Somatic mutant clones colonize the human esophagus with age

Martincorena I.[1] Fowler JC.[1] *et al.* [1 2 3 4 5]

[1]Wellcome Sanger Institute, Hinxton, Cambridge, UK [2]MRC Unit Cancer unit, Hutchinson-MRC Research centre, University of Cambridge, UK [3]Dept Surgery and Cambridge NIHR Biomedical Research Centre, Biomedical Campus, University of Cambridge, Cambridge, UK [4]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, UK [5]Department of Haematology, University of Cambridge, Cambridge, UK

Declan Bennett

National University of Ireland, Galway, Ireland

November 8, 2018

Table of contents

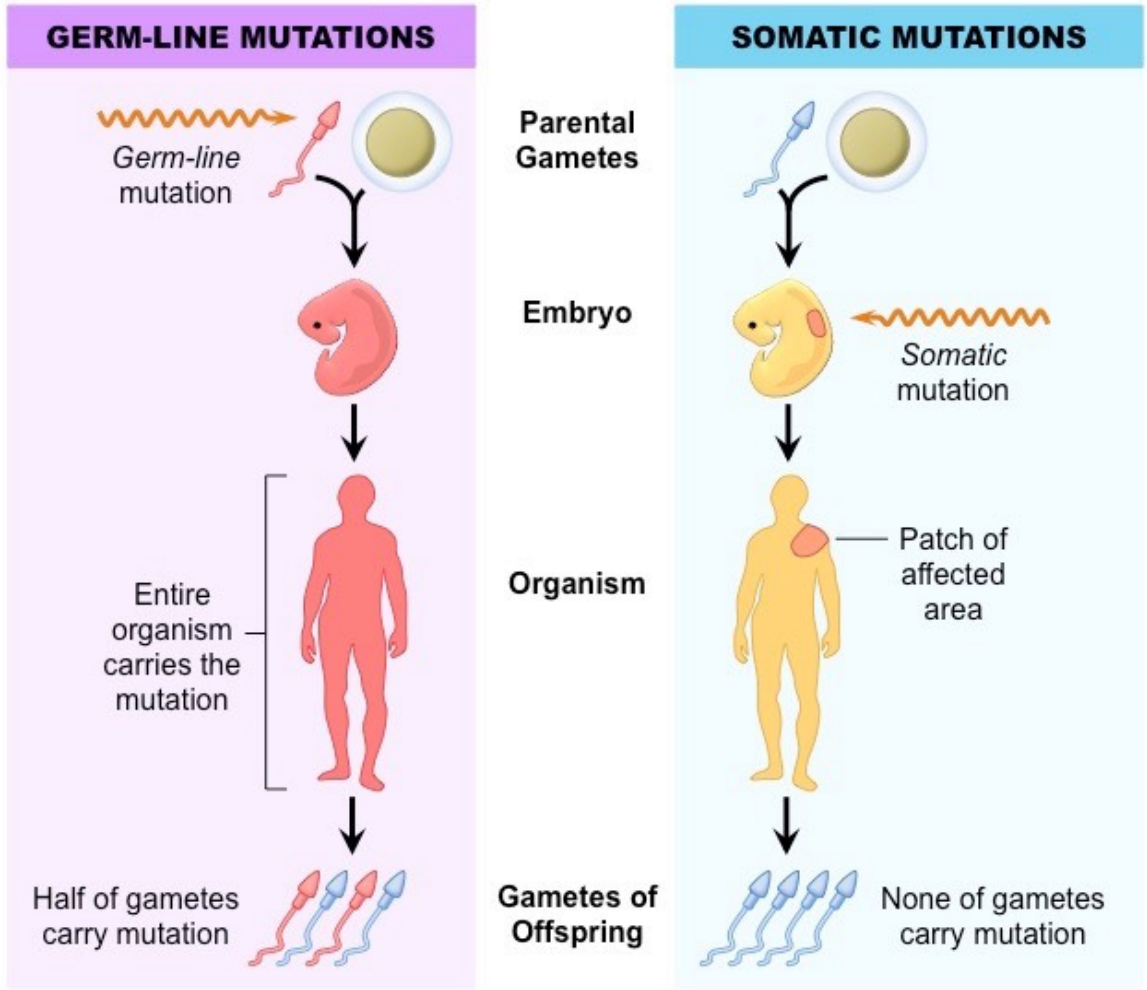
Overview of somatic mutation

What is a Somatic mutation?

Overview of somatic mutation

What is a Somatic mutation?

- A Somatic mutation is the collective term for mutations that arise in non-germline cells

