Research Article

Cancer Epidemiology, Biomarkers & Prevention

Aristolochic Acid in the Etiology of Renal Cell Carcinoma

Margaret L. Hoang[1], Chung-Hsin Chen[2], Pau-Chung Chen[3], Nicholas J. Roberts[1], Kathleen G. Dickman[4,5], Byeong Hwa Yun[6], Robert J. Turesky[6], Yeong-Shiau Pu[2], Bert Vogelstein[1], Nickolas Papadopoulos[1], Arthur P. Grollman[4,5], Kenneth W. Kinzler[1], and Thomas A. Rosenquist[4]

Abstract

Background: *Aristolochia* species used in the practice of traditional herbal medicine contains aristolochic acid (AA), an established human carcinogen contributing to urothelial carcinomas of the upper urinary tract. AA binds covalently to genomic DNA, forming aristolactam (AL)–DNA adducts. Here we investigated whether AA is also an etiologic factor in clear cell renal cell carcinoma (ccRCC).

Methods: We conducted a population-based case–control study to investigate the linkage between *Aristolochia* prescription history, cumulative AA consumption, and ccRCC incidence in Taiwan (5,709 cases and 22,836 matched controls). The presence and level of mutagenic dA-AL-I adducts were determined in the kidney DNA of 51 Taiwanese ccRCC patients. The whole-exome sequences of ccRCC tumors from 10 Taiwanese ccRCC patients with prior exposure to AA were determined.

Results: Cumulative ingestion of more than 250 mg of AA increased risk of ccRCC (OR, 1.25), and we detected dA-AL-I adducts in 76% of Taiwanese ccRCC patients. Furthermore, the distinctive AA mutational signature was evident in six of 10 sequenced ccRCC exomes from Taiwanese patients.

Conclusions: This study strongly suggests that AA contributes to the etiology of certain RCCs.

Impact: The current study offers compelling evidence implicating AA in a significant fraction of the RCC arising in Taiwan and illustrates the power of integrating epidemiologic, molecular, and genetic data in the investigation of cancer etiology. *Cancer Epidemiol Biomarkers Prev; 25(12); 1600–8. ©2016 AACR.*

Introduction

Genome-wide sequencing has the capacity to link mutational signatures to specific mutagenic agents. Notable examples include C-to-T transitions in pyrimidine dimers induced by UV radiation (1) and C-to-A transversions induced by tobacco exposure (2). A recent bioinformatics survey of tumor types occurring primarily in Western populations identified at least 21 mutational signatures (3). However, for most of these putative signatures, the exogenous or endogenous agent responsible has not been identified. This gap could, in principle, be filled through molecular epidemiologic studies linking exposure to agents with the mutational signature observed.

Another remarkable example of a mutational signature associated with a specific carcinogen is provided by data on upper tract urothelial carcinomas (UTUC) in Taiwan and the Balkans. These tumors harbor a high content of A-to-T transversions, affecting adenines genome wide within a set of specific trinucleotide sequences (4, 5). This mutational signature has been linked to aristolochic acid (AA), a nitrophenanthrene carboxylic acid found in *Aristolochia* species (6) used worldwide in the practice of traditional herbal medicine (7). Following metabolic activation, AA forms aristolactam (AL)–DNA adducts that serve as specific biomarkers of exposure to AA (8). Remarkably, AL–DNA adducts can be detected in normal human tissues decades after exposure to AA (9–11). These adducts were also found in the normal kidney tissues of individuals in Taiwan (12), a country with one of the highest rates of UTUC in the world. In Taiwan, prescription data reveal that one in three residents had ingested herbs containing AA (13).

In rodents exposed to AA, promutagenic AL–DNA adducts are found in kidney, liver, forestomach, and bladder (14, 15). Thus, we hypothesized that AA contributes to cancers of the corresponding tissues in humans. In fact, signatures resembling the AA mutational signature have been reported to occur in bladder cancer (16), a small fraction of hepatocellular carcinomas in China (5, 17), Japan (18), and the United States (18), and in intrahepatic cholangiocarcinomas from China (19), and

[1]Ludwig Center for Cancer Genetics and Therapeutics and The Howard Hughes Medical Institute at Johns Hopkins Kimmel Cancer Center, Baltimore, Maryland. [2]Department of Urology, National Taiwan University Hospital and College of Medicine, Taipei, Taiwan. [3]Department of Occupational and Environmental Medicine, National Taiwan University Hospital and Institute of Occupational Medicine and Industrial Hygiene, National Taiwan University, Taipei, Taiwan. [4]Department of Pharmacological Sciences, Stony Brook University, Stony Brook, New York. [5]Department of Medicine, Stony Brook University, Stony Brook, New York. [6]Masonic Cancer Center and Department of Medicinal Chemistry, University of Minnesota, Minneapolis, Minnesota.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

M.L. Hoang and C.-H. Chen contributed equally to this article.

Corresponding Authors: Thomas A. Rosenquist, Department of Pharmacological Sciences, Stony Brook University, Nicolls Road, Stony Brook, NY 11794-8651. Phone: 631-444-8054; Fax: 631-444-3218; E-mail: Thomas.rosenquist@stonybrook.edu; and Kenneth W. Kinzler, Ludwig Center for Cancer Genetics and Therapeutics and The Howard Hughes Medical Institute at Johns Hopkins Kimmel Cancer Center, 1650 Orleans St., Rm. 520, Baltimore, MD 21231. E-mail: kinzlke@jhmi.edu

doi: 10.1158/1055-9965.EPI-16-0219

in clear cell renal cell carcinomas (ccRCC) in China (20), Romania (21), and the endemic region of Croatia (22). AL–DNA adducts have also been detected in kidney DNA from Romanian ccRCC patients (23). Thus, AA could be involved in the initiation and/or progression of tumors in these tissues. However, for hepatocellular carcinoma, intrahepatic cholangiocarcinoma, and ccRCC, integrated epidemiologic and molecular evidence linking AA exposure to the incidence of these cancers has not been available.

In the current study, we bridge this gap by applying both epidemiologic and molecular approaches to the Taiwanese population, where a high fraction of the population has been exposed to AA (13). Specifically, we provide strong evidence linking AA exposure to ccRCC pathogenesis in this population by demonstrating an association between AA exposure and ccRCC incidence, by measuring AL–DNA adducts in the renal cortex of ccRCC patients and by detection of the AA mutational signature in ccRCC tumors from patients previously exposed to AA.

