Non-invasive Detection of Aristolochic Acid Exposure Using Ultra-Sensitive Duplex Sequencing™

Fang Yin Lo[1] – Charles Valentine[1] – Elizabeth Schmidt[1] – Lindsey Williams[1] – Arnoud Boot[2] – Steve Rozen[2] – Jesse Salk[1]

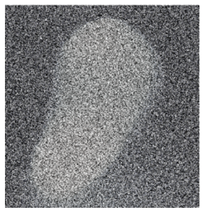
TwinStrand Biosciences, Seattle, WA[1] – Duke-NUS Graduate Medical School, Cancer and Stem Cell Biology, Singapore

Introduction

Aristolochic acids (AA) are a potent class of mutagen that can be found in plants of the genera *Aristolochia* and *Asarum*. Exposures to AA are caused by intentional ingestion of traditional herbal medicines containing *Aristolochia* or through inadvertent consumption from contaminated crops. The carcinogenic potential and epidemiologic connection of AA exposure with upper-tract urothelial cancers (UTUC), liver cancer, and kidney failure have been known for years. Whole exome sequencing shows that 78% of HCCs from Taiwan exhibit distinct mutation signatures of AA exposure. Exposure to AA is geographically widespread, therefore, there is a major opportunity for primary and secondary prevention of AA-associated diseases if it were possible to detect the exposure through non-invasively sampled body fluids such as urine sediment or blood.

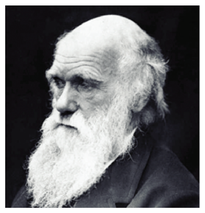
Duplex Sequencing is a tag-based NGS error-correction method that enables the detection of ultra-rare mutations that exist at the mutation frequency of normal tissues (1×10^{-8}). We applied Duplex Sequencing to the tissues and liquid biopsies of cancer patients with known AA-exposure status. We demonstrate that Duplex Sequencing is a promising approach for the detection of mutagenic signatures caused by environmental carcinogens, and we foresee Duplex Sequencing as being a powerful tool to be used for measuring life-integrated carcinogenic processes and cancer risk.

Sequencing Errors Obscure Truth



Next-Generation Sequencing (NGS)

Single Strand Error-Corrected NGS

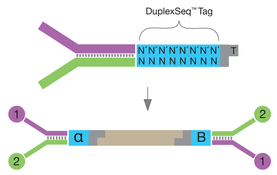


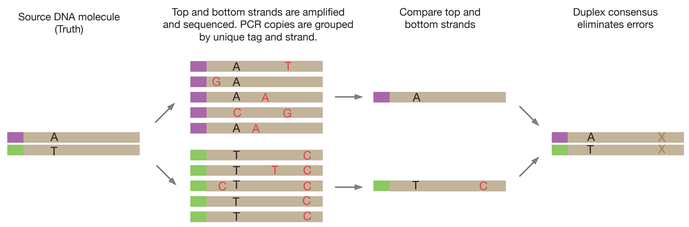
Duplex Sequencing

TwinStrand Duplex Sequencing™ Technology

A DuplexSeq™ Adapter has:

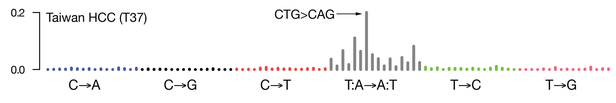
1. Identical (or relatable) degenerate tags in each strand.
2. An asymmetry allowing independent strand identification.





MKT0005 v1.0 For Research Use Only. Not for use in diagnostic procedures. ©2020 TwinStrand Biosciences, Inc. All rights reserved. All trademarks are the property of TwinStrand Biosciences, Inc. or their respective owners.

Exome Sequencing of Tumors Reveal Distinct Aristolochic Acid Trinucleotide Mutational Signature

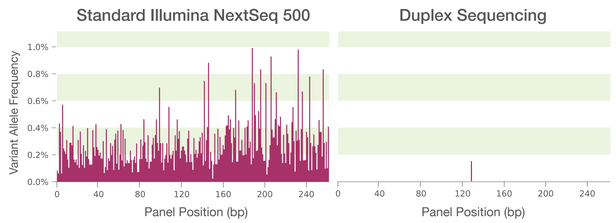


Alvin W. T. Ng *et al.*, Sci Transl Med 2017;9

Exome sequencing of hepatocellular carcinomas (HCC) from AA-exposed patients in Taiwan have revealed significant levels of the canonical AA mutational signature. This pattern is highly specific and indicative of AA exposure and has been cataloged in the COSMIC mutation signatures set as SBS22. The AA signature is marked by a high abundance of T:A→A:T transversions with a peak of mutations occuring in the trinucleotide context CTG and, to a lesser extent, CTA and ATG contexts.

