

CANCER

Aristolochic acids and their derivatives are widely implicated in liver cancers in Taiwan and throughout Asia

Alvin W. T. Ng,^{1,2,3*} Song Ling Poon,^{4*} Mi Ni Huang,^{1,2} Jing Quan Lim,^{4,5} Arnoud Boot,^{1,2} Willie Yu,^{1,2} Yuka Suzuki,^{1,2} Saranya Thangaraju,⁴ Cedric C. Y. Ng,⁴ Patrick Tan,^{2,6,7,8} See-Tong Pang,⁹ Hao-Yi Huang,¹⁰ Ming-Chin Yu,¹¹ Po-Huang Lee,¹² Sen-Yung Hsieh,^{10†} Alex Y. Chang,^{13†} Bin T. Teh,^{2,4,7,14†} Steven G. Rozen^{1,2,3,7†}

Copyright © 2017
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim
to original U.S.
Government Works

Many traditional pharmacopeias include *Aristolochia* and related plants, which contain nephrotoxins and mutagens in the form of aristolochic acids and similar compounds (collectively, AA). AA is implicated in multiple cancer types, sometimes with very high mutational burdens, especially in upper tract urothelial cancers (UTUCs). AA-associated kidney failure and UTUCs are prevalent in Taiwan, but AA's role in hepatocellular carcinomas (HCCs) there remains unexplored. Therefore, we sequenced the whole exomes of 98 HCCs from two hospitals in Taiwan and found that 78% showed the distinctive mutational signature of AA exposure, accounting for most of the nonsilent mutations in known cancer driver genes. We then searched for the AA signature in 1400 HCCs from diverse geographic regions. Consistent with exposure through known herbal medicines, 47% of Chinese HCCs showed the signature, albeit with lower mutation loads than in Taiwan. In addition, 29% of HCCs from Southeast Asia showed the signature. The AA signature was also detected in 13 and 2.7% of HCCs from Korea and Japan as well as in 4.8 and 1.7% of HCCs from North America and Europe, respectively, excluding one U.S. hospital where 22% of 87 "Asian" HCCs had the signature. Thus, AA exposure is geographically widespread. Asia, especially Taiwan, appears to be much more extensively affected, which is consistent with other evidence of patterns of AA exposure. We propose that additional measures aimed at primary prevention through avoidance of AA exposure and investigation of possible approaches to secondary prevention are warranted.

INTRODUCTION

Mutational signature analysis provides a molecular epidemiological tool for detecting environmental exposures that cause cancers (1–5). This has important implications for public health by providing evidence to substantiate causal links between exposures and tumors, providing opportunities for primary and secondary prevention. Mutational signature analysis may also affect clinical oncology in situations where identifiable mutagenic exposures suggest specific cancer risks or preferred treatments.

Mutational signature analysis has been particularly helpful in illuminating the epidemiology of tumors associated with aristolochic acids and their derivatives (collectively, AA). Among these compounds, the *in vitro* toxicity and mutagenicity of aristolochic acids and aristolactams have been most intensively studied (6–8). AAs include potent mutagens and nephrotoxins present in plants in the genera

Aristolochia and *Asarum*, as well as related plants (6, 7). Many of these plants are used as herbal medicines (9–18). AA mutagenesis is thought to stem from the formation of bulky adducts on purines (19–21). For reasons that are imperfectly understood, but possibly related to better repair of AA-guanine adducts, more accurate translesion synthesis across AA-guanine adducts, or both, AA induces adenine-to-thymine (A>T) mutations almost exclusively (9, 20, 22–24).

In the early 1990s, inadvertent treatment with AA-containing herbs at a Belgian weight loss clinic caused kidney failure in ~100 women (25), many of whom later developed bladder and upper tract urothelial carcinomas (UTUCs) (9). Subsequently, additional reports of kidney failure and urothelial cancers due to AA poisoning appeared, and it emerged that AA was also responsible for Balkan endemic nephropathy (9). Taiwan also emerged as a hot spot for AA exposure based on prescription records and high rates of kidney failure and UTUCs, which are likely to be partly due to AA exposure (26–30). More recently, mutational signature analysis and other lines of evidence have suggested that AA mutagenesis may be widespread in terms of both geography and types of cancer affected (10). In particular, after we and others described a distinctive mutational signature of AA exposure in the genomes of UTUCs from Taiwan (29, 30), this signature was also found in bladder carcinomas (BCs) from Taiwan and other regions (31), renal cell carcinomas (RCCs) from Taiwan and the Balkans (32–34), intrahepatic bile duct carcinomas from China (35), bile duct carcinomas from Singapore (36), and hepatocellular carcinomas (HCCs) from China, Vietnam, and other Southeast Asian countries (29, 37–39).

Inference of high rates of AA exposure in Taiwan is based on the following evidence: (i) prescription records indicating about one-third of the population exposed to AA (28), (ii) high rates of UTUCs and co-association of kidney failure and UTUCs (40, 41), (iii) presence of AA-DNA adducts associated with UTUCs and RCCs (26, 34), and (iv) presence of the AA mutational signature in UTUCs, BCs, and

¹Centre for Computational Biology, Duke-NUS Medical School, Singapore 169857, Singapore. ²Programme in Cancer and Stem Cell Biology, Duke-NUS Medical School, Singapore 169857, Singapore. ³NUS Graduate School for Integrative Sciences and Engineering, Singapore 117456, Singapore. ⁴Laboratory of Cancer Epigenome, Division of Medical Science, National Cancer Centre Singapore, Singapore 169610, Singapore. ⁵Lymphoma Genomic Translational Research Laboratory, Division of Medical Oncology, National Cancer Centre Singapore, Singapore 169610, Singapore. ⁶Cancer Science Institute of Singapore, National University of Singapore, Singapore 117599, Singapore. ⁷SingHealth/Duke-NUS Precision Medicine Institute, Singapore 169609, Singapore. ⁸Genome Institute of Singapore, Singapore 138672, Singapore. ⁹Division of Urooncology, Department of Urology, Chang Gung University and Memorial Hospital, Linkou, Taoyuan 33305, Taiwan. ¹⁰Department of Gastroenterology and Hepatology, Chang Gung Memorial Hospital, Linkou, Taoyuan 33305, Taiwan. ¹¹Department of General Surgery, Chang Gung Memorial Hospital, Linkou, Taoyuan 33305, Taiwan. ¹²Department of Surgery, National Taiwan University, Taipei 10051, Taiwan. ¹³Johns Hopkins Singapore, Singapore 308433, Singapore. ¹⁴Institute of Molecular and Cell Biology, Singapore 138673, Singapore.

*These authors contributed equally to this work.

†Corresponding author. Email: siming.shia@msa.hinet.net (S.-Y.H.); alexchang@imc.jhmi.edu (A.Y.C.); teh.bin.tean@singhealth.com.sg (B.T.T.); steve.rozen@duke-nus.edu.sg (S.G.R.)

RCCs from Taiwan (29–31, 34). However, despite the high amount of AA exposure in Taiwan and reports of the AA mutational signature in HCCs from China and other areas, AA exposure in Taiwan HCCs remains unexplored.

RESULTS

Overview of somatic changes in 98 HCCs from Taiwan

To investigate the possible presence of the AA mutational signature in HCCs from Taiwan, we sequenced the exomes of 98 HCCs and matched nonmalignant tissues from two hospitals (table S1). Tumor tissue was obtained from nonconsecutive patients, and inclusion in this study was solely based on the availability of adequate DNA. Tumors were not selected based on suspicion of AA exposure.

We sequenced whole exomes, with a mean of 95% targeted tumor bases with $\geq 30\times$ coverage (table S2). We detected a total of 26,805 somatic single-base substitution (SBS) mutations across the HCCs (median, 167 SBS per tumor; interquartile range, 103 to 316), with an estimated false discovery rate (FDR) of 1.9% (tables S3 and S4). We detected a total of 648 short insertions or deletions (indels; median, 6 indels per tumor; interquartile range, 3 to 9), with an estimated FDR of 3.2% (tables S3 and S5).

In total, 10,174 genes harbored nonsilent SBS mutations (table S4). Driver analysis with MutSigCV (42) and 20/20+ (43) identified 16 significantly mutated genes (tables S6 to S8). The most commonly mutated genes—*TP53*, *CTNNB1*, *ALB*, and *AXIN1*—were also the most commonly mutated in the recent TCGA (The Cancer Genome Atlas) report (37), but the proportions of tumors with mutations in these genes were higher in the Taiwan HCCs (table S6). Among these genes, it has been proposed that *ALB* inactivation may promote cancer development by “diverting energy into cancer-relevant metabolic pathways” (37, 44). An additional gene identified by this analysis was *IRF2*, which was previously reported to act as a tumor suppressor in HCC (45). Of the other genes, two have not been identified as likely drivers in previous genome- or exome-wide resequencing of HCCs, and other evidence of their roles in cancer is absent or very limited, suggesting a lack of functional roles in HCC (table S6). Previously observed genomically amplified oncogenes and deleted tumor suppressors were also amplified or deleted in the Taiwan HCCs (fig. S1). These genes included the amplified oncogene *CCND1* and the deleted tumor suppressor *RB1* (37, 39, 46–49).

High rates of the AA mutational signature in Taiwan HCCs

The mutational spectra of most of the HCCs from Taiwan showed marked evidence of AA exposure, in the form of high proportions of A:T>T:A mutations in the trinucleotide contexts characteristic of AA-exposed tumors and cell lines (Fig. 1, A to D, and fig. S2) (29–31), although some HCCs did not show this evidence (Fig. 1E and fig. S2). The trinucleotide contexts characteristic of AA exposure included a prominent peak at 5'-CTG-3' (5'-CAG-3' on the complementary strand). There was also a notable excess of A>T mutations on the nontranscribed strands of genes, which is characteristic of AA-induced mutations in other tumor types and in cell lines (29–35). Principal components analysis clustered the majority of the Taiwan HCCs away from other HCCs and with previously reported AA-associated UTUCs (29, 30) and BCs (31) and with AA-exposed cell lines (Fig. 1F) (29).

To systematically assess the extent of AA exposure across the 98 HCCs, we developed the mSigAct (mutational signature activity) software. mSigAct provides a signature presence test to infer whether

observed mutation spectra are better explained with a contribution from the AA mutational signature [Catalogue of Somatic Mutations in Cancer (COSMIC) signature 22] than without. We developed mSigAct because, to our knowledge, current approaches, most of which are based on nonnegative matrix factorization (NMF), do not support statistical inference of the presence or absence of a signature (3, 4, 50, 51). Briefly, the mSigAct test starts by generating optimal coefficients for reconstruction of the observed spectrum using the mutational signatures previously detected in HCCs. The test first does this without the AA signature (null hypothesis) and then with the AA signature (alternative hypothesis). The test then carries out a standard likelihood ratio test on these two hypotheses. Supplementary Materials and Methods and codes S1 and S2 provide details on the test, its evaluation on synthetic data, and the code. mSigAct revealed strong evidence of AA exposure in 76 of the 98 HCCs (78% with FDR < 0.05; Fig. 2, Table 1, and table S9). Among tumors with the AA signature, there was a median of 2.26 AA signature mutations/Mb (mean, 4.94 AA signature mutations/Mb).

