Signatures of mutational processes in human cancer

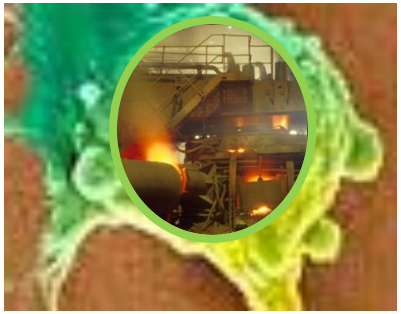
wellcome trust sanger institute

International Cancer Genome Consortium

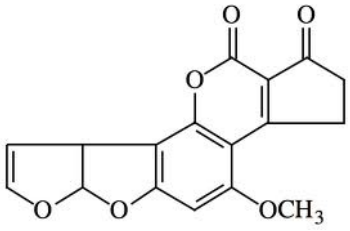
Insights from cancer genome sequencing

- Mutated cancer genes
- Biological processes implicated in cancer development
- Identification of drug targets
- Evolution of subclones
- Evolution of metastases
- Infectious pathogens involved in oncogenesis
- Prediction of outcome and response to therapy
- Circulating DNA for monitoring and early detection
- Mutational processes operative in carcinogenesis

Somatic mutations occur in all cells of the body throughout life



.....ATCGGGAATCGGACCCGATG.....



TP53 mutations in human cancer

TP53 mutations in human cancer

C>T

C>A

C>G

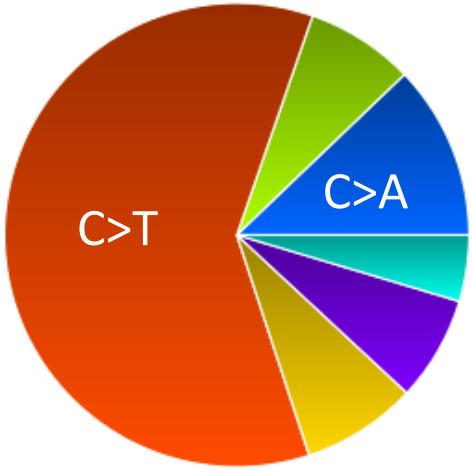
T>A

T>C

T>G

TP53 mutations in human cancer

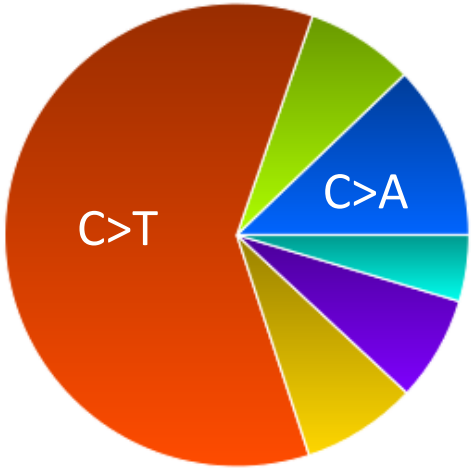
686 skin cancers



Ultraviolet light
causes C>T mutations

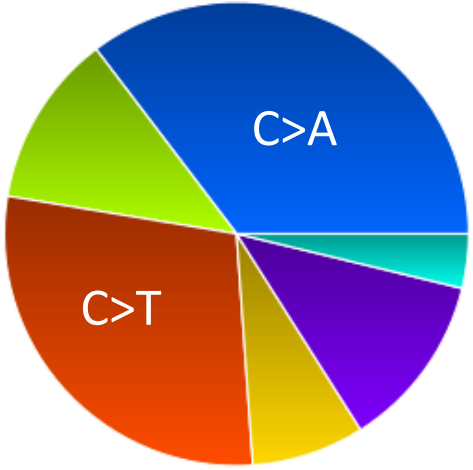
TP53 mutations in human cancer

686 skin cancers



Ultraviolet light causes C>T mutations

1647 lung cancers



Tobacco carcinogens cause C>A mutations

From fertilised egg to cancer cell

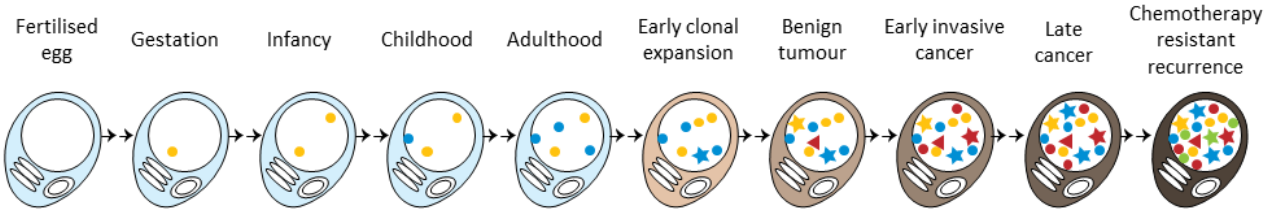
Chemotherapy resistant recurrence

From fertilised egg to cancer cell

Chemotherapy resistant recurrence

From fertilised egg to cancer cell

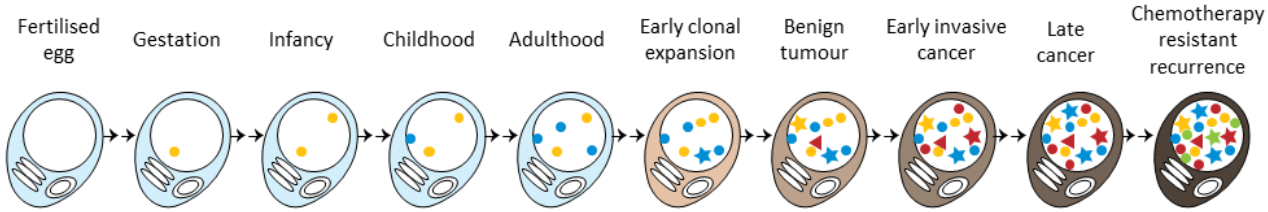
Mutational processes



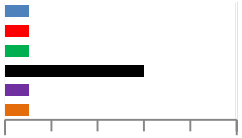
A

From fertilised egg to cancer cell

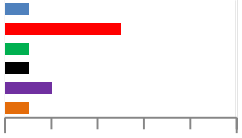
Mutational processes



A

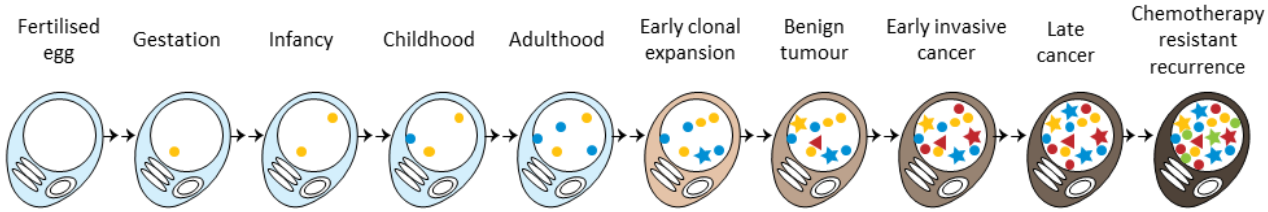


B

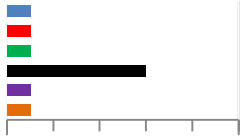


From fertilised egg to cancer cell

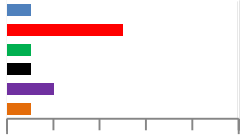
Mutational processes



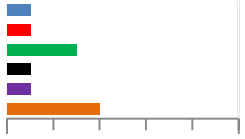
A



B



C



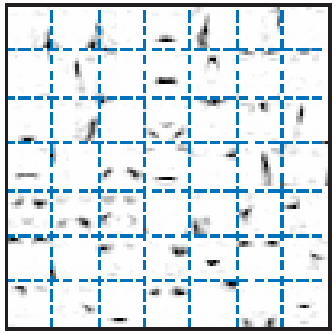
From fertilised egg to cancer cell

Non-negative matrix factorization (NMF)

Learning the parts of objects by non-negative matrix factorization

Daniel D. Lee* **& H. Sebastian Seung*†**

** Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA*
† Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

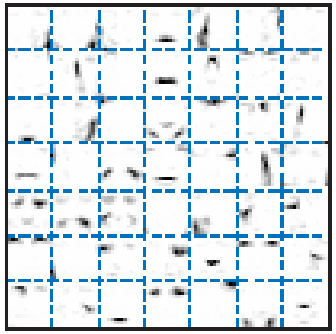


Non-negative matrix factorization (NMF)

Learning the parts of objects by non-negative matrix factorization

Daniel D. Lee* **& H. Sebastian Seung*†**

** Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA*
† Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA



Cell

Mutational Processes Molding the Genomes of 21 Breast Cancers

Serena Nik-Zainal,[1] Ludmil B. Alexandrov,[1] David C. Wedge,[1] Peter Van Loo,[1,2,3] Christopher D. Greenman,[1,4,5]

Cell PRESS Open ACCESS

Cell Reports
Resource

Deciphering Signatures of Mutational Processes Operative in Human Cancer

Ludmil B. Alexandrov,[1] Serena Nik-Zainal,[1] David C. Wedge,[1] Peter J. Campbell,[1,2,3] and Michael R. Stratton[1,*]

ARTICLE

doi:10.1038/nature12477

Signatures of mutational processes in human cancer

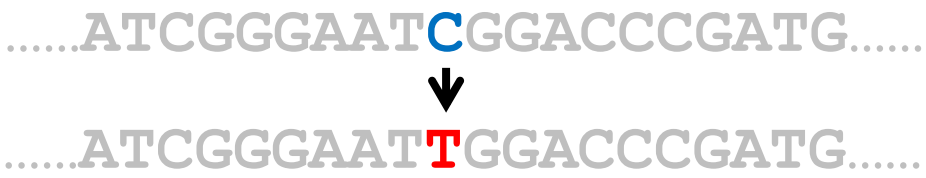
A list of authors and their affiliations appears at the end of the paper

Classification of base substitution mutations

C>T
C>A
C>G
T>A
T>C
T>G

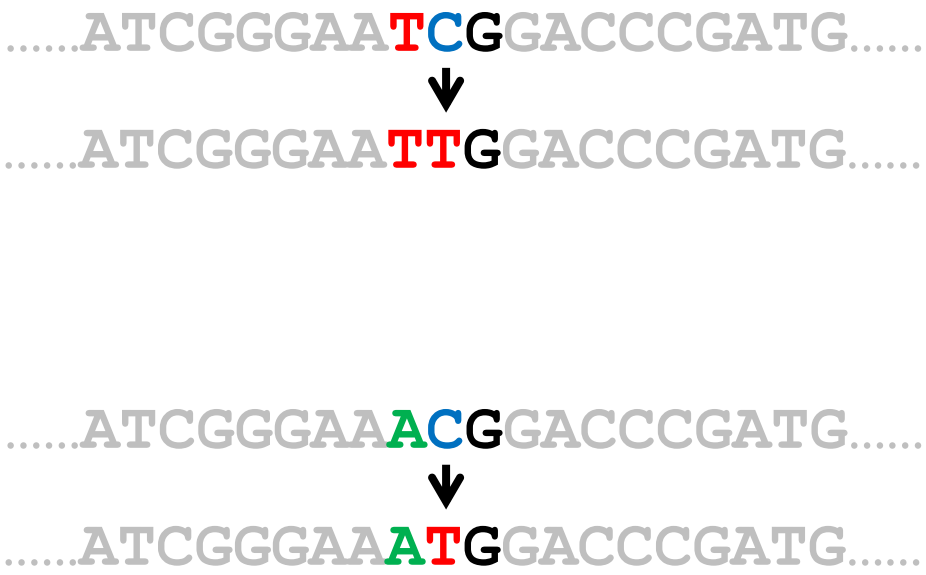
6 mutation classes

Classification of base substitution mutations

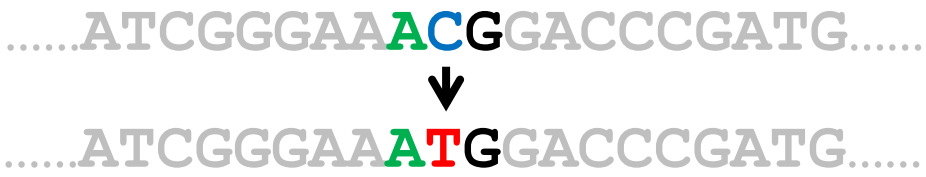


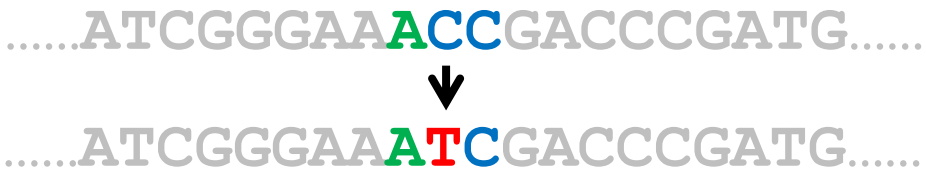
Classification of base substitution mutations

