# Mutational signatures:
# What caused the mutations in these cancers?
# Why do we care?

Steven G Rozen

steverozen@gmail.com

Duke-NUS Medical School, Singapore

InCoB 2020

19th International Conference on Bioinformatics

Virtual meeting

2020 Nov 28 (35 min)

(updated 2020 Dec 24)

# Outline

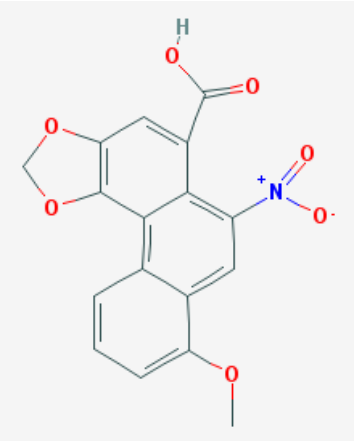
- What are mutational signatures and what are they good for?
- Computational analysis of mutational signatures – state of the art
- Important unsolved problems in mutational signature analysis
- Summary

# Outline

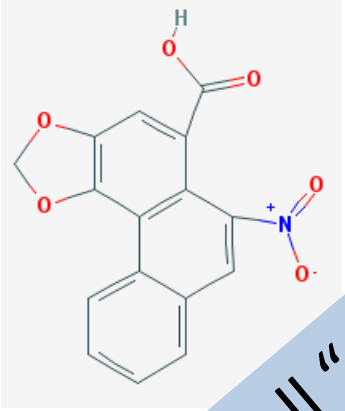
- What are mutational signatures and what are they good for?
- Computational analysis of mutational signatures – state of the art
- Important unsolved problems in mutational signature analysis
- Summary

# Aristolochic acids and relatives **"AA"**

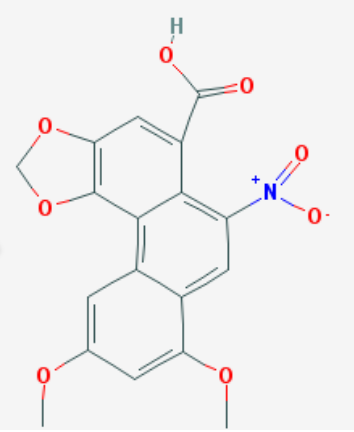
Aristolochic acid I



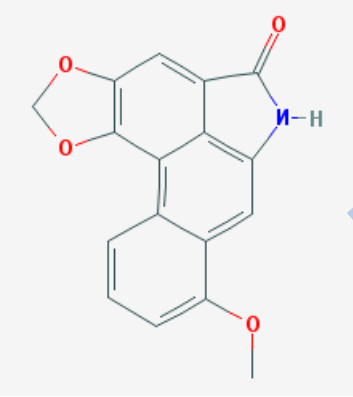
Aristolochic acid II



Aristolochic acid IV

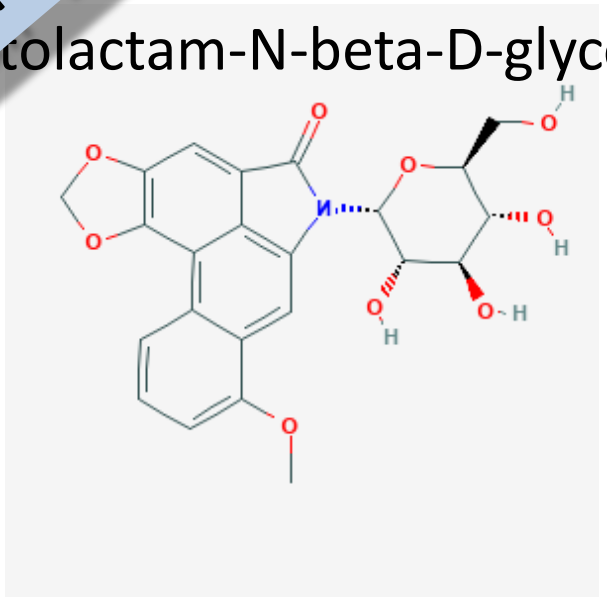


Aristolactam I, proximal mutagen

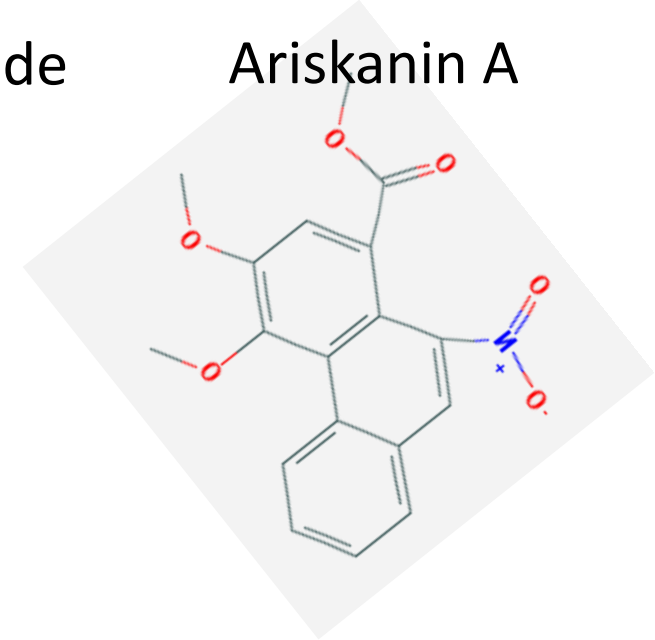


We will call them all "**AA**"

tolactam-N-beta-D-glycoside



Ariskanin A

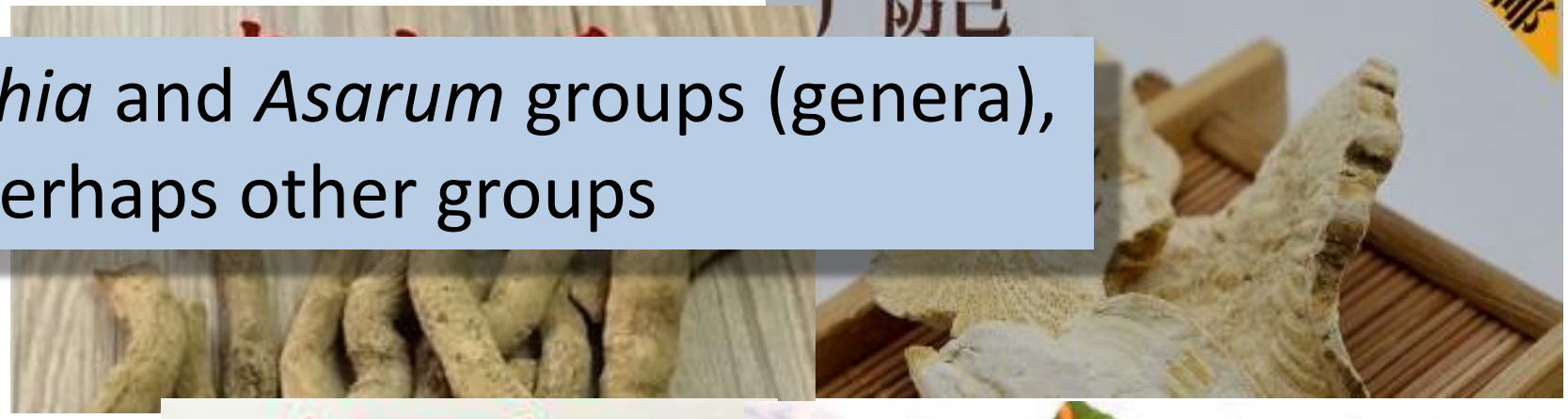


# Plants with AA widely used as herbal medicine

(Occurs naturally in these plants)

Plants in *Aristolochia* and *Asarum* groups (genera), and perhaps other groups




# Not just Chinese herbal medicine

DukeNUS Medical School

## Romania

**Mărul-lupului, beneficii. Combate CANCERUL şi ULCERUL, previne CĂDEREA părului şi tratează HEMOROIZII**

16 feb 2015 / 16:34 Salveaza PDF Comentarii

*Aristolochia clematitis*

Combats cancer, ulcers, prevents hair loss and hemorrhoids

## India

ayurveda cart | Search Herbs and Medicines | India

CLASSICAL MEDICINES | HERBS | MEDICINE BY DISORDERS | SINGLE HERB FORMULATION | PANCHKARMA OILS | ESSENTIAL OILS

Home > Ishwari >

Ishwari (ARISTOLOCHIA INDICA)

Be the first

Sold and fu

Ask Seller a

*Aristolochia indica*

Please select the varition of this herb you would like to purchase, Single Unit is **350 gm**

| PRODUCT NAME | PRICE | QTY |
|---|---|---|
| **Ishwari's Raw Herb** it is raw plant which is 99% dried | Rs.750 | 0 |
| **Ishwari's Powder** it is the powder of Raw herb | Rs.1,030 | 0 |
| **Ishwari's Oil** For external applications | Rs.1,500 | 0 |
| **Ishwari's Extract** Extract is prepared by water/solvent based extraction process, it is the most potent form | Rs.3,750 | 0 |

## Brazil


*Aristolochia cymbifera*

# AA: well-known nephrotoxins

XV.

Arbeiten aus dem pharmakologischen Institut der deutschen Universität zu Prag.

**29. Ueber das Aristolochin, einen giftigen Bestandtheil der Aristolochia-Arten.**

Von

Dr. Julius Pohl,
Assistent des Instituts.

1891

# 1990s: AA emerges as a human health problem

- ~100 young women with end stage kidney disease
- Single weight-loss clinic in Europe
- 汉防己, **hàn** fáng jǐ replaced by 广防己, **guǎng** fáng jǐ

# 漢防己 / 汉防己, **hàn** fáng jǐ (genus *Stephania* – no AA)

## Fang Ji (Fen) 汉防己

from: $1.25

Chinese Herb: Fang Ji (Fen)
(Stephania Root)

Fang Ji (Fen) acts to dispel wind and dampness to relieve pain and promotes diuresis.

Quantity

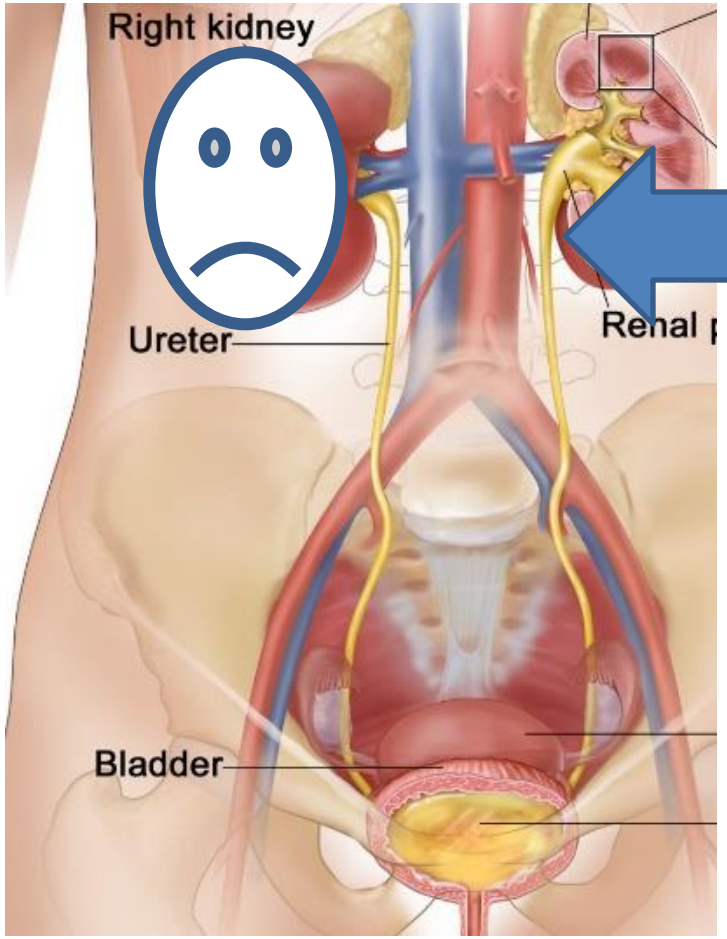
Choose an option

1 + -

Add to cart

# 廣防己/广防己, **guǎng** fáng jǐ (genus *Aristolochia* – has AA)

# Many of the AA-kidney-failure victims developed upper tract urothelial cancer




Urothelial cancer in ureter (tube that drains kidney)

(Nortier et al., 2000, NEJM)

# Quick review: somatic mutations

- Every time our cells divide, mutations arise
- Mutations not present in the fertilized egg are **somatic mutations**
- Our focus will be somatic mutations

# We detect somatic mutations by DNA sequencing normal cells

Normal

…ACGTCCTAGTCAAAATCGAGCC…

# We detect somatic mutations by DNA sequencing normal cells and tumour cells

Normal

…ACGTCCTAGTCAAAATCGAGCC…

Tumour

…ACGTCCTAGTCTAAATCGAGCC…

We detect somatic mutations by DNA sequencing normal cells and tumour cells

We look for the variants found only in the tumor

Normal

Tumour

…ACGTCCTAGTCAAAATCGAGCC…

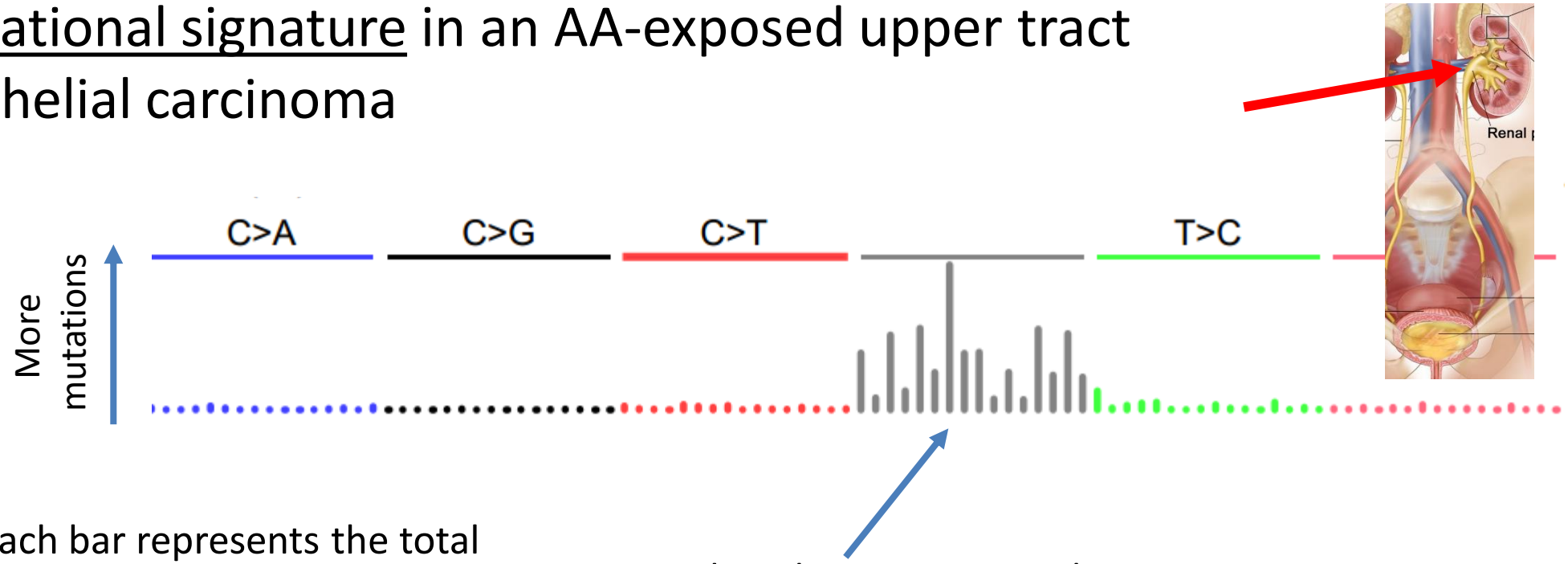
…ACGTCCTAGTCTAAATCGAGCC…

Somatic mutation from **A** > **T**

CAA > CTA ➡ trinucleotide context: information in neighboring bases

- We look at all mutations (most are not causing the cancer; they are innocent bystanders)
- We want to know: what mutated (changed) the DNA?