As a further check on the mSigAct signature presence test, we also analyzed the 98 Taiwan HCCs with the NMF procedure in (3, 4) (code S1). The signature extracted by NMF had a Pearson correlation coefficient of 0.997 and a cosine similarity of 0.997 with the AA signature (COSMIC signature 22; fig. S3 and table S10). We also used NMF to detect the presence or absence of the AA signature and compared the results for this to those from the mSigAct signature presence test. The two procedures were concordant for 90 tumors (code S1). NMF identified eight putatively AA-exposed HCCs that mSigAct did not identify (T18, T41, T50, T53, T57, T61, T68, and T92; fig. S2). Thus, the mSigAct signature presence test was more conservative; the tumors identified by NMF but not mSigAct had very low numbers of A>T mutations (all but one ≤ 15) in backgrounds of relatively high numbers of other mutations, making it difficult to be confident of AA exposure (code S1). We would propose that this is the desired characteristic, that is, it is preferable to err on the side of undercalling rather than overcalling the presence of the AA signature. Furthermore, testing on synthetic data also indicated that the mSigAct signature presence test had better sensitivity and specificity (Supplementary Materials and Methods and code S2).

We examined associations between the extent of exposure to the AA signature and multiple clinical and epidemiological variables, namely, hospital, cirrhosis status, hepatitis B carrier status, hepatitis C carrier status, status as carrier of either hepatitis virus, diagnosis before or after the medicinal use of some AA-containing plants was banned in Taiwan in 2003, gender, date of diagnosis, and age at diagnosis (fig. S4). Of these, without correction for multiple hypothesis testing, AA exposure differed significantly by gender and age at diagnosis. There was a weak association of increased AA exposure with age (Spearman's $\rho = 0.28$, $P = 0.008$). In addition, AA mutation numbers were higher in females than in males (median, 176 versus 55 AA signature mutations per HCC; $P = 0.015$ by two-sided Wilcoxon rank sum test). After considering multiple hypothesis testing, the Benjamini-Hochberg FDRs for both gender and age were 0.065. Although AA mutation numbers were not statistically higher in women than in men, we note other evidence of more exposure to AA-containing herbs among women: In Taiwan before the ban, exposures were 31.6 person-years per 1000 for women compared to 25.9 for men (28). We also note that, because only 10 HCCs were hepatitis-negative, these data did not offer an opportunity to investigate interactions between hepatitis and AA exposure.

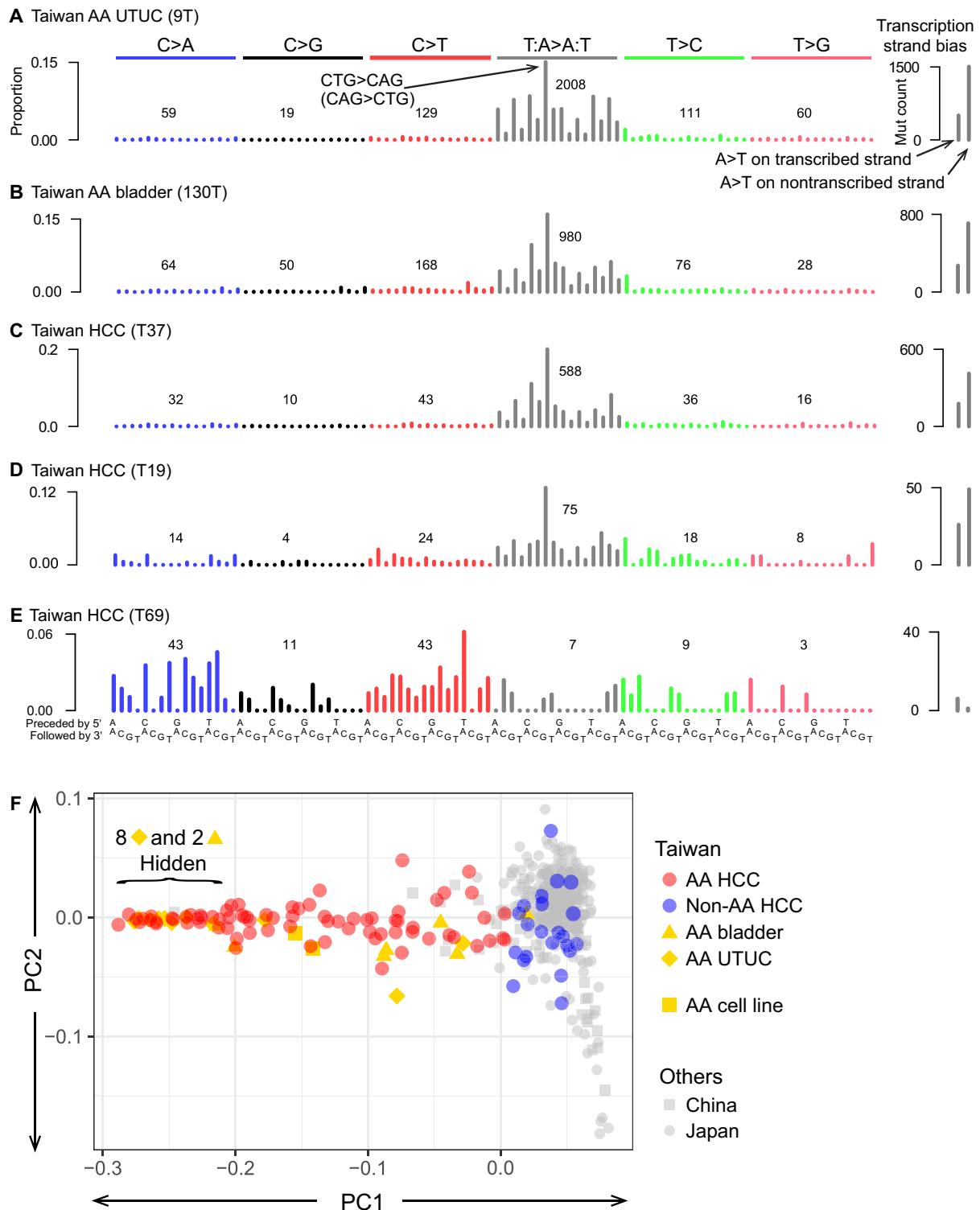


Fig. 1. Evidence of AA exposure in Taiwan HCCs. (A and B) Sample exome spectra of individual AA-exposed UTUCs (A) and BCs (B) from Taiwan. (C and D) Sample exome spectra of individual Taiwan HCCs with high (C) and moderate (D) levels of the AA signature. (E) Sample Taiwan HCC without AA signature. In the major plots in (A) to (E), each bar indicates the proportion of mutations in a particular trinucleotide context. In the AA signature (A to D), the overwhelming majority of mutations are T:A>A:T. By convention, mutations are shown as T>A (for example) rather than A>T, although AA mutations are physical consequences of adducts on adenines that cause A>T mutations (9, 20, 22–24). In tumors strongly mutagenized by AA, the most prominent peak is at CTG>CAG (CAG>CTG on the complementary strand), indicated in (A), often with additional prominent peaks at CTA>CAA and ATG>AAG. Small plots at right in (A) to (E) show transcription strand bias. Mut count, mutation count. (F) Mutation spectra-based principal components analysis of HCCs from Taiwan, China (52), and Japan (53), plus AA-exposed UTUCs (29) and BCs (31) and an AA-exposed cell line (29). The most distinguishable features are the T>A mutations induced by AA, which are reflected in PC1. PC1 explains 35% of the variance, and PC2 explains 5.5%.

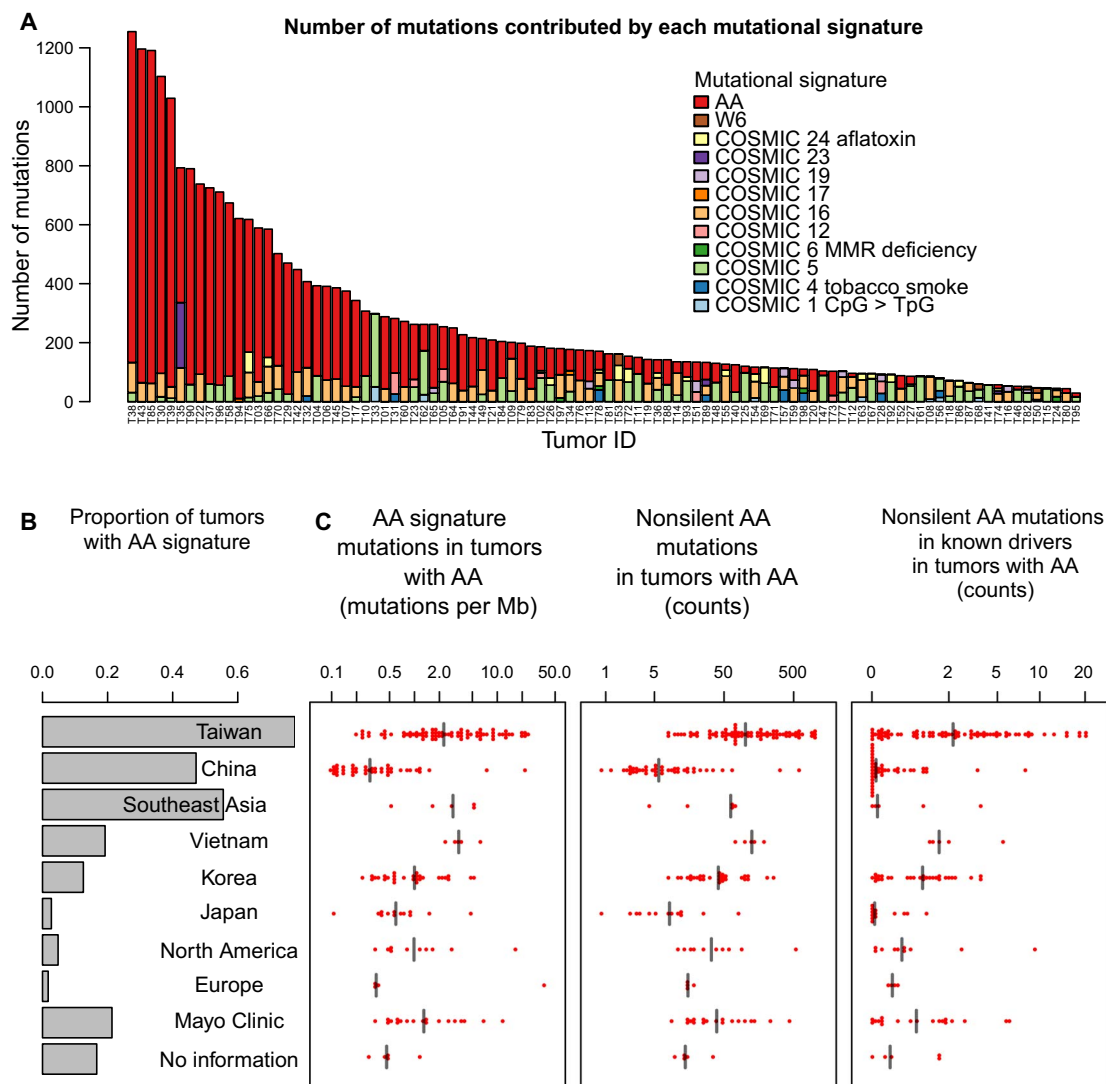


Fig. 2. Mutational signature exposures in Taiwan HCCs and summary of AA signature mutations. (A) Estimated numbers of mutations due to each mutational signature in each HCC. AA is COSMIC signature 22. W6 is from (53). COSMIC signatures 4 and 24 reflect known exogenous risk factors for HCC: tobacco smoking and aflatoxin exposure, respectively. MMR, mismatch repair. (B) Proportions of tumors with the AA signature in various groups of HCCs. "Southeast Asia" indicates Southeast Asia excluding Vietnam; "Mayo Clinic" denotes a group of HCCs from patients treated at that clinic for whom there was no country information and who we speculate may have traveled from Asia for treatment; "No information" denotes TCGA HCCs from biobanks for which there is no information on geographic origin. (C) Densities and counts of AA signature mutations among tumors with the AA signature. Each mutation is associated with a weighted assignment of the probability that it was caused by the AA signature (see Materials and Methods). The weighted count of AA signature mutations is the sum of these probabilities across all mutations in the tumors. The geographical regions indicated at the right of (B) also apply to (C).