Classification of base substitution mutations



Classification of base substitution mutations



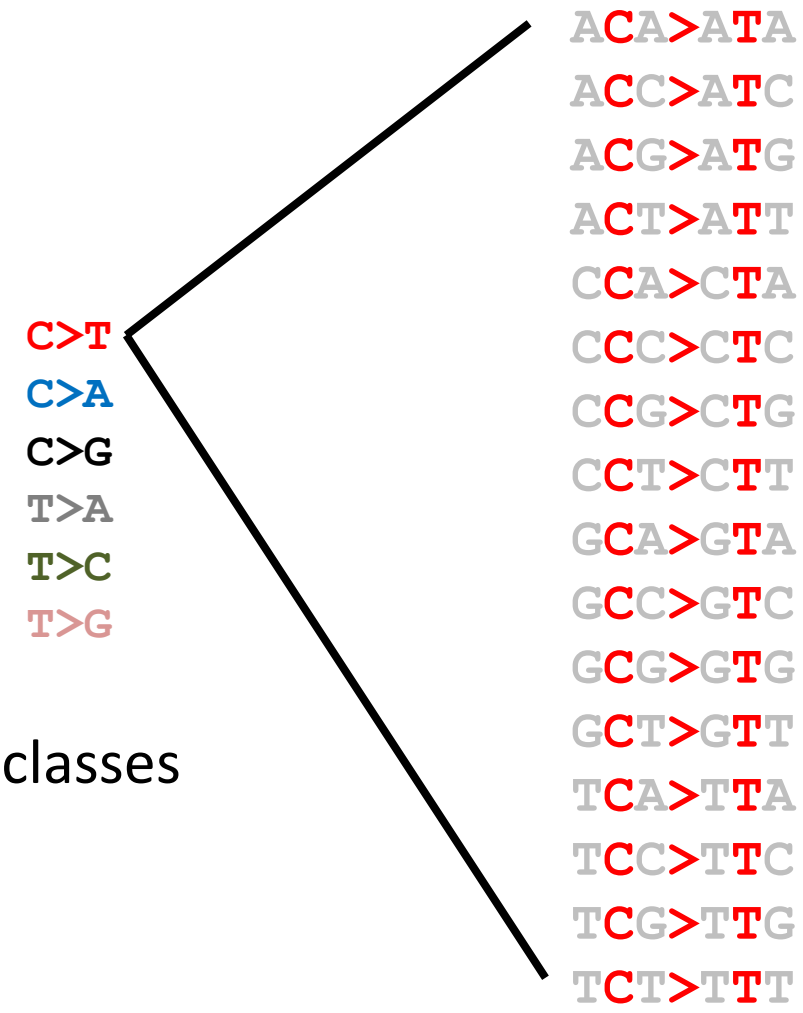


Classification of base substitution mutations

C>T
C>A
C>G
T>A
T>C
T>G

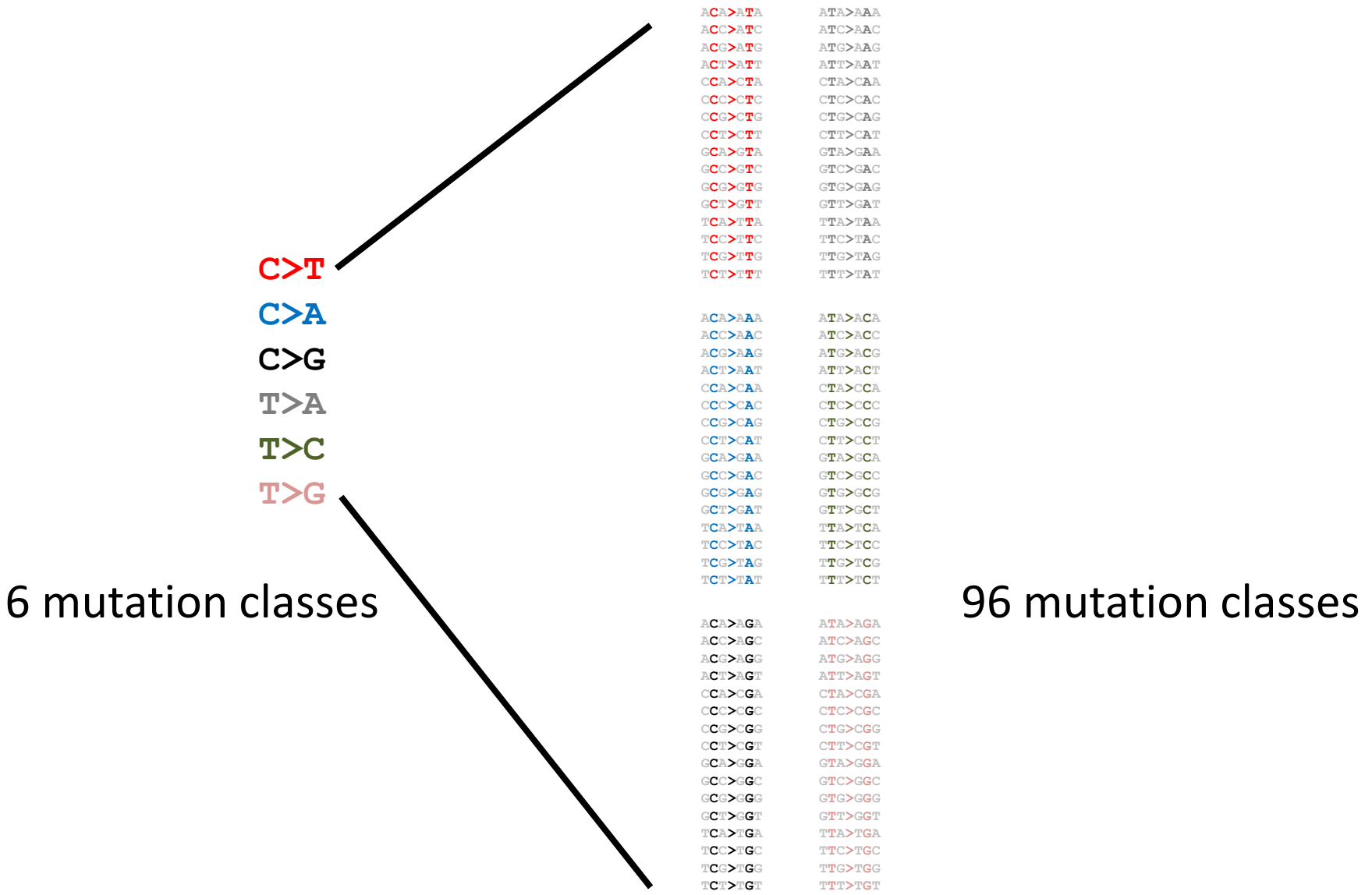
6 mutation classes

Classification of base substitution mutations



6 mutation classes

Classification of base substitution mutations



Extracting mutational signatures from human cancers

7,042 cancer cases of 30 cancer types

Extracting mutational signatures from human cancers

7,042 cancer cases of 30 cancer types

7,042 mutation catalogues = 4,942,984 somatic mutations

Extracting mutational signatures from human cancers

7,042 cancer cases of 30 cancer types

↓

7,042 mutation catalogues = 4,942,984 somatic mutations

↓

Apply Non-negative Matrix Factorisation (NMF)

Extracting mutational signatures from human cancers

7,042 cancer cases of 30 cancer types

↓

7,042 mutation catalogues = 4,942,984 somatic mutations

↓

Apply Non-negative Matrix Factorisation (NMF)

↓

Extract mutational signatures

Extracting mutational signatures from human cancers

7,042 cancer cases of 30 cancer types

↓

7,042 mutation catalogues = 4,942,984 somatic mutations

↓

Apply Non-negative Matrix Factorisation (NMF)

↓

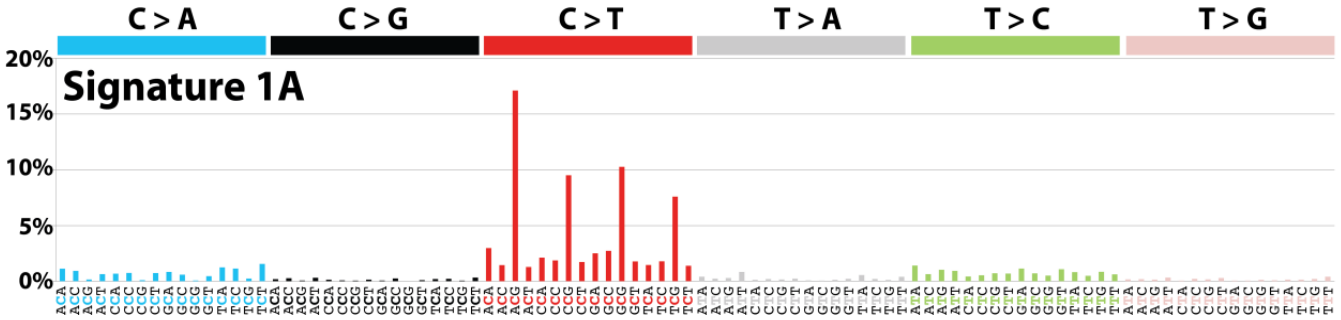
Extract mutational signatures

↓

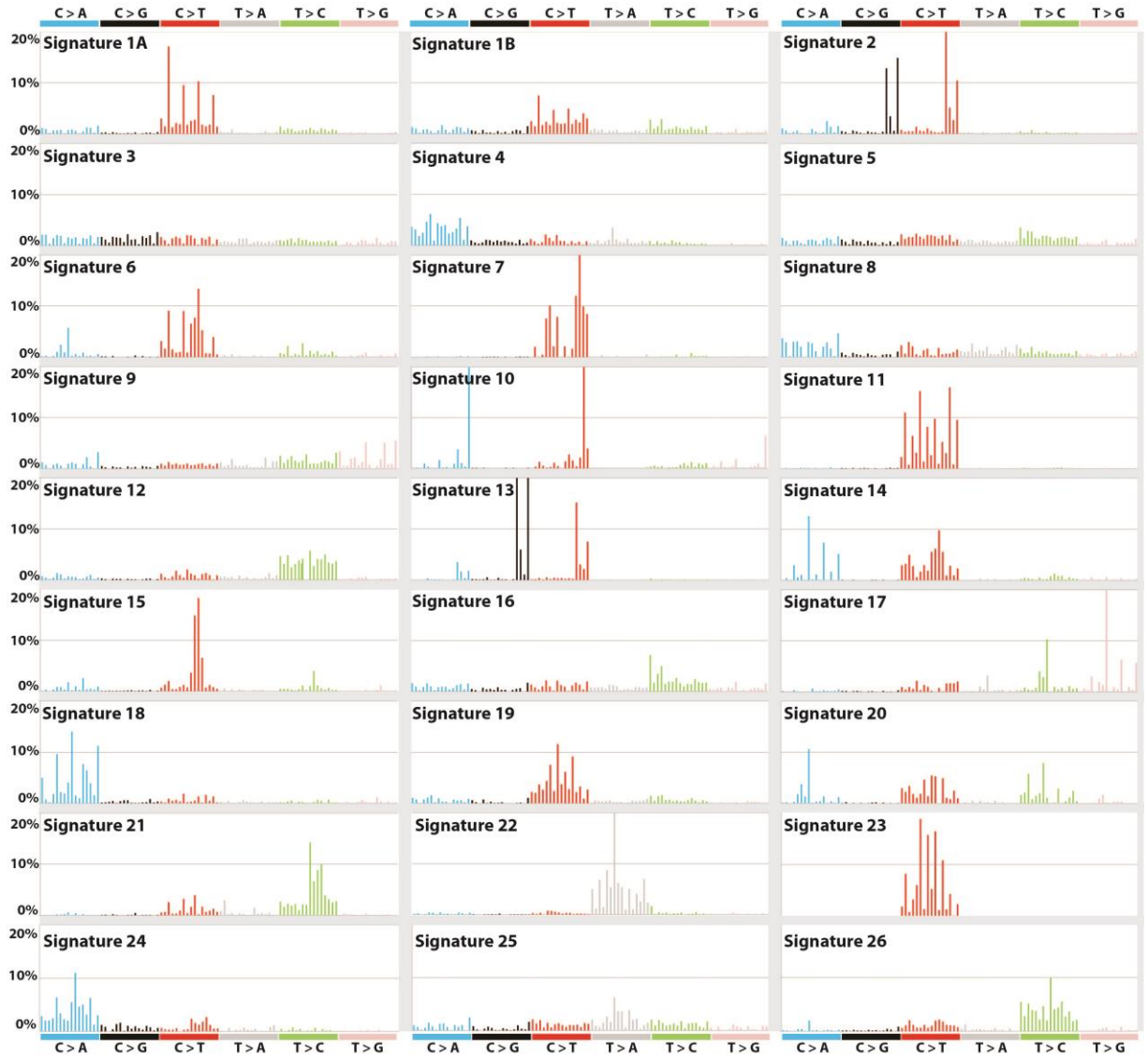
Estimate the contribution of each mutational signature to the mutational catalogue of each of the 7,042 cancer genomes