Materials and Methods

The research protocols were reviewed and approved by the Institutional Review Boards of Stony Brook University (Stony Brook, NY) and the National Taiwan University (Taipei, Taiwan).

Prescription database analysis

The Taiwan National Health Insurance (NHI) covers >96% of Taiwanese residents and reimburses cost of prescriptions, including Chinese herbal products. RCC cases were identified from the NHI catastrophic illness registry using ICD-9 code 189.0. A total of 5,709 RCC patients and 22,836 controls were enrolled from 1999 to 2008. The index date for each case was the date of diagnosis with RCC, and we defined the index date of control subjects as the date of RCC diagnosis of their corresponding cases. RCC patients with other cancers (ICD-9 140–208) or kidney transplant (V42.0) were excluded from this study. Each case was paired with 4 controls randomly selected from the insured population, which were matched by sex, age, income, and urbanization.

The Taiwan NHI regularly reimbursed enrollees for the cost of prescribed Chinese herbs medicines containing AA, which were prescribed extensively before the ban in 2003. During our investigation period, the detailed herbal prescriptions, including regimen, dosage, duration, and prescription date were uploaded to NHI for obtaining payment. All prescribed herbal medications were covered by the Taiwan NHI and required a doctor's prescription. Usage of AA-containing herbs in each case and control was determined and calculated from Taiwan NHI database from January 1997 to October 2003. Cumulative AA exposure dosage was also calculated as described in a previous study that demonstrated a positive association between AA consumption and risk of urothelial carcinoma (24).

Known risk factors of RCCs were treated as potential confounders defined by the following diagnoses recorded between January 1, 1997, and 1 year before the diagnosis of RCCs or index dates: hypertension (ICD-9 401), diabetes (250), and hyperlipidemia (272), chronic obstructive pulmonary disease (491, 492, 496), chronic hepatitis C infection (070.7, 070.41, 070.44, 070.51, 070.54, V02.62), chronic kidney disease (585), cystic kidney disease (753.1), and kidney stones (592.0, 592.1, 592.9). Logistic regression was used to assess RCC risk based on the cumulative dose of AA. The ORs and 95% confidence intervals for RCC were calculated and estimated as crude and adjusted for covariates, including sex, age, monthly income, urbanization level, hypertension, diabetes, hyperlipidemia, chronic obstructive pulmonary disease, chronic hepatitis C infection, chronic kidney disease, cystic kidney disease, kidney stones, aspirin, NSAIDs, and acetaminophen. Of these covariates, only sex, monthly income, urbanization level, chronic obstructive pulmonary disease, chronic hepatitis C infection, aspirin, NSAIDs, and acetaminophen affected the results. All of these analyses were conducted using SAS statistical software (version 9.2; SAS Institute).

Genomic DNA for sequencing and adduct analysis

RCC patients undergoing radical nephrectomy at National Taiwan University Hospital (Taipei, Taiwan) between December 1998 and May 2007 were enrolled. Patients with previous radiotherapy or systemic chemotherapy were excluded. Tissue specimens were sampled immediately after surgery. Renal cortical samples were taken from normal-appearing cortex far-removed from renal tumors. Tumor tissues were sampled so as to avoid inclusion of surrounding normal tissues. All fresh tissue samples were placed in aseptic vials, snap frozen in liquid nitrogen, and then stored at −80°C till DNA extraction as described previously (12).

Whole-exome sequencing

Sequencing and mutational analysis methods are in Supplementary Material and as described by Hoang and colleagues (4).

Mass spectrometric determination of AL–DNA adducts

AL–DNA adduct concentrations in 5 μg of kidney DNA were determined by ultraperformance liquid chromatography–electrospray ionization/multistage scan mass spectrometry as described previously (25, 26).

Results

Evidence of exposure to AA-containing herbs in patients with ccRCC in Taiwan

In Taiwan, 39% of the population received prescriptions for *Aristolochia*-containing remedies during the years 1997 to 2003 (13). The results (Table 1) indicate an adjusted OR of 1.25 (1.004–1.547) for ccRCC in persons consuming more than 250 mg of AA during the period of 1997 to 2003. Each 100 mg of AA consumed contributed 1.03 (1.005–1.054) to the OR. As our study only monitored exposure during the 6-year period for which prescription records are available, it is likely that OR underestimates the impact of AA on RCC.

AL–DNA adduct levels in renal tissues of ccRCC patients

Molecular evidence that Taiwanese ccRCC patients were exposed to AA was obtained using a quantitative mass spectrometric method to measure dA-AL-I-DNA adduct levels. Genomic DNA was isolated from the nonneoplastic renal cortical tissue of 51 Taiwanese RCC patients (Supplementary Table S1). AA exposure induces deoxyadenosine (dA) and deoxyguanosine (dG) adducts derived from AAI and AAII (Fig. 1); the dA-AL-I adduct is resistant to DNA repair and can be detected in renal cortex DNA decades after exposure. We detected AL–DNA adducts in 39 of 51 (76%) patients tested (Fig. 1). In samples

Table 1. Population-based case–control study of consumption of herbal products containing AA and ccRCC incidence in Taiwan

Chinese herbal products containing aristolochic acid	Cases ($n = 5{,}709$) No. (%)	Controls ($n = 22{,}836$) No. (%)	Crude OR (95% CI)	Adjusted OR[a] (95% CI)
All patients				
No	3,520 (61.66)	14,281 (62.54)	1	1
Yes	2,189 (38.34)	8,555 (37.46)	1.038 (0.978–1.102)	0.977 (0.918 to 1.04)
Cumulative AA, mg				
0	3,520 (61.66)	14,281 (62.54)	1	1
1–125	1,937 (33.93)	7,681 (33.64)	1.023 (0.962–1.089)	0.967 (0.907–1.032)
126–250	134 (2.35)	527 (2.31)	1.032 (0.85–1.251)	0.936 (0.769–1.14)
>250	118 (2.07)	347 (1.52)	1.38 (1.116–1.706)	1.246 (1.004–1.547)
Each 100 mg increase			1.04 (1.016–1066)	1.029 (1.005–1.054)

Abbreviation: CI, confidence interval.

[a]Adjusted for monthly income, urbanization level, hypertension, diabetes, hyperlipidemia, chronic obstructive pulmonary disease, chronic hepatitis C infection, chronic kidney disease, cystic kidney disease, kidney stones, sickle cell disease, aspirin, NSAIDs, and acetaminophen.

with detectable amounts of adducts, levels ranged from 0.3 to 258 adducts per 10^8 bases; the average and median values were 36.3 and 20.8 adducts per 10^8 bases, respectively. Thus, these data indicate that at least 76% of the patients in our cohort consumed AA-containing herbs. Importantly, in these individuals, dA–AL adducts were found within the same tissue that generated the tumor.