The AA mutational signature in HCCs from other regions

Given the high prevalence of the AA signature in Taiwan HCCs, we examined publicly available data comprising 1400 HCCs (Table 1). These included data from China, Japan, Korea, and several countries in Southeast Asia (37–39, 46–48, 52, 53), as well as data from North America and Europe (37, 49) as negative controls with likely rare AA exposure.

We detected the AA signature in 42 of 89 HCCs (47%) from China (Figs. 2 to 4; Table 1; fig. S5, A and B; and table S11). Among the HCCs from earlier studies (47, 52), the mSigAct signature presence test detected many more affected HCCs than we were able to identify previously (29). Overall, however, AA signature mutation burdens were

lower in China (median, 0.29 AA signature mutations/Mb) than in Taiwan (median, 2.26 AA signature mutations/Mb).

We detected the AA signature in five of nine HCCs from various countries in Southeast Asia other than Vietnam (56%; fig. S5, C and D) (39). Among the HCCs with the AA signature, the median mutation burden was high (2.9 AA signature mutations/Mb). We also detected the signature in 5 of 26 HCCs from Vietnam (19%) (37), with a high median mutation burden of 3.4 AA signature mutations/Mb (fig. S5E). We also detected the AA signature in lower proportions of the HCCs from Korea and Japan (Table 1; Figs. 2, B and C, and 3; and fig. S5, F to H).

We analyzed TCGA data (37) from areas other than Vietnam in several subgroups (fig. S5I). In the largest subgroup, North America, we

Table 1. Summary of AA signature mutations in HCCs.

Geographic origin	Number of HCCs	Median (all SBS mutations/Mb)	HCCs with AA signature		Weighted counts* of AA signature SBS mutations among HCCs with the signature				Number of HCCs with nonsilent AA signature SBS mutations in known drivers	P value for difference from Taiwan in proportion of HCCs with the AA signature (Fisher's exact test)		
			Number of HCCs	Proportion of HCCs	SBS mutations/Mb	Number of nonsilent SBS mutations		Mean				
						Median	Mean	Median			Mean	
New data, this paper												
Taiwan	98	3.33	76	0.78	2.26	4.94	101.6	223.3	2.19	3.94	57	—
Previously published data												
China	89	1.94	42	0.47	0.29	1.07	5.8	32.6	0.06	0.50	9	2.0×10^{-5}
Southeast Asia (excluding Vietnam)	9	5.74	5	0.56	2.92	3.07	62.3	43.1	0.08	0.97	2	0.22 [†]
Vietnam	26	3.10	5	0.19	3.42	3.71	125.4	126.8	1.61	2.36	5	7.3×10^{-8}
Korea	231	1.78	29	0.13	1.00	1.29	41.4	56.2	1.06	1.18	19	2.1×10^{-30}
Japan	477	4.82	13	0.027	0.60	0.94	8.2	13.5	0.04	0.23	3	2.2×10^{-60}
North America	209	2.36	10	0.048	0.99	2.60	32.9	84.4	0.53	1.51	5	2.6×10^{-40}
Europe	230	1.67	4	0.017	0.35	9.72	15.2	16.0	0.34	0.35	0	7.3×10^{-49}
Mayo Clinic	89	2.94	19	0.21	1.30	2.27	39.2	77.0	0.89	1.46	11	6.0×10^{-15}
No information	30	2.79	5	0.17	0.46	0.57	13.9	17.1	0.30	0.76	2	2.6×10^{-9}
Total non-Taiwan	1400 [‡]											

*We ascribe the AA signature to a proportion of each SBS as described in Materials and Methods. Because the AA signature is dominated by A:T>T:A SBS, when the activities of other signatures with these SBS mutations are low or absent, this approximates the number of A:T>T:A SBS mutations. The weighted counts are the sums of the proportions of SBS mutations ascribed to the AA signature. †P = 7.1×10^{-7} , for all Southeast Asia, including Vietnam. ‡Includes 10 additional TCGA HCCs: 1 Russia, 5 Brazil, 4 South Korea; 1 Brazil HCC has the AA signature.

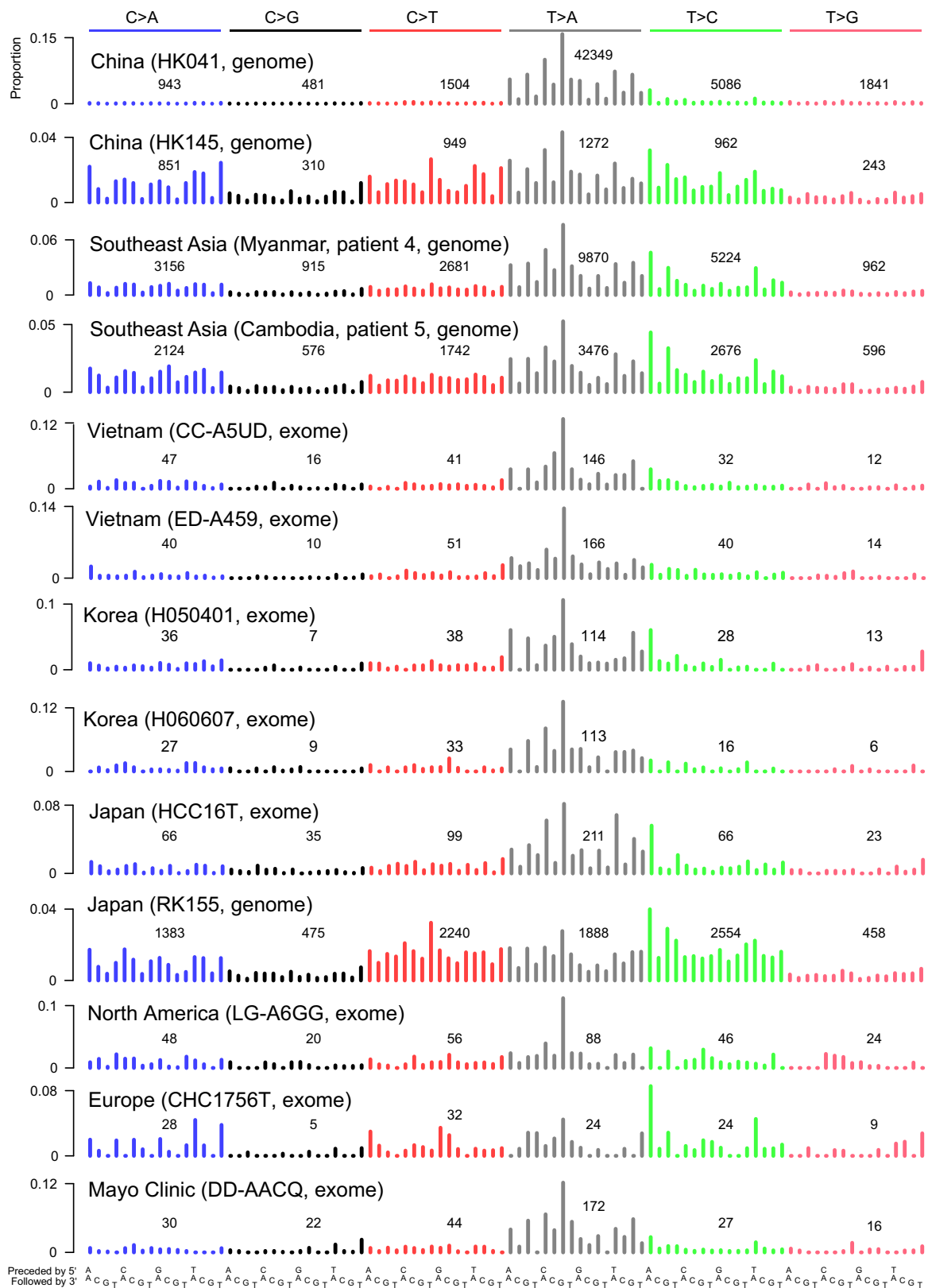


Fig. 3. Sample spectra of HCCs with the AA signature. Display conventions are the same as in Fig. 1.

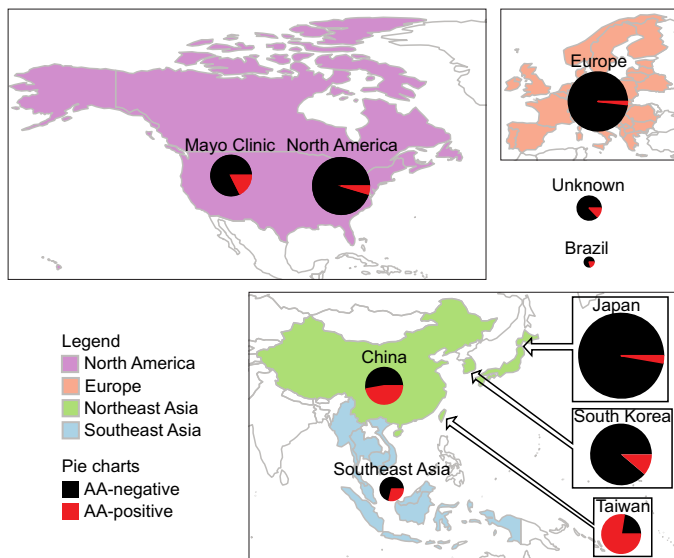


Fig. 4. Global distribution of mutagenesis associated with aristolochic acid and derivatives in HCCs. The pie chart labeled “Southeast Asia” includes both Vietnam and the other Southeast Asian HCCs. Pie chart areas are proportional to the number of HCCs in the given group.

detected AA signature mutations in 10 of 209 HCCs (5%; Table 1 and fig. S5J). Among HCCs from North America, the proportion with the AA signature from “Asian” patients (2/20) was not significantly different from non-Asian patients (8/189).

We also detected AA signature mutations in 4 of 230 HCCs from Europe (1.7%; fig. S5K). This low proportion is consistent with the rarity of reports of AA exposure in Europe outside of the Balkans and the Belgian poisoning incident in the 1990s (9, 25, 32, 33). Furthermore, the median AA mutation burden was low (0.35 AA signature mutations/Mb), although one HCC with likely DNA mismatch repair deficiency had many more mutations.