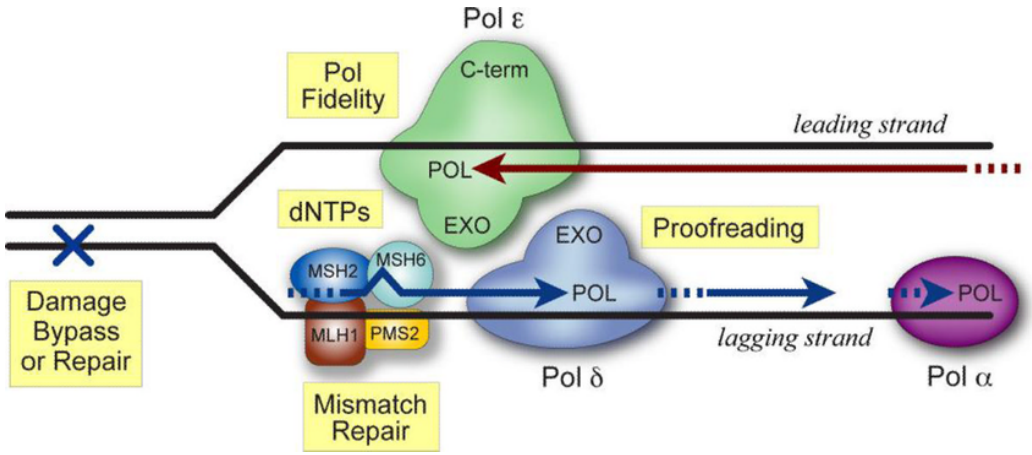


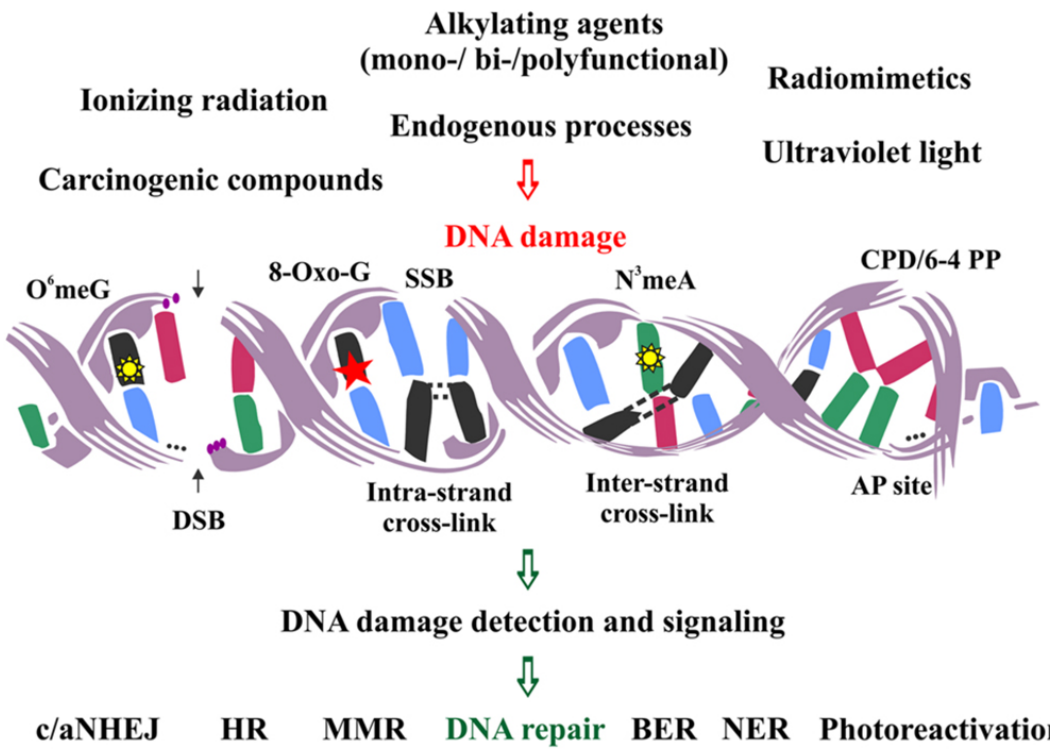
Overview of somatic mutation

How do somatic mutation arise?

Overview of somatic mutation

How do somatic mutation arise?





Overview of somatic mutation

Why study somatic mutation?

Overview of somatic mutation

Why study somatic mutation?
Somatic mutation is the driving process behind cancer as well as the ageing process

Overview of somatic mutation

Why study somatic mutation?
Somatic mutation is the driving process behind cancer as well as the ageing process

