Lung cancer susceptibility among atomic bomb survivors in relation to CA repeat number polymorphism of epidermal growth factor receptor gene and radiation dose

Kengo Yoshida*, Kei Nakachi, Kazue Imai, John B. Cologne¹, Yasuharu Niwa, Yoichiro Kusunoki and Tomonori Hayashi

Department of Radioepidemiology/Molecular Epidemiology and ¹Department of Statistics, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami Ward, Hiroshima City 732-0815, Japan

*To whom correspondence should be addressed. Tel: +81 82 261 3131; Fax: +81 82 261 3170; Email: kyoshi@rerf.or.jp

Lung cancer is a leading cause of cancer death worldwide. Prevention could be improved by identifying susceptible individuals as well as improving understanding of interactions between genes and etiological environmental agents, including radiation exposure. The epidermal growth factor receptor (EGFR)-signaling pathway, regulating cellular radiation sensitivity, is an oncogenic cascade involved in lung cancer, especially adenocarcinoma. The cytosine adenine (CA) repeat number polymorphism in the first intron of EGFR has been shown to be inversely correlated with EGFR production. It is hypothesized that CA repeat number may modulate individual susceptibility to lung cancer. Thus, we carried out a case–cohort study within the Japanese atomic bomb (A-bomb) survivor cohort to evaluate a possible association of CA repeat polymorphism with lung cancer risk in radiation-exposed or negligibly exposed (<5 mGy) A-bomb survivors. First, by dividing study subjects into Short and Long genotypes, defined as the summed CA repeat number of two alleles ≤37 and ≥38, respectively, we found that the Short genotype was significantly associated with an increased risk of lung cancer, specifically adenocarcinoma, among negligibly exposed subjects. Next, we found that prior radiation exposure significantly enhanced lung cancer risk of survivors with the Long genotype, whereas the risk for the Short genotype did not show any significant increase with radiation dose, resulting in indistinguishable risks between these genotypes at a high radiation dose. Our findings imply that the EGFR pathway plays a crucial role in assessing individual susceptibility to lung adenocarcinoma in relation to radiation exposure.

Introduction

Despite focused research over decades, lung cancer remains a leading cause of cancer deaths worldwide as well as in Japan (1). Effective prevention of lung cancer could be improved by identifying susceptible individuals as well as seeking a comprehensive understanding of lung carcinogenesis and potentially causal environmental agents. It is conceivable that lung cancer risk is influenced by low-penetration genetic variation, which may contribute to gene–environment interactions (2,3). Causal environmental factors include cigarette smoking and other factors, including exposure to ionizing radiation (4–6). Epidemiological studies on atomic bomb (A-bomb) survivors conducted by the Radiation Effects Research Foundation (RERF) have revealed a radiation dose-dependent increase in incidence of certain cancers, especially lung cancer (7).

Few molecular epidemiology studies have assessed interindividual variation in radiation sensitivity in relation to lung cancer development among A-bomb survivors. Our previous study showed that large interindividual variation in somatic mutability in response to prior A-bomb irradiation was associated with subsequent development of all cancers combined, in terms of erythrocyte glycophorin A-mutant frequency, which suggests the existence of genetic factors involved in radiation sensitivity and cancer susceptibility (8). Considering the implications of those results, we began searching for gene polymorphisms that might be responsible for individual susceptibility to radiation-associated lung cancer.

The initial event in the development of lung adenocarcinoma is thought to be alteration of the epidermal growth factor receptor (EGFR) gene (9). Mutations in EGFR are frequently observed in patients with lung adenocarcinoma (10,11). Gene amplification of EGFR is also frequently observed in patients with adenocarcinoma as well as squamous cell carcinoma of the lung, resulting in enhanced EGFR levels that lead to poor clinical prognosis and a chemo/radio-resistant phenotype (9,12,13). Therefore, the altered EGFR-signaling pathway, which turns on downstream pathways such as Ras-Raf-Mek and PI3K-Akt, is thought to contribute to the development of non-small-cell lung cancer in general through enhanced cell proliferation, inhibition of apoptosis, invasion and metastasis (4,9).