Whole-exome sequencing of ccRCCs

We hypothesized that if AA exposure directly contributed to ccRCC tumorigenesis, then the tumor genome should harbor the characteristic AA mutational signature. To test this hypothesis, we performed whole-exome sequencing on tumor and nontumor DNA pairs from 5 men and 5 women who had been exposed to AA. Each patient selected had significant amounts of AL–DNA adducts in their renal DNA (Fig. 1). From the sequencing data, we identified an average of 22,158 known SNPs per individual (Supplementary Tables S2–S4). We next estimated the neoplastic cell content of the samples by determining the average mutation allele fraction across all genes. The median estimate of neoplastic cell content was 44% (range, 26–78). Tumor purity in ccRCC samples is generally low, as reported in

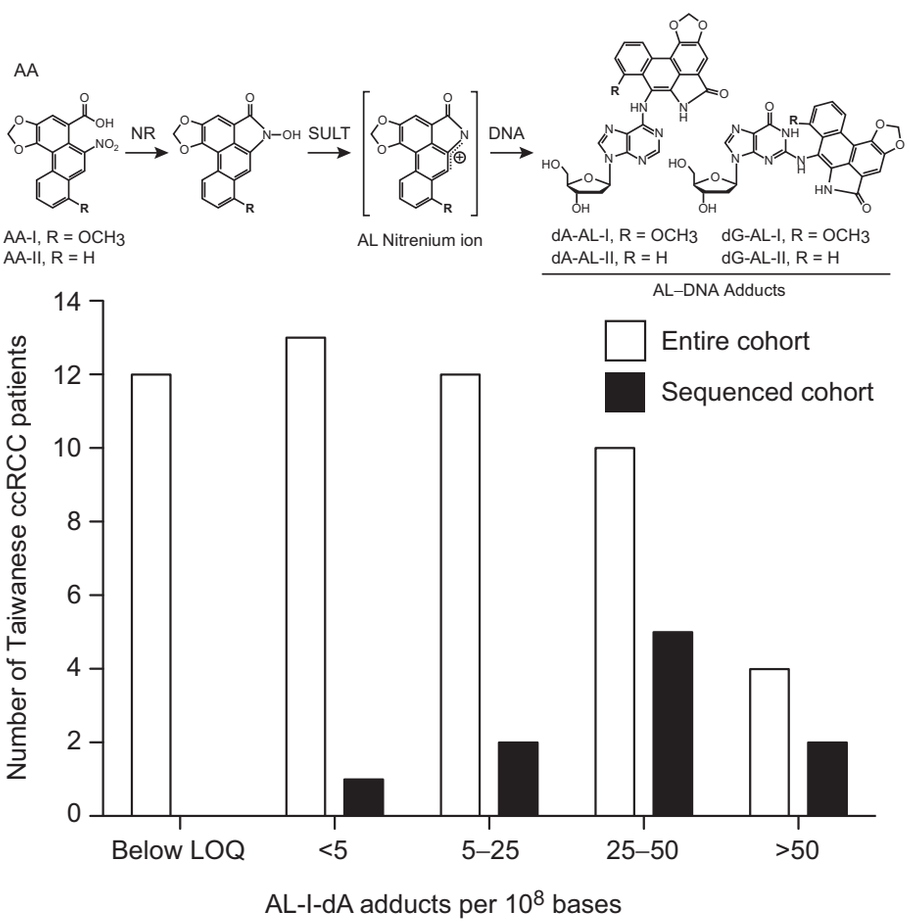

Figure 1.
Direct measurement of AL–DNA adducts in kidney DNA of Taiwanese ccRCC patients. Top, diagram of metabolic activation of AA by nitroreductase (NR) and sulfotransferase (SULT). AL nitrenium ion intermediates react with DNA to generate AL–DNA adducts at dA or dG; bottom, histogram of amount of dA-AL-I adduct (7-(deoxyadenosine-N^6-yl) aristolactam I) determined by UPLC-ESI/MS^3 analysis of nucleotide digestions of nontumor kidney DNA. White bars, all patients in the cohort ($n = 51$ patients); black bars, 10 patients with tumors sequenced in this study. LOQ, limit of quantification.

Table 2. Characteristics of Taiwanese ccRCCs

Sample	Patient age (years)	Gender	AL-I-dA adducts per 10^8 nts	% neoplastic content[a]	Number of somatic mutations		*VHL* mutation status		Other driver alterations at 3p21-25			3p copy number change	AA mutational signature			
					SBS	Indels	Nucleotide (hg19)	Protein consequence	Gene	Nucleotide (hg19)	Protein consequence		% A:T>T:A	A>T nontranscribed-to-transcribed ratio	% of splice acceptor mutations	AA signature in tumor?
RCC6	44	M	21.2	78%	323	11	chr3:10183788C>A	86P>H	ND	ND	ND	LOH	66%	2.3 (155/67)	1.8%	Yes
RCC48	60	F	137	52%	225	14	chr3:10191497C>T	164Q>X	*PBRM1*	chr3:52661296A>T	480R>X	LOH	62%	1.9 (97/50)	2.5%	Yes
									SETD2	chr3:47084192A>T	IVS5-2; splice acceptor					
									BAP1	chr3:52436414C>T	694Q>X					
RCC93	43	F	49.9	32%	18	0	ND	ND	ND	ND	ND	LOH	44%	1.0 (4/4)	0%	Yes
RCC90	64	M	41	44%	124	8	chr3:10191543-1019155delACATCGTCA[b]	Frameshift	ND	ND	ND	LOH	44%	1.5 (35/23)	4.5%	Yes
RCC53	67	F	6.9	42%	83	4	chr3:10191489delC	Frameshift	ND	ND	ND	LOH	39%	2.4 (24/10)	4.6%	Yes
RCC65	59	M	106.1	45%	70	3	chr3:10188204delA	Frameshift	ND	ND	ND	LOH	26%	2.8 (14/5)	4.1%	Yes
RCC49	45	M	38	26%	24	2	ND	ND	ND	ND	ND	ND	12%	2.0 (2/1)	0%	No
RCC67	50	F	27.5	34%	70	8	chr3:10188196A>T[b]	IVS2-2; splice acceptor	ND	ND	ND	ND	10%	0.6 (3/5)	0%	No
RCC15	84	M	2.3	62%	81	7	chr3:10191475A>T	156Y>X	*PBRM1*	chr3:52598223delC	Frameshift	ND	9%	0.3 (2/6)	1.1%	No
RCC86	54	F	44.7	29%	123	5	ND	ND	ND	ND	ND	ND	6%	1.0 (4/4)	0.8%	No

Abbreviations: F, female; LOH, loss of heterozygosity; M, male; ND, not detected; nts, nucleotides.
[a]Estimated from exome sequencing data using (distinct mutation count/distinct coverage)*2*100 averaged across all genes.
[b]Identified from Sanger sequencing data.

