Risk of lung adenocarcinoma from smoking and radiation arises in distinct molecular pathways

Noemi Castelletti[1,*], Jan Christian Kaiser[1,*], Kyoji Furukawa[2], Helmut Küchenhoff[3], and Georgios T. Stathopoulos[4,5]

[1] Institute of Radiation Protection (ISS), Helmholtz Zentrum München; Ingolstädter Landstraße 1, 85764, Neuherberg, Bavaria, Germany.
[2] Department of Statistics, Radiation Effects Research Foundation (RERF); 5-2 Hijiyama Park, Minami-Ku, 723-0815, Hiroshima City, Japan.
[3] Department of Statistics, Ludwig-Maximilian University (LMU) Munich; Akademiestraße 1, 80799, Munich, Bavaria, Germany.
[4] Laboratory for Molecular Respiratory Carcinogenesis, Department of Physiology, Faculty of Medicine; University of Patras; 1 Asklepiou Str., 26504, Rio, Achaia, Greece.
[5] Comprehensive Pneumology Center (CPC) and Institute for Lung Biology and Disease (iLBD), University Hospital, Ludwig-Maximilian University (LMU) and Helmholtz Zentrum München, Member of the German Center for Lung Research (DZL); Max-Lebsche-Platz 31, 81377, Munich, Bavaria, Germany.

* **Corresponding Authors to whome request for reprints and correspondence should be addressed:** Noemi Castelletti and Jan Christian Kaiser; Institute of Radiation Protection, Helmholtz Zentrum München; Building 34, Ground Floor, Rooms 0110/0124, Ingolstädter Landstraße 1, 85764, Neuherberg, Germany; Phone: +49-089-3187-2133/4028; Fax: +49-089-3187-3363; E-mail: noemi.castelletti@helmholtz-muenchen.de, christian.kaiser@helmholtz-muenchen.de

Authors' contributions: N.C. performed most analyses, developed the mechanistic model, wrote parts of the paper draft, edited and finalized the submitted version of the paper; J.C.K. conceived the main idea and supervised the study, wrote parts of the paper draft; K.F. analyzed LSS data and helped developing the descriptive models; H.K. guided the statistical analysis of survival data and wrote parts of the paper; and G.T.S. performed molecular analyses, wrote portions of the paper draft and helped editing and finalizing the submitted version of the paper. All authors edited and concur with the submitted manuscript.

Financial support: N.C. was supported by the PASSOS project (FKZ 02NUK026, Personalisierte Abschätzung von Spätfolgen nach Strahlenexposition und Orientierungshilfe für Strahlenanwendungen in der Medizin) funded by the German Federal Ministry of Education and Research (BMBF) in the framework of Kompetenzverbund Strahlenforschung (KVSF). G.T.S. was supported by European Research Council 2010 Starting Independent Investigator and 2015 Proof of Concept Grants (grant numbers # 260524 and # 679345, respectively).

Running head: Radiation and receptor mutations in lung adenocarcinoma

Descriptors: 9.28 Lung Cancer: Epidemiology

Word count, manuscript: 2,993

At a Glance Commentary:

Conflict of interest: The authors declare no conflict of interest.

ABSTRACT

Rationale: Although KRAS mutations of lung adenocarcinoma (LADC) are associated with smoking, little is known on other exposure-oncogene associations and on the cause of LADC in never smokers. We aimed to identify distinct molecular pathways to LADC hypothesizing that different inciting agents cause different driver mutations.

Objectives: To identify molecular pathways to LADC and to develop molecular risk prediction models.

Methods: We examined clinicopathologic features and genomic signatures of environmental exposures in a large LADC dataset (the Campbell dataset), designed a molecular mechanistic risk model (M3) of LADC, and applied it to incidence data of Japanese atom-bomb survivors.

Measurements and Main Results: Both clinical and model analysis identified two distinct molecular pathways to LADC: one unique to transmembrane receptor-mutant patients that displayed robust signatures of radiation exposure and one shared between sub-membrane transducer-mutant patients and patients with no evident driver mutation that carried the signature of smoking. M3 predicts receptor and transducer mutation frequencies and, together with molecular evidence, supports the unknown link between receptor-mutant LADC with radiation.

Conclusions: Using molecular and exposure data from two unique patient cohorts, we show that receptor-mutant lung adenocarcinomas of never smokers are likely caused by radiation along a molecular pathway distinct from that of smokers. In addition, we develop molecular mechanistic models for prediction of lung adenocarcinoma risk from smoking and radiation.

Word count, Abstract: 219

MANUSCRIPT TEXT

INTRODUCTION

Word count, introduction: 364

Lung adenocarcinoma (LADC) is the number one cancer killer worldwide (1, 2). LADC is mainly caused by tobacco smoke, but also occurs in never smokers possibly due to both anthropogenic and environmental irradiation exposures (3-5). The comprehensive genomic characterization of LADC from Caucasian and Asian patients has identified mutations in major driver oncogenes such as *KRAS*, *EGFR*, and others, with different frequencies observed in different populations (6, 7). However, a biological concept explaining the relative contributions of cigarette smoking and radiation exposures to LADC incidence, as well as the reason for the different mutation frequencies in different patient populations remain unknown (3, 5).

State-of-the-art epidemiological risk estimates from smoking and radiation exposure merely establish statistical associations without explicitly considering pathogenic processes and molecular data: molecular biology and epidemiology lack a common interface (8). Here we bridge this gap by applying molecular mechanistic models (M^3) of carcinogenesis as tools to harness molecular data of LADC. M^3 treat carcinogenesis as a progression of cell-based key events on the pathway to malignancy and can detect in cancer incidence imprints from molecular events on recorded hazard or survival rates (9).

Using comprehensive genomic datasets from Eastern and Western LADC patient populations, we determine two distinct molecular pathways to LADC: one unique to *EGFR*-, and other transmembrane receptor-mutant (R^{MUT}) patients and one shared between *KRAS*-, and other sub-membrane transducer-mutant (T^{MUT}) patients. While T^{MUT} tumor tissues expectedly display

genomic signatures of tobacco smoke exposure, we identify for the first time genomic imprints of radiation exposure in R^{MUT} patients with LADC.

Deploying unparalleled information of smoke and irradiation exposure from the Life Span Study (LSS) of Japanese atomic bomb survivors, we develop M^3 for LADC-risk-estimation by molecular pathway. M^3 for LADC (M^3_{LADC}) accurately reproduces the observed LADC incidence in the LSS with moderately improved goodness-of-fit compared to standard epidemiological models. Amazingly, M^3_{LADC} predicts for the first time the different mutation frequencies actually observed in different populations (10), a fact open to direct validation since for the LSS genomic data are not yet available. Importantly, M^3_{LADC} harnesses firm biological evidence for a close association between R^{MUT} LADC with environmental radiation and T^{MUT} LADC with smoking for the explanation of observational data (11).

METHODS

Statistical analysis of molecular data: Mutation rates of 660 patients with LADC from the US (6) and 101 from China (7) were extracted from the primary publications. Individual patient clinical, exposure, and mutation data from the US cohort were downloaded from the primary publication (6) and manually analyzed. Clinicopathologic and molecular data from (6) and (7) were examined for normality by Kolmogorov-Smirnov test, were found to be not normally distributed, and are hence presented as median with Tukey's whiskers (boxes: interquartile range; bars: 50% extreme quartiles) and raw data points (dots). Differences in frequencies were examined by Fischer's exact or χ^2 tests, and in medians of non-normally distributed variables by Kruskal-Wallis non-parametric analyses of variance with Dunn's post-tests. Survival was

examined by Kaplan-Meier estimates with log-rank tests. Probability (*P*) is two-tailed and $P < 0.05$ was considered significant. Statistics and plots from clinicopathologic and molecular data were done on Prism v5.0 (GraphPad, La Jolla, CA). Univariate multinomial regression analysis of clinicopathologic and molecular data from (6) stratified by molecular pathways using T^{MUT} as the reference category due to its largest size was done with R* (https://www.r-project.org/).

The LSS cohort of Japanese atomic bomb survivors: The LSS cohort has been the primary epidemiological basis for evaluating the long-term health effects of radiation, dominated by 0–4 Gy gamma rays of low linear energy transfer (LET). It includes about 94,000 survivors who were in Hiroshima and Nagasaki at the time of bombing and about 27,000 who were temporarily away at that time, and whose mortality and cancer incidence have been followed up since 1950 and 1958, respectively (12). In the current analysis we used the imputed data of (13). To put the results of mechanistic modelling into perspective, a descriptive risk model (8) has been applied to the imputed data (see Supplementary Methods). Supplementary Table EI summarizes the LSS cohort data broken down by sex and smoking status.

