Every year of smoking causes DNA mutations that make cancer more likely

This graphic shows the average number of mutations produced each year in a given type of cell by smoking a pack of cigarettes every day.

Attention smokers: For every year that you continue your pack-a-day habit, the DNA in every cell in your lungs acquires about 150 new mutations.

Some of those mutations may be harmless, but the more there are, the greater the risk that one or more of them will wind up causing cancer.

The threat doesn’t stop there, according to a study in Friday’s edition of the journal Science. After a year of smoking a pack of cigarettes each day, the cells in the larynx pick up roughly 97 new mutations, those in the pharynx accumulate 39 new mutations, and cells in the oral cavity gain 23 new mutations.
Even organs with no direct exposure to tobacco smoke appear to be affected. The researchers counted about 18 new mutations in every bladder cell and six new mutations in every liver cell for each “pack-year” that smokers smoked.

The findings are based on a genetic analysis of 5,243 cancers, including 2,490 from smokers and 1,063 from patients who said they had never smoked tobacco cigarettes.

At this point, it should come as no surprise that cigarettes are bad for your health. The U.S. surgeon general has warned about the dangers of smoking for more than 50 years. The American Cancer Society recently calculated that in a single year, at least 167,133 cancer deaths in the U.S. could be blamed on smoking.

But the new study offers a clearer picture of how smoking does its deadly damage.

The researchers used powerful supercomputers to compare thousands of cancer genome sequences. The computers grouped the sequences into about 20 distinct categories, or “mutational signatures.” Mutations tied to five of these signatures were more common in tumors from smokers than in tumors from nonsmokers.

One of the signatures involves a specific DNA nucleobase change — instead of a C for cytosine, there was an A for adenine — that “is very similar” to the change that occurs in the lab when cells are exposed to benzo[a]pyrene, a compound that the International Agency for Research on Cancer says is carcinogenic to humans.
Most of the lung and larynx cancers obtained from smokers had this type of mutation, the researchers reported. They also found that the signature was more common among smokers than nonsmokers.

Another mutational signature was characterized by Cs that should have been Ts (thymine) and vice versa. Although these changes can be found in all kinds of cancers, the signature was 1.3 to 5.1 times more common in tumors from smokers than in tumors from nonsmokers, according to the study.

The researchers said they think these kinds of mutations have the effect of speeding up the “clock” inside of cells. The faster a cellular clock runs, the more chances the cell’s DNA has to mutate.

“The way tobacco smoking causes cancer is more complex than we thought,” study coauthor Mike Stratton, director of the Wellcome Trust Sanger Institute in England, said in a statement.
quit smoking and addiction

Medical EXPOSE
http://www.medicalexpose.com/
Assessing smoke damage in cancer genomes

We have known for over 60 years that smoking tobacco is one of the most avoidable risk factors for cancer. Yet the detailed mechanisms by which tobacco smoke damages the genome and creates the mutations that ultimately cause cancer are still not fully understood. Alexandrov et al. examined mutational signatures and DNA methylation changes in over 5000 genome sequences from 17 different cancer types linked to smoking (see the Perspective by Pfeifer). They found a complex pattern of mutational signatures. Only cancers originating in tissues directly exposed to smoke showed a signature characteristic of the known tobacco carcinogen benzo[a]pyrene. One mysterious signature was shared by all smoking-associated cancers but is of unknown origin. Smoking had only a modest effect on DNA methylation.
extents in different cancers. One of these signatures, mainly found in cancers derived from tissues directly exposed to tobacco smoke, is attributable to misreplication of DNA damage caused by tobacco carcinogens. Others likely reflect indirect activation of DNA editing by APOBEC cytidine deaminases and of an endogenous clocklike mutational process. Smoking is associated with limited differences in methylation. The results are consistent with the proposition that smoking increases cancer risk by increasing the somatic mutation load, although direct evidence for this mechanism is lacking in some smoking-related cancer types.

Tobacco smoking has been associated with at least 17 types of human cancer (Table 1) and claims the lives of more than 6 million people every year (1–4). Tobacco smoke is a complex mixture of chemicals, among which at least 60 are carcinogens (5). Many of these are thought to cause cancer by inducing DNA damage that, if misreplicated, leads to an increased burden of somatic mutations and, hence, an elevated chance of acquiring driver mutations in cancer genes. Such damage often occurs in the form of covalent bonding of metabolically activated reactive species of the carcinogen to DNA bases, termed DNA adducts (6). Tissues directly exposed to tobacco smoke (e.g., lung), as well as some tissues not directly exposed (e.g., bladder), show elevated levels of DNA adducts in smokers and, thus, evidence of exposure to carcinogenic components of tobacco smoke (7, 8).

Table 1 Mutational signatures and cancer types associated with tobacco smoking.

Information about the age-adjusted odds ratios for current male smokers to develop cancer is taken from (2–4). Odds ratios for small cell lung cancer, squamous cell lung cancer, and lung adenocarcinoma are for an average daily dose of more than 30 cigarettes. Odds ratios for cervical and ovarian cancers are for current female smokers. Detailed information about all mutation types, all mutational signatures, and DNA methylation is provided in table S2. Nomenclature for signature identification numbers is consistent with the COSMIC database (http://cancer.sanger.ac.uk/cosmic/signatures). The numbers of smokers and nonsmokers are unknown (i.e., not reported in the original studies) for acute myeloid leukemia, stomach, ovarian, and colorectal cancers. The patterns of all mutational signatures with elevated mutation burden in smokers are displayed in Fig. 2B. N/A denotes lack of smoking annotation for a given cancer type. Asterisks indicate that a signature correlates with pack years smoked in a cancer type. N.S. reflects cancer types without statistically significant elevation of mutational signatures. The odds ratio for all cancer types is not provided.