The Cancer Genome Atlas (TCGA) exome sequencing study of 417 ccRCCs (median 54% ± 14%; ref. 27). Because of the high coverage of our whole-exome sequencing, this did not pose a limitation to the identification of somatic mutations. The average high quality coverage of each base in the targeted region was 96-fold; 93% of the targeted region contained at least 10 reads. Furthermore, we observed no correlation between neoplastic cell content and the total number of somatic mutations identified for each tumor (Table 2).

Identification of somatic mutations

We used stringent criteria to identify a total of 1,204 somatic mutations with a median of 87 per tumor (range, 18–334; Fig. 2A; Supplementary Table S2). For each tumor, any mutation in the *VHL* and *PBRM1* genes, and six to nine randomly chosen

Figure 2.
Somatic mutational burden is elevated in AA-exposed ccRCCs. **A,** Scatterplot of the number of somatic substitutions in each ccRCC identified from whole-exome sequencing of AA-exposed Taiwanese ($n = 10$ tumors) and the publicly available TCGA dataset ($n = 417$ tumors). Red lines, median with interquartile range. **B,** The distribution of mutation consequences of AA-exposed ccRCC (black bars, $n = 1{,}204$ mutations) and TCGA ccRCC (white bars, $n = 26{,}193$ mutations).

mutations, were selected for validation by Sanger sequencing (Supplementary Table S3). For clonal somatic mutations with mutant fractions greater than or equal to 20%, 91% percent (46/51) of mutations crossvalidated, indicating a low false-positive rate. In contrast, 36 randomly chosen changes with mutant fractions below 20%, which are more likely to represent subclonal mutations, were validated in only 15 cases (41.7%). This lower validation fraction might reflect a higher false-positive rate for subclonal mutations but is more likely to reflect the lower sensitivity of Sanger sequencing. Of the total number of mutations, 94.8% (1,141/1,204) were single-base substitutions (SBS), while 5.2% were indels (63/1,204). Nonsynonymous mutations (protein altering) accounted for 78.9% of total mutations with a median of 65 nonsynonymous mutations per tumor (range, 14–278). The ratio of nonsynonymous-to-synonymous mutations was not significantly different between AA-exposed ccRCC at 3.7 (950/254) and TCGA ccRCCs at 3.5 (20,323/5,870; Fig. 2B, $P = 0.3$, two-sided χ^2), indicating that AA exposure did not significantly alter this genome-wide parameter.

Nonsynonymous mutations were found in 853 genes, including several genes found to be significantly mutated in prior studies of ccRCC. *VHL* was the most frequently mutated driver gene (7/10 tumors) in our AA-exposed ccRCCs (Table 2). The second most frequent was *PBRM1*, with inactivating mutations in 2 of 10 tumors. *VHL* and *PBRM1* were also the two most frequent significantly mutated driver genes in the TCGA ccRCC study and occurred at a similar mutation frequency as our cohort (53% *VHL*, 34% PBRM1; ref. 27). We also observed mutations in other known ccRCC driver genes, including the tumor suppressors *SETD2*, *BAP1*, *GNB2L1*, and *EPAS1*, as well as an oncogenic mutation in *PIK3CA* (Supplementary Table S2). Furthermore, we compared the allelic fractions from SNPs in the tumor and normal pairs to identify regions of somatic allelic imbalance. Loss of chromosome 3p is the most frequent somatic alteration in ccRCC (28), as this region contains a number of driver genes, including *VHL*. Indeed, we detected chromosome 3p arm loss in 6 of 10 tumors (Supplementary Fig. S1). These data suggest that AA-exposed and TCGA individuals share common ccRCC driver genes.

Mutational patterns of SBSs in ccRCC

We next asked whether the AA-exposed ccRCC tumors harbored the main feature of the AA mutational signature, an elevated frequency of A-to-T transversions. Indeed, the fraction of A-to-T transversion mutations in the 10 AA-exposed ccRCC was significantly higher (33 ± 23%, mean ± SD) than in the TCGA ccRCCs (10.7 ± 4.4%; $P < 0.0001$, two-tailed t test; Fig. 3A). However, we observed a wide distribution of A-to-T transversion fraction in our 10 AA-exposed ccRCCs (range, 7%–69%) compared with TCGA ccRCCs (range, 0%–26%).

We took advantage of the TCGA ccRCC cohort data to define a stringent criterion for the expected fraction of A-to-T transversions. Two standard deviations above the TCGA ccRCC cohort average was set as a threshold, and tumors with A-to-T transversion fractions above this value (>20%) were designated as having elevated A-to-T. In 6 tumors of our AA-exposed cohort, the A-to-T fraction exceeded 20% of the SBSs (Fig. 3B). Note that 2.4% of the TCGA ccRCCs (10/417 tumors) has an A-to-T fraction greater than 20%. However, without the complementary AL–DNA adduct analysis to implicate AA, the elevated A-to-T fraction could have been due to other factors, including compounds whose metabolites lead to A-to-T mutations (29).

We compared the SBS mutational spectra of the following three ccRCC cohorts: (i) AA-exposed individuals with high A-to-T fractions ($n = 6$); (ii) AA-exposed individuals with low A-to-T fractions ($n = 4$); and (iii) TCGA individuals ($n = 417$; Fig. 3C). The difference between the high A-to-T set and the TCGA dataset was primarily in the excess of A-to-T transversion mutations among the SBSs. After the A-to-T class is removed, the relative ranking of the other classes of SBSs is similar to that seen in the TCGA samples. This pattern in AA-associated ccRCC, an excess of A-to-T transversions, is reminiscent of that observed in AA-associated UTUC.