The landscape of mutational signatures across human cancer

Mutational signatures in human cancer

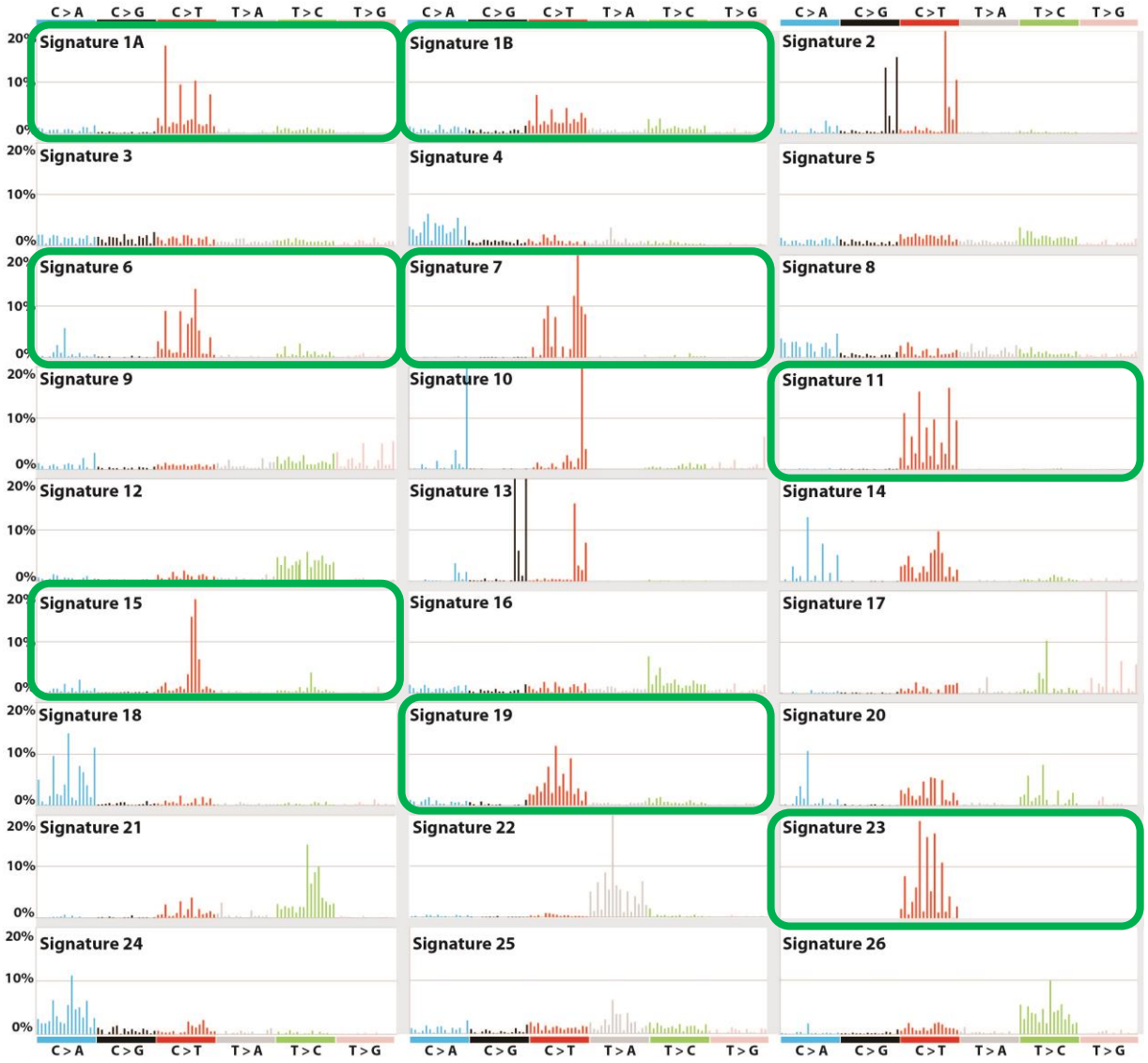


Mutational signatures in human cancer



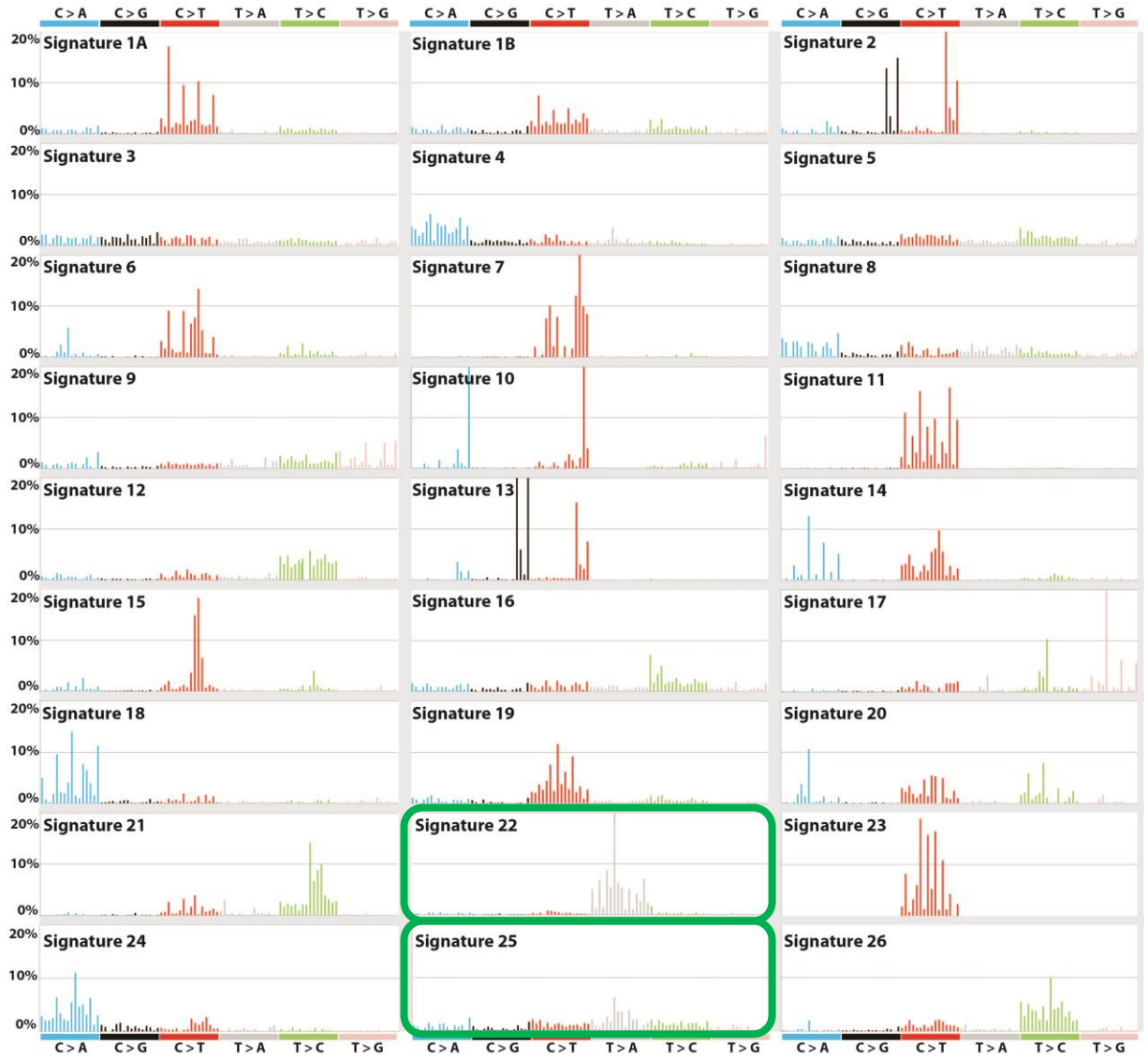
Mutational signatures in human cancer

Mutational signatures in human cancer



Mutational signatures in human cancer

Mutational signatures in human cancer



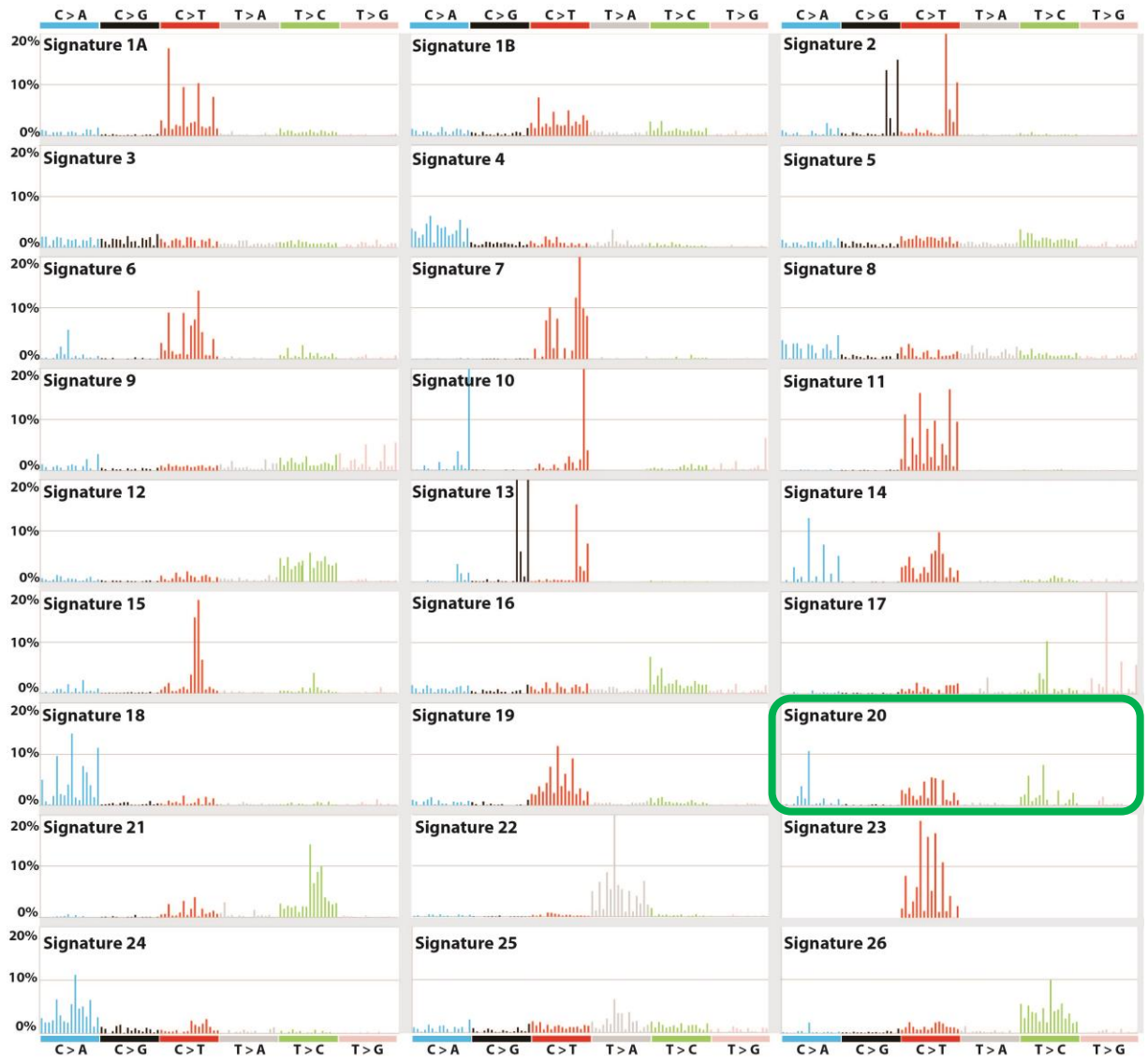
Mutational signatures in human cancer

Mutational signatures in human cancer