Within the TCGA data (37), there were 89 HCCs from the Mayo Clinic for which the “Country” field had no data, and almost all of these (87) had “Ethnicity” listed as Asian. Among these, 19 (21%) had the AA signature (fig. S5L). Given the high prevalence of the signature and relatively high numbers of AA signature mutations in these HCCs (median, 1.3 AA signature mutations/Mb), we speculate that some of these patients may have traveled from Asia for treatment. In addition, there were 30 HCCs from biobanks for which no Country information was available. Of these, 20 were listed as Asian, and 5 (25%) of these had the AA signature, whereas none of the non-Asian HCCs had the AA signature (fig. S5M). Finally, one of five HCCs from Brazil with non-Asian ethnicity (20%) showed the AA signature (fig. S5N).

The effects of the AA signature were especially prominent in Taiwan: A higher proportion of HCCs from Taiwan showed the AA signature than in any group other than the nine HCCs from Southeast Asia (not including the HCCs from Vietnam; Table 1). Nevertheless, this analysis of publicly available data showed widespread AA exposure in East and Southeast Asia and in self-identified Asians elsewhere.

AA signature mutations in known cancer drivers

Our initial analyses with MutSigCV and 20/20+ did not reveal any strong possibilities for previously unknown driver genes in the Taiwan HCCs, but many genes listed in the Cancer Gene Census as known cancer drivers (table S12) were affected by nonsilent mutations ascribed

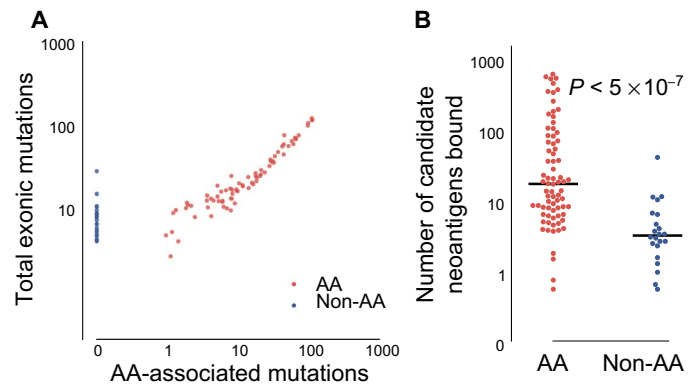


Fig. 5. High burdens of AA signature mutations and predicted immunogenicity in Taiwan HCCs. (A) AA signature mutations constitute the majority of mutations in most Taiwan HCCs affected by the AA mutational signature. (B) Many more in silico predicted candidate neoantigens in AA HCCs than non-AA HCCs; P value by Wilcoxon rank sum test.

to the AA signature (54). Across all Taiwan HCCs, the AA signature accounted for 59% (299 of 505) of nonsilent mutations in known driver genes (table S9). Among the Taiwan HCCs, 57 had a nonsilent AA mutation in ≥ 1 known driver (Fig. 2C, Table 1, and table S9). Among HCCs with the AA signature, two genes, *TP53* and *LRP1B*, were mutated frequently by both $A>T$ and by non- $A>T$ mutations (39 and 27 total nonsilent mutations, respectively, of which 48 and 63% were AA signature mutations; table S13). Recurrent mutations in *LRP1B* could be due to its large size (4599 amino acids; UniProt accession code Q9NZR2). It was not identified as a driver in our MutSigCV and 20/20+ analysis, and experimental evidence that it can function as a tumor suppressor is limited (55, 56). Several known tumor suppressors harbored predominantly AA signature mutations (table S13). Three of these are WNT-related tumor suppressors: *AXIN1*, *AXIN2*, and *APC*. Three others—*ARID1A*, *ARIDB*, and *SETD2*—are involved in chromatin remodeling, as is the oncogene *KMT2A*.

Tumors with the AA signature from regions other than Taiwan also had driver genes harboring nonsilent AA signature mutations (Table 1, Fig. 2C, and table S13). For example, 19 of the 29 AA-affected HCCs from Korea and all 5 of the AA-affected HCCs from Vietnam had AA signature mutations in known driver genes.

Clonality analysis of the Taiwan HCCs that had the AA signature suggested that AA mutations are predominantly early events, which is consistent with exposure before carcinogenesis (fig. S6). However, some AA signature mutations were subclonal, indicating that AA-associated mutagenesis, and presumably AA exposure, continued during tumor development and growth. Phylogenetic analysis based on multi-sector sequencing of HCCs from China in (38) showed that most AA mutations were truncal (found in all regions of the tumors), but some were subclonal, suggesting additional exposure to AA after initiation of carcinogenesis. A reanalysis of HCCs treated in Singapore showed a similar pattern of predominantly truncal AA signature mutations in four of the five AA-affected tumors (table S14) (39).

Potential for immunotherapy in high-AA-burden HCCs

A large proportion of HCCs in Taiwan bore heavy burdens of AA signature mutations (Fig. 2C and Table 1) and thus may be good candidates for immune checkpoint inhibitors (57). To investigate this possibility, we predicted neopeptides arising from nonsynonymous mutations and then predicted neopeptides binding to patient-specific human leukocyte

antigen (HLA) types. HCCs with the AA signature had many more predicted neoantigens (median, 146.5) than the remaining Taiwan HCCs (median, 60; $P < 2 \times 10^{-8}$ by Wilcoxon rank sum test; Fig. 5).

DISCUSSION

Some AA-containing herbal remedies have been officially prohibited in Taiwan since 2003, and we looked for evidence of whether this ban reduced exposure. We detected no significant difference in the prevalence of the AA signature or in the numbers of AA signature mutations in HCCs diagnosed before and after 2003. There are a number of possible, nonexclusive explanations. One possibility is that the decline in incidence of AA-associated HCCs may simply be lagging behind reduced AA exposure. There is precedent for this in tobacco-associated lung cancer. In the United States, male death rates roughly doubled in the 25 years after the 1964 Surgeon General's Report before beginning to decline in the 1990s, presumably as a result of tobacco suppression efforts that began decades earlier (58). In this context, we also note that AA-DNA adducts are extremely persistent (59).

Another possible explanation for unchanged prevalence of AA signature or numbers of AA signature mutations after the 2003 ban would be ongoing exposure to AA-containing herbal remedies. This could arise in various ways. (i) AA-containing plants may have still been prescribed by traditional Chinese medicine practitioners after the ban; this was documented to be the case in the first year after the ban (28). (ii) The nomenclature and labeling of products is confusing and error-prone; the herbs are often bought in formulations rather than individually, and in some traditional formulations, innocuous herbs can be replaced by herbs containing AA (12, 16, 28). (iii) Plants known to contain high concentrations of aristolochic acids are easily available on the internet, sometimes labeled correctly and sometimes incorrectly (table S15) (12, 16). (iv) Some plants containing AA were not banned in Taiwan. In particular, plants in the genus *Asarum*, collectively termed *xi xin* in Mandarin Chinese, were not banned and were the most commonly included plants in prescriptions surveyed in reference (28). As an example, powdered *xi xin* products manufactured in Taiwan and China were recently recalled in Singapore because they contained aristolochic acid I (fig. S7) (60, 61). Although reports of the presence and concentrations of aristolochic acids, aristolactams, and related compounds in *xi xin* are limited, available literature suggests that the concentrations are extremely variable and sometimes high (6, 12, 14, 62). An additional area of concern are aristolactams, which are not as intensively studied as the aristolochic acids but are thought to be the immediately mutagenic metabolites of aristolochic acids that interact directly with DNA (21).

There is strong evidence that the mutational signature that we and others have consistently observed is caused by aristolochic acids, possibly in conjunction with related compounds. Nevertheless, we cannot exclude the formal possibility that chemicals unrelated to aristolochic acids, aristolactams, and derivatives might also induce a mutational signature resembling the AA signature. However, at present, no such chemical is known, and many groups have independently and reproducibly detected a consistent AA mutational signature in UTUCs, BCs, RCCs, and bile duct cancers (table S10) (3, 29–37, 49, 63). Evidence that this signature is caused by AA and related compounds includes the mutational spectra of AA-treated cell lines (29) and the signature's association with AA-DNA adducts (30, 34) and with AA-related nephropathy in Taiwan and the Balkans. Furthermore, animal studies have shown that AA adducts, and presumably AA mutagenesis, occur in the liver (64–66).

To summarize the findings of this study, mutational signature analysis implicated AA exposure in 78% of HCCs from Taiwan. The AA signature was much more prevalent, and the number and proportion of mutations were notably higher in Taiwan than in most other regions. At the same time, AA exposure was found in cohorts from all Asian countries examined and in 22% of Asian patients treated at the Mayo Clinic. Across Taiwan HCCs, 299 of 505 nonsilent mutations in known driver genes were ascribed to the AA signature. Among the 76 AA-affected Taiwan HCCs, 57 had ≥ 1 nonsilent AA mutation in a known driver, and among the 133 AA-affected HCCs from elsewhere, 56 had ≥ 1 nonsilent AA mutation in a known driver gene, suggesting an active role for AA in the origins of these HCCs.

The findings here indicate that exposure to aristolochic acids and their derivatives is geographically widespread, implying substantial opportunities for primary and secondary prevention (Fig. 4). Medicinal use of AA-containing plants is only lightly regulated in many jurisdictions. The plants are not banned outright in China (67), and even in Taiwan, to the best of our understanding, only specific plants, rather than any plant and product containing AA or its derivatives, are regulated. Strikingly, *xi xin*, the most commonly prescribed herb before 2003, is not banned (28). In the United States, sale of AA-containing herbs is unregulated provided that they are correctly labeled and there are no claims of health benefits (68). Furthermore, plants containing aristolochic acid and its derivatives are readily available for sale on the internet (table S15).

In light of the wide availability of AA-containing plants, education and public awareness are paramount for primary prevention. In addition, the traditional nomenclature is confusing, making it difficult for consumers and suppliers to be sure of plant identification or of the constituents of multiherb preparations; there is ample evidence that mislabeling is common (12, 16, 25, 28). This latter point suggests that more thorough methods for testing herbal products, such as chromatographic fingerprinting, combined with regulatory oversight of the supply chain could also help reduce exposure (69). Secondary prevention might take the form of enhanced screening for AA-associated cancers or for kidney disease in patients suspected or known to be exposed to AA.

MATERIALS AND METHODS

Study design

This was designed as an exploratory retrospective study, because it was not known in advance whether HCCs from Taiwan would show evidence of the AA mutational signature or what the signature's prevalence might be. As noted above, tumor tissue was obtained from nonconsecutive patients, and inclusion of HCCs from Taiwan in this study was solely based on availability of adequate DNA and was not selected based on suspicion of AA exposure. After we discovered prevalent AA signature mutations in HCCs from Taiwan, we extended the study to publicly available somatic mutation data from 1400 tumors.

Patients and preparation of clinical samples

HCCs were diagnosed and classified by histopathological examination of surgically excised tumors, according to the World Health Organization classification system. Snap-frozen liver tumor tissues and matched normal samples (whole blood) from patients with HCC were obtained from Chang Gung Memorial Hospital Taiwan (21 patients) and National Taiwan University (77 patients). The human samples were sourced ethically with informed consent, and their research use was in accordance with the protocols approved by the Chang Gung Memorial Hospital

(103-6534C) and National Taiwan University Institutional Review Boards and Human Research Ethics Committees. Table S1 provides clinicopathological data and information on sequencing.