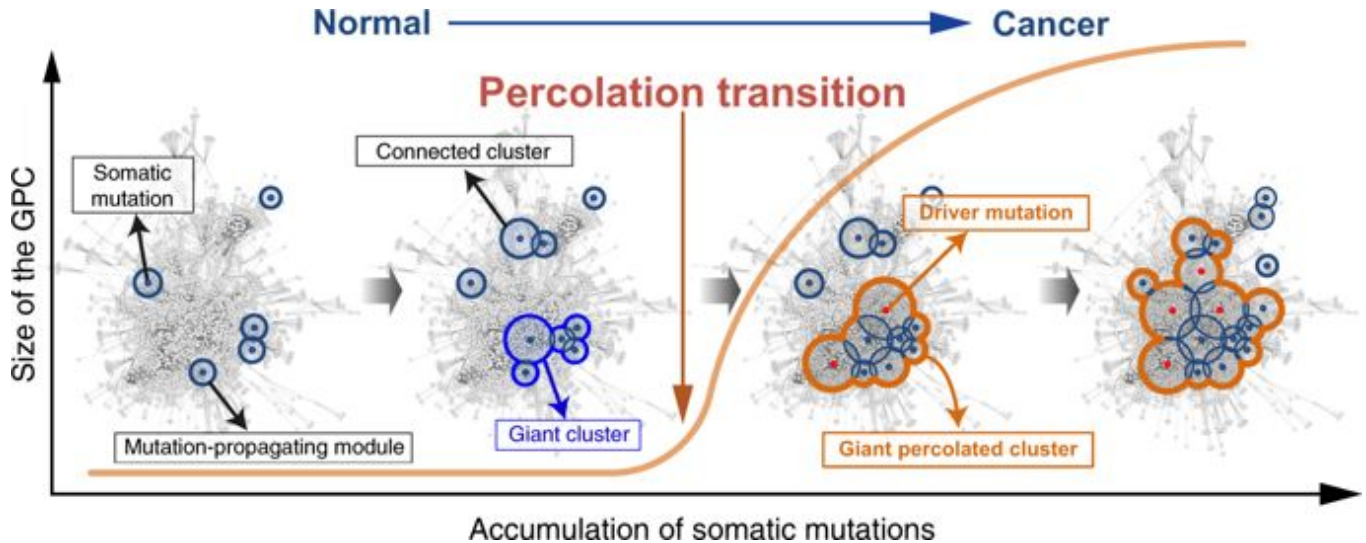


Figure: Shin *et al.* 2017

Overview of somatic mutation

Somatic mutation is intrinsic in the ageing process

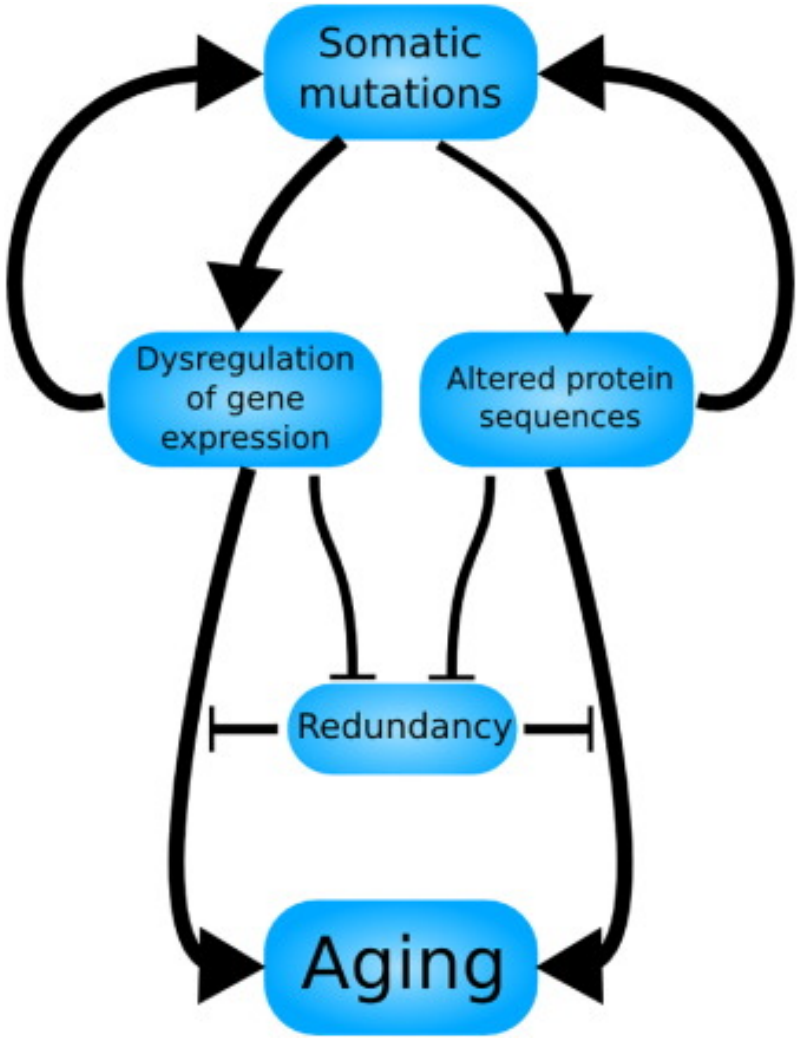
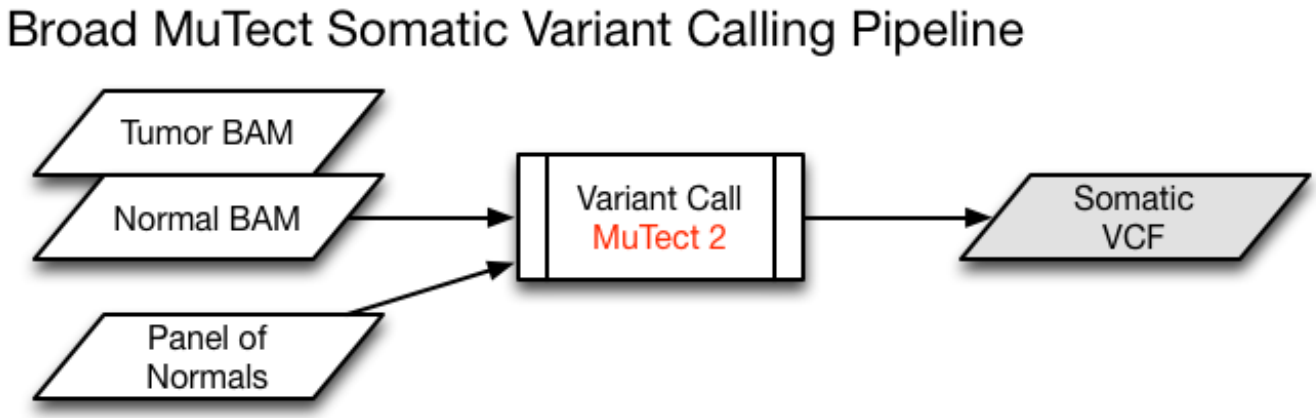


Figure: Mulholland *et al.* 2017

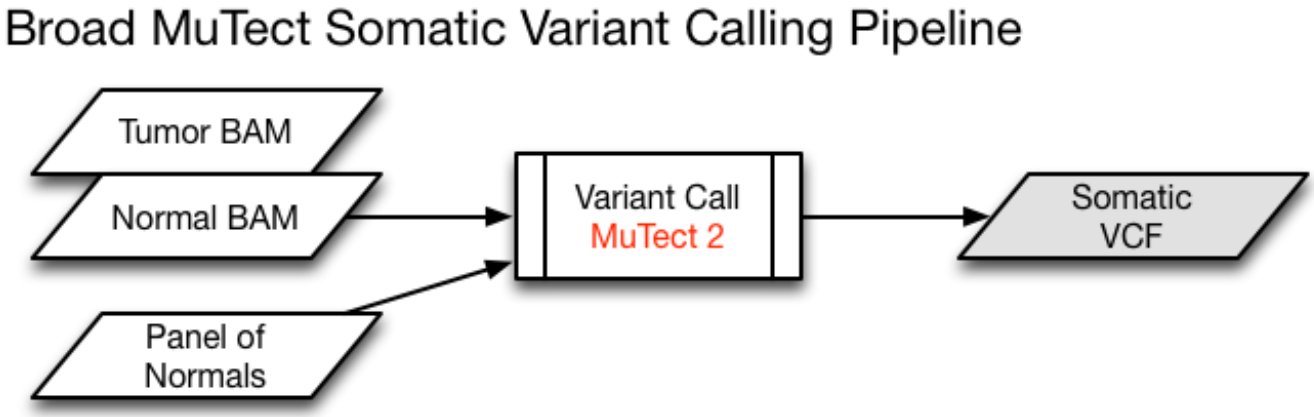
Variant Calling

In cancer somatic mutations are called by comparing against the healthy tissue of the same sample 'Tumour-Normal matched pair'



Variant Calling

In cancer somatic mutations are called by comparing against the healthy tissue of the same sample 'Tumour-Normal matched pair'



Detecting somatic variants in healthy samples?

ShearwaterML algorithm

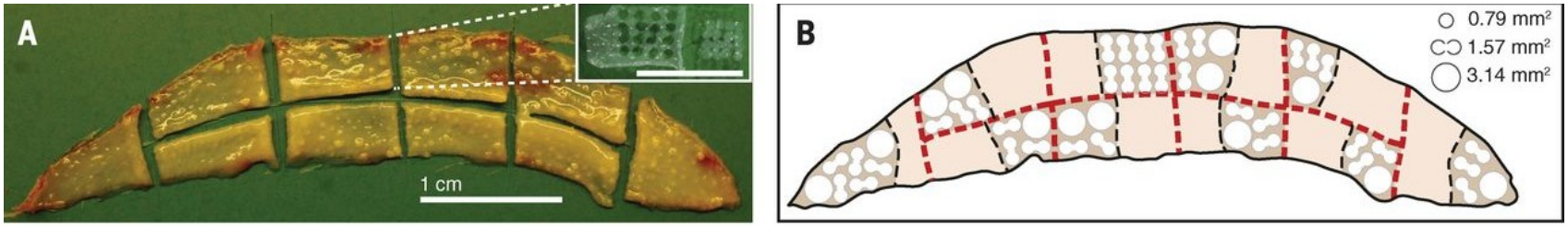


Figure: Martincorena *et al.* 2015

In deep sequenced biopsies a variation of the Shearwater algorithm can be applied.
This uses multiple matched normal samples from the same individual to create an error model per nucleotide increasing sequencing coverage by a factor $n_\jmath = \sum_{\iota=1}^{N} n_{\iota\jmath}$