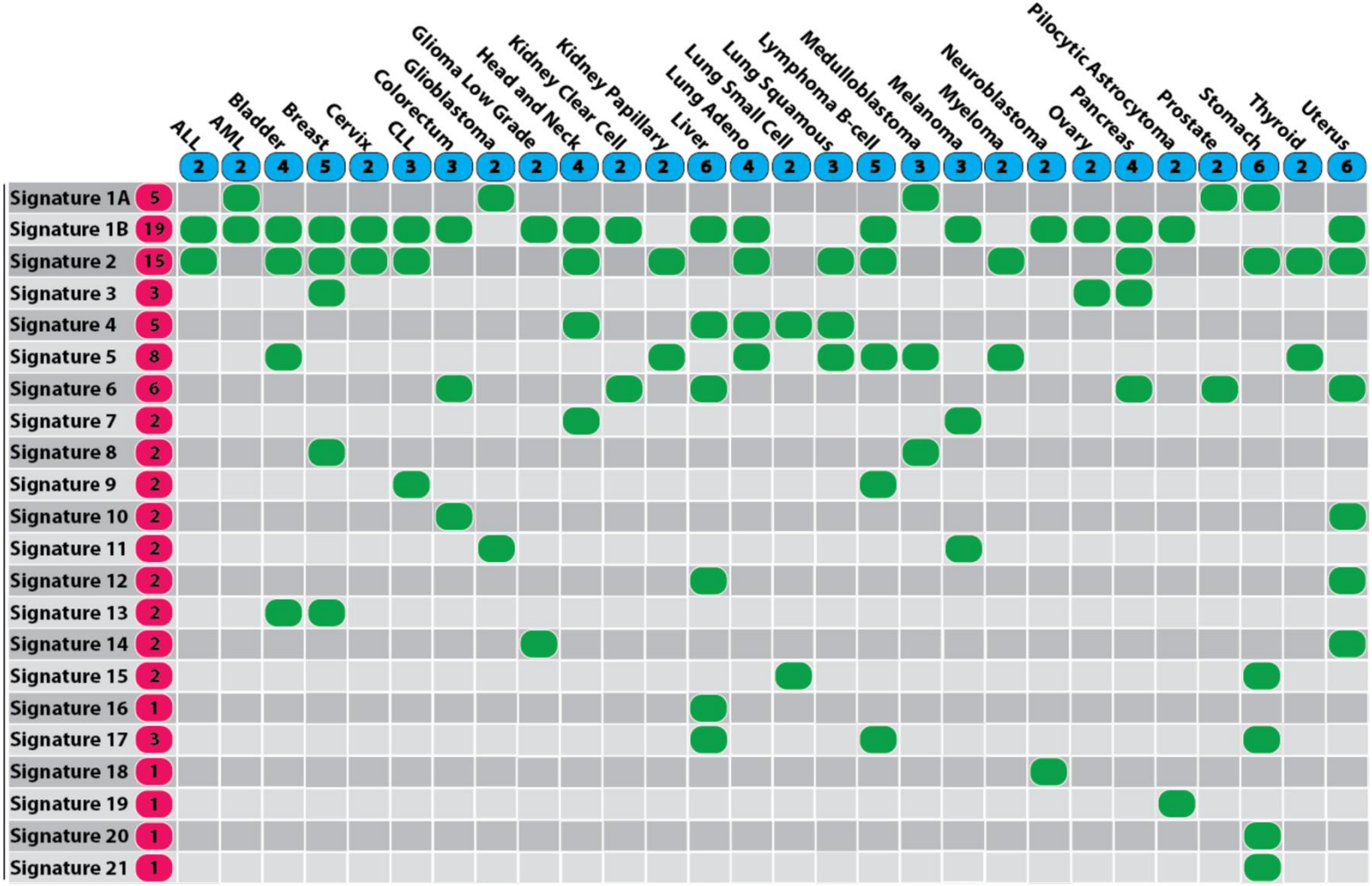
Mutational signatures in human cancer

Mutational signatures in human cancer

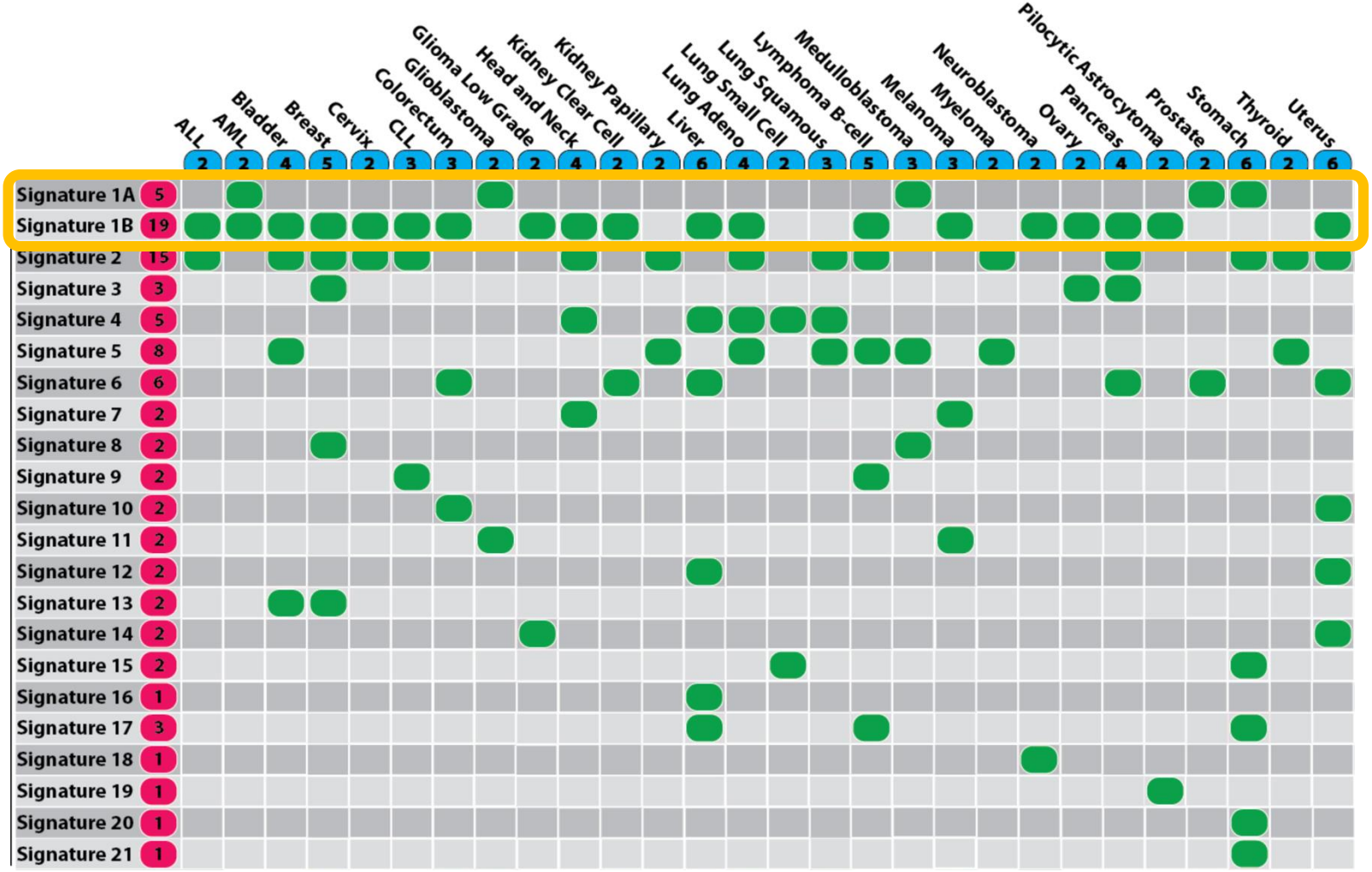
Mutational signatures in human cancer



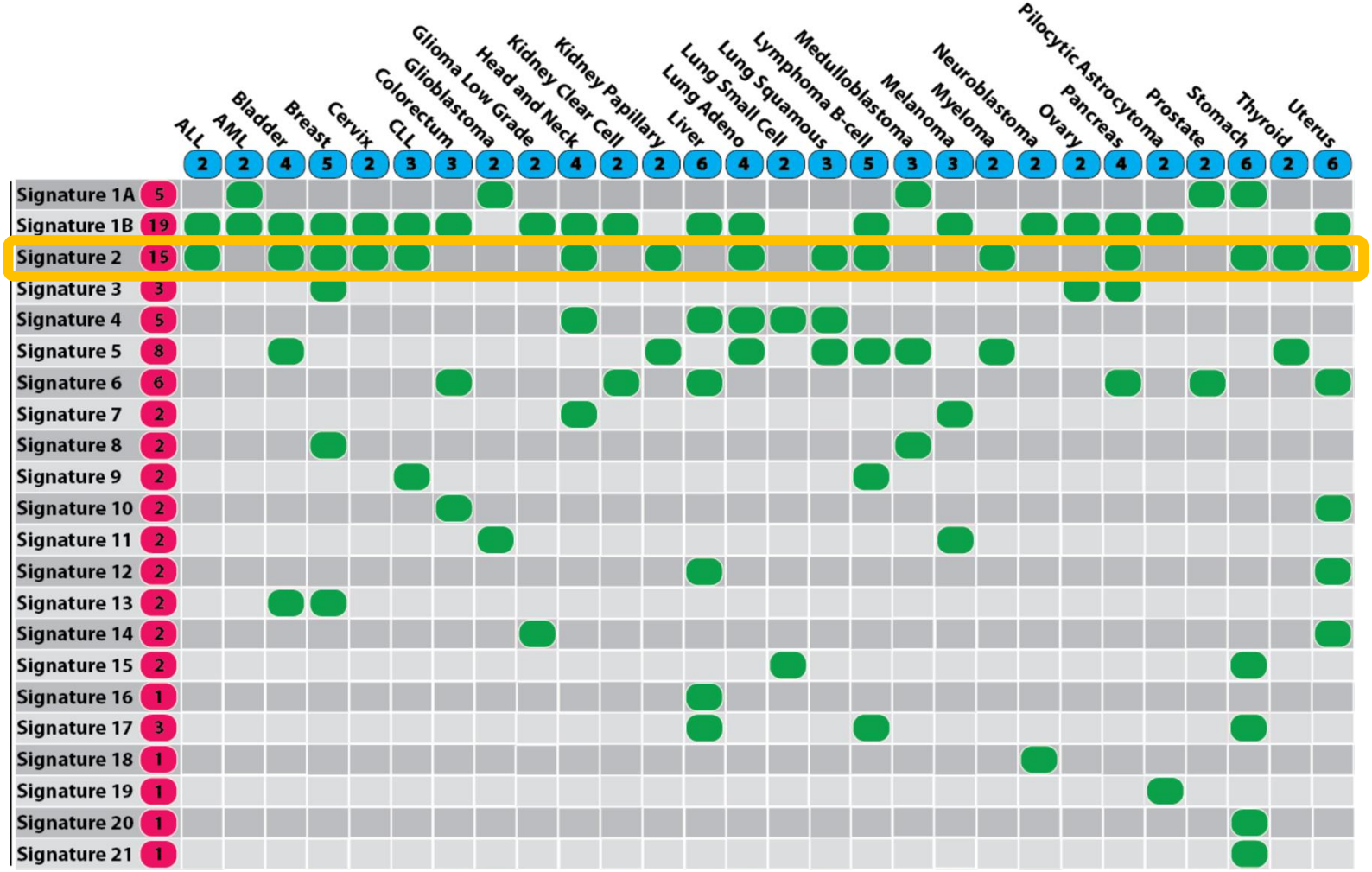
Mutational signatures by cancer type



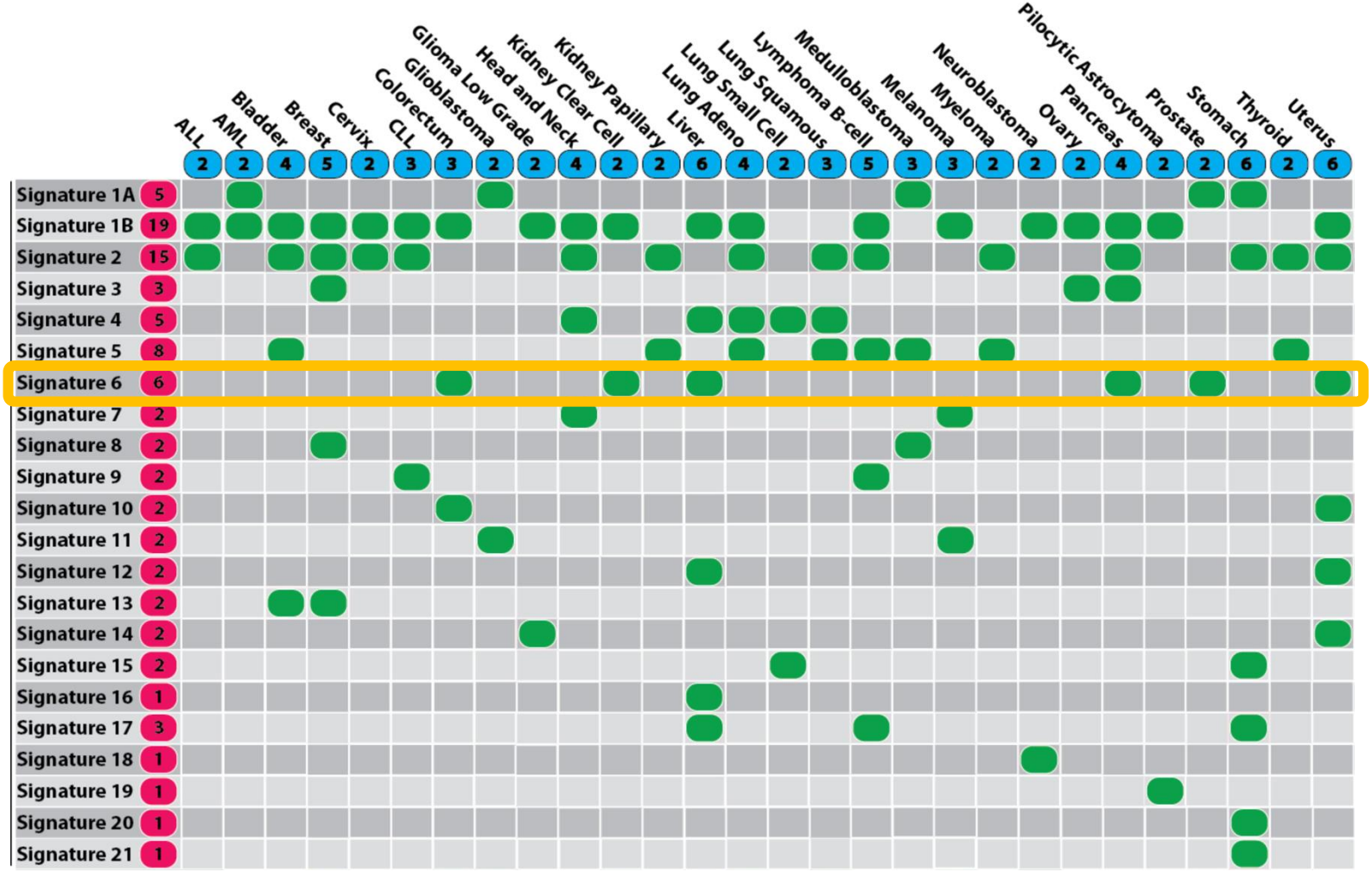
Mutational signatures by cancer type



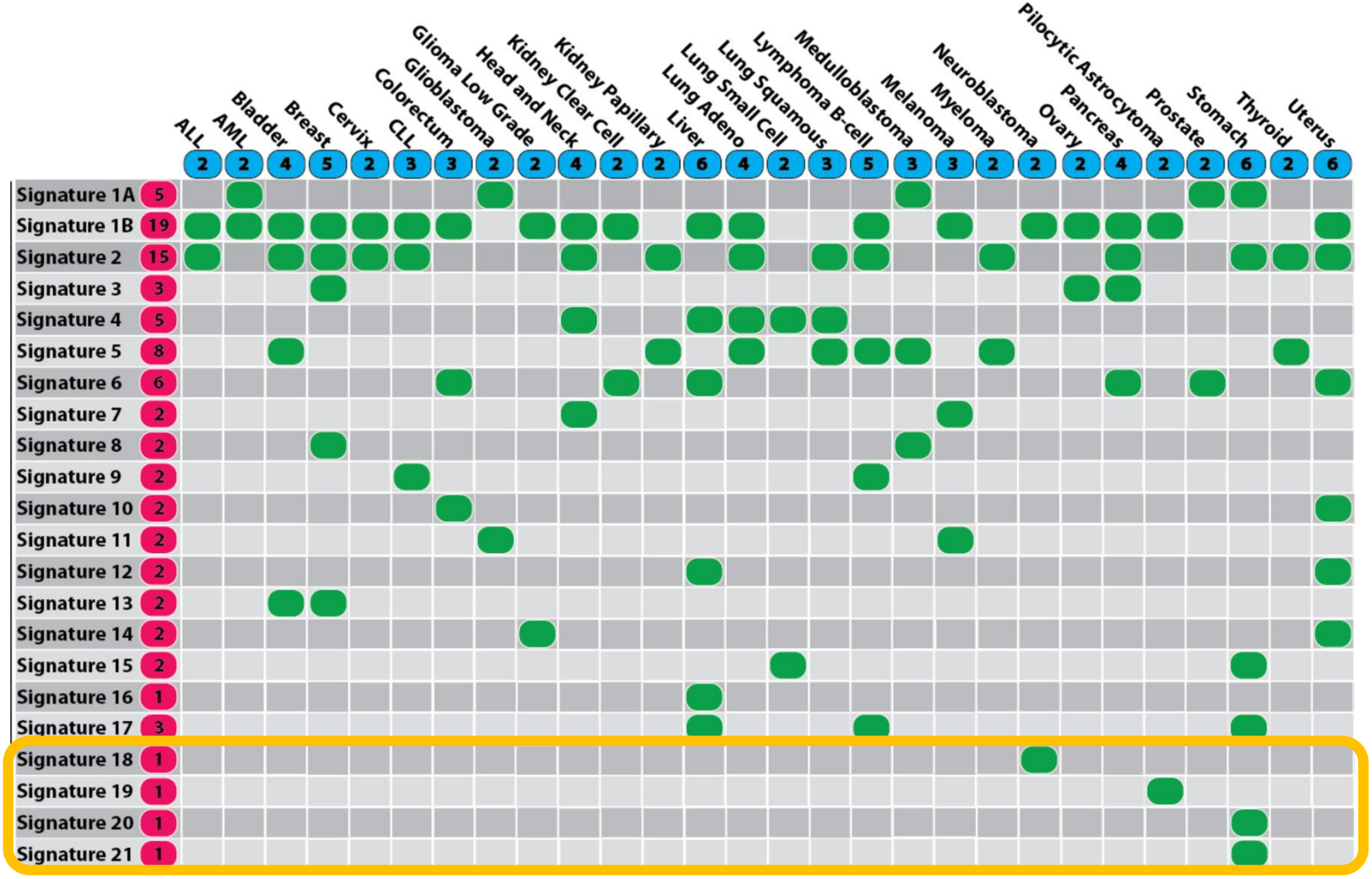
Mutational signatures by cancer type



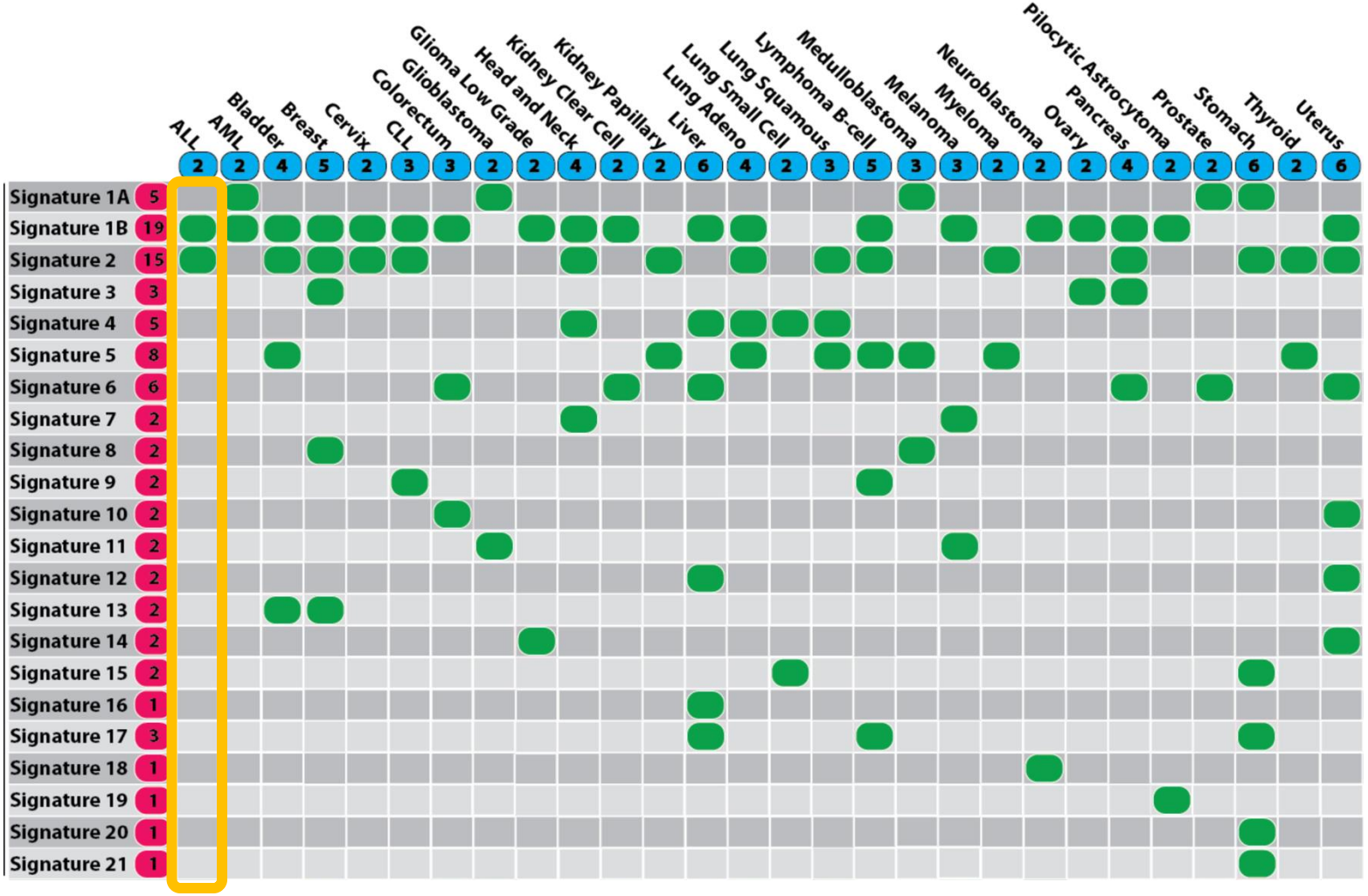