NGS Accuracy is Required to Detect *in vivo* Mutagenesis Without Prior Clonal Amplification

We developed a Duplex Sequencing based assay to test whether error-corrected NGS is capable of detecting rare somatic mutations in the normal non-neoplastic tissues of patients that were confirmed to have AA exposure through either whole genome sequencing (WGS) or whole exome sequencing (WES) of tumors elsewhere. We sought to postively identify mutational patterns in those patients through liquid biopsy (urine sediment or blood) that could also be attributed to AA.



The error rate of standard NGS is ~0.1% which creates a background that obscures rare variants. Common sources of error include sequencer artifacts, PCR misincorporations, and DNA damaging events such as 8-oxoguanine adducts, deaminated cytosines, and abasic sites. Duplex Sequencing overcomes these errors by forming consensus among PCR duplicates from the same source DNA strand, leaving only a true biological signal of rare DNA mutation.

Experimental Design

6 patients not exposed to AA

Tumor (UTUC) x 6
Normal (Urothelium) x 6
Blood x 0
Urine x 0
Kidney x 0

6 patients exposed to AA

Tumor (UTUC) x 6
Normal (Urothelium) x 5
Blood x 5
Urine x 3
Kidney x 1

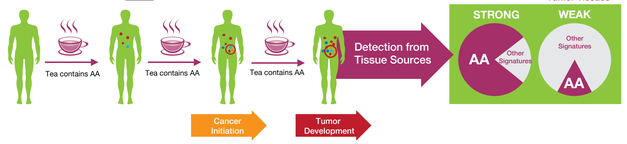
Duplex Sequencing with Human Mutagenesis Panel

Total 32 samples

Evidence for AA Exposure is Modulated By Lifestyle

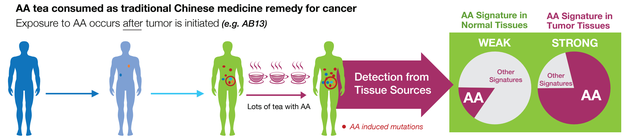
Individuals are exposed to AA through herbal teas containing *Aristolochia* leaves

Chronic exposure to AA before tumor starts development *(e.g. AB19 and AB34)*



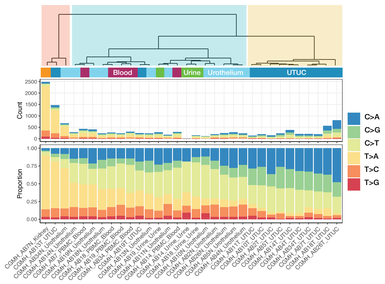
AA tea consumed as traditional Chinese medicine remedy for cancer

Exposure to AA occurs after tumor is initiated *(e.g. AB13)*



Simple Base Substitution Spectra

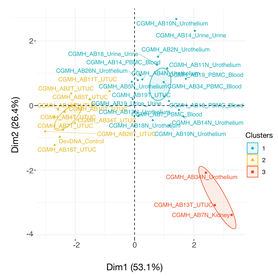
Unsupervised Clustering of Simple Spectra Reveal 3 Clusters



The base substitution spectra for all samples from all patients was clustered using Ward's algorithm with the distance metric being a cosine similarity between the spectra. Three distinct groups were revealed after the tree was cut.

One outgroup contained only three samples and was highly correlated with patterns of AA mutagenesis. Nearly all tumor tissues cluster together in a second outgroup. Finally, all the remaining samples across diverse tissue types form a single ingroup in which no observable patterns emerge.

The Principal Components of 3 Simple Spectra Clusters

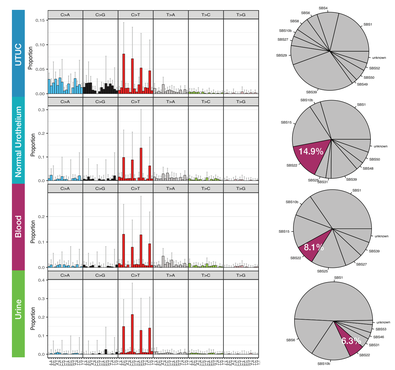


Principal component analysis of base substitution spectra also reveals three clusters and shows the relationship between samples. The variance that drives Cluster 3 can be explained by strong features of AA mutagenesis and high mutation frequencies. Cluster 2 consists of 10 tumor samples and have increased proportions of C→A and C→G mismatches suggesting higher rates of endogenous damage. Finally, 19 out of the 32 samples are located in Cluster 1 and feature patterns of spontaneous deamination from methylated cytosines. This is characteristic of normal aging or in the cases of AA exposure, faint exposure signals.

