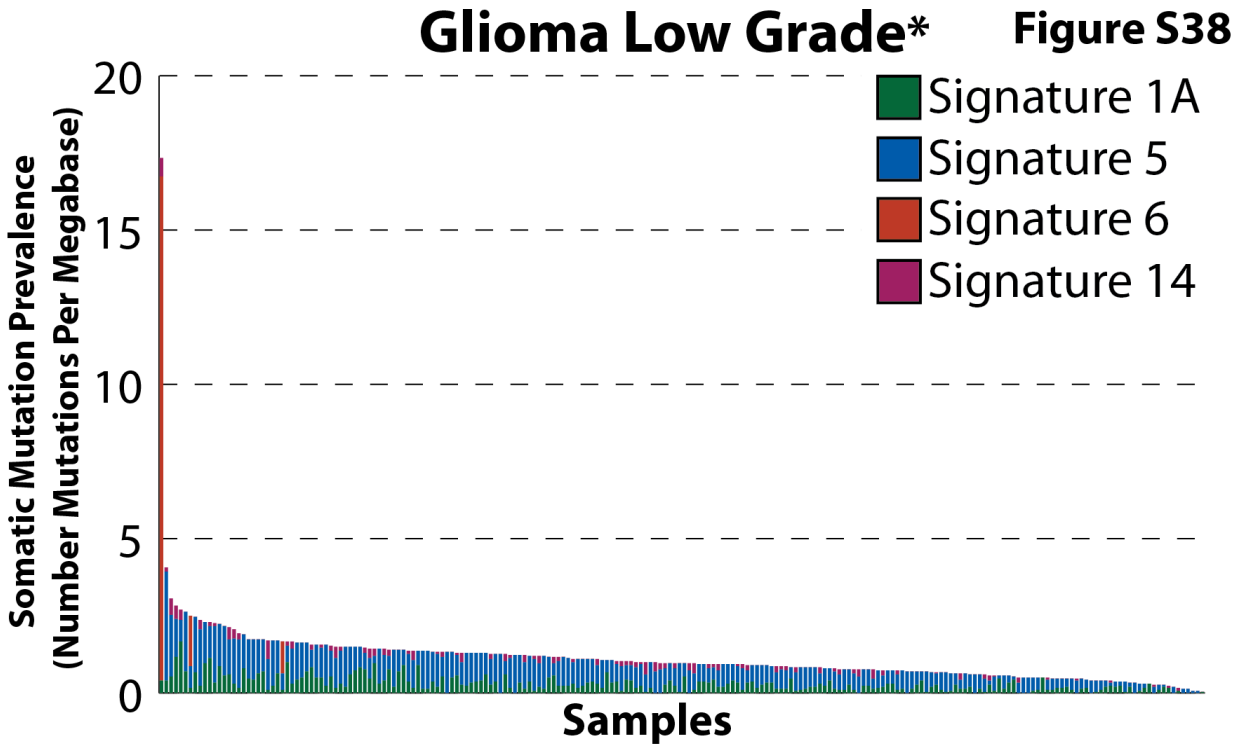
Supplementary Figure 35. Contributions of the signatures of mutational processes operative in colorectal cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.



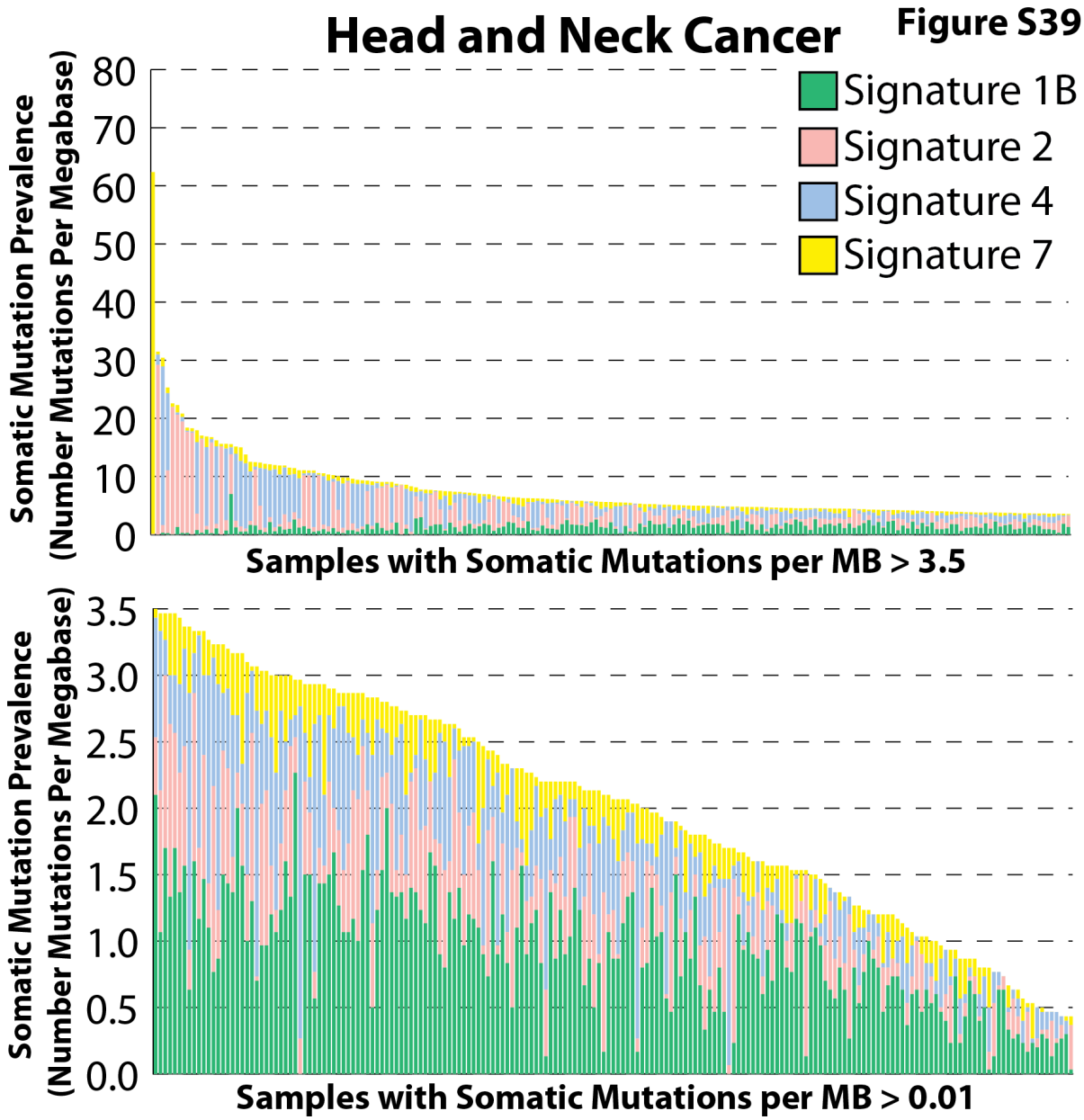
Supplementary Figure 36. Contributions of the signatures of mutational processes operative in esophageal cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.



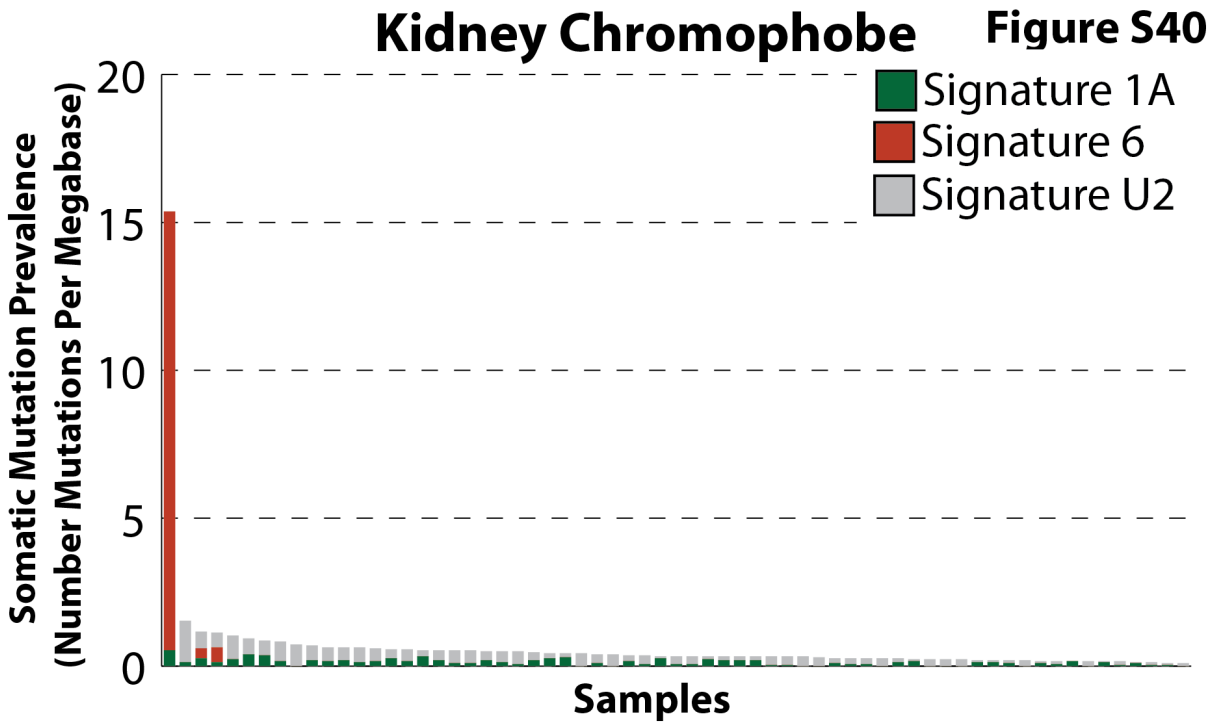
Supplementary Figure 37. Contributions of the signatures of mutational processes operative in glioblastoma multiforme. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.



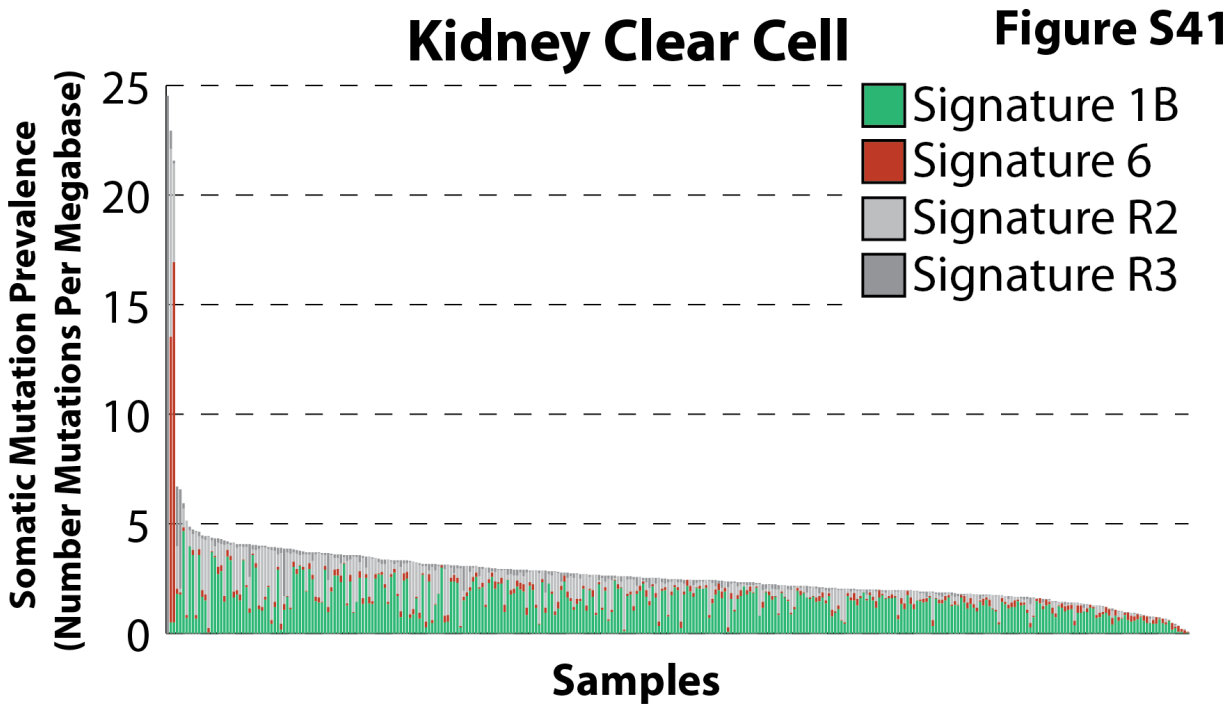
Supplementary Figure 38. Contributions of the signatures of mutational processes operative in glioma low grade. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples. *One hypermutator sample purely of Signature 14 (440 mutations per MB) is not displayed.



Supplementary Figure 39. Contributions of the signatures of mutational processes operative in head and neck cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

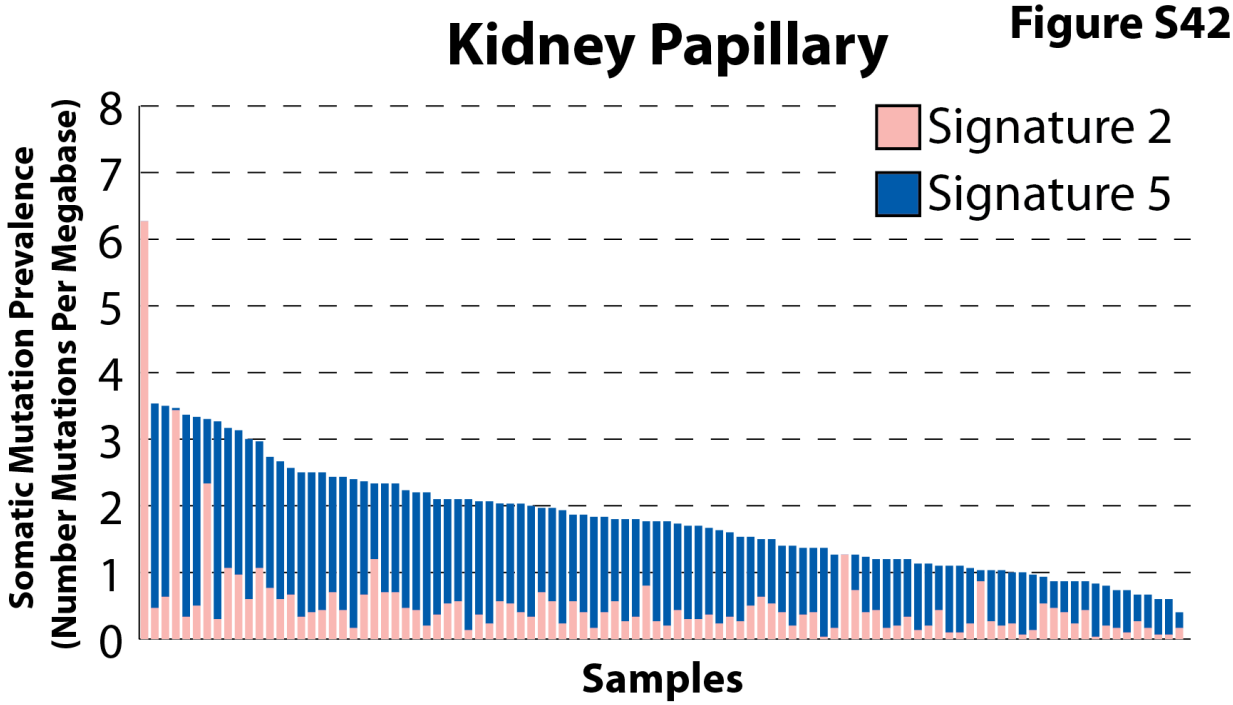


Supplementary Figure 40. Contributions of the signatures of mutational processes operative in kidney chromophobe. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.



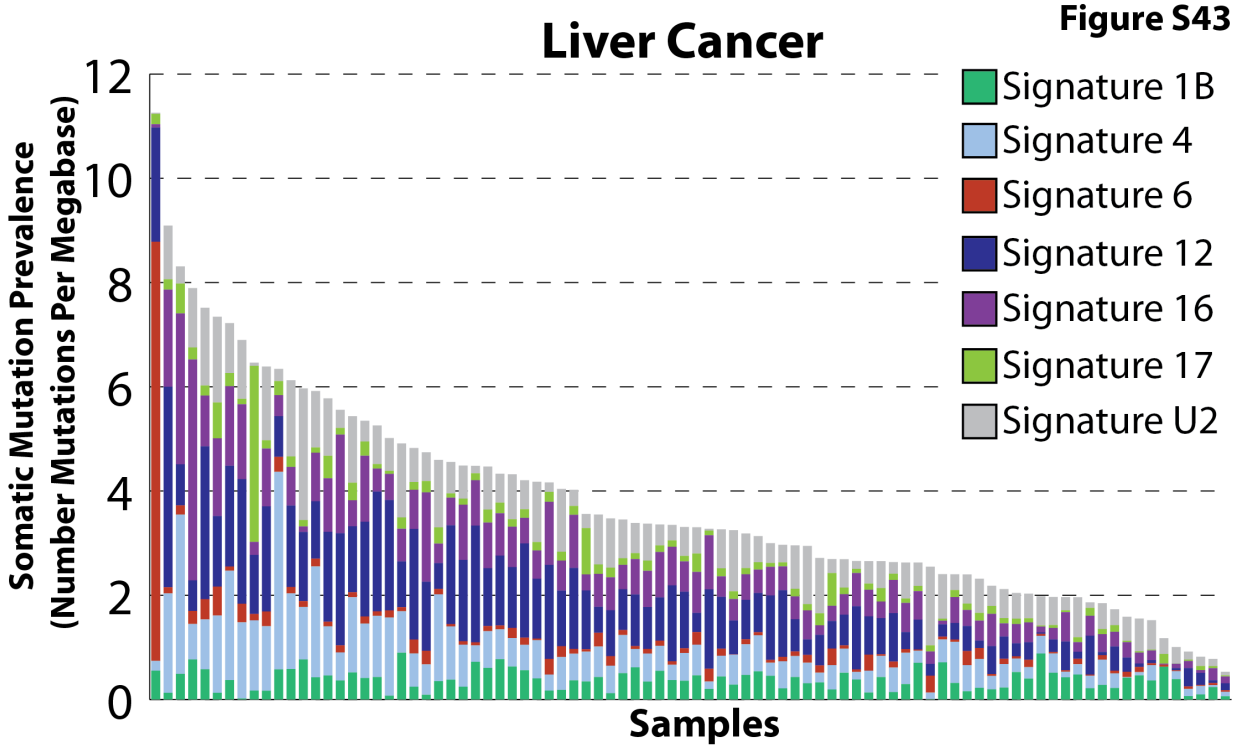
Supplementary Figure 41. Contributions of the signatures of mutational processes operative in kidney clear cell. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

Figure S42



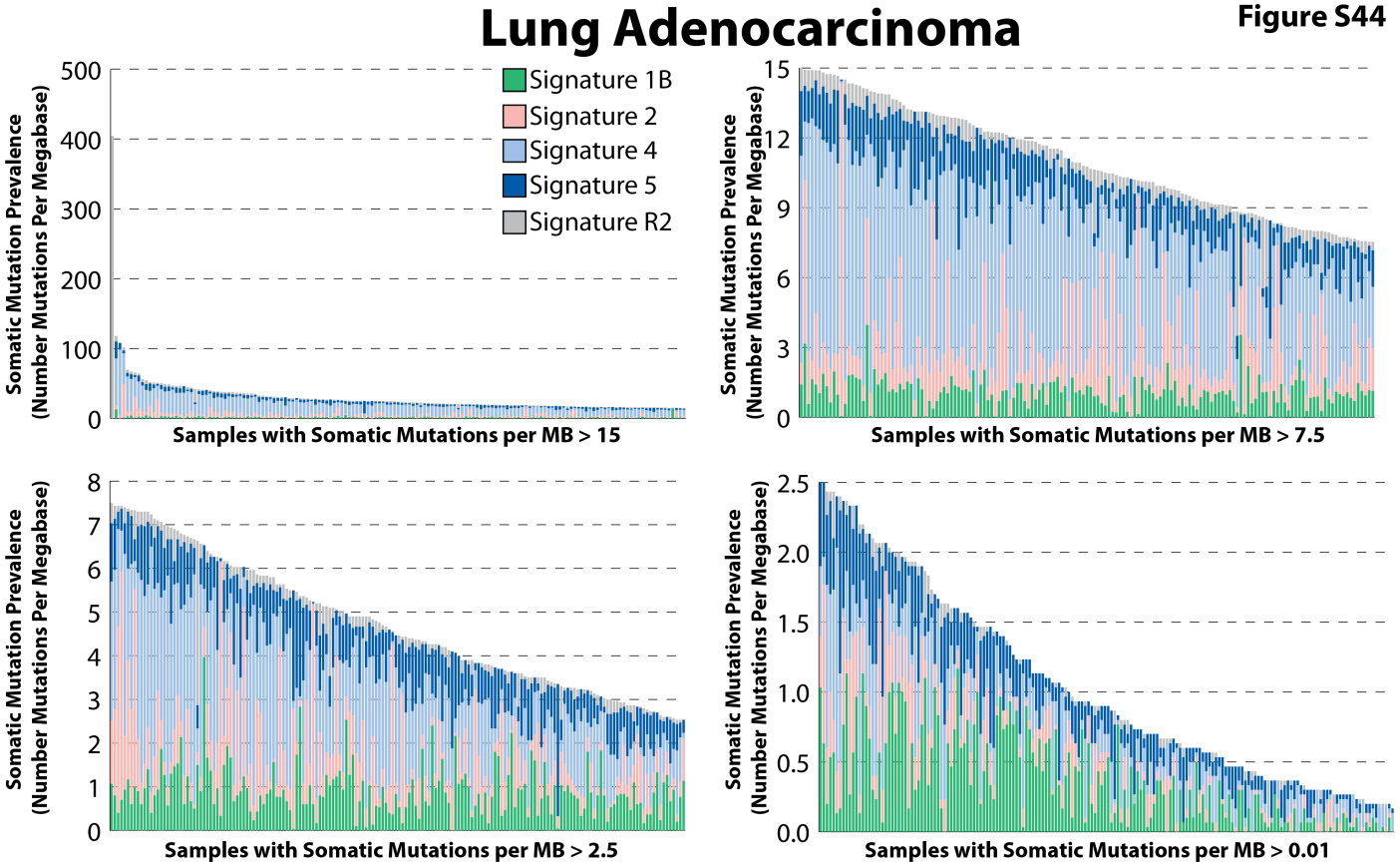
Supplementary Figure 42. Contributions of the signatures of mutational processes operative in kidney papillary. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

Figure S43

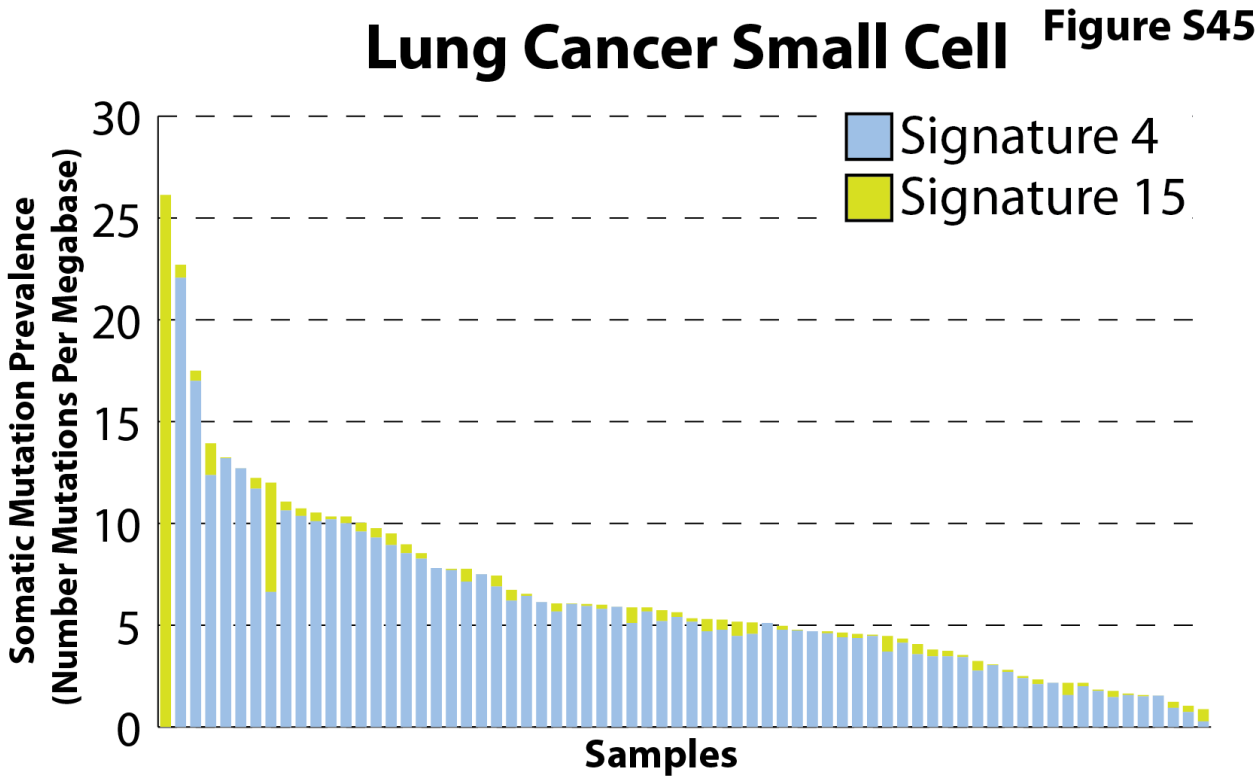


Supplementary Figure 43. Contributions of the signatures of mutational processes operative in liver cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