Mechanistic risk modeling: Mechanistic models have long been applied for the analysis of radio-epidemiological cohorts (14). For the present study the two molecular pathways (R^{MUT} vs. T^{MUT}) to LADC determine the conceptual model design as fundamental feature. With this constrain and due to the fact that (15) argue that two/three driver mutations are involved in LADC, we considered only two- and three-stage clonal expansion models as candidates for both molecular pathways (Table EII). The conceptual design of the final preferred two-path LADC-model is shown in Figure 3A. Table EIII lists the parameters as means over 50 data sets for both pathways. Central risk estimates were calculated with the parameters of Table EIII. The model equations and explanations are outlined in Table EIV. Smoking and radiation exposure are

assumed to change biological parameters in mechanistic risk models. We tested actions on the rate ν of initiating mutations and the net clonal expansion rate γ using several functional forms: linear, linear-quadratic and linear-exponential responses. For smoking, model parameters were increased at smoking initiation and remained elevated for current smokers until end of follow-up. Baseline values were retained when past-smokers quit. Judged by goodness-of-fit, the main biological effects of smoking and radiation enhanced clonal expansion. In the T^{MUT} pathway, smoking intensity *smkint* linearly enhances the clonal expansion rate $\gamma_T = \alpha_T - \beta_T(S) - \mu_T = \gamma_{T0}[1 + g_S \cdot smkint \cdot \exp(-\kappa \cdot smkint)]$ during a period of constant smoking intensity with an attenuated effect for high smoking intensity. In the R^{MUT} pathway, a radiation dose D linearly enhances the clonal expansion rate $\gamma_R = \alpha_R - \beta_R(R) - \mu_R = \gamma_{R0}[1 + g_R \cdot D]$ after exposure for life. Since both pathways apply the same constant value of the stochasticity parameter $\delta = \alpha\mu$, increase of clonal expansion is solely caused by reduced cell inactivation. In any mechanistic analysis, standard epidemiological models (descriptive models) are indispensable to put the results of mechanistic model into perspective. Here, the descriptive model is inspired by the excess relative risk (ERR) of (8). An explanation of the model is presented in the Supplementary Methods and in Tables EV-EVI.

RESULTS

Identification of two causally and molecularly distinct pathways to LADC development

To identify possible clinical and/or molecular clusters of patients with LADC, we initially analyzed all data available from 660 Caucasian patients with LADC classified by driver oncogene (6). In addition to the available clinical information, total single nucleotide variant

(SNV) rates, insertion/deletion (indel) rates, copy number alteration (CNA) indices (calculated as the square root of the sum of all CNA squares of each tumor), as well as the contribution of established genomic signatures of environmental exposures were examined. These included a UV-related signature of C>T at TpCpC or CpCpC (COSMIC Signature 7, abbreviated SI7), a smoking-related signature of C>A transversions (SI4), a DNA mismatch repair (MMR) signature of C>T at GpCpG (SI15/SI6), two APOBEC-related signatures of C>G or C>T at TpCpT or TpCpA (SI13 and SI2), and a COSMIC signature 5 (SI5) with putative "molecular clock" properties (6, 16). In addition, we calculated the indel/SNV ratios, since such high ratios were found elsewhere to represent a direct molecular imprint of iatrogenic γ-radiation (γ-IR) (17). Grouping of the 660 patients by the most frequent drivers (every driver with $n \geq 10$ patients available was examined) revealed that patients with *EGFR* ($n = 86$), *ERBB2* ($n = 17$), *MET* ($n = 22$), and *ALK/RET/ROS1* (pooled to achieve $n = 14$) mutations [hereafter collectively referred to as receptor-mutant (R^{MUT})] were different from patients with *KRAS* ($n = 210$), *BRAF* ($n = 37$), *ARHGAP35* ($n = 13$), and *NF1* ($n = 58$) mutations [hereafter collectively referred to as transducer-mutant (T^{MUT})]. To this end, R^{MUT} patients displayed lower SNV and indel rates, and decreased smoking exposure evident by lower transversion rates and decreased activity of the smoking-related SI4 compared with T^{MUT} patients. At the same time, R^{MUT} patients were more frequently female, and displayed increased activities of UV light-related SI7, of DNA MMR-related SI15/SI6, and of SI5 putatively reflecting molecular clock properties compared with T^{MUT} patients. Interestingly, R^{MUT} patients had higher indel/SNV ratios compared with T^{MUT} patients, indicating a molecular signature of γ-IR exposure (17). CNA indices were comparable across patients with different drivers, except from *ALK/RET/ROS1*-fused patients that collectively displayed lower CNA indices compared with all other patients (Figures 1A, 1B). Based on this

finding, we grouped US patients (6) and 101 LADC obtained from Asian patients (7) into R^{MUT}, T^{MUT}, and oncogene wild-type (O^{WT}; patient without R^{MUT} or T^{MUT}) groups, hypothesizing that these three groupings may represent distinct molecular pathways to LADC (Figure 1C). Individual mutation prevalence varied widely between East and West, translating into different frequencies of these pathways in Caucasian and Asian LADC (Figure 1D). A fact that has to be taken into account since the molecular analysis is done with American patients and the model analysis with a Japanese cohort. We next sought to compare the molecular profiles of the three candidate molecular pathways LADC to identify potential similarities and differences. Interestingly, R^{MUT} LADC appeared distinct, while T^{MUT} and O^{WT} LADC were similar by all parameters examined except CNA index (Figures 2A, 2B). This was also evident from univariate multinomial logistic regression analyses that showed a general pattern of O^{WT} LADC trending with T^{MUT} LADC (Figure 2C). In the case of R^{MUT} LADC, 13 of the 18 analyzed covariables trended different from the reference category T^{MUT} with high significance (Figure 2C). These findings indicated the existence of two distinct molecular pathways to LADC that bear different genomic marks of environmental exposures: one unique to R^{MUT} patients that features robust imprints of γ-IR and the associated DNA MMR (18), and one shared between T^{MUT} and O^{WT} patients (hereafter referred to as T^{MUT}) with genomic marks of smoking exposure (Figure 2D). Interestingly, the R^{MUT} pathway contained patients with *ALK/RET/ROS1*-fusions, which were recently shown to dose-dependently culminate from γ-IR in thyroid cancer (19).

M^3_{LADC} for risk-prediction of LADC from smoking and radiation: M^3_{LADC} development is detailed in Materials and Methods and Supplementary Tables EII-EIV and is graphically represented in Figure 3A. M^3_{LADC} clearly revealed the two molecular pathways (R^{MUT} versus T^{MUT}) in observational incidence data of the LSS (Figure 3B) although no genomic information

of the LSS is available. For cigarette smoking, clonal expansion in the T^{MUT} pathway was identified as the main biological target: smoking-related inactivation of initiated cells increased the net clonal growth rate γ_T for pre-neoplastic lesions. Sex-specific response curves exhibited markedly different shapes (Supplementary Figure E1). For men, the growth rate increased almost linearly up to a smoking intensity of 20 cigarettes/day and flattened thereafter. Clonal growth in women reacted much stronger to low-smoking intensity. The growth reduction after a peak at about 10 cigarettes/day is biologically not plausible but might be caused by a reporting bias (Supplementary Figure E1). The main radiation effect occurred in the R^{MUT} pathway. An acute radiation pulse yielded a linear permanent increase of the net clonal expansion γ_R pointing to lifelong radiation-induced inflammation caused by genetic damage. Summarizing, the main impact of smoking and radiation took effect in distinct molecular pathways without noticeable synergy. For risk assessment, this particular biological action is better reflected in the excess absolute risk (EAR) compared to the excess relative risk (ERR). Figure EII presents a comparison of baseline hazard rates and hazard rates between the two molecular pathways. Figures 4 and 5 depict the EAR depending on smoking and radiation for pertinent exposure scenarios. In Supplementary Figures E3 and E4 the additive effect from both agents on the EAR is shown. Supplementary Figures E5 and E6 give the sex-specific ERR from smoking and radiation, respectively. Figure 6 presents a pathway-specific breakdown of expected LADC cases in different exposure groups for smoking and radiation.