Mutational signatures by cancer type



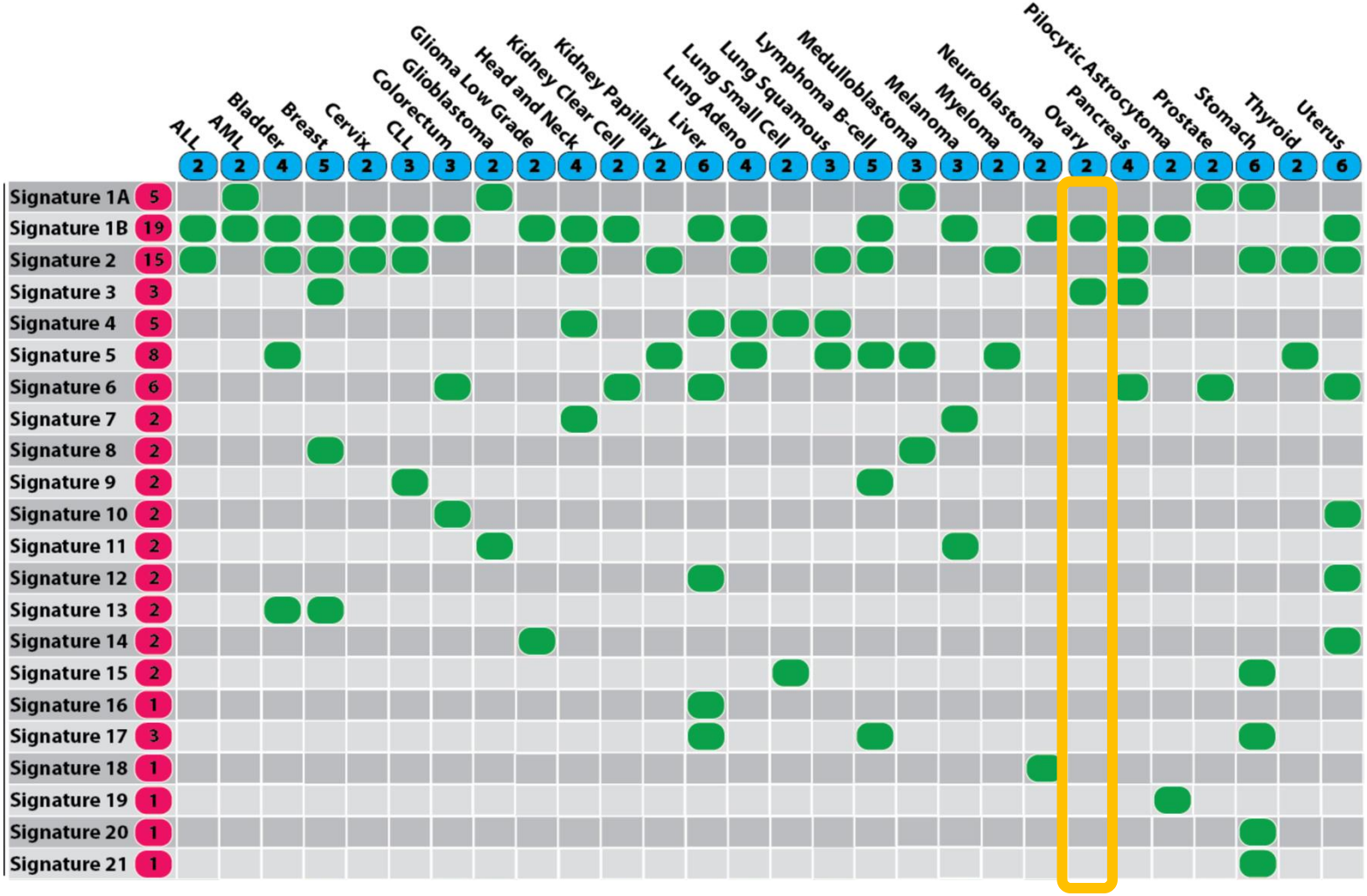
Mutational signatures by cancer type



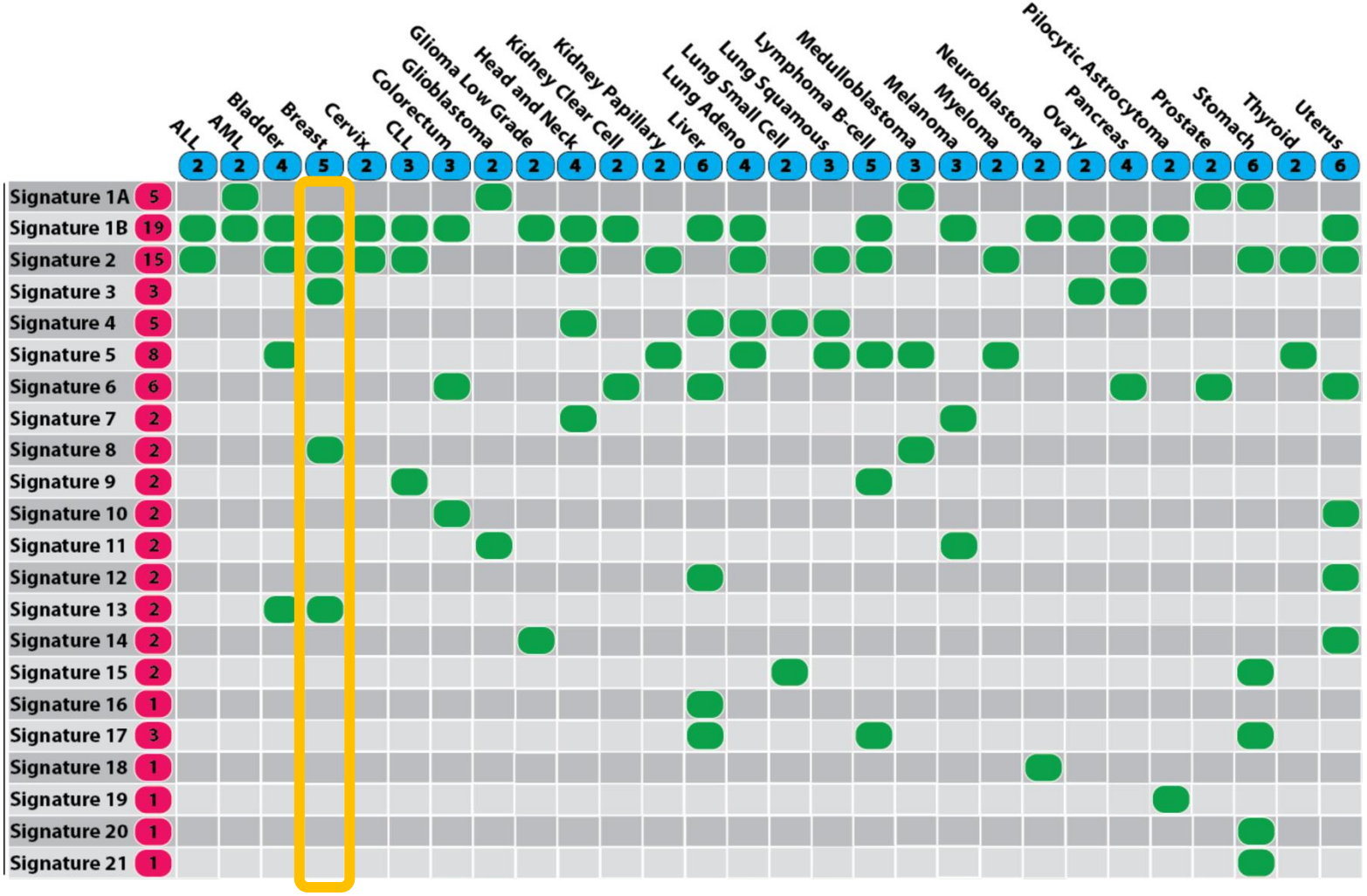
Mutational signatures by cancer type



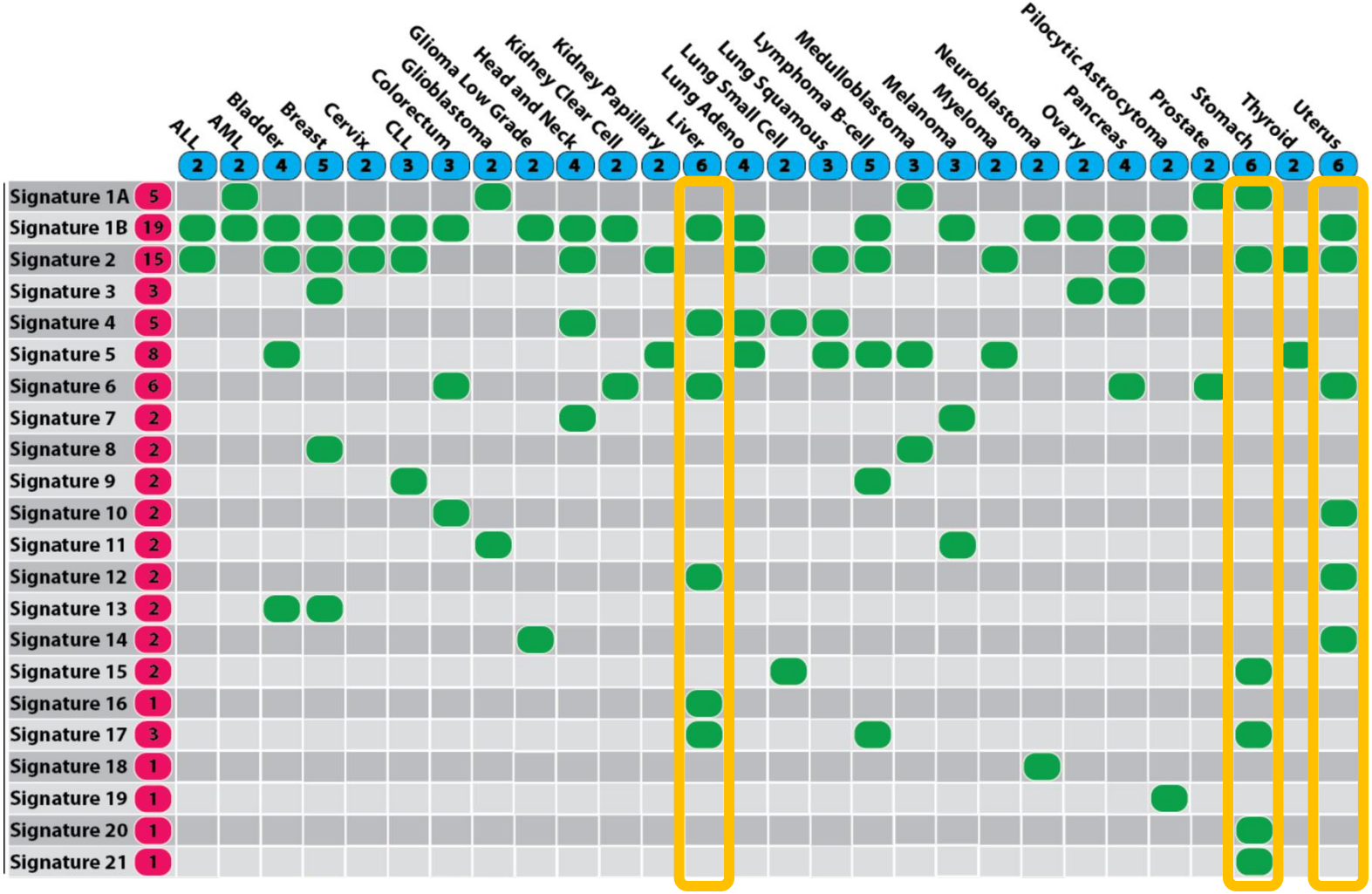
Mutational signatures by cancer type



Mutational signatures by cancer type

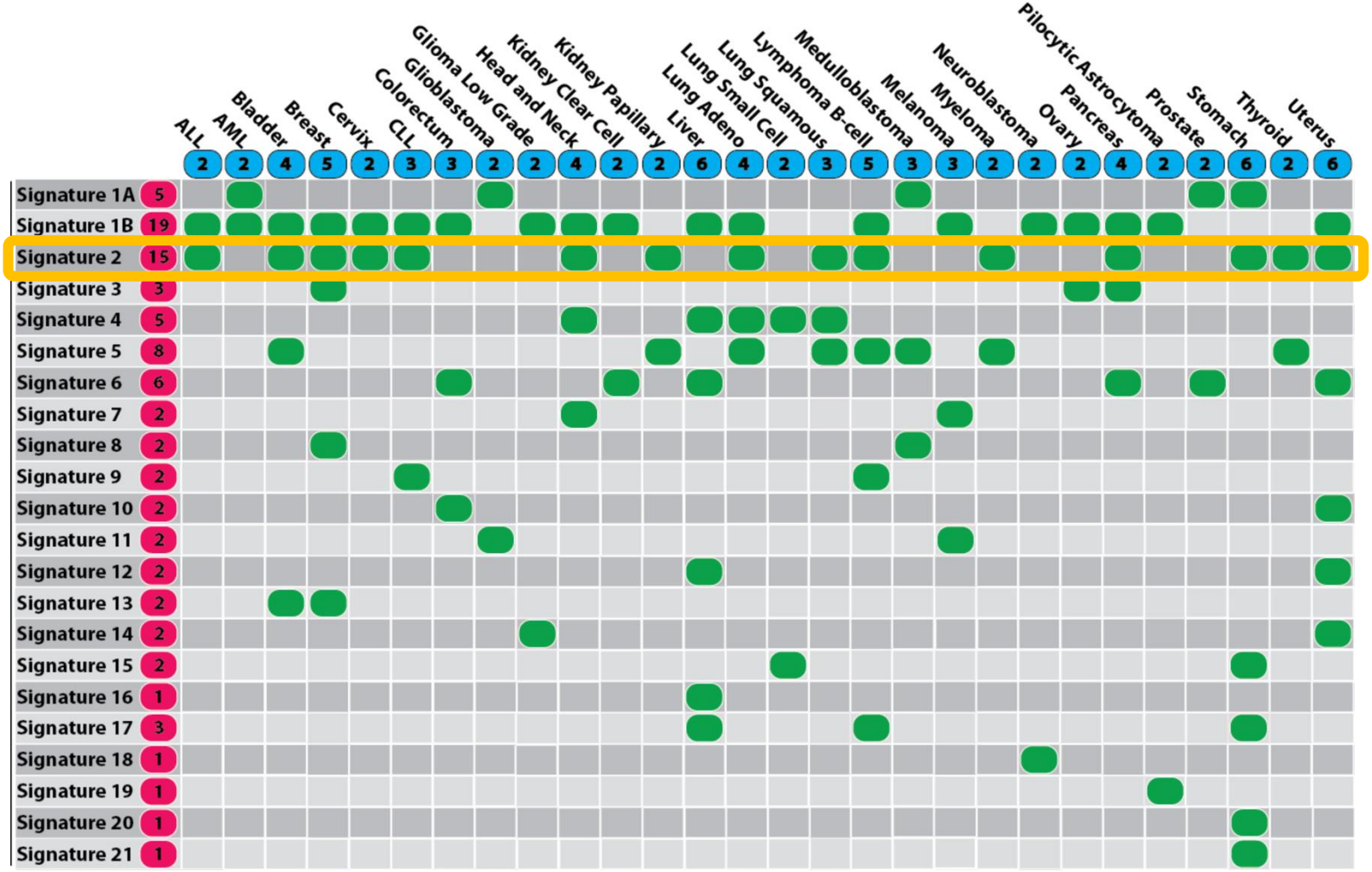


Mutational signatures by cancer type



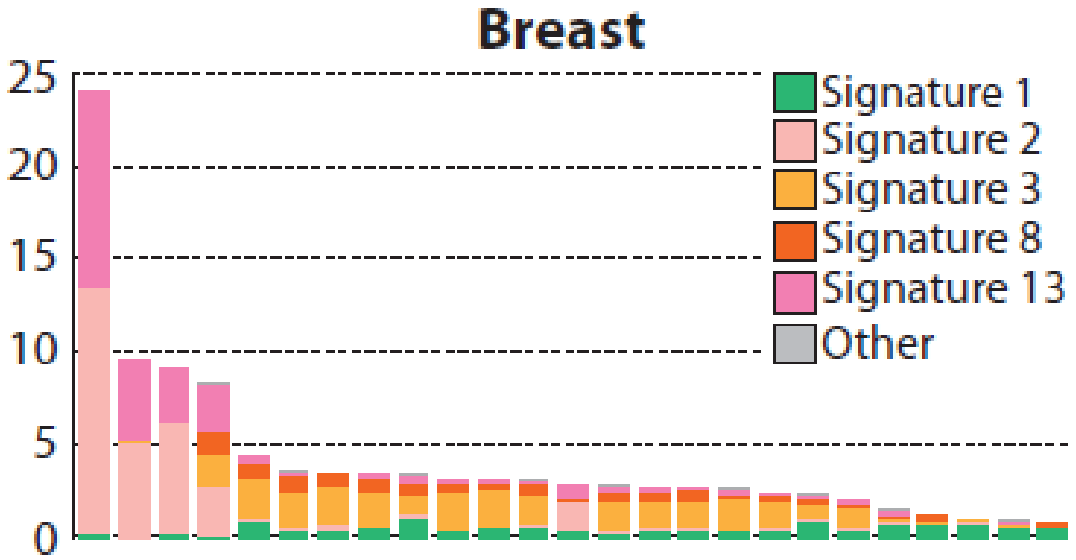
Mutational processes underlying mutational signatures