Whole-exome sequencing and mutation identification

Samples were captured with the Agilent SureSelect V5 exome panels. Paired 101–base pair reads were generated on HiSeq 2500 sequencers. BWA-MEM aligned reads to the human reference genome (hg19) (70), SAMtools removed polymerase chain reaction duplicates (71), and Qualimap 2 computed quality control metrics (72). Candidate somatic mutations were initially identified by three callers: GATK (73), Strelka (74), and MuTect (75). SBS called by ≥ 2 callers and small indels called by both GATK and Strelka were curated for downstream analysis. Examination of $\geq 1\%$ of SBS calls and $\geq 1\%$ indel calls from each tumor-normal pair (with ≥ 1 mutation from each tumor-normal pair) in IGV indicated an FDR of 6/312 (1.9%) of SBS calls and 3/93 (3.2%) of indel calls (table S3). Indels were rare, and their length distribution was unremarkable (fig. S8).

Sources of publicly available HCC somatic mutation data

These were as follows: (i) whole-genome sequence (WGS) from 78 HCCs from China (47, 52), in which we had previously noted the AA signature in 11 tumors (29), downloaded read data from <ftp.ncbi.nlm.nih.gov/sra/sra-instant/reads/ByStudy/sra/ERP/ERP001/ERP001196/> and realigned and called as described (76); (ii) whole-exome sequence (WES) from 11 HCCs from China (38), variant calls downloaded from the supplementary information for that paper (http://cancerres.aacrjournals.org/highwire/filestream/334829/field_highwire_adjunct_files/1/173141_1_supp_3805627_mj6gp5.xls); (iii) WES and WGS from nine HCCs in patients from various Southeast Asian countries, treated in Singapore (39), variant calls downloaded from the supplementary information for that paper; (iv) WES from 231 HCCs from Korea (48), downloaded from cBioPortal (www.cbioportal.org/; 8 February 2017) (77); (v) WGS from 264 HCCs from Japan (53), downloaded from the International Cancer Genome Consortium (ICGC) data portal (<https://dcc.icgc.org/>; March 2015, release 18, before publication of the paper); (vi) 213 WES from Japan, in which a mutational signature that, in retrospect, appears to be a merge of the AA signature with another signature had been reported (46), downloaded from the ICGC data portal (<https://dcc.icgc.org/>; March 2015, release 18); (vii) WES from 230 HCCs from France, Spain, and Italy, which likely are not regions of widespread AA exposure (49), downloaded from the ICGC data portal (<https://dcc.icgc.org/>, March, 2015, release 18); (viii) WES from 364 HCCs from North America, Vietnam, and a few other regions (fig. S5I) (37), downloaded from the supplementary information from that publication.

Analysis of mutational signatures

We used the R (www.r-project.org/) function `prcomp` to compute the first two principal components over 96-channel mutation spectra (Fig. 1F). The sources of the data were as follows: UTUCs and AA-exposed cell lines (29), AA BCs (31), and HCCs from Japan (53) and China (47, 52). Supplementary Materials and Methods and code S1 describe in detail the mSigAct signature presence test and its comparison to the NMF approach in (3, 4).

Allocating mutational signature contributions to mutations (“weighted mutation counts”)

For a given tumor, we allocated the partial contribution of each signature to each type of mutation as follows. Let $t_1 \dots t_{96}$ be the 96 strand

invariant mutation types in trinucleotide context, namely, ACA>AAA, ACA>AGA, ACA>ATA, CCA>CAA, ..., TTT>TAT, TTT>TCT, and TTT>TGT (see also the labels at the bottom of Fig. 1E). Let n_s be the number of mutational signatures, and let $e_1 \dots e_{n_s}$ be the exposures of the tumor to each of the signatures. Let $\pi(t_j, S_i)$ be the proportion of mutation type t_j in signature S_i , with $1 = \sum_{j=1}^{96} \pi(t_j, S_i)$. Then, in a given tumor, we define the partial contribution of S_i to each instance of t_j as

$$\pi(t_j, S_i) \cdot e_i / \sum_{k=1}^{n_s} \pi(t_j, S_k) \cdot e_k$$

in which the denominator ensures that the partial credits sum to 1.

AA signature mutations in driver genes

We used MutSigCV (42) and 20/20+ (43) to identify candidate drivers in the Taiwan AA HCCs (tables S6 to S8). The MutSigCV preprocessor was used to generate the MAF file, and MutSigCV was run with default parameters. 20/20+ was run with default parameters except using the “pretrained 2020plus_100k.Rdata” classifier.

To analyze nonsilent mutations in known cancer driver genes, we used the Cancer Gene Census (<http://cancer.sanger.ac.uk/census>; downloaded 3 January 2017) and selected the 159 genes listed as “oncogene” or “TSG” (tumor suppressor gene) in the “Role in cancer” column (table S12) (54).

Neoantigen prediction

Nonsynonymous somatic variants were annotated by wANNOVAR (<http://wannovar.wglab.org/>) (78), and a custom script generated all possible 9–amino acid sequences containing the mutated residue. In silico HLA typing of individual tumors was carried out using OptiType for major histocompatibility complex class I genes and used for HLA allele-specific peptide binding predictions (79). NetMHC4.0 and NetMHCpan2.8 were used to predict peptide binding (80, 81). Rank parameters > 2 were considered nonbinding, and those ≤ 2 were considered binding, as suggested in (80).

SUPPLEMENTARY MATERIALS

www.sciencetranslationalmedicine.org/cgi/content/full/9/412/eaan6446/DC1
Materials and Methods

Fig. S1. GISTIC analysis of significant amplifications and deletions in Taiwan HCCs.

Fig. S2. Mutational spectra of all 98 individual Taiwan HCCs.

Fig. S3. Comparison of COSMIC signature 22 with AA mutational signature extracted from all Taiwan HCCs.

Fig. S4. Associations between the number of AA signature mutations and clinical and epidemiological variables.

Fig. S5. Mutational spectra of all individual HCCs with the AA signature from publicly available data.

Fig. S6. Examples of clonal and subclonal AA SBS mutations in Taiwan HCCs.

Fig. S7. Two recall notices from Singapore for xi xin products containing aristolochic acid I.

Fig. S8. Length distributions of small indels in 98 Taiwan HCC exomes.

Fig. S9. Workflow for generating synthetic mutation data for testing.

Fig. S10. Receiver operating characteristics of LA-NMF for AA signature detection.

Fig. S11. Receiver operating characteristics for AA detection by mSigAct and LA-NMF.

Fig. S12. Correlations of AA exposure assigned by mSigAct and LA-NMF.

Table S1. Clinicopathological parameters and statistics on sequencing for 98 HCCs and matched normal tissues from Taiwan.

Table S2. Percent targeted bases at $\geq 30\times$ coverage.

Table S3. FDR estimated from IGV screenshots.

Table S4. Somatic SBS mutations in Taiwan HCCs.

Table S5. Somatic indel mutations in Taiwan HCCs.

Table S6. Drivers identified by MutSigCV and 20/20+ in 98 Taiwan HCCs.

Table S7. MutSigCV output for 98 Taiwan HCCs.

Table S8. 20/20+ output for 98 Taiwan HCCs.

Table S9. AA signature mutations and effects on driver genes in 98 Taiwan HCCs.
 Table S10. Comparison of LA-NMF-extracted AA signatures with COSMIC 22.
 Table S11. List of AA signature-positive HCCs from publicly available data.
 Table S12. Known oncogene and tumor suppressor drivers from COSMIC Cancer Gene Census.
 Table S13. Nonsilent mutations in known cancer driver genes plus genes identified by MutSigCV or 20/20+.
 Table S14. Subclonality analysis of AA mutations in published HCC multiregion sequencing studies.
 Table S15. Likely AA-containing plants for sale on the internet.
 Table S16. Selecting the negative binomial dispersion parameter for mSigAct.
 Table S17. True- and false-positive rates for detection of the AA signature by mSigAct and LA-NMF.
 Table S18. Comparison of detection of the AA signature by mSigAct and LA-NMF on 1400 publicly available HCC spectra.
 Code S1. Code for analyses presented in this paper, including mSigAct.v0.8.R and mSigTools.v0.7.R.
 Code S2. Analysis and tests of HCCs with mSigAct and the NMF procedure from (3, 4).
 References (82–90)