DISCUSSION

LADC management and outcomes largely rely on tumor genotype (20). However, current prediction models of LADC do not provide molecularly stratified risks. We used molecular data from Caucasian and Asian patients with LADC to reveal two broad molecular fingerprints of the disease likely caused by different environmental exposures: one unique to patients with mutations in transmembrane receptors (R^{MUT}) featuring imprints of irradiation and another shared by patients with mutations in signal transduction genes and by patients with no known oncogene mutations (T^{MUT}) displaying the molecular signature of tobacco smoking. This information was applied to observational data of LADC incidence in Japanese atomic bomb survivors with known radiation/smoking exposure but unknown mutation status for the development of the first M^3_{LADC} for LADC risk prediction stratified by two modelled molecular pathways. The provided molecular risk prediction can be tested in the LSS in the future, and explains for the first time the different mutation frequencies in Eastern populations based on smoke and radiation exposures. More importantly, our combined genomic and epidemiologic analyses provide the first mechanistic link between irradiation exposure and receptor mutations in LADC, including *EGFR* mutations and *ALK/RET/ROS1* fusions.

Just like standard epidemiological risk models, M^3_{LADC} accurately reproduced LADC incidence in the LSS, albeit with moderately improved goodness-of-fit. Lubin et al. (21) analyzed a European lung cancer cohort with detailed smoking information using a generalized linear model in logistic regression. In their Figure 4, the sex-independent exposure response for LADC is measured in units of ERR/pack-year and shows remarkable agreement with our results for current male smokers (Supplementary Figure E5). As a striking new feature, M^3_{LADC} clearly identified the two molecular pathways that emerged from the molecular analysis. Importantly,

the predictive power of M^3_{LADC} can be subject to rigorous validation by future measurements of the mutation status in LADC tissue of LSS patients.

Previous molecular studies underpin the biological plausibility of M^3_{LADC}. *KRAS* mutations are more common in smokers (6) and are suspected to confer resistance to radiotherapy (22), which is consistent with the lack of a radiation response in the T^{MUT} pathway in our study. Thus, the main contribution of radiation to LADC incidence is imparted via the R^{MUT} pathway and a possible contribution from the T^{MUT} pathway is too small for quantification. To date, the risk factor that drives LADC development in never smokers is unknown, while these patients exhibit higher frequencies of *EGFR* mutations and *EML4/ALK* fusions (3-6). Here we show that radiation may drive disease development in these patients and provide a risk prediction model for this molecular class of LADC. The genomic signatures of radiation and of MMR of radiation-induced DNA strand breaks was enriched in R^{MUT} tumors (16). R^{MUT} tumors also displayed elevated indel/SNV ratios, shown elsewhere to be a hallmark of γ-IR-induced secondary cancers (17). Moreover, gene fusions such as *EML4/ALK*, *KIF5B/RET*, and *CD74/ROS1*, included here in the R^{MUT} pathway, have been linked with irradiation in other cancers (18, 19). These observations correspond to the radiation response of the R^{MUT} pathway as the most relevant radiation effect proposed by M^3_{LADC}. Hence we link for the first time radiation exposure to a molecular subset of LADC using molecular and epidemiologic evidence.

Smoking is linked with *KRAS*-mutant LADC and US patient analysis showed enhanced mutation rates in ever smokers of the T^{MUT} pathway (6). In a mechanistic model, this observation should generate an increase of initiating mutations in smokers. However, M^3_{LADC} works without such a plausible smoking effect because improvement in goodness-of-fit was inferior compared to smoking action on clonal expansion. Hence, the main biological mechanism of smoking on

LADC incidence is associated with enhanced clonal growth. Initiated cells exhibit a growth advantage over healthy cells due to reduced cell death possibly caused by smoking-associated chronic inflammation. Hence our data build on the known linkage between smoking and *KRAS*-mutant LADC by expanding this link to T^{MUT} and O^{WT} LADC, and by pinning the effects of smoke in time: at early time-points of smoking exposure. These results are relevant and important for the design of future chemoprevention strategies aimed to halt disease progression in smokers.

M^3_{LADC} also explains the higher susceptibility of women to smoke, evident by the current LADC pandemic in women (1). A study of *EGFR* and *KRAS* mutations in 3000 LADC of Caucasian patients revealed a higher susceptibility of women to smoking exposure for *KRAS*-mutant cancers (23). These findings are in line with a stronger increase of the smoking risk in the T^{MUT} pathway for female light smokers compared to male light smokers. Our results are concordant to the aforementioned study and can likely be explained by genetic predisposition of women to persistent smoke-induced DNA damage, notwithstanding the possibility for sex-related differences in innate immune responses to tobacco smoke and its carcinogens, as those observed in inbred strains of mice (24).

Risk prediction models, which are informed by adequate bioassays in addition to epidemiological variables, can predict lung cancer risk with high accuracy (25). They do lack, however, a link between environmental agents and molecular risk stratification, which is provided by $M^3_{LADC.}$ For example, this link suggests no elevated LADC risk even for heavy smokers in CT screening. It can be exploited in retrospective assessment to pin down the agent causing LADC based on the molecular profile of diseased tissue.

In conclusion, our study answers a longstanding question on the biological origins of age-risk patterns for LADC from concomitant exposure to smoking and radiation. To describe such patterns, standard epidemiological models must inevitably rely on a vague implementation of synergistic effects, which are commonly couched in mathematical terms as either “additive” or “multiplicative” sometimes with further generalizations (8, 26, 27). We have shown here that smoking and radiation drive the development of LADC along different molecular pathways with negligible interaction for doses below 4 Gy by projecting signatures of environmental exposure into epidemiological cohorts. The M^3_{LADC} approach provides a powerful tool for harnessing molecular data to improve studies of risk assessment and prediction in radiation protection and clinical applications. Our approach is of clinical relevance, because we solidify cause-effect relationships in LADC development by integrating molecular and epidemiologic data. The cause of LADC can be detected from their molecular alterations and the share of LADC with specific alterations can be predicted using the M^3_{LADC} model with huge medical and socioeconomic implications.

ACKNOWLEDGEMENTS

This work could not be done without a strong cooperation with the Radiation Effect Research Foundation (RERF) that provided the data and helped in the analysis. The RERF of Hiroshima and Nagasaki, Japan is a public interest foundation founded by the Japanese Ministery of Health, Labour and Welfare (MHLW) and the US department of Energy (DOE). The authors have no conflict of interest. We thank Cristoforo Simonetto and Ignacio Zaballa for fruitful discussions on mechanistic model conception and development.

REFERENCES

1. Global Burden of Disease Cancer C, Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, Brenner H, Dicker DJ, Chimed-Orchir O, Dandona R, Dandona L, Fleming T, Forouzanfar MH, Hancock J, Hay RJ, Hunter-Merrill R, Huynh C, Hosgood HD, Johnson CO, Jonas JB, Khubchandani J, Kumar GA, Kutz M, Lan Q, Larson HJ, Liang X, Lim SS, Lopez AD, MacIntyre MF, Marczak L, Marquez N, Mokdad AH, Pinho C, Pourmalek F, Salomon JA, Sanabria JR, Sandar L, Sartorius B, Schwartz SM, Shackelford KA, Shibuya K, Stanaway J, Steiner C, Sun J, Takahashi K, Vollset SE, Vos T, Wagner JA, Wang H, Westerman R, Zeeb H, Zoeckler L, Abd-Allah F, Ahmed MB, Alabed S, Alam NK, Aldhahri SF, Alem G, Alemayohu MA, Ali R, Al-Raddadi R, Amare A, Amoako Y, Artaman A, Asayesh H, Atnafu N, Awasthi A, Saleem HB, Barac A, Bedi N, Bensenor I, Berhane A, Bernabe E, Betsu B, Binagwaho A, Boneya D, Campos-Nonato I, Castaneda-Orjuela C, Catala-Lopez F, Chiang P, Chibueze C, Chitheer A, Choi JY, Cowie B, Damtew S, das Neves J, Dey S, Dharmaratne S, Dhillon P, Ding E, Driscoll T, Ekwueme D, Endries AY, Farvid M, Farzadfar F, Fernandes J, Fischer F, TT GH, Gebru A, Gopalani S, Hailu A, Horino M, Horita N, Husseini A, Huybrechts I, Inoue M, Islami F, Jakovljevic M, James S, Javanbakht M, Jee SH, Kasaeian A, Kedir MS, Khader YS, Khang YH, Kim D, Leigh J, Linn S, Lunevicius R, El Razek HMA, Malekzadeh R, Malta DC, Marcenes W, Markos D, Melaku YA, Meles KG, Mendoza W, Mengiste DT, Meretoja TJ, Miller TR, Mohammad KA, Mohammadi A, Mohammed S, Moradi-Lakeh M, Nagel G, Nand D, Le Nguyen Q, Nolte S, Ogbo FA, Oladimeji KE, Oren E, Pa M, Park EK, Pereira DM, Plass D, Qorbani M, Radfar A, Rafay A, Rahman M, Rana SM, Soreide K, Satpathy M, Sawhney M, Sepanlou SG,

Shaikh MA, She J, Shiue I, Shore HR, Shrime MG, So S, Soneji S, Stathopoulou V, Stroumpoulis K, Sufiyan MB, Sykes BL, Tabares-Seisdedos R, Tadese F, Tedla BA, Tessema GA, Thakur JS, Tran BX, Ukwaja KN, Uzochukwu BSC, Vlassov VV, Weiderpass E, Wubshet Terefe M, Yebyo HG, Yimam HH, Yonemoto N, Younis MZ, Yu C, Zaidi Z, Zaki MES, Zenebe ZM, Murray CJL, Naghavi M. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. JAMA Oncol 2017; 3: 524-548.