View this table:
- View popup
- View inline

Each biological process causing mutations in somatic cells leaves a mutational signature (9). Many cancers have a somatic mutation in the TP53 gene, and catalogs of TP53 mutations compiled two decades ago enabled early exploration of these signatures (10), showing that lung cancers from smokers have more C>A transversions...
than lung cancers from nonsmokers (11–14). To investigate mutational signatures using the thousands of mutation catalogs generated by systematic cancer genome sequencing, we recently described a framework in which each base substitution signature is characterized using a 96-mutation classification that includes the six substitution types together with the bases immediately 5’ and 3’ to the mutated base (15). The analysis extracts mutational signatures from mutation catalogs and estimates the number of mutations contributed by each signature to each cancer genome (15). Using this approach, more than 30 different base substitution signatures have been identified (16–18).

In this study, we examined 5243 cancer genome sequences (4633 exomes and 610 whole genomes) of cancer classes for which smoking increases risk, with the goal of identifying mutational signatures and methylation changes associated with tobacco smoking (table S1). Of the samples we studied, 2490 were reported to be from tobacco smokers and 1063 from never-smokers (Table 1). Thus, we were able to investigate the mutational consequences of smoking by comparing somatic mutations and methylation in smokers versus nonsmokers for lung, larynx, pharynx, oral cavity, esophageal, bladder, liver, cervical, kidney, and pancreatic cancers (Fig. 1 and table S2).

**Fig. 1** Comparison between tobacco smokers and lifelong nonsmokers.

Bars are used to display average values for numbers of somatic substitutions per megabase (MB), numbers of indels per megabase, numbers of dinucleotide mutations per megabase (Dinucs), numbers of breakpoints per megabase (Breaks), fraction of the genome that shows copy-number changes (Aberr.), and numbers of mutations per megabase attributed to mutational signatures found in multiple cancer types associated
with tobacco smoking. Light gray bars represent nonsmokers, whereas dark gray bars are for smokers. Comparisons between smokers and nonsmokers for all features, including mutational signatures specific for a cancer type and overall DNA methylation, are provided in table S2. Error bars correspond to 95% confidence intervals for each feature. Each q value is based on a two-sample Kolmogorov-Smirnov test corrected for multiple hypothesis testing for all features in a cancer type. Cancer types are ordered based on their age-adjusted odds ratios for smoking, as provided in Table 1. Data for numbers of breakpoints per megabase and fraction of the genome that shows copy-number changes were not available for liver cancer and small cell lung cancer. Adeno, adenocarcinoma; Esophag., esophagus. Note that the presented data include only a few cases (<10) of nonsmokers for small cell lung cancer, squamous cell lung cancer, and cancer of the larynx.

We first compared total numbers of base substitutions, small insertions and deletions (indels), and genomic rearrangements. The total number of base substitutions was higher in smokers compared with nonsmokers for all cancer types together (q-value < 0.05) and, for individual cancer types, in lung adenocarcinoma, larynx, liver, and kidney cancers (table S2). Total numbers of indels were higher in smokers compared with nonsmokers in lung adenocarcinoma and liver cancer (table S2). The whole-genome–sequenced cases allowed comparison of genome rearrangements between smokers and nonsmokers in pancreatic and liver cancer, where no differences were found (table S2). However, subchromosomal copy-number changes entail genomic rearrangement and can serve as surrogates for rearrangements. Lung adenocarcinomas from smokers exhibited more copy-number aberrations than those from nonsmokers (table S2).

We then extracted mutational signatures, estimated the contributions of each signature to each cancer, and compared the numbers of mutations attributable to each signature in smokers and nonsmokers. Increases in smokers compared with nonsmokers were seen for signatures 2, 4, 5, 13, and 16 [the mutational signature nomenclature is that used in the Catalogue of Somatic Mutations in Cancer (COSMIC) and in (16–18)]. There was sufficient statistical power to show that these increases were of clonal mutations (mutations present in all cells of each cancer) for signatures 4 and 5 (q < 0.05), as expected if these mutations are due to cigarette smoke exposure before neoplastic change (supplementary text).

Signature 4 is characterized mainly by C>A mutations with smaller contributions from other base substitution classes (Fig. 2B and fig. S1). This signature was found only in cancer types in which tobacco smoking increases risk and mainly in those derived from epithelia directly exposed to tobacco smoke (figs. S2 and S3). Signature 4 is very similar to the mutational signature induced in vitro by exposing cells to benzo[a]pyrene (cosine similarity = 0.94) (Fig. 2B and fig. S3), a tobacco smoke carcinogen (19). The similarity extends to the presence of a transcriptional strand bias indicative of transcription-coupled nucleotide excision repair (NER) of bulky DNA adducts on guanine (fig. S1), the proposed mechanism of DNA damage by benzo[a]pyrene. Thus, signature 4 is likely the direct mutational consequence of misreplication of DNA damage induced by tobacco carcinogens.
Fig. 2 Mutational signatures associated with tobacco smoking.

(A) Each panel contains 25 randomly selected cancer genomes (represented by individual bars) from either smokers or nonsmokers in a given cancer type. The y axes reflect numbers of somatic mutations per megabase. Each bar is colored proportionately to the number of mutations per megabase attributed to the mutational signatures found in that sample. The naming of mutational signatures is consistent with previous reports (16–18). (B) Each panel contains the pattern of a mutational signature associated with tobacco smoking. Signatures are depicted using a 96-substitution classification defined by the substitution type and sequence context immediately 5’ and 3’ to the mutated base. Different colors are used to display the various types of substitutions. The percentages of mutations attributed to specific substitution types are on the y axes, whereas the x axes display different types of substitutions. Mutational signatures are depicted based on the trinucleotide frequency of the whole human genome. Signatures 2, 4, 5, 13, and 16 are extracted from cancers associated with tobacco smoking. The signature of benzo[a]pyrene is based on in vitro experimental data (19). Numerical values for these mutational signatures are provided in table S6.