REFERENCES AND NOTES

1. S. L. Poon, J. R. McPherson, P. Tan, B. T. Teh, S. G. Rozen, Mutation signatures of carcinogen exposure: Genome-wide detection and new opportunities for cancer prevention. *Genome Med.* **6**, 24 (2014).
2. G. P. Pfeifer, How the environment shapes cancer genomes. *Curr. Opin. Oncol.* **27**, 71–77 (2015).
3. L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, S. A. J. R. Aparicio, S. Behjati, A. V. Biankin, G. R. Bignell, N. Bolli, A. Borg, A.-L. Børresen-Dale, S. Boyault, B. Burkhardt, A. P. Butler, C. Caldas, H. R. Davies, C. Desmedt, R. Eils, J. E. Eyfjörð, J. A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilicic, S. Imbeaud, M. Imielinski, N. Jäger, D. T. W. Jones, D. Jones, S. Knappskog, M. Kool, S. R. Lakhani, C. López-Otin, S. Martin, N. C. Munshi, H. Nakamura, P. A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J. V. Pearson, X. S. Puente, K. Raine, M. Ramakrishna, A. L. Richardson, J. Richter, P. Rosenstiel, M. Schlessner, T. N. Schumacher, P. N. Span, J. W. Teague, Y. Totoki, A. N. J. Tutt, L. R. Valdes-Mas, M. M. van Buuren, L. van 't Veer, A. Vincent-Salomon, N. Waddell, L. R. Yates; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MML-Seq Consortium; ICGC PedBrain, J. Zucman-Rossi, P. A. Futreal, U. McDermott, P. Lichter, M. Meyerson, S. M. Grimmond, R. Siebert, E. Campo, T. Shibata, S. M. Pfister, P. J. Campbell, M. R. Stratton, Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
4. L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, P. J. Campbell, M. R. Stratton, Deciphering signatures of mutational processes operative in human cancer. *Cell Rep.* **3**, 246–259 (2013).
5. L. B. Alexandrov, M. R. Stratton, Mutational signatures: The patterns of somatic mutations hidden in cancer genomes. *Curr. Opin. Genet. Dev.* **24**, 52–60 (2014).
6. J. Michl, O. Bello, G. C. Kite, M. S. J. Simmonds, M. Heinrich, Medicinally used asarum species: High-resolution LC-MS analysis of the aristolochic acid analogues and in vitro toxicity screening in HK-2 cells. *Front. Pharmacol.* **8**, 215 (2017).
7. J. Michl, M. J. Ingrouille, M. S. Simmonds, M. Heinrich, Naturally occurring aristolochic acid analogues and their toxicities. *Nat. Prod. Rep.* **31**, 676–693 (2014).
8. U. Mengs, Tumour induction in mice following exposure to aristolochic acid. *Arch. Toxicol.* **61**, 504–505 (1988).
9. F. D. Debelle, J.-L. Vanherweghem, J. L. Nortier, Aristolochic acid nephropathy: A worldwide problem. *Kidney Int.* **74**, 158–169 (2008).
10. A. P. Grollman, Aristolochic acid nephropathy: Harbinger of a global iatrogenic disease. *Environ. Mol. Mutagen.* **54**, 1–7 (2013).
11. J. Michl, H. M. Jennings, G. C. Kite, M. J. Ingrouille, M. S. J. Simmonds, M. Heinrich, Is aristolochic acid nephropathy a widespread problem in developing countries?: A case study of *Aristolochia indica* L. in Bangladesh using an ethnobotanical-phytochemical approach. *J. Ethnopharmacol.* **149**, 235–244 (2013).
12. L. Vaclavik, A. J. Krynetsky, J. I. Rader, Quantification of aristolochic acids I and II in herbal dietary supplements by ultra-high-performance liquid chromatography-multistage fragmentation mass spectrometry. *Food Addit. Contam. Part A* **31**, 784–791 (2014).
13. H.-Y. Yang, P.-C. Chen, J.-D. Wang, Chinese herbs containing aristolochic acid associated with renal failure and urothelial carcinoma: A review from epidemiologic observations to causal inference. *Biomed. Res. Int.* **2014**, 569325 (2014).
14. T. P. Cheung, C. Xue, K. Leung, K. Chan, C. G. Li, Aristolochic acids detected in some raw Chinese medicinal herbs and manufactured herbal products—A consequence of inappropriate nomenclature and imprecise labelling?. *Clin. Toxicol.* **44**, 371–378 (2006).
15. M.-N. Lai, S.-M. Wang, P.-C. Chen, Y.-Y. Chen, J.-D. Wang, Population-based case-control study of Chinese herbal products containing aristolochic acid and urinary tract cancer risk. *J. Natl. Cancer Inst.* **102**, 179–186 (2010).
16. M. J. Martena, J. C. A. van der Wielen, L. F. J. van de Laak, E. J. M. Konings, H. N. de Groot, I. M. C. M. Rietjens, Enforcement of the ban on aristolochic acid in Chinese traditional herbal preparations on the Dutch market. *Anal. Bioanal. Chem.* **389**, 263–275 (2007).
17. International Agency for Research on Cancer, Some traditional herbal medicines, some mycotoxins, naphthalene and styrene, in *IARC Monographs On The Evaluation Of Carcinogenic Risks To Humans* (IARC Press, 2002), vol. 82.
18. M. Heinrich, J. Chan, S. Wanke, C. Neinhuis, M. S. J. Simmonds, Local uses of *Aristolochia* species and content of nephrotoxic aristolochic acid 1 and 2—A global assessment based on bibliographic sources. *J. Ethnopharmacol.* **125**, 108–144 (2009).
19. *Final Report on Carcinogens Background Document for Aristolochic Acids* (U.S. Department of Health and Human Services, Public Health Services, Public Health Services, 2008); https://ntp.niehs.nih.gov/ntp/roc/twelfth/2010/finalbds/aristolochic_acids_final_508.pdf.
20. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, Some traditional herbal medicines, some mycotoxins, naphthalene and styrene, in *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans* (IARC Press, 2002).
21. M. Stiborová, E. Frei, V. M. Arlt, H. H. Schmeiser, Metabolic activation of carcinogenic aristolochic acid, a risk factor for Balkan endemic nephropathy. *Mutat. Res.* **658**, 55–67 (2008).
22. V. S. Sidorenko, J.-E. Yeo, R. R. Bonala, F. Johnson, O. D. Schärer, A. P. Grollman, Lack of recognition by global-genome nucleotide excision repair accounts for the high mutagenicity and persistence of aristolactam-DNA adducts. *Nucleic Acids Res.* **40**, 2494–2505 (2012).
23. T. H. Broschard, M. Wiessler, C.-W. von der Lieth, H. H. Schmeiser, Translesional synthesis on DNA templates containing site-specifically placed deoxyadenosine and deoxyguanosine adducts formed by the plant carcinogen aristolochic acid. *Carcinogenesis* **15**, 2331–2340 (1994).
24. S. Attaluri, R. R. Bonala, I.-Y. Yang, M. A. Lukin, Y. Wen, A. P. Grollman, M. Moriya, C. R. Iden, F. Johnson, DNA adducts of aristolochic acid II: Total synthesis and site-specific mutagenesis studies in mammalian cells. *Nucleic Acids Res.* **38**, 339–352 (2010).
25. J.-L. Vanherweghem, Misuse of herbal remedies: The case of an outbreak of terminal renal failure in Belgium (Chinese herbs nephropathy). *J. Altern. Complement. Med.* **4**, 9–13 (1998).
26. C.-H. Chen, K. G. Dickman, M. Moriya, J. Zavadil, V. S. Sidorenko, K. L. Edwards, D. V. Gnatenko, L. Wu, R. J. Turesky, X.-R. Wu, Y.-S. Pu, A. P. Grollman, Aristolochic acid-associated urothelial cancer in Taiwan. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8241–8246 (2012).
27. C. J. Chen, L. Y. Wang, S. N. Lu, M. H. Wu, S. L. You, Y. J. Zhang, L. W. Wang, R. M. Santella, Elevated aflatoxin exposure and increased risk of hepatocellular carcinoma. *Hepatology* **24**, 38–42 (1996).
28. S.-C. Hsieh, I.-H. Lin, W.-L. Tseng, C.-H. Lee, J.-D. Wang, 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.* **3**, 13 (2008).
29. S. L. Poon, S.-T. Pang, J. R. McPherson, W. Yu, K. K. Huang, P. Guan, W.-H. Weng, E. Y. Siew, Y. Liu, H. L. Heng, S. C. Chong, A. Gan, S. T. Tay, W. K. Lim, I. Cutcutache, D. Huang, L. D. Ler, M.-L. Nairismagi, M. H. Lee, Y.-H. Chang, K.-J. Yu, W. Chan-on, B.-K. Li, Y.-F. Yuan, C.-N. Qian, K.-F. Ng, C.-F. Wu, C.-L. Hsu, R. M. Bunte, M. R. Stratton, P. A. Futreal, W.-K. Sung, C. K. Chuang, C. K. Ong, S. G. Rozen, P. Tan, B. T. Teh, Genome-wide mutational signatures of aristolochic acid and its application as a screening tool. *Sci. Transl. Med.* **5**, 197ra101 (2013).
30. M. L. Hoang, C.-H. Chen, V. S. Sidorenko, J. He, K. G. Dickman, B. H. Yun, M. Moriya, N. Niknafs, C. Douville, R. Karchin, R. J. Turesky, Y.-S. Pu, B. Vogelstein, N. Papadopoulos, A. P. Grollman, K. W. Kinzler, T. A. Rosenquist, Mutational signature of aristolochic acid exposure as revealed by whole-exome sequencing. *Sci. Transl. Med.* **5**, 197ra102 (2013).
31. S. L. Poon, M. N. Huang, Y. Choo, J. R. McPherson, W. Yu, H. L. Heng, A. Gan, S. S. Myint, E. Y. Siew, L. D. Ler, L. Guatt, W.-H. Weng, C.-K. Chuang, J. S. P. Yuen, S.-T. Pang, P. Tan, B. T. Teh, S. G. Rozen, Mutation signatures implicate aristolochic acid in bladder cancer development. *Genome Med.* **7**, 38 (2015).
32. B. Jelaković, X. Castells, K. Tomić, M. Ardin, S. Karanović, J. Zavadil, Renal cell carcinomas of chronic kidney disease patients harbor the mutational signature of carcinogenic aristolochic acid. *Int. J. Cancer* **136**, 2967–2972 (2014).
33. G. Scelo, Y. Riazalhosseini, L. Greger, L. Letourneau, M. González-Porta, M. B. Wozniak, M. Bourgey, P. Harnden, L. Egevad, S. M. Jackson, M. Karimzadeh, M. Arseneault, P. Lepage, A. How-Kit, A. Daunay, V. Renault, H. Blanché, E. Tubacher, J. Sehmoun, J. Viksna, E. Celms, M. Opanis, A. Zarins, N. S. Vasudev, M. Seywright, B. Abedi-Ardekani, C. Carreira, P. J. Selby, J. J. Cartledge, G. Byrnes, J. Zavadil, J. Su, I. Holcatova, A. Brisuda, D. Zaridze, A. Moukeria, L. Foretova, M. Navratilova, D. Mates, V. Jinga, A. Artemov, A. Nedoluzhko, A. Mazur, S. Rastorguev, E. Boulygina, S. Heath, M. Gut, M.-T. Bihoreau,