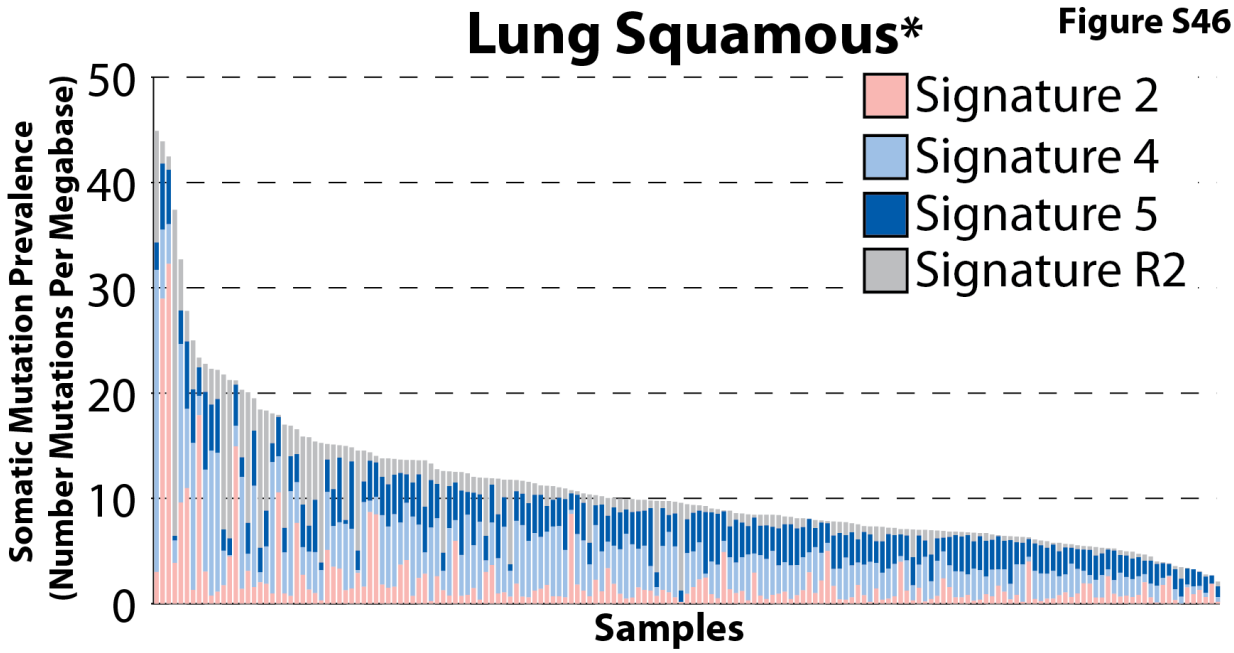
Figure S44



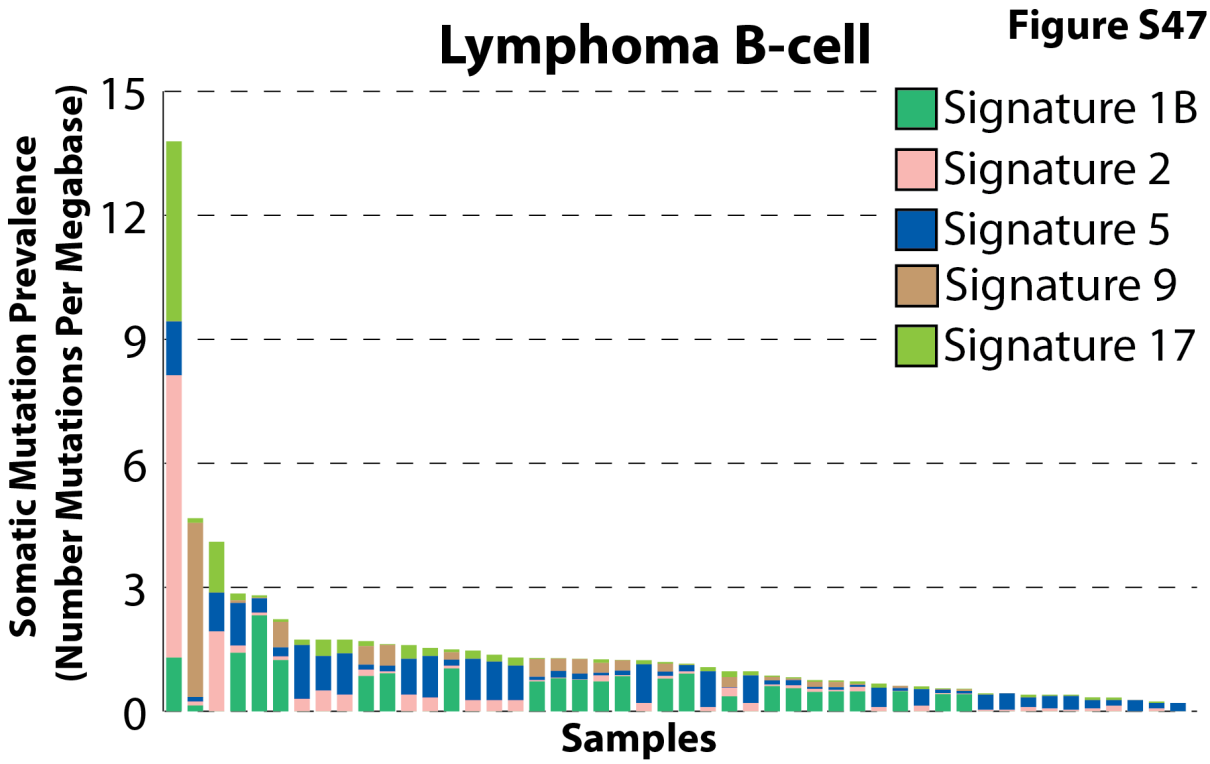
Supplementary Figure 44. Contributions of the signatures of mutational processes operative in lung adenocarcinoma. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.



Supplementary Figure 45. Contributions of the signatures of mutational processes operative in lung cancer small cell. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

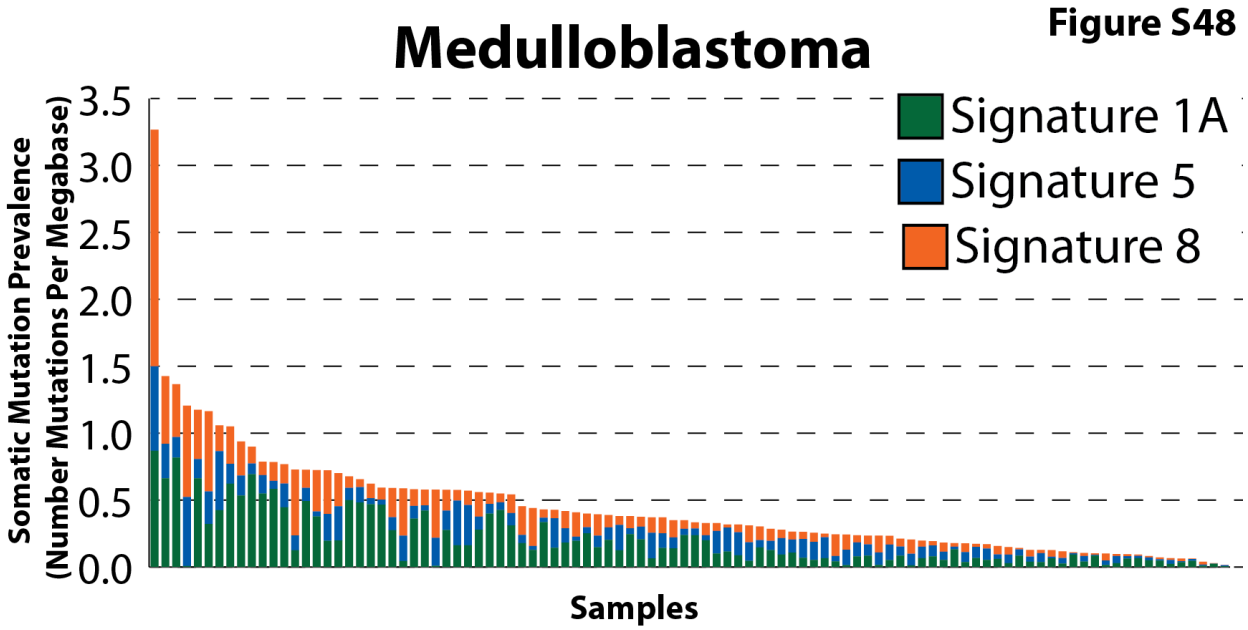


Supplementary Figure 46. Contributions of the signatures of mutational processes operative in lung squamous. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples. *One hypermutator sample purely of Signature 7 (122 mutations per MB) is not displayed. Signature 7 is associated with UV-light, an unlikely carcinogen for lung cancer. As such, we believe that this lung sample is most likely either a melanoma metastasis or a mis-annotated sample. Thus, the association between Signature 7 and lung squamous has not been displayed in Figure 3.



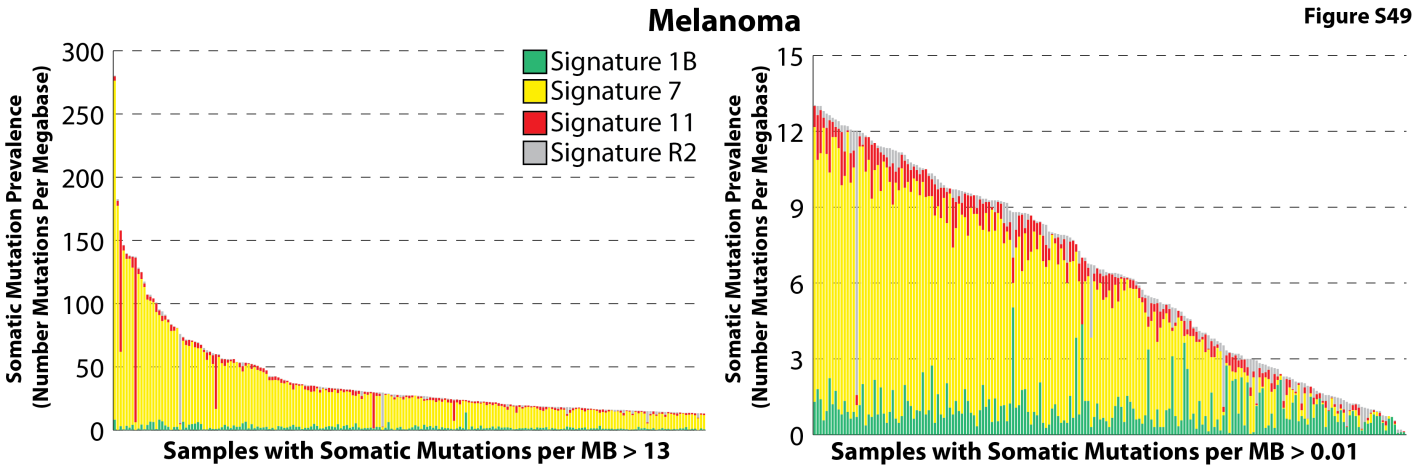
Supplementary Figure 47. Contributions of the signatures of mutational processes operative in lymphoma B-cell. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

Figure S48

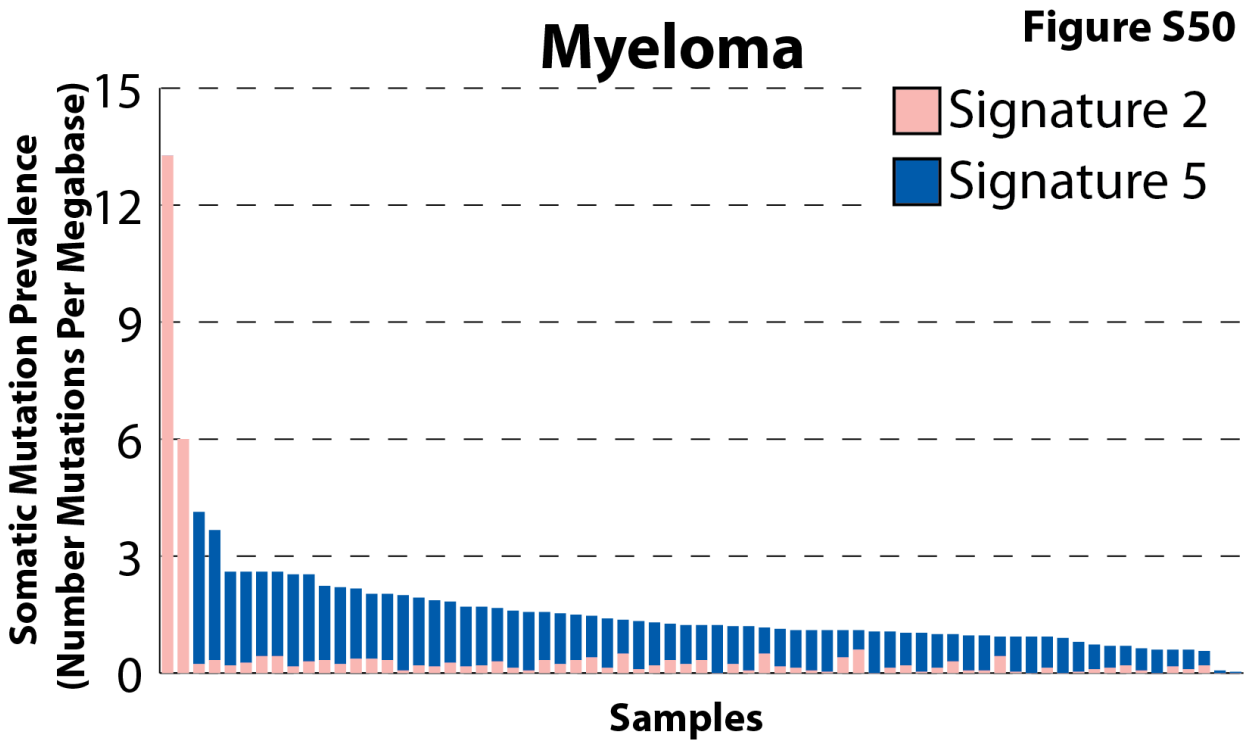


Supplementary Figure 48. Contributions of the signatures of mutational processes operative in medulloblastoma. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

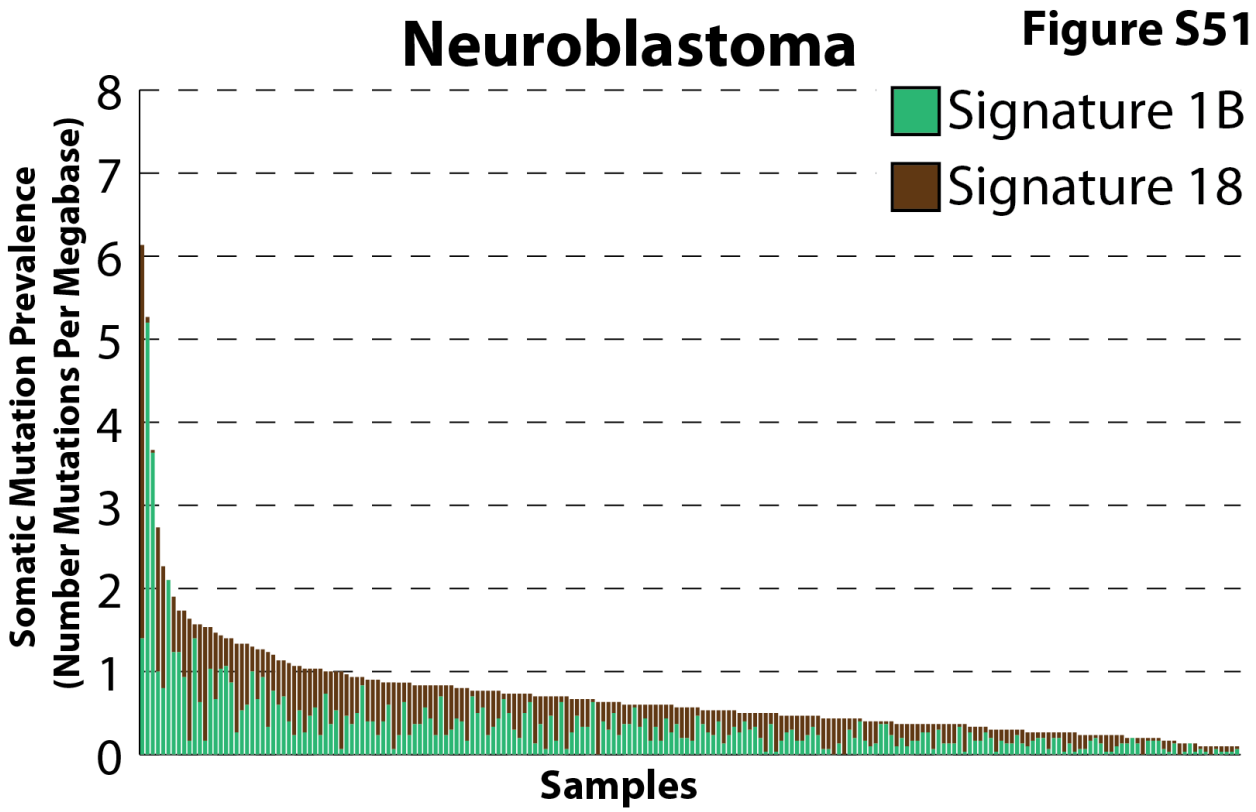
Figure S49



Supplementary Figure 49. Contributions of the signatures of mutational processes operative in melanoma. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

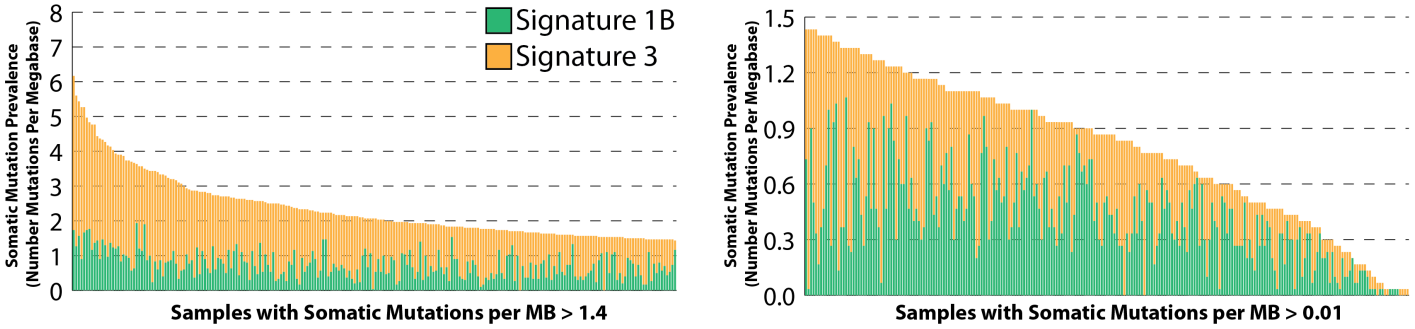


Supplementary Figure 50. Contributions of the signatures of mutational processes operative in myeloma. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.



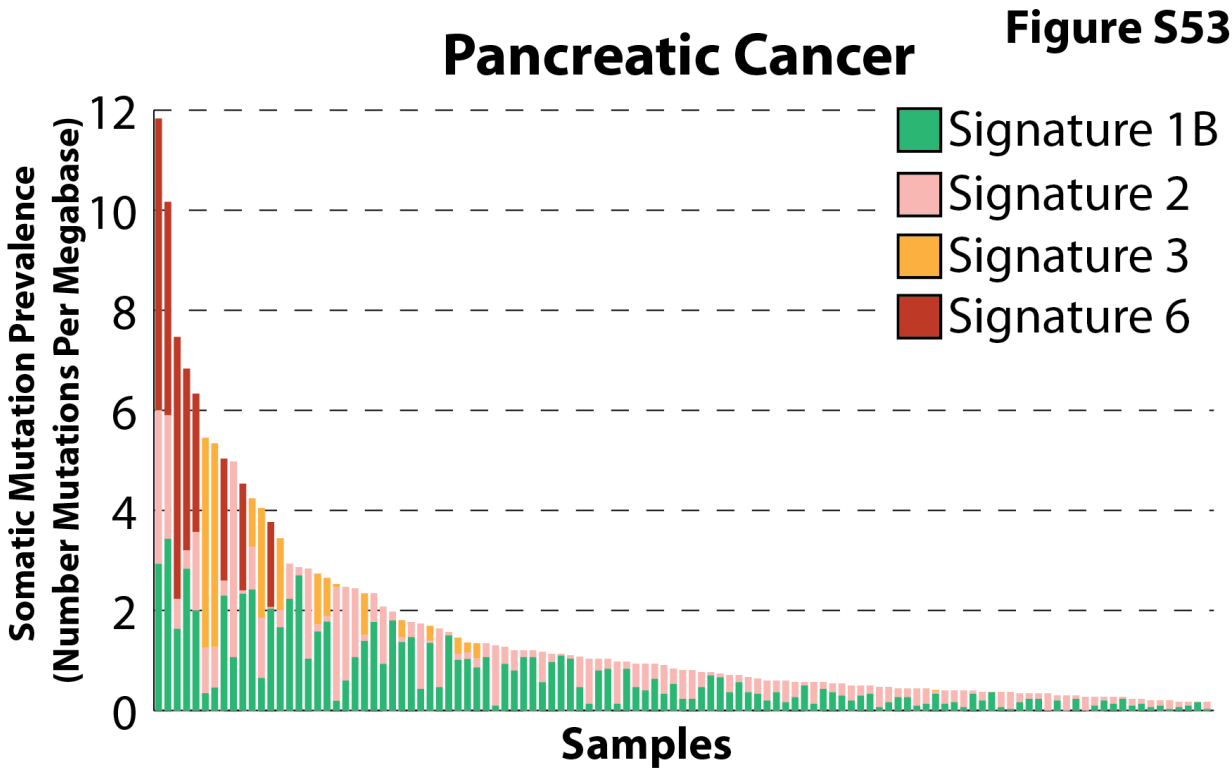
Supplementary Figure 51. Contributions of the signatures of mutational processes operative in neuroblastoma. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

Figure S52

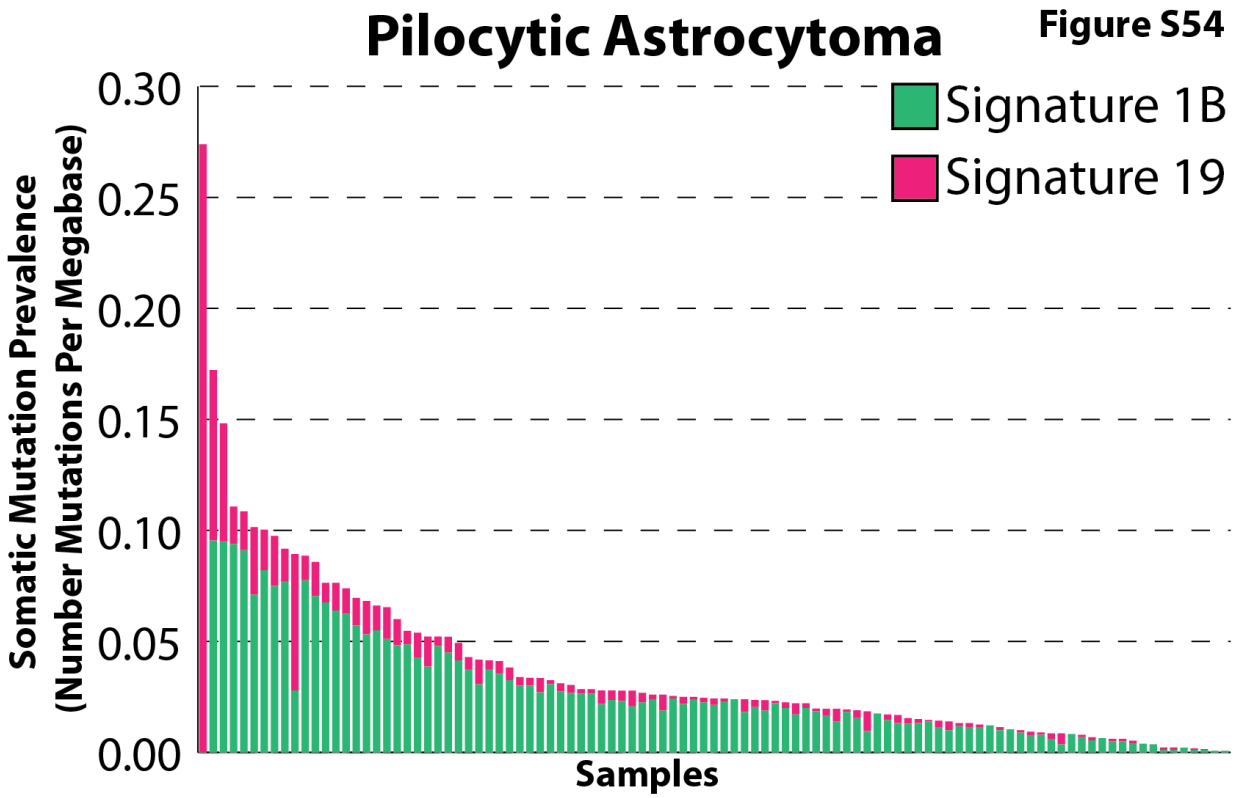


Supplementary Figure 52. Contributions of the signatures of mutational processes operative in ovarian cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

Figure S53

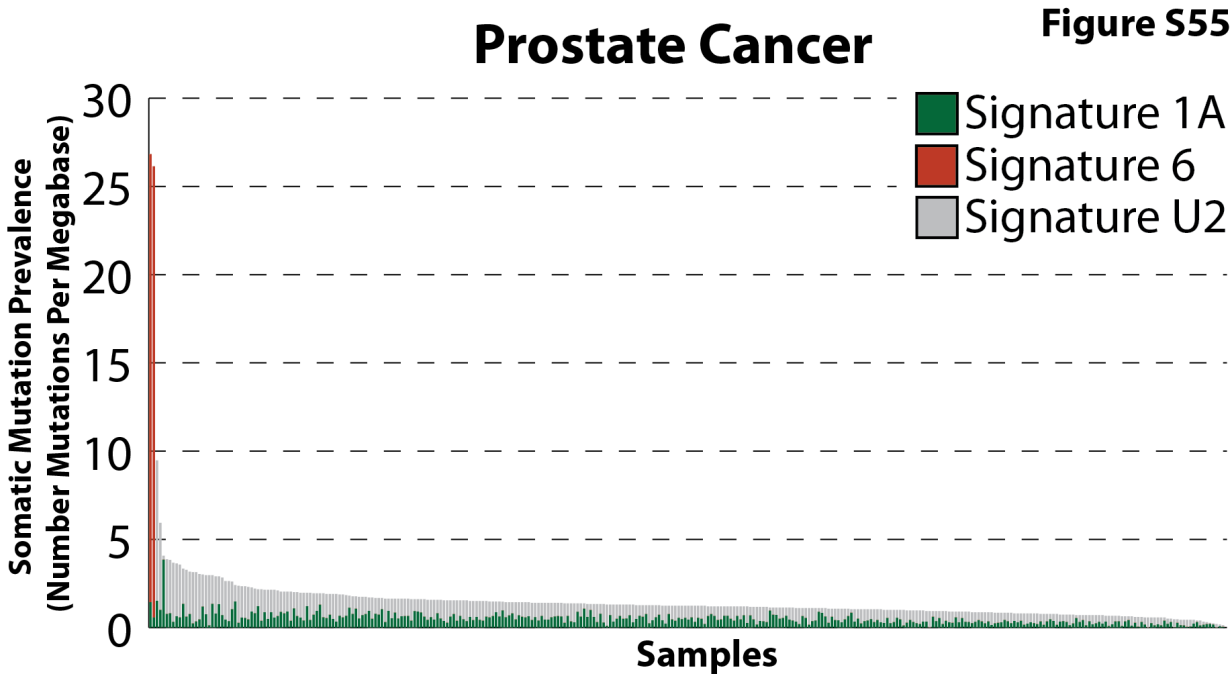


Supplementary Figure 53. Contributions of the signatures of mutational processes operative in pancreatic cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.