# Mutational signature in an AA-exposed upper tract urothelial carcinoma



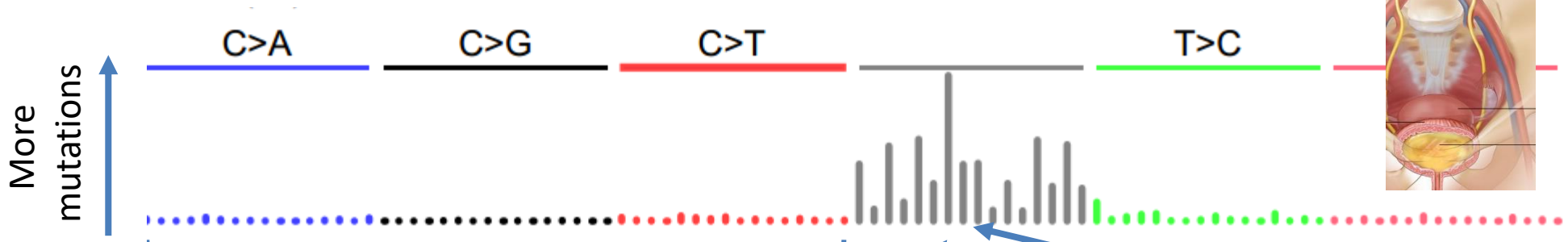

Each bar represents the total number of single nucleotide substitutions in a particular trinucleotide context

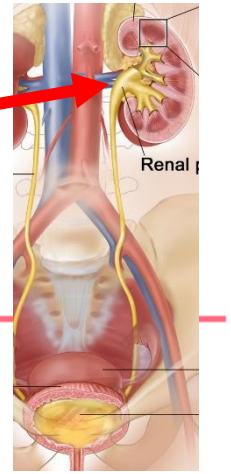
Highest bar represents the number of C**A**G to C**T**G mutations across the entire genome

The higher the bar, the more C**A**G to C**T**G mutations

Poon et al, Science Translational Medicine, 2013

# Mutational signature in an AA-exposed upper tract urothelial carcinoma




AAG > ATG mutations
Not as many

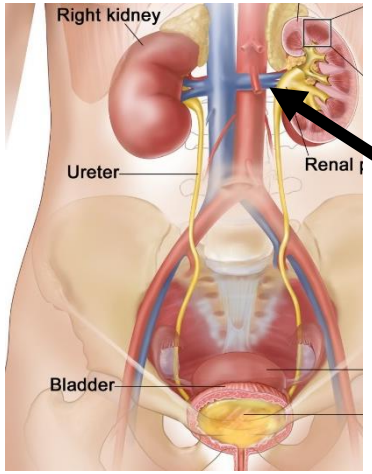
Mutations from C to A, C to G, C to T
Almost none

Highest bar represents the number of CAG to CTG mutations across the entire genome

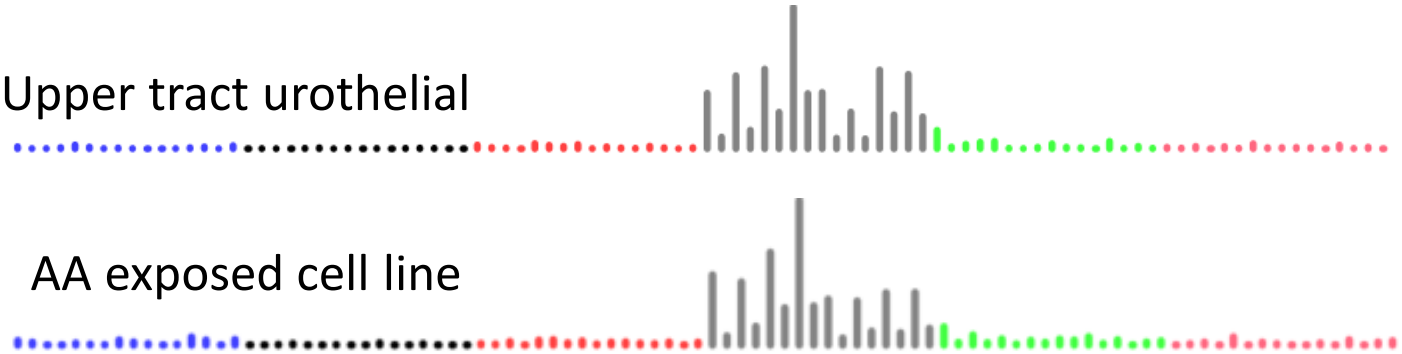
The higher the bar, the more CAG to CTG mutations

Poon et al, Science Translational Medicine, 2013

# 6 years ago (2013):
# AA in upper tract urothelial cancer and cell lines



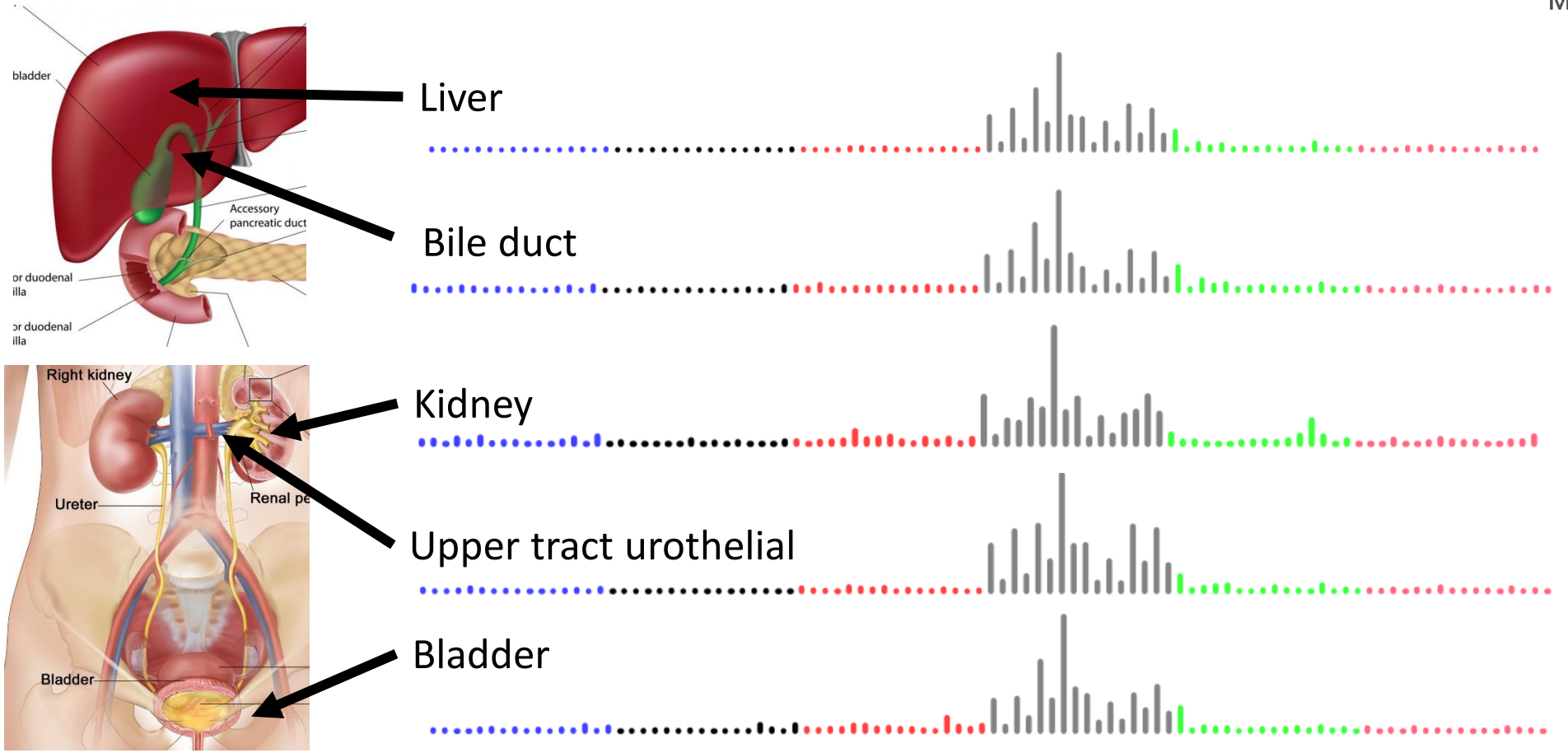

Upper tract urothelial



AA exposed cell line

Poon et al., 2013, Hoang et al. 2013

# 3 years ago:
# Mutational signatures showed AA in multiple tumor types



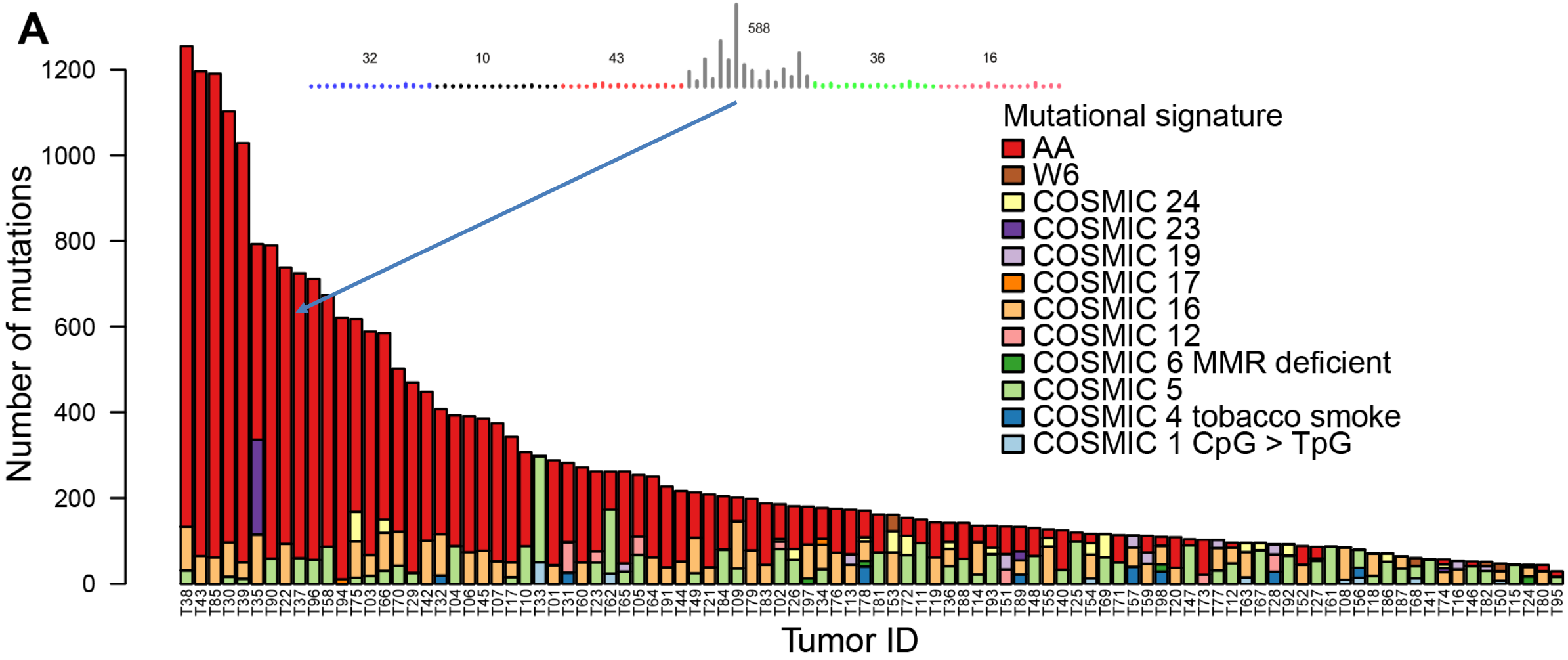

Poon et al., 2013 and subsequent data (liver - hepatocellular carcinoma [HCC])
Zou et al., 2015, Jusakul et al., 2017 (bile duct carcinoma)
Scelo et al., 2014, Jelakovic et al., 2014 (kidney/renal cell carcinoma)
Poon et al., 2013, Hoang et al., 2013, many others (upper tract urothelial carcinoma)
Poon et al., 2015, others (bladder urothelial carcinoma)

# Taiwan a known hotspot for AA exposure but AA in liver cancer there not studied

Chen, …., Grollman, PNAS, 2012

Map created by Freepik

# 78% of Taiwan liver cancers (HCCs) had the AA signature




Ng, Poon, et al., Oct 2017 DOI: 10.1126/scitranslmed.aan6446

DukeNUS Medical School

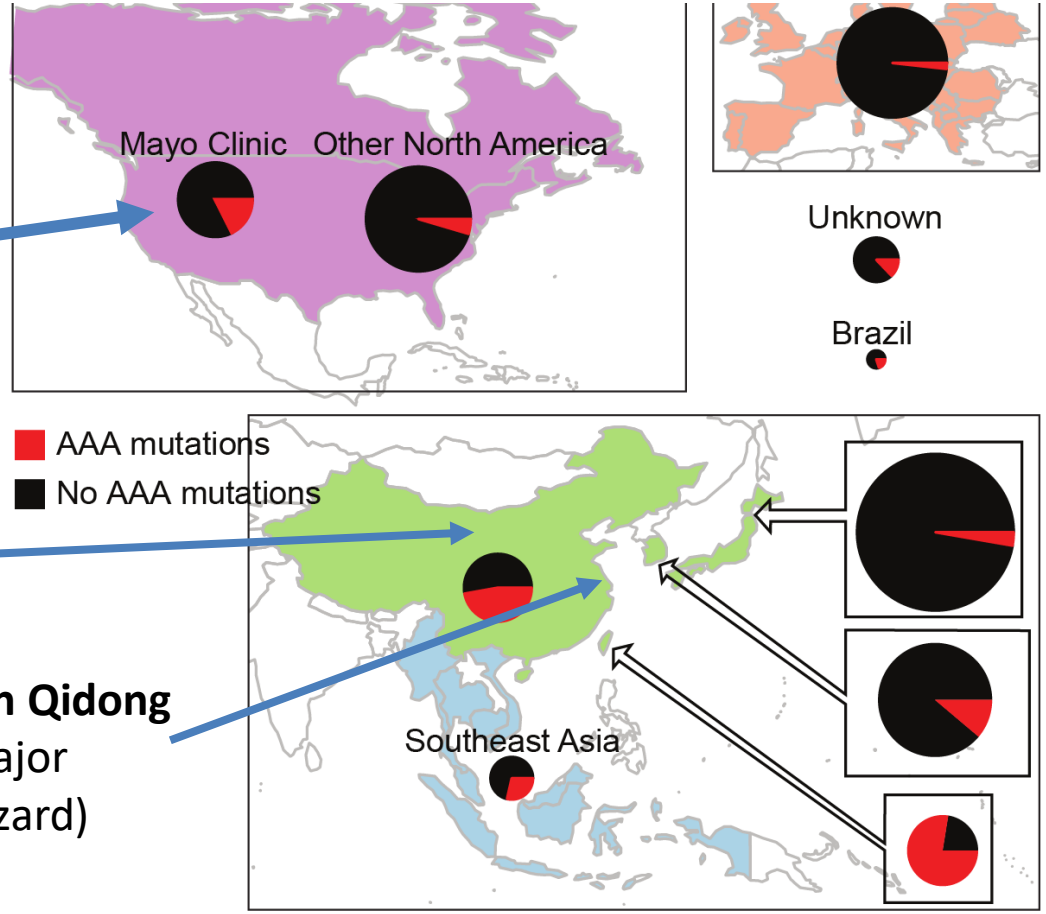
# How extensive is AA exposure in > 1,600 liver cancers around the world?