- D. Lechner, M. Foglio, I. G. Gut, K. Skryabin, E. Prokhortchouk, A. Cambon-Thomsen, J. Rung, G. Bourque, P. Brennan, J. Tost, R. E. Banks, A. Brazma, G. M. Lathrop, Variation in genomic landscape of clear cell renal cell carcinoma across Europe. *Nat. Commun.* **5** (2014).
34. M. L. Hoang, C.-H. Chen, P.-C. Chen, N. J. Roberts, K. G. Dickman, B. H. Yun, R. J. Turesky, Y.-S. Pu, B. Vogelstein, N. Papadopoulos, A. P. Grollman, K. W. Kinzler, T. A. Rosenquist, Aristolochic acid in the etiology of renal cell carcinoma. *Cancer Epidemiol. Biomarkers Prev.* **25**, 1600–1608 (2016).
35. S. Zou, J. Li, H. Zhou, C. Frech, X. Jiang, J. S. Chu, X. Zhao, Y. Li, Q. Li, H. Wang, Mutational landscape of intrahepatic cholangiocarcinoma. *Nat. Commun.* **5**, 5696 (2014).
36. A. Jusakul, I. Cutcutache, C. H. Yong, J. Q. Lim, M. N. Huang, N. Padmanabhan, V. Nellore, S. Kongpetch, A. W. T. Ng, L. M. Ng, S. P. Choo, S. S. Myint, R. Thanan, S. Nagarajan, W. K. Lim, C. C. Y. Ng, A. Boot, M. Liu, C. K. Ong, V. Rajasegaran, S. Lie, A. S. T. Lim, T. H. Lim, J. Tan, J. L. Loh, J. R. McPherson, N. Khuntikeo, V. Bhudhisawasdi, P. Yongvanit, S. Wongkham, Y. Totoki, H. Nakamura, Y. Arai, S. Yamasaki, P. K.-H. Chow, A. Y. F. Chung, L. L. P. J. Ooi, K. H. Lim, S. Dima, D. G. Duda, I. Popescu, P. Broet, S.-Y. Hsieh, M.-C. Yu, A. Scarpa, J. Lai, D.-X. Luo, A. L. Carvalho, A. L. Vettore, H. Rhee, Y. N. Park, L. B. Alexandrov, R. Gordán, S. G. Rozen, T. Shibata, C. Pairojkul, B. T. Teh, P. Tan, Whole-genome and epigenomic landscapes of etiologically distinct subtypes of cholangiocarcinoma. *Cancer Discov.* (2017).
37. The Cancer Genome Atlas Research Network, Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell* **169**, 1327–1341.e23 (2017).
38. D.-C. Lin, A. Mayakonda, H. Q. Dinh, P. Huang, L. Lin, X. Liu, L.-w. Ding, J. Wang, B. P. Berman, E.-W. Song, D. Yin, H. P. Koeffler, Genomic and epigenomic heterogeneity of hepatocellular carcinoma. *Cancer Res.* **77**, 2255–2265 (2017).
39. W. Zhai, T. K. Lim, T. Zhang, S. T. Phang, Z. Tiang, P. Guan, M. H. Ng, J. Q. Lim, F. Yao, Z. Li, P. Y. Ng, J. Yan, B. K. Goh, A. Y. Chung, S. P. Choo, C. C. Khor, W. W. Soon, K. W. Sung, R. S. Foo, P. K. Chow, The spatial organization of intra-tumour heterogeneity and evolutionary trajectories of metastases in hepatocellular carcinoma. *Nat. Commun.* **8**, 4565 (2017).
40. J.-Y. Guh, H.-C. Chen, J.-F. Tsai, L.-Y. Chuang, Herbal therapy is associated with the risk of CKD in adults not using analgesics in Taiwan. *Am. J. Kidney Dis.* **49**, 626–633 (2007).
41. J.-S. Chen, C.-L. Lu, L.-C. Huang, C.-H. Shen, S. C.-C. Chen, Chronic kidney disease is associated with upper tract urothelial carcinoma: A nationwide population-based cohort study in taiwan. *Medicine* **95**, e3255 (2016).
42. M. S. Lawrence, P. Stojanov, P. Polak, G. V. Kryukov, K. Cibulskis, A. Sivachenko, S. L. Carter, C. Stewart, C. H. Mermel, S. A. Roberts, A. Kiezun, P. S. Hammerman, A. McKenna, Y. Drier, L. Zou, A. H. Ramos, T. J. Pugh, N. Stransky, E. Helman, J. Kim, C. Sougnez, L. Ambrogio, E. Nickerson, E. Shefler, M. L. Cortés, D. Auclair, G. Saksena, D. Voet, M. Noble, D. DiCara, P. Lin, L. Lichtenstein, D. I. Heiman, T. Fennell, M. Imielinski, B. Hernandez, E. Hodis, S. Baca, A. M. Dulak, J. Lohr, D.-A. Landau, C. J. Wu, J. Melendez-Zajgla, A. Hidalgo-Miranda, A. Koren, S. A. McCarroll, J. Mora, R. S. Lee, B. Crompton, R. Onofrio, M. Parkin, W. Winckler, K. Ardlie, S. B. Gabriel, C. W. M. Roberts, J. A. Biegel, K. Stegmaier, A. J. Bass, L. A. Garraway, M. Meyerson, T. R. Golub, D. A. Gordenin, S. Sunyaev, E. S. Lander, G. Getz, Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
43. C. J. Tokheim, N. Papadopoulos, K. W. Kinzler, B. Vogelstein, R. Karchin, Evaluating the evaluation of cancer driver genes. *Proc. Natl. Acad. Sci. U.S.A.* **143**, 14330–14335 (2016).
44. J. Fernandez-Banet, N. P. Lee, K. T. Chan, H. Gao, X. Liu, W.-K. Sung, W. Tan, S. T. Fan, R. T. Poon, S. Li, Decoding complex patterns of genomic rearrangement in hepatocellular carcinoma. *Genomics* **103**, 189–203 (2014).
45. C. Guichard, G. Amaddeo, S. Imbeaud, Y. Ladeiro, L. Pelletier, I. B. Maad, J. Calderaro, P. Bioulac-Sage, M. Letexier, F. Degos, B. Clément, C. Balabaud, E. Chevet, A. Laurent, G. Couchy, E. Letouzé, F. Calvo, J. Zucman-Rossi, Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat. Genet.* **44**, 694–698 (2012).
46. Y. Totoki, K. Tatsuno, K. R. Covington, H. Ueda, C. J. Creighton, M. Kato, S. Tsuji, L. A. Donehower, B. L. Slagle, H. Nakamura, S. Yamamoto, E. Shinbrot, N. Hama, M. Lehmkuhl, F. Hosoda, Y. Arai, K. Walker, M. Dahdouli, K. Gotoh, G. Nagae, M.-C. Gingras, D. M. Muzny, H. Ojima, K. Shimada, Y. Midorikawa, J. A. Goss, R. Cotton, A. Hayashi, J. Shibahara, S. Ishikawa, J. Guiteau, M. Tanaka, T. Urushidate, S. Ohashi, N. Okada, H. Doddapaneni, M. Wang, Y. Zhu, H. Dinth, T. Okusaka, N. Kokudo, T. Kosuge, T. Takayama, M. Fukayama, R. A. Gibbs, D. A. Wheeler, H. Aburatani, T. Shibata, Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. *Nat. Genet.* **46**, 1267–1273 (2014).
47. Z. Kan, H. Zheng, X. Liu, S. Li, T. D. Barber, Z. Gong, H. Gao, K. Hao, M. D. Willard, J. Xu, R. Hauptschein, P. A. Rejto, J. Fernandez, G. Wang, Q. Zhang, B. Wang, R. Chen, J. Wang, N. P. Lee, W. Zhou, Z. Lin, Z. Peng, K. Yi, S. Chen, L. Li, X. Fan, J. Yang, R. Ye, J. Ju, K. Wang, H. Estrella, S. Deng, P. Wei, M. Qiu, I. H. Wulur, J. Liu, M. E. Ehsani, C. Zhang, A. Loboda, W. K. Sung, A. Aggarwal, R. T. Poon, S. T. Fan, J. Wang, J. Hardwick, C. Reinhard, H. Dai, Y. Li, J. M. Luk, M. Mao, Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma. *Genome Res.* **23**, 1422–1433 (2013).
48. S. M. Ahn, S. J. Jang, J. H. Shim, D. Kim, S. M. Hong, C. O. Sung, D. Baek, F. Haq, A. A. Ansari, S. Y. Lee, S.-M. Chun, S. Choi, H.-J. Choi, J. Kim, S. Kim, S. Hwang, Y.-J. Lee, J.-e. Lee, W.-r. Jung, H. Y. Jang, E. Yang, W.-K. Sung, N. P. Lee, M. Mao, C. Lee, J. Zucman-Rossi, E. Yu, H. C. Lee, G. Kong, Genomic portrait of resectable hepatocellular carcinomas: Implications of *RB1* and *FGF19* aberrations for patient stratification. *Hepatology* **60**, 1972–1982 (2014).
49. K. Schulze, S. Imbeaud, E. Letouzé, L. B. Alexandrov, J. Calderaro, S. Rebouissou, G. Couchy, C. Meiller, J. Shinde, F. Soysouvanh, A.-L. Calatayud, R. Pinyol, L. Pelletier, C. Balabaud, A. Laurent, J.-F. Blanc, V. Mazzaferro, F. Calvo, A. Villanueva, J.-C. Nault, P. Bioulac-Sage, M. R. Stratton, J. M. Llovet, J. Zucman-Rossi, Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat. Genet.* **47**, 505–511 (2015).
50. A. Fischer, C. J. R. Illingworth, P. J. Campbell, V. Mustonen, EMu: Probabilistic inference of mutational processes and their localization in the cancer genome. *Genome Biol.* **14**, R39 (2013).
51. R. Rosenthal, N. McGranahan, J. Herrero, B. S. Taylor, C. Swanton, DeconstructSigs: Delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* **17**, 31 (2016).
52. W.-K. Sung, H. Zheng, S. Li, R. Chen, X. Liu, Y. Li, N. P. Lee, W. H. Lee, P. N. Ariyaratne, C. Tennakoon, F. H. Mulawadi, K. F. Wong, A. M. Liu, R. T. Poon, S. T. Fan, K. L. Chan, Z. Gong, Y. Hu, Z. Lin, G. Wang, Q. Zhang, T. D. Barber, W.-C. Chou, A. Aggarwal, K. Hao, W. Zhou, C. Zhang, J. Hardwick, C. Buser, J. Xu, Z. Kan, H. Dai, M. Mao, C. Reinhard, J. Wang, J. M. Luk, Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nat. Genet.* **44**, 765–769 (2012).
53. A. Fujimoto, M. Furuta, Y. Totoki, T. Tsunoda, M. Kato, Y. Shiraiishi, H. Tanaka, H. Taniguchi, Y. Kawakami, M. Ueno, K. Gotoh, S.-i. Arizumi, C. P. Wardell, S. Hayami, T. Nakamura, H. Aikata, K. Arihori, K. A. Borojevich, T. Abe, K. Nakano, K. Maejima, A. Sasaki-Oku, A. Ohsawa, T. Shibuya, H. Nakamura, N. Hama, F. Hosoda, Y. Arai, S. Ohashi, T. Urushidate, G. Nagae, S. Yamamoto, H. Ueda, K. Tatsuno, H. Ojima, N. Hiraoka, T. Okusaka, M. Kubo, S. Marubashi, T. Yamada, S. Hirano, M. Yamamoto, H. Ohdan, K. Shimada, O. Ishikawa, H. Yamaue, K. Chayama, S. Miyano, H. Aburatani, T. Shibata, H. Nakagawa, Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. *Nat. Genet.* **48**, 500–509 (2016).
54. P. A. Futreal, L. Coin, M. Marshall, T. Down, T. Hubbard, R. Wooster, N. Rahman, M. R. Stratton, A census of human cancer genes. *Nat. Rev. Cancer* **4**, 177–183 (2004).
55. A. G. Beer, C. Zenzmaier, M. Schreinlechner, J. Haas, M. F. Dietrich, J. Herz, P. Marschang, Expression of a recombinant full-length LRP1B receptor in human non-small cell lung cancer cells confirms the postulated growth-suppressing function of this large LDL receptor family member. *Oncotarget* **7**, 68721–68733 (2016).
56. Z. Wang, P. Sun, C. Gao, J. Chen, J. Li, Z. Chen, M. Xu, J. Shao, Y. Zhang, J. Xie, Down-regulation of LRP1B in colon cancer promoted the growth and migration of cancer cells. *Exp. Cell Res.* **357**, 1–8 (2017).
57. D. T. Le, J. N. Uram, H. Wang, B. R. Bartlett, H. Kemberling, A. D. Eyring, A. D. Skora, B. S. Luber, N. S. Azad, D. Laheru, B. Biedrzycki, R. C. Donehower, A. Zaheer, G. A. Fisher, T. S. Crocenzi, J. J. Lee, S. M. Duffy, R. M. Goldberg, A. de la Chapelle, M. Koshiji, F. Bhajjee, T. Huebner, R. H. Hruban, L. D. Wood, N. Cuka, D. M. Pardoll, N. Papadopoulos, K. W. Kinzler, S. Zhou, T. C. Cornish, J. M. Taube, R. A. Anders, J. R. Eshleman, B. Vogelstein, L. A. Diaz Jr., PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015).
58. R. Siegel, D. Naishadham, A. Jemal, Cancer statistics, 2012. *CA Cancer J. Clin.* **62**, 10–29 (2012).
59. H. H. Schmeiser, J. L. Nortier, R. Singh, G. Gamboa da Costa, J. Sennesael, E. Cassuto-Viguier, D. Ambrosetti, S. Rorive, A. Pozdzik, D. H. Phillips, M. Stiborova, V. M. Arlt, 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* **135**, 502–507 (2014).
60. Health Sciences Authority, ST XI XIN product recall (2014); www.hsa.gov.sg/content/hsa/en/Health_Products_Regulation/Safety_Information_and_Product_Recalls/Product_Recalls/2014/ST_XI_XIN.html.
61. Health Sciences Authority, E-Fong Brand Xi Xin product recall (2014); www.hsa.gov.sg/content/hsa/en/Health_Products_Regulation/Safety_Information_and_Product_Recalls/Product_Recalls/2014/e-fong-brand-xi-xin.html.
62. T.-T. Jong, M.-R. Lee, S.-S. Hsiao, J.-L. Hsai, T.-S. Wu, S. T. Chiang, S.-Q. Cai, Analysis of aristolochic acid in nine sources of Xixin, a traditional Chinese medicine, by liquid chromatography/atmospheric pressure chemical ionization/tandem mass spectrometry. *J. Pharm. Biomed. Anal.* **33**, 831–837 (2003).
63. Wellcome Trust Sanger Institute, COSMIC: Signatures of Mutational Processes in Human Cancer (2016); <http://cancer.sanger.ac.uk/cosmic/signatures>.
64. R. C. Fernando, H. H. Schmeiser, H. R. Scherf, M. Wiessler, 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.*, 167–171 (1992).
65. S. Shibutani, H. Dong, N. Suzuki, S. Ueda, F. Miller, A. P. Grollman, Selective toxicity of aristolochic acids I and II. *Drug Metab. Dispos.* **35**, 1217–1222 (2007).