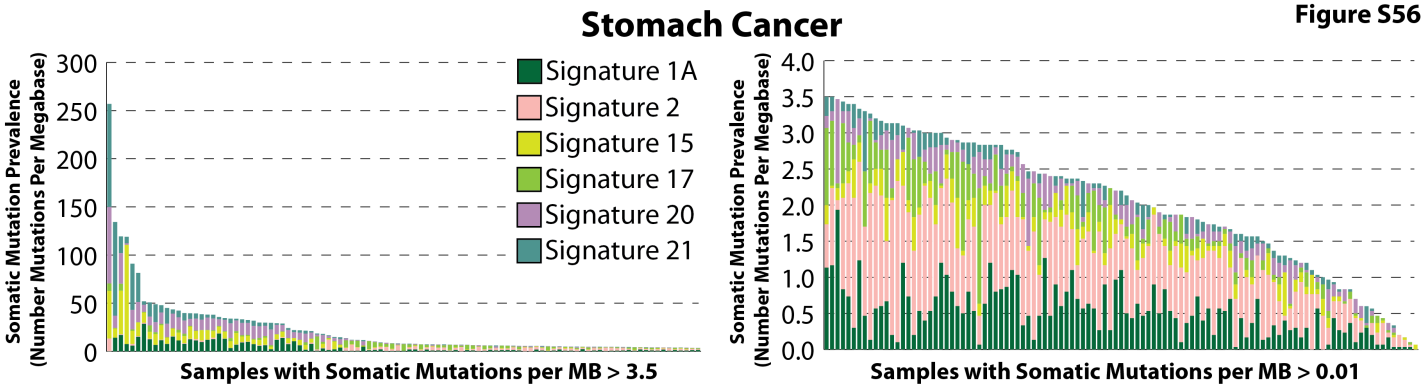
Supplementary Figure 54. Contributions of the signatures of mutational processes operative in pilocytic astrocytoma. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples. Note that this dataset is predominantly composed of pilocytic astrocytomas but also includes a small number of other pediatric low-grade gliomas and pediatric low-grade glioneuronal tumors

Figure S55



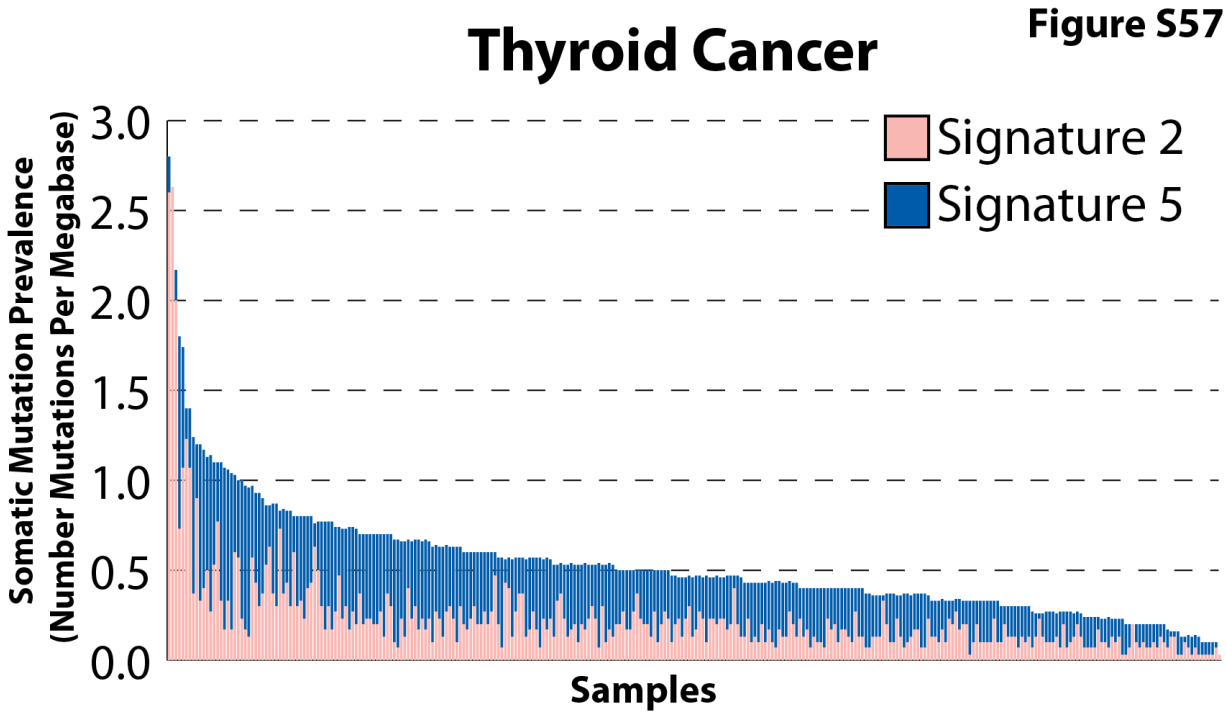
Supplementary Figure 55. Contributions of the signatures of mutational processes operative in prostate cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

Figure S56

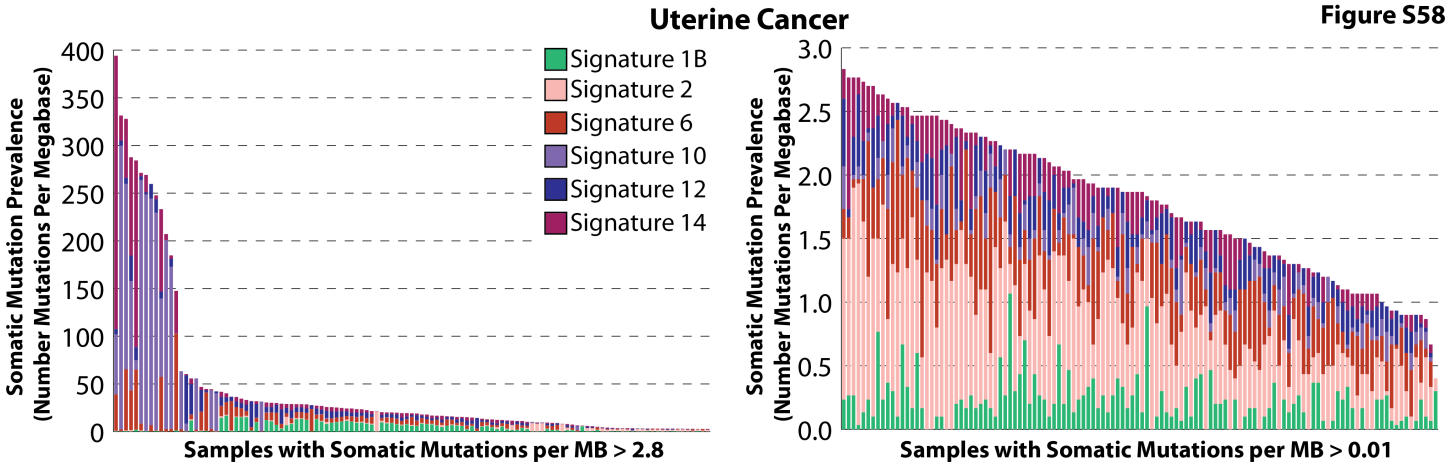


Supplementary Figure 56. Contributions of the signatures of mutational processes operative in stomach cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.

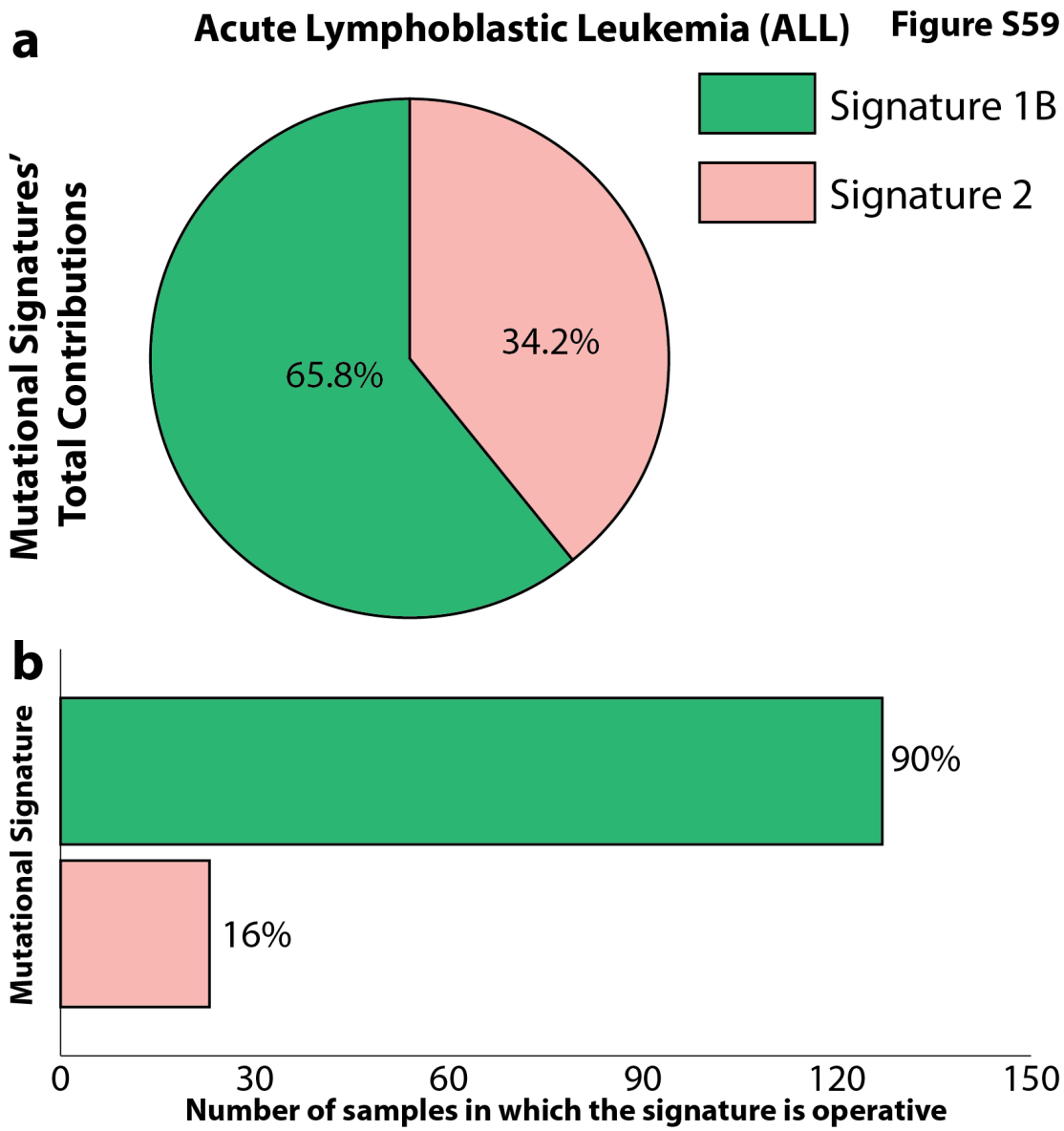
Figure S57



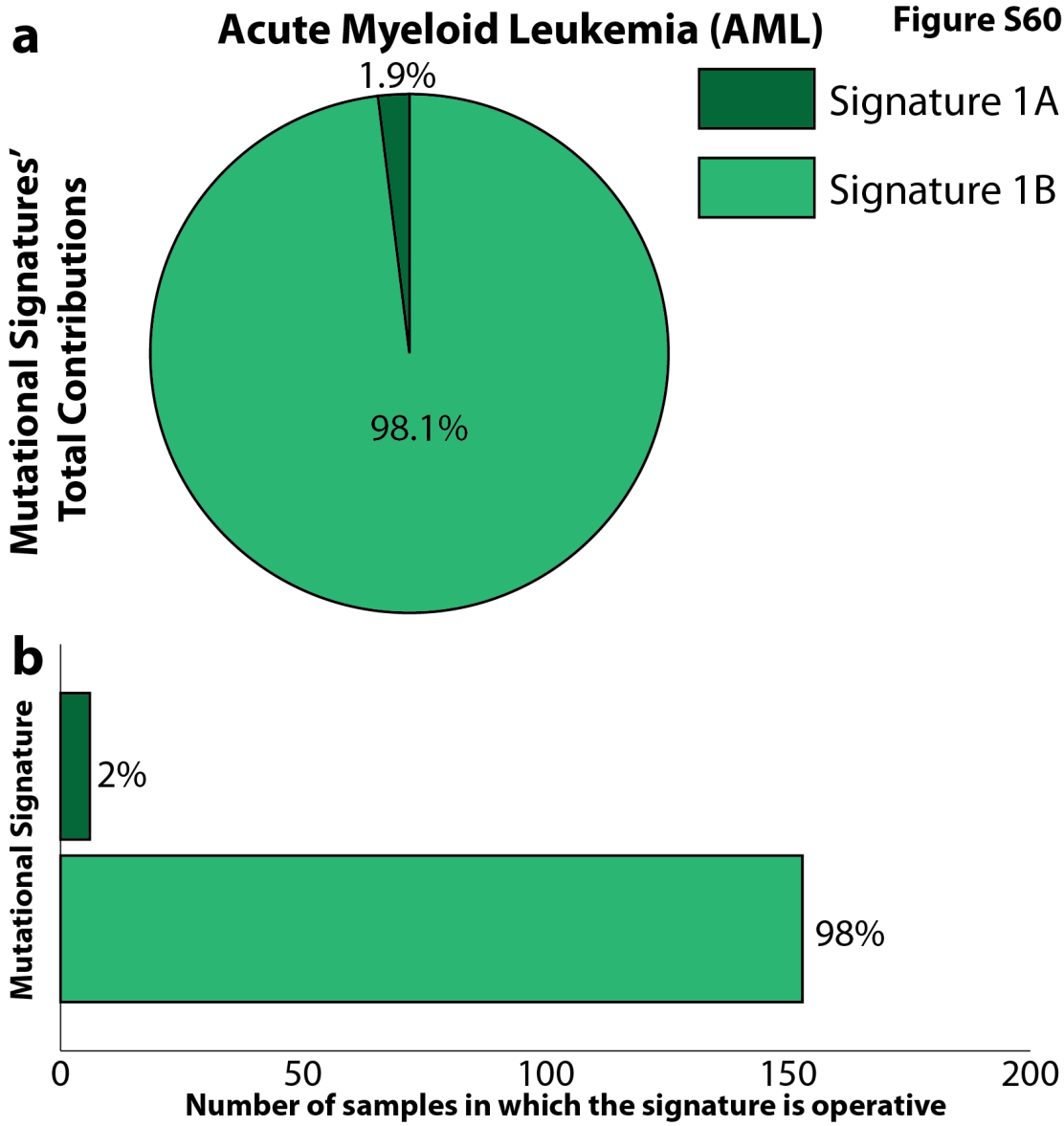
Supplementary Figure 57. Contributions of the signatures of mutational processes operative in thyroid cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.



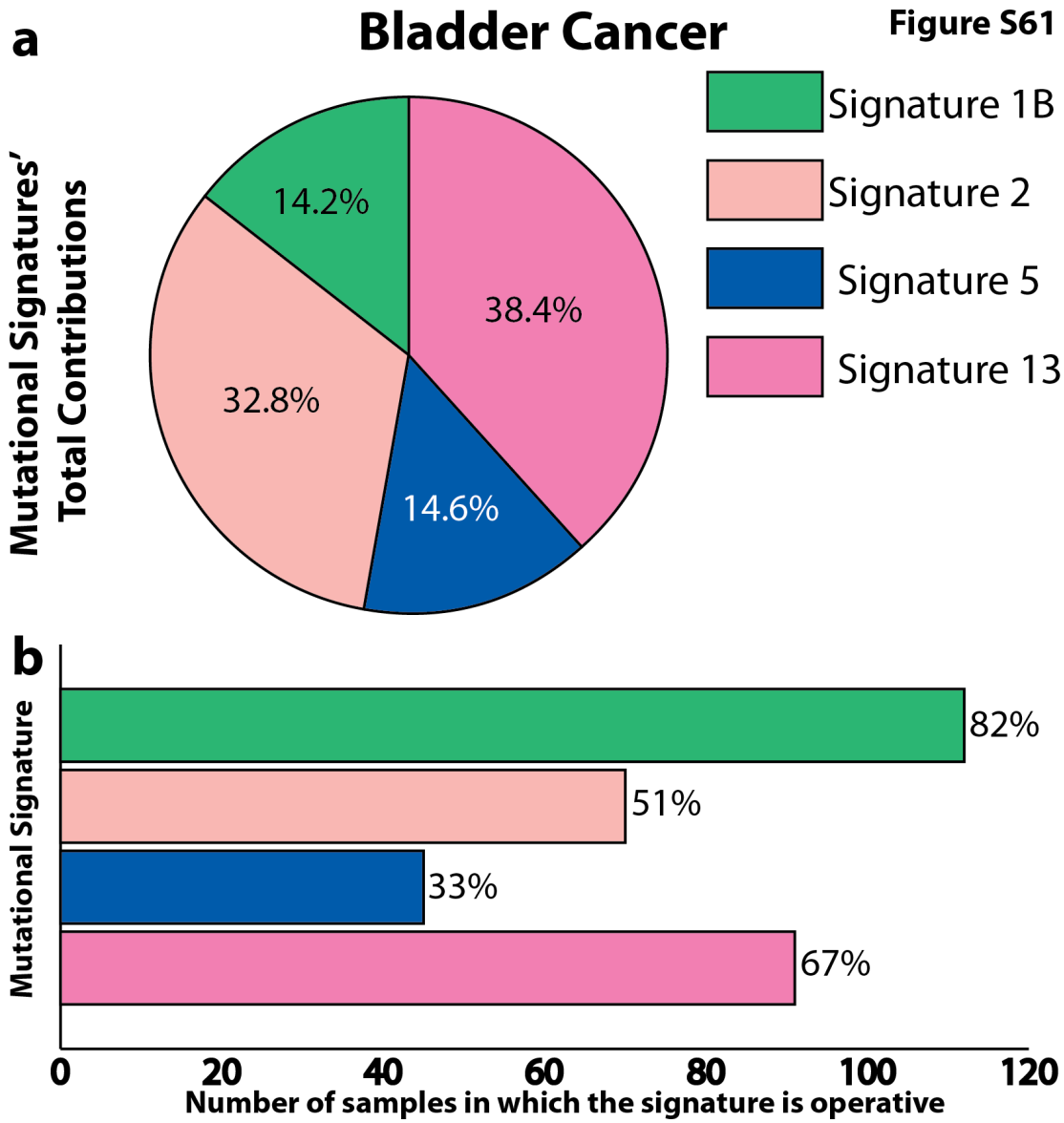
Supplementary Figure 58. Contributions of the signatures of mutational processes operative in uterine cancer. Samples are displayed on the horizontal axis, sorted in descending order based on the numbers of somatic mutations per megabase found in each sample. The somatic mutation prevalence is displayed on the vertical axis. Mutational signatures are displayed in distinct colors, consistent in all figures. For clarity, several panels are provided (and clearly labeled) when the number of samples is too high or the somatic prevalence differs significantly between samples.



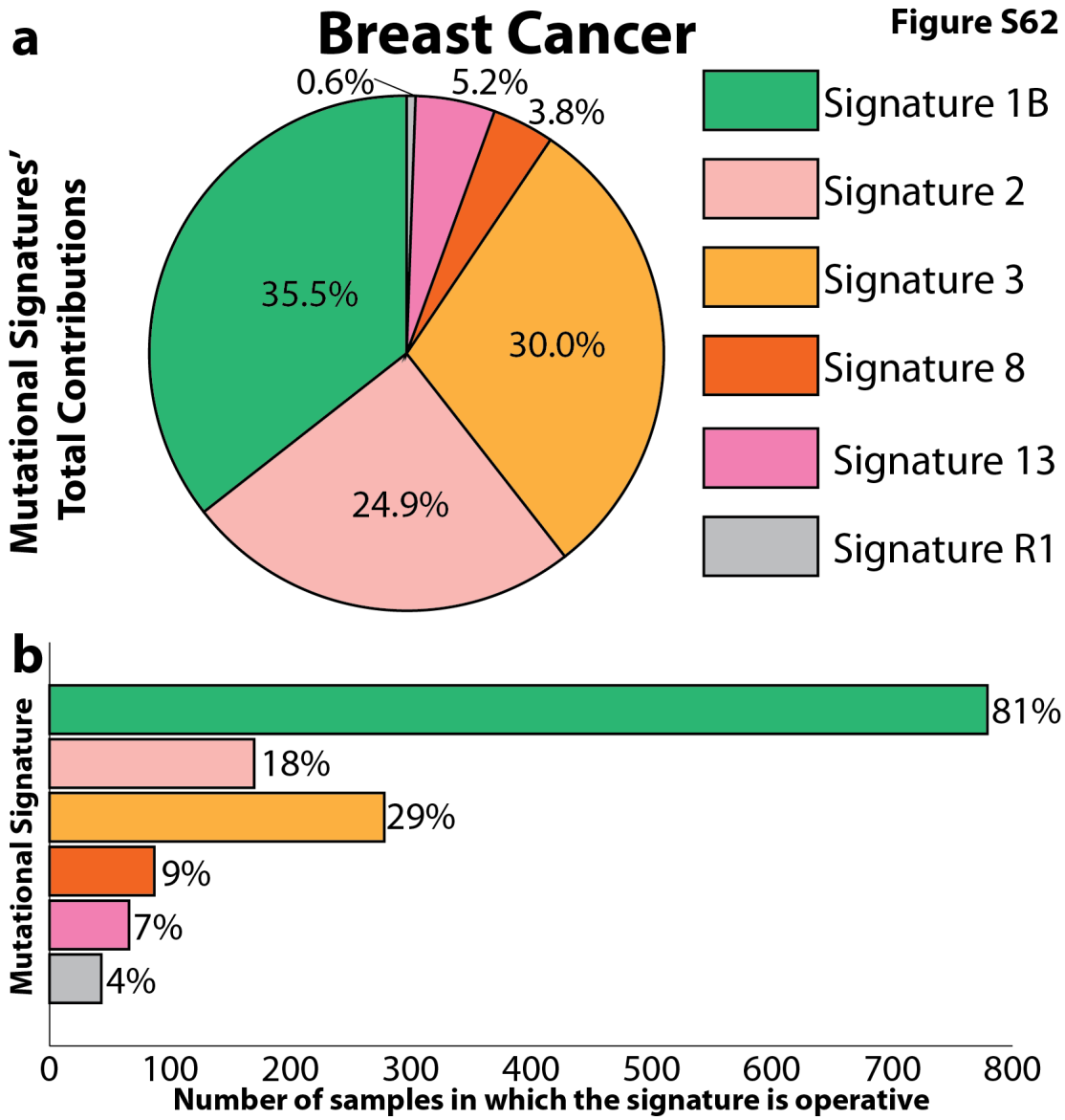
Supplementary Figure 59. Summary of the contributions of the signatures of mutational processes operative in ALL. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



Supplementary Figure 60. Summary of the contributions of the signatures of mutational processes operative in AML. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.