# Prevalent AA exposure across > 1,600 liver cancers

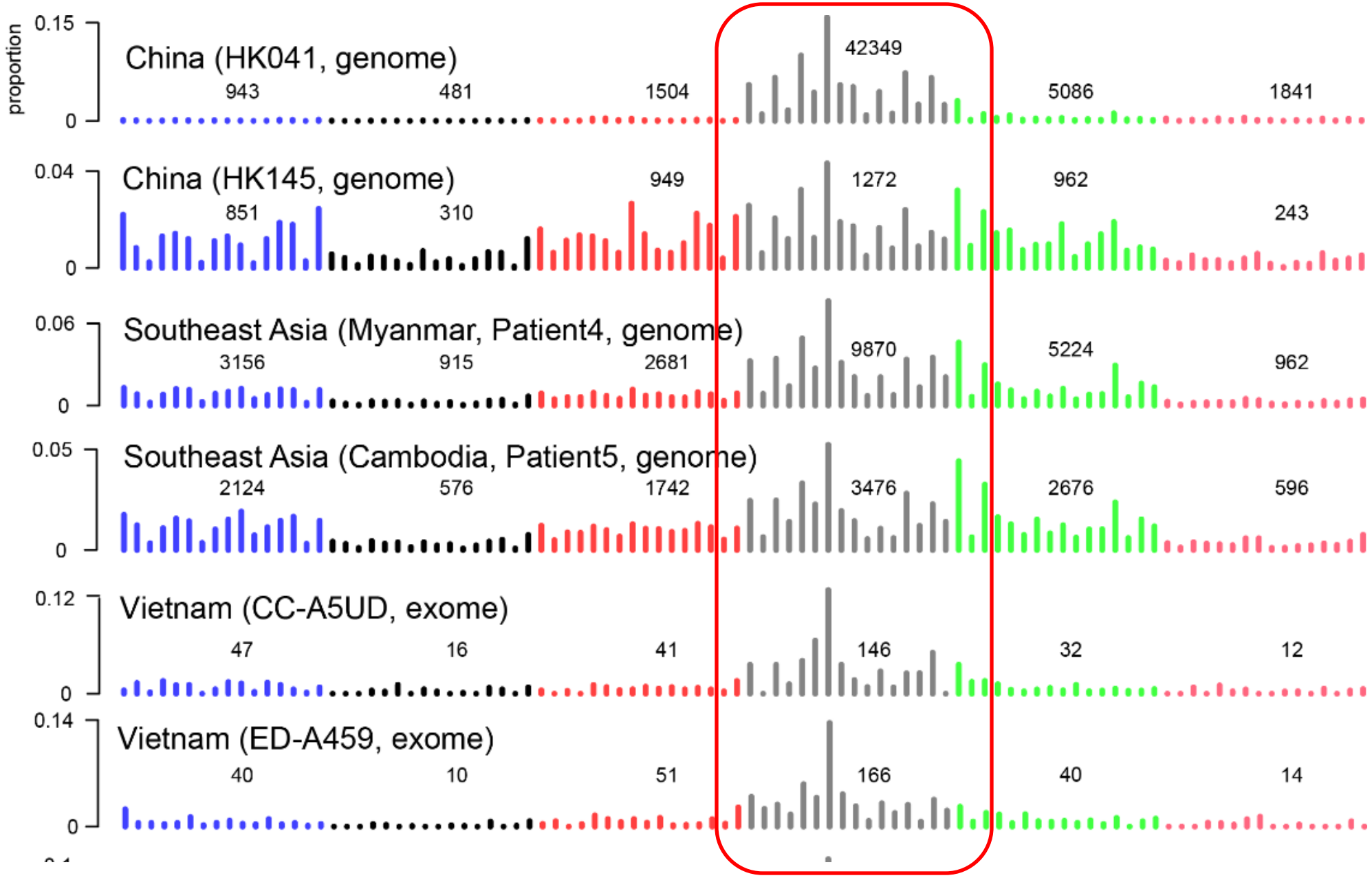
"Asian" patients with no "country" – Overseas patients?

**~50/100 (~50%) combined HCC and bile duct cancer in China**

**12/47 (25%) in Qidong** (aflatoxin a major competing hazard)




Ng, Poon, et al., Oct 2017 DOI: 10.1126/stranslmed.aan6446 and subsequent publications; Zhang et al., 2017, Gastroenterology, 10.1053/j.gastro.2017.03.024; Xue et al., 2019, Cancer Cell,10.1016/j.ccell.2019.04.007




DukeNUS Medical School

# AA can cause liver cancer

Evidence…

Known mutagen and urinary tract carcinogen

Multiple known cancer driver genes with AA mutations

Proportion of AA exposed liver cancers in Taiwan (78%) much greater than proportion of population exposed to AA herbs, (33%)

AA adducts liver tissue of HCC patients

Linear relationship between AA dose and risk of liver cancer in hepatitis-B-infected and hepatitis-C-infected patients

AA causes liver cancer in mice

"Asian" patien
no "country" –
Overseas patients?

**~50/100 (~50%) combined HCC and bile duct cancer in China**

**12/47 (25%) in Qidong** (aflatoxin a major competing hazard)

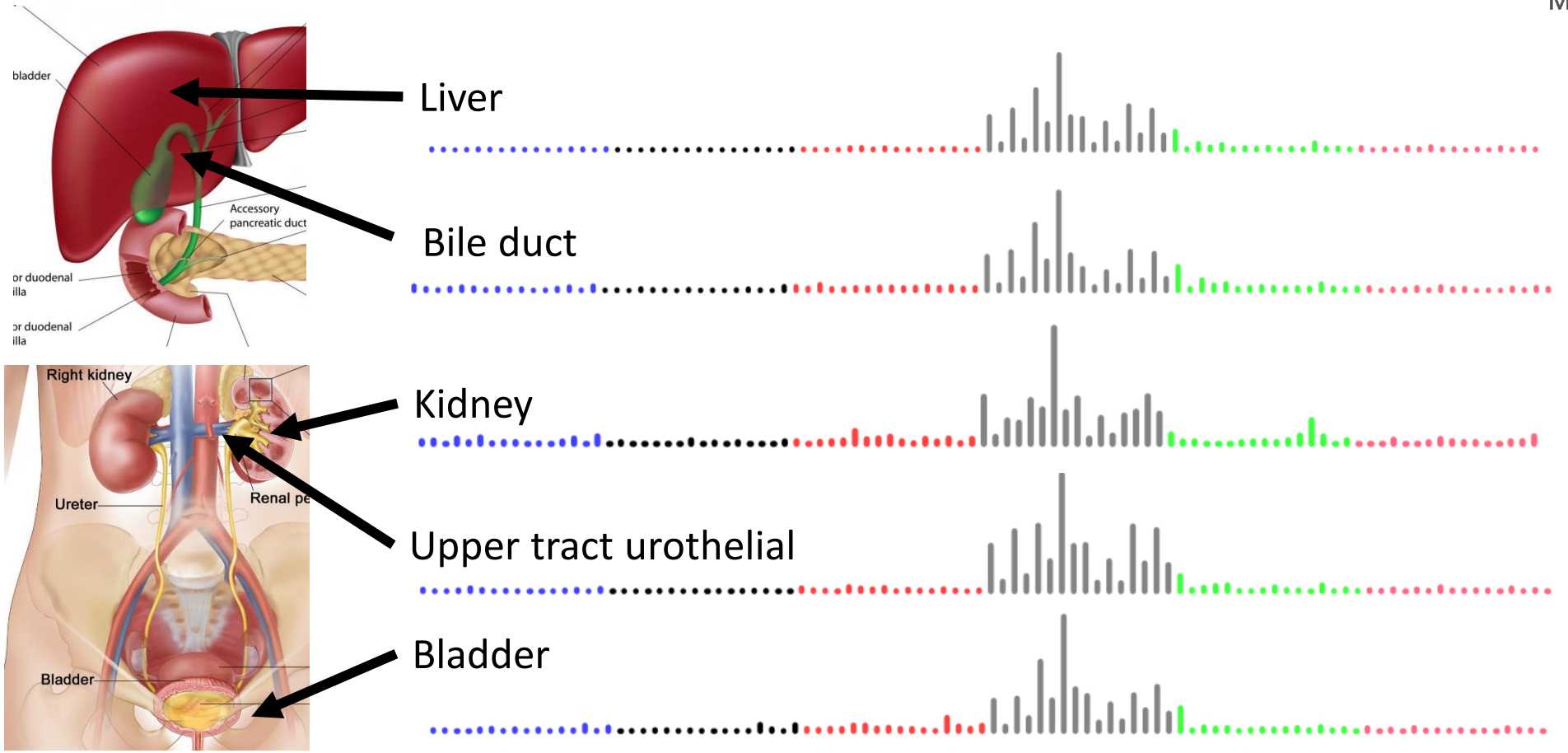
No AAA

Ng, Poon, et al., Oct 2017 DOI: 10.1126/stranslmed.aan6446 and subsequent publications; Zhang et al., 2017, Gastroenterology, 10.1053/j.gastro.2017.03.024; Xue et al., 2019, Cancer Cell,10.1016/j.ccell.2019.04.007，Hsieh et al., 2008, 10.1186/1749-8546-3-13) Chen et al., 2018 (*International Journal of Cancer,* 10.1002/ijc.31544) Chen et al., 2019 (*Cancer Epidemiology, Biomarkers & Prevention*, 10.1158/1055-9965.EPI-19-0023），Lu et al. 2019 Hepatology, 2019, 10.1002/hep.30863 (mouse study)

# Outline

- What are mutational signatures and what are they good for?
- Computational analysis of mutational signatures – state of the art
- Important unsolved problems in mutational signature analysis
- Summary

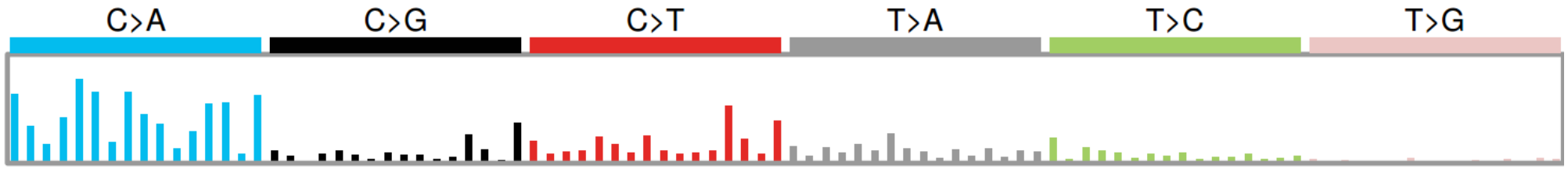
# Sometimes a spectrum is dominated by 1 signature
# (Herbal medicine aristolochic acid mutations in multiple cancer types)



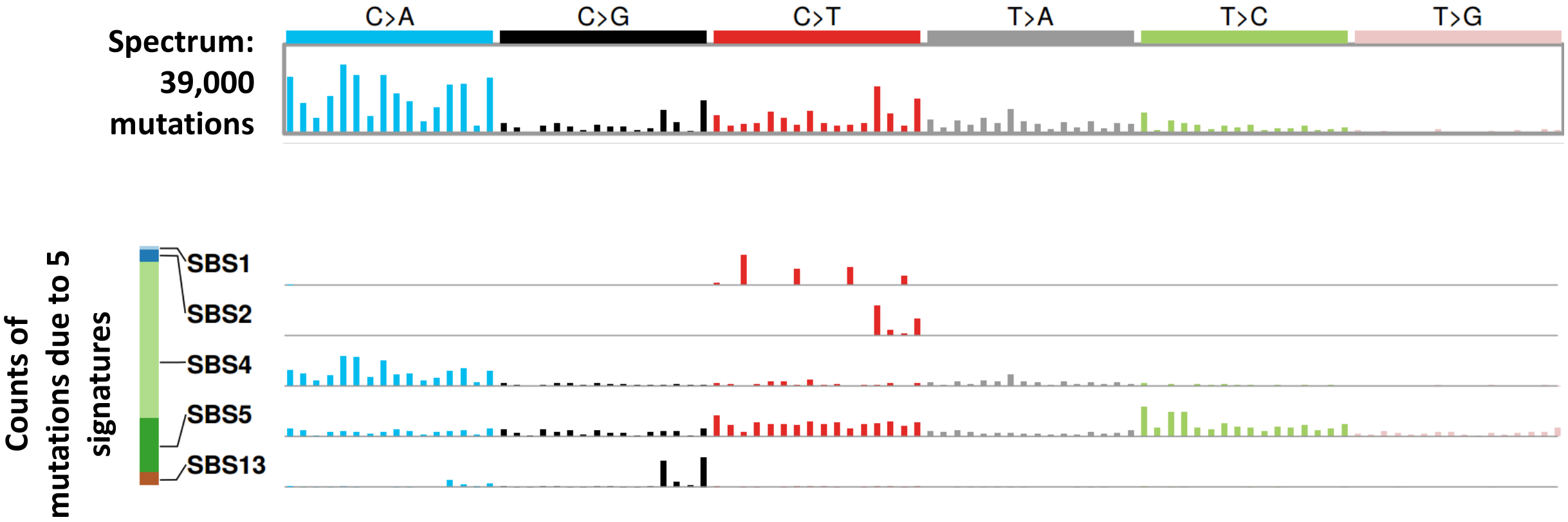

Poon et al., 2013 and subsequent data (liver - hepatocellular carcinoma [HCC])
Zou et al., 2015, Jusakul et al., 2017 (bile duct carcinoma)
Scelo et al., 2014, Jelakovic et al., 2014 (kidney/renal cell carcinoma)
Poon et al., 2013, Hoang et al., 2013, many others (upper tract urothelial carcinoma)
Poon et al., 2015, others (bladder urothelial carcinoma)

# More often, spectra consist of superimposed mutations from multiple processes

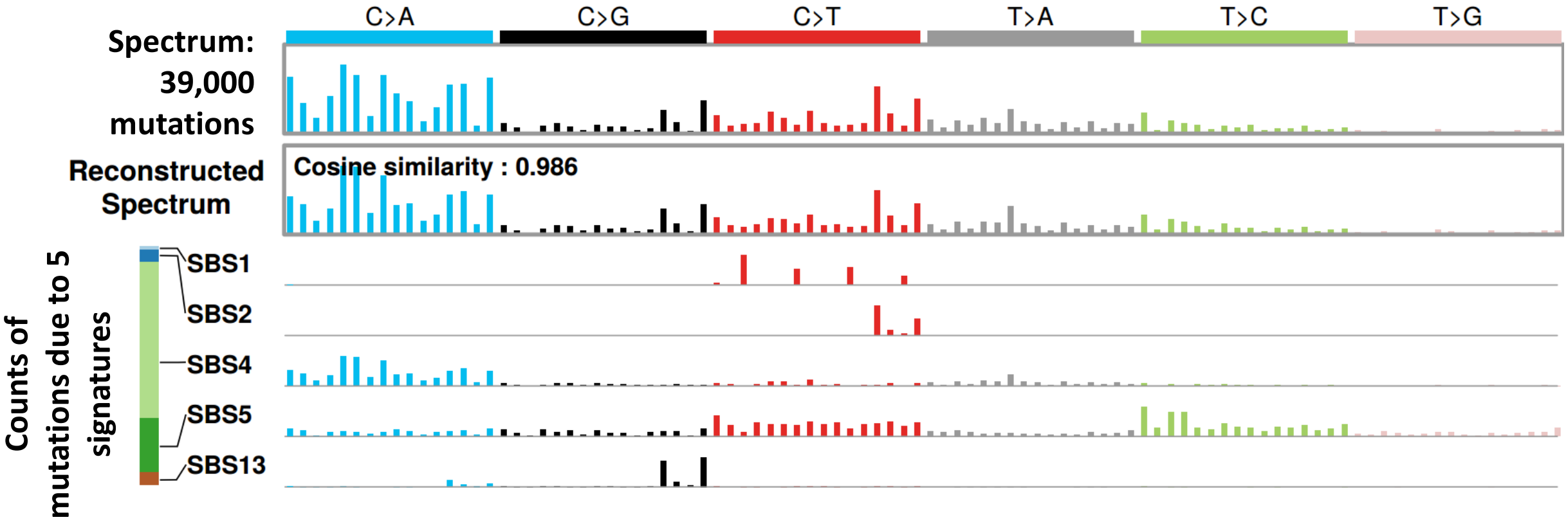
**Spectrum: 39,000 mutations**



# More often, spectra consist of superimposed mutations from multiple processes



# More often, spectra consist of superimposed mutations from multiple processes



# What questions do we want to answer?

- **"Extraction" / Discovery** Given a large number of spectra, what mutational signatures are present? (I.e. mutational signatures are latent variables to be discovered.) And how many mutations are caused by each extracted signature in each spectrum? (**"Attribution"**)
- **"Attribution only"** Given **known** mutational signatures and one or more spectra, how many mutations in each spectrum were caused by each signature?
- **"Signature presence test"** Given known mutational signatures, what is the evidence that a signature of interest is present in a spectrum?