Most lung and larynx cancers from smokers had many signature 4 mutations. Signature 4 mutations occurred more often in cancers from smokers compared with nonsmokers in all cancer types together (table S2) and in lung squamous, lung adenocarcinoma, and larynx cancers (table S2). This finding largely accounts for differences in total numbers of base substitutions (Table 1). In nonsmokers, 13.8% of lung cancers showed many signature 4 mutations (Fig. 2A; >1 mutation per megabase), which may be due to passive smoking, misreporting of smoking habits, or annotation errors. Signature 4 mutations were also detected in cancers of the oral cavity, pharynx, and esophagus, albeit in much smaller numbers than in lung and larynx cancers, perhaps because of reduced exposure to tobacco smoke or more efficient clearance. Differences in mutation
burden attributed to signature 4 between smokers and nonsmokers were not observed in these cancer types (Fig. 1). Signature 4 mutations were found at low levels in cancers of the liver, an organ not directly exposed to tobacco smoke, and were elevated in smokers versus nonsmokers (Fig. 1).

Signature 4 was not extracted from bladder, cervical, kidney, or pancreatic cancers, despite the known risks conferred by smoking and the presence of many smokers in these series. Additionally, this mutational signature was not extracted from cancers of the stomach, colorectum, and ovary, nor from acute myeloid leukemia (in the analyzed series, the smoking status of patients with these cancers was unknown, but it is likely that many have been smokers). The tissues from which all of these cancer types are derived are not directly exposed to tobacco smoke. Simulations indicate that the lack of signature 4 is not due to statistical limitations (supplementary text and fig. S4). The absence of signature 4 suggests that misreplication of direct DNA damage due to tobacco smoke constituents does not contribute substantially to mutation burden in these cancers, even though DNA adducts indicative of tobacco-induced DNA damage are present in the tissues from which they arise (7).

Signatures 2 and 13 are characterized by C>T and C>G mutations, respectively, at TpC dinucleotides and have been attributed to overactive DNA editing by APOBEC deaminases (20, 21). The cause of the overactivity in most cancers has not been established, although APOBECs are implicated in the cellular response to the entrance of foreign DNA, retrotransposon movement, and local inflammation (22). Signatures 2 and 13 showed more mutations in smokers versus nonsmokers with lung adenocarcinoma (table S2). Because these signatures are found in many other cancer types, where they are apparently unrelated to tobacco smoking, it seems unlikely that the signature 2 and 13 mutations associated with smoking in lung adenocarcinoma are direct consequences of misreplication of DNA damage induced by tobacco smoke. More plausibly, the cellular machinery underlying signatures 2 and 13 is activated by tobacco smoke, perhaps as a result of inflammation arising from the deposition of particulate matter or by indirect consequences of DNA damage.

Signature 5 is characterized by mutations distributed across all 96 subtypes of base substitution, with a predominance of T>C and C>T mutations (Fig. 2B) and evidence of transcriptional strand bias for T>C mutations (18). Signature 5 is found in all cancer types, including those unrelated to tobacco smoking, and in most cancer samples. It is “clocklike” in that the number of mutations attributable to this signature correlates with age at the time of diagnosis in many cancer types (17). Signature 5, together with signature 1, is thought to contribute to mutation accumulation in most normal somatic cells and in the germline (17, 23). The mechanisms underlying signature 5 are not well understood, although an enrichment of signature 5 mutations was found in bladder cancers harboring inactivating mutations in ERCC2, which encodes a component of NER (24).

Signature 5 (or a similar signature that is difficult to differentiate from signature 5 because of the relatively flat profiles of these signatures) was increased by a factor of 1.3 to 5.1 (q < 0.05; table S2) in smokers versus nonsmokers in all cancer types.
together and in lung squamous, lung adenocarcinoma, larynx, pharynx, oral cavity, esophageal squamous, bladder, liver, and kidney cancers. The association of smoking with signature 5 mutations across these nine cancer types therefore includes some for which the risks conferred by smoking are modest and for which normal progenitor cells are not directly exposed to cigarette smoke (Table 1). Given the clocklike nature of signature 5 (17), its presence in the human germline (23), its ubiquity in cancer types unrelated to tobacco smoking (18), and its widespread occurrence in nonsmokers, it seems unlikely that signature 5 mutations associated with tobacco smoking are direct consequences of misreplication of DNA damaged by tobacco carcinogens. It is more plausible that smoking affects the machinery generating signature 5 mutations (24).

Presumably as a consequence of the effects of smoking, signature 5 mutations correlated with age at the time of diagnosis in nonsmokers (P = 0.001) but not in smokers (P = 0.59).