Etiologies of mutational signatures

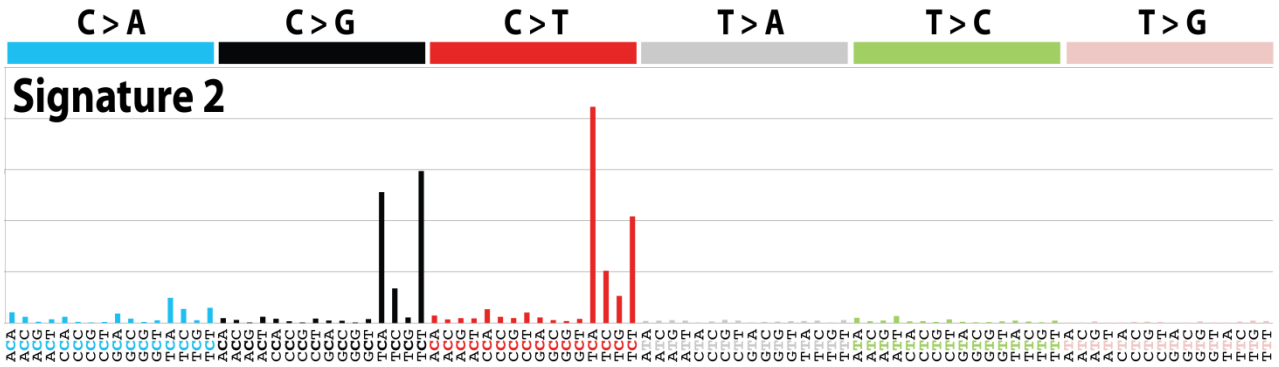


Contributions of mutation signatures to individual cancer cases

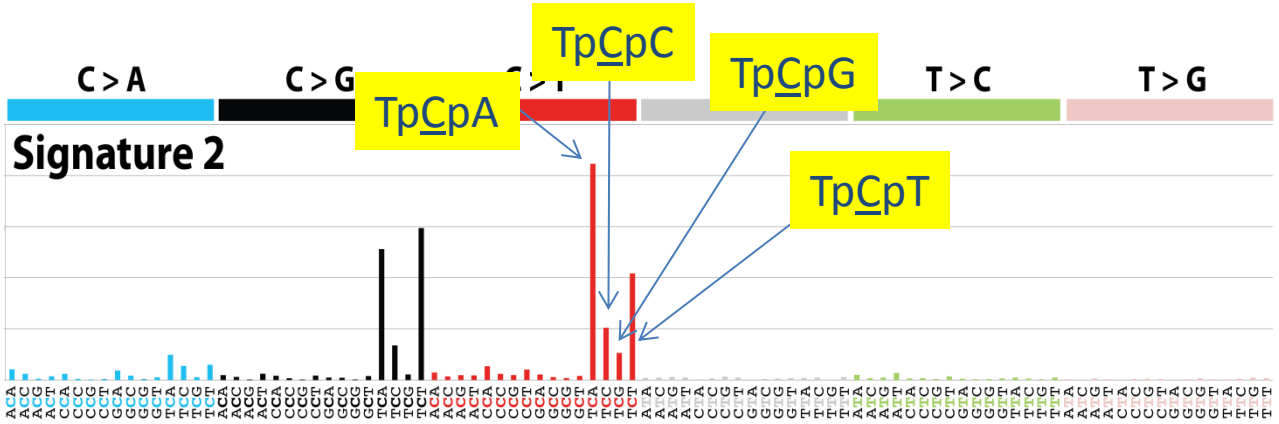
Mutations/Mb



Signature 2 is characterised by C>T and C>G substitutions



Signature 2 is characterised by C>T and C>G substitutions at TpCpN trinucleotides



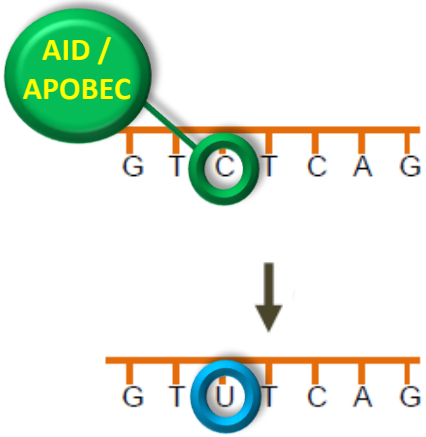
What is the mutational process underlying Signature 2?

- Deamination of cytosine by one of the family of AID/APOBEC enzymes?

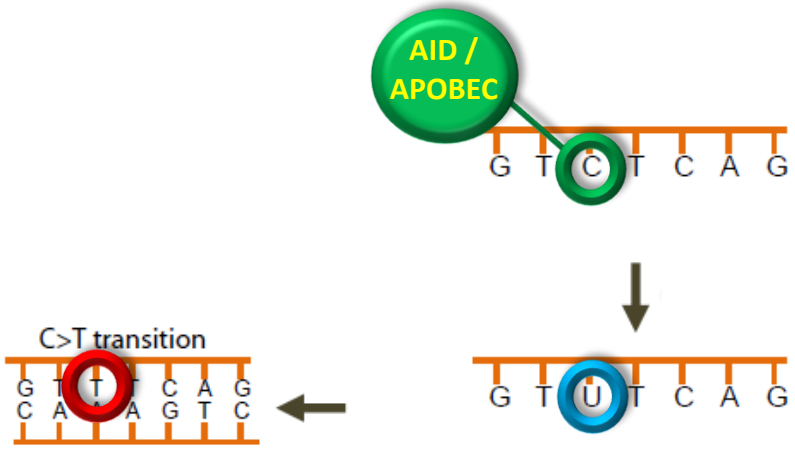
- The family includes
  - AID
  - APOBEC1
  - APOBEC2
  - APOBEC3A-H
  - APOBEC4

DNA editing by the AID/APOBEC family of cytidine deaminases

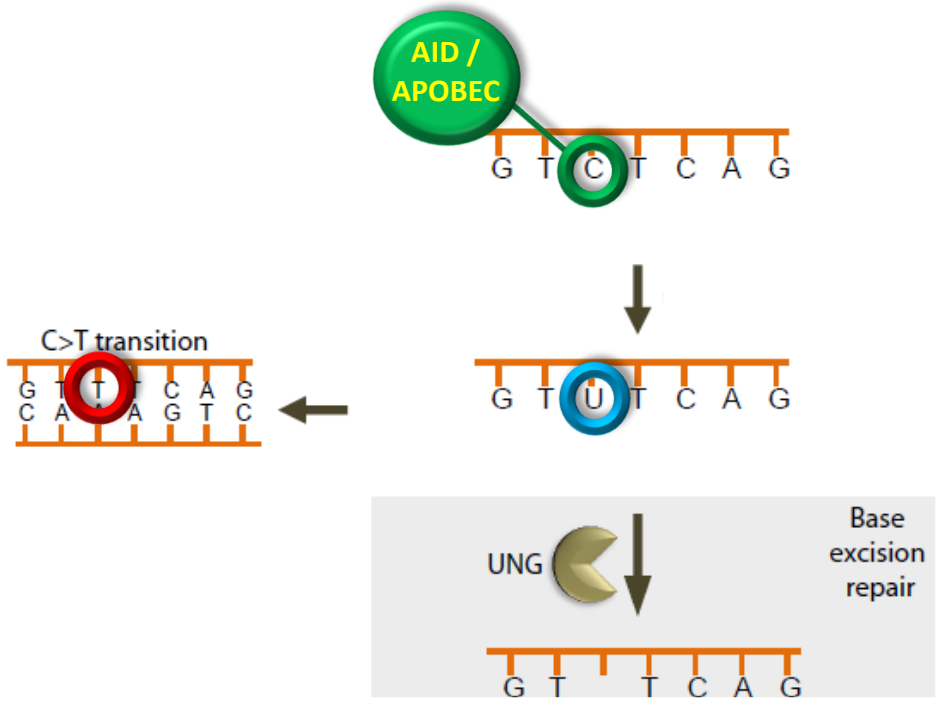
DNA editing by the AID/APOBEC family of cytidine deaminases



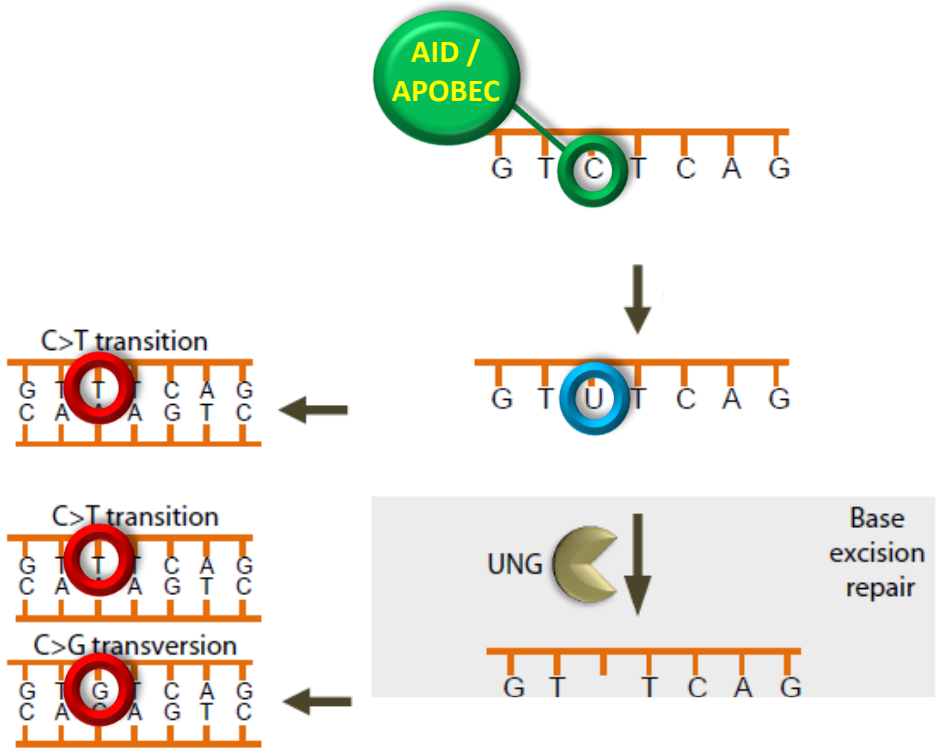
DNA editing by the AID/APOBEC family of cytidine deaminases



DNA editing by the AID/APOBEC family of cytidine deaminases



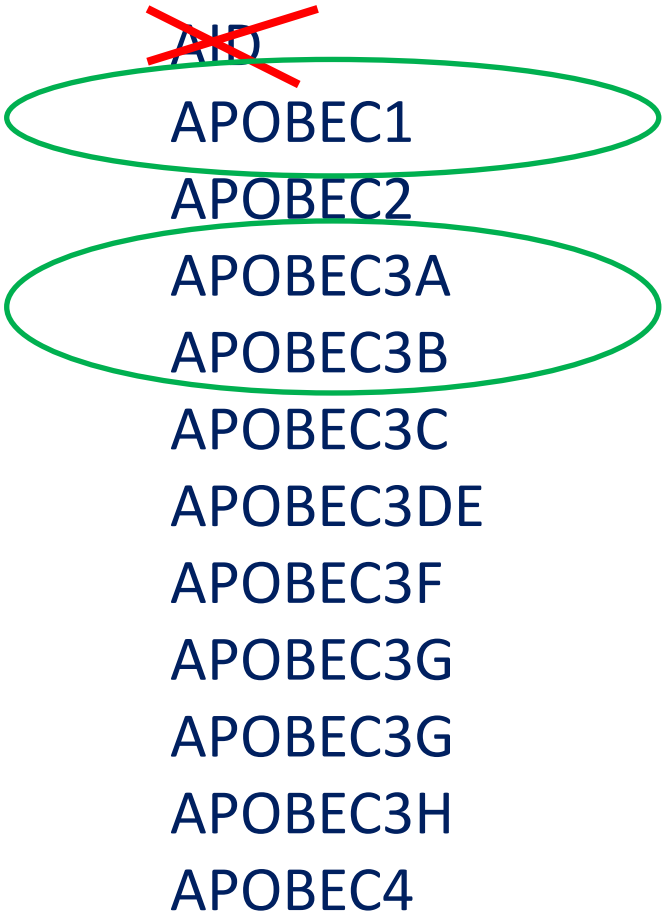
DNA editing by the AID/APOBEC family of cytidine deaminases



The AID / APOBEC family of cytidine deaminases have normal functions that require DNA editing

- AID plays a central role in somatic hypermutation and class switch recombination at the immunoglobulin loci
- APOBEC3A-H mutate HIV, Hepatitis B virus and retrotransposons to restrict their activity and replication

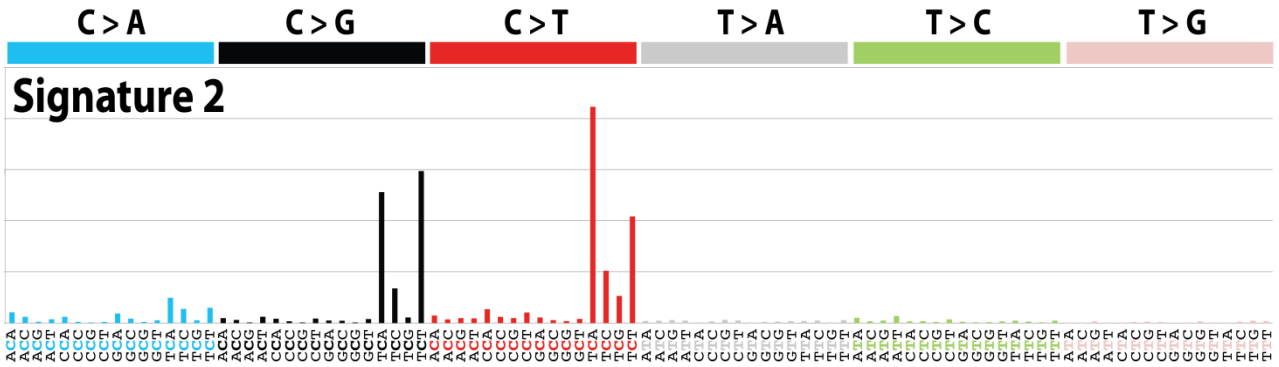
Which member(s) of the family is responsible for Signature 2?