AA mutational signature

We examined 488 A-to-T mutations compiled from the 6 ccRCC cases with an excess fraction of A-to-T transversions for key features of the AA mutational signature observed in AA-associated UTUC. dA–AL adducts are repaired efficiently by transcription-coupled DNA repair but only poorly by global genome repair (30). Indeed, the deoxyadenosine residue mutated in A-to-T transversions was found primarily on the nontranscribed strand in these ccRCCs (Fig. 4A). In contrast, we observed no strand preference for the deoxyadenosine residue among A-to-T transversions in the TCGA dataset (Fig. 4A). The A>T strand biases in AA-associated ccRCCs and AA-associated UTUCs were 2.1-fold and 2.6-fold, respectively (4).

The sequence context of A-to-T mutations was probed by tabulating the bases neighboring the mutated deoxyadenines. Figure 4B shows the frequencies of the 16 potential trinucleotides in which a dA residue in the central position is mutated to form an A-to-T transversion. The 5′CpApG trinucleotide includes 29% of the A-to-T mutations (141 A>T in CpApG of 488 XpApX) in the Taiwanese ccRCCs compared with only 11% (273/2,575) in the TCGA dataset ($P = 2.3^{-27}$, two-sided χ^2). In Taiwanese AA-associated UTUCs, the CpApG trinucleotide was also found in similar proportion, 33% (3,365/10,326) in AA-associated UTUCs compared with 10% (503/5,250) in controls (4).

Exon splice acceptor sites almost always occur at ApG sequences and are most frequently found in a CpApG context. Consistent with an increased A-to-T preference for the CpApG context, the frequency of mutations in the splice acceptor was elevated in ccRCC compared with TCGA (2.8% vs. 1.0% of substitutions, respectively; Fig. 4C). The ratio of the number of mutations at splice acceptor to splice donor sites was 3.4-fold (24/7) in the Taiwanese ccRCCs. In contrast, there was no splice site mutation preference observed among ccRCCs in the TCGA dataset (1.1-fold, 269/251; $P = 0.0021$ two-sided Fisher exact test). The excess of splice acceptor mutations was less than that observed in AA-associated UTUC exome sequences (6.9-fold; ref. 4). In summary, all significant features of the AA mutational signature reported for AA-associated UTUCs were found in this set of Taiwanese ccRCCs and were highly statistically significant. We made no attempt to match our cohort to the TCGA with respect to known ccRCC risk factors, such as smoking history and BMI, but as the TCGA set completely lacks the AA mutational signature, we can confidently conclude that these known risk factors, represented abundantly in the TCGA set, do not produce the mutational signature we are attributing to AA in ccRCC.

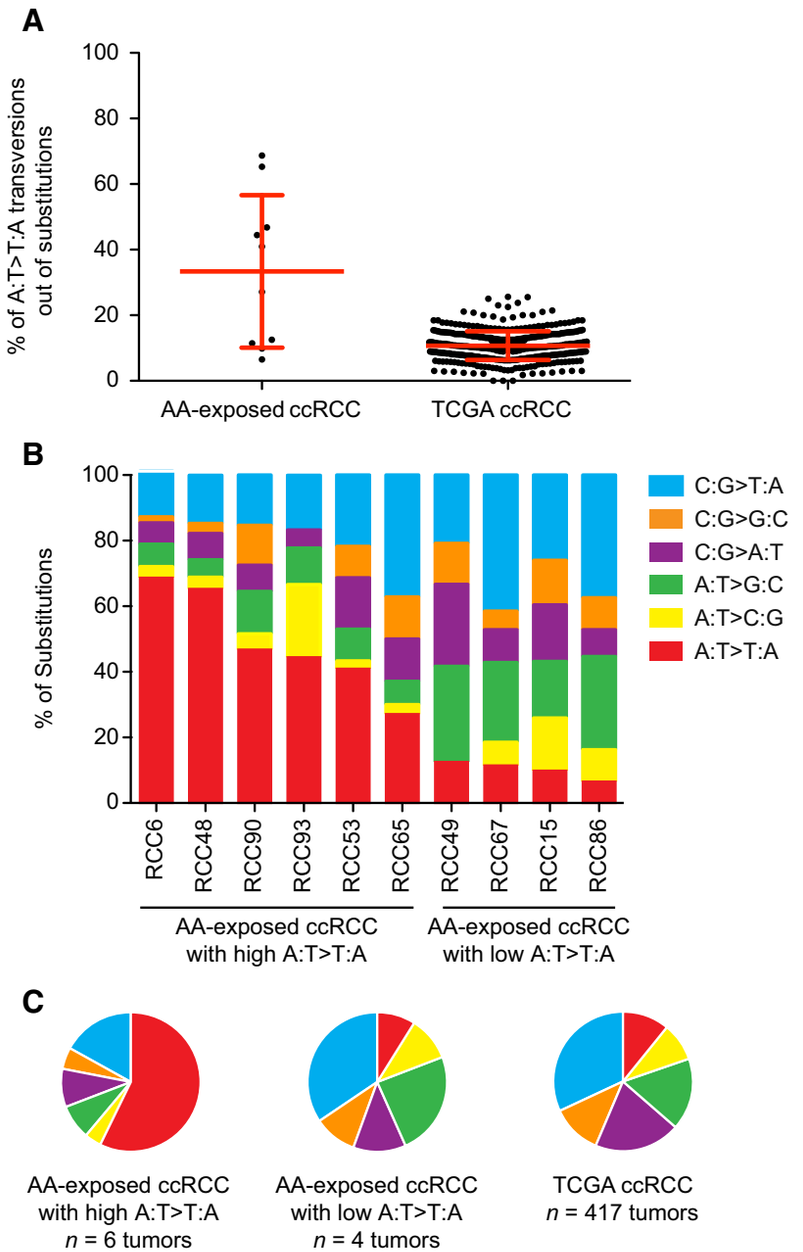

Figure 3.

Subset of AA-exposed ccRCCs have elevated levels of A:T>T:A transversions. **A,** Scatterplot of fractions of A:T>T:A transversions, based on the six substitution types of 10 AA-exposed Taiwanese ccRCCs, compared with 417 TCGA ccRCCs. Red line for AA-exposed ccRCC is 33 ± 23% (mean ± SD) and for TCGA, ccRCC is 11 ± 4.4%. **B,** Stacked bar chart with frequencies of each of the six substitution types in 10 AA-exposed Taiwanese ccRCC indicated on the *x*-axis. Tumors arranged by decreasing fraction of A-to-T transversions from left to right. Elevated level of A-to-T determined as greater than two SDs above A-to-T average of TCGA ccRCC in **A** (>20%). AA-exposed ccRCC with "high" and "low" A-to-T fractions indicated. **C,** Pie graphs showing mutational spectrum of AA-exposed Taiwanese ccRCCs with high (left, $n = 843$ mutations) and low (middle, $n = 298$ mutations) A-to-T fractions compared with TCGA ccRCC (right, $n = 24559$). Legend for substitution types is indicated in **B**.