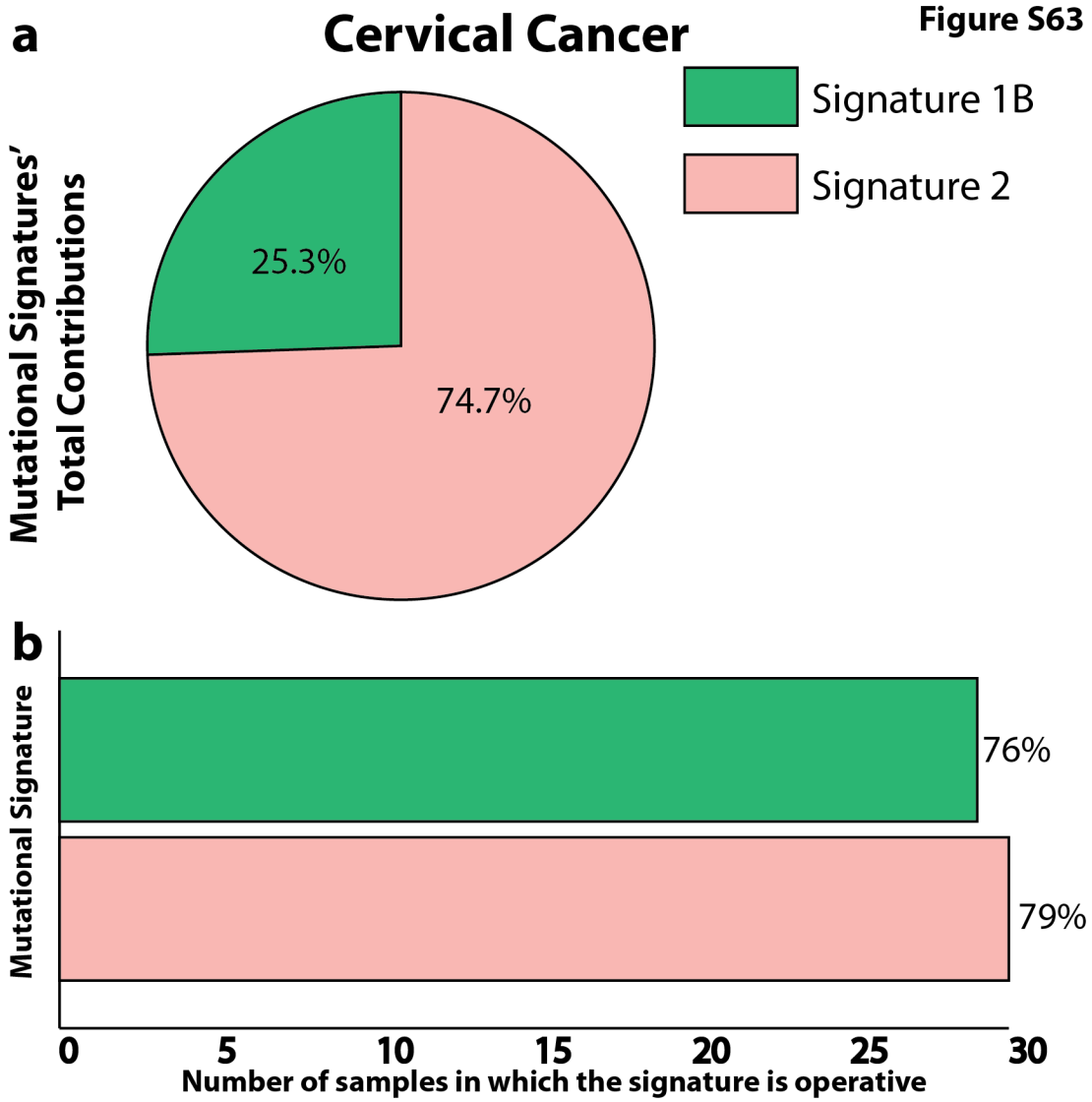


Supplementary Figure 61. Summary of the contributions of the signatures of mutational processes operative in bladder cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.

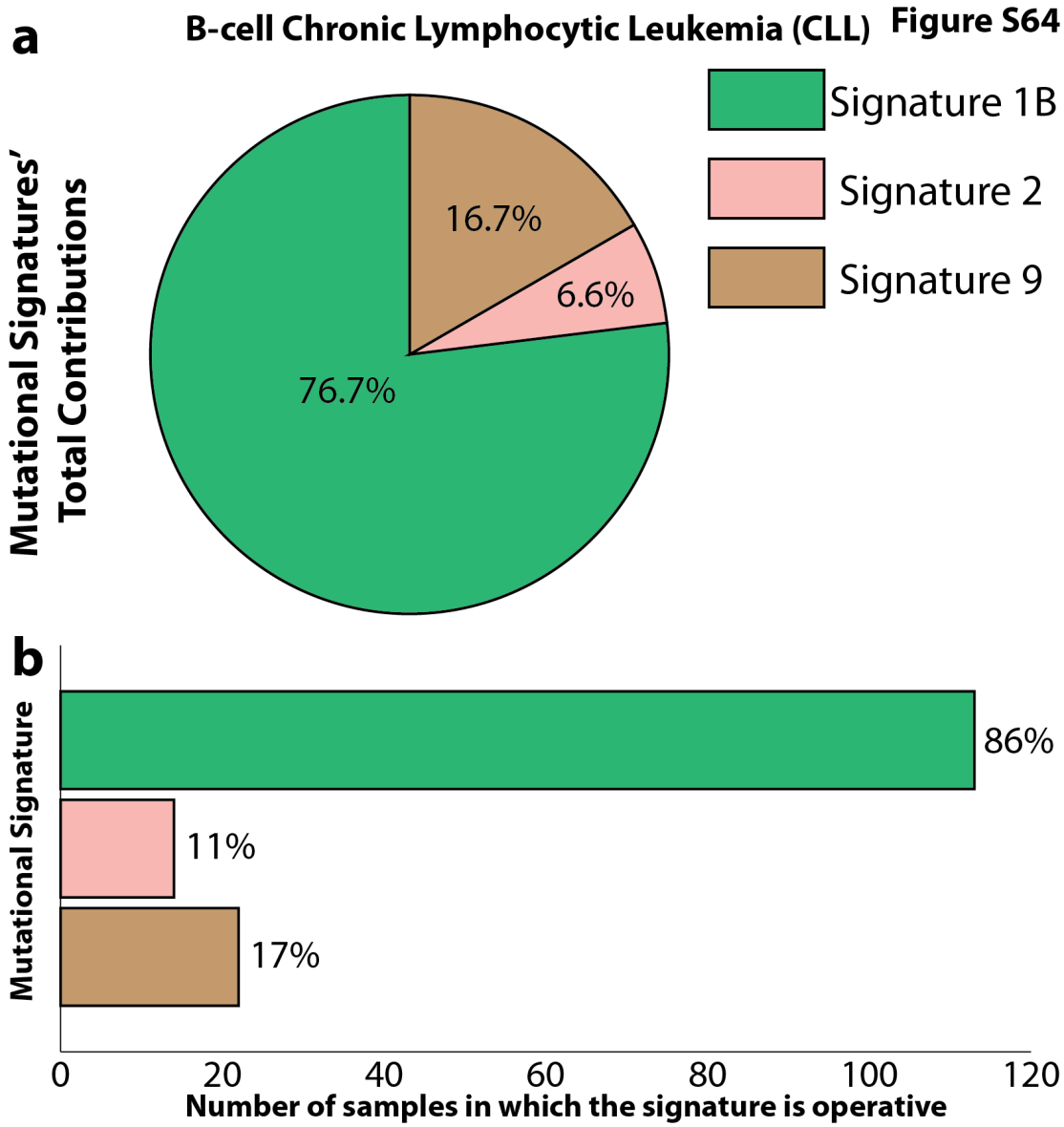


Supplementary Figure 62. Summary of the contributions of the signatures of mutational processes operative in breast cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.

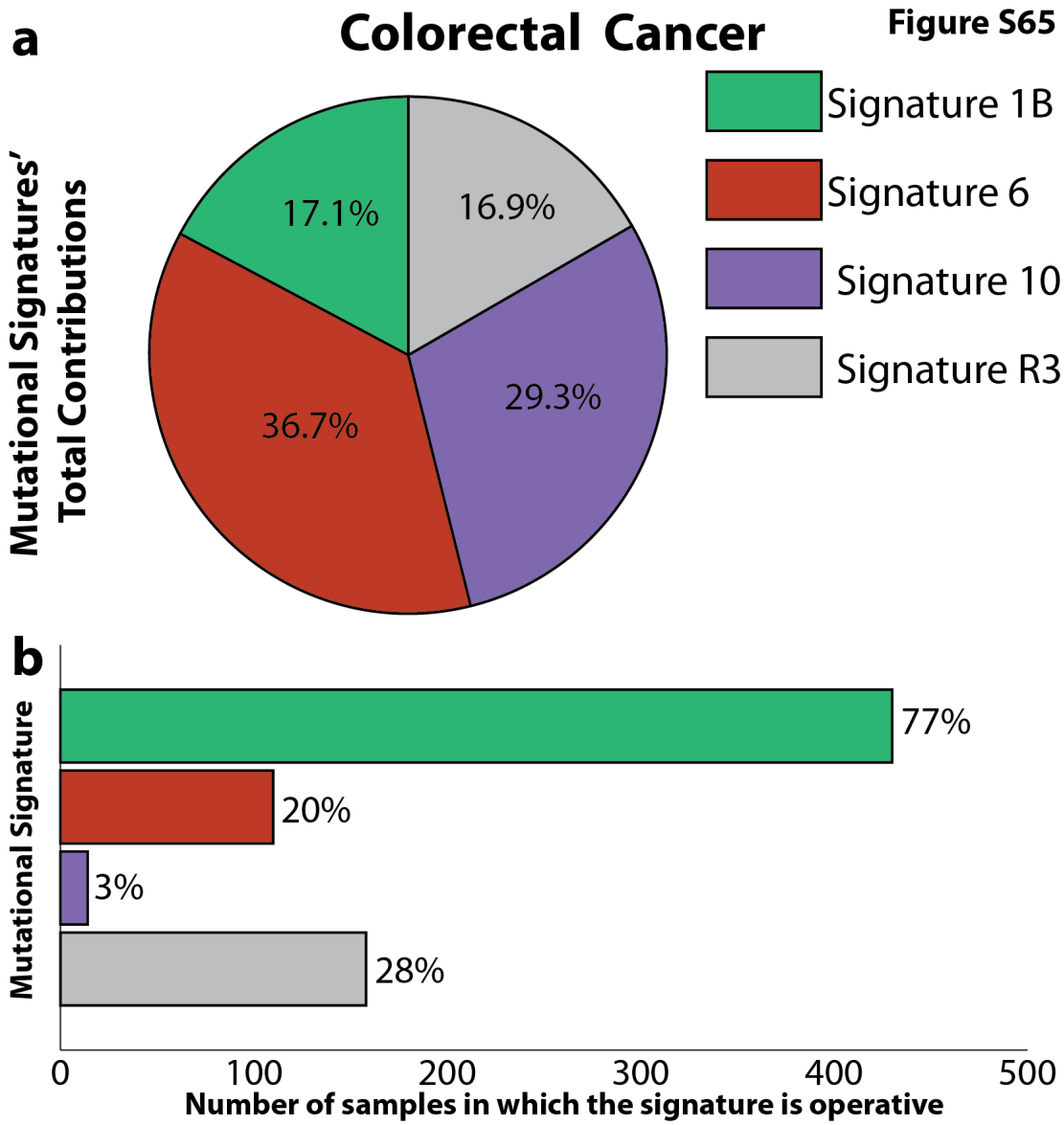
Figure S63



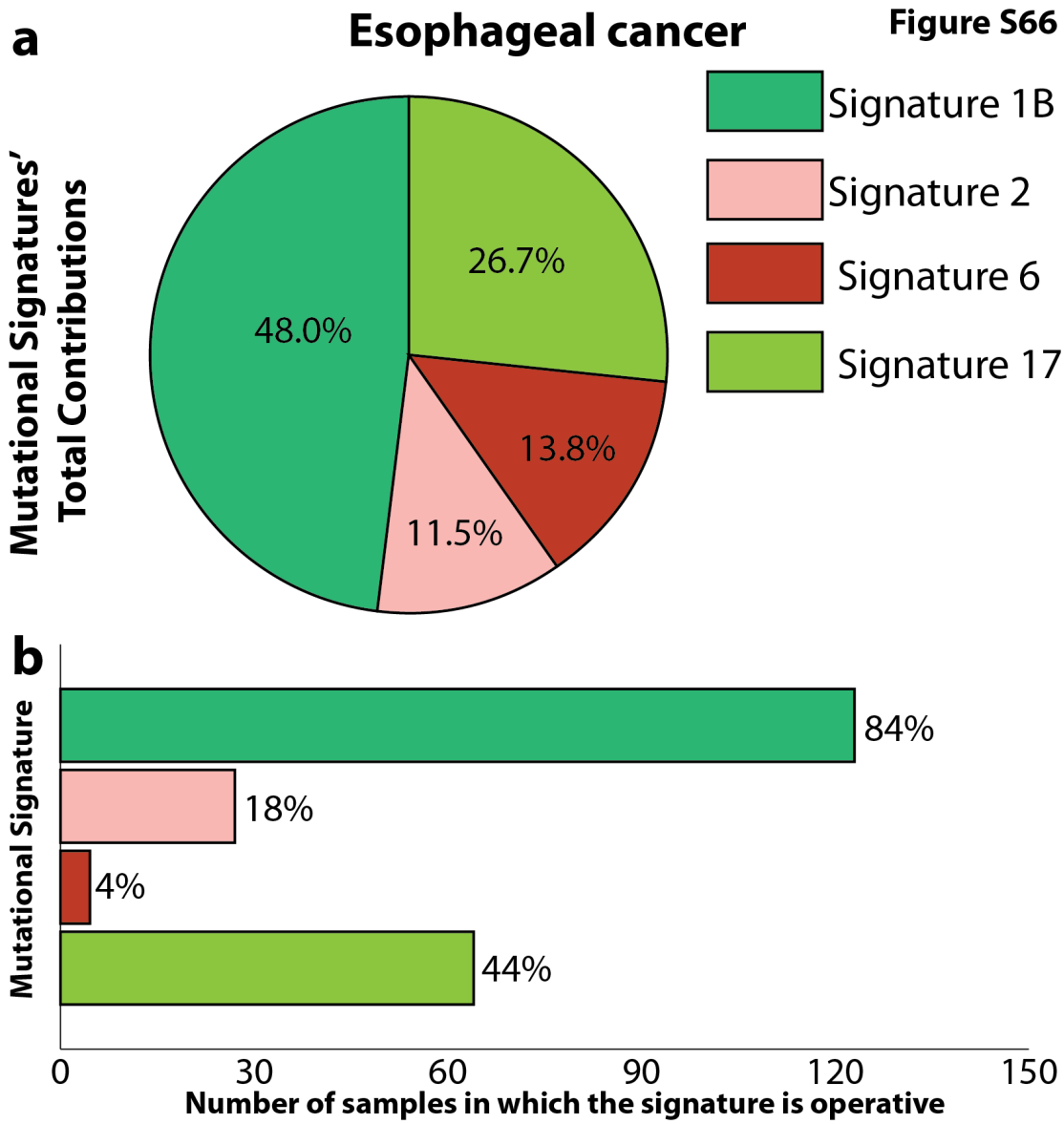
Supplementary Figure 63. Summary of the contributions of the signatures of mutational processes operative in cervical cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



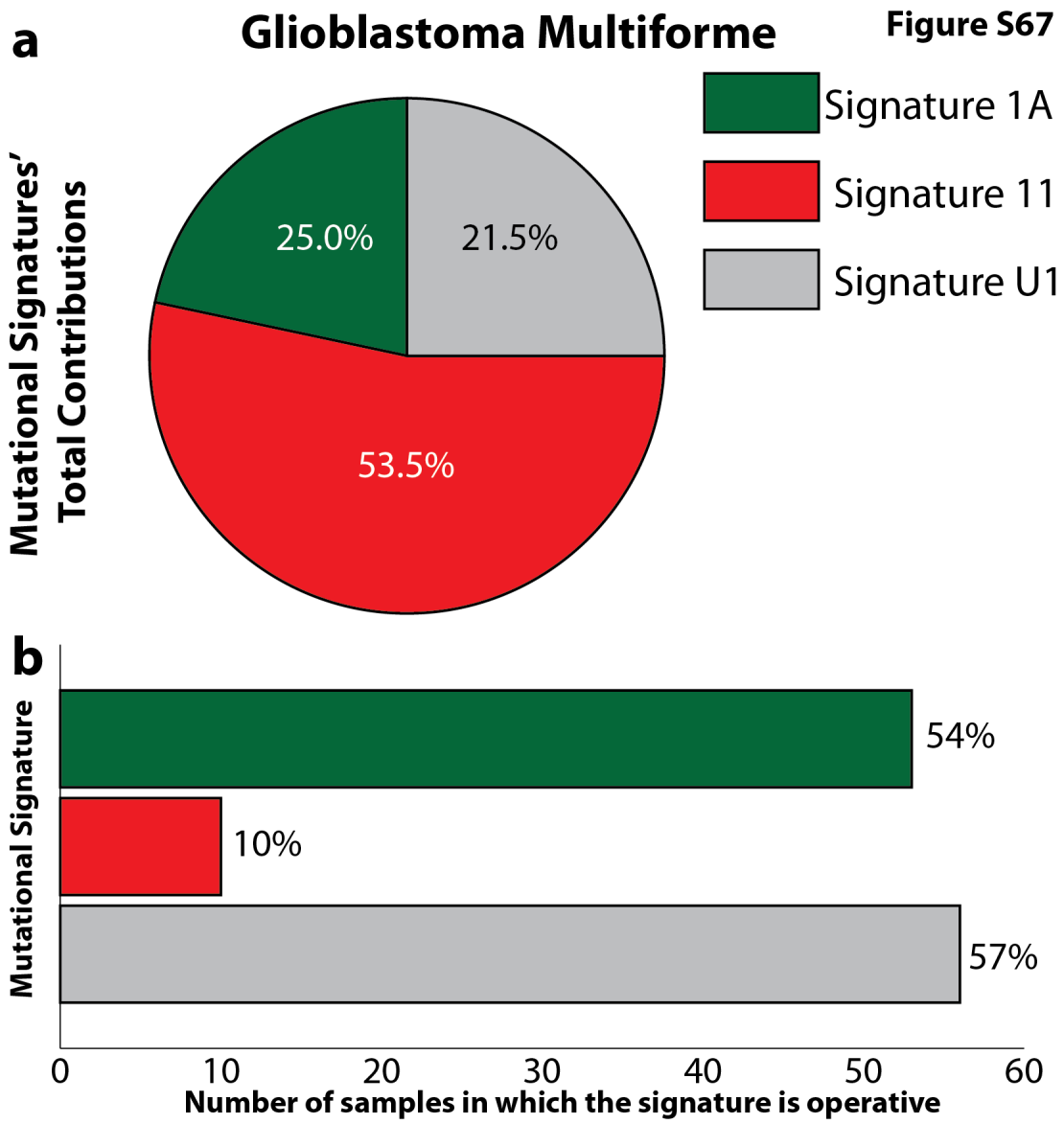
Supplementary Figure 64. Summary of the contributions of the signatures of mutational processes operative in CLL. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



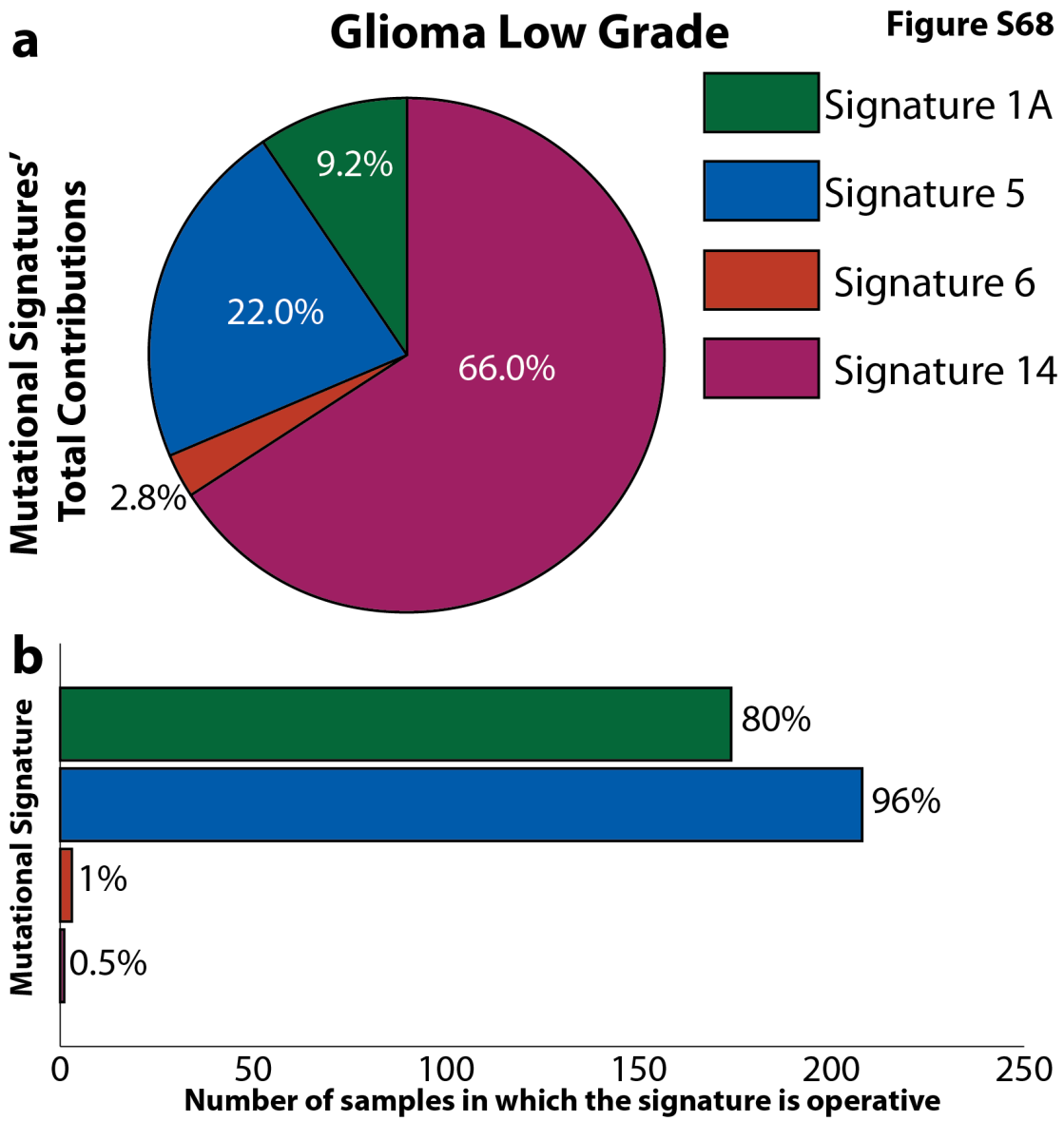
Supplementary Figure 65. Summary of the contributions of the signatures of mutational processes operative in colorectal cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



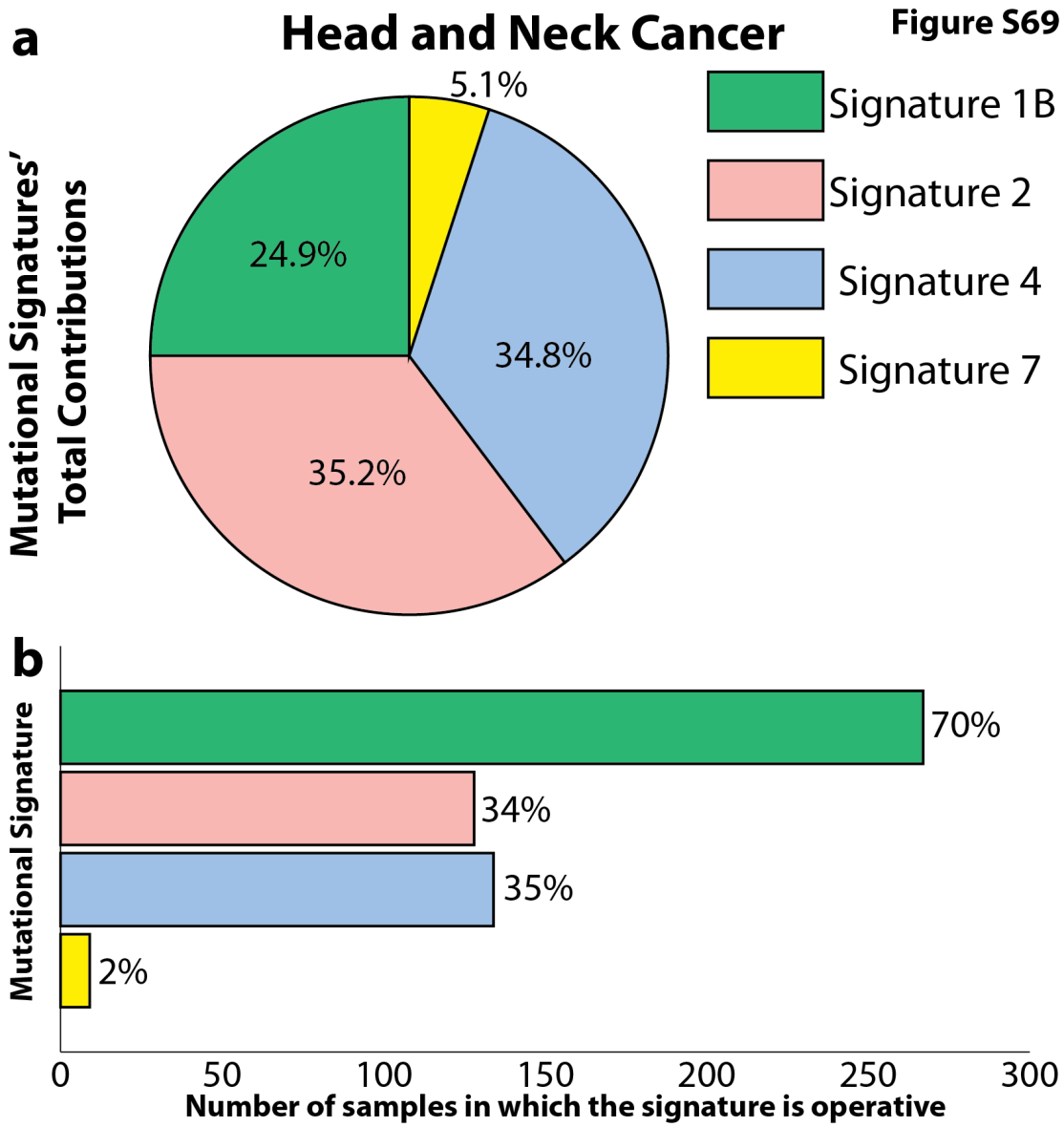
Supplementary Figure 66. Summary of the contributions of the signatures of mutational processes operative in esophageal cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



Supplementary Figure 67. Summary of the contributions of the signatures of mutational processes operative in glioblastoma multiforme. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.