What switches on intense APOBEC activity?

- Mutation or amplification of the APOBEC genes?
- Overexpression of APOBECs?
- Response to viral entry, retrotransposon movement or inflammation?

Is Signature 2 collateral damage on the cellular genome of APOBEC activation due to foreign DNA?



Known or speculative causes of mutational signatures

Deamination of 5-methyl cytosine, age

Known or speculative causes of mutational signatures

Tobacco

Known or speculative causes of mutational signatures

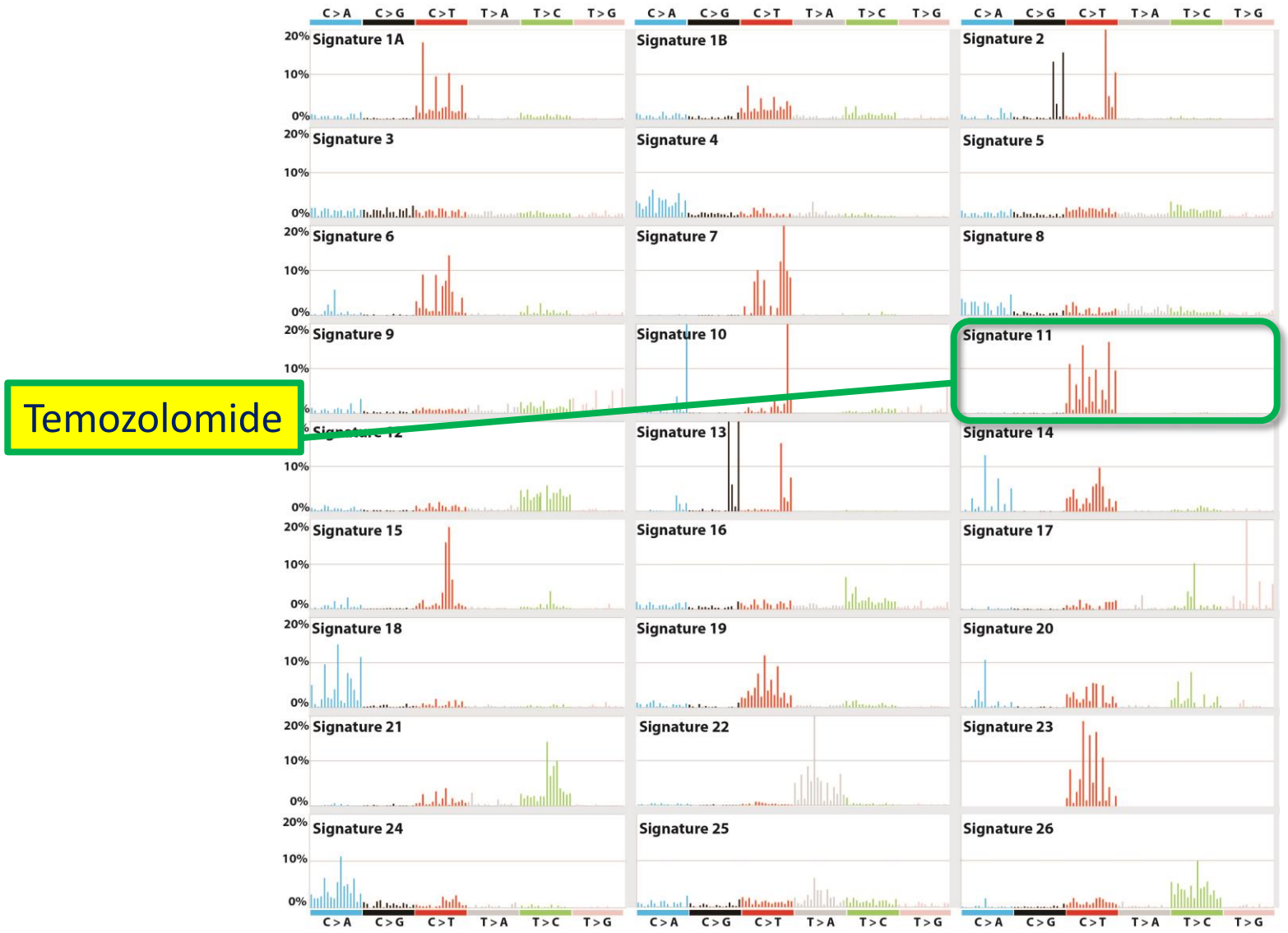
Known or speculative causes of mutational signatures

Aflatoxin

Known or speculative causes of mutational signatures

Aristolochic acid

Known or speculative causes of mutational signatures



Known or speculative causes of mutational signatures

Defective DNA mismatch repair

Known or speculative causes of mutational signatures

Defective BRCA1, BRCA2, homologous recombination repair

Known or speculative causes of mutational signatures

Defective polymerase epsilon activity

Known or speculative causes of mutational signatures

APOBEC cytosine deamination

Known or speculative causes of mutational signatures

Elucidating the mutational processes underlying mutational signatures

Association of a signature with other features of the cancer:

- epidemiology
- gene expression
- mutated genes

Elucidating the mutational processes underlying mutational signatures

Association of a signature with other features of the cancer:

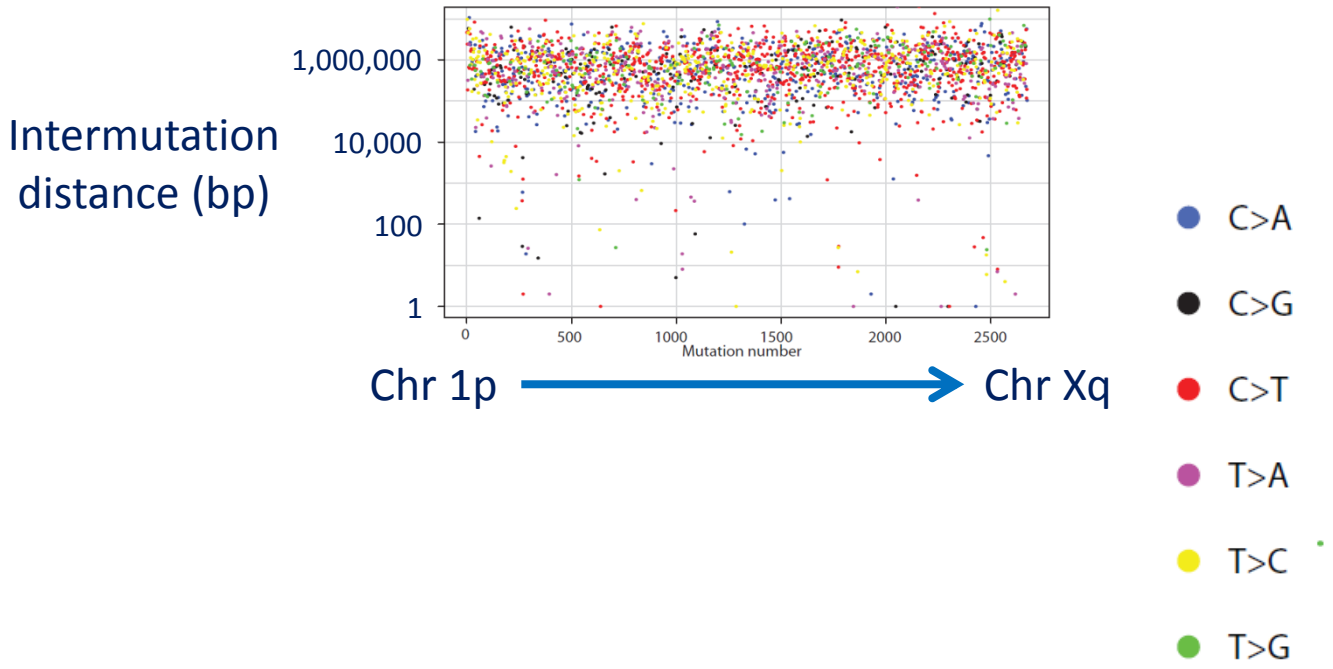
- epidemiology
- gene expression
- mutated genes

Acquiring a compendium of signatures from known exposures in experimental systems:

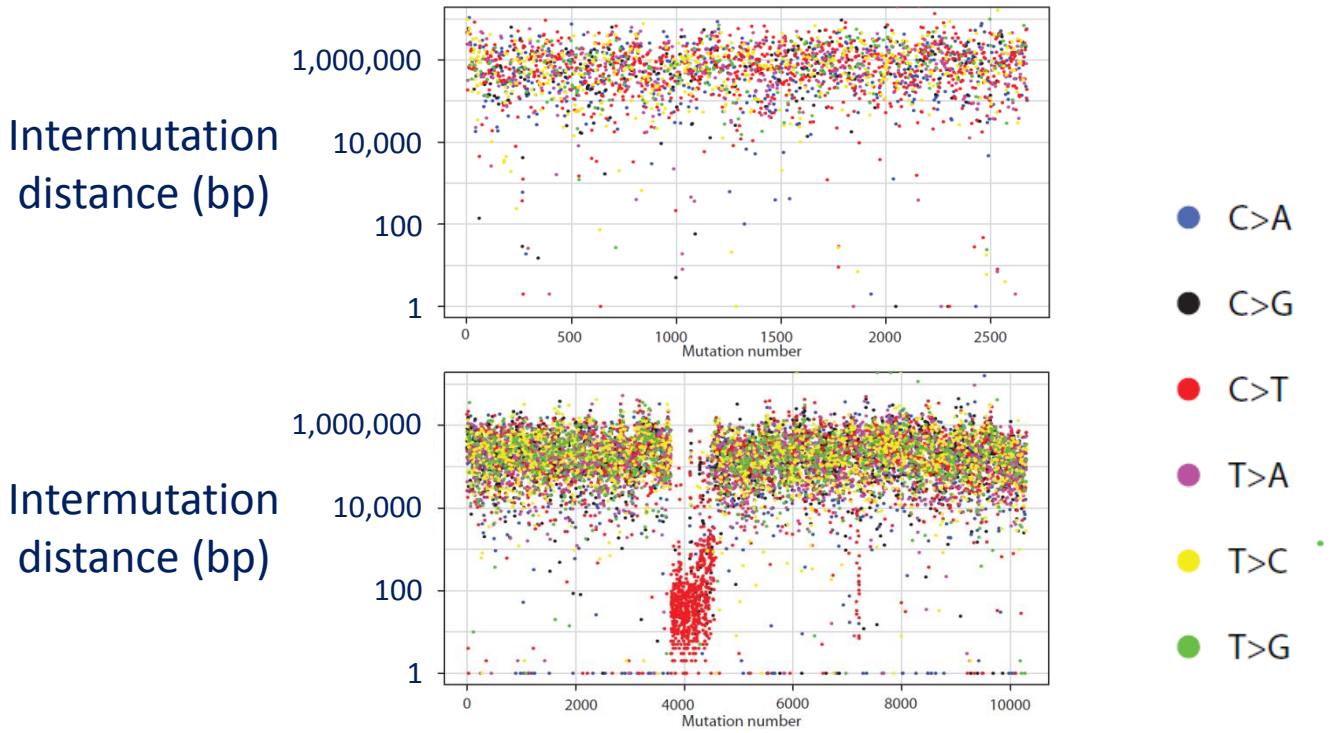
- mutagens
- defective DNA maintenance

Localised hypermutation: *kataegis*

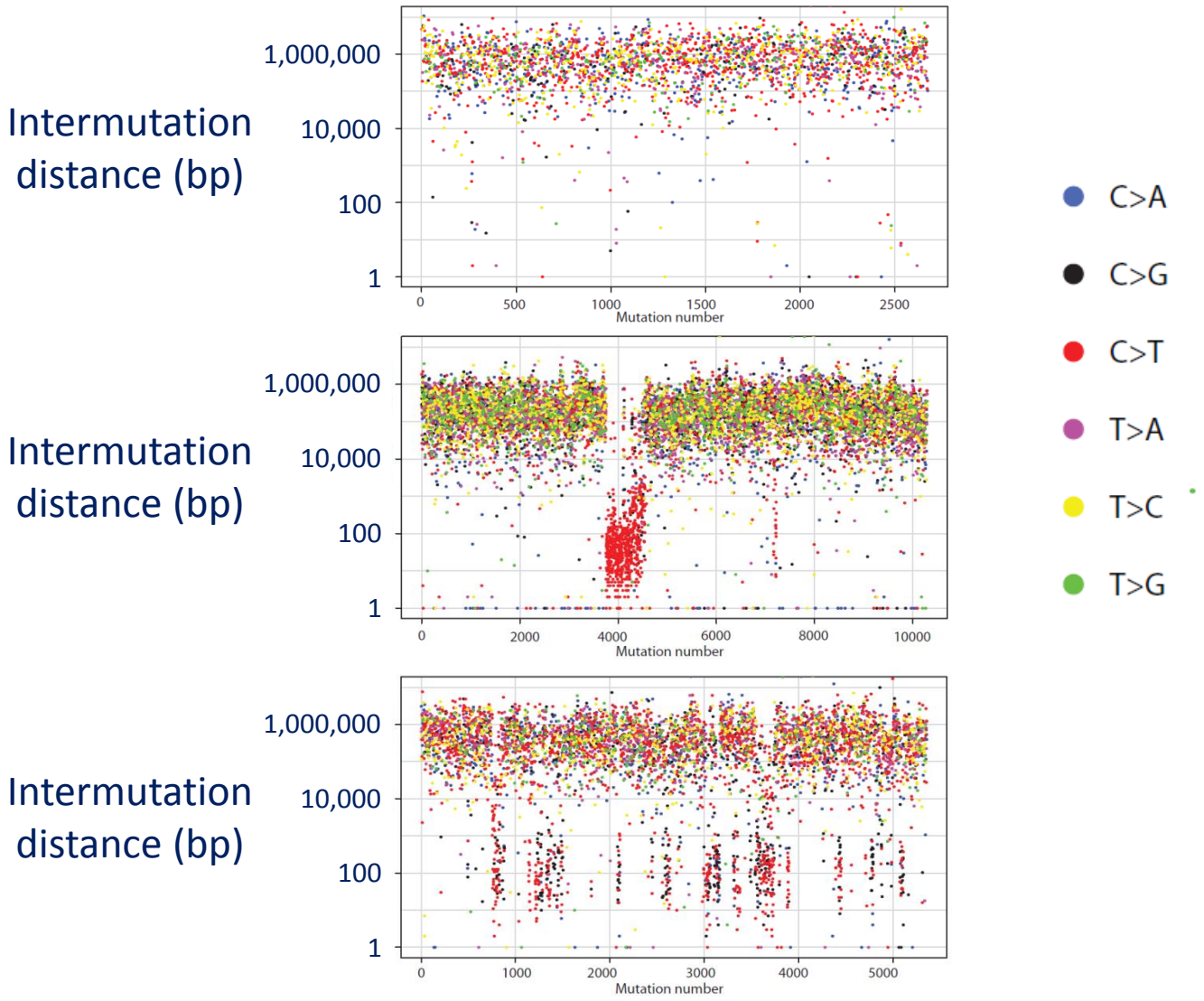
Foci of substitution hypermutation, *kataegis,* occur in cancer genomes