2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin 2015; 65: 87-108.

3. Alberg AJ, Brock MV, Ford JG, Samet JM, Spivack SD. Epidemiology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 2013; 143: e1S-29S.

4. Hecht SS. Tobacco smoke carcinogens and lung cancer. J Natl Cancer Inst 1999; 91: 1194-1210.

5. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers--a different disease. Nat Rev Cancer 2007; 7: 778-790.

6. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, Shukla SA, Guo G, Brooks AN, Murray BA, Imielinski M, Hu X, Ling S, Akbani R, Rosenberg M, Cibulskis C, Ramachandran A, Collisson EA, Kwiatkowski DJ, Lawrence MS, Weinstein JN, Verhaak RG, Wu CJ, Hammerman PS, Cherniack AD, Getz G, Cancer Genome Atlas Research N, Artyomov MN, Schreiber R, Govindan R, Meyerson M.

Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. Nat Genet 2016; 48: 607-616.

7. Wu K, Zhang X, Li F, Xiao D, Hou Y, Zhu S, Liu D, Ye X, Ye M, Yang J, Shao L, Pan H, Lu N, Yu Y, Liu L, Li J, Huang L, Tang H, Deng Q, Zheng Y, Peng L, Liu G, Gu X, He P, Gu Y, Lin W, He H, Xie G, Liang H, An N, Wang H, Teixeira M, Vieira J, Liang W, Zhao X, Peng Z, Mu F, Zhang X, Xu X, Yang H, Kristiansen K, Wang J, Zhong N, Wang J, Pan-Hammarstrom Q, He J. Frequent alterations in cytoskeleton remodelling genes in primary and metastatic lung adenocarcinomas. Nat Commun 2015; 6: 10131.
8. Egawa H, Furukawa K, Preston D, Funamoto S, Yonehara S, Matsuo T, Tokuoka S, Suyama A, Ozasa K, Kodama K, Mabuchi K. Radiation and smoking effects on lung cancer incidence by histological types among atomic bomb survivors. Radiat Res 2012; 178: 191-201.
9. Ruhm W, Eidemuller M, Kaiser JC. Biologically-based mechanistic models of radiation-related carcinogenesis applied to epidemiological data. Int J Radiat Biol 2017: 1-25.
10. Takamochi K, Oh S, Suzuki K. Differences in EGFR and KRAS mutation spectra in lung adenocarcinoma of never and heavy smokers. Oncol Lett 2013; 6: 1207-1212.
11. Moolgavkar SH. Carcinogenesis models: an overview. Basic Life Sci 1991; 58: 387-396; discussion 396-389.
12. Ozasa K, Shimizu Y, Suyama A, Kasagi F, Soda M, Grant EJ, Sakata R, Sugiyama H, Kodama K. Studies of the mortality of atomic bomb survivors, Report 14, 1950-2003: an overview of cancer and noncancer diseases. Radiat Res 2012; 177: 229-243.
13. Furukawa K, Preston DL, Misumi M, Cullings HM. Handling incomplete smoking history data in survival analysis. Stat Methods Med Res 2014.

14. Ruhm W, Eidemuller M, Kaiser JC. Biologically-based mechanistic models of radiation-related carcinogenesis applied to epidemiological data. Int J Radiat Biol 2017; 93: 1093-1117.

15. Tomasetti C, Marchionni L, Nowak MA, Parmigiani G, Vogelstein B. Only three driver gene mutations are required for the development of lung and colorectal cancers. Proc Natl Acad Sci U S A 2015; 112: 118-123.

16. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjord JE, Foekens JA, Greaves M, Hosoda F, Hutter B, Ilicic T, Imbeaud S, Imielinski M, Jager N, Jones DT, Jones D, Knappskog S, Kool M, Lakhani SR, Lopez-Otin C, Martin S, Munshi NC, Nakamura H, Northcott PA, Pajic M, Papaemmanuil E, Paradiso A, Pearson JV, Puente XS, Raine K, Ramakrishna M, Richardson AL, Richter J, Rosenstiel P, Schlesner M, Schumacher TN, Span PN, Teague JW, Totoki Y, Tutt AN, Valdes-Mas R, van Buuren MM, van 't Veer L, Vincent-Salomon A, Waddell N, Yates LR, Australian Pancreatic Cancer Genome I, Consortium IBC, Consortium IM-S, PedBrain I, Zucman-Rossi J, Futreal PA, McDermott U, Lichter P, Meyerson M, Grimmond SM, Siebert R, Campo E, Shibata T, Pfister SM, Campbell PJ, Stratton MR. Signatures of mutational processes in human cancer. Nature 2013; 500: 415-421.

17. Behjati S, Gundem G, Wedge DC, Roberts ND, Tarpey PS, Cooke SL, Van Loo P, Alexandrov LB, Ramakrishna M, Davies H, Nik-Zainal S, Hardy C, Latimer C, Raine KM, Stebbings L, Menzies A, Jones D, Shepherd R, Butler AP, Teague JW, Jorgensen M, Khatri B, Pillay N, Shlien A, Futreal PA, Badie C, Group IP, McDermott U, Bova

GS, Richardson AL, Flanagan AM, Stratton MR, Campbell PJ. Mutational signatures of ionizing radiation in second malignancies. Nat Commun 2016; 7: 12605.

18. Seki Y, Mizukami T, Kohno T. Molecular Process Producing Oncogene Fusion in Lung Cancer Cells by Illegitimate Repair of DNA Double-Strand Breaks. Biomolecules 2015; 5: 2464-2476.

19. Efanov AA, Brenner AV, Bogdanova TI, Kelly LM, Liu P, Little MP, Wald AI, Hatch M, Zurnadzy LY, Nikiforova MN, Drozdovitch V, Leeman-Neill R, Mabuchi K, Tronko MD, Chanock SJ, Nikiforov YE. Investigation of the Relationship Between Radiation Dose and Gene Mutations and Fusions in Post-Chernobyl Thyroid Cancer. J Natl Cancer Inst 2018; 110: 371-378.

20. Reck M, Heigener DF, Mok T, Soria JC, Rabe KF. Management of non-small-cell lung cancer: recent developments. Lancet 2013; 382: 709-719.

21. Lubin JH, Caporaso NE. Cigarette smoking and lung cancer: modeling total exposure and intensity. Cancer Epidemiol Biomarkers Prev 2006; 15: 517-523.

22. Wang M, Kern AM, Hulskotter M, Greninger P, Singh A, Pan Y, Chowdhury D, Krause M, Baumann M, Benes CH, Efstathiou JA, Settleman J, Willers H. EGFR-mediated chromatin condensation protects KRAS-mutant cancer cells against ionizing radiation. Cancer Res 2014; 74: 2825-2834.

23. Dogan S, Shen R, Ang DC, Johnson ML, D'Angelo SP, Paik PK, Brzostowski EB, Riely GJ, Kris MG, Zakowski MF, Ladanyi M. Molecular epidemiology of EGFR and KRAS mutations in 3,026 lung adenocarcinomas: higher susceptibility of women to smoking-related KRAS-mutant cancers. Clin Cancer Res 2012; 18: 6169-6177.

24. Stathopoulos GT, Sherrill TP, Cheng DS, Scoggins RM, Han W, Polosukhin VV, Connelly L, Yull FE, Fingleton B, Blackwell TS. Epithelial NF-kappaB activation promotes urethane-induced lung carcinogenesis. Proc Natl Acad Sci U S A 2007; 104: 18514-18519.

25. El-Zein RA, Schabath MB, Etzel CJ, Lopez MS, Franklin JD, Spitz MR. Cytokinesis-blocked micronucleus assay as a novel biomarker for lung cancer risk. Cancer Res 2006; 66: 6449-6456.