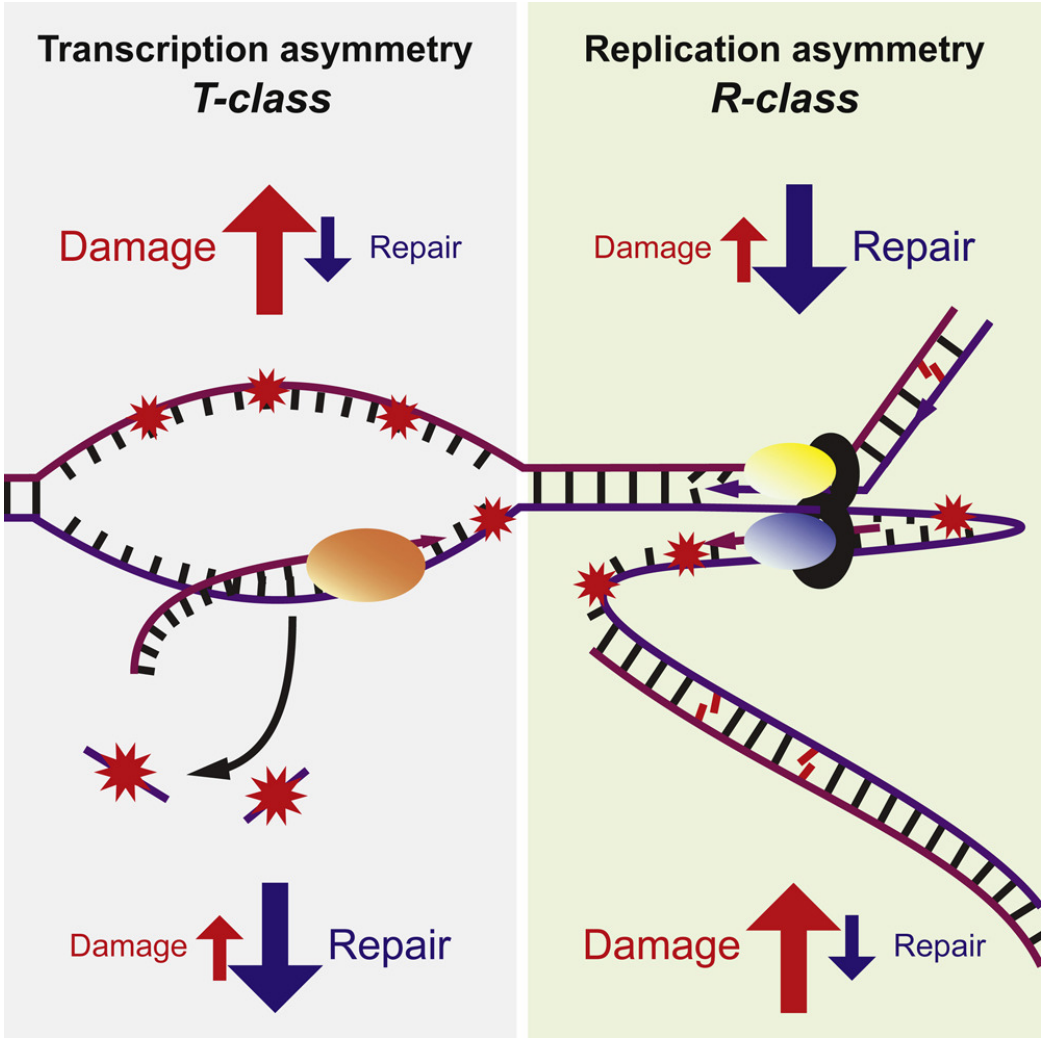
$$\mathcal{L}(H_0)=\mathcal{L}(\nu_{\iota\jmath\kappa},\rho \mid X_{\iota\jmath\kappa},n_{\iota\jmath})*\mathcal{L}(\nu'_{\iota\jmath\kappa},\, \rho \mid X'_{\iota\jmath\kappa},\, n'_{\iota\jmath})*\mathcal{L}(\nu_{\jmath\kappa},\, \rho \mid x_{\jmath\kappa},\, n_\jmath)*\mathcal{L}(\nu'_{\jmath\kappa},\, \rho \mid x'_{\jmath\kappa},\, n'_\jmath)$$

$$\mathcal{L}(H_1)=\mathcal{L}(\nu_{\iota\jmath\kappa},\rho \mid X_{\iota\jmath\kappa},n_{\iota\jmath})*\mathcal{L}(\nu'_{\iota\jmath\kappa},\, \rho \mid X'_{\iota\jmath\kappa},\, n'_{\iota\jmath})*\mathcal{L}(\mu_{\jmath\kappa},\, \rho \mid x_{\jmath\kappa},\, n_\jmath)*\mathcal{L}(\mu'_{\jmath\kappa},\, \rho \mid x'_{\jmath\kappa},\, n'_\jmath)$$

ShearwaterML algorithm

High coverage allows for exclusion of SNVs and sequencing artefacts such as PCR errors and DNA damage

Orientation with respect to the transcribed strand



Nucleotide context

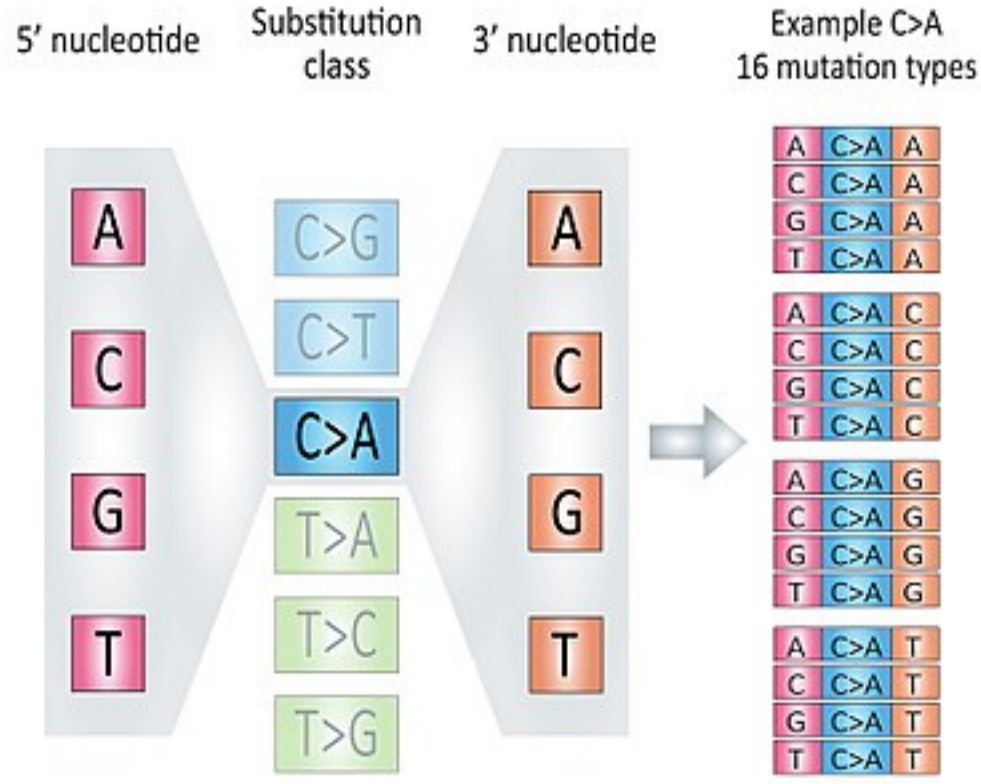


Figure: Alexandrov *et al.* 2018

96 context model

When aligned with the the transcribed strand this model can be extended to 192 contexts

Enrichement of Mutational patterns/signatures

Mutational patterns can be used to identify molecular mutational signatures e.g. Tobacco smoke signature

There are 30 cancer related signatures.