Discussion

The widespread exposure of the Taiwanese population to herbs containing AA has been well documented through analysis of the national prescription database (13). Moreover, exposure to AA has been directly demonstrated via the detection of AL–DNA adducts in target tissues and by the AA mutational signature in UTUC (12). The incidences of kidney and upper tract urinary cancer are similar in the Taiwanese population. The age-specific rate (ASR, adjusted to the world standard population), as reported in the Taiwan Cancer Registry, for renal pelvis and ureter cancer (primarily UTUC) from 2003 to 2007 was 2.9 per 10^5 years in men and 3.3 per 10^5 years in women, whereas the ASR for kidney cancers (primarily ccRCC) was 3.3 per 10^5 years in men and 1.9 per 10^5 years in women.

Similar to what had been previously found with urothelial carcinoma, there is a dose-dependent relationship between consumption of AA-containing herbs and ccRCC incidence. We could document exposure to AA through prescribed medicines in 38.3% of Taiwanese with ccRCC. Furthermore, the presence of AL–DNA adducts revealed the exposure of 75% of Taiwanese ccRCC patients to this potent and highly persistent human carcinogen, indicating additional exposure beyond the estimate based on prescriptions reimbursed by NHI. This underestimate is expected in assessing life-long exposure by a 7-year window. Other contributors are potential misclassification of patients that obtained

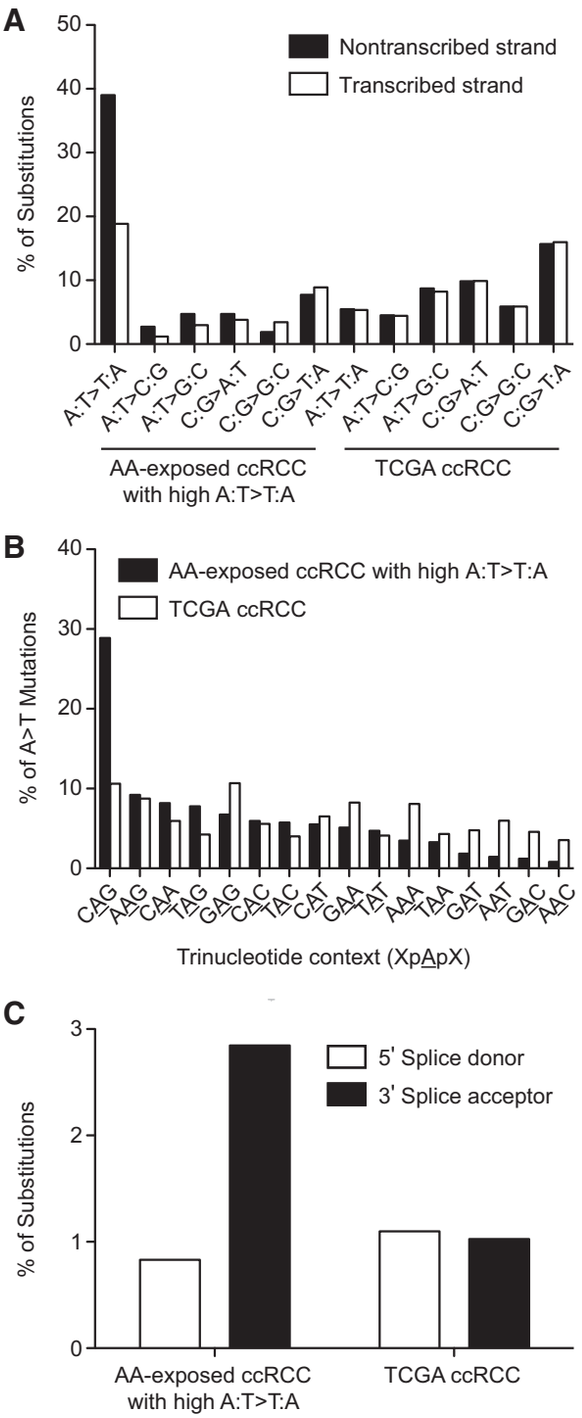

Figure 4.

AA mutagenesis patterns in Taiwanese ccRCC tumors with greater than 20% A-to-T. **A,** Percentages of the six classes of substitutions indicated on the *x*-axis for AA-exposed ccRCCs (left, 6 tumors, $n = 843$ mutations) compared with TCGA ccRCCs (right, 417 tumors, $n = 24{,}559$ mutations). Each substitution type subdivided by nontranscribed strand (black bar) or transcribed strand (white bar). **B,** Percentages of A>T (same as T>A) mutations within each of the indicated trinucleotide sequences in AA-exposed ccRCCs (black bars, $n = 488$ A>Ts) and TCGA ccRCC dataset (white bars, $n = 2{,}575$ A>Ts). The middle A (underlined) is the mutated base (A>T) in each trinucleotide sequence shown on the *x*-axis. Note that A>T in CpApG is the same as T>A in CpTpG and is accounted for in these data. **C,** Bar graph showing frequency of mutations in 5′ splice donors (black bars) or 3′ splice acceptor (white bars) in AA-exposed ccRCCs compared with TCGA ccRCCs. Mutations were counted only in the canonical 5′ splice donor (GT, IVS+1 and IVS+2) and 3′ splice acceptor (AT, IVS-1, IVS-2) on the nontranscribed strand.

herbal remedies outside of the NHI system. This same underestimate has affected previous epidemiologic studies of AA and UTUC, utilizing the prescription database. Lai and colleagues (24) reported that 28.7% of UTUC patients had been reimbursed for prescriptions containing AA, but Chen and colleagues (12) reported that 60% of UTUC patients had AL–DNA adducts. Whole-exome sequencing confirmed that the AA mutational signature was present in 6 of 10 ccRCC patients studied. Previous examination of the *TP53* gene in Taiwanese RCC did not detect the AA mutational signature due to limited sample size; only three *TP53* mutations were detected in 25 RCCs (12). From the fraction of patients whose tumors had the AA mutational signature and the distribution of AL adducts in our cohort, we estimate that AA plays a role in the etiology of at least 30% of Taiwanese ccRCC.