Signature 16 is predominantly characterized by T>C mutations at ApT dinucleotides (Fig. 2B); exhibits a strong transcriptional strand bias consistent with almost all damage occurring on adenine (fig. S5); and, thus far, has been detected only in liver cancer. The underlying mutational process is currently unknown. Signature 16 exhibited a higher mutation burden in smokers versus nonsmokers with liver cancer (table S2). For smokers with lung, larynx, pharynx, oral cavity, esophageal, bladder, liver, cervical, kidney, and pancreatic cancers, quantitative data on cumulative exposure to tobacco smoke were available (table S1). Total numbers of base substitution mutations were positively correlated with pack years smoked (1 pack year is defined as smoking one pack per day for 1 year) for all cancer types together (q < 0.05) and for lung adenocarcinoma (table S3). For individual mutational signatures, correlations with pack years smoked were found in multiple cancer types for signatures 4 and 5 (table S3). Signature 4 correlated with pack years in lung squamous, lung adenocarcinoma, larynx, and liver cancers. Signature 5 correlated with pack years in all cancers together, as well as in lung adenocarcinoma, pharynx, oral cavity, and bladder cancers (table S3). In lung adenocarcinoma, correlations with pack years smoked were also observed for signatures 2 and 13. The rates of these correlations allow estimation of the approximate numbers of mutations accumulated in a normal cell of each tissue due to smoking a pack of cigarettes a day for a year: lung, 150 mutations; larynx, 97; pharynx, 39; oral cavity 23; bladder, 18; liver, 6 (table S3).

Consistent with our results, previous studies have reported higher numbers of total base substitutions in lung adenocarcinoma in smokers versus nonsmokers (mainly due to C>A substitutions) (25, 26). The same is true of signatures 4 and 5 in lung adenocarcinoma (18), signature 4 in liver cancer (27), and signature 5 in bladder cancer (24).

Differential methylation of the DNA of normal cells of smokers compared to nonsmokers has been reported (28). Using data from methylation arrays, each containing ~470,000 of the ~28 million CpG sites in the human genome, we evaluated whether differences in methylation are found in cancers. Overall levels of CpG methylation in DNA from cancers were similar in smokers and nonsmokers for all cancer types (fig. S6).
Individual CpGs were differentially methylated (>5% difference) in only two cancer types: 369 CpGs were hypomethylated and 65 were hypermethylated in lung adenocarcinoma, with five hypomethylated and three hypermethylated in oral cancer (Fig. 3 and fig. S7). CpGs exhibiting differences in methylation clustered in certain genes but were not associated with known cancer genes more than expected by chance, nor with genes hypomethylated in normal blood or buccal cells of tobacco smokers (fig. S8 and tables S4 and S5) (28). Therefore, with the exception of lung cancer, CpG methylation showed limited differences between the cancers of smokers and nonsmokers (Fig. 3).

Fig. 3Differentially methylated individual CpGs in tobacco smokers across cancers associated with tobacco smoking.

Each dot represents an individual CpG. The x axes reflect differences in methylation between lifelong nonsmokers and smokers, where positive values correspond to hypermethylation and negative values to hypomethylation. The y axes depict levels of statistical significance. Results satisfying a Bonferroni threshold of $10^{-7}$ (above the red line) are considered statistically significant.
The genomes of smoking-associated cancers permit reassessment of our understanding of how tobacco smoke causes cancer. Consistent with the proposition that an increased mutation load caused by tobacco smoke contributes to increased cancer risk, the total mutation burden is elevated in smokers versus nonsmokers with lung adenocarcinoma, larynx, liver, and kidney cancers. However, differences in total mutation burden were not observed in the other smoking-associated cancer types and, in some, there were no statistically significant smoking-associated differences in mutation load, signatures, or DNA methylation. Caution should be exercised in the interpretation of the latter observations. In addition to limitations of statistical power, multiple rounds of clonal expansion over many years are often required for development of a symptomatic cancer. It is thus conceivable that, in the normal tissues from which smoking-associated cancer types originate, there are more somatic mutations (or differences in methylation) in smokers than in nonsmokers but that these differences become obscured during the intervening clonal evolution. Moreover, some theoretical models predict that relatively small differences in mutation burden caused by smoking in preneoplastic cells could account for the observed increases in cancer risks (29). Other models indicate that differences in mutation burden between smokers and nonsmokers need not be observed in the final cancers (supplementary text and fig. S6). Thus, increased somatic mutation loads in precancerous tissues may still explain the smoking-induced risks of most cancers, although other mechanisms have been proposed (30, 31).

However, the generation of increased somatic mutation burden by tobacco smoking appears to be mechanistically complex. Smoking correlates with increases in base substitutions of multiple mutational signatures, together with increases in indels and copy-number changes. The extent to which these distinct mutational processes operate differs between tissue types (at least partially depending on the degree of direct exposure to tobacco smoke), and their mechanisms range from misreplication of DNA damage caused by tobacco smoke constituents to activation of more generally operative mutational processes. Although we cannot exclude roles for covariate behaviors of smokers or differences in the biology of cancers arising in smokers compared with nonsmokers, smoking itself is most plausibly the cause of these differences.

Supplementary Materials

www.sciencemag.org/content/354/6312/618/suppl/DC1

Materials and Methods

Supplementary Text

Figs. S1 to S10

Tables S1 to S6

References (32–54)
References and Notes

1. B. Secretan,
2. K. Straif,
3. R. Baan,
4. Y. Grosse,
5. F. El Ghissassi,
6. V. Bouvard,
7. L. Benbrahim-Tallaa,
8. N. Guha,
9. C. Freeman,
10. L. Galichet,
11. V. Cogliano,
12. WHO International Agency for Research on Cancer Monograph Working Group