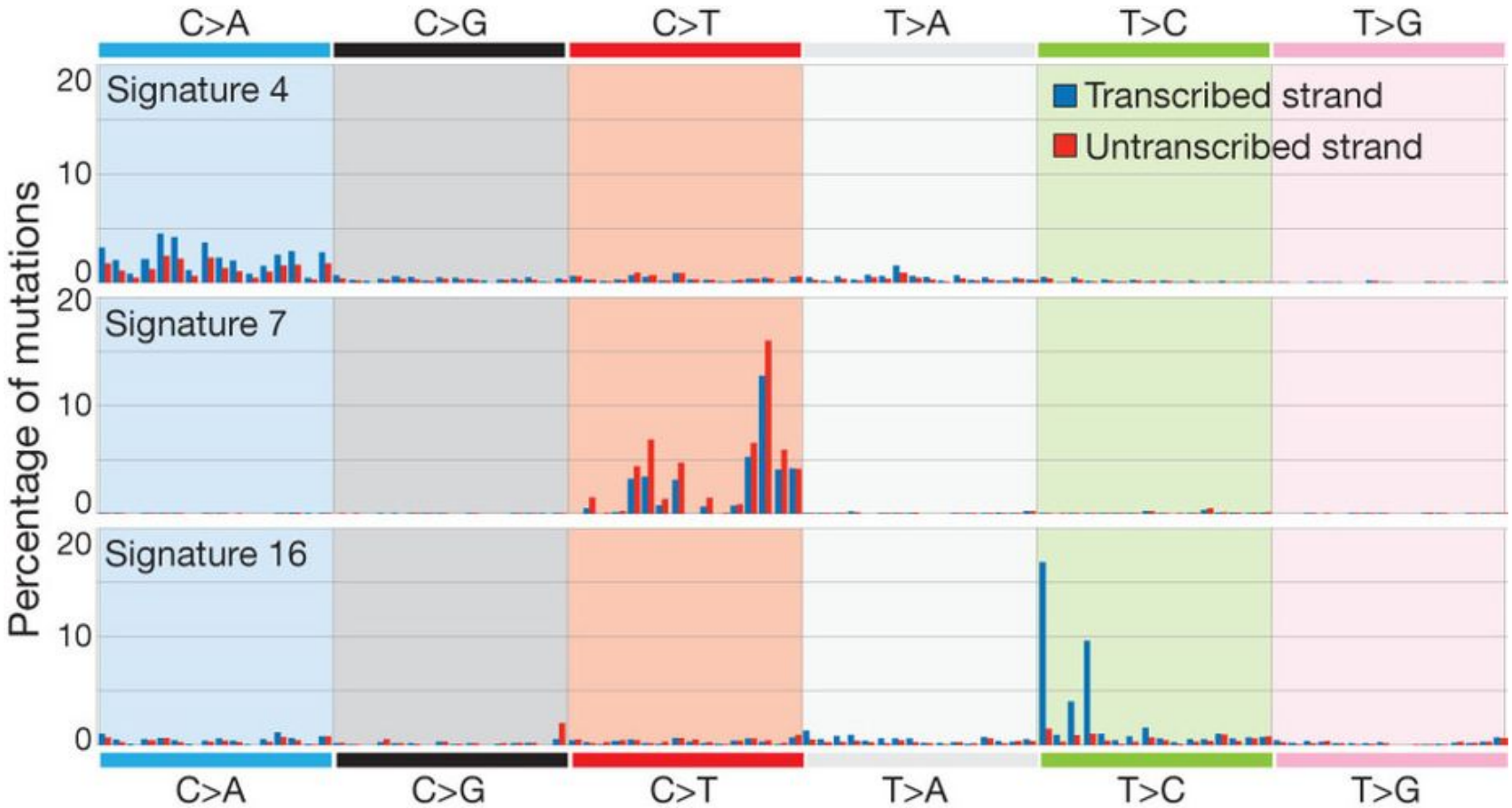
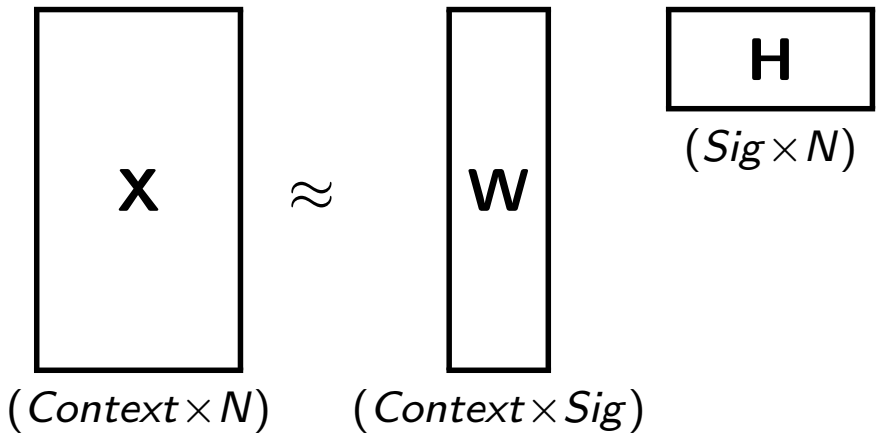


Figure: Alexandrov *et al.* 2013

NMF

Nonnegative matrix factorization



Here X is a counts matrix of the 192 contexts per sample.