26. Furukawa K, Preston DL, Lonn S, Funamoto S, Yonehara S, Matsuo T, Egawa H, Tokuoka S, Ozasa K, Kasagi F, Kodama K, Mabuchi K. Radiation and smoking effects on lung cancer incidence among atomic bomb survivors. Radiat Res 2010; 174: 72-82.

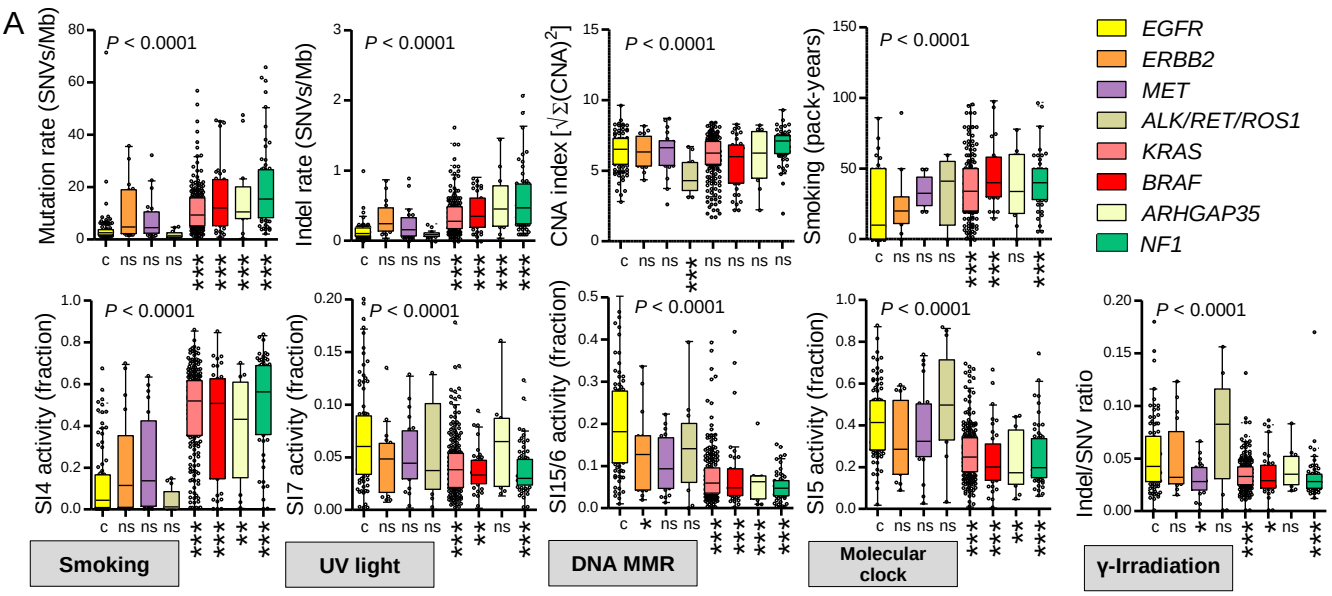
27. Cahoon EK, Preston DL, Pierce DA, Grant E, Brenner AV, Mabuchi K, Utada M, Ozasa K. Lung, Laryngeal and Other Respiratory Cancer Incidence among Japanese Atomic Bomb Survivors: An Updated Analysis from 1958 through 2009. Radiat Res 2017; 187: 538-548.

28. Yatabe Y, Borczuk AC, Powell CA. Do all lung adenocarcinomas follow a stepwise progression? Lung Cancer 2011; 74: 7-11.

FIGURE LEGENDS

Figure 1. Identification of broad molecular pathways to lung adenocarcinoma. (A, B) Single nucleotide variant (SNV) rates, insertion/deletion (indel) rates, copy number alteration (CNA) indices, smoking exposure, sex, genomic signatures of environmental carcinogen-induced base changes in the trinucleotide context (SI), indel/SNV ratios, and transversion status of 660 patients with lung adenocarcinoma (LADC) from the USA (6) grouped by the most frequent driver mutations. Significances $P \geq 0.05$, $P < 0.05$, $P < 0.01$, and $P < 0.001$ are coded as ns, *, **, and ***, respectively. (A) Data are given as raw data points, median ± Tukey's whiskers (lines: median; boxes: interquartile range; bars: 50% extreme quartiles). *P*, probabilities by Kruskal-Wallis test. Significances for comparison with *EGFR*-mutant control group (c) by Dunn's post-tests. (B) Data are given as number of patients (*n*). Color scale indicated frequency per row. *P*, probabilities by χ^2 test. Significances for comparison with *EGFR*-mutant control group (c) by χ^2 or Fischer's exact tests. Sample sizes were *EGFR* (*n* = 86), *ERBB2* (*n* = 17), *MET* (*n* = 22), *ALK/RET/ROS1* (pooled *n* = 14), *KRAS* (*n* = 210), *BRAF* (*n* = 37), *ARHGAP35* (*n* = 13), and *NF1* (*n* = 58). **(C)** Proposed grouping of US LADC patients (6) according to driver mutation into receptor-mutant (R^{MUT}), transducer-mutant (T^{MUT}), and oncogene-wild type (O^{WT}) molecular pathways. **(D)** Mutation rates and molecular pathway classification of 660 US LADC patients (6) and 101 LADC patients from China (7). *P*, probability by χ^2 test.

Figure 1



B

	Sex		Smoking			Transversion	
n	Male	Female	Never	Former	Current	Low	High
EGFR	18	68^{c}	41	34	8^{c}	69	17^{c}
ERBB2	8	9*	5	9	2^{ns}	11	6^{ns}
MET	13	9***	8	11	2^{ns}	14	8^{ns}
ALK/RET/ROS1	5	9^{ns}	8	5	1^{ns}	14	0^{ns}
KRAS	99	111***	11	142	44***	22	188***
BRAF	22	14***	0	23	9*	9	28***
ARHGAP35	5	8^{ns}	1	9	3^{ns}	3	10***
NF1	26	32**	1	31	19***	5	53***
χ² P	< 0.0001		< 0.0001			< 0.0001	

Frequency by driver (%): 0-20, 20-40, 40-60, 60-80, 80-100

C

EGFR, *ERBB2*, *MET*, *ALK/RET/ROS1*, *FGFR1*, *NTRK2*
Receptor-mutant R^{MUT} (*n* = 146)

KRAS, *BRAF*, *ARHGAP35*, *NF1*, *ARAF*, *HRAS*, *MAP2K1*, *NRAS*, *RAF1*, *RASA1*, *SOS1*, *VAV1*, *RIT1*
Transducer-mutant T^{MUT} (*n* = 353)

No known oncogenic driver
Oncogene wild-type O^{WT} (*n* = 161)

D

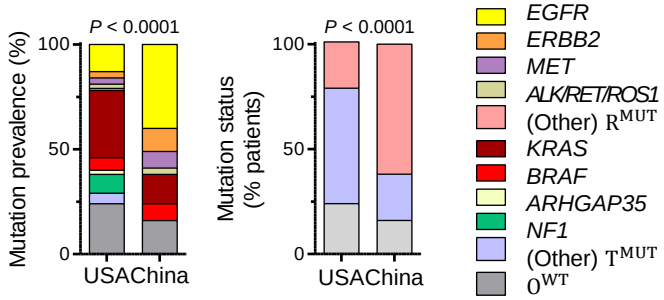
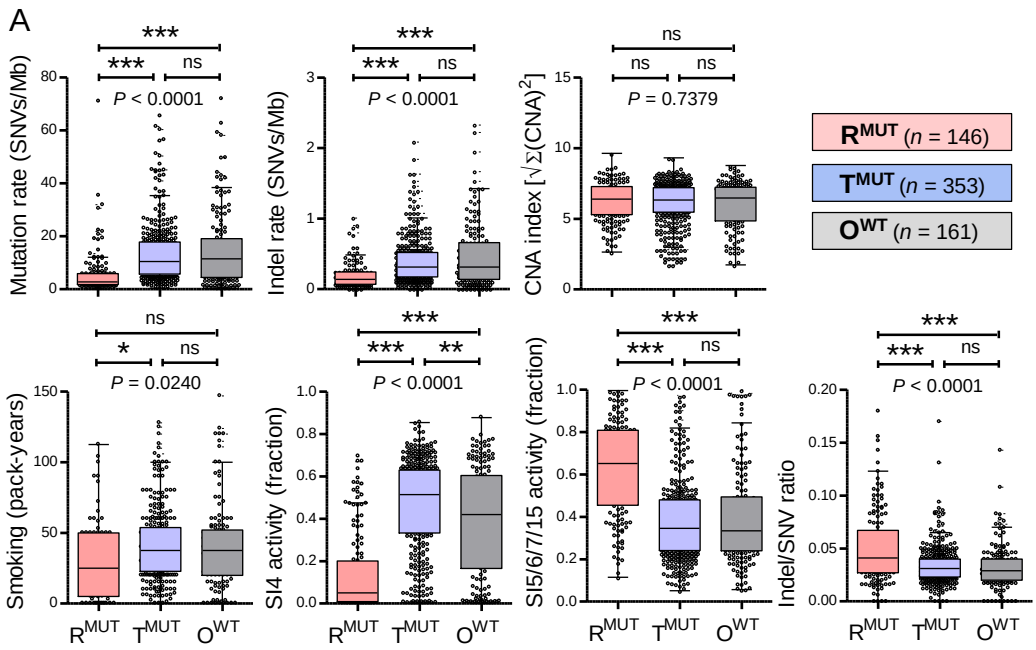