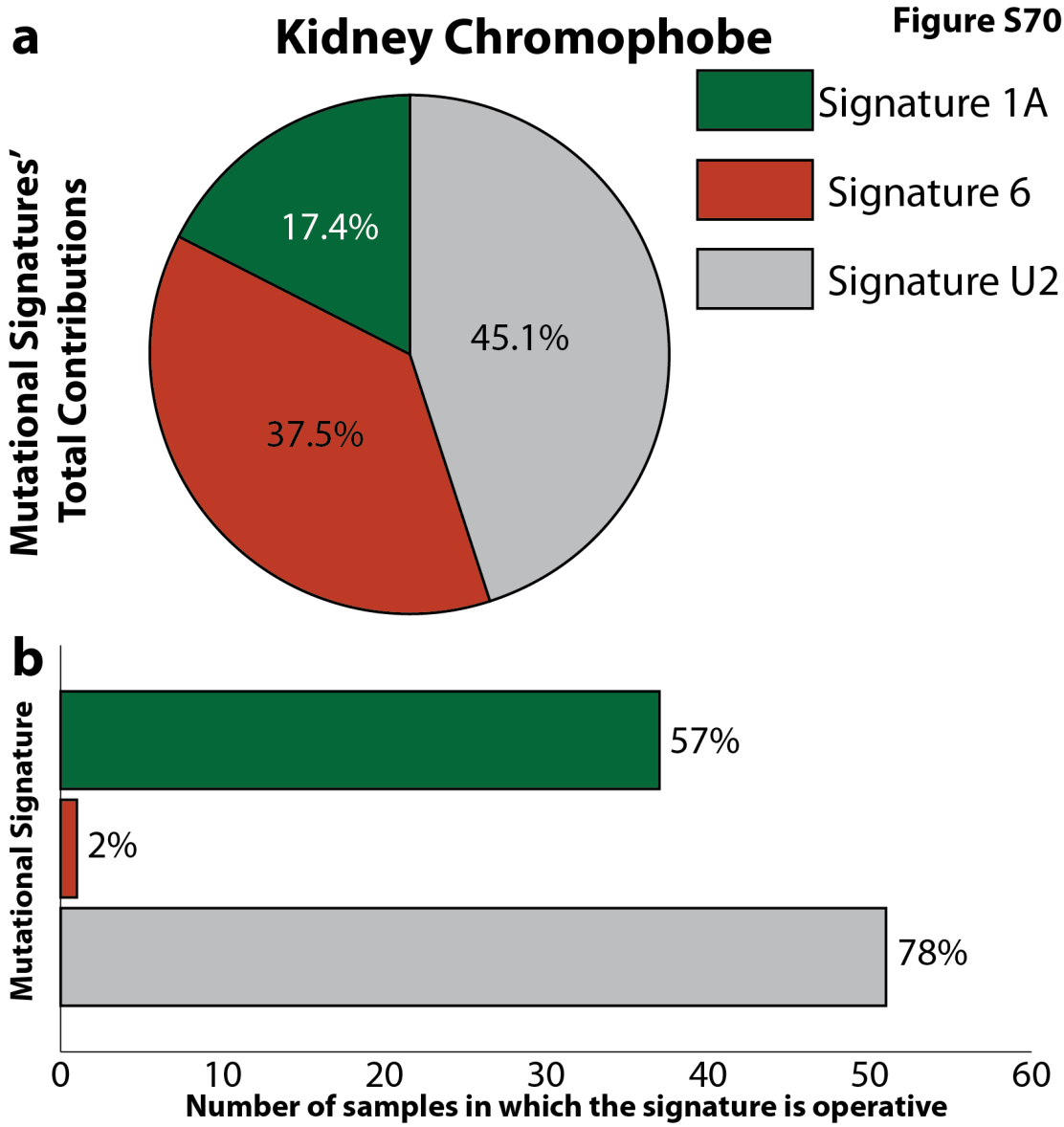


Supplementary Figure 68. Summary of the contributions of the signatures of mutational processes operative in glioma low grade. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.

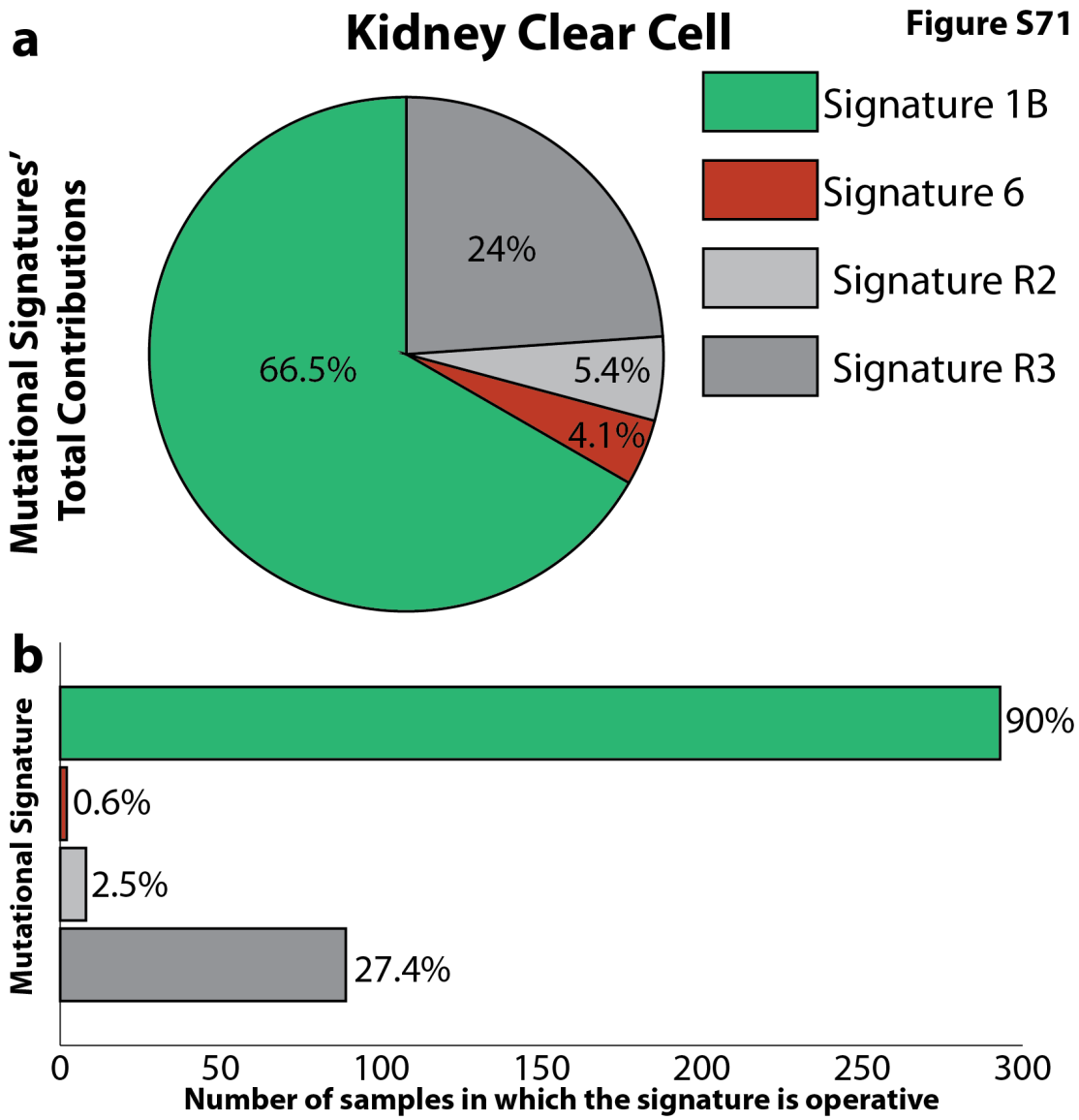


Supplementary Figure 69. Summary of the contributions of the signatures of mutational processes operative in head and neck cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.

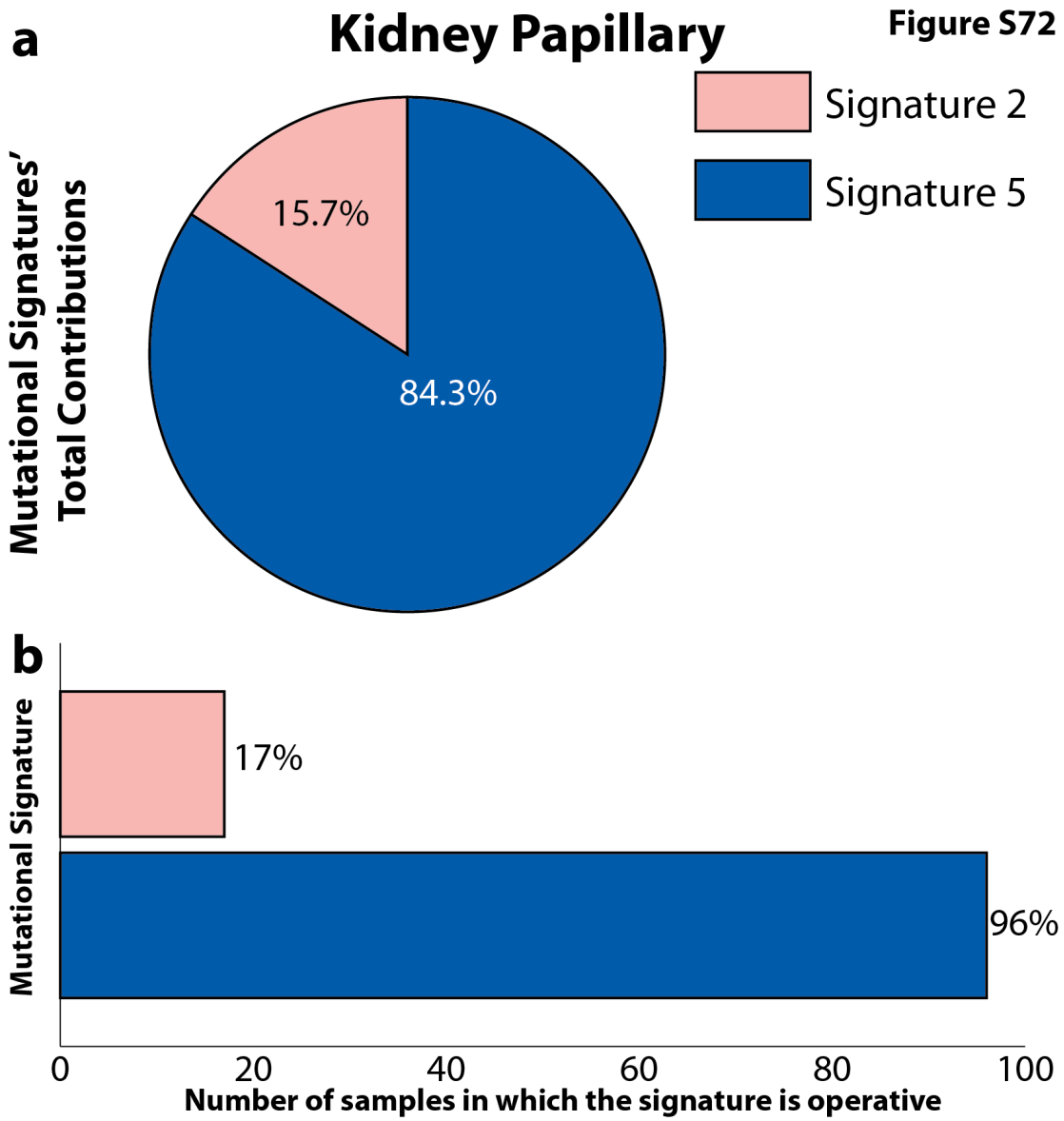
Figure S70



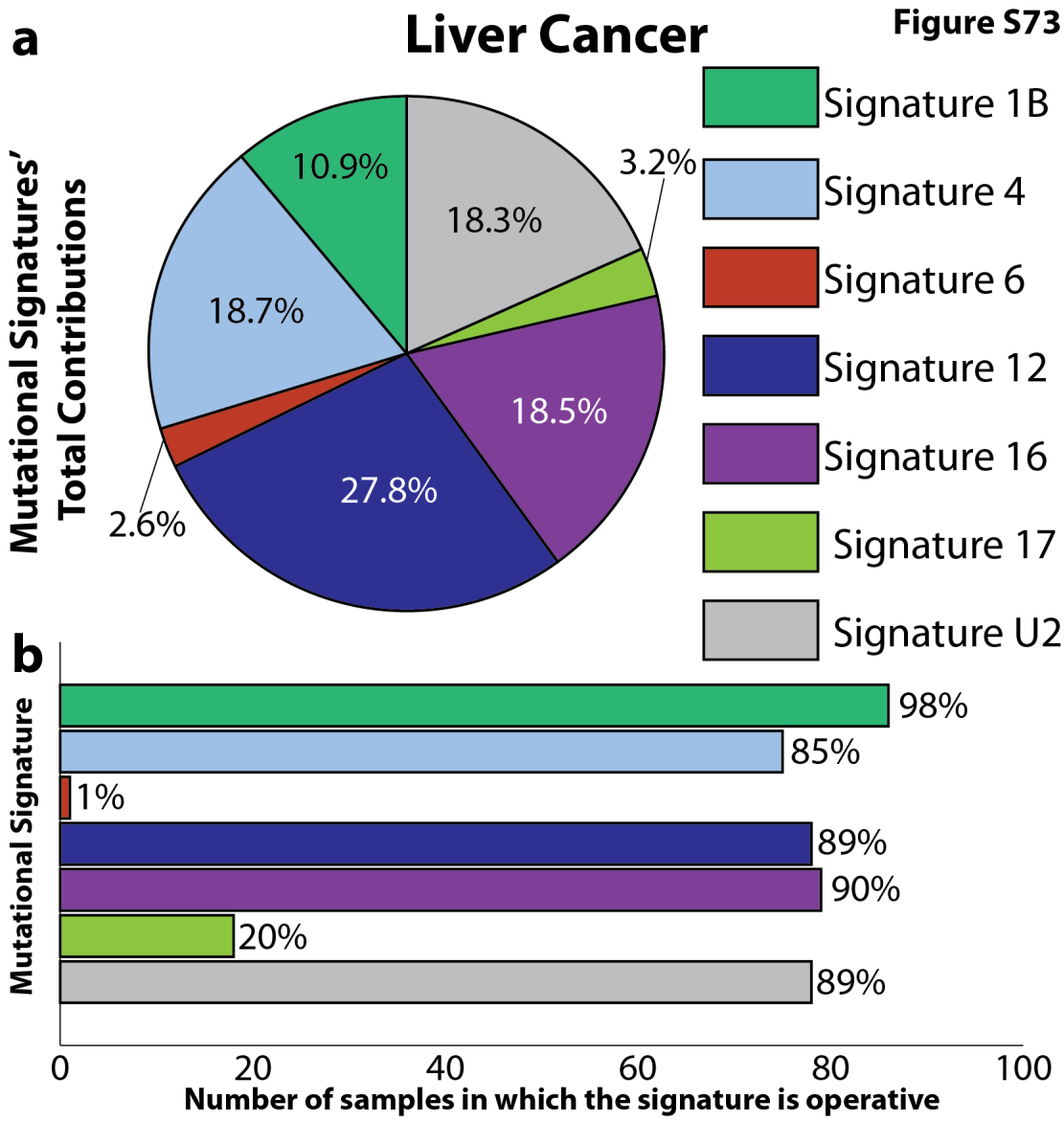
Supplementary Figure 70. Summary of the contributions of the signatures of mutational processes operative in kidney chromophobe. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



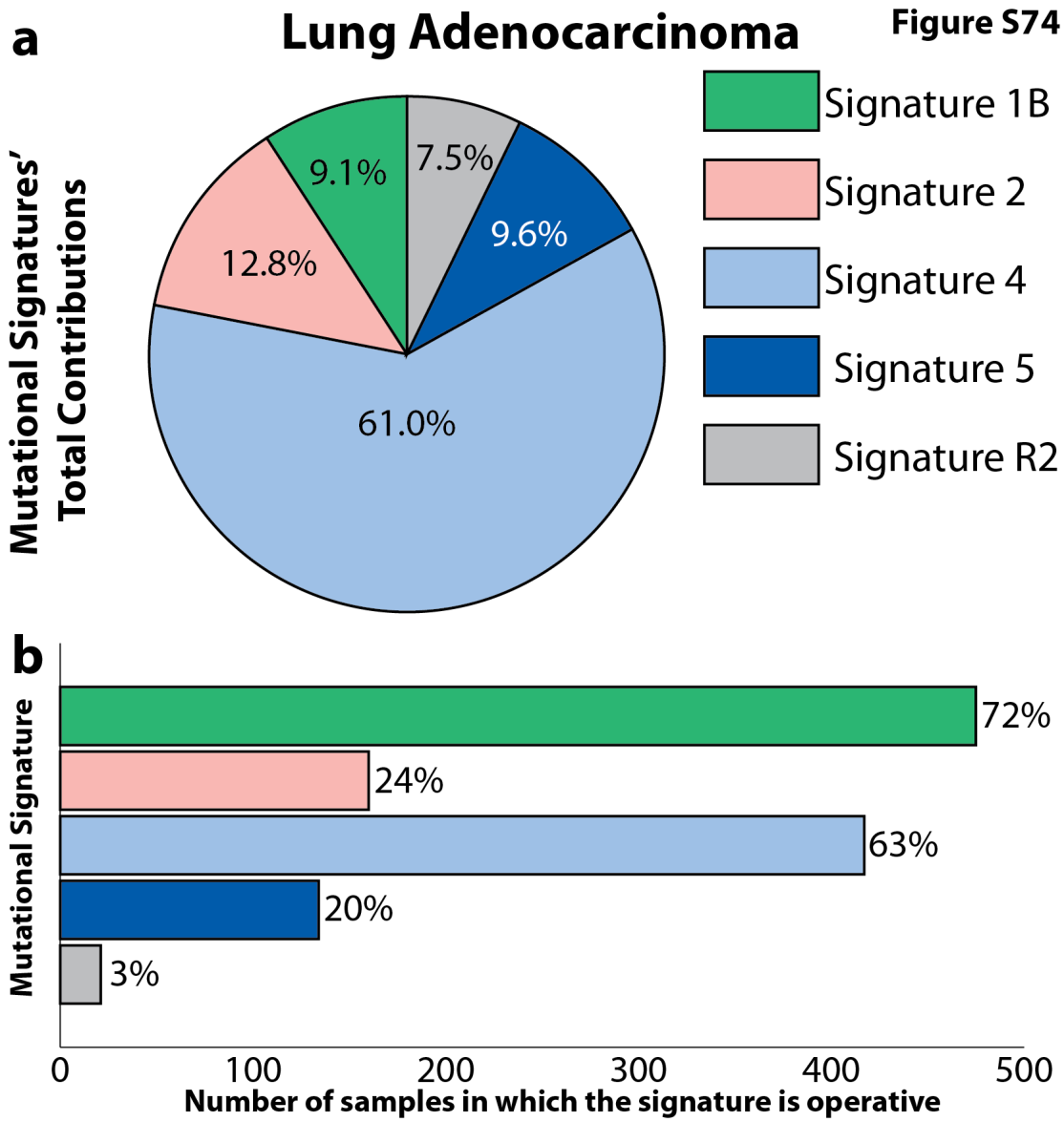
Supplementary Figure 71. Summary of the contributions of the signatures of mutational processes operative in kidney clear cell. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



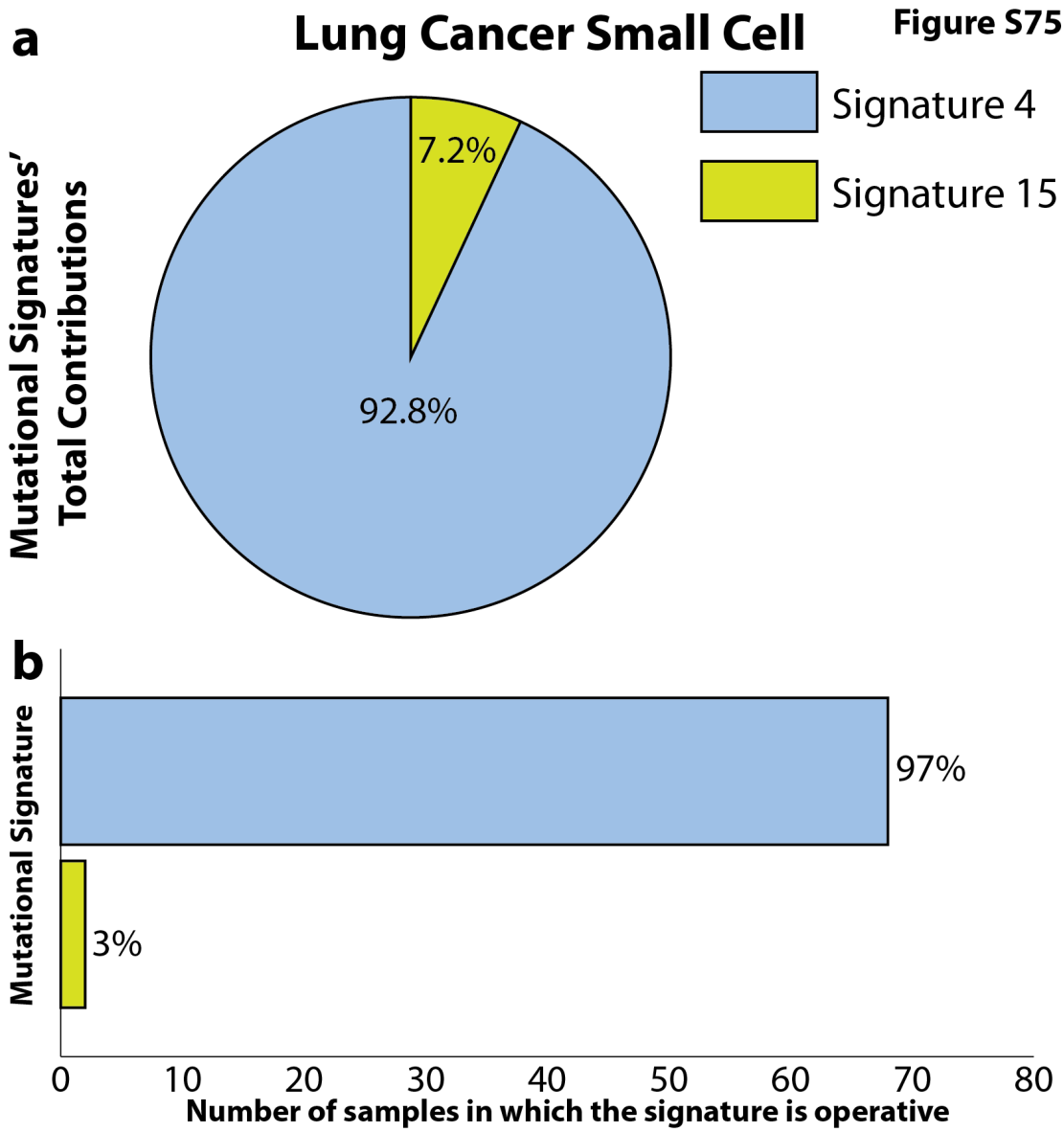
Supplementary Figure 72. Summary of the contributions of the signatures of mutational processes operative in kidney papillary. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