NMF

Choosing the correct number of signatures

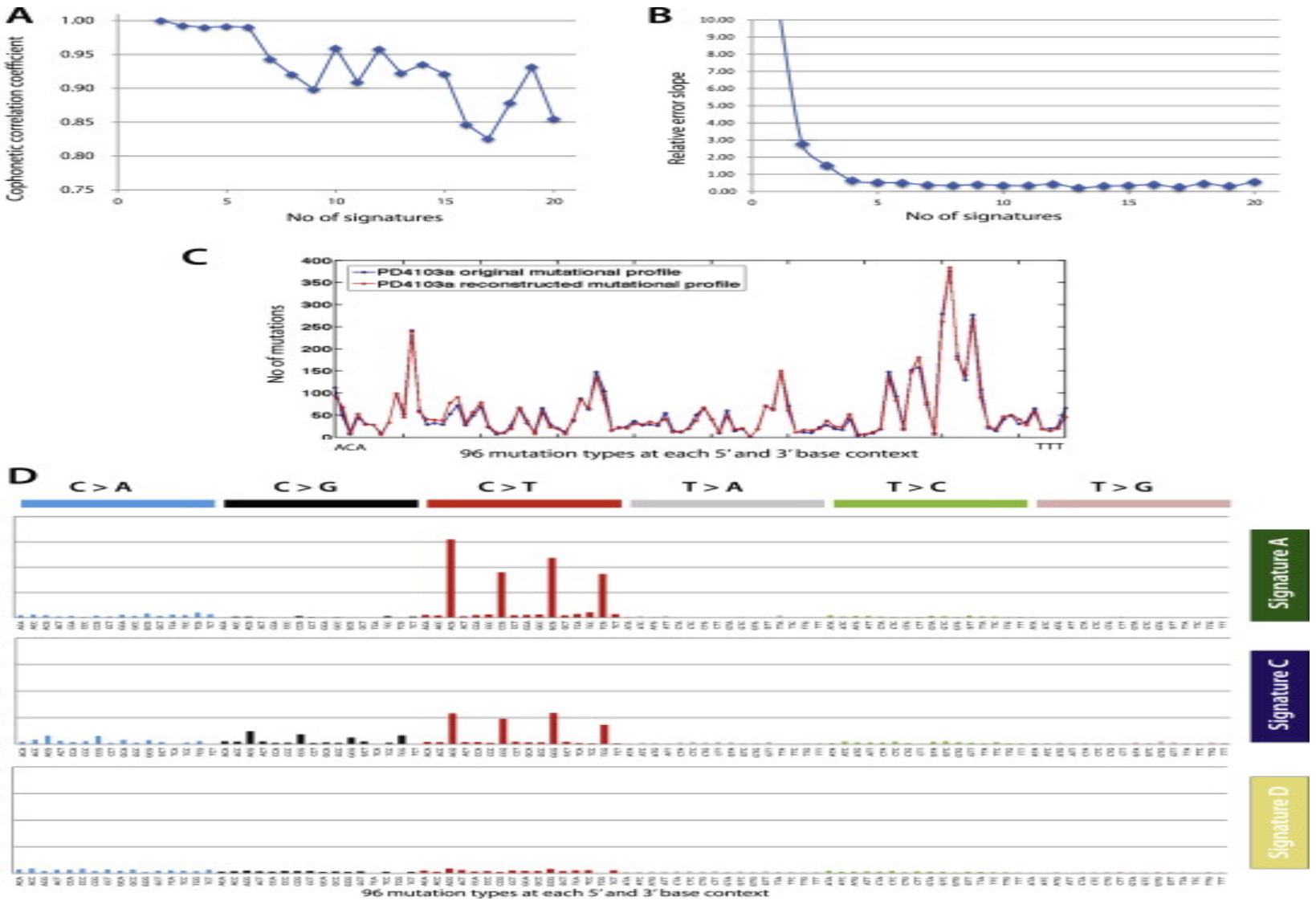


Figure: Nik-Zainal *et al.* 2012

Defining regions under selective constraints

Positive selection

Is the process of how advantageous alleles sweep a population (organisms/cells)

Negative/Purifying selection

Is the process of how disadvantageous alleles are removed a population

Neutral selection

A loci is said to be evolving neutrally if there is a lack of either positive or negative selection

How do we measure selection

Assumption

In coding regions, somatic mutations effect the three nucleotide codon resulting in no change to the amino acid, encoding a new a new amino acid or signalling for premature termination of translation. These can be grouped into 2 categories; synonymous, k_s and non-synonymous, k_a.

We can infer the presence of positive,negative or neutral selection using the ratio of non-synonymous to synonymous mutation

$$\omega = \frac{k_a}{k_s}$$

$$\omega \begin{cases} > 1 & \text{loci is under positive selection} \\ = 1 & \text{loci is evolving neutrally} \\ < 1 & \text{loci is under negative selection} \end{cases}$$

Somatic mutation in healthy skin cells

RESEARCH ARTICLES

TUMOR EVOLUTION

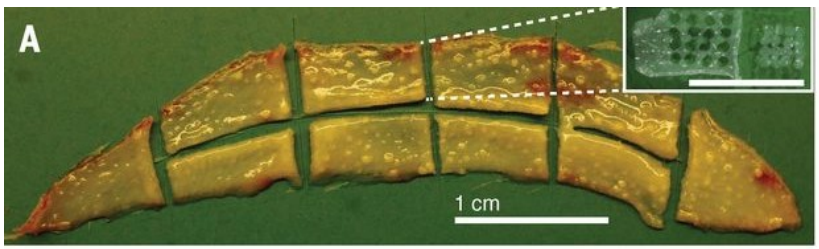
High burden and pervasive positive selection of somatic mutations in normal human skin

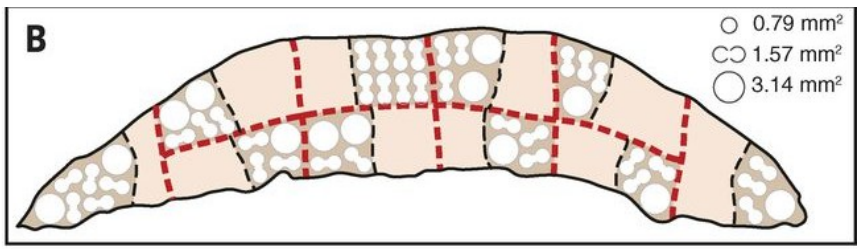
Iñigo Martincorena,[1] Amit Roshan,[2] Moritz Gerstung,[1] Peter Ellis,[1] Peter Van Loo,[1,3,4] Stuart McLaren,[1] David C. Wedge,[1] Anthony Fullam,[1] Ludmil B. Alexandrov,[1] Jose M. Tubio,[1] Lucy Stebbings,[1] Andrew Menzies,[1] Sara Widaa,[1] Michael R. Stratton,[1] Philip H. Jones,[2*] Peter J. Campbell[1,5*]