# What questions do we want to answer?

- **"Extraction" / Discovery** Given a large number of spectra, what mutational signatures are present? (I.e. mutational signatures are latent variables to be discovered.) And how many mutations are caused by each extracted signature in each spectrum? (**"Attribution"**)
- **"Attribution only"** Given **known** mutational signatures and one or more spectra, how many mutations in each spectrum were caused by each signature?
- **"Signature presence test"** Given known mutational signatures, what is the evidence that a signature of interest is present in a spectrum?

# Many signatures have been extracted and many have been confirmed by additional evidence


...

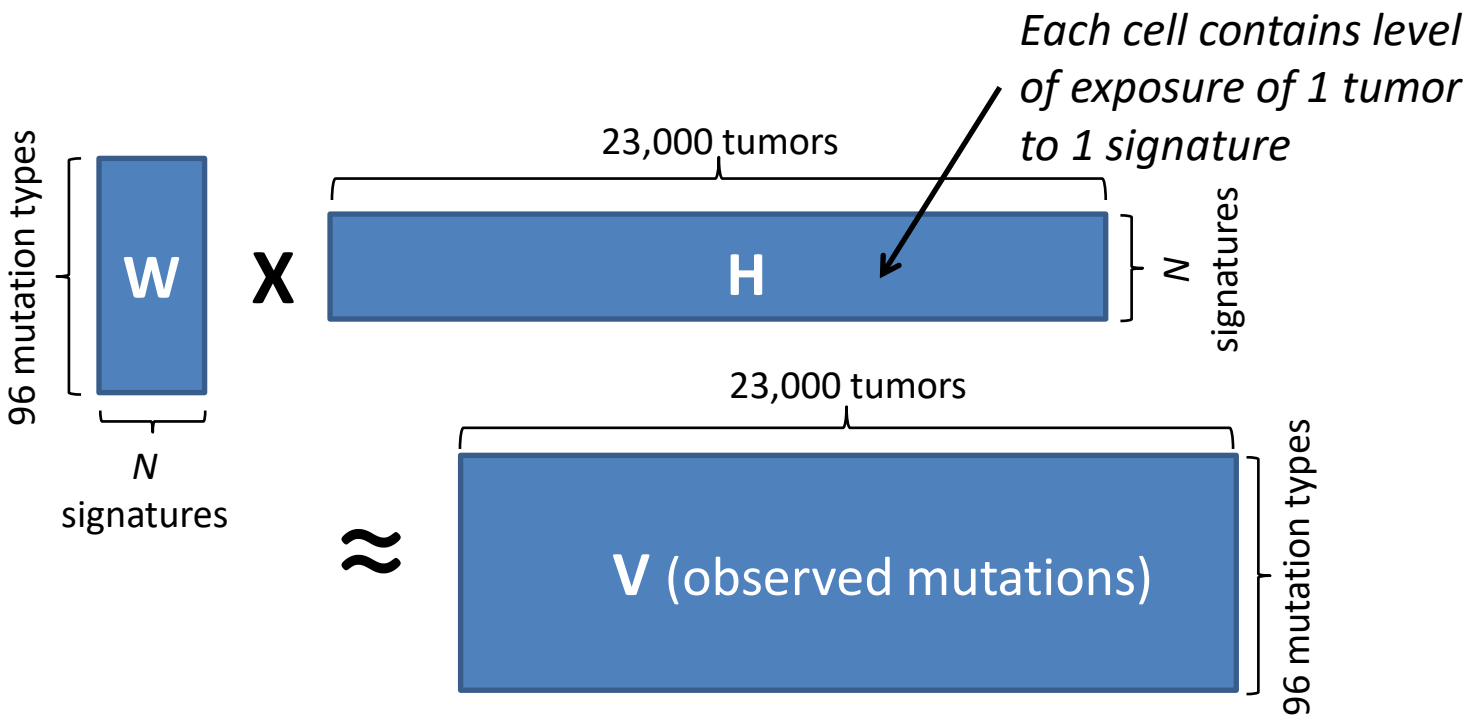
and many more

https://cancer.sanger.ac.uk/cosmic/signatures/index.tt from Alexandrov, L.B., Kim, J., Haradhvala, N.J. *et al.* The repertoire of mutational signatures in human cancer. *Nature* **578,** 94–101 (2020) https://doi.org/10.1038/s41586-020-1943-3

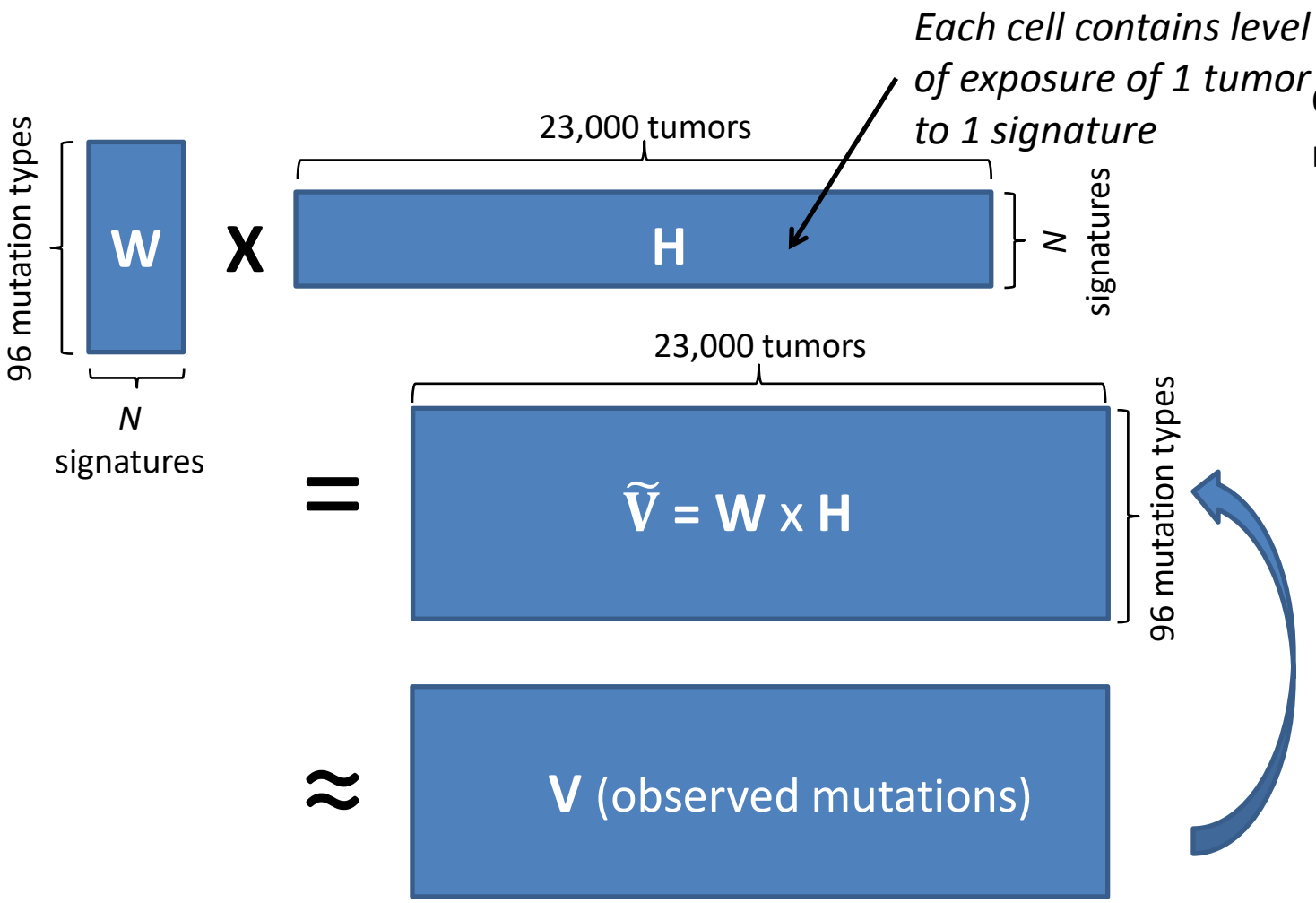
# Data mining can tease apart signatures in large sets of spectra

- There are approaches based on:
  - Non-negative-factorization (NMF)
  - Probabilistic topic models
    - (will not discuss, see Nicola Roberts' thesis https://www.repository.cam.ac.uk/handle/1810/275454 and https://github.com/steverozen/mSigHdp )
- ***Unsupervised machine learning***
- ***All methods seem to face similar challenges***
- ***Signature discovery is not a purely algorithmic process***
- The granularity of extracted signatures and the type of mutations considered depend on the larger questions you are considering

# Many techniques for extraction are based on non-negative matrix factorization (NMF)



# Many techniques for mutational signature extraction are based on non-negative matrix factorization (NMF)




Objective fn is often Frobenius norm of $D = \widetilde{V} - V$:

$$\sqrt{\sum_{m \in (\text{mutation types})} \sum_{t \in \text{tumors}} D_{m,t}^2}$$

Objective function:
How "close" is $\widetilde{V}$ to V?

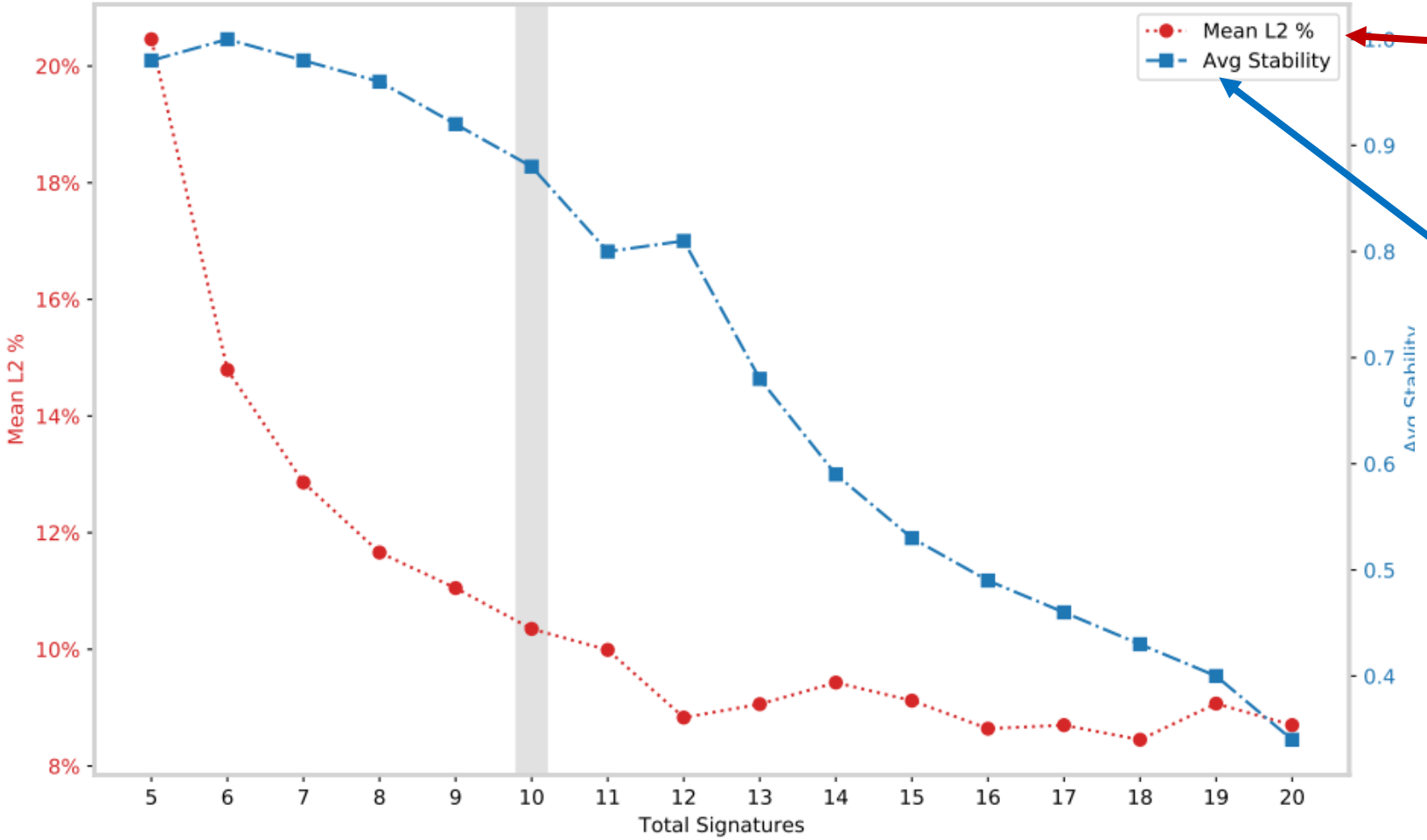
# General observations on NMF

- NMF is a collection of algorithms and techniques
- The best approximate factorization depends on the objective function
- As for many unsupervised learning approaches, determining the number of items (signatures) to discover is challenging
- Need to avoid overfitting (so often do multiple factorizations on e.g. bootstrapped data)

# NMF example: SigProfilerExtractor

- Supersedes the MATLAB code used in Alexandrov 2013 and 2020
- Latest release: https://pypi.org/project/sigproextractor/
  - https://github.com/AlexandrovLab
  - Documentation at https://osf.io/t6j7u/wiki/3.%20Using%20the%20Tool%20-%20Output/
- More info in Alexandrov et al 2013 and 2020:
  - http://dx.doi.org/10.1016/j.celrep.2012.12.008
  - https://doi.org/10.1038/s41586-020-1943-3

# SigProfilerExtractor: Selecting the number of signatures




Mean L2% is the average across samples of the percent Euclidian reconstruction error (low is good)

Average stability is the mean across NMF iterations of silhouette coefficient of the clusters of proto-signatures (indicates how well the proto-signatures are clustered into signatures – high is good)

# Known issues with approaches based on non-negative matrix factorization (NMF)

- The number of signatures is a judgement call (some approaches deal with this)
- Weak signatures cannot be discovered in background of strong signatures
- Signatures can be imperfectly separated ("bleeding" and "stealing" between signatures)
- More generally, recovered signatures depend strongly on exact set of tumours studied
- Lowest error reconstruction often involves many signatures, some with very low contributions, and often not biologically plausible (see following slides)
- ***Signature discovery (extraction) is not a purely algorithmic process***

In practice, discovering signatures ("extraction") needs to be separated from estimating exposures ("attribution")

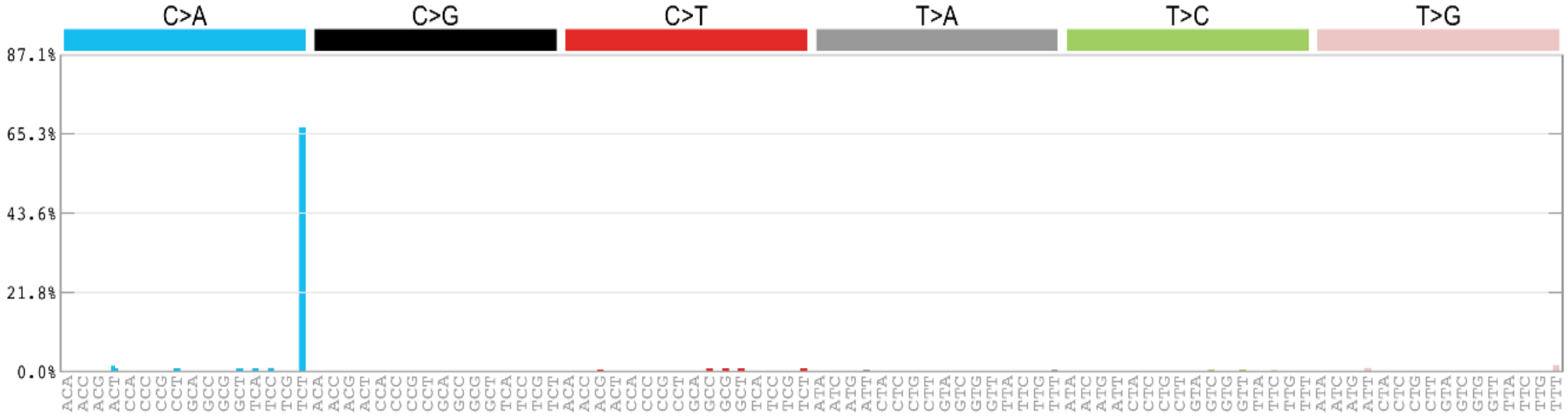
!