It is noteworthy that EGFR signaling facilitates cellular resistance to radiation exposure and that EGFR and its signaling cascade are induced by radiation exposure even in the absence of ligand binding (12,14,15). EGFR is thought to be a key molecule in the development of lung cancer in the general population as well as that of radiation-associated lung cancer found in A-bomb survivors. We therefore conducted a case–cohort study within a cohort of A-bomb survivors to assess the relationship between a functional polymorphism of EGFR and radiation-associated lung cancer.

The 5′-regulatory sequence of EGFR contains two functional gene polymorphisms that are markedly associated with the transcriptional activity of this gene: −216G/T and a highly polymorphic microsatellite sequence consisting of cytosine adenine (CA)-dinucleotide repeats (16,17). In this study, we focused on the CA repeat polymorphism in the first intron of EGFR because −216G/T variant alleles are infrequent in Asian populations (16,18). The number of CA repeats also shows substantial variation by ethnicity; larger numbers are found in East Asians than in persons of European descent or African-Americans (18–20). As demonstrated by in vitro and in vivo studies, CA repeat number is inversely correlated with messenger RNA or protein expression (21–23). Consequently, the association between CA repeat number and risk of various cancers has been examined by several groups, although the results on lung cancer have been inconsistent (24,25). This study aims to address: (i) whether EGFR CA repeat number is associated with risk of lung cancer development in the Japanese population who had not been exposed to A-bomb irradiation and (ii) how prior exposure to A-bomb irradiation influences the risk of radiation-associated lung cancer among A-bomb survivors in conjunction with the CA repeat polymorphism.

Subjects and methods

Study subjects

RERF's predecessor research organization, the Atomic Bomb Casualty Commission, established two major cohort studies to assess the health effects of A-bomb radiation exposure (26): the Life Span Study started in 1950 with ~120 000 members in Hiroshima and Nagasaki, including 93 000 A-bomb survivors and the Adult Health Study (AHS) including ~25 000 members who were selected from the Life Span Study and received biennial medical examinations beginning in 1958. The Immunology Study began in the AHS cohort in December 1981 with the aim of investigating radiation effects on the immune systems of A-bomb survivors; blood samples to examine immune-related biomarkers were collected from 9385 AHS participants who visited RERF for medical examinations and who donated blood samples for this study in 1981–2006. Additionally, 7131 persons among the 9385 AHS participants

Abbreviations: A-bomb, atomic bomb; AHS, Adult Health Study; CA, cytosine adenine; EGFR, epidermal growth factor receptor; Gy, Gray; PCR, polymerase chain reaction; RERF, Radiation Effects Research Foundation; RR, relative risk.
K. Yoshida et al.

Donated blood samples (peripheral lymphocytes and/or blood absorbed on paper disks) from which DNA could be extracted.

A total of 4764 participants were selected for the Immunogenome cohort study to assess relationships between cancer development and gene polymorphisms among A-bomb survivors, focusing on immune-related genes. The inclusion criteria were as follows: age <80 years at the time of blood collection; radiation dose information available; no prior cancer diagnosis at the time of blood collection and informed consent obtained for extracting and using DNA for research purposes (living members) or approval of the RERF Ethics Committee for Genome Research (deceased members who died prior to giving informed consent). Among the 4764 cohort members, 1061 incident cancers were identified from the Hiroshima and Nagasaki Tumor Registries, diagnosed between 1981 and 2001. Those included cancers of the stomach (n = 227), colon (n = 165), rectum (n = 53), liver (n = 115), lung (n = 124), breast (n = 90) and thyroid (n = 47).

Within the cohort, we defined a sub-cohort for this case–cohort study consisting of 2160 members who were randomly selected, a sampling rate of 0.45. A total of 486 incident cancers, including 62 lung cancer cases, were included in this sub-cohort. Cases were all 124 members who were diagnosed with lung cancer between 1981 and 2001. Their time of entry into the cohort was the year when the blood donation for the Immunology Study was first made. By histological type of lung cancer, the cases consisted of 66 (52.4%) adenocarcinomas, 21 (16.7%) squamous cell carcinomas, 10 (7.9%) small cell carcinomas and 29 (23.0%) other types including large cell carcinomas. In addition, two multiple cancer cases were included; those histological types were counted individually. This study was approved by the RERF Ethics Committee for Genome Research.