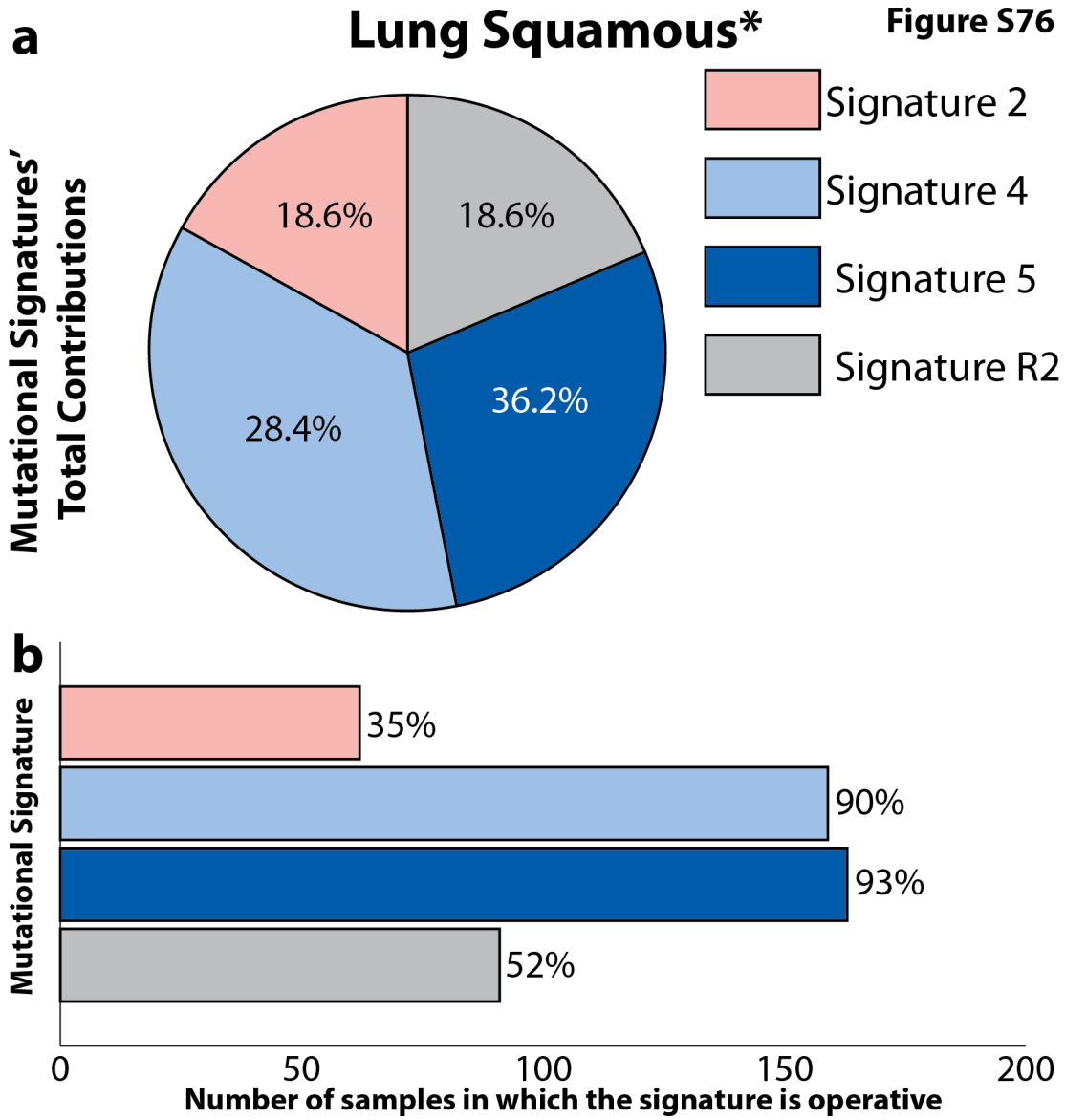
Supplementary Figure 73. Summary of the contributions of the signatures of mutational processes operative in liver cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



Supplementary Figure 74. Summary of the contributions of the signatures of mutational processes operative in lung adenocarcinoma. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.

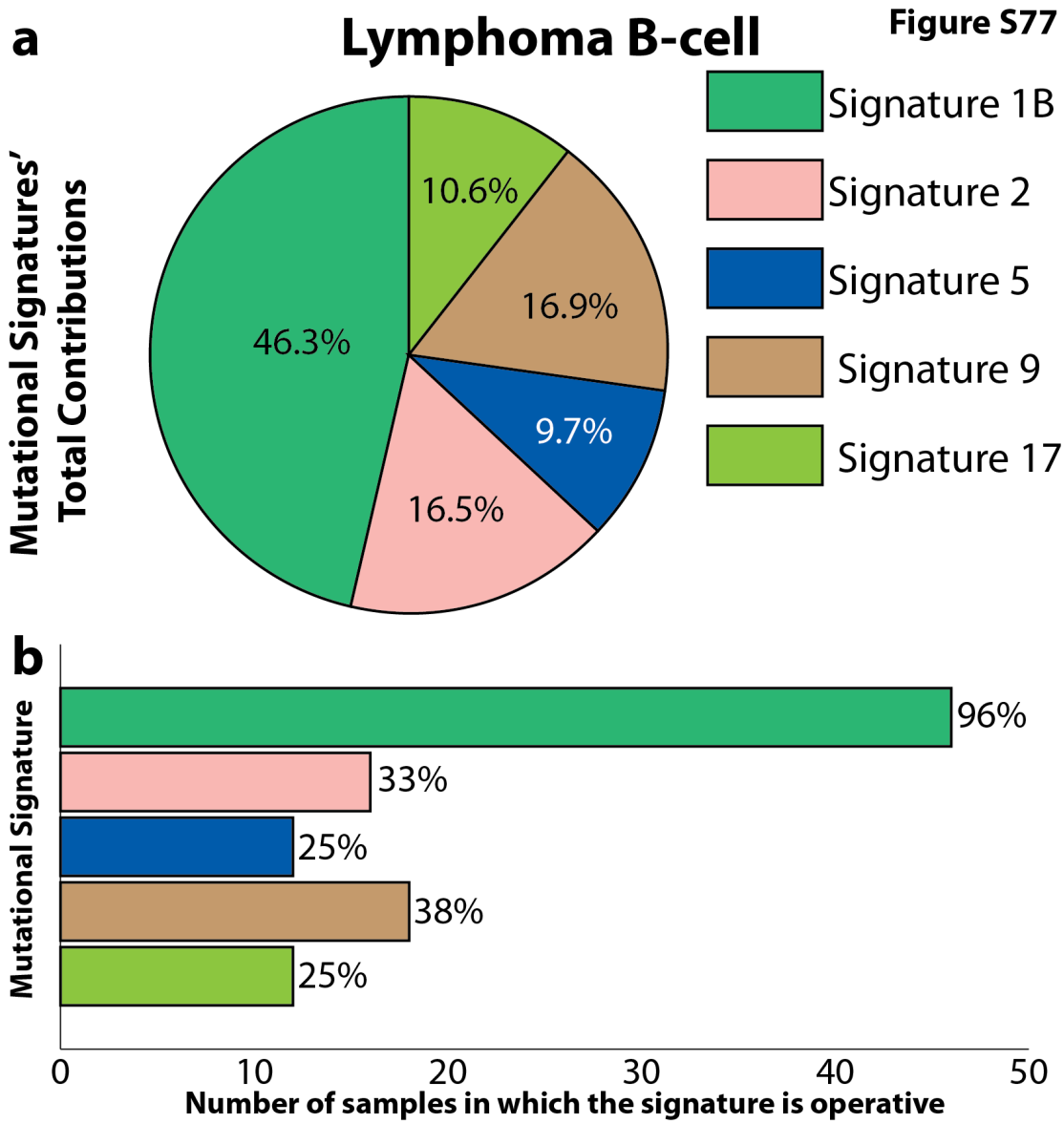


Supplementary Figure 75. Summary of the contributions of the signatures of mutational processes operative in lung cancer small cell. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.

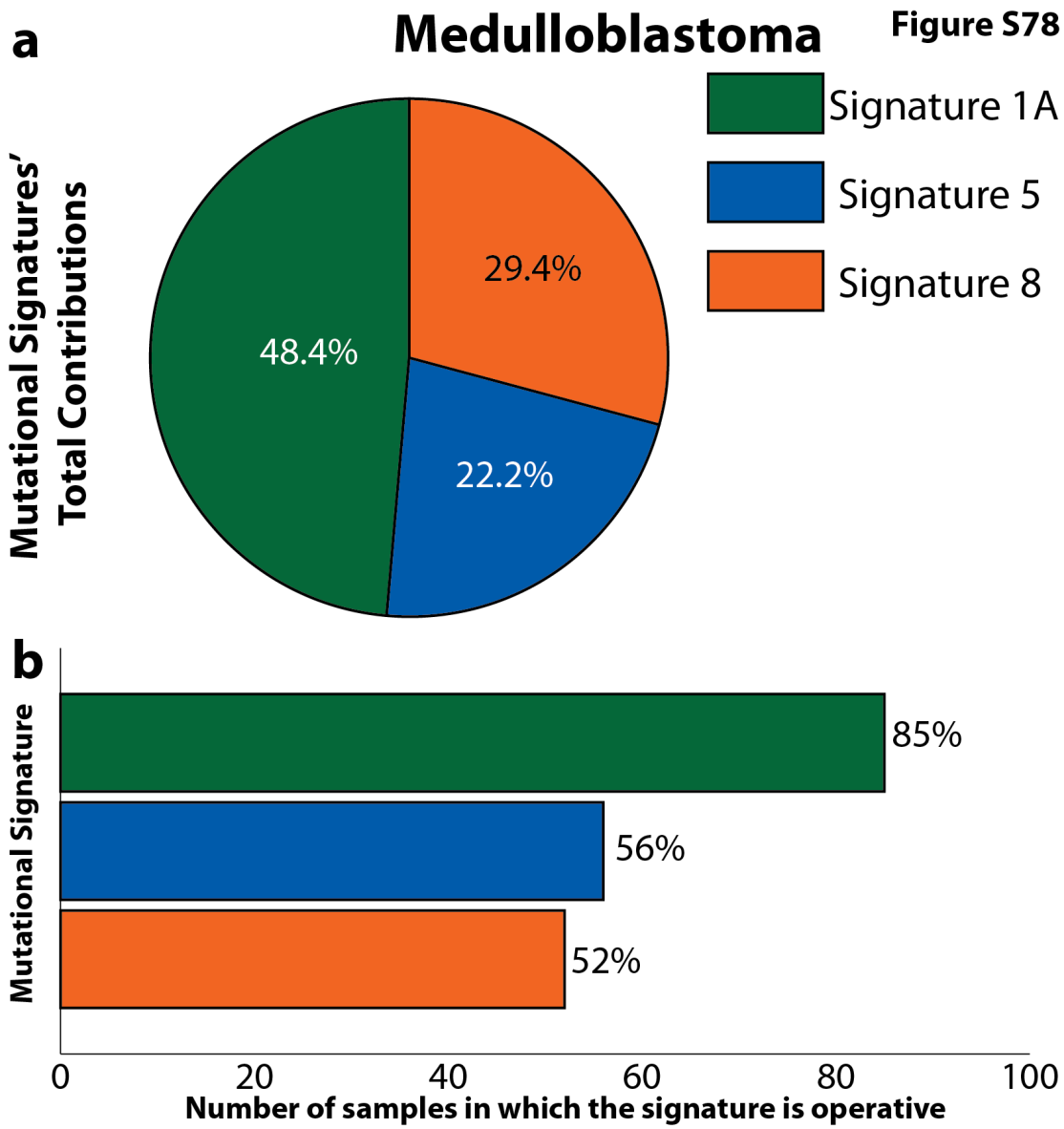


Supplementary Figure 76. Summary of the contributions of the signatures of mutational processes operative in lung squamous. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures. *One hypermutator sample purely of Signature 7 (122 mutations per MB) is not included in the

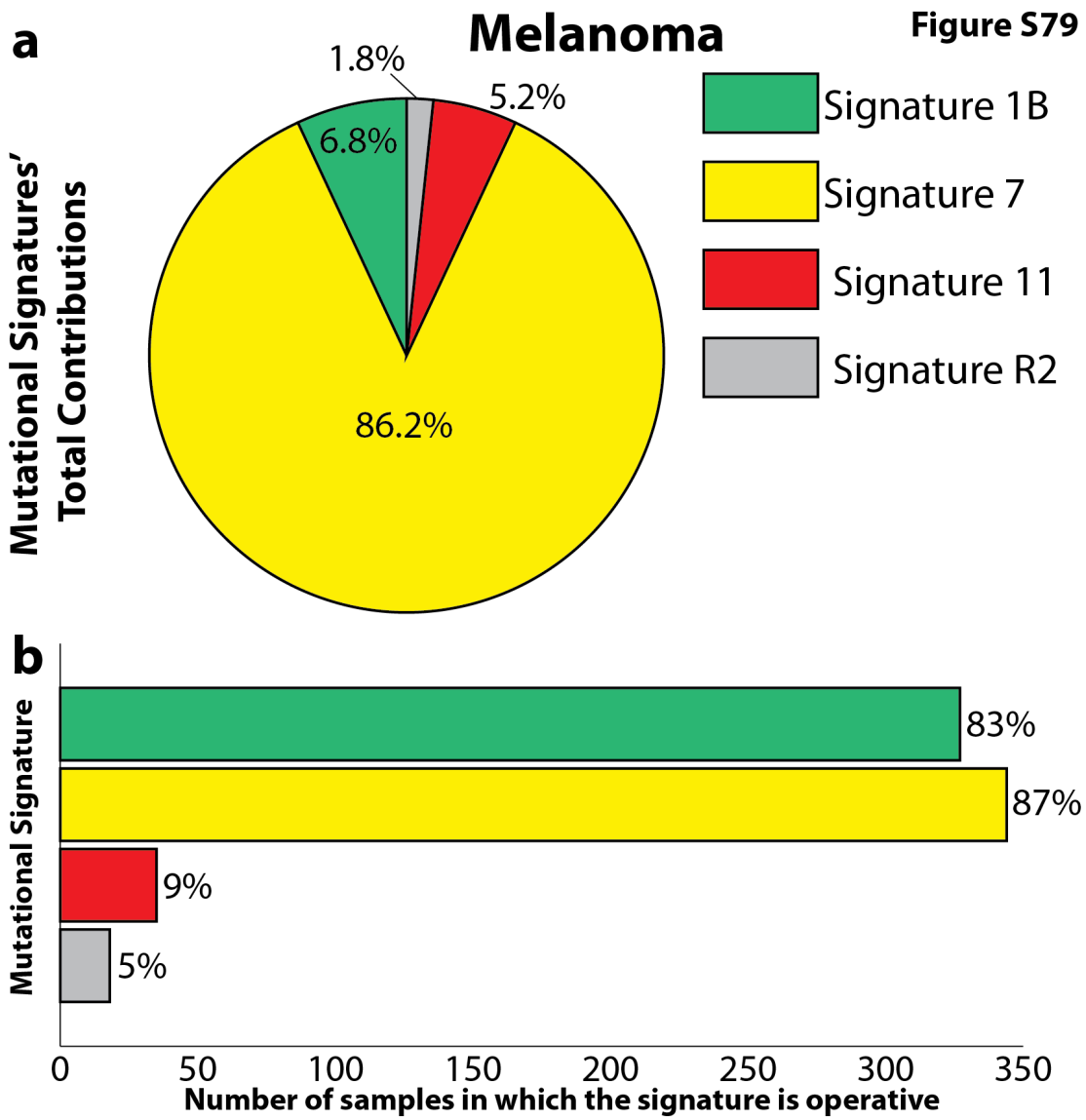
summary data. Signature 7 is associated with UV-light, an unlikely carcinogen for lung cancer. As such, we believe that this lung sample is most likely either a melanoma metastasis or a mis-annotated sample. Thus, the association between Signature 7 and lung squamous has not been displayed in Figure 3.



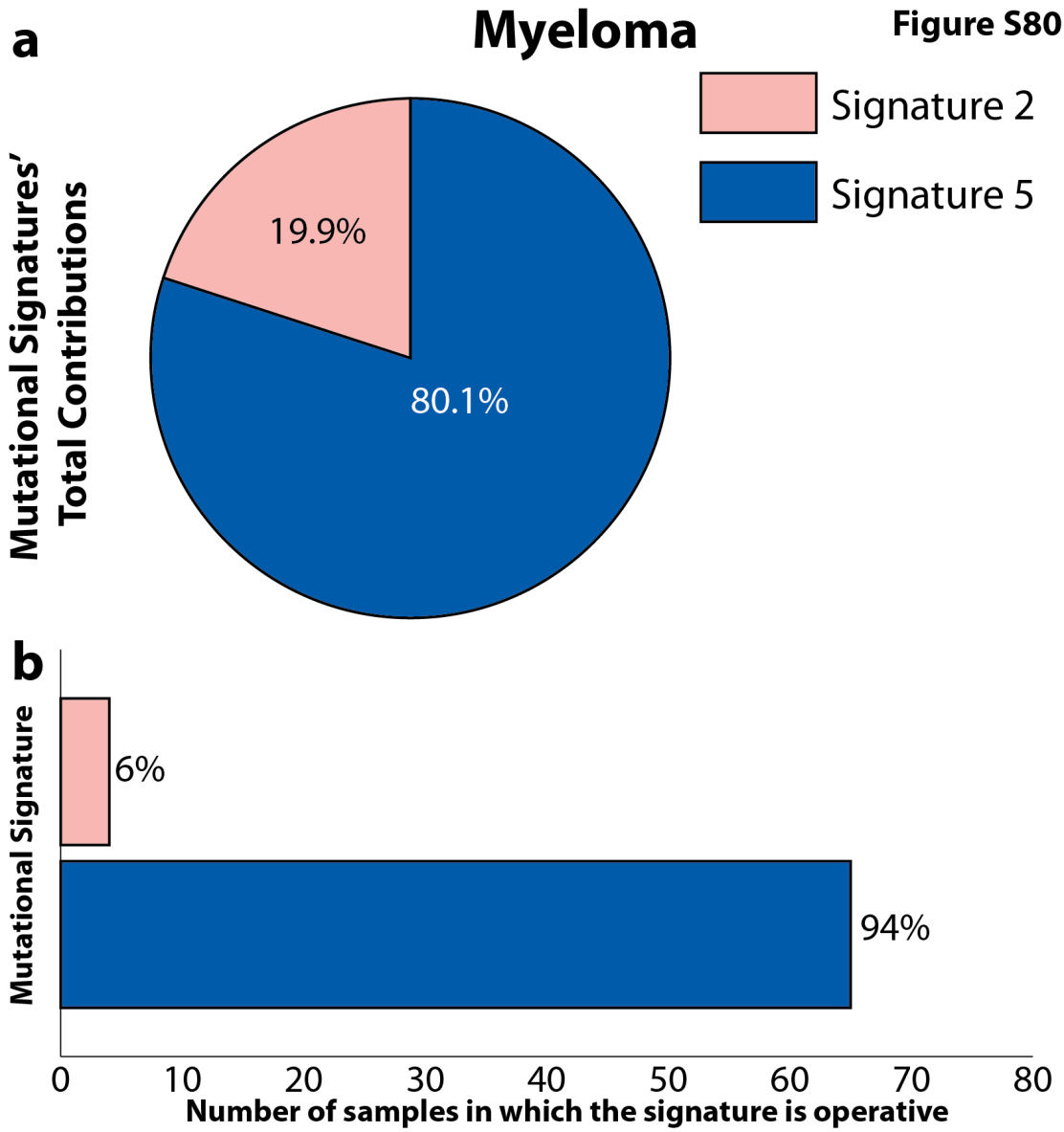
Supplementary Figure 77. Summary of the contributions of the signatures of mutational processes operative in lymphoma B-cell. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



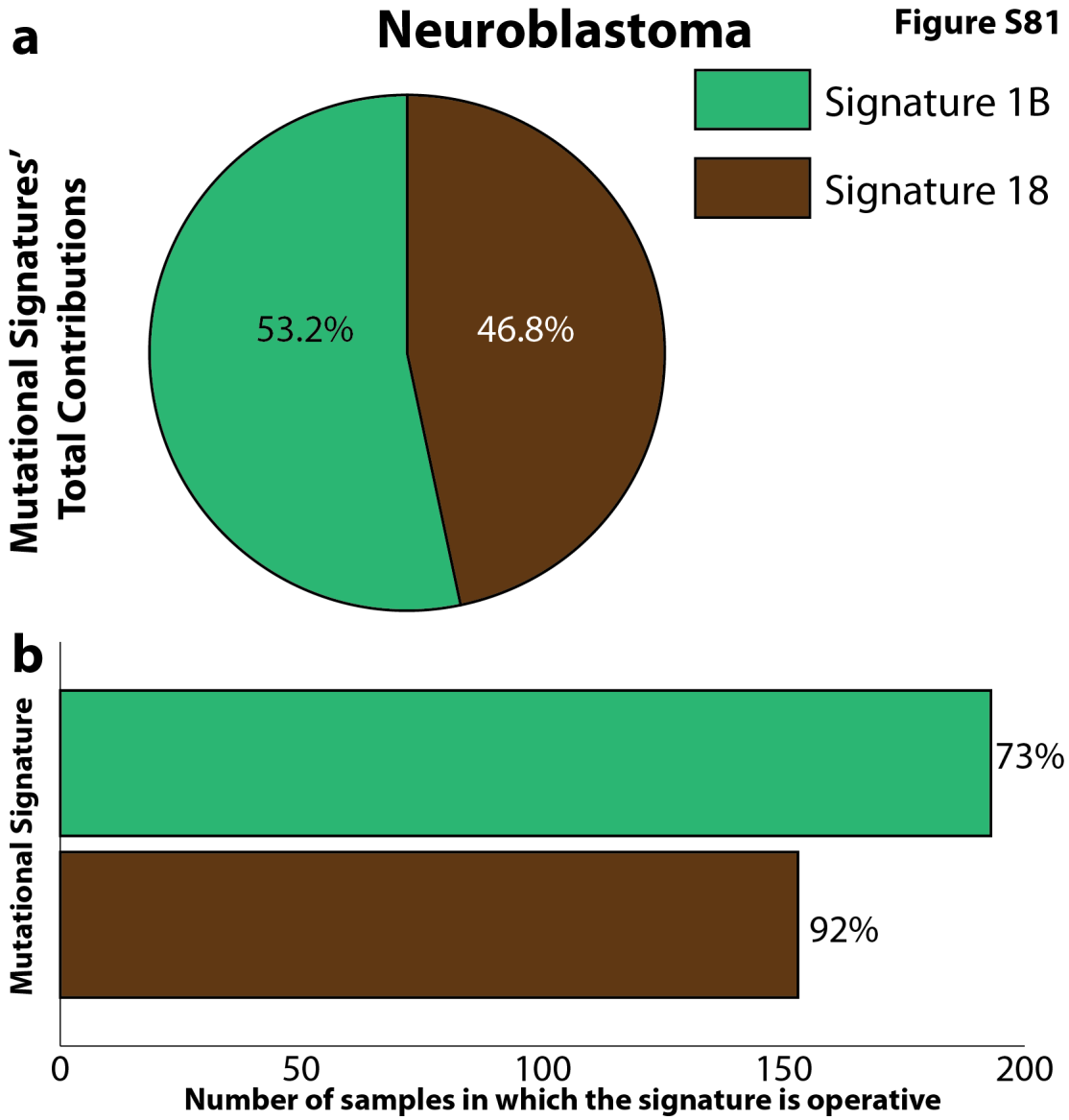
Supplementary Figure 78. Summary of the contributions of the signatures of mutational processes operative in medulloblastoma. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