Figure 2. Clinical and molecular characteristics of 660 US LADC patients stratified by molecular pathway. (A, B) Single nucleotide variant (SNV) rates, insertion/deletion (indel) rates, copy number alteration (CNA) indices, smoking exposure, sex, genomic signatures of environmental carcinogen-induced base changes in the trinucleotide context (SI), indel/SNV ratios, and transversion status of 660 patients with lung adenocarcinoma (LADC) from the USA (6) grouped by receptor-mutant (R^{MUT}), transducer-mutant (T^{MUT}), and oncogene-wild-type (O^{WT}) molecular pathways. Significances $P \geq 0.05$, $P < 0.05$, $P < 0.01$, and $P < 0.001$ are coded as ns, *, **, and ***, respectively. (A) Data are given as raw data points, median ± Tukey's whiskers (lines: median; boxes: interquartile range; bars: 50% extreme quartiles). *P*, probabilities by Kruskal-Wallis test. Significances are given for the indicated comparisons by Dunn's post-tests. (B) Data are given as number of patients (*n*). Color scale indicated frequency per row. *P*, probabilities by χ^2 test. Significances are given for the indicated comparisons by χ^2 or Fischer's exact tests. Sample sizes were *EGFR* (*n* = 86), *ERBB2* (*n* = 17), *MET* (*n* = 22), *ALK/RET/ROS1* (pooled *n* = 14), *KRAS* (*n* = 210), *BRAF* (*n* = 37), *ARHGAP35* (*n* = 13), and *NF1* (*n* = 58). **(C)** Points represent regression coefficients divided by their standard errors in univariate multinomial regression. 18 clinical and molecular variables of 660 US patients with LADC (6) stratified by molecular pathway were analyzed. Position on x-axis denotes deviation from the estimate in reference group T^{MUT}. Significance of deviation from the reference is color-coded (red: R^{MUT}; black: O^{WT}): ns, *, **, and ***: $P \geq 0.05$, $P < 0.05$, $P < 0.01$, and $P < 0.001$, respectively, for the indicated variables. **(D)** Schematic of the two proposed molecular pathways to LADC and the main risk factors for each pathway.

Figure 2

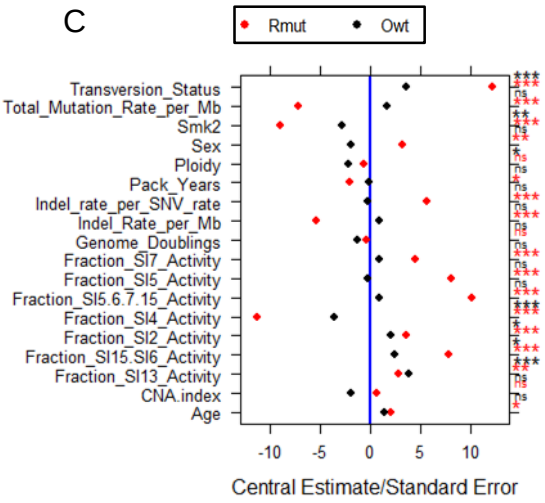


B

	Sex		Smoking			Transversion	
n	Male	Female	Never	Former	Current	Low	High
R^{MUT}	49	97[c]	62	63	15[c]	109	37[c]
T^{MUT}	174	178**	14	229	83***	44	309***
O^{WT}	94	67***	17	85	46***	41	120***
χ^2 *P*	< 0.0001			< 0.0001		< 0.0001	

0-20 | 20-40 | 40-60 | 60-80 | 80-100

Frequency by pathway (%)



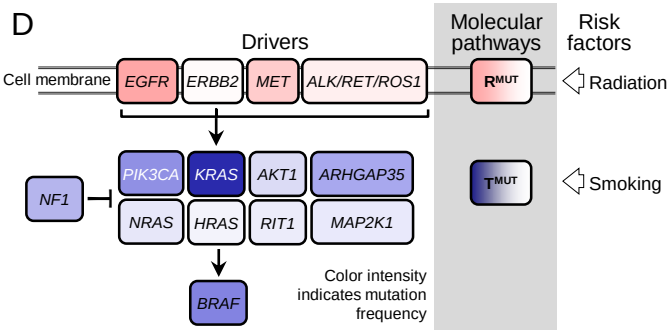
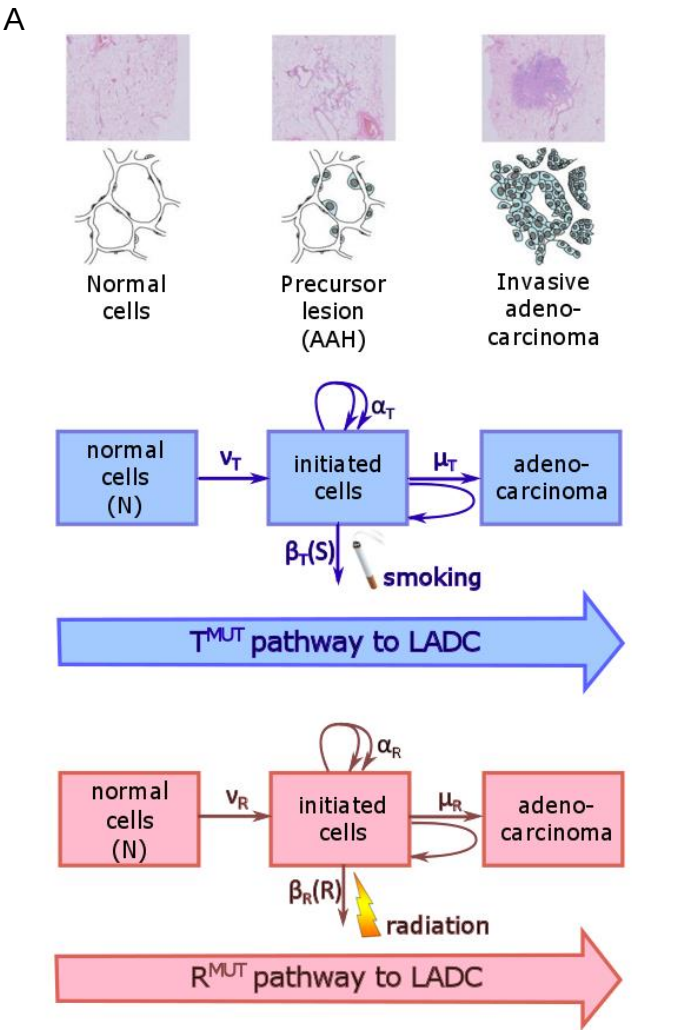


Figure 3. Molecular prediction of LADC risk stratified by molecular pathway. (A) Top: Histological progression from normal cells over atypical adenomatous hyperplasia (AAH) as precursor lesions to invasive LADC [modified figure from Yatabe, Borczuk and Powell (28)]. Bottom: Model implementation with two distinct molecular pathways pertaining to either T^{MUT} or R^{MUT} with two versions of the Two Stage Clonal Expansion (TSCE) model. Boxes represent cells in states with defined molecular properties. Arrows represent rates of transition between cell states. In both pathways a tiny fraction of a large number of N healthy cells incurs early molecular changes with yearly rates ν. Initiated cells may either divide symmetrically with rates α or become inactivated with rates β. The final transformation stage summarizes a sequence of complex processes with effective rate μ. Both agents of smoking and radiation cause the acceleration of clonal expansion by reduced cell inactivation. See model details in the Materials and Methods, mathematical model implementation and parameter estimates are given in the Supplementary Tables EIII and EIV. **(B)** Crude rate and predicted hazard (LADC cases in 10,000 persons per year) from the preferred mechanistic model (Mech) for the LSS cohort in 5 year-age groups from 40-45 up to 80-85 years. The model clearly distinguishes pathway-specific hazards. The hazard of R^{MUT}-related LADC cases peaks at age 70 yr. The hazard in the T^{MUT} pathway becomes dominant at old ages. This is a model prediction of the LSS cohort without any genomic data.