Although the AA mutational signature clearly documents exposure and a mutational impact, we cannot rule out that AA could promote tumorigenesis through genotoxicity-independent mechanisms. Indeed, AA is a potent human nephrotoxin. Although subclinical renal effects of AA have not been investigated in the general population, a minority of people, estimated to be approximately 5%, are highly sensitive to the nephrotoxic effects of AA (31). Cellular necrosis induced by AA, followed by tissue regeneration, may provide a promitogenic environment for proliferation of tumor-initiating cells.

Potential AA mutational signature in other RCC sequences

Recent whole-exome sequencing studies of RCC have described subsets of tumors with a high incidence of A-to-T transversions. One study determined the sequences in a small cohort of Chinese RCC tumors (20). In this analysis, one of 10 tumors possessed an elevated frequency of A-to-T mutations. As in Taiwan, the use of *Aristolochia* herbal remedies is widespread in China, and many cases of aristolochic acid nephropathy have been reported there (32).

In other studies, DNA was obtained from RCC cases in England, Russia, the Czech Republic, and Romania (21) and also from the endemic region of Croatia (22). The mutation spectrum in RCCs from England, Russia, and the Czech Republic resembled that seen in the TCGA RCC study (drawn from an American population); however, the AA mutational signature predominated a majority of the Romanian RCCs (12/14) and 4 of 8 Croatian RCCs. Although Romania has long been known to harbor regions of endemic nephropathy, the RCC patients whose tumors were sequenced apparently did not reside in the endemic areas. However, the therapeutic use of *Aristolochia* herbs is known to occur in Romania (33), and a follow-up study did establish the presence of AL–DNA adducts in Romanian RCC patients (23).

The fraction of individuals with AL–DNA adducts is similar among Taiwanese patients with UTUC and RCC (12). Although our case–control study showed that patients who took a cumulative dosage of more than 250 mg of AA had an increased RCC

risk (crude OR = 1.4), a similar study reported a higher risk of urothelial carcinoma (bladder cancer and/or UTUC) among Taiwanese ingesting similar amounts of AA (crude OR = 1.9; ref. 24). Also, the number of A-to-T transversions per exome is higher in UTUC with the AA mutational signature than AA-related RCC (median = 188 per exome in UTUC and 46 per exome in RCC). Although these results indicate that renal tissue is sensitive to the carcinogenic effects of AA, they also suggest that the urothelium may be more sensitive to these effects. An alternative explanation is that the nephrotoxic effects of AA dominate renal tumorigenesis, and thus, AA-induced RCC develops only in the minority of people sensitive to the nephrotoxic effects of AA. Regardless of the precise mechanism, the current study offers compelling evidence implicating AA in a significant fraction of the RCC arising in Taiwan and illustrates the power of integrating epidemiologic, molecular, and genetic data in the investigation of cancer etiology.

Disclosure of Potential Conflicts of Interest

B. Vogelstein has ownership interest (including patents) in PapGene, Inc. and PGDx, is a consultant/advisory board member for PapGene, Inc., PGDx, Morphotek, Inc., and Sysmex Inostics, Inc., and has provided expert testimony for Johns Hopkins University (licensed invention). N. Papadopoulos has ownership interest (including patents) in and is a consultant/advisory board member for PapGene, Inc. and PGDx. K.W. Kinzler has ownership interest (including patents) in Gene and Testing Patents, PapGene, Inc., and PGDx and is a consultant/advisory board member for Morphotek and Sysmex Inostics. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: M.L. Hoang, C.-H. Chen, P.-C. Chen, Y.-S. Pu, B. Vogelstein, N. Papadopoulos, A.P. Grollman, K.W. Kinzler, T.A. Rosenquist

Development of methodology: M.L. Hoang, P.-C. Chen, N.J. Roberts, B.H. Yun, R.J. Turesky, A.P. Grollman, K.W. Kinzler, T.A. Rosenquist

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): M.L. Hoang, C.-H. Chen, P.-C. Chen, K.G. Dickman, B.H. Yun, Y.-S. Pu, B. Vogelstein, N. Papadopoulos, K.W. Kinzler

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): M.L. Hoang, C.-H. Chen, P.-C. Chen, N.J. Roberts, B.H. Yun, R.J. Turesky, N. Papadopoulos, K.W. Kinzler, T.A. Rosenquist

Writing, review, and/or revision of the manuscript: M.L. Hoang, C.-H. Chen, P.-C. Chen, N.J. Roberts, B.H. Yun, B. Vogelstein, N. Papadopoulos, A.P. Grollman, K.W. Kinzler, T.A. Rosenquist

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M.L. Hoang, C.-H. Chen, K.G. Dickman, B.H. Yun, Y.-S. Pu, K.W. Kinzler

Study supervision: Y.-S. Pu, A.P. Grollman, K.W. Kinzler, T.A. Rosenquist

Acknowledgments

Chi-Jen Chen and Yao-Hsu Yang from the Department of Urology, National Taiwan University College of Medicine and Hospital (Taipei, Taiwan) helped perform statistical analysis of the Taiwan National Health Insurance database.

Grant Support

This research was supported, in part, by Taiwan National Science Council grants 101-2314-B-002-027-MY3 (to C.H. Chen) and 101-2314-B-002-023-MY3 (to Y.S. Pu). Additional support included the Virginia and D.K. Ludwig Fund for Cancer Research, Commonwealth Foundation, and NIH grants CA057345 (to K.W. Kinzler), CA190889 (to N.J. Roberts), and ES019564 (to R.J. Turesky), HHMI (to B.V. Vogelstein), and Marsha and Henry Laufer (to A.P. Grollman).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 22, 2016; revised July 20, 2016; accepted July 20, 2016; published OnlineFirst August 23, 2016.