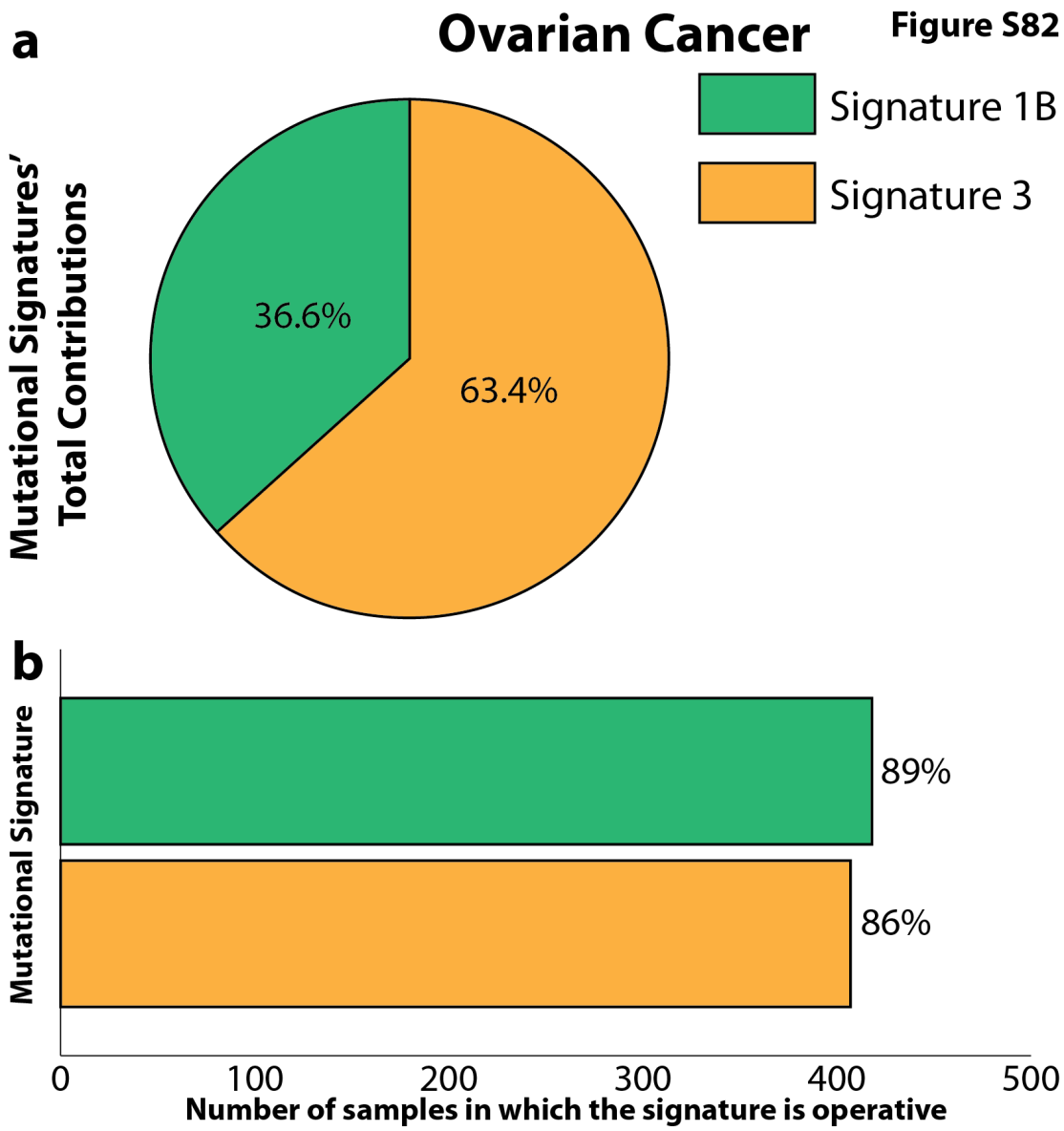
Supplementary Figure 79. Summary of the contributions of the signatures of mutational processes operative in melanoma. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



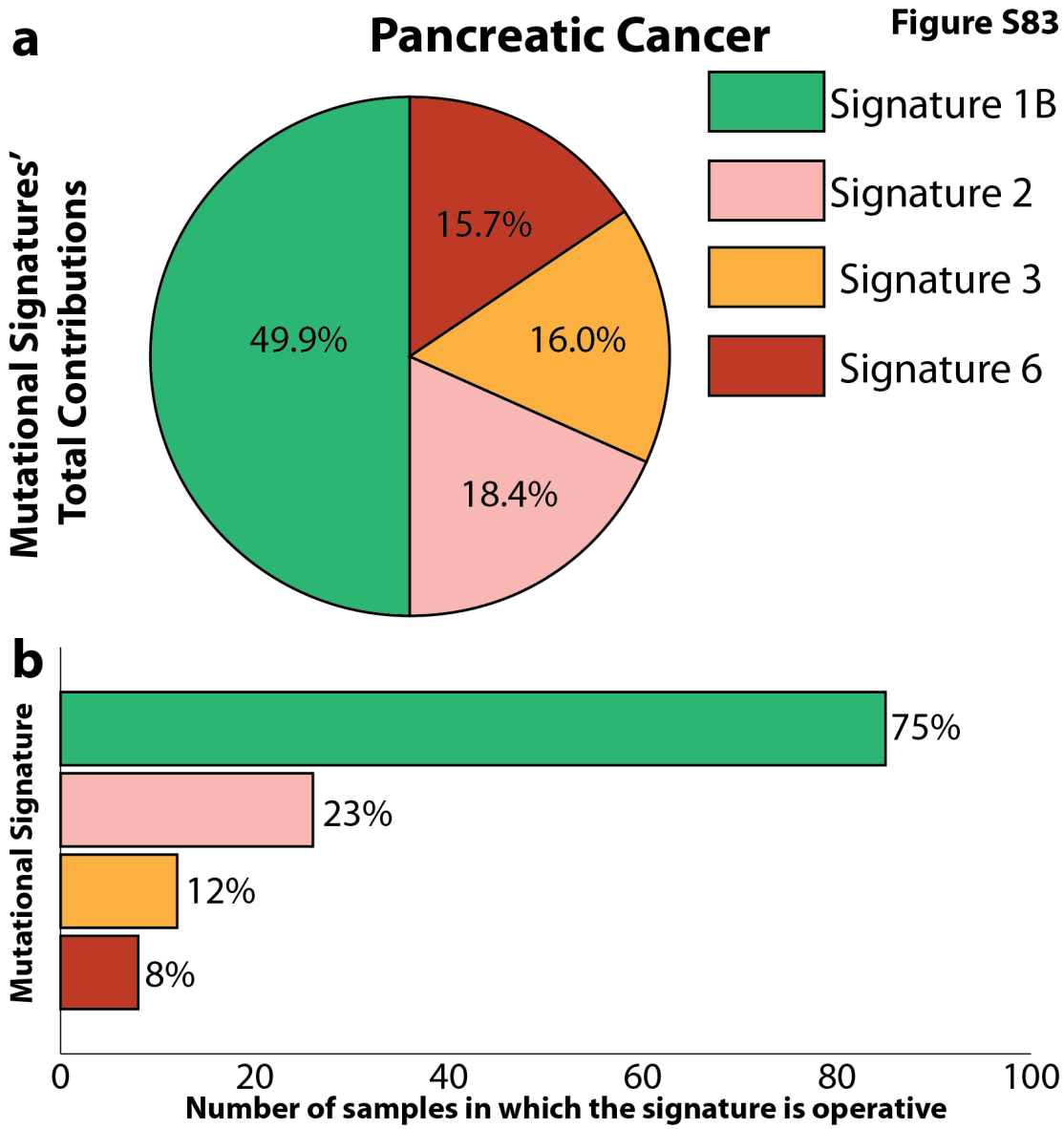
Supplementary Figure 80. Summary of the contributions of the signatures of mutational processes operative in myeloma. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



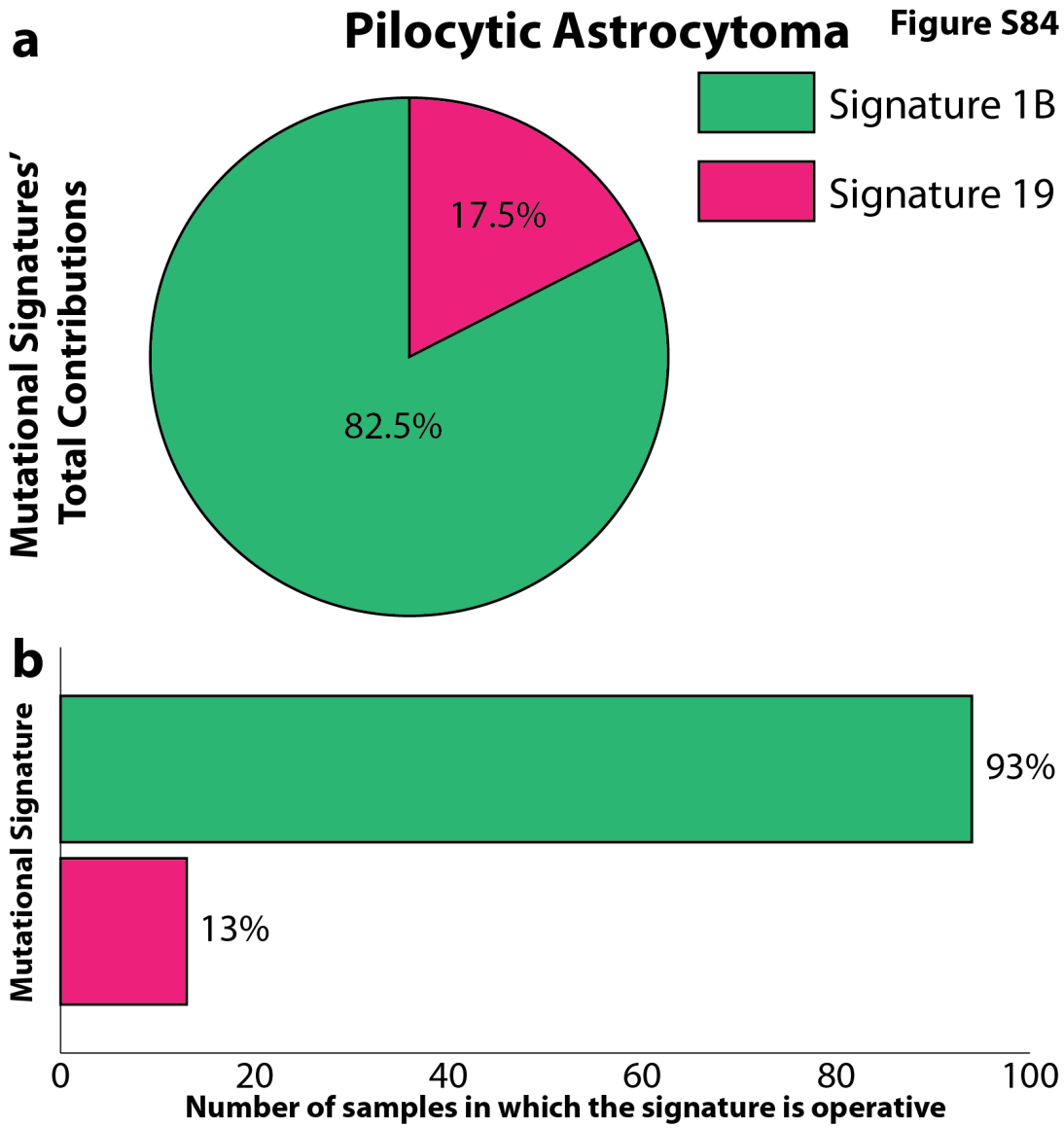
Supplementary Figure 81. Summary of the contributions of the signatures of mutational processes operative in neuroblastoma. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



Supplementary Figure 82. Summary of the contributions of the signatures of mutational processes operative in ovarian cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.

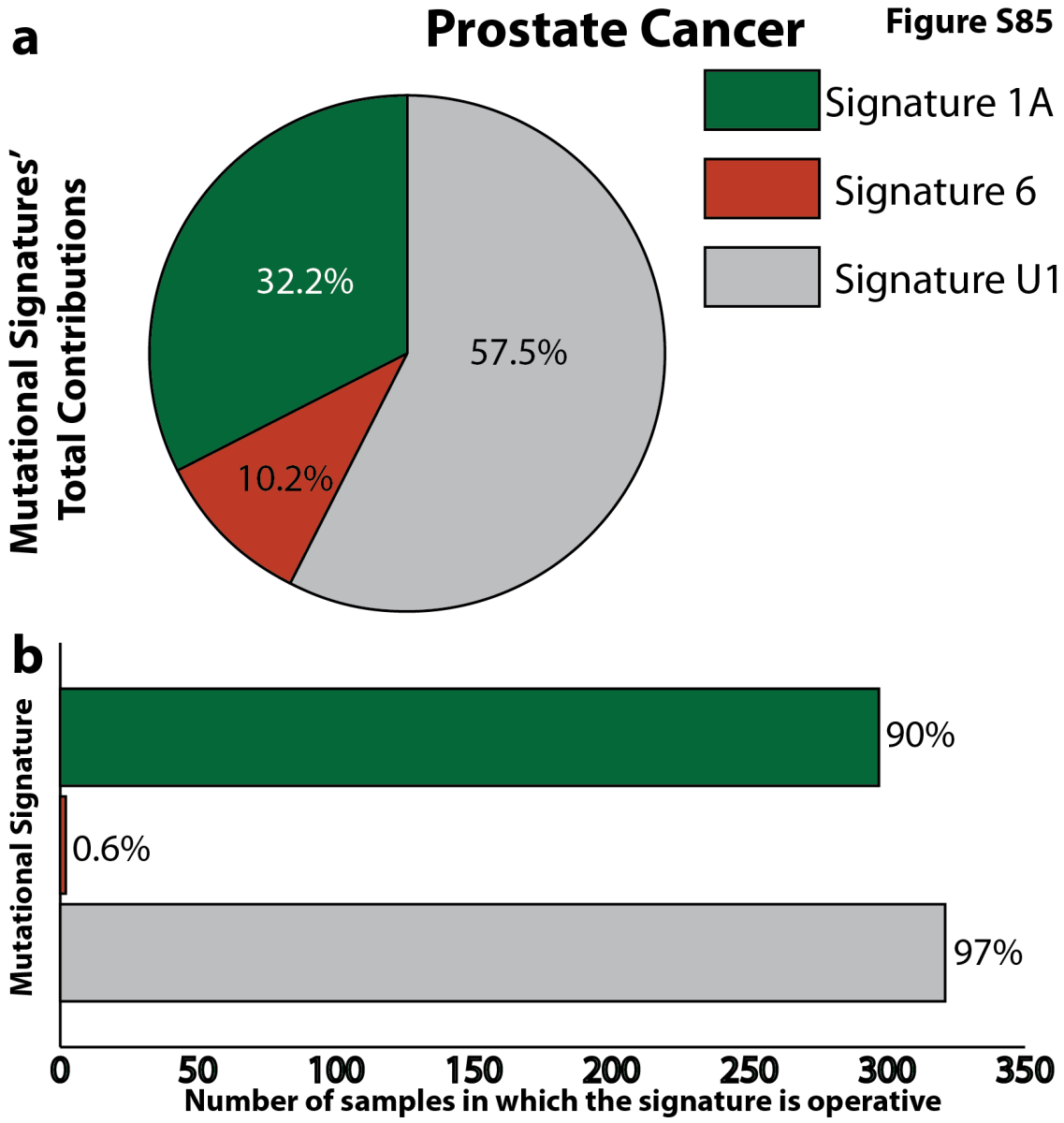


Supplementary Figure 83. Summary of the contributions of the signatures of mutational processes operative in pancreatic cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.

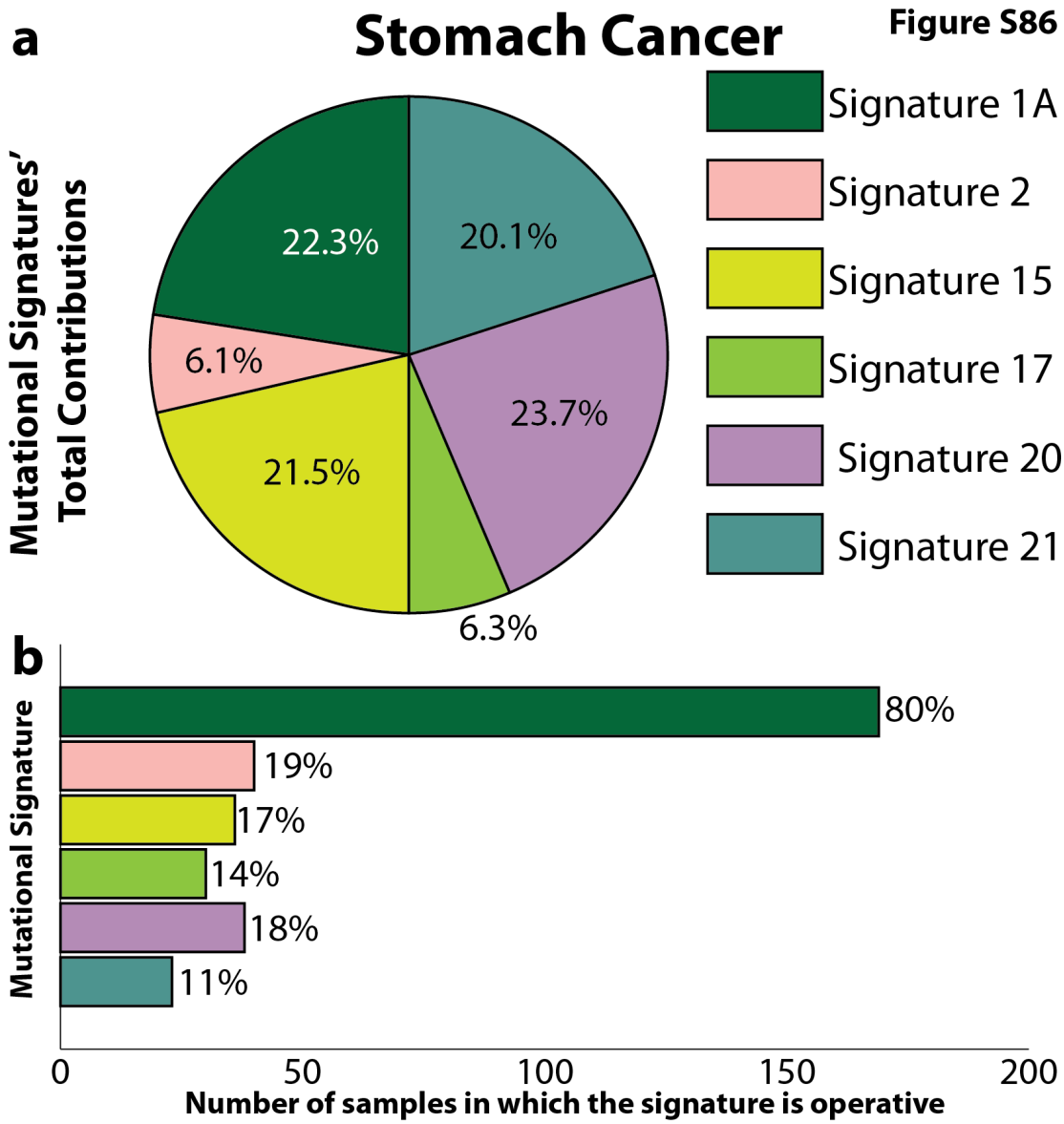


Supplementary Figure 84. Summary of the contributions of the signatures of mutational processes operative in pilocytic astrocytoma. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures. Note that this dataset is predominantly composed of pilocytic astrocytomas but also

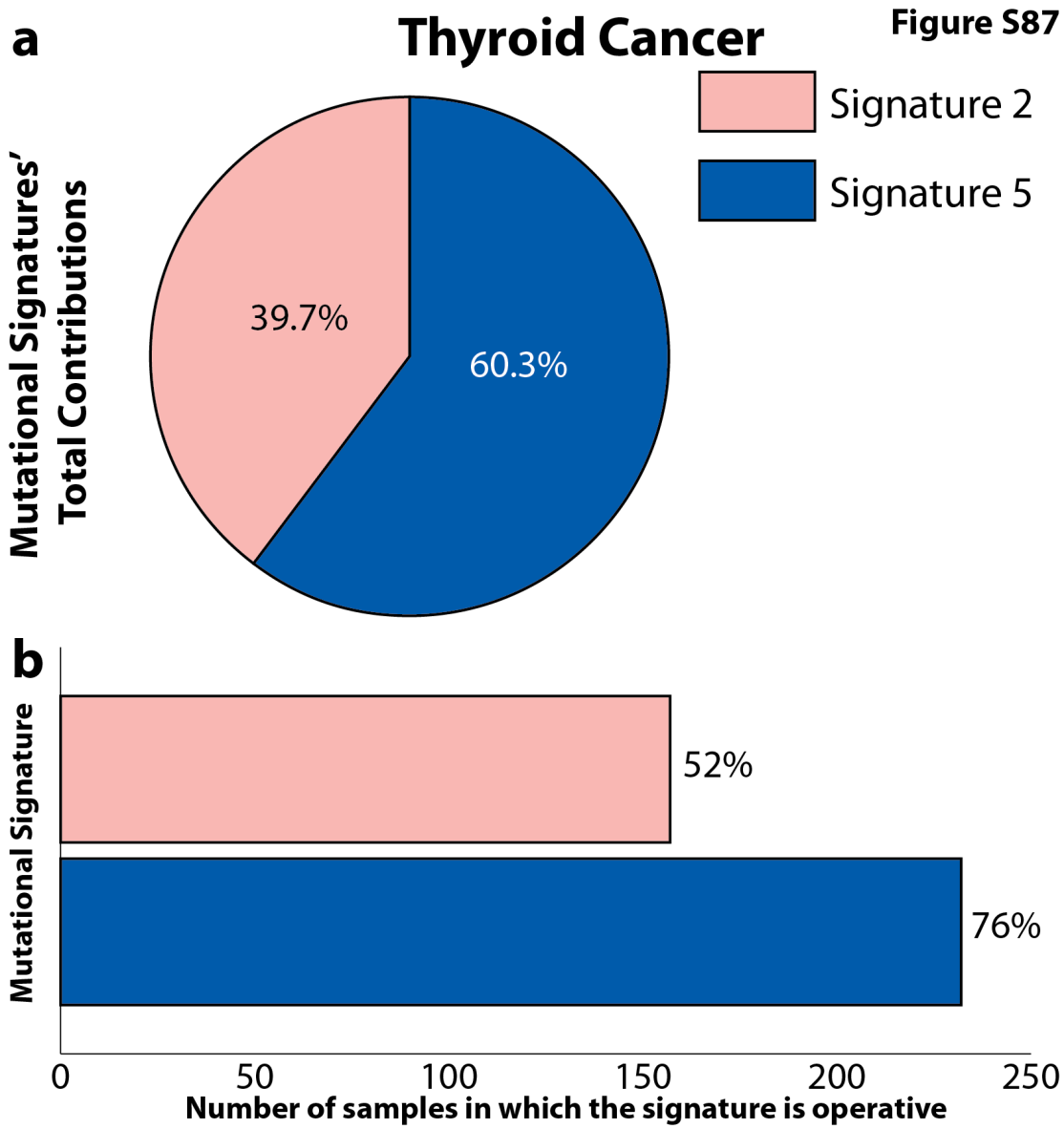
includes a small number of other pediatric low-grade gliomas and pediatric low-grade glioneuronal tumors.



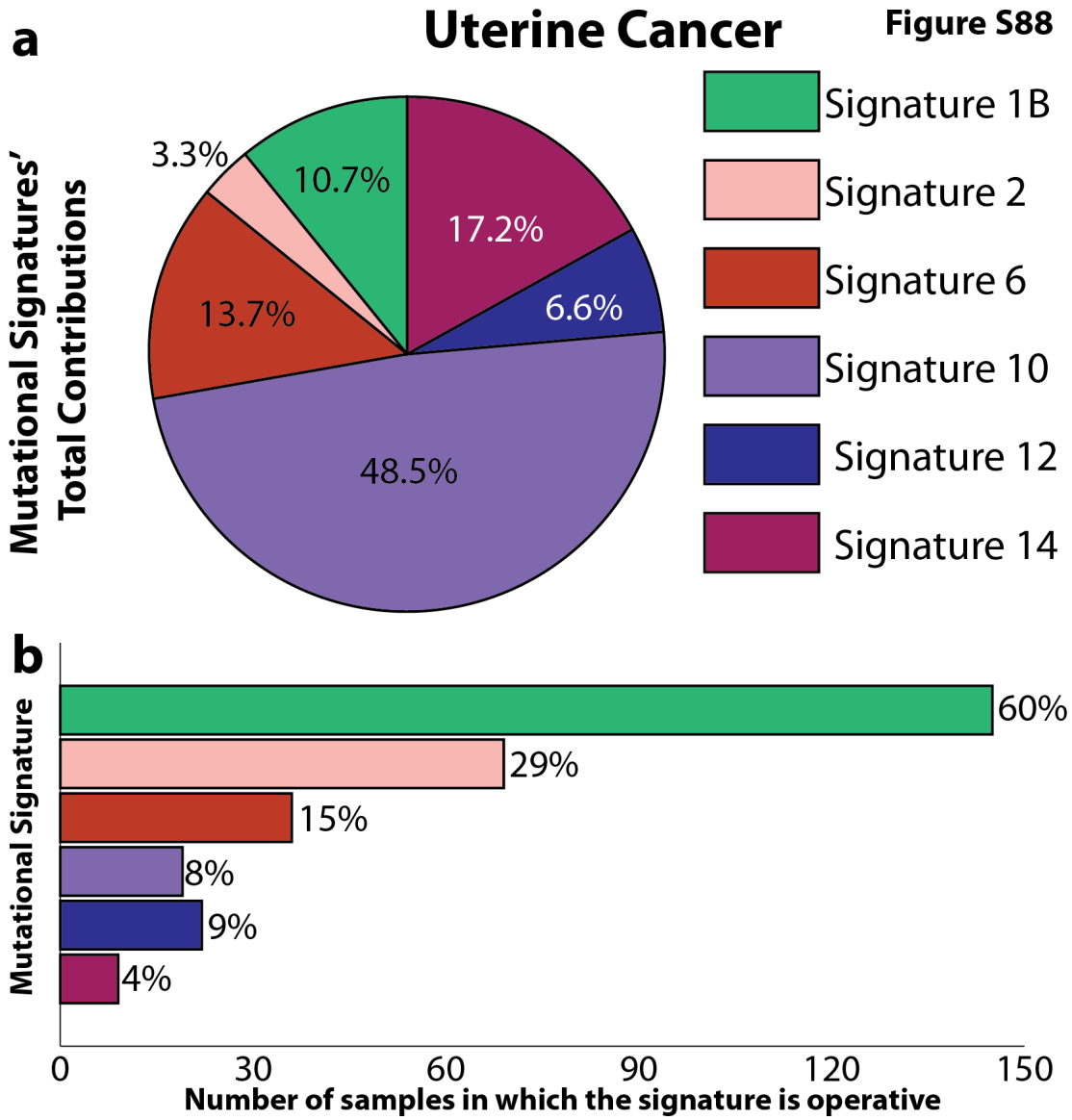
Supplementary Figure 85. Summary of the contributions of the signatures of mutational processes operative in prostate cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



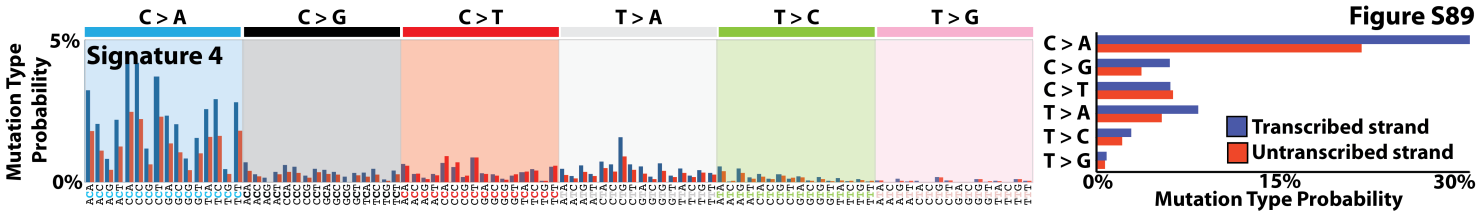
Supplementary Figure 86. Summary of the contributions of the signatures of mutational processes operative in stomach cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



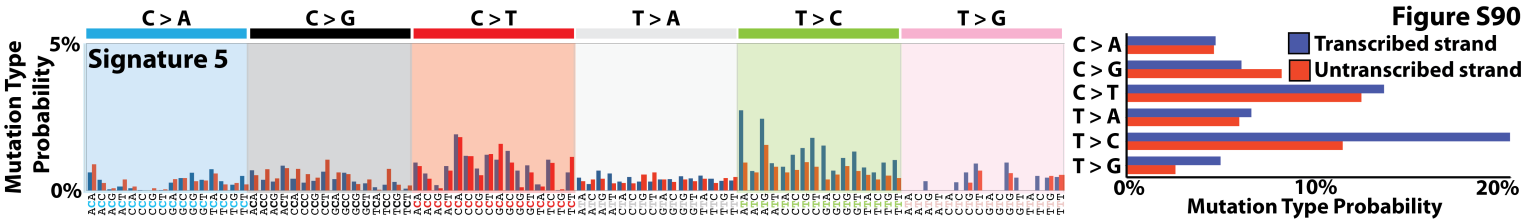
Supplementary Figure 87. Summary of the contributions of the signatures of mutational processes operative in thyroid cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



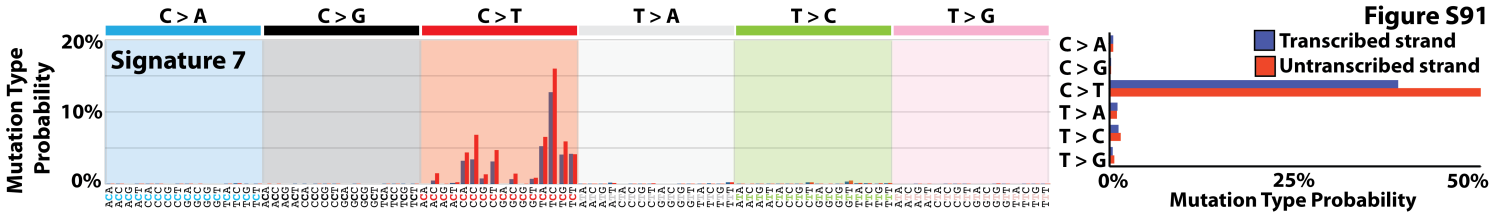
Supplementary Figure 88. Summary of the contributions of the signatures of mutational processes operative in uterine cancer. ***(a)*** Percentage of total mutations contributed by each of the operative mutational signatures. ***(b)*** Percentage and number of samples in which each mutational signature contributes significant number of somatic mutations. For most signatures, significant number of mutations in a sample is defined as more than 100 substitutions or more than 25% of all mutations in that sample. Mutational signatures are displayed in distinct colors, consistent in both panels and all other figures.