66. W. Chan, H. Yue, W. T. Poon, Y.-W. Chan, O. J. Schmitz, D. W. J. Kwong, R. N. S. Wong, Z. Cai, Quantification of aristolochic acid-derived DNA adducts in rat kidney and liver by using liquid chromatography–electrospray ionization mass spectrometry. *Mutat. Res.* **646**, 17–24 (2008).
67. China Food and Drug Administration, Notice concerning strengthened supervision of guang fang ji and six other herbs and their pharmaceutical preparations (2004); www.sda.gov.cn/WS01/CL0844/10242.html.
68. United States Congress, Dietary Supplement Health and Education Act of 1994 (1994); www.gpo.gov/fdsys/pkg/STATUTE-108/pdf/STATUTE-108-Pg4325.pdf.
69. C. L. Wah, S. C. Hock, T. K. Yun, Current scientific status and regulatory control of traditional/herbal medicinal products: Globalization challenges. *Pharm. Eng.* **32**, 1–15 (2012).
70. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
71. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
72. K. Okonechnikov, A. Conesa, F. García-Alcalde, Qualimap 2: Advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292–294 (2016).
73. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytksy, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
74. C. T. Saunders, W. S. W. Wong, S. Swamy, J. Becq, L. J. Murray, R. K. Cheetham, Strelka: Accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* **28**, 1811–1817 (2012).
75. K. Cibulskis, M. S. Lawrence, S. L. Carter, A. Sivachenko, D. Jaffe, C. Sougnez, S. Gabriel, M. Meyerson, E. S. Lander, G. Getz, Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
76. M. N. Huang, W. Yu, W. W. Teoh, M. Ardin, A. Jusakul, A. W. T. Ng, A. Boot, B. Abedi-Ardekani, S. Villar, S. S. Myint, R. Othman, S. L. Poon, A. Heguy, M. Olivier, M. Hollstein, P. Tan, B. T. Teh, K. Sabapathy, J. Zavadil, S. G. Rozen, Genome-scale mutational signatures of aflatoxin in cells, mice, and human tumors. *Genome Res.* **27**, 1475–1486 (2017).
77. E. Cerami, J. Gao, U. Dogrusoz, B. E. Gross, S. O. Sumer, B. A. Aksoy, A. Jacobsen, C. J. Byrne, M. L. Heuer, E. Larsson, Y. Antipin, B. Reva, A. P. Goldberg, C. Sander, N. Schultz, The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).
78. H. Yang, K. Wang, Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat. Protoc.* **10**, 1556–1566 (2015).
79. A. Szolek, B. Schubert, C. Mohr, M. Sturm, M. Feldhahn, O. Kohlbacher, OptiType: Precision HLA typing from next-generation sequencing data. *Bioinformatics* **30**, 3310–3316 (2014).
80. M. Nielsen, M. Andreatta, NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med.* **8**, 33 (2016).
81. M. Andreatta, M. Nielsen, Gapped sequence alignment using artificial neural networks: Application to the MHC class I system. *Bioinformatics* **32**, 511–517 (2016).
82. V. Y. Tan, C. Févotte, Automatic relevance determination in nonnegative matrix factorization with the β -divergence. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 1592–1605 (2013).
83. S. Kasar, J. Kim, R. Improgo, G. Tiao, P. Polak, N. Haradhvala, M. S. Lawrence, A. Kiezun, S. M. Fernandes, S. Bahl, C. Sougnez, S. Gabriel, E. S. Lander, H. T. Kim, G. Getz, J. R. Brown, Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat. Commun.* **6**, 8866 (2015).
84. J. S. Gehring, B. Fischer, M. Lawrence, W. Huber, SomaticSignatures: Inferring mutational signatures from single-nucleotide variants. *Bioinformatics* **31**, 3673–3675 (2015).
85. M. Ardin, V. Cahais, X. Castells, L. Bouaoun, G. Byrnes, Z. Herceg, J. Zavadil, M. Olivier, MutSpec: A Galaxy toolbox for streamlined analyses of somatic mutation spectra in human and mouse cancer genomes. *BMC Bioinformatics* **17**, 170 (2016).
86. R. A. Rosales, R. D. Drummond, R. Valieris, E. Dias-Neto, I. T. da Silva, signeR: An empirical Bayesian approach to mutational signature discovery. *Bioinformatics* **33**, 8–16 (2016).
87. A. Ozerov, C. Févotte, R. Blouet, J.-L. Durrieu, Multichannel nonnegative tensor factorization with structured constraints for user-guided audio source separation, 2011 *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, 22 to 27 May 2011 (IEEE, 2011).
88. M. J. Powell, *Advances in Optimization and Numerical Analysis* (Springer, 1994), pp. 51–67.
89. A. Pascual-Montano, J. M. Carazo, K. Kochi, D. Lehmann, R. D. Pascual-Marqui, Nonsmooth nonnegative matrix factorization (nsNMF). *IEEE Trans. Pattern Anal. Mach. Intell.* **28**, 403–415 (2006).
90. C. H. Mermel, S. E. Schumacher, B. Hill, M. L. Meyerson, R. Beroukhim, G. Getz, GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).

Acknowledgments: We thank M. Chua, W. K. Lim, and Y. C. Tang for comments on the manuscript; S. S. Myint, J. L. Loh, and R. Vikneswari for technical assistance; and S.-C. Ho for translations of regulatory and informational documents from Taiwan. The results here are partly based on data generated by the TCGA Research Network (<http://cancergenome.nih.gov/>).

Funding: Funding was provided by the Singapore National Medical Research Council (NMRC/CIRG/1422/2015) to S.G.R., the Singapore Ministry of Health via the Duke-NUS Signature Research Programmes, and the Chang Gung Medical Foundation (CMRPG3F0031-3) to S.Y.H.

Author contributions: A.Y.C., B.T.T., H.-Y.H., M.-C.Y., C.C.Y.N., P.-H.L., P.T., S.G.R., S.L.P., S.-T.P., and S.-Y.H. planned the project and coordinated laboratory work. A.W.T.N., J.Q.L., M.N.H., S.G.R., S.T., W.Y., and Y.S. carried out the bioinformatics analysis. A.B., A.W.T.N., M.N.H., S.G.R., S.L.P., and W.Y. created the figures and drafted the manuscript. A.W.T.N., B.T.T., S.G.R., A.B., S.L.P., and S.-Y.H. edited the manuscript. A.W.T.N. and S.G.R. organized the final manuscript preparation. A.Y.C., H.-Y.H., M.-C.Y., P.-H.L., and S.-Y.H. contributed patient samples and clinical information. B.T.T., S.-Y.H., and S.G.R. provided funding for this study.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: Sequencing data are available at the European Genome-phenome Archive (EGA; www.ebi.ac.uk/ega/) under accession EGAS00001002301. Additional large data sets (full output from ASCAT and GISTIC and IGV screenshots used in estimating FDRs) are available at <https://doi.org/10.5281/zenodo.898817> (72 Mb).

Submitted 12 May 2017
Resubmitted 31 July 2017
Accepted 25 September 2017
Published 18 October 2017
10.1126/scitranslmed.aan6446

Citation: A. W. T. Ng, S. L. Poon, M. N. Huang, J. Q. Lim, A. Boot, W. Yu, Y. Suzuki, S. Thangaraju, C. C. Y. Ng, P. Tan, S.-T. Pang, H.-Y. Huang, M.-C. Yu, P.-H. Lee, S.-Y. Hsieh, A. Y. Chang, B. T. Teh, S. G. Rozen, Aristolochic acids and their derivatives are widely implicated in liver cancers in Taiwan and throughout Asia. *Sci. Transl. Med.* **9**, ean6446 (2017).

Science Translational Medicine

Aristolochic acids and their derivatives are widely implicated in liver cancers in Taiwan and throughout Asia

Alvin W. T. Ng, Song Ling Poon, Mi Ni Huang, Jing Quan Lim, Arnoud Boot, Willie Yu, Yuka Suzuki, Saranya Thangaraju, Cedric C. Y. Ng, Patrick Tan, See-Tong Pang, Hao-Yi Huang, Ming-Chin Yu, Po-Huang Lee, Sen-Yung Hsieh, Alex Y. Chang, Bin T. Teh and Steven G. Rozen

Sci Transl Med 9, eaan6446.
DOI: 10.1126/scitranslmed.aan6446

The dark side of an herbal medicine

Aristolochic acid, an herbal compound found in many traditional medicines, had been previously linked to kidney failure, as well as cancers of the urinary tract. Because of these known toxicities, herbs containing this compound have been restricted or banned in some countries, but it is still available on the internet and in alternate formulations. By analyzing numerous samples from Taiwan and other countries in Asia and elsewhere, Ng *et al.* demonstrated the effects of aristolochic acid in hepatocellular carcinoma, a much more common tumor type. The authors showed that the use of this drug remains widespread in Asia and particularly in Taiwan, and that it appears to increase the risk of multiple different cancer types.

ARTICLE TOOLS

<http://stm.sciencemag.org/content/9/412/eaan6446>

SUPPLEMENTARY MATERIALS

<http://stm.sciencemag.org/content/suppl/2017/10/16/9.412.eaan6446.DC1>

RELATED CONTENT

<http://stm.sciencemag.org/content/scitransmed/5/197/197ra101.full>
<http://stm.sciencemag.org/content/scitransmed/5/197/197pc5.abstract>
<http://stm.sciencemag.org/content/scitransmed/5/197/197ra102.full>
<http://stm.sciencemag.org/content/scitransmed/5/197/197fs31.full>
<http://stm.sciencemag.org/content/scitransmed/10/437/eaat3741.full>
<http://stm.sciencemag.org/content/scitransmed/10/437/eaap9840.full>
<http://science.sciencemag.org/content/sci/360/6391/eaan5931.full>
<http://stm.sciencemag.org/content/scitransmed/10/457/eaar7939.full>

REFERENCES

This article cites 76 articles, 10 of which you can access for free
<http://stm.sciencemag.org/content/9/412/eaan6446#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

Science Translational Medicine (ISSN 1946-6242) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Translational Medicine* is a registered trademark of AAAS.

Copyright © 2017 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works