**Determination of CA repeat number**

The number of polymorphic CA repeats in the first intron of *EGFR* was determined by polymerase chain reaction (PCR)-based fragment length analysis using fluorescent primers along with DNA direct sequencing (18,24). First, genomic DNA was extracted from peripheral blood cells with proteinase K digestion and a QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany) and subjected to whole genome amplification (GenomiPhi DNA Amplification Kit, GE Healthcare, Little Chalfont, Buckinghamshire, UK). Then, 50–200 ng of the genómically amplified genomic DNA was extracted from peripheral blood cells with proteinase K digestion and a QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany) and subjected to whole genome amplification (GenomiPhi DNA Amplification Kit, GE Healthcare, Little Chalfont, Buckinghamshire, UK). Then, 50–200 ng of the amplified DNA was used for PCR with both carbocyanine fluorochrome-labeled and unlabeled forward primers (5'-GGGCTCAGCAGCAACTCTTC-3') and unlabeled reverse primers (5'-AAAGCCAGACTGCGCTGATG-3'). The 10 µl PCR reaction mixture contained 10× PCR buffer (Sigma–Aldrich, St Louis, MO), 0.5 U Taq DNA polymerase (Sigma–Aldrich), 0.2 mM each of deoxyxycytosine triphosphate, 2 mM MgCl₂, 200 mM each unlabeled primer and 20 nM labeled primer. PCR cycle conditions consisted of an initial denaturation step at 94°C for 5 min followed by 40 cycles of 30 s at 94°C, 30 s at 65°C and 60 s at 72°C, with a final elongation step at 72°C for 5 min. After 1 µl of PCR product and 0.5 µl Genescan 500 ROX molecular weight standard (Applied Biosystems, Foster City, CA) were denatured in 10 µl denaturing formamide, the number of CA repeats was determined using an ABI 3100 genetic analyzer (version 14.0, SPSS, Chicago, IL).

**Statistical analysis**

The data were sampled according to the case–cohort design, which does not require a rare disease assumption (27). The unweighted case-cohort approach was used for analysis (28). Relative risks for cancer incidence were estimated using the Cox proportional hazard model in terms of either follow-up (years) or age as the underlying time axis. Analyses were performed with SPSS (version 14.0, SPSS, Chicago, IL).

All models included adjustment for age at the time of blood collection, gender, city (Hiroshima versus Nagasaki), smoking status (number of cigarettes per day) and radiation dose. Information on smoking was collected at the time of blood collection. A-bomb radiation dose in Gray (Gy) was estimated using the DS02 dosimetry system (29), based on weighted skin dose computed as the gamma dose plus 10 times the neutron dose. A radiation dose <0.005 Gy was called non-exposed when performing analyses based on dose group.

**Results**

Table I shows characteristics of cases and sub-cohort members. Cases evidenced higher proportions of males, smokers and persons with the highest radiation doses compared with sub-cohort members.

**Table I. Characteristics of the study population within the RERF AHS cohort**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cases</th>
<th>Sub-cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>124 (100)</td>
<td>2160 (100)</td>
</tr>
<tr>
<td>Age at entry</td>
<td>59 (48–74)</td>
<td>56 (41–75)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>68 (54.8)</td>
<td>779 (36.1)</td>
</tr>
<tr>
<td>Female</td>
<td>56 (45.2)</td>
<td>1381 (63.9)</td>
</tr>
<tr>
<td>City</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hiroshima</td>
<td>84 (67.7)</td>
<td>1432 (66.3)</td>
</tr>
<tr>
<td>Nagasaki</td>
<td>40 (32.3)</td>
<td>728 (33.7)</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-smoker</td>
<td>49 (39.5)</td>
<td>1571 (72.7)</td>
</tr>
<tr>
<td>Smoker</td>
<td>75 (60.5)</td>
<td>627 (27.3)</td>
</tr>
<tr>
<td>Unknown</td>
<td>11 (8.9)</td>
<td>78 (3.6)</td>
</tr>
<tr>
<td>Radiation dose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5 mGy</td>
<td>41 (33.1)</td>
<td>906 (41.9)</td>
</tr>
<tr>
<td>5–712</td>
<td>37 (29.8)</td>
<td>627 (29.0)</td>
</tr>
<tr>
<td>≥712</td>
<td>46 (37.1)</td>
<td>627 (29.0)</td>
</tr>
<tr>
<td>Histological type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>66 (52.4)</td>
<td></td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>21 (16.7)</td>
<td></td>
</tr>
<tr>
<td>Small cell carcinoma</td>
<td>10 (7.9)</td>
<td></td>
</tr>
<tr>
<td>Other types</td>
<td>29 (23.0)</td>
<td></td>
</tr>
</tbody>
</table>