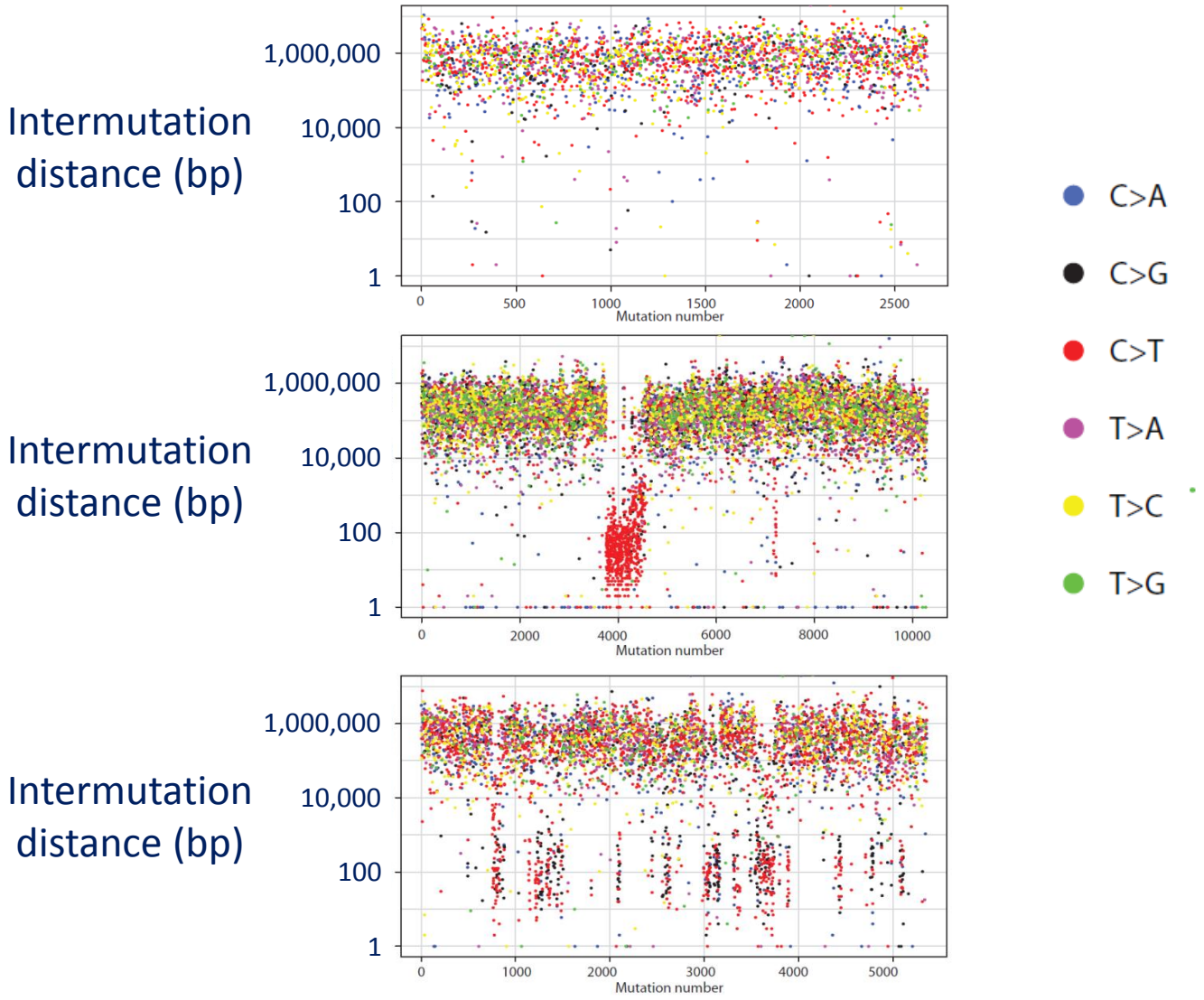
Foci of substitution hypermutation, *kataegis,* occur in cancer genomes



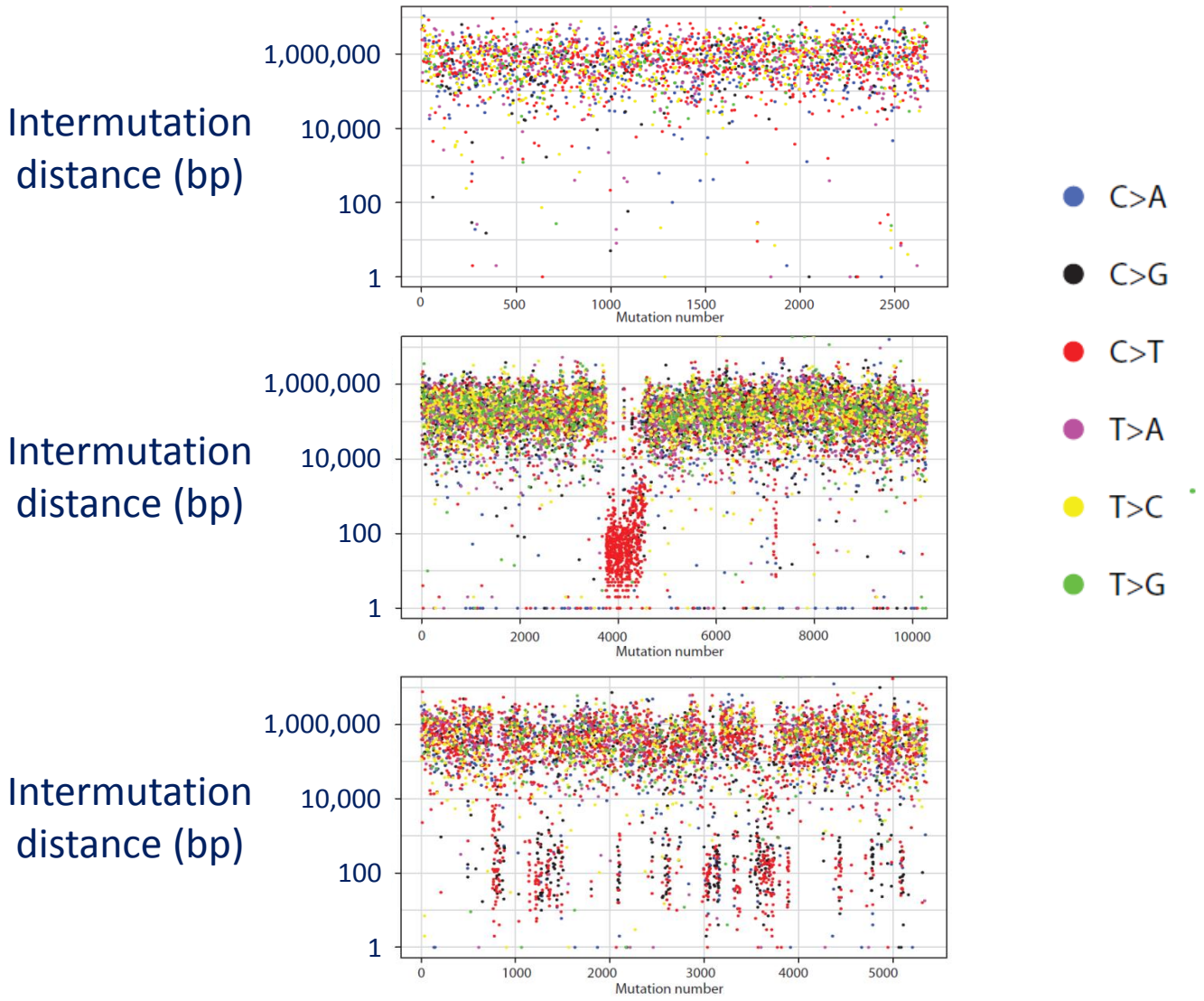
Foci of substitution hypermutation, *kataegis,* occur in cancer genomes



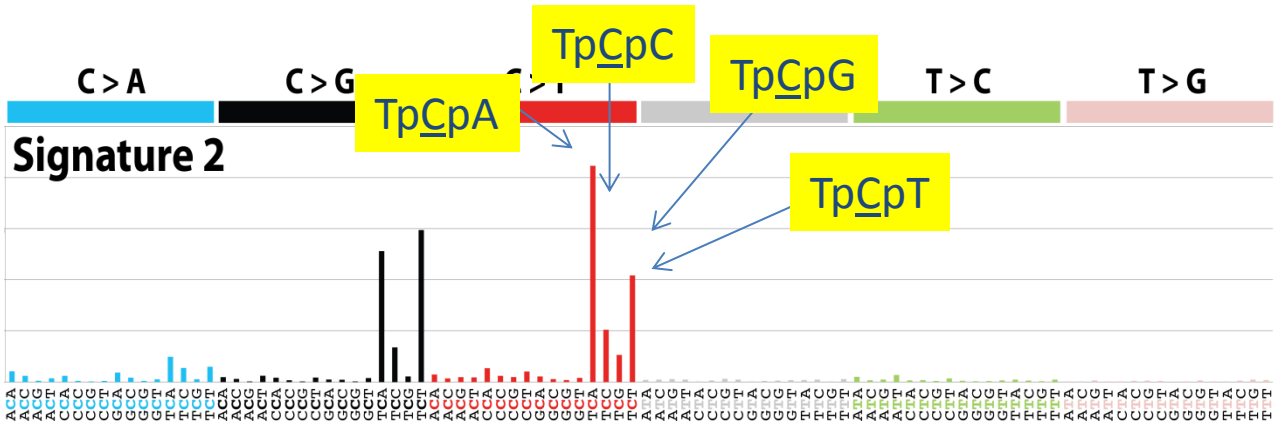
Mutations in regions of *kataegis* are almost all C>T or C>G



Mutations in regions of *kataegis* are almost all at TpCpN trinucleotides



Mutations in regions of *kataegis* are almost all at TpCpN trinucleotides



Are APOBECs responsible for *kataegis*?

Are APOBECs responsible for *kataegis*?

- Overexpression of APOBECs in yeast generates *kataegis*

Are APOBECs responsible for *kataegis*?

- Overexpression of APOBECs in yeast generates *kataegis*

What directs an APOBEC to a particular part of the genome to generate *kataegis*?

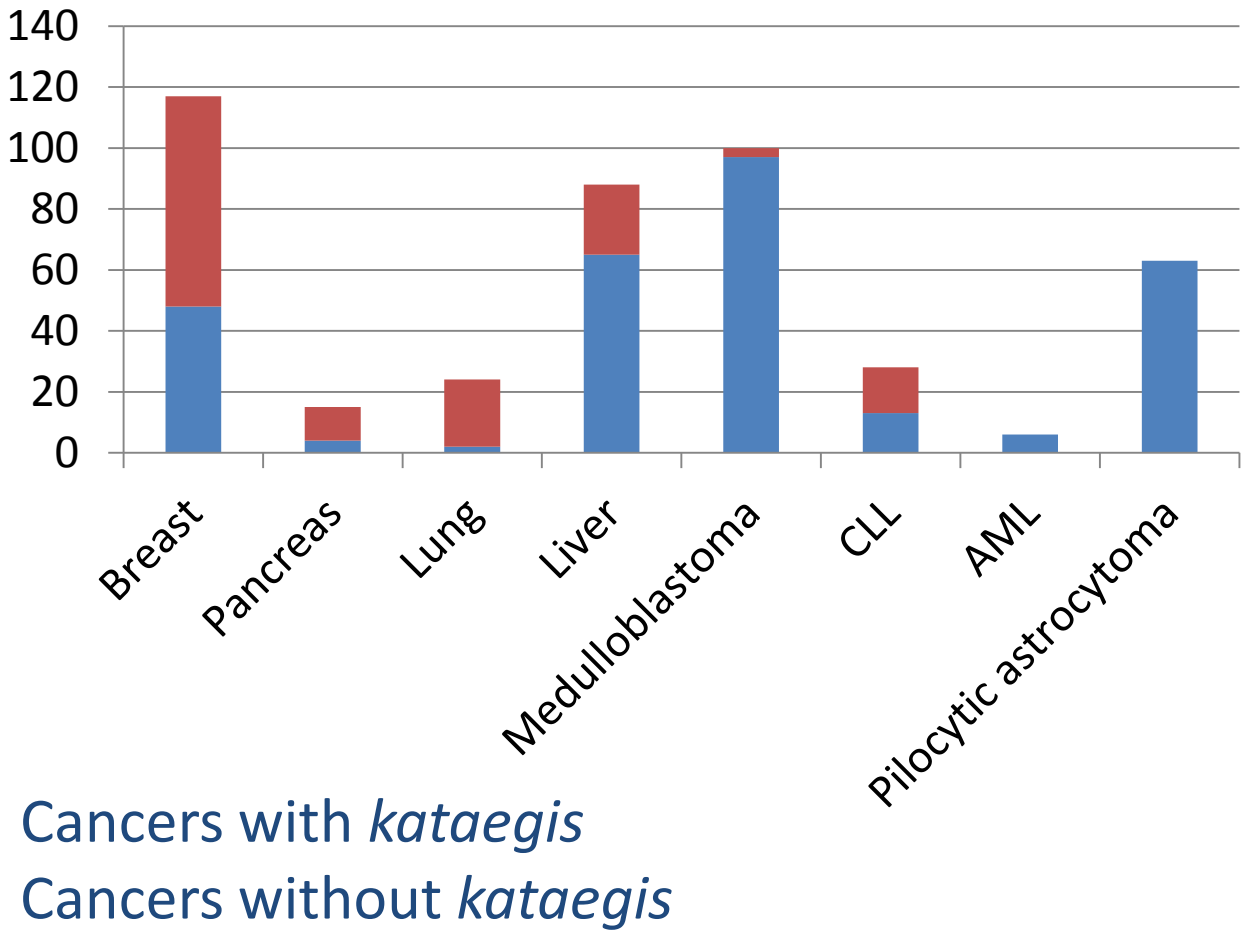
Are APOBECs responsible for *kataegis*?

- Overexpression of APOBECs in yeast generates *kataegis*

What directs an APOBEC to a particular part of the genome to generate *kataegis*?

- Introduction of a double strand break in yeast markedly increases the frequency of *kataegis* in the vicinity

Kataegis is present in many cancer types



Summary

- There are more than 20 genome-wide mutational signatures across human cancer
- A signature likely representing cytosine deamination due to APOBEC activity is common
- *Kataegis*, localised hypermutation, is found in many cancers
- Some signatures are due to known mutagenic exposures or known defects in DNA maintenance but the cause of many is unknown
- Understanding of the mutational processes underlying these signatures will inform on cancer causation, prevention and treatment

Serena Nik-Zainal

Ludmil Alexandrov

BREAST CANCER WORKING GROUP

Sam Aparicio
Alan Ashworth
Ake Borg
Anne-Lise Borresen-Dale
Carlos Caldas
Doug Easton
Diana Eccles
Ian Ellis
Jorunn Eyfjord
John Foekens
Louise Jones
Jocelyne Jacquemier
Jorge Reis-Filho
Sunil Lakhani
Mike Lee
Larry Norton

Angelo Paradiso
Martine Piccart
Jorge Reis-Filho
Andrea Richardson
Anne Salomon
Christos Sotiriou
Paul Spellman
Henk Stunnenberg
Fred Sweep
Benita Tan
Gilles Thomas
Andy Tutt
Laura Van t' Veer
Marc Van de Vijver

Ludmil Alexandrov
Serena Nik-Zainal
Manasa Ramakrishna
Helen Davies
David Wedge
Andy Futreal
Peter Campbell

Ben Taylor
Christina Rada
Michael Neuberger

Sean Grimmond
Reiner Siebert
Stefan Pfister
Peter Lichter
Elias Campo
Carlos Lopez-Ortin
Xose Puente
Tatsuhiro Shibata

International Cancer Genome Consortium

THE CANCER GENOME ATLAS

wellcome trust sanger institute