Somatic mutation in healthy skin cells

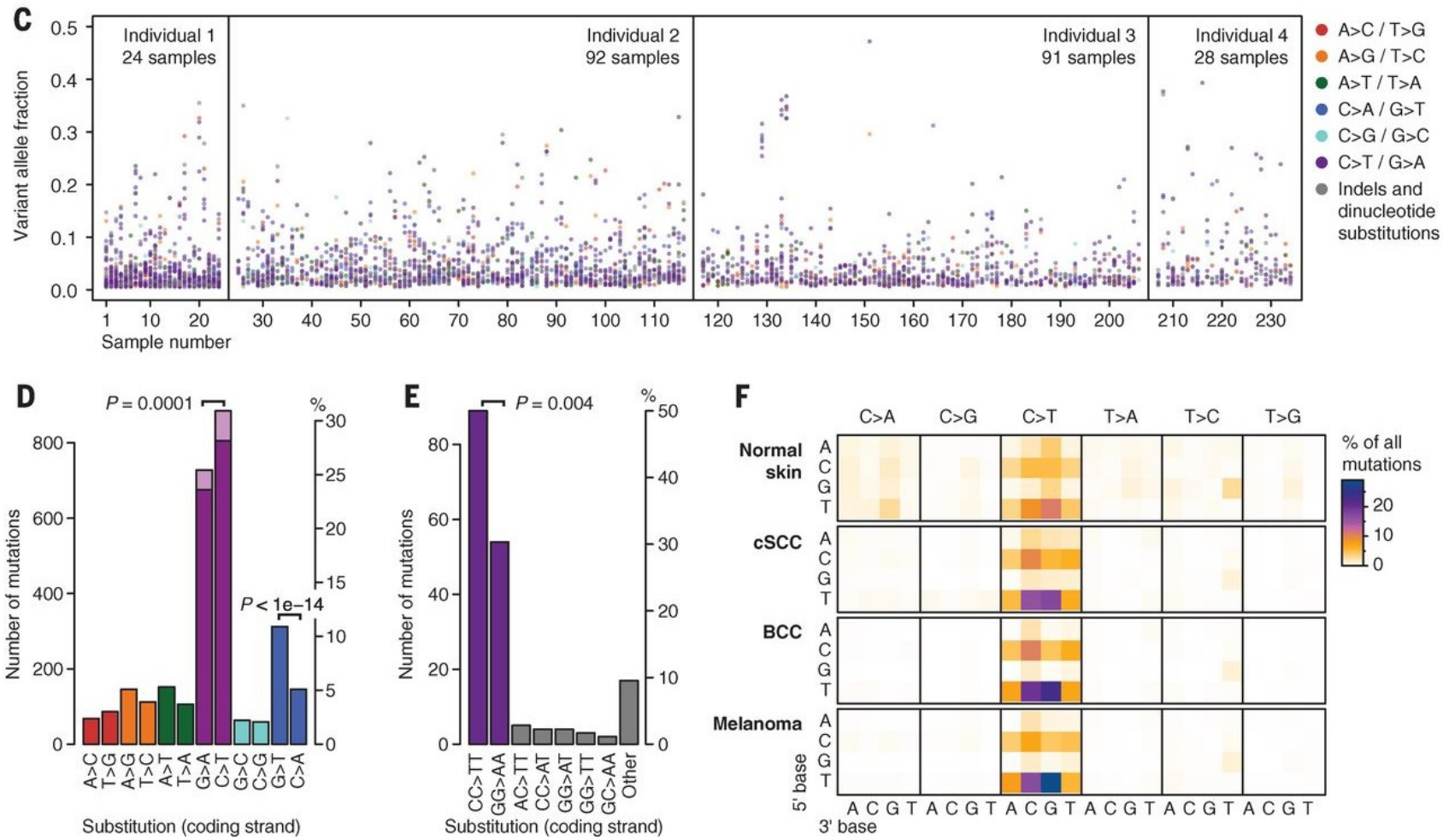
Experimental overview

- 234 samples from 4 healthy donor eyelids
- Ultra-deep sequencing of 74 cancer genes with an average depth of $\approx 500x$
- Modified Shearwater algorithm allowed for detection of mutants in 1% of biopsied cells