*aNumber (%). bMedian (5–95% percentiles). c712 mGy: median dose in exposed sub-cohort members. dTwo multiple primary cancer cases are included (one is adenocarcinoma and adenosquamous carcinoma, the other is adenosquamous carcinoma). Those histological types are counted individually.

In the analysis to calculate lung cancer risk, we used two methods to genotype the CA repeat polymorphism according to previous studies (25,30): method 1 sums repeat numbers of two alleles in each individual and defines his or her genotype as Long (sum ≥38) or Short (sum ≤37) (Figure 2); method 2 defines the two alleles separately as Long (allele repeat number ≥18) or Short (repeat number ≤17) and combines them for each individual as long/long, long/short or short/short. We report the results from method 1 genotyping in the text (for results from method 2 genotyping, refer supplementary Tables SI–SIII, available at Carcinogenesis Online).

As a preliminary to risk estimation, we confirmed that age, gender, city, smoking status and radiation dose did not influence the distribution of the CA repeat number in the sub-cohort population (data not shown). First, overall relative risk (RR) of lung cancer for the Short genotype in all subjects was evaluated by Cox regression analysis with adjustment for age, gender, city, smoking and radiation dose, taking the Long genotype as a reference: a significant increase in risk was found for the Short genotype (RR: 1.79, 95% CI: 1.14–2.82, Table II).

Next, we stratified the subjects by radiation dose (<5, 5–712 and ≥712 mGy) and evaluated the relative risks for the genotypes using two models: risk ratios for the genotypes within the same dose group (without considering radiation effects on risk) and relative risks for...
the genotypes as functions of radiation dose (Long genotype in the non-exposed population used as a reference). When looking at risk ratios (RR1 in Table III), the Short genotype revealed a significantly increased risk of lung cancer in the non-exposed population (RR: 1.76, 95% CI: 0.88–3.53 for 5–712 mGy and RR: 2.60, 95% CI: 1.36–4.96 for ≥712 mGy), whereas that for the Long genotype remained unchanged (RR: 2.43, 95% CI: 1.23–4.79). This genetic effect on risk disappeared in the exposed population. Relative risk for genotype combined with radiation dose (RR2 in Table III) showed that risk for the Long genotype increased with increasing radiation dose (RR: 1.76, 95% CI: 0.88–3.53 for 5–712 mGy and RR: 2.60, 95% CI: 1.36–4.96 for ≥712 mGy), whereas that for the Short genotype remained unchanged (RR: 2.11, 95% CI: 1.09–4.08 for 5 mGy and RR: 2.45, 95% CI: 1.22–4.92 for ≥712 mGy) independently of radiation exposure status or dose, resulting in almost identical relative risks for both genotypes at the highest dose.

Finally, we extended the analysis to histological type of lung cancer (adenocarcinoma or squamous cell carcinoma). With lung adenocarcinoma, there was a significant increase in risk for the Short genotype only in the non-exposed population (RR: 3.33, 95% CI: 1.36–4.96 for ≥712 mGy) as well as a radiation-associated elevation in risk for the Long genotype (RR: 2.60, 95% CI: 1.36–4.96 for ≥712 mGy). With lung squamous cell carcinoma, there were no significant risk differences among these genotypes (data not shown).