CrossRef Medline Web of Science Google Scholar

2. S. S. Lim,
3. T. Vos,
4. A. D. Flaxman,
5. G. Danaei,
6. K. Shibuya,
7. H. Adair-Rohani,
8. M. Amann,
9. H. R. Anderson,
10. K. G. Andrews,
11. M. Aryee,
12. C. Atkinson,
13. L. J. Bacchus,
14. A. N. Bahalim,
15. K. Balakrishnan,
16. J. Balmes,
17. S. Barker-Collo,
18. A. Baxter,
19. M. L. Bell,
20. J. D. Blore,
21. F. Blyth,
22. C. Bonner,
23. G. Borges,
24. R. Bourne,
25. M. Boussinesq,
26. M. Brauer,
27. P. Brooks,
28. N. G. Bruce,
28. B. Brunekreef,
29. C. Bryan-Hancock,
30. C. Bucello,
31. R. Buchbinder,
32. F. Bull,
33. R. T. Burnett,
34. T. E. Byers,
35. B. Calabria,
36. J. Carapetis,
37. E. Carnahan,
38. Z. Chafe,
39. F. Charlson,
40. H. Chen,
41. J. S. Chen,
42. A. T. Cheng,
43. J. C. Child,
44. A. Cohen,
45. K. E. Colson,
46. B. C. Cowie,
47. S. Darby,
48. S. Darling,
49. A. Davis,
50. L. Degenhardt,
51. F. Dentener,
52. D. C. Des Jarlais,
53. K. Devries,
54. M. Dherani,
55. E. L. Ding,
56. E. R. Dorsey,
57. T. Driscoll,
58. K. Edmond,
59. S. E. Ali,
60. R. E. Engell,
61. P. J. Erwin,
62. S. Fahimi,
63. G. Falder,
64. F. Farzadfar,
65. A. Ferrari,
66. M. M. Finucane,
67. S. Flaxman,
68. F. G. Fowkes,
69. G. Freedman,
70. M. K. Freeman,
71. E. Gakidou,
72. S. Ghosh,
73. E. Giovannucci,
74. G. Gmel,
75. K. Graham,
76. R. Grainger,
77. B. Grant,
78. D. Gunnell,
79. H. R. Gutierrez,
80. W. Hall,
81. H. W. Hoek,
82. A. Hogan,
83. H. D. Hosgood 3rd,
84. D. Hoy,
85. H. Hu,
86. B. J. Hubbell,
87. S. J. Hutchings,
88. S. E. Ibeanusi,
89. G. L. Jacklyn,
90. R. Jasrasaria,
91. J. B. Jonas,
92. H. Kan,
93. J. A. Kanis,
94. N. Kassebaum,
95. N. Kawakami,
96. Y. H. Khang,
97. S. Khatibzadeh,
98. J. P. Khoo,
99. C. Kok,
100. F. Laden,
101. R. Laloo,
102. Q. Lan,
103. T. Lathlean,
104. J. L. Leasher,
105. J. Leigh,
106. Y. Li,
107. J. K. Lin,
108. S. E. Lipshultz,
109. S. London,
110. R. Lozano,
111. Y. Lu,
112. J. Mak,
113. R. Malekzadeh,
114. L. Mallinger,
115. W. Marcenes,
116. L. March,
117. R. Marks,
118. R. Martin,
119. P. McGale,
120. J. McGrath,
121. S. Mehta,
122. G. A. Mensah,
123. T. R. Merriman,
124. R. Micha,
125. C. Michaud,
126. V. Mishra,
127. K. Mohd Hanafiah,
128. A. A. Mokdad,
129. L. Morawska,
D. Mozaffarian,
T. Murphy,
M. Naghavi,
B. Neal,
P. K. Nelson,
J. M. Nolla,
R. Norman,
C. Olives,
S. B. Omer,
J. Orchard,
R. Osborne,
B. Ostro,
A. Page,
K. D. Pandey,
C. D. Parry,
E. Passmore,
J. Patra,
N. Pearce,
P. M. Pelizzari,
M. Petzold,
M. R. Phillips,
D. Pope,
C. A. Pope 3rd,
J. Powles,
M. Rao,
H. Razavi,
E. A. Rehfues,
J. T. Rehm,
B. Ritz,
F. P. Rivara,
T. Roberts,
C. Robinson,
J. A. Rodriguez-Portales,
I. Romieu,
R. Room,
L. C. Rosenfeld,
A. Roy,
L. Rushton,
J. A. Salomon,
U. Sampson,
L. Sanchez-Riera,
E. Sanman,
A. Sapkota,
S. Seedat,
P. Shi,
K. Shield,
R. Shivakoti,
G. M. Singh,
D. A. Sleet,
E. Smith,
K. R. Smith,


3.

14. L. Simonato,  
15. C. Fortes,  
16. J. Siemiatycki,  
17. M.-E. Parent,  
18. D. Consonni,  
19. M. T. Landi,  
20. N. Caporaso,  
21. D. Zaridze,  
22. A. Cassidy,  
23. N. Szeszenia-Dabrowska,  
24. P. Rudnai,  
25. J. Lissowska,  
26. I. Stücker,  
27. E. Fabianova,  
28. R. S. Dumitru,  
29. V. Bencko,  
30. L. Foretova,  
31. V. Janout,  
32. C. M. Rudin,  
33. P. Brennan,  
34. P. Boffetta,  
35. K. Straif,  
36. T. Brüning


CrossRef Medline Web of Science Google Scholar


CrossRef Medline Web of Science Google Scholar

1. A. Agudo,  
2. C. Bonet,  
3. N. Travier,  
4. C. A. González,  
5. P. Vineis,  
6. H. B. Bueno-de-Mesquita,  
7. D. Trichopoulos,  
8. P. Boffetta,  
9. F. Clavel-Chapelon,  
10. M.-C. Boutron-Ruault,  
11. R. Kaaks,  
12. A. Lukanova,  
13. M. Schütze,  
14. H. Boeing,  
15. A. Tjonneland,  
16. J. Halkjaer,  
17. K. Overvad,  
18. C. C. Dahm,  
19. J. R. Quirós,  
20. M.-J. Sánchez,


Abstract/FREE Full Text Google Scholar

5. 