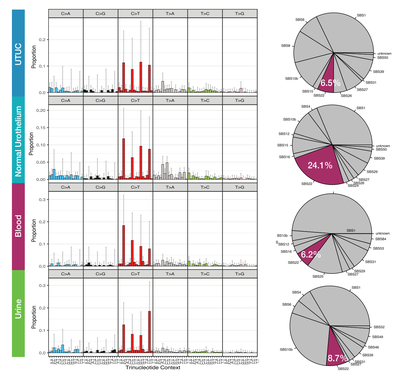
Trinucleotide Spectra Visualization and Deconvolution

AA Signature is Positively Detected in Blood and Urine from Exposed Patients

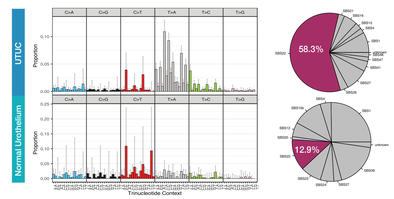
Patient AB18



Patient AB19



Patient AB13



Trinucleotide spectra, and deconvolution of those spectra through non-negative matrix factorization, result in a complex assignment of predicted signature contributions to the overall mutation landscape for each sample. Each sample was deconvolved using the 67 COSMIC base substitution signatures as a reference database. In 3 highlighted patients that were confirmed to have been exposed to AA through WES/WGS, we detect traces of AA mutagenesis in the normal, non-tumor, tissues at rare somatic frequencies. Contributions of AA signature was observed in all available non-invasively sampled body fluids as well as 5 normal urothelium samples from AA exposed subjects. Most exposed patients (4 out of 5) had AA signature at higher abundance in normal tissues as compared to tumors. Curiously, in patient AB13, the AA signature is stronger in tumor samples.

AA Mutational Signature was Detected in Most Normal Samples from AA-Exposed Individuals

Subjects	AA Signature Detected by WES/WGS	AA Signature Detected by Duplex Sequencing				
	Tumor	Tumor	Paired Normal	Blood	Urine	Kidney
AB10	No	No	No	N/A	N/A	N/A
AB13	No	Yes	Yes	N/A	N/A	N/A
AB2	No	No	No	N/A	N/A	N/A
AB26	No	No	No	N/A	N/A	N/A
AB4	No	No	No	N/A	N/A	N/A
AB5	No	No	No	N/A	N/A	N/A
AB11	Yes	No	No	N/A	N/A	N/A
AB14	Yes	No	Yes	Yes	Yes	N/A
AB18	Yes	No	Yes	Yes	Yes	N/A
AB19	Yes	Yes	Yes	Yes	Yes	N/A
AB34	Yes	Yes	Yes	Yes	N/A	N/A
AB7	Yes	No	N/A	Yes	N/A	Yes

Patient AA exposure status was initially determined through WES/WGS sequencing of clonal somatic mutations from tumor samples only. Using a Duplex Sequencing mutagenesis assay, we detect traces of AA signature in 6 patients including one with no prior exposure status. The proportion of mutations attributed to AA exposure ranged from 2.5% in blood to 67.8% in kidney. These mutations, on the order of one-in-a-million, were exposure-related and not tumor derived which was validated by comparing somatic genotypes between tissue and tumor. Patients presenting higher AA signature in the tumor like AB13 are likely to have been exposed to AA during tumourigenesis instead of before disease which may be driven by AA's use as an herbal remedy and supposed treatment for cancer in Taiwan. Non-invasive samples from blood and urine for patients without exposure were not available at the time of study and are being procured.

Conclusions

Duplex-Sequencing has the potential to identify at-risk members of the public who have been exposed to Aristolochic Acid or other environmental carcinogenic agents:

- A mutational signature was detected successfully in all liquid biopsy samples of AA-exposed patients. This is the first time we are aware that an AA mutation signature has been detected in non-invasively sampled body fluids.

Duplex-Sequencing measures the life integrated exposure of mutagenic compounds or processes:

- The mutational signature of spontaneous deamination, which is primarily caused by aging, was detected in all but 3 samples – 2 tumor samples and 1 urine sample that had very low DNA quantity.
- In the blood of patient AB14, ~34% of the mutations were correlated with platinum drug treatment, and, indeed, this patient was later confirmed to have gone through a cisplatin regiment and was unique to this cohort.

Contact: https://twinstrandbio.com/contact/