# Sparsity and plausibility

- ~~Signatures that, when combined, give the most accurate reconstruction of a spectrum are the best ones~~

- Adding more signatures in tiny amounts usually improves reconstruction even if activity of signatures is biologically implausible
- Not useful for deciding if a mutational signature is present to a biologically meaningful extent

Extreme example:
A set of 96 signatures, each of one single base substitution in trinucleotide context will yield perfect reconstruction and zero biological insight



# Signature discovery is not a purely algorithmic process: How to assess results (1)

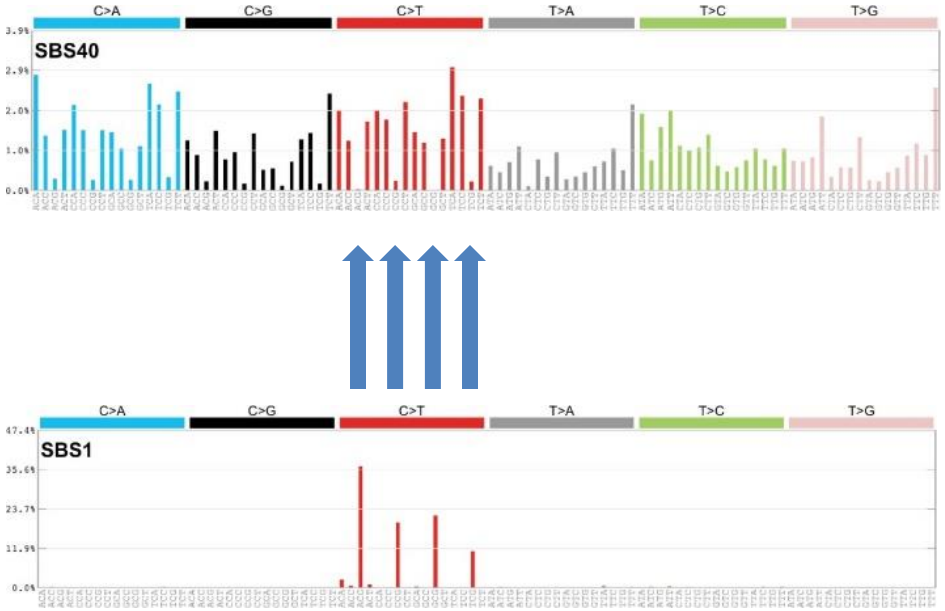
- The input data have been examined for sequencing and mapping artifacts (some which are well-known); mapped reads supporting unusual patterns have been examined in a read-alignment viewer or have been experimentally verified
- Look for external supporting evidence
  - Correlation with known or expected mutagenic exposures, eg
    - cigarette smoking
    - Aflatoxins
    - AA-containing herbs
    - Haloalkanes
    - Age
    - UV radiation
  - Correlation with known or expected genetic causes, eg
    - polymerase epsilon proofreading defects
    - mismatch repair deficiency
    - homologous recombination repair deficiency (BRCA)
  - Can be tied to known biochemical processes (eg guanine adducts and signatures with C:G > N:N mutations)

# Signature discovery is not a purely algorithmic process: How to assess results (2)

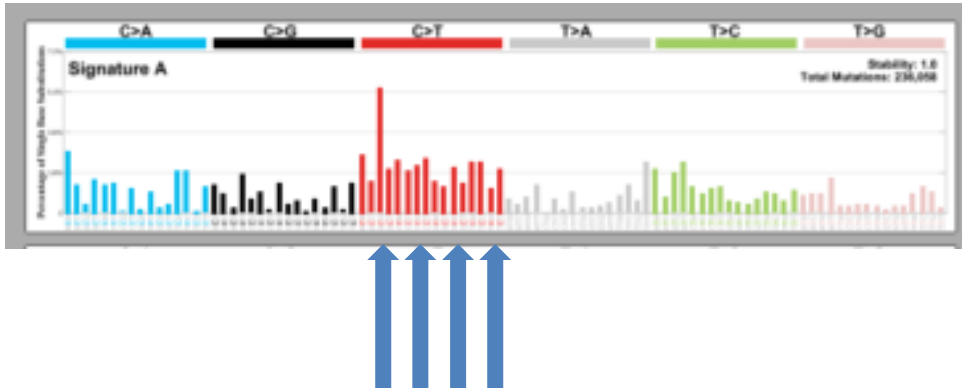
- Association with genomic features that interact with DNA repair
  - Transcription strand bias
  - Replication strand / timing
  - Homopolymers (indels)
- Look for supporting evidence within the data set
  - Samples dominated by a single signature
  - Signature is consistently deciphered from multiple independent datasets using different techniques or hyperparameters
  - Absence of known problems (next slides)

# Known problems: bleeding

Ground truth (synthetic data)

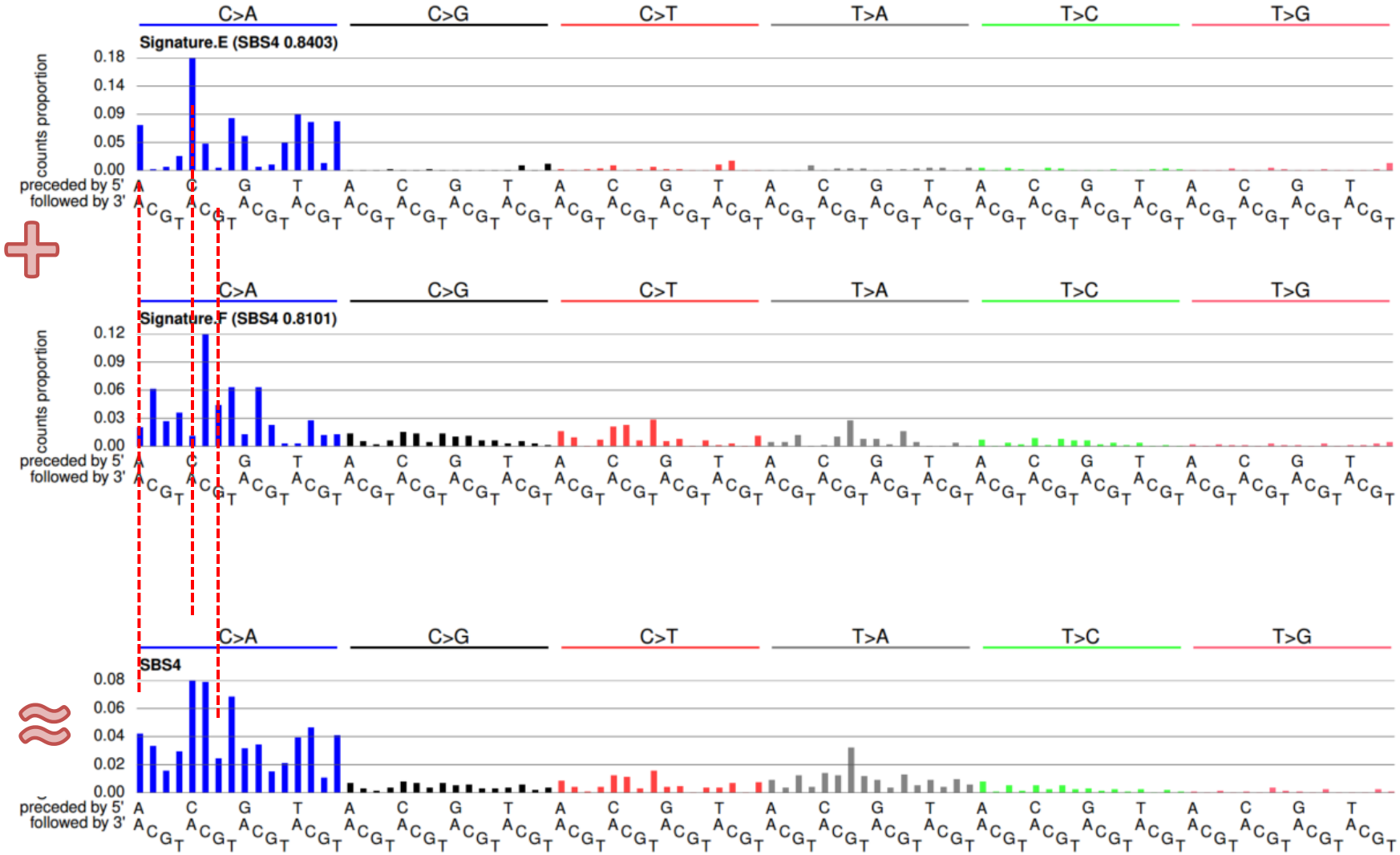



Extracted




"Bleeding" or incomplete separation of some CG > TG mutations from SBS1 to SBS40-like extracted signature

# Known problem: over-splitting (from a different synthetic data set)




2 extracted signatures

1 ground-truth signature

# Outline

- What are mutational signatures and what are they good for?
- Computational analysis of mutational signatures – state of the art
- Important unsolved problems in mutational signature analysis
- Summary

# What questions do we want to answer?

- **"Extraction" / Discovery** Given a large number of spectra, what mutational signatures are present? (I.e. mutational signatures are latent variables to be discovered.) And how many mutations are caused by each extracted signature in each spectrum? (**"Attribution"**)
- **"Attribution only"** Given **known** mutational signatures and one or more spectra, how many mutations in each spectrum were caused by each signature?
- **"Signature presence test"** Given known mutational signatures, what is the evidence that a signature of interest is present in a spectrum?

# What questions do we want to answer?

- **"Extraction" / Discovery** Given a large number of spectra, what mutational signatures are present? (I.e. mutational signatures are latent variables to be discovered.) And how many mutations are caused by each extracted signature in each spectrum? (**"Attribution"**)
- **"Attribution only"** Given **known** mutational signatures and one or more spectra, how many mutations in each spectrum were caused by each signature?
- **"Signature presence test"** Given known mutational signatures, what is the evidence that a signature of interest is present in a spectrum?

# General issues with attribution

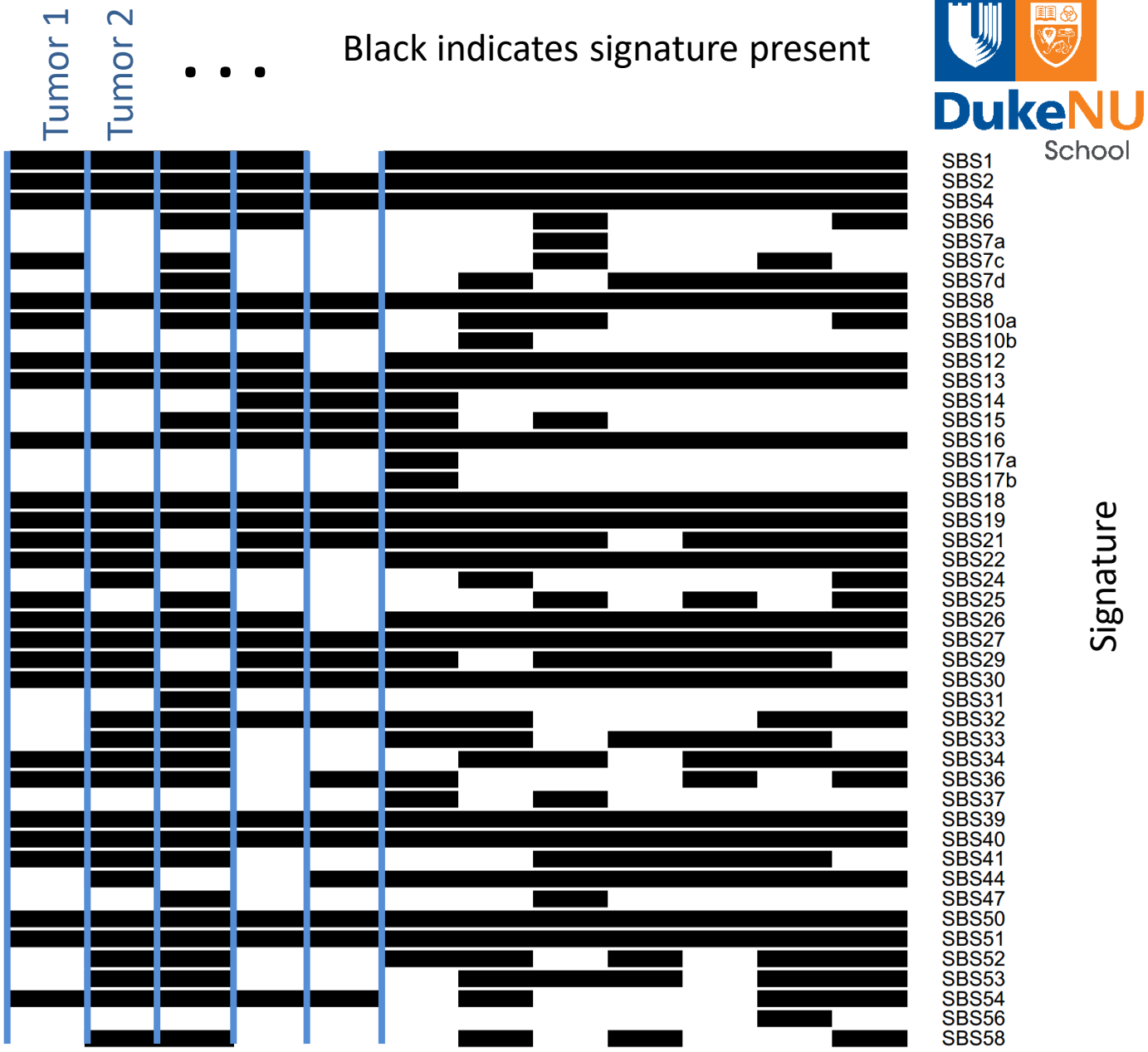
- Not studied very thoroughly
- With enough signatures, adding more signatures almost always improves reconstructions
- Issues of sparsity (tiny exposures to many signatures)
- Issues of biological plausibility
- Common examples:
  - ultraviolet signatures in tumors with no possibility of UV exposure
  - Single base substitution signatures of microsatellite instability in tumors that clearly do not have microsatellite instability
  - Signatures of defective polymerase epsilon proofreading (which generates very high numbers of mutations) with very low mutation counts (and no defect in the polymerase epsilon gene).