1. S. S. Hecht


CrossRef Medline Web of Science Google Scholar

6. 


7. 

1. D. H. Phillips


Abstract/FREE Full Text Google Scholar
8.  
1. D. H. Phillips,  
2. S. Venitt  

CrossRefMedlineWeb of ScienceGoogle Scholar

9.  
1. L. B. Alexandrov,  
2. M. R. Stratton  

CrossRefMedlineGoogle Scholar

10.  
1. P. Hainaut,  
2. M. Hollstein  

CrossRefMedlineGoogle Scholar

11.  
1. M. F. Denissenko,  
2. A. Pao,  
3. M. Tang,  
4. G. P. Pfeifer  

Abstract/FREE Full TextGoogle Scholar

12.  
1. G. P. Pfeifer,  
2. M. F. Denissenko  

CrossRefMedlineWeb of ScienceGoogle Scholar

13.  
1. L. E. Smith,  
2. M. F. Denissenko,
3. W. P. Bennett, 
4. H. Li, 
5. S. Amin, 
6. M. Tang, 
7. G. P. Pfeifer


Abstract/FREE Full TextGoogle Scholar

14. 

1. F. Le Calvez, 
2. A. Mukeria, 
3. J. D. Hunt, 
4. O. Kelm, 
5. R. J. Hung, 
6. P. Tanière, 
7. P. Brennan, 
8. P. Boffetta, 
9. D. G. Zaridze, 
10. P. Hainaut


Abstract/FREE Full TextGoogle Scholar

15. 

1. L. B. Alexandrov, 
2. S. Nik-Zainal, 
3. D. C. Wedge, 
4. P. J. Campbell, 
5. M. R. Stratton


CrossRefMedlineGoogle Scholar

16. 

1. L. B. Alexandrov


Abstract/FREE Full TextGoogle Scholar

17. 

1. L. B. Alexandrov,
2. P. H. Jones,
3. D. C. Wedge,
4. J. E. Sale,
5. P. J. Campbell,
6. S. Nik-Zainal,
7. M. R. Stratton


CrossRef Medline Google Scholar

18. e

1. L. B. Alexandrov,
2. S. Nik-Zainal,
3. D. C. Wedge,
4. S. A. J. R. Aparicio,
5. S. Behjati,
6. A. V. Biankin,
7. G. R. Bignell,
8. N. Bolli,
9. A. Borg,
10. A.-L. Børresen-Dale,
11. S. Boyault,
12. B. Burkhardt,
13. A. P. Butler,
14. C. Caldas,
15. H. R. Davies,
16. C. Desmedt,
17. R. Eils,
18. J. E. Eyfjörd,
19. J. A. Foekens,
20. M. Greaves,
21. F. Hosoda,
22. B. Hutter,
23. T. Ilicic,
24. S. Imbeaud,
25. M. Imielinski,
26. N. Jäger,
27. D. T. W. Jones,
28. D. Jones,
29. S. Knappskog,
30. M. Kool,
31. S. R. Lakhani,
32. C. López-Otín,
33. S. Martin,
34. N. C. Munshi,
35. H. Nakamura,
36. P. A. Northcott,
37. M. Pajic,
38. E. Papaemmanuil,
39. A. Paradiso,
40. J. V. Pearson,
41. X. S. Puente,
42. K. Raine,
43. M. Ramakrishna,
44. A. L. Richardson,
45. J. Richter,
46. P. Rosenstiel,
47. M. Schlesner,
48. T. N. Schumacher,
49. P. N. Span,
50. J. W. Teague,
51. Y. Totoki,
52. A. N. J. Tutt,
53. R. Valdés-Mas,
54. M. M. van Buuren,
55. L. van ’t Veer,
56. A. Vincent-Salomon,
57. N. Waddell,
58. L. R. Yates,
59. J. Zucman-Rossi,
60. P. A. Futreal,
61. U. McDermott,
62. P. Lichter,
63. M. Meyerson,
64. S. M. Grimmond,
65. R. Siebert,
66. E. Campo,
67. T. Shibata,
68. S. M. Pfister,
69. P. J. Campbell,
70. M. R. Stratton,
71. Australian Pancreatic Cancer Genome Initiative,
72. ICGC Breast Cancer Consortium,
73. ICGC MMML-Seq Consortium,
74. ICGC PedBrain


CrossRefMedlineWeb of ScienceGoogle Scholar

19. el

1. S. Nik-Zainal,
2. J. E. Kucab,
3. S. Morganella,
4. D. Glodzik,
5. L. B. Alexandrov,
6. V. M. Arlt,
7. A. Weninger,
8. M. Hollstein,
9. M. R. Stratton,
10. D. H. Phillips

Abstract/FREE Full TextGoogle Scholar

20. 

1. S. Nik-Zainal,
2. L. B. Alexandrov,
3. D. C. Wedge,
4. P. Van Loo,
5. C. D. Greenman,
6. K. Raine,
7. D. Jones,
8. J. Hinton,
9. J. Marshall,
10. L. A. Stebbings,
11. A. Menzies,
12. S. Martin,
13. K. Leung,
14. L. Chen,
15. C. Leroy,
16. M. Ramakrishna,
17. R. Rance,
18. K. W. Lau,
19. L. J. Mudie,
20. I. Varela,
21. D. J. McBride,
22. G. R. Bignell,
23. S. L. Cooke,
24. A. Shlien,
25. J. Gamble,
26. I. Whitmore,
27. M. Maddison,
28. P. S. Tarpey,
29. H. R. Davies,
30. E. Papaemmanuil,
31. P. J. Stephens,
32. S. McLaren,
33. A. P. Butler,
34. J. W. Teague,
35. G. Jönsson,
36. J. E. Garber,
37. D. Silver,
38. P. Miron,
39. A. Fatima,
40. S. Boyault,
41. A. Langerød,
42. A. Tutt,
Breast Cancer Working Group of the International Cancer Genome Consortium


R. Rahbari,
2. A. Wuster,
3. S. J. Lindsay,
4. R. J. Hardwick,
5. L. B. Alexandrov,
6. S. Al Turki,
7. A. Dominiczak,
8. A. Morris,
9. D. Porteous,
10. B. Smith,
11. M. R. Stratton,
12. M. E. Hurles,
13. UK10K Consortium

CrossRef Medline Google Scholar

24. 