Discussion

We investigated the association between CA repeat number polymorphism in the EGFR gene and lung cancer risk using a case–cohort study setting, with the aim of elucidating risk of radiation-associated lung cancer in individuals with different EGFR genotypes. In the radiation dose-stratified analysis, we found that the Short genotype (the smaller sums of CA repeat numbers of two alleles) was significantly associated with an increased risk of lung cancer in survivors who had not been exposed to radiation (Table III). When we further examined the risk of lung cancer by histological types, adenocarcinoma, but not squamous cell carcinoma, showed a significant association between CA repeat genotypes (Long and Short) and lung cancer risk (Table IV). That is consistent with the clinico-pathological observation that adenocarcinoma is the most frequent histology in lung cancer of non-smokers, with frequent EGFR mutations detected.
EGFR e712 mGy: median dose in exposed sub-cohort members.

(30–50%) (31). The results suggest that the EGFR CA repeat number polymorphism may work as an indicator of individuals susceptible to lung cancer, specifically adenocarcinoma, in the Japanese population who have not been exposed to ionizing radiation. One limitation of this study is that no information about the EGFR mutation status of the cases was available. Population-based risk estimation of EGFR-mutated lung cancer for the CA repeat number polymorphism is therefore a goal of future research. Interestingly, the shorter CA repeat length has been reported to be associated with EGFR mutations among lung adenocarcinoma patients (23).

When looking at radiation dose effects on lung cancer risk, A-bomb survivors with the Long genotype were found to have a significant elevation of lung cancer risk that increased with increasing radiation dose, whereas the risk among those with the Short genotype did not change with radiation dose (Tables III and IV). As a result, the differences in risk between the Long and Short genotypes decreased with increasing radiation dose. Compared with the Short genotype, the Long genotype, presumably possessed by more than half of the Japanese population, is hypothesized to confer increased susceptibility to lung cancer after radiation exposure, although the baseline risk at no radiation dose is lower than that with the Short genotype. Our findings imply that individuals with the Long genotype, who are genetically at low risk of lung cancer when they are not exposed to radiation, may have to be more cautious about medical or occupational radiation exposure to minimize the cancer risk, which would otherwise show an elevation due to their high sensitivity to radiation.

We attempted two different methods of summarizing genotypes: summing CA repeat numbers of two alleles in each individual (resulting in Long or Short genotype) or combining allele-specific genotypes (resulting in long/long, long/short or short/short genotype). Although both methods produced similar results (Tables II–IV compared with supplementary Tables SI–SIII are available at Carcinogenesis Online), results from the latter method seem to be less clear, being in part due to an insufficient number of cases, especially of the short/short genotype. This also suggests that both alleles may combine to contribute to the risk of lung cancer, not in an allele-specific way (i.e. no preference for a contributory allele). In addition, analyses in this study were based on the weighted skin dose as an approximation for air dose. There were no substantial differences in the relative risks obtained using skin dose and lung dose, as shown by RR2 (2.45, 95% CI: 1.22–4.92, Table III) for Short genotype in the highest skin dose group compared with RR2 (2.45, 95% CI: 1.25–4.83) in the highest lung dose group.

Although the precise mechanism linking the CA repeat polymorphism and lung cancer risk remains to be elucidated, the shorter CA repeat length has been reported to be associated with increased EGFR transcription activity—supposedly by alteration in the repressor protein binding property (21,22,32)—and also with an increased EGFR mutation rate (23). Thus, the increased risk for the Short genotype observed in this study for non-exposed individuals may be in part ascribed to the inherent high EGFR production ability. Furthermore, it is probably that there is an undefined transcriptional mechanism that contributes to the more efficient EGFR production in individuals with the Long genotype than with the Short genotype as radiation dose increases, which results in an elevated lung cancer risk for the Long genotype with radiation dose. Further work, including in vitro studies, is required to establish the mechanistic link between the CA repeat polymorphism, EGFR production and radiation exposure in lung carcinogenesis.

**Supplementary material**

Supplementary Tables I–III can be found at http://carcin.oxfordjournals.org/

**Funding**


**Acknowledgements**

The Radiation Effects Research Foundation, Hiroshima and Nagasaki, Japan, is a private non-profit foundation funded by the Japanese Ministry of Health, Labour and Welfare and the US Department of Energy, the latter in part through the National Academy of Sciences.

**Conflict of Interest Statement**: None declared.

**References**


Received September 3, 2009; revised September 3, 2009; accepted October 8, 2009