# Example: several lung cancers (squamous)

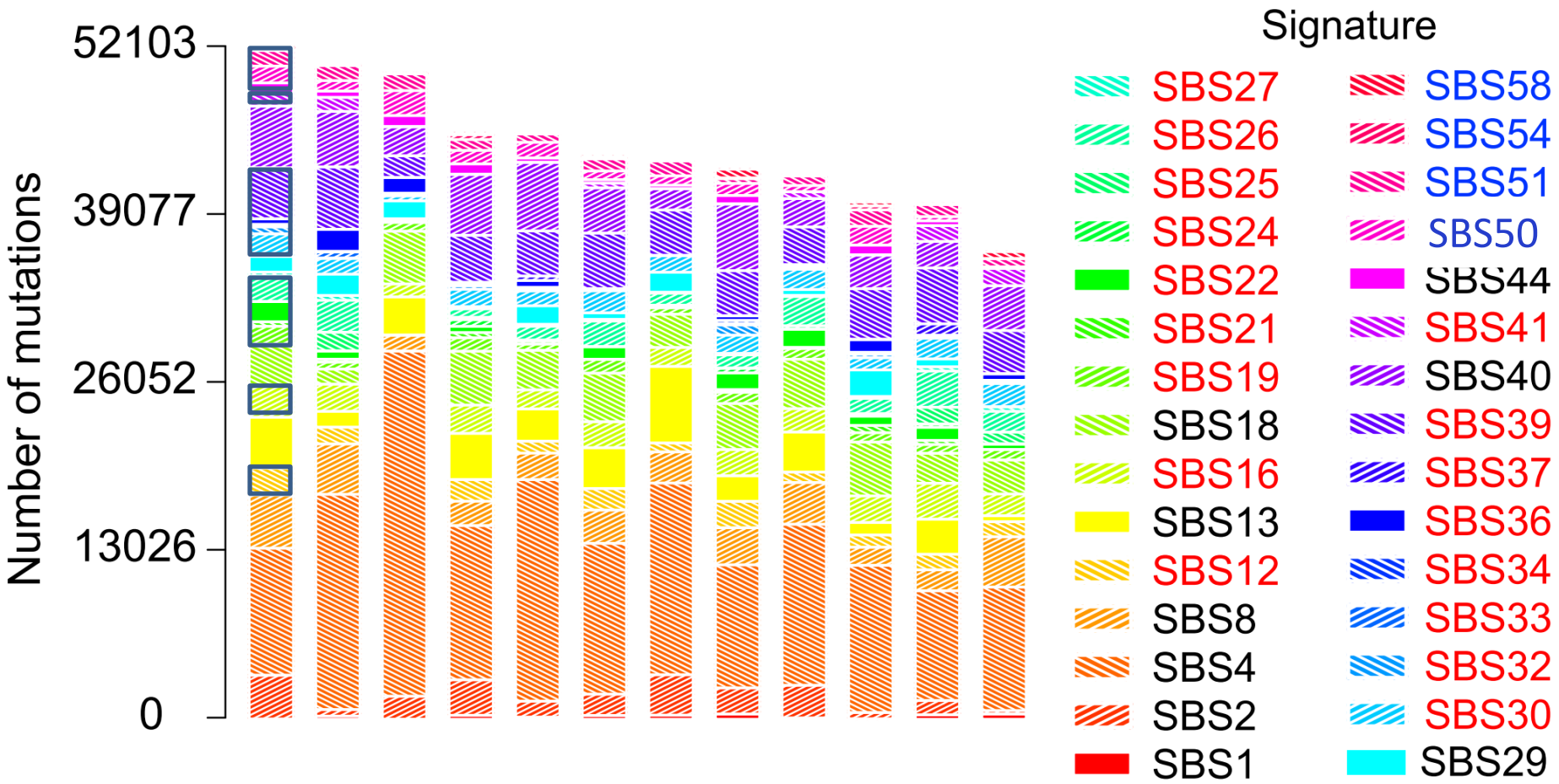
We simply optimized coefficients of all known signatures to minimize reconstruction error (45 signatures assigned)

Good reconstructions but bad models of reality

(Signatures from Alexandrov, L.B., Kim, J., Haradhvala, N.J. *et al.* The repertoire of mutational signatures in human cancer. *Nature* **578,** 94–101 (2020). https://doi.org/10.1038/s41586-020-1943-3)

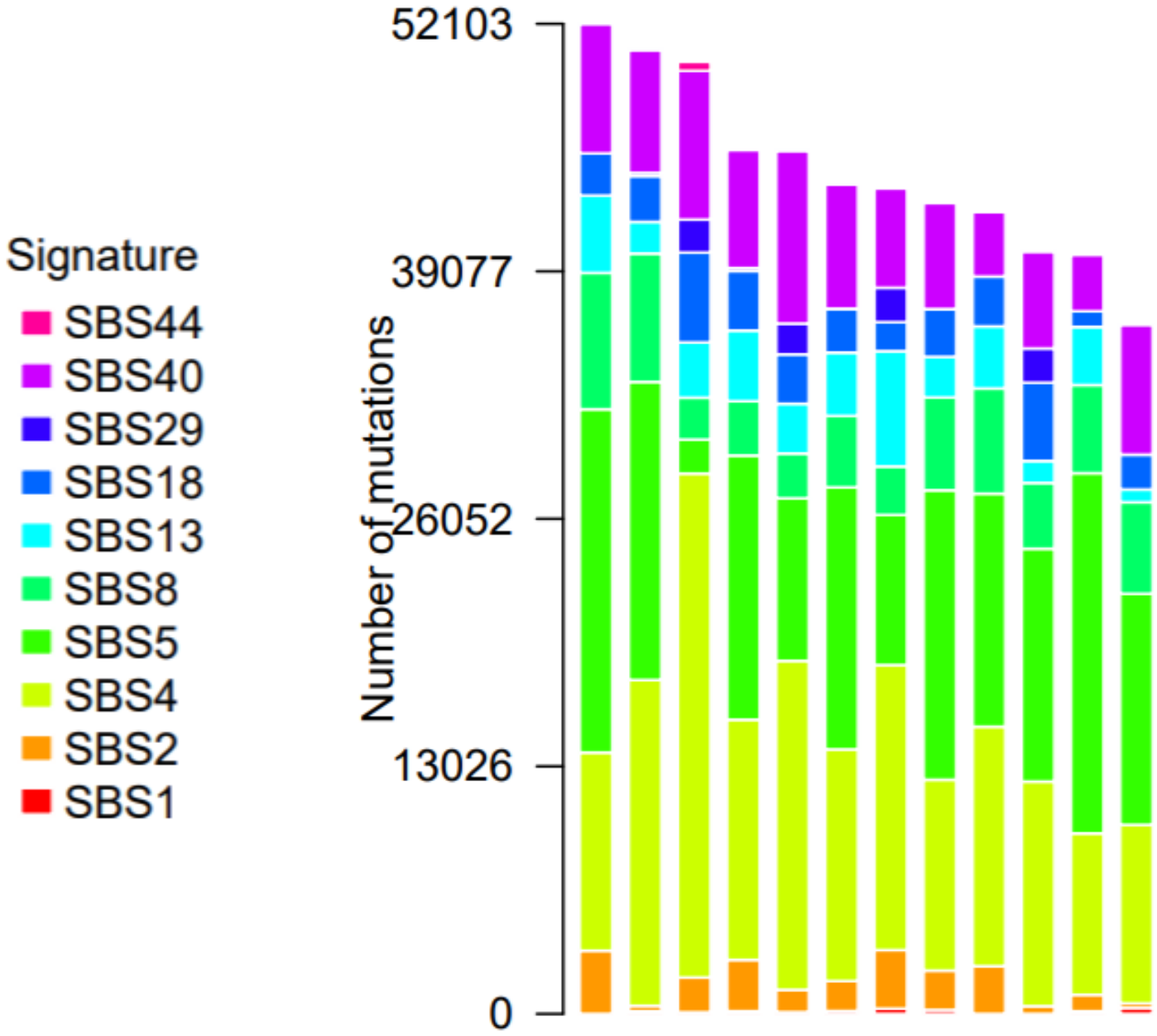


Impose some sparsity by requiring minimum number of mutations (30 signatures assigned)
Still a good reconstruction but  a bad model of reality



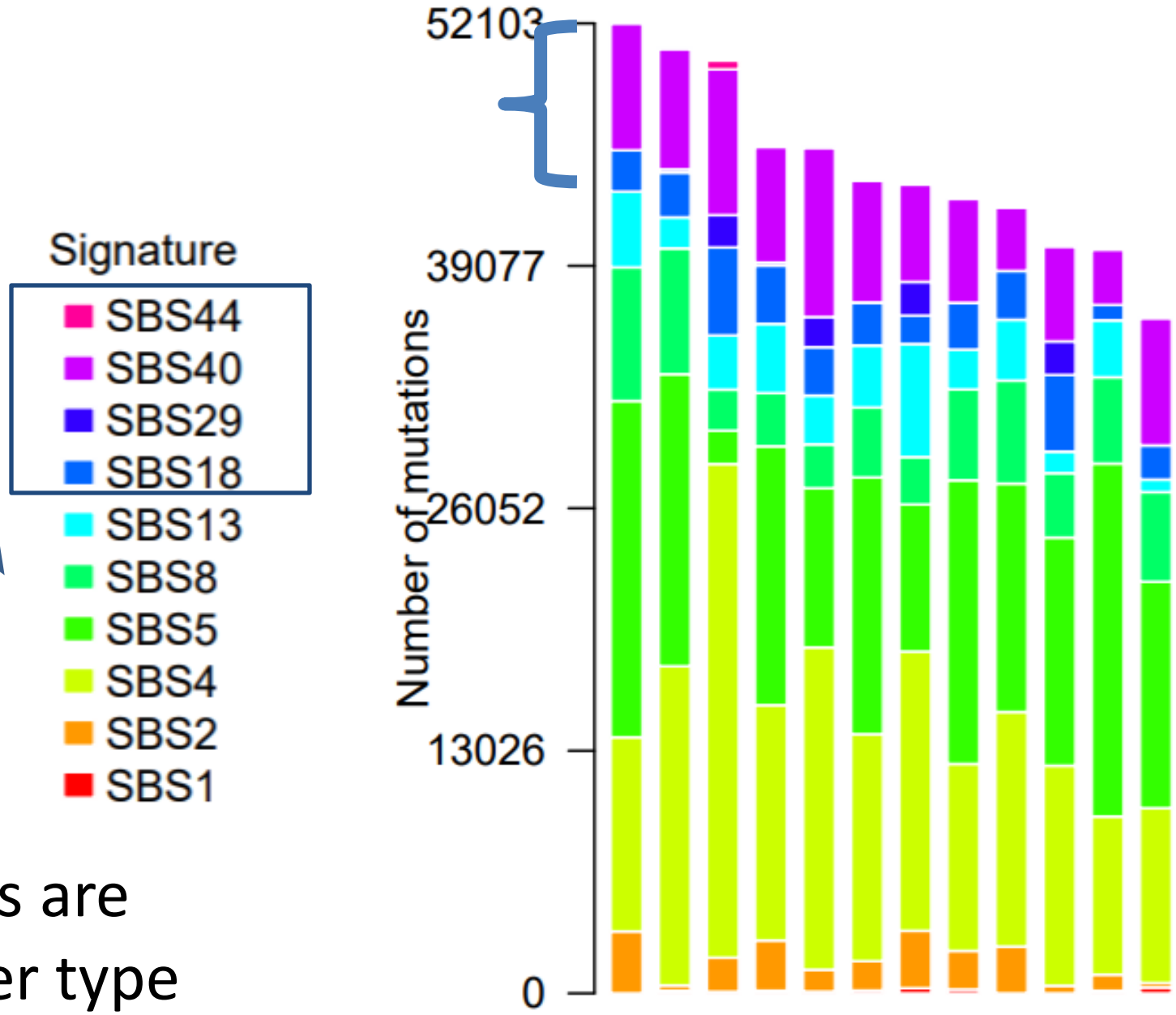
Considering only mutations found in lung cancer (squamous)

Much better



Considering only mutations found in lung cancer (squamous)

Much better




But these signatures are very rare in this cancer type

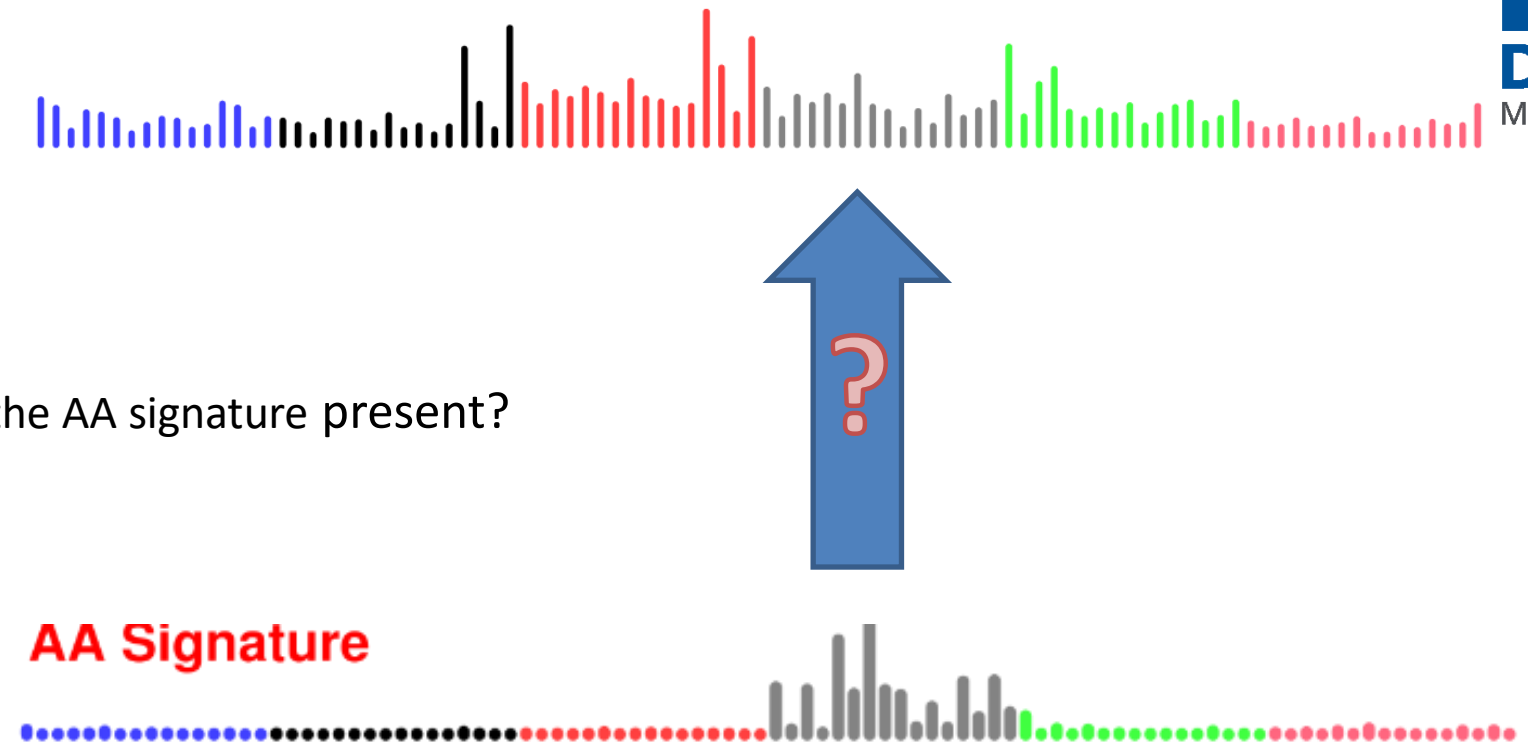
# What questions do we want to answer?

- **"Extraction" / Discovery** Given a large number of spectra, what mutational signatures are present? (I.e. mutational signatures are latent variables to be discovered.) And how many mutations are caused by each extracted signature in each spectrum? (**"Attribution"**)
- **"Attribution only"** Given **known** mutational signatures and one or more spectra, how many mutations in each spectrum were caused by each signature?

Detour

**"Signature presence test"** Given known mutational signatures, what is the evidence that a signature of interest is present in a spectrum?

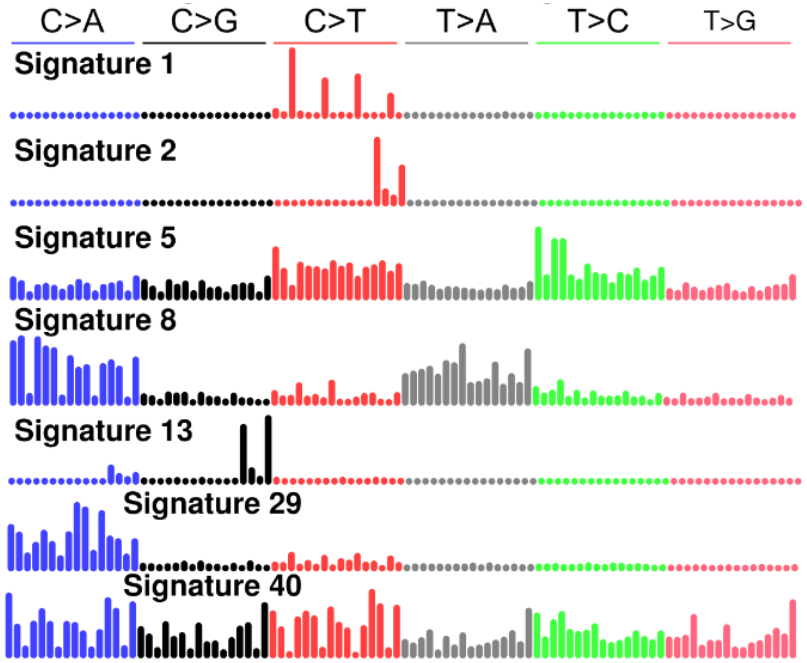
# Signature presence test




Example, motivated by our study of AA exposure in liver cancer
Was this tumour exposed to AA?