Figure 3


A
Normal cells
Precursor lesion (AAH)
Invasive adeno-carcinoma
normal cells (N)
initiated cells
adeno-carcinoma
smoking
T^MUT pathway to LADC
normal cells (N)
initiated cells
adeno-carcinoma
radiation
R^MUT pathway to LADC

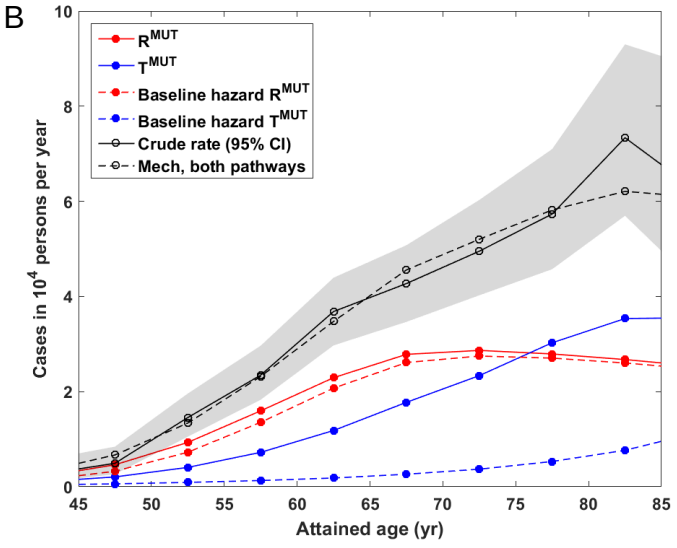

B
R^MUT
T^MUT
Baseline hazard R^MUT
Baseline hazard T^MUT
Crude rate (95% CI)
Mech, both pathways
Cases in 10^4 persons per year
Attained age (yr)

Figure 4. Excess absolute rates (EARs as cases in 10,000 persons per year) from M^3_{LADC} **(Mech) for smoking-induced lung adenocarcinoma (LADC) in the Japanese life span study (LSS) for lifelong smokers starting at age 20 yr.** The EAR is determined by the sex-dependent linear-exponential response to the smoking intensity which increases the clonal expansion rate in the T^{MUT} pathway independent of radiation (Supplementary Figure E1, top). To eliminate the influence for city of residence person-year weighted city means are used. Bivariate EAR dependence on attained age and smoking intensity for female smokers **(A)** and male smokers **(B)**. Panels **(C)** and **(D)** depict cross-sectional cuts to panels (A), (B) for attained ages of 50, 60 and 70 yr. Panels **(E)** and **(F)** depict cross-sectional cuts to panels (A), (B) for 5 cigs/day (males and females) and 20 cigs/day (males only). Female smokers of 5 cigs/day and male smokers of 20 cigs/day possess about the same risk. The EAR from a descriptive risk model (Desc) is shown for comparison.

Figure 4

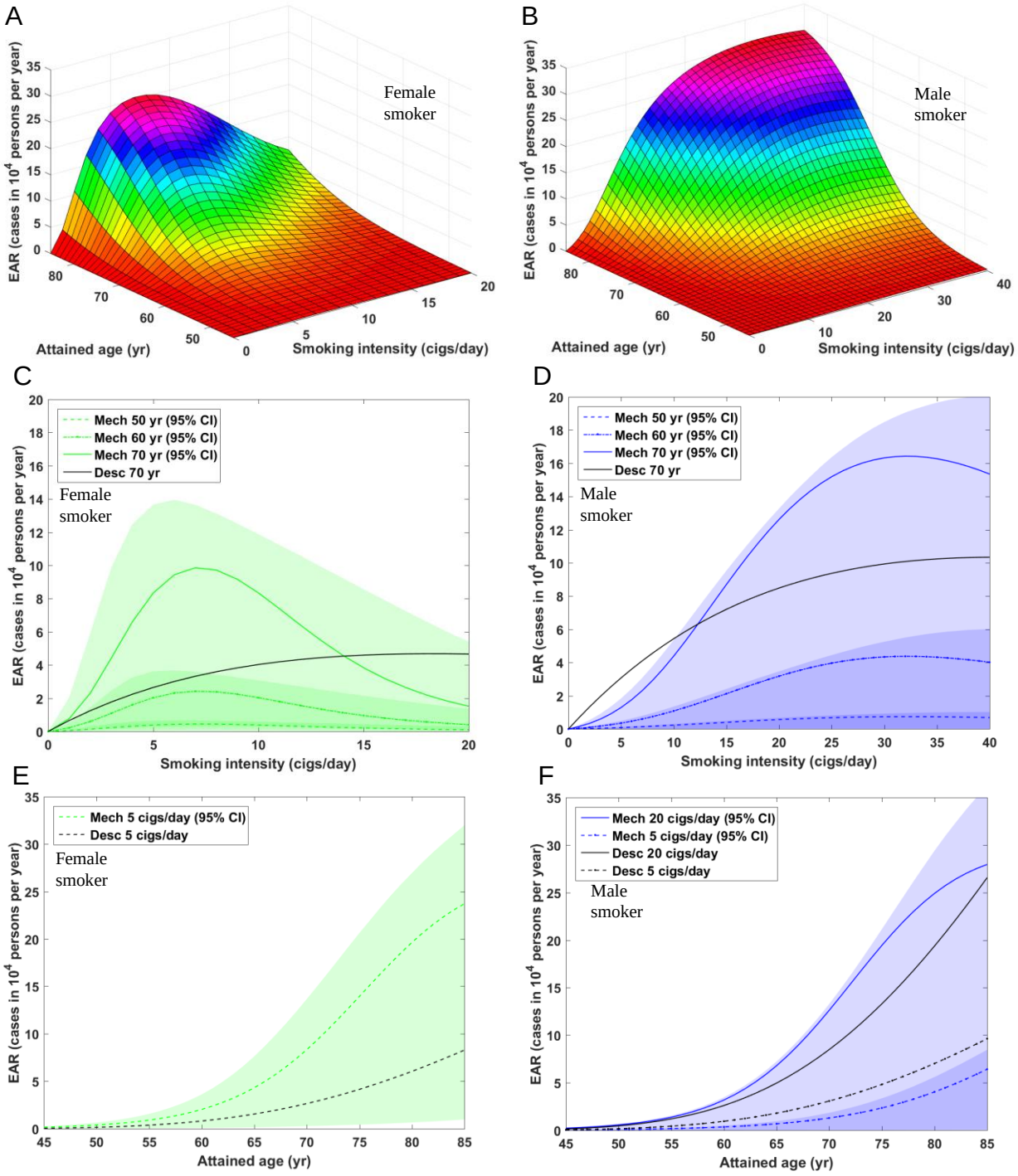

A
Female smoker
EAR (cases in 10^4 persons per year)
Attained age (yr)
Smoking intensity (cigs/day)
B
Male smoker
EAR (cases in 10^4 persons per year)
Attained age (yr)
Smoking intensity (cigs/day)
C
Mech 50 yr (95% CI)
Mech 60 yr (95% CI)
Mech 70 yr (95% CI)
Desc 70 yr
Female smoker
EAR (cases in 10^4 persons per year)
Smoking intensity (cigs/day)
D
Mech 50 yr (95% CI)
Mech 60 yr (95% CI)
Mech 70 yr (95% CI)
Desc 70 yr
Male smoker
EAR (cases in 10^4 persons per year)
Smoking intensity (cigs/day)
E
Mech 5 cigs/day (95% CI)
Desc 5 cigs/day
Female smoker
EAR (cases in 10^4 persons per year)
Attained age (yr)
F
Mech 20 cigs/day (95% CI)
Mech 5 cigs/day (95% CI)
Desc 20 cigs/day
Desc 5 cigs/day
Male smoker
EAR (cases in 10^4 persons per year)
Attained age (yr)

Figure 5. Excess absolute rates (EARs as cases in 10,000 persons per year) from M^3_{LADC} (Mech) for radiation-induced lung adenocarcinoma (LADC) in the Japanese life span study (LSS) cohort for a person exposed at 30 yr. The EAR is determined by the linear permanent response to an acute radiation pulse, which increases the clonal expansion rate in the R^{MUT} pathway independent of sex and smoking status (Supplementary Figure E1, bottom). To eliminate the influence for city of residence person-year weighted city means are used. **(A)** Bivariate EAR dependence on attained age and lung dose. The radiation risk maximizes at about 55 years for high lung dose. **(B)** Cross-sectional cuts to panel (A) for attained ages 50, 60 and 70 years. Over the dose range 0–4 Gy the EAR responds non-linearly to a lifelong radiation-induced linear response of the clonal expansion rate in the R^{MUT} pathway. **(C)** Cross-sectional cuts to panel (A) for lung doses 0.5, 1 and 2 Gy. The radiation-induced EAR peaks at decreasing age with increasing value. The EAR from a descriptive risk model (Desc) is shown for comparison.