References

1. Brash DE, Rudolph JA, Simon JA, Lin A, McKenna GJ, Baden HP, et al. A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. Proc Natl Acad Sci U S A 1991;88:10124–8.
2. Hollstein M, Sidransky D, Vogelstein B, Harris CC. p53 mutations in human cancers. Science 1991;253:49–53.
3. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415–21.
4. Hoang ML, Chen CH, Sidorenko VS, He J, Dickman KG, Yun BH, et al. Mutational signature of aristolochic acid exposure as revealed by whole-exome sequencing. Sci Transl Med 2013;5:197ra02.
5. Poon SL, Pang ST, McPherson JR, Yu W, Huang KK, Guan P, et al. Genome-wide mutational signatures of aristolochic acid and its application as a screening tool. Sci Transl Med 2013;5:197ra01.
6. Kumar V, Poonam, Prasad AK, Parmar VS. Naturally occurring aristolactams, aristolochic acids and dioxoaporphines and their biological activities. Nat Prod Rep 2003;20:565–83.
7. Grollman AP, Scarborough J, Jelakovic B. Aristolochic acid nephropathy: an environmental and iatrogenic disease. In: Fishbein J, editor. Advances in molecular toxicology. Amsterdam, the Netherlands: Elsevier; 2009. p. 211–22.
8. Schmeiser HH, Schoepe KB, Wiessler M. DNA adduct formation of aristolochic acid I and II in vitro and in vivo. Carcinogenesis 1988;9:297–303.
9. Grollman AP, Shibutani S, Moriya M, Miller F, Wu L, Moll U, et al. Aristolochic acid and the etiology of endemic (Balkan) nephropathy. Proc Natl Acad Sci U S A 2007;104:12129–34.
10. Jelakovic B, Karanovic S, Vukovic-Lela I, Miller F, Edwards KL, Nikolic J, et al. Aristolactam-DNA adducts are a biomarker of environmental exposure to aristolochic acid. Kidney Int 2012;81:559–67.
11. Schmeiser HH, Nortier JL, Singh R, Gamboa da Costa G, Sennesael J, Cassuto-Viguier E, et al. Exceptionally long-term persistence of DNA adducts formed by carcinogenic aristolochic acid I in renal tissue from patients with aristolochic acid nephropathy. Int J Cancer 2014;135:502–7.
12. Chen CH, Dickman KG, Moriya M, Zavadil J, Sidorenko VS, Edwards KL, et al. Aristolochic acid-associated urothelial cancer in Taiwan. Proc Natl Acad Sci U S A 2012;109:8241–6.
13. Hsieh SC, Lin IH, Tseng WL, Lee CH, Wang JD. Prescription profile of potentially aristolochic acid containing Chinese herbal products: an analysis of National Health Insurance data in Taiwan between 1997 and 2003. Chin Med 2008;3:13.
14. Fernando RC, Schmeiser HH, Scherf HR, Wiessler M. Formation and persistence of specific purine DNA adducts by 32P-postlabelling in target and non-target organs of rats treated with aristolochic acid I. IARC Sci Publ 1993;124:167–71.
15. Shibutani S, Dong H, Suzuki N, Ueda S, Miller F, Grollman AP. Selective toxicity of aristolochic acids I and II. Drug Metab Dispos 2007;35:1217–22.
16. Poon SL, Huang MN, Choo Y, McPherson JR, Yu W, Heng HL, et al. Mutation signatures implicate aristolochic acid in bladder cancer development. Genome Med 2015;7:38.
17. Huang J, Deng Q, Wang Q, Li KY, Dai JH, Li N, et al. Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. Nat Genet 2012;44:1117–21.
18. Totoki Y, Tatsuno K, Covington KR, Ueda H, Creighton CJ, Kato M, et al. Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. Nat Genet 2014;46:1267–73.
19. Zou S, Li J, Zhou H, Frech C, Jiang X, Chu JS, et al. Mutational landscape of intrahepatic cholangiocarcinoma. Nat Commun 2014;5:5696.
20. Guo G, Gui Y, Gao S, Tang A, Hu X, Huang Y, et al. Frequent mutations of genes encoding ubiquitin-mediated proteolysis pathway components in clear cell renal cell carcinoma. Nat Genet 2012;44:17–9.

21. Scelo G, Riazalhosseini Y, Greger L, Letourneau L, Gonzalez-Porta M, Wozniak MB, et al. Variation in genomic landscape of clear cell renal cell carcinoma across Europe. Nat Commun 2014;5:5135.
22. Jelakovic B, Castells X, Tomic K, Ardin M, Karanovic S, Zavadil J. Renal cell carcinomas of chronic kidney disease patients harbor the mutational signature of carcinogenic aristolochic acid. Int J Cancer 2015;136:2967–72.
23. Turesky RJ, Yun BH, Brennan P, Mates D, Jinga V, Harnden P, et al. Aristolochic acid exposure in Romania and implications for renal cell carcinoma. Br J Cancer 2016;114:76–80.
24. Lai MN, Wang SM, Chen PC, Chen YY, Wang JD. Population-based case-control study of Chinese herbal products containing aristolochic acid and urinary tract cancer risk. J Natl Cancer Inst 2010;102:179–86.
25. Goodenough AK, Schut HA, Turesky RJ. Novel LC-ESI/MS/MS(n) method for the characterization and quantification of 2'-deoxyguanosine adducts of the dietary carcinogen 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine by 2-D linear quadrupole ion trap mass spectrometry. Chem Res Toxicol 2007;20:263–76.
26. Yun BH, Rosenquist TA, Sidorenko V, Iden CR, Chen CH, Pu YS, et al. Biomonitoring of aristolactam-DNA adducts in human tissues using ultra-performance liquid chromatography/ion-trap mass spectrometry. Chem Res Toxicol 2012;25:1119–31.
27. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 2013;499:43–9.
28. Carroll PR, Murty VV, Reuter V, Jhanwar S, Fair WR, Whitmore WF, et al. Abnormalities at chromosome region 3p12–14 characterize clear cell renal carcinoma. Cancer Genet Cytogenet 1987;26:253–9.
29. Nassar D, Latil M, Boeckx B, Lambrechts D, Blanpain C. Genomic landscape of carcinogen-induced and genetically induced mouse skin squamous cell carcinoma. Nat Med 2015;21:946–54.
30. Sidorenko VS, Yeo JE, Bonala RR, Johnson F, Scharer OD, Grollman AP. Lack of recognition by global-genome nucleotide excision repair accounts for the high mutagenicity and persistence of aristolactam-DNA adducts. Nucleic Acids Res 2012;40:2494–505.
31. Debelle FD, Vanherweghem JL, Nortier JL. Aristolochic acid nephropathy: a worldwide problem. Kidney Int 2008;74:158–69.
32. Yang L, Su T, Li XM, Wang X, Cai SQ, Meng LQ, et al. Aristolochic acid nephropathy: variation in presentation and prognosis. Nephrol Dial Transplant 2012;27:292–8.
33. Gluhovschi G, Margineanu F, Kaycsa A, Velciov S, Gluhovschi C, Bob F, et al. Therapeutic remedies based on *Aristolochia clematitis* in the main foci of Balkan endemic nephropathy in Romania. Nephron Clin Pract 2010;116:c36–46.