# Signature presence test

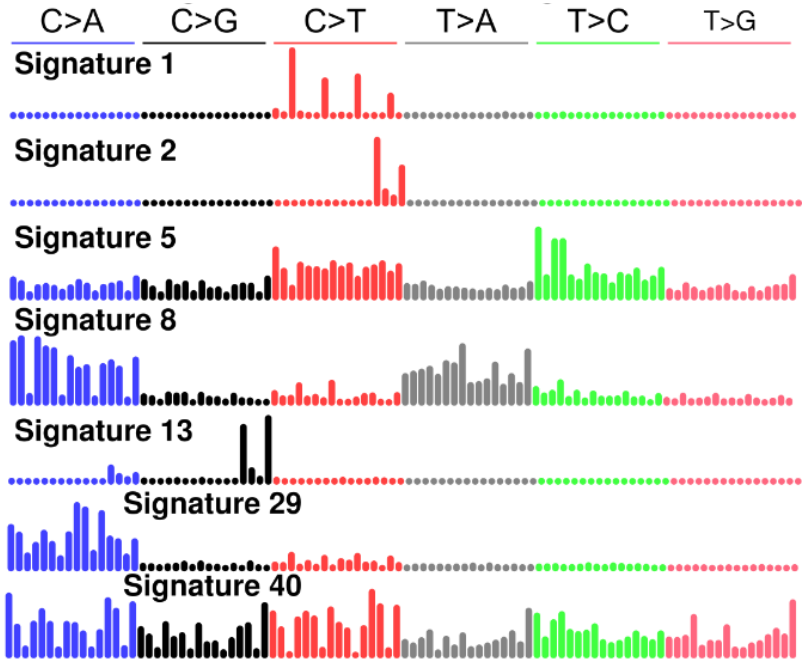
Best reconstruction without AA




Log likelihood of reconstruction -1040

# Signature presence test

Best reconstruction without AA

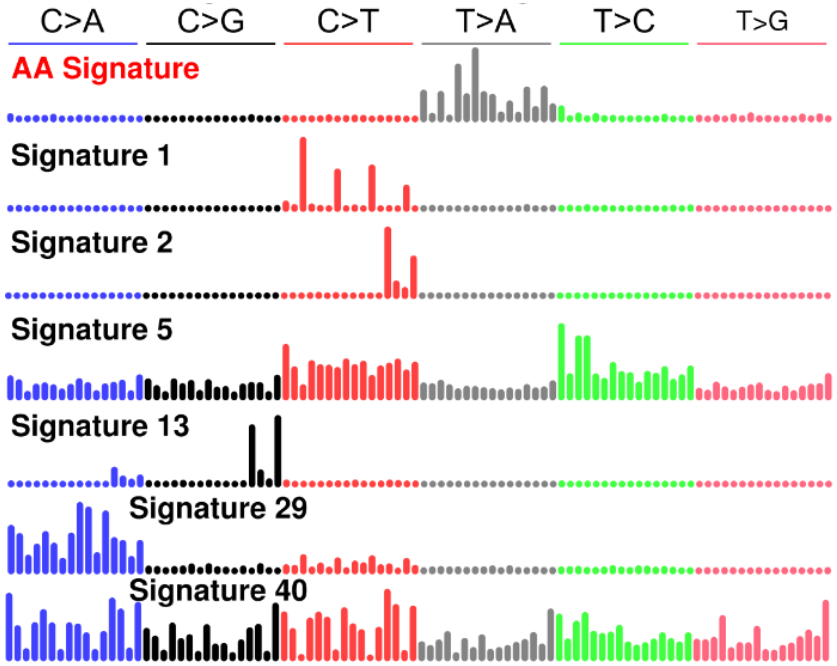



The likelihood of the reconstruction is the probability that negative binomial resampling from the reconstruction yields the spectrum

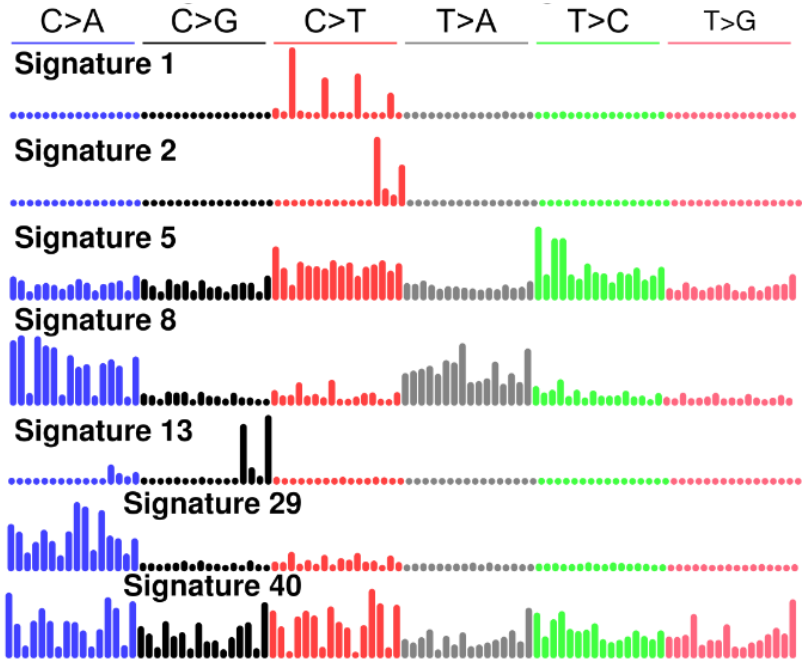
Log likelihood of reconstruction -1040

# Signature presence test

Best reconstruction ***WITH*** AA




Best reconstruction without AA




Log likelihood of reconstruction -1021

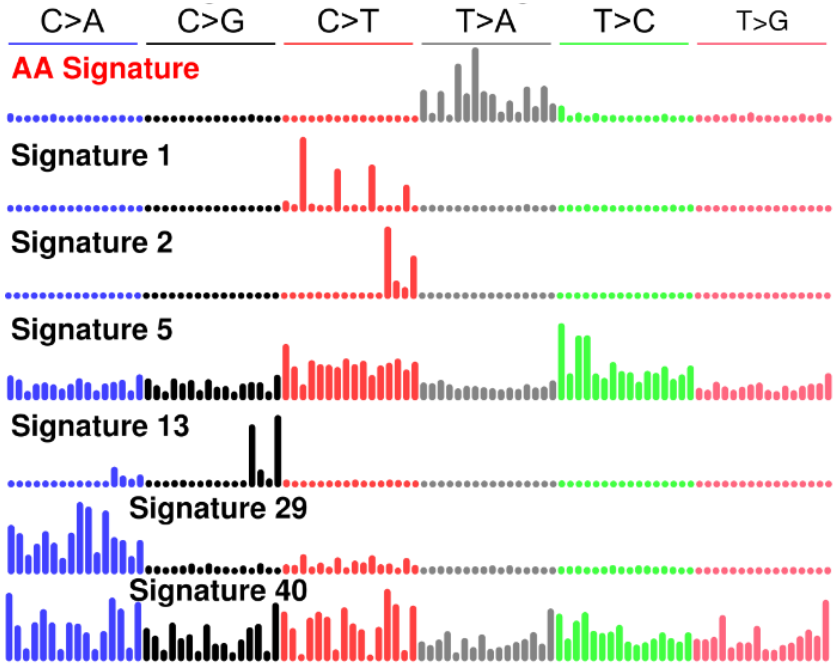
Log likelihood of reconstruction -1040

# Signature presence test

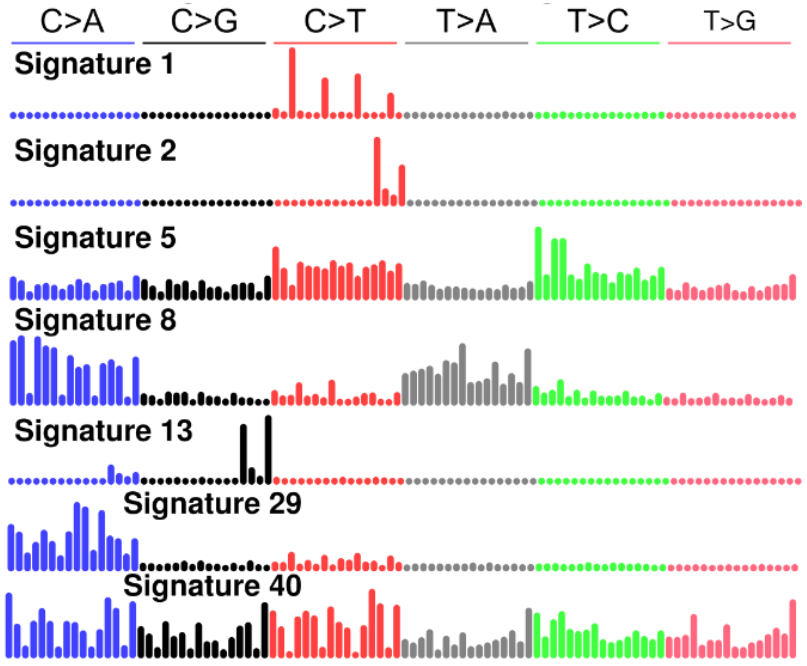
Likelihood ratio test $p < 10^{-9}$

Note that we have added a statistical concept

Best reconstruction ***WITH*** AA




Best reconstruction without AA




Log likelihood of reconstruction -1021

Log likelihood of reconstruction -1040

We reject the null hypothesis that the reconstruction of the spectrum without the AA signature is as good as the reconstruction with the AA signature

That is, we need the AA signature to plausibly explain the spectrum

Can extend the signature presence test to sets of signatures

"Is this ***set*** of signatures needed to plausibly account for this spectrum?"

This means we can find all subsets of signatures that **can** plausibly account for a spectrum (usually more than one such subset)

We can then search among subsets of signatures that can plausibly explain the spectrum to find the “best” subset

# Example of multiple plausible reconstructions

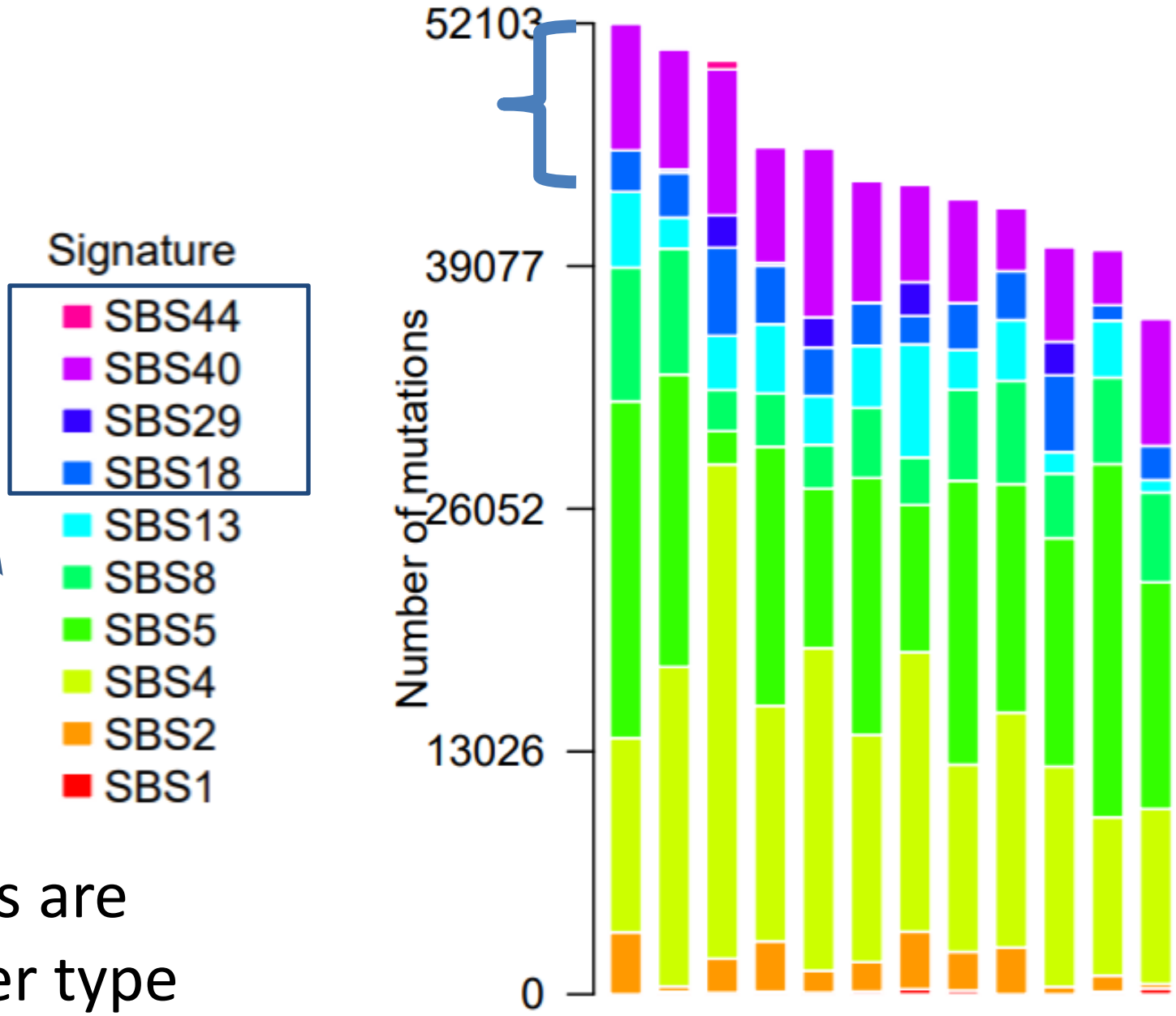
- One squamous lung cancer
- Possible signatures SBS1, SBS2, SBS4, SBS5, SBS8, SBS13, SBS18, SBS29, SBS40, SBS44
- Can remove ~60 subsets of signatures and still get plausible reconstruction
- For example, can remove
  - SBS29
  - SBS18, SBS29
  - SBS18, SBS29, SBS40, SBS44,
  - SBS1, SBS18, SBS29, SBS40, SBS44

# What questions do we want to answer?

- **"Extraction" / Discovery** Given a large number of spectra, what mutational signatures are present? (I.e. mutational signatures are latent variables to be discovered.) And how many mutations are caused by each extracted signature in each spectrum? (**"Attribution"**)
- **"Attribution only"** Given **known** mutational signatures and one or more spectra, how many mutations in each spectrum were caused by each signature?
- **"Signature presence test"** Given known mutational signatures, what is the evidence that a signature of interest is present in a spectrum?