Figure 5

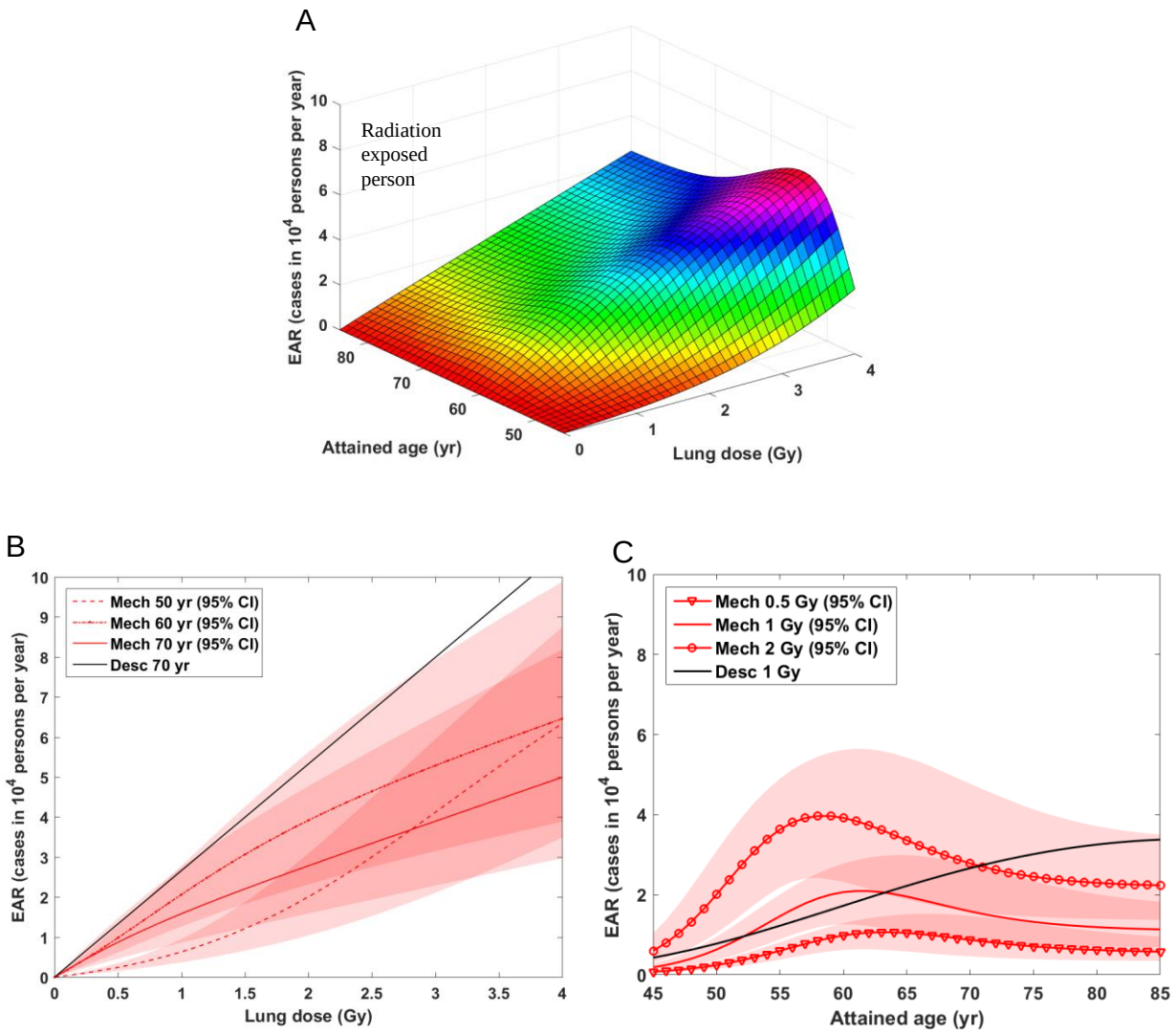

A
Radiation exposed person
EAR (cases in 10^4 persons per year)
Attained age (yr)
Lung dose (Gy)
B
Mech 50 yr (95% CI)
Mech 60 yr (95% CI)
Mech 70 yr (95% CI)
Desc 70 yr
EAR (cases in 10^4 persons per year)
Lung dose (Gy)
C
Mech 0.5 Gy (95% CI)
Mech 1 Gy (95% CI)
Mech 2 Gy (95% CI)
Desc 1 Gy
EAR (cases in 10^4 persons per year)
Attained age (yr)

Figure 6. M^3_{LADC} **estimates for the breakdown of 636 LADC cases (% of 636 cases) from the life span study (LSS) cohort in modelled molecular pathways R^{MUT} and T^{MUT} cross-tabulated with exposure groups for smoking and radiation**. Refined resolution in exposure subgroups of low (5-100 mGy) and moderate (100+ mGy) radiation dose, and light (1-10 cigs/day), moderate (11-20 cigs/day) and heavy (20+ cigs/day) smoking intensity is made. Female smokers fall mostly in the light category. In each subgroup observed cases are estimated well by the model. Exposure group numbers (bold-faced) add up to total numbers (bold-faced) in the bottom line. Exposure subgroup numbers add up to group numbers. Note that M^3_{LADC} estimates are derived from LADC incidence data in the LSS without genotyping. Model estimations for numbers and shares of cases in each molecular pathway would be directly accessible to measurements.

Figure 6

Radiation intensity (mGy)	Smoking intensity (cigs/day)	Observed cases (%)	Estimated cases (%)	R^{MUT} estimation (%)	Radiation induced estimated cases (%)	T^{MUT} estimation (%)	Smoking induced estimated cases (%)	Spontaneous estimated cases (%)
0-5	=0	**137 (21)**	**139 (22)**	**120 (19)**	**0 (0)**	**19 (3)**	**0 (0)**	**139 (22)**
5+	=0	**121 (19)**	**116 (18)**	**103 (16)**	**19 (3)**	**13 (2)**	**0 (0)**	**97 (15)**
5-100		57 (9)	61 (9)	53 (8)	1 (0)	8 (1)	0 (0)	60 (9)
100+		64 (10)	55 (9)	50 (8)	18 (3)	5 (1)	0 (0)	37 (6)
0-5	>0	**209 (33)**	**209 (33)**	**74 (12)**	**0 (0)**	**135 (21)**	**124 (20)**	**86 (13)**
	1-10	41 (7)	43 (7)	19 (3)	0 (0)	24 (4)	21 (3)	22 (3)
	10-20	109 (17)	105 (16)	37 (6)	0 (0)	68 (11)	63 (10)	42 (7)
	20+	59 (9)	61 (9)	18 (3)	0 (0)	43 (6)	40 (6)	21 (3)
5+	>0	**169 (27)**	**172 (27)**	**71 (11)**	**16 (3)**	**101 (16)**	**93 (14)**	**63 (10)**
5-100		*87 (14)*	*92 (14)*	*33 (5)*	*1 (0)*	*59 (9)*	*54 (8)*	*37 (6)*
	1-10	24 (4)	20 (3)	9 (1)	0 (0)	11 (2)	9 (1)	11 (2)
	10-20	43 (7)	45 (7)	16 (3)	1 (0)	29 (4)	27 (4)	17 (3)
	20+	20 (3)	27 (4)	8 (1)	0 (0)	19 (3)	18 (3)	9 (1)
100+		*82 (13)*	*80 (13)*	*38 (6)*	*15 (3)*	*42 (7)*	*39 (6)*	*26 (4)*
	1-10	21 (3)	18 (3)	10 (2)	3 (1)	8 (1)	9 (1)	6 (1)
	10-20	34 (6)	34 (6)	15 (2)	6 (1)	19 (4)	16 (3)	12 (2)
	20+	27 (4)	28 (4)	13 (2)	6 (1)	15 (2)	14 (2)	8 (1)
Total		**636 (100)**	**636 (100)**	**368 (58)**	**35 (6)**	**268 (42)**	**217 (34)**	**384 (60)**