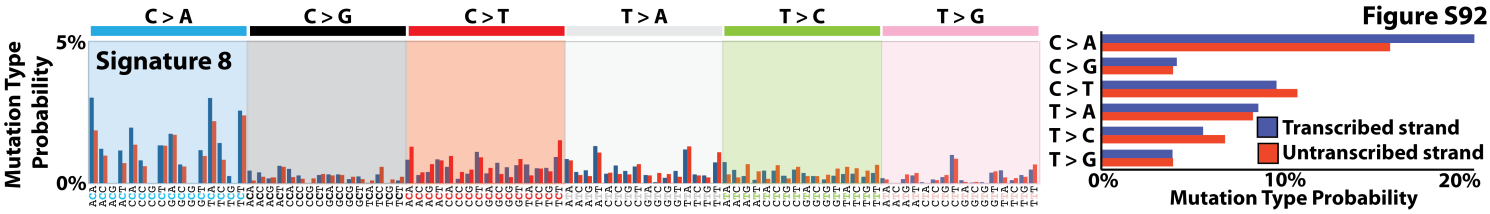
Supplementary Figure 89. Transcriptional strand bias for Signature 4. Mutational signature is displayed using a 192 substitution classification incorporating the substitution type, the sequence context immediately 5' and 3' to the mutated base and whether the mutated base (in pyrimidine context) is on the transcribed or untranscribed strand. The panels for each of the six types of substitutions as well as the mutated base are displayed in different colors. Mutations on the transcribed pyrimidine strand are displayed in blue while mutations on the untranscribed strand are displayed in red.



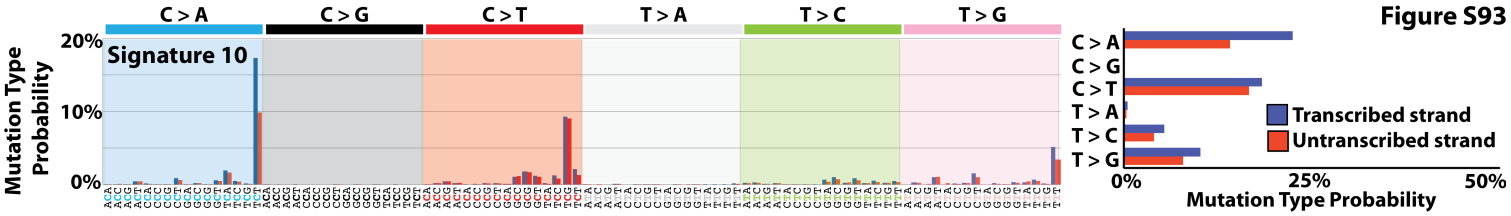
Supplementary Figure 90. Transcriptional strand bias for Signature 5. Mutational signature is displayed using a 192 substitution classification incorporating the substitution type, the sequence context immediately 5' and 3' to the mutated base and whether the mutated base (in pyrimidine context) is on the transcribed or untranscribed strand. The panels for each of the six types of substitutions as well as the mutated base are displayed in different colors. Mutations on the transcribed pyrimidine strand are displayed in blue while mutations on the untranscribed strand are displayed in red.



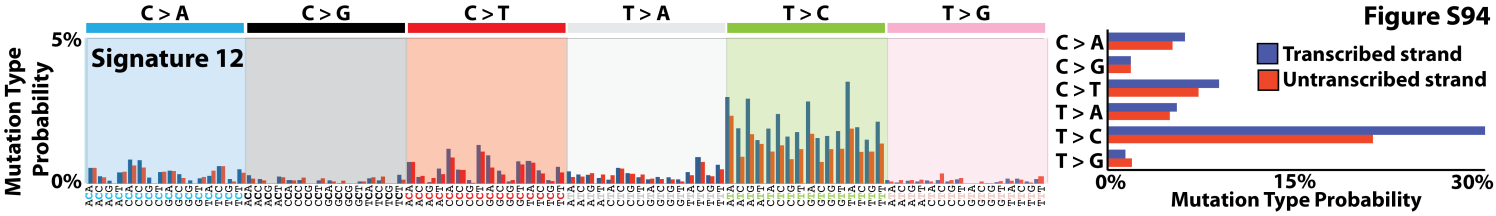
Supplementary Figure 91. Transcriptional strand bias for Signature 7. Mutational signature is displayed using a 192 substitution classification incorporating the substitution type, the sequence context immediately 5' and 3' to the mutated base and whether the mutated base (in pyrimidine context) is on the transcribed or untranscribed strand. The panels for each of the six types of substitutions as well as the mutated base are displayed in different colors. Mutations on the transcribed pyrimidine strand are displayed in blue while mutations on the untranscribed strand are displayed in red.



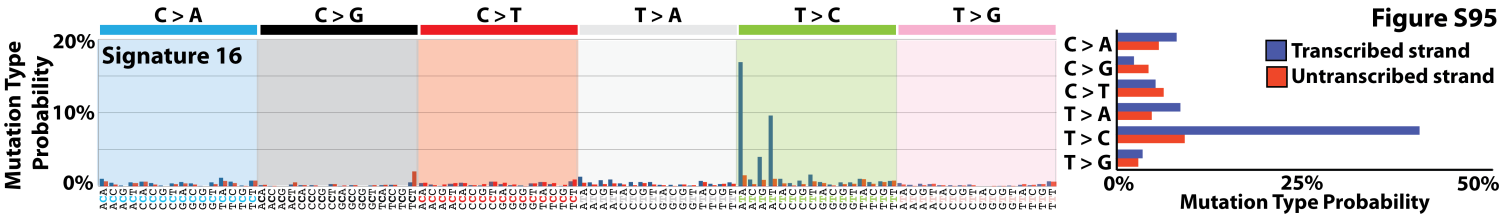
Supplementary Figure 92. Transcriptional strand bias for Signature 8. Mutational signature is displayed using a 192 substitution classification incorporating the substitution type, the sequence context immediately 5' and 3' to the mutated base and whether the mutated base (in pyrimidine context) is on the transcribed or untranscribed strand. The panels for each of the six types of substitutions as well as the mutated base are displayed in different colors. Mutations on the transcribed pyrimidine strand are displayed in blue while mutations on the untranscribed strand are displayed in red.



Supplementary Figure 93. Transcriptional strand bias for Signature 10. Mutational signature is displayed using a 192 substitution classification incorporating the substitution type, the sequence context immediately 5' and 3' to the mutated base and whether the mutated base (in pyrimidine context) is on the transcribed or untranscribed strand. The panels for each of the six types of substitutions as well as the mutated base are displayed in different colors. Mutations on the transcribed pyrimidine strand are displayed in blue while mutations on the untranscribed strand are displayed in red.

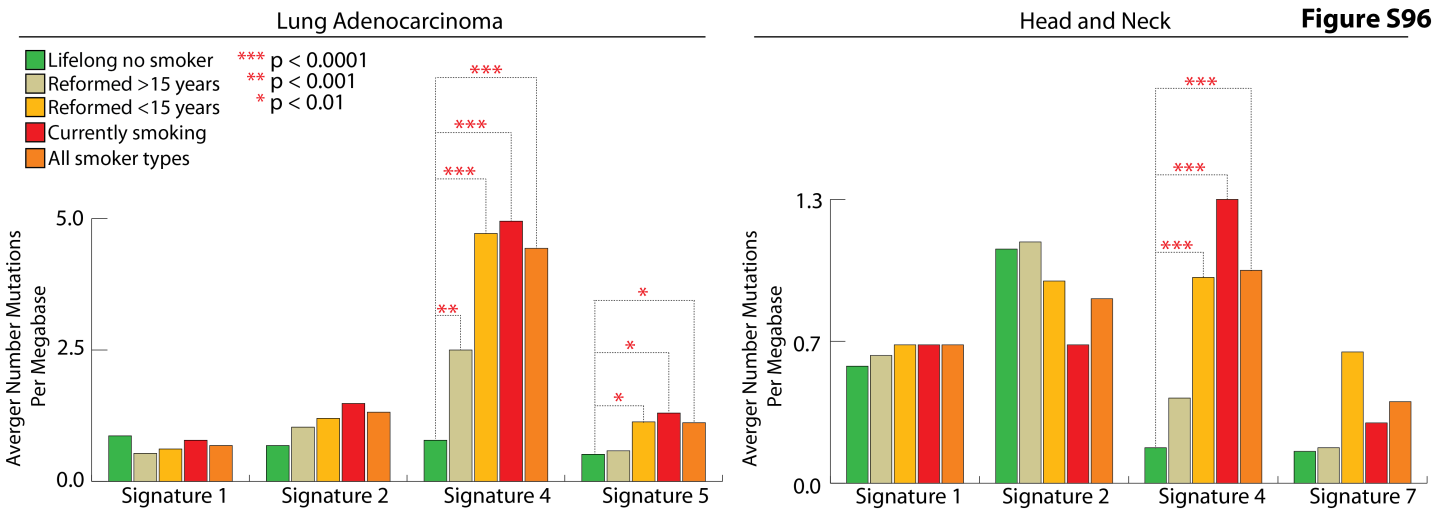


Supplementary Figure 94. Transcriptional strand bias for Signature 12. Mutational signature is displayed using a 192 substitution classification incorporating the substitution type, the sequence context immediately 5' and 3' to the mutated base and whether the mutated base (in pyrimidine context) is on the transcribed or untranscribed strand. The panels for each of the six types of substitutions as well as the mutated base are displayed in different colors. Mutations on the transcribed pyrimidine strand are displayed in blue while mutations on the untranscribed strand are displayed in red.



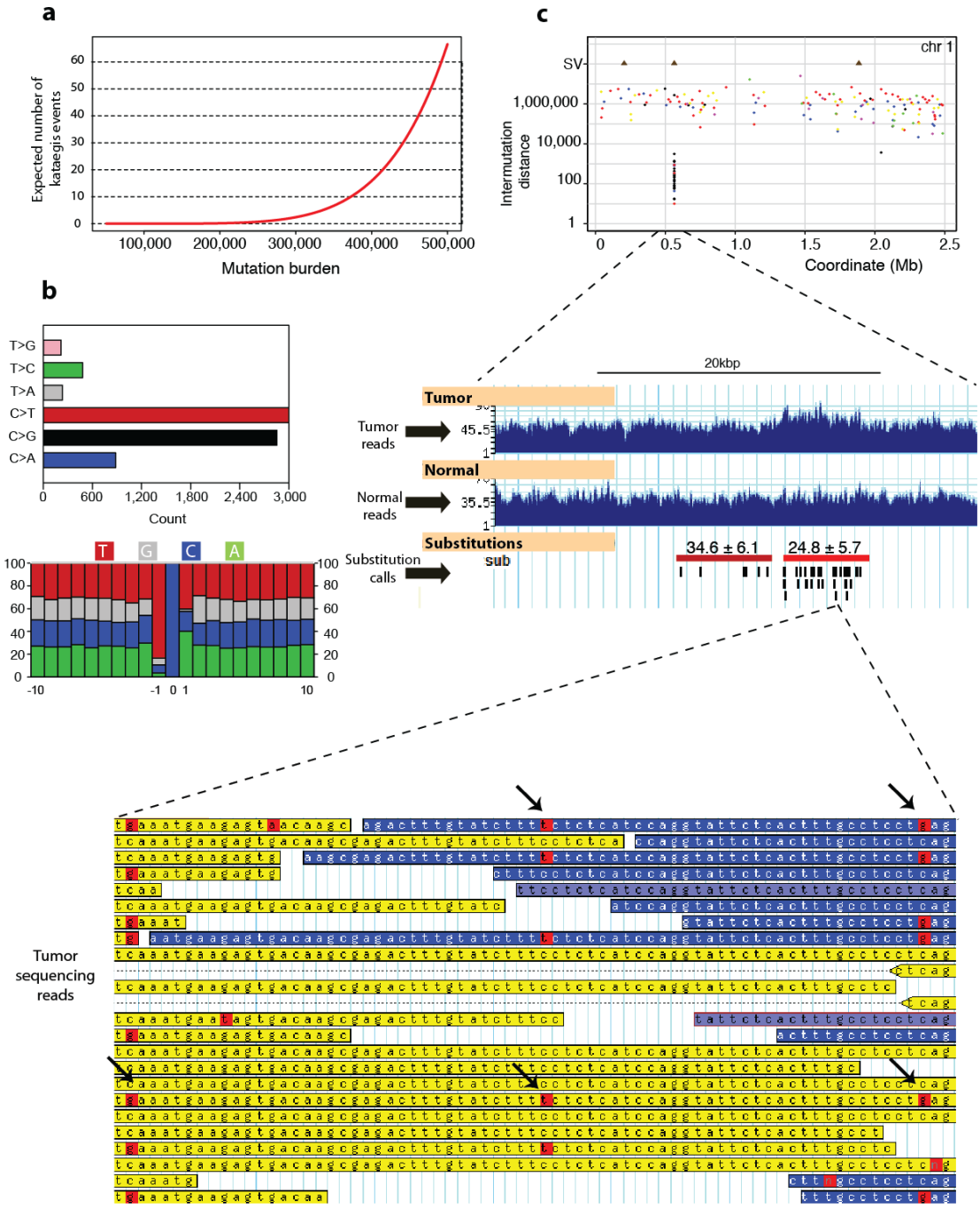
Supplementary Figure 95. Transcriptional strand bias for Signature 16. Mutational signature is displayed using a 192 substitution classification incorporating the substitution type, the sequence context immediately 5' and 3' to the mutated base and whether the mutated base (in pyrimidine context) is on the transcribed or untranscribed strand. The panels for each of the six types of substitutions as well as the mutated base are displayed in different colors. Mutations on the transcribed pyrimidine strand are displayed in blue while mutations on the untranscribed strand are displayed in red.

Figure S96



Supplementary Figure 96. Associating exposures of mutational signatures and cigarette smoking in samples from lung adenocarcinomas and head and neck cancers. Statistical analysis was performed as described in the online methods.

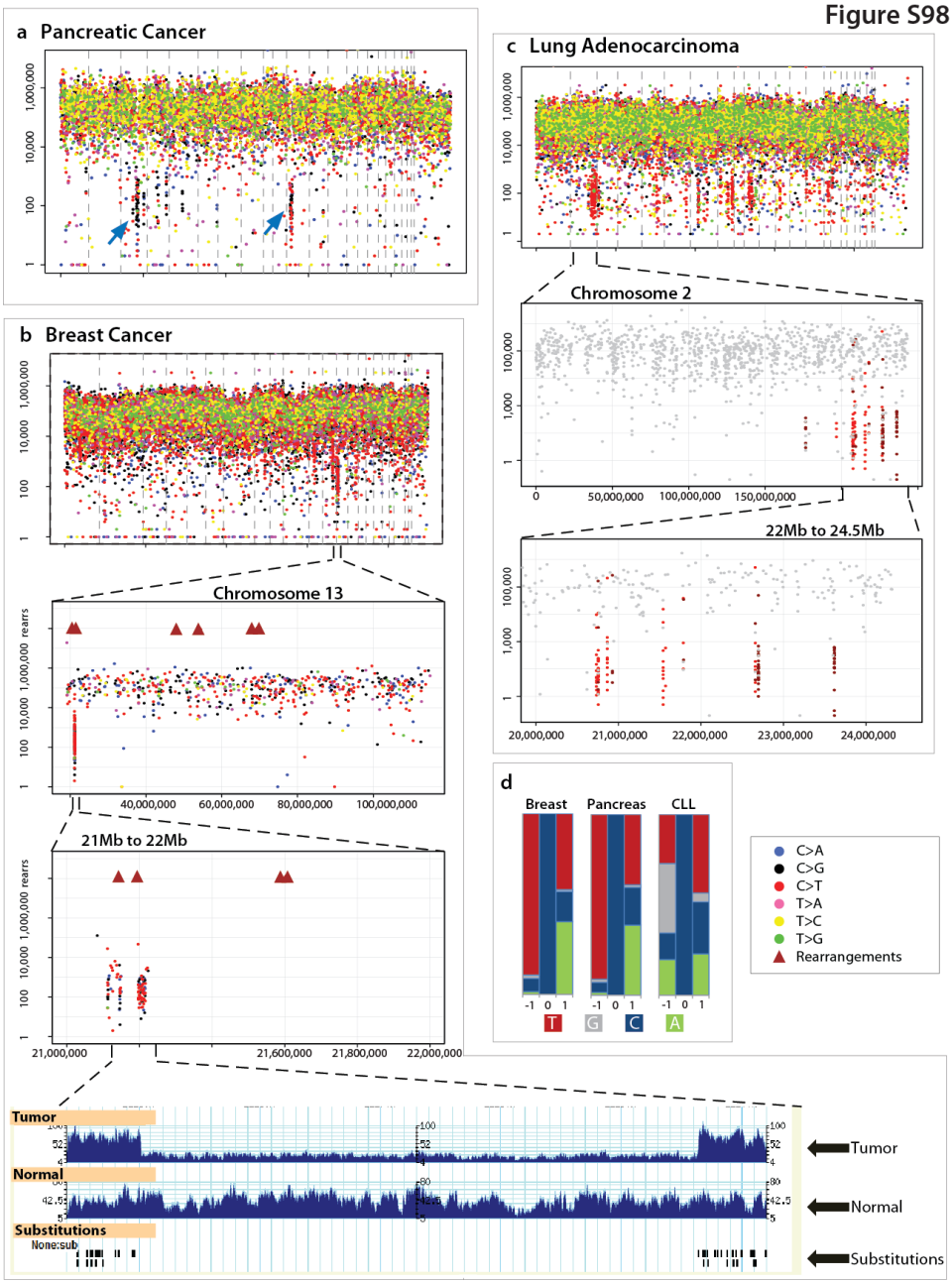
Figure S97



Supplementary Figure 97. Identifying kataegis events. ***(a)*** The expected number of *kataegis* events identified in genomes containing between 100,000 and 500,000 mutations. ***(b)*** Mutational spectra and sequence context of foci of *kataegis* that were called via the PCF-based method. ***(c)*** The relationship between *kataegis* foci and somatic structural variation. Marked co-localization of *kataegis* foci with

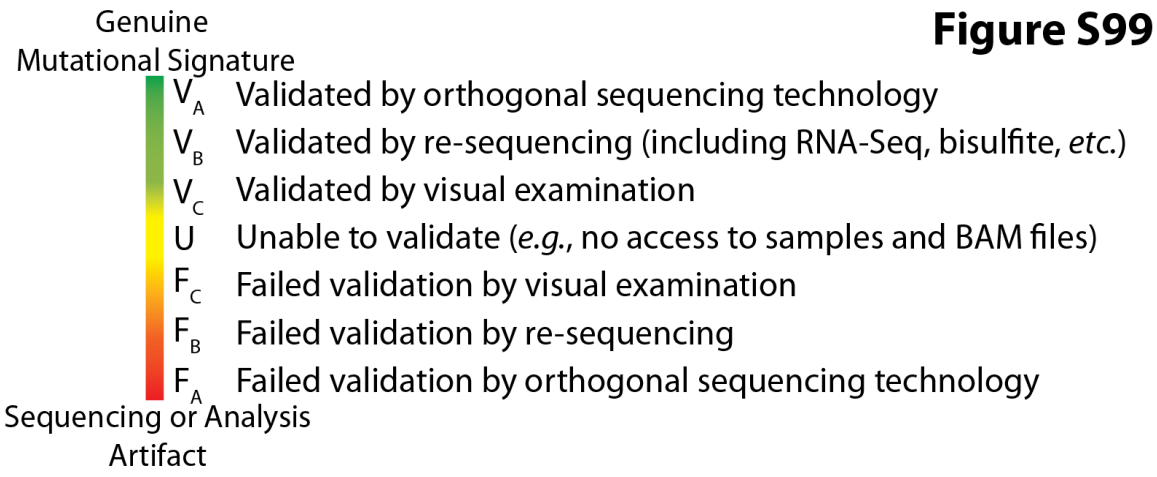
rearrangements is seen in this cancer. Seen at increasing resolution, apart from this cluster around a tandem duplication, no other mutations were seen for 2Mb flanking this region. Note the difference in variant allele fraction (mean ± standard deviation provided) between the variants within the region of tandem duplication (red bar) and the region flanking it (brown bar). Short reads demonstrating the presence of mutations (arrowed) in *cis* on the same sequencing read. Blue reads (forward reads), yellow reads (reverse reads), red bases represent mismatches relative the reference genome. No mismatches were seen in the matched normal at these sites.

Figure S98



Supplementary Figure 98. *Kataegis* is present in a wide range of cancer types. *(a)* C>G and C>T *kataegis* is present on chromosomes 3 and 11 (arrowed) in this genome-wide rainfall plot of a pancreatic cancer. ***(b)*** The relationship between foci of *kataegis* and rearrangements in a breast cancer is explored on chromosome 13 at increasing resolution until the 50kb window (lowest panel) which demonstrates the contrast between the read coverage in the tumour and the normal. A 40kb deletion

in the tumour is flanked by *kataegis* substitutions. ***(c)*** A lung cancer demonstrating many foci of *kataegis*. Mutations in grey are all mutations not involved in *kataegis*. Mutations in *kataegis* are coloured red for C>T and dark red for G>A. The foci on chromosome 2 therefore demonstrate processivity, or strand-coordinated mutations within individual kataegis microclusters. ***(d)*** Representation of 5' and 3' bases of all C>X mutations involved in *kataegis* in three cancer types showing the predilection for a TpC context in breast and pancreatic cancer. This contrasted to mutation clusters in the CLL samples, which represented the only recurrent foci across cancers and is believed to be due to the activity of AID. This is supported by the enrichment of ApC and GpC in sequence context of CLL kataegis. Dashed vertical lines represent the next chromosome. Log intermutation distance on the y-axis.



Supplementary Figure 99. Types of statuses for validating mutational signatures. Please refer to online methods for our strategy for validating mutational signatures.