Considering only mutations found in lung cancer (squamous)

Much better



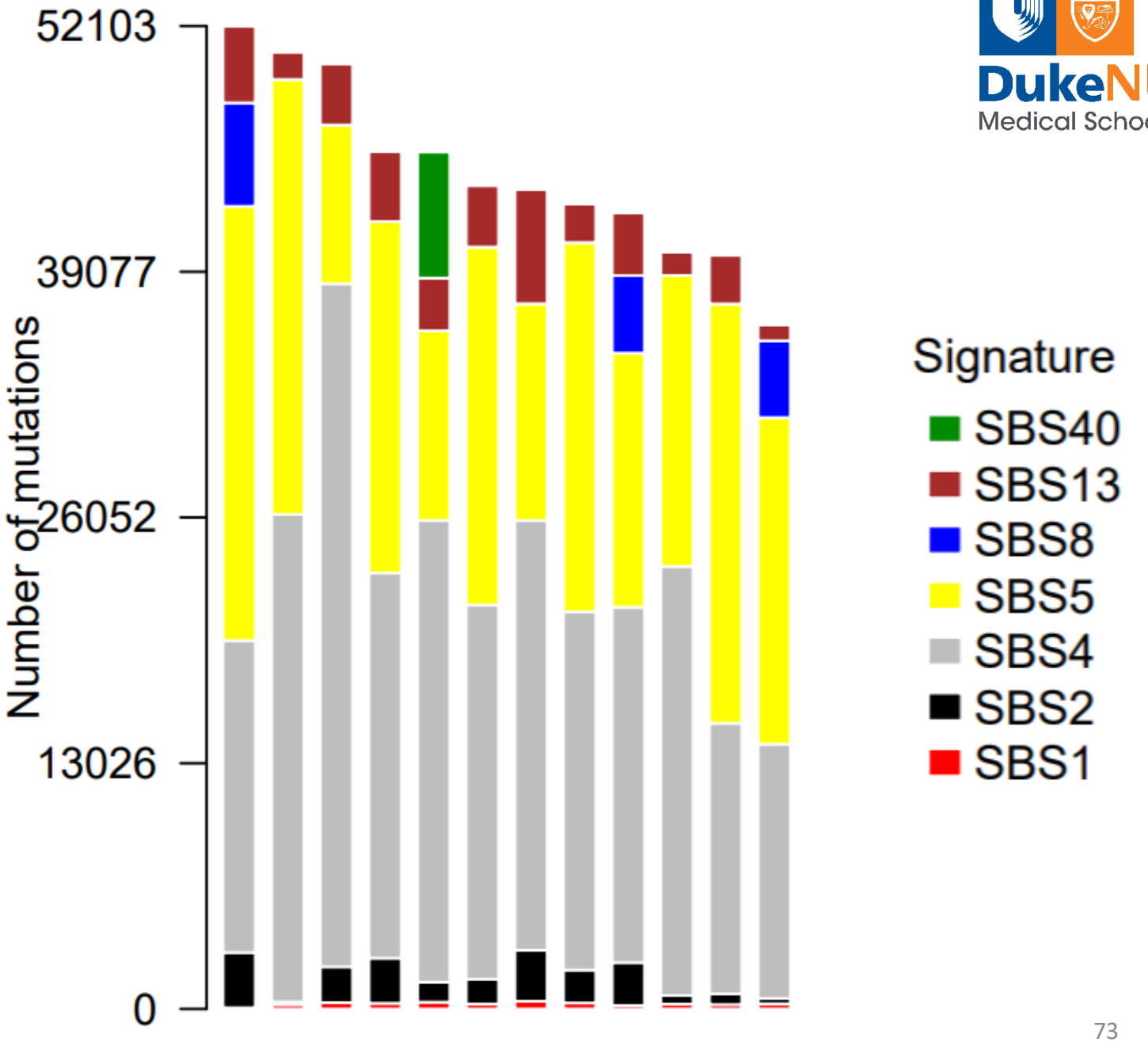

But these signatures are very rare in this cancer type

# A notion of the "best" subset among the subsets that can plausibly explain the spectrum

- We "Maximum A Posteriori Probability" to find a best attribution (subset of signatures)
- We maximize $P(D|M)P(M)$, where
- $P(D|M)$ is the likelihood of the attribution – the probability of the spectrum given the attribution
- For $P(M)$ we use the product of the probabilities that each signature is present or absent in the given cancer type

# Very rare signatures in squamous cell lung cancer (SBS18, SBS29, SBS44) no longer attributed

SBS40 (found in 5% of squamous lung cancer) probably needs more investigation



# Outline

- What are mutational signatures and what are they good for?
- Computational analysis of mutational signatures – state of the art
- Important unsolved problems in mutational signature analysis
- Summary

# Summary

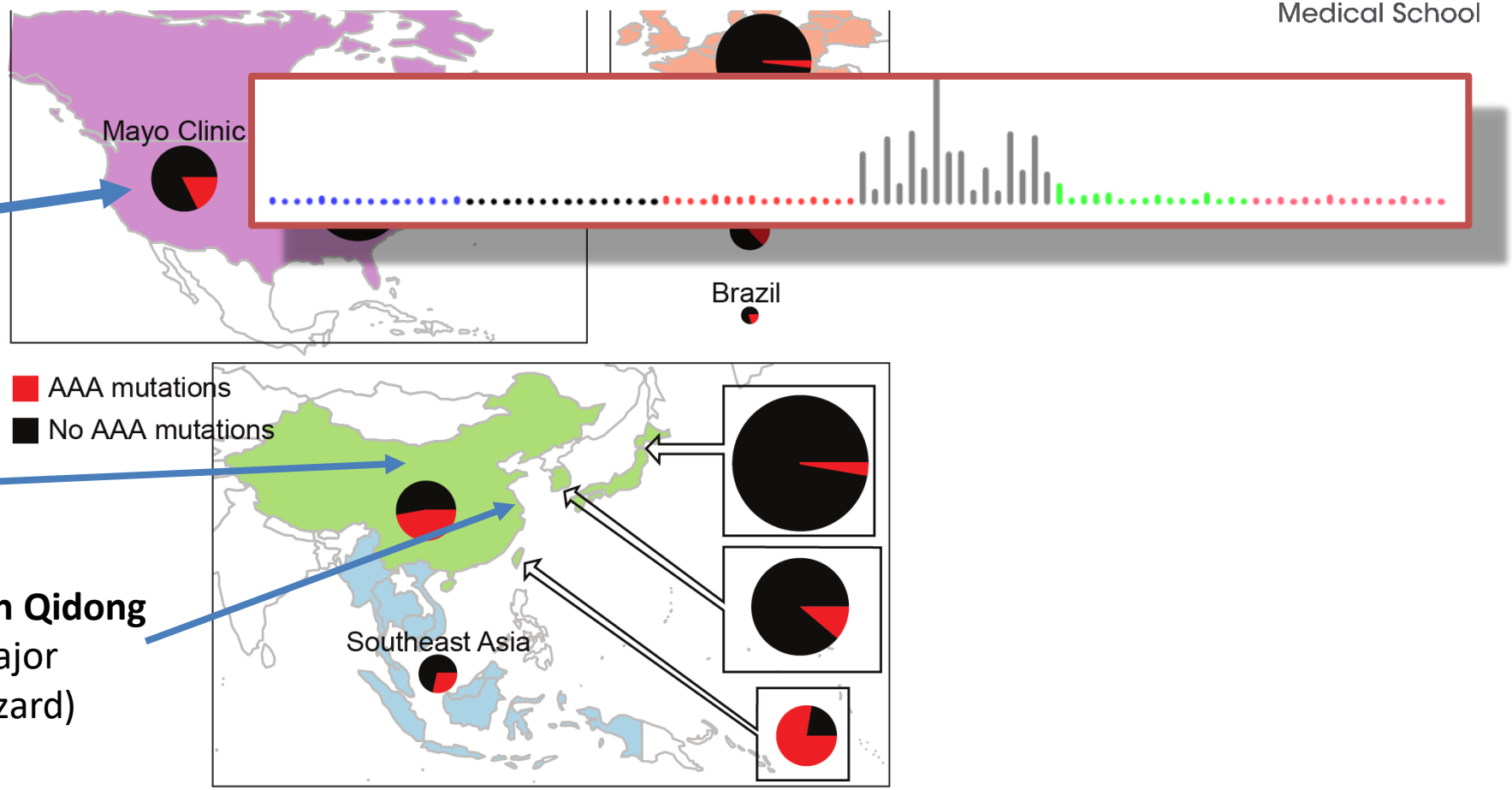
- In the overview…
  - Explained the concepts of mutational signatures
  - Gave an example of mutational signatures in molecular epidemiology: mutational signatures revealed the role of AA – aristolochic acid – in causing liver cancer and other cancers

# Prevalent AA exposure across > 1,600 liver cancers

"Asian" patients with no "country" – Overseas patients?

**~50/100 (~50%) combined HCC and bile duct cancer in China**

**12/47 (25%) in Qidong** (aflatoxin a major competing hazard)




Ng, Poon, et al., Oct 2017 DOI: 10.1126/stranslmed.aan6446 and subsequent publications; Zhang et al., 2017, Gastroenterology, 10.1053/j.gastro.2017.03.024; Xue et al., 2019, Cancer Cell,10.1016/j.ccell.2019.04.007

# Summary: What questions do we want to answer?

- **Extraction (Discovery) of new signatures**
  - Discussed approaches based on NMF (and mentioned but did not discuss approaches based on probabilistic topic models)
  - Discussed challenges (including deciding on the number of signatures) and pitfalls and the fact that discovering signatures is not a purely algorithm process
- **Attribution of known signatures**
  - Looked at issues of sparsity
  - Observed that often many attributions can plausibly reconstruct a given spectrum
  - Which suggested that we need to formally and informally incorporate prior knowledge
  - Tried to capture some of this prior knowledge using by evaluating the set of plausible attributions using maximum a posteriori probability estimates
- **Signature presence tests**
  - Use a likelihood ratio test to decide if a particular signature or set of signatures is plausibly required to reconstruct an observed spectrum

# Acknowledgements for AA in Liver Cancer


**Chang Gung Memorial Hospital**

**Sen-Yung Hsieh (谢森永)**
Hao-Yi Huang
Ming-Chin Yu
**Po-Huang Lee**
**Jacob See-Tong Pang**

**Singapore (Duke NUS, National Cancer Centre Singapore, and Johns Hopkins Singapore)**

**Arnoud Boot**
**Alvin NG Wei Tian (黄伟添)**
**Song Ling POON**
Mi Ni HUANG
Jing Quan LIM
**Nanhai Jiang**
Shenli Zhang
Szu-Chi Ho
Willie Yu
Yuka Suzuki
Cedric C. Y. Ng
**Patrick Tan**
**Bin Tean TEH**
**Alex Chang**

**Funding**

Singapore National Medical Research Council

Singapore Ministry of Health via the Duke-NUS Signature Research Programmes

Chang Gung Medical Foundation


**https://tinyurl.com/aa-liver-cancer**

# General acknowledgements


- The International Cancer Genome and The Cancer Genome Atlas Pan-Cancer Analysis of Whole Genomes Consortium (PCAWG)
- The PCAWG Mutational Signatures Analysis Working Group
  - Co-lead Mike Stratton, many contributors including Ludmil Alexandrov, Jaegil Kim, Gaddy Getz, Nick Haradhvala, and many many others, see https://doi.org/10.1038/s41586-020-1943-3
- My lab
  - Arnoud Boot, Szu-Chi HO, Mini Huang, Nanhai Jiang, John McPherson, LIU Mo, Alvin Wei Tian Ng, WU Yang, Willie Yu, Shenli Zhang

# Caution: Conceptual Hazard

These analytical tools are not oracles
They do not algorithmically reveal a Platonic truth
Their analyses need to be assessed by humans in the light of all available evidence

# “All models are wrong but some are useful”

George E. P. Box (1979), "Robustness in the strategy of scientific model building", in Launer & Wilkinson, *Robustness in Statistics*

I hope you will find some of these models useful

More research is needed!

Thank you and questions?

steverozen@gmail.com

# Backup Slides

Medicinal uses of species of the genus *Aristolochia*.

| Use | Number of citations, total ($n$) = 1445 | % of citations of the total number of citations |
|---|---|---|
| Cardiovascular | 29 | 2.0 |
| Central nervous system | 99 | 6.8 |
| Bites and poison | 147 | 10.2 |
| Dermatology | 80 | 5.5 |
| Endocrinology | 17 | 1.2 |
| Gastrointestinal | 215 | 14.9 |
| Gynaecology including STDs | 113 | 7.8 |
| Infectious diseases | 78 | 5.4 |
| Musculoskeletal | 67 | 4.6 |
| Respiratory | 67 | 4.6 |
| Nephrology | 73 | 5.1 |
| Parasitology | 85 | 5.9 |
| Veterinary uses | 8 | 0.6 |
| Miscellaneous including general 'medicinal use' | 227 | 15.7 |

M. Heinrich et al 2009
doi:10.1016/j.jep.2009.05.028

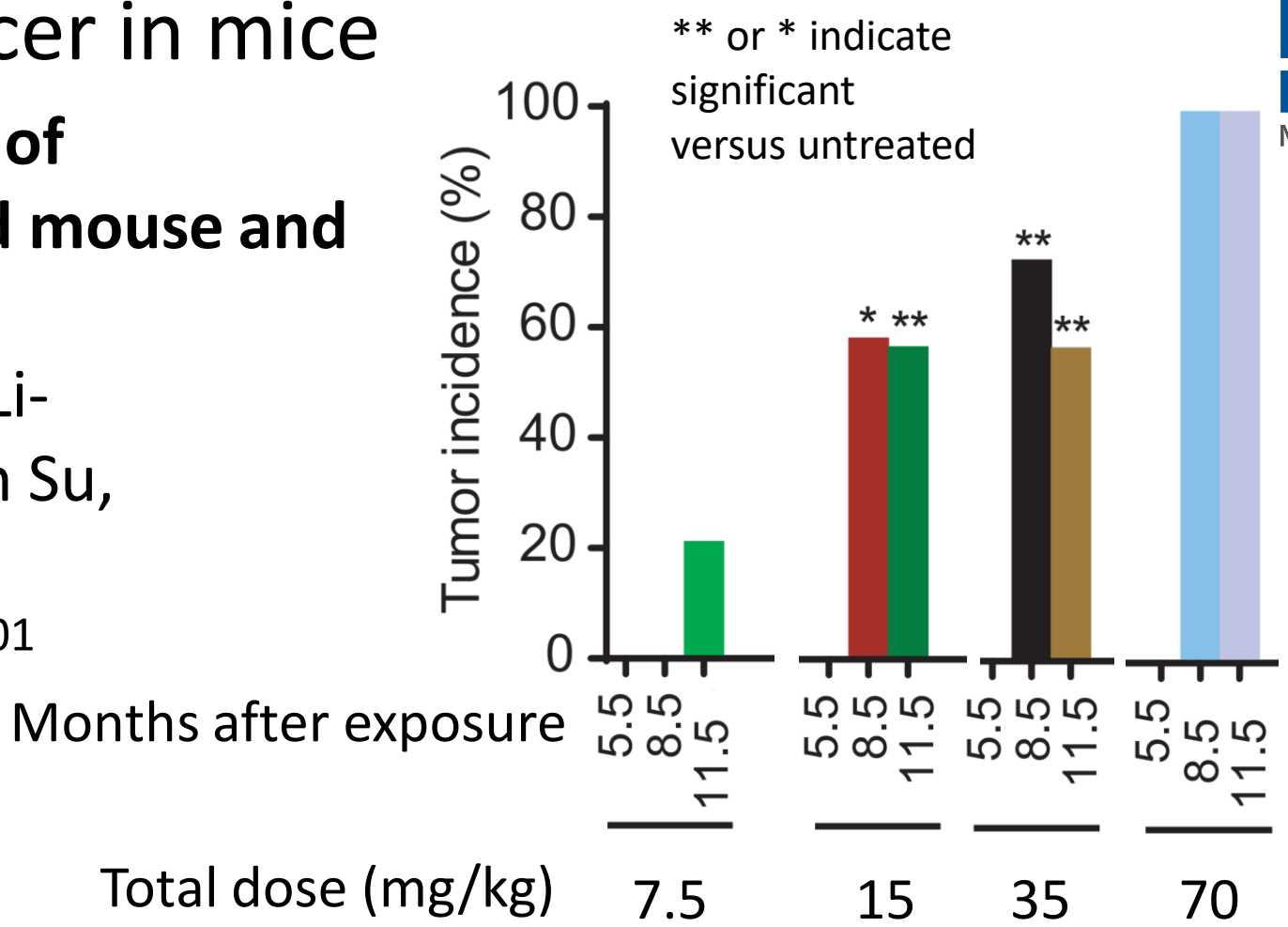
# AA causes liver cancer in mice

**The mutational features of aristolochic acid-induced mouse and human liver cancers**

Zhao-Ning Lu, Qing Luo, Li-Nan Zhao, Yi Shi, Xian-Bin Su, Ze-Guang Han

**doi:** https://doi.org/10.1101/507301

Dec 28, 2018




(Re-drawn from Lu et al.)