1. J. Kim,
2. K. W. Mouw,
3. P. Polak,
4. L. Z. Braunstein,
5. A. Kamburov,
6. G. Tiao,
7. D. J. Kwiatkowski,
8. J. E. Rosenberg,
9. E. M. Van Allen,
10. A. D. D’Andrea,
11. G. Getz

CrossRef Medline Google Scholar

25. 

1. R. Govindan,
2. L. Ding,
3. M. Griffith,
4. J. Subramanian,
5. N. D. Dees,
6. K. L. Kanchi,
7. C. A. Maher,
8. R. Fulton,
9. L. Fulton,
10. J. Wallis,
11. K. Chen,
12. J. Walker,
13. S. McDonald,
14. R. Bose,
15. D. Ornitz,
16. D. Xiong,
17. M. You,
18. D. J. Dooling,
19. M. Watson,
20. E. R. Mardis,
21. R. K. Wilson


26. M. Imielinski,
27. A. H. Berger,
28. P. S. Hammerman,
29. B. Hernandez,
30. T. J. Pugh,
31. E. Hodis,
32. J. Cho,
33. J. Suh,
34. M. Capelletti,
35. A. Sivachenko,
36. C. Sougnez,
37. D. Auclair,
38. M. S. Lawrence,
39. P. Stojanov,
40. K. Cibulskis,
41. K. Choi,
42. L. de Waal,
43. T. Sharifnia,
44. A. Brooks,
45. H. Greulich,
46. S. Banerji,
47. T. Zander,
48. D. Seidel,
49. F. Leenders,
50. S. Ansén,
51. C. Ludwig,
52. W. Engel-Riedel,
53. E. Stoelben,
54. J. Wolf,
55. C. Goparju,
56. K. Thompson,
57. W. Winckler,
58. D. Kwiatkowski,
59. B. E. Johnson,
60. P. A. Jänne,
61. V. A. Miller,
62. W. Pao,
38. W. D. Travis,
39. H. I. Pass,
40. S. B. Gabriel,
41. E. S. Lander,
42. R. K. Thomas,
43. L. A. Garraway,
44. G. Getz,
45. M. Meyerson


27. 
1. A. Fujimoto,
2. M. Furuta,
3. Y. Totoki,
4. T. Tsunoda,
5. M. Kato,
6. Y. Shiraishi,
7. H. Tanaka,
8. H. Taniguchi,
9. Y. Kawakami,
10. M. Ueno,
11. K. Gotoh,
12. S. Ariizumi,
13. C. P. Wardell,
14. S. Hayami,
15. T. Nakamura,
16. H. Aikata,
17. K. Arihiro,
18. K. A. Boroevich,
19. T. Abe,
20. K. Nakano,
21. K. Maejima,
22. A. Sasaki-Oku,
23. A. Ohsawa,
24. T. Shibuya,
25. H. Nakamura,
26. N. Hama,
27. F. Hosoda,
28. Y. Arai,
29. S. Ohashi,
30. T. Urushidate,
31. G. Nagae,
32. S. Yamamoto,
33. H. Ueda,
34. K. Tatsuno,
35. H. Ojima,
36. N. Hiraoka,
37. T. Okusaka,
38. M. Kubo,
39. S. Marubashi,
40. T. Yamada,
41. S. Hirano,
42. M. Yamamoto,
43. H. Ohdan,
44. K. Shimada,
45. O. Ishikawa,
46. H. Yamaue,
47. K. Chayama,
48. S. Miyano,
49. H. Aburatani,
50. T. Shibata,
51. H. Nakagawa


28. 

1. A. E. Teschendorff,
2. Z. Yang,
3. A. Wong,
4. C. P. Pipinikas,
5. Y. Jiao,
6. A. Jones,
7. S. Anjum,
8. R. Hardy,
9. H. B. Salvesen,
10. C. Thirlwell,
11. S. M. Janes,
12. D. Kuh,
13. M. Widschwendter


29. 

1. C. Tomasetti,
2. L. Marchionni,
3. M. A. Nowak,
4. G. Parmigiani,
5. B. Vogelstein

30. M. Sopori


31. H. Rubin


34.

1. W. Fu,
2. T. D. O’Connor,
3. G. Jun,
4. H. M. Kang,
5. G. Abecasis,
6. S. M. Leal,
7. S. Gabriel,
8. M. J. Rieder,
9. D. Altshuler,
10. J. Shendure,
11. D. A. Nickerson,
12. M. J. Bamshad,
13. J. M. Akey,
14. NHLBI Exome Sequencing Project


35.

1. L. B. Alexandrov,
2. S. Nik-Zainal,
3. H. C. Siu,
4. S. Y. Leung,
5. M. R. Stratton


36.

1. K. Schulze,
2. S. Imbeaud,
3. E. Letouzé,
4. L. B. Alexandrov,
5. J. Calderaro,
6. S. Rehouissou,
7. G. Couchy,
8. C. Meiller,
9. J. Shinde,
10. F. Soysouvanh,
11. A.-L. Calatayud,
12. R. Pinotol,
13. L. Pelletier,
14. C. Balabaud,
15. A. Laurent,
16. J.-F. Blanc,  
17. V. Mazzaferro,  
18. F. Calvo,  
19. A. Villanueva,  
20. J.-C. Nault,  
21. P. Bioulac-Sage,  
22. M. R. Stratton,  
23. J. M. Llovet,  
24. J. Zucman-Rossi


37.

1. S. Behjati,  
2. M. Huch,  
3. R. van Boxtel,  
4. W. Karthaus,  
5. D. C. Wedge,  
6. A. U. Tamuri,  
7. I. Martincorena,  
8. M. Petljak,  
9. L. B. Alexandrov,  
10. G. Gundem,  
11. P. S. Tarpey,  
12. S. Roerink,  
13. J. Blokker,  
14. M. Maddison,  
15. L. Mudie,  
16. B. Robinson,  
17. S. Nik-Zainal,  
18. P. Campbell,  
19. N. Goldman,  
20. M. van de Wetering,  
21. E. Cuppen,  
22. H. Clevers,  
23. M. R. Stratton


38.

1. N. Bolli,  
2. H. Avet-Loiseau,  
3. D. C. Wedge,  
4. P. Van Loo,
40.

1. E. P. Murchison,
2. D. C. Wedge,
3. L. B. Alexandrov,
4. B. Fu,
5. I. Martincorena,
6. Z. Ning,
7. J. M. C. Tubio,
8. E. I. Werner,
9. J. Allen,
10. A. B. De Nardi,
11. E. M. Donelan,
12. G. Marino,
13. A. Fassati,
14. P. J. Campbell,
15. F. Yang,
16. A. Burt,
17. R. A. Weiss,
18. M. R. Stratton


41.
1. T. Helleday,
2. S. Eshtad,
3. S. Nik-Zainal


42.
1. S. Nik-Zainal,
2. D. C. Wedge,
3. L. B. Alexandrov,
4. M. Petljak,
5. A. P. Butler,
6. N. Bolli,
7. H. R. Davies,
8. S. Knappskog,
9. S. Martin,
10. E. Papaemmanuil,
11. M. Ramakrishna,
12. A. Shlien,
13. I. Simonic,
14. Y. Xue,
15. C. Tyler-Smith,
16. P. J. Campbell,
17. M. R. Stratton

12. Y. S. Ju,
13. M. Ramakrishna,
14. H. K. Haugland,
15. P. K. Lillegård,
16. S. Nik-Zainal,
17. S. McLaren,
18. A. Butler,
19. S. Martin,
20. D. Glodzik,
21. A. Menzies,
22. K. Raine,
23. J. Hinton,
24. D. Jones,
25. L. J. Mudie,
26. B. Jiang,
27. D. Vincent,
28. A. Greene-Colozzi,
29. P.-Y. Adnet,
30. A. Fatima,
31. M. Maetens,
32. M. Ignatiadis,
33. M. R. Stratton,
34. C. Sotiriou,
35. A. L. Richardson,
36. P. E. Lønning,
37. D. C. Wedge,
38. P. J. Campbell


Analysis of mutational signatures in exomes from B-cell lymphoma cell lines suggest APOBEC3 family members to be involved in the pathogenesis of primary effusion lymphoma. Leukemia 29, 1612–1615 (2015). doi:10.1038/leu.2015.22 pmid:25650088


1. C. O. Nordling


53.

1. P. Armitage,
2. R. Doll


54. e

1. D. T. Silverman,
2. P. Hartge,
3. A. S. Morrison,
4. S. S. Devesa


55. Acknowledgments: This work was supported by the Wellcome Trust (grant 098051). S.N.-Z. is a Wellcome-Beit Prize Fellow and is supported through a Wellcome Trust Intermediate Fellowship (grant WT100183MA). P.J.C. is personally funded through a Wellcome Trust Senior Clinical Research Fellowship (grant WT088340MA). M.R.S. is a paid advisor for GRAIL, a company developing technologies for sequencing of circulating tumor DNA for the purpose of early cancer detection. L.B.A. is personally supported through a J. Robert Oppenheimer Fellowship at Los Alamos National Laboratory. This research used resources provided by the Los Alamos National Laboratory Institutional Computing Program, which is supported by the U.S. Department of Energy (DOE) National Nuclear Security Administration under contract no. DE-AC52-06NA25396. Research performed at Los Alamos National Laboratory was carried out under the auspices of the National Nuclear Security Administration of the DOE. This work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (grant FC001202), the UK MRC (grant FC001202), and the Wellcome Trust (grant FC001202). P.V.L. is a Winton Group Leader in recognition of the Winton Charitable Foundation’s support toward the establishment of The Francis Crick Institute. D.H.P. is funded by Cancer Research UK (grant C313/A14329), the Wellcome Trust (grants 101126/Z/13/Z and 101126/B/13/Z), the National Institute for Health Research (NIHR) Health Protection Research Unit in Health Impact of Environmental Hazards at King’s College London in partnership with PHE [the views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health, or PHE], and by the project EXPOSOMICS (grant agreement 308610-FP7) (European Commission). P.V. was partially supported by the project EXPOSOMICS (grant agreement 308610-FP7) (European Commission). Y.T. and T.S. are supported by the Practical Research for Innovative Cancer Control from Japan Agency for Medical Research and Development (grant 15ck0106094h0002) and National Cancer Center Research and
Development Funds (26-A-5). We thank The Cancer Genome Atlas, the International Cancer Genome Consortium, and the authors of all studies cited in table S1 for providing free access to their somatic mutational data.