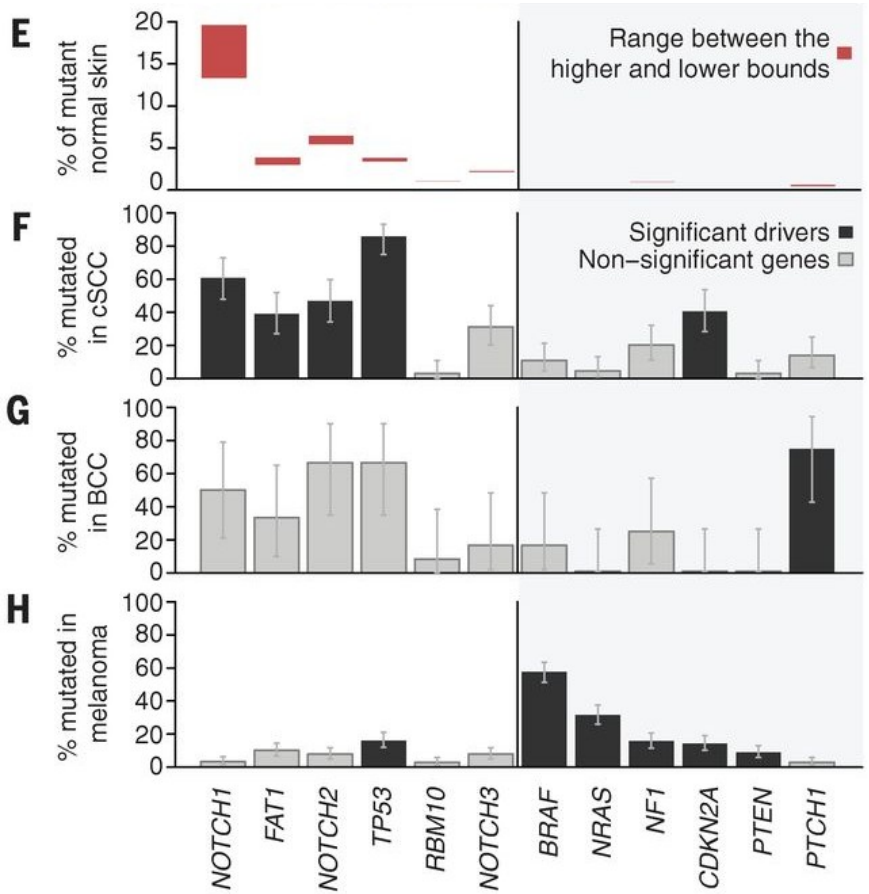


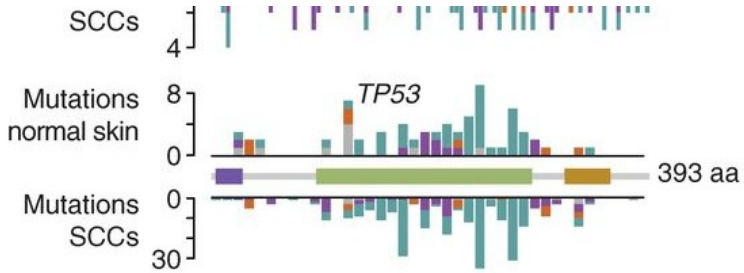


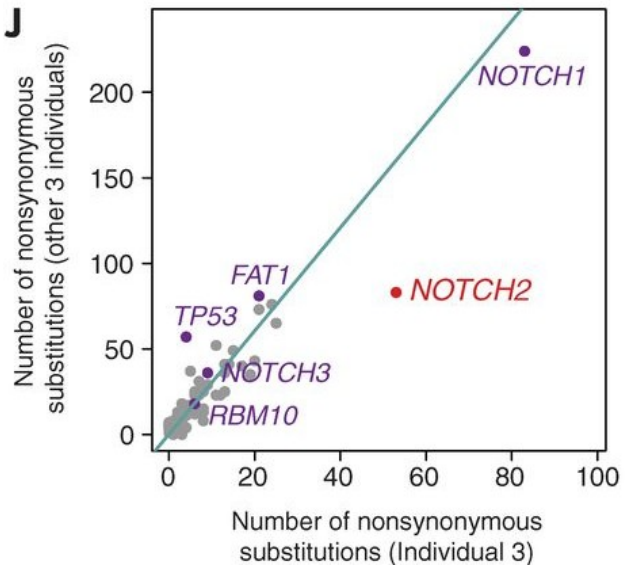
Somatic mutation in healthy skin cells



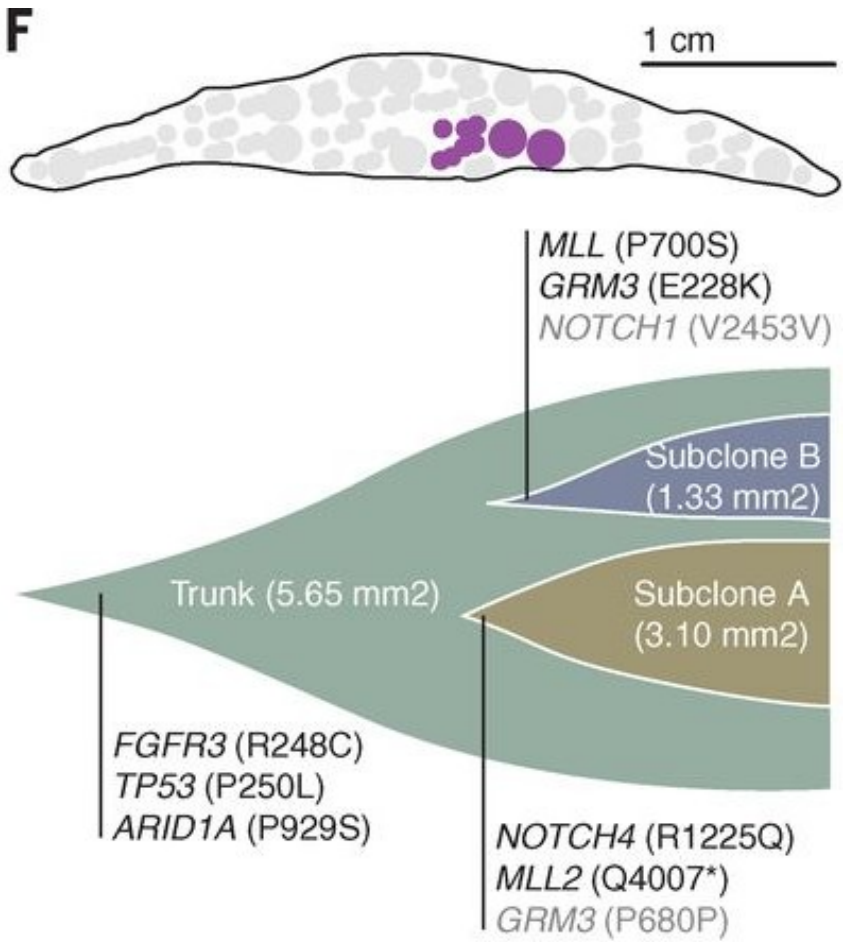
Somatic mutation in healthy skin cells

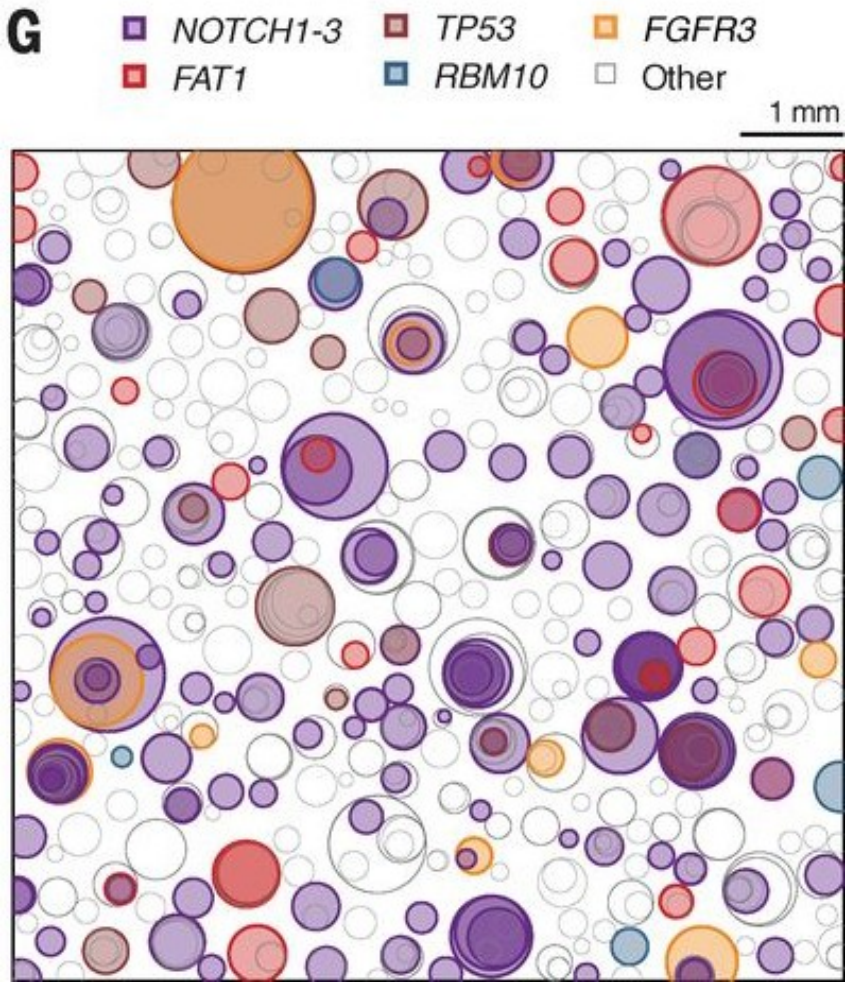






Somatic mutation in healthy skin cells





Somatic mutation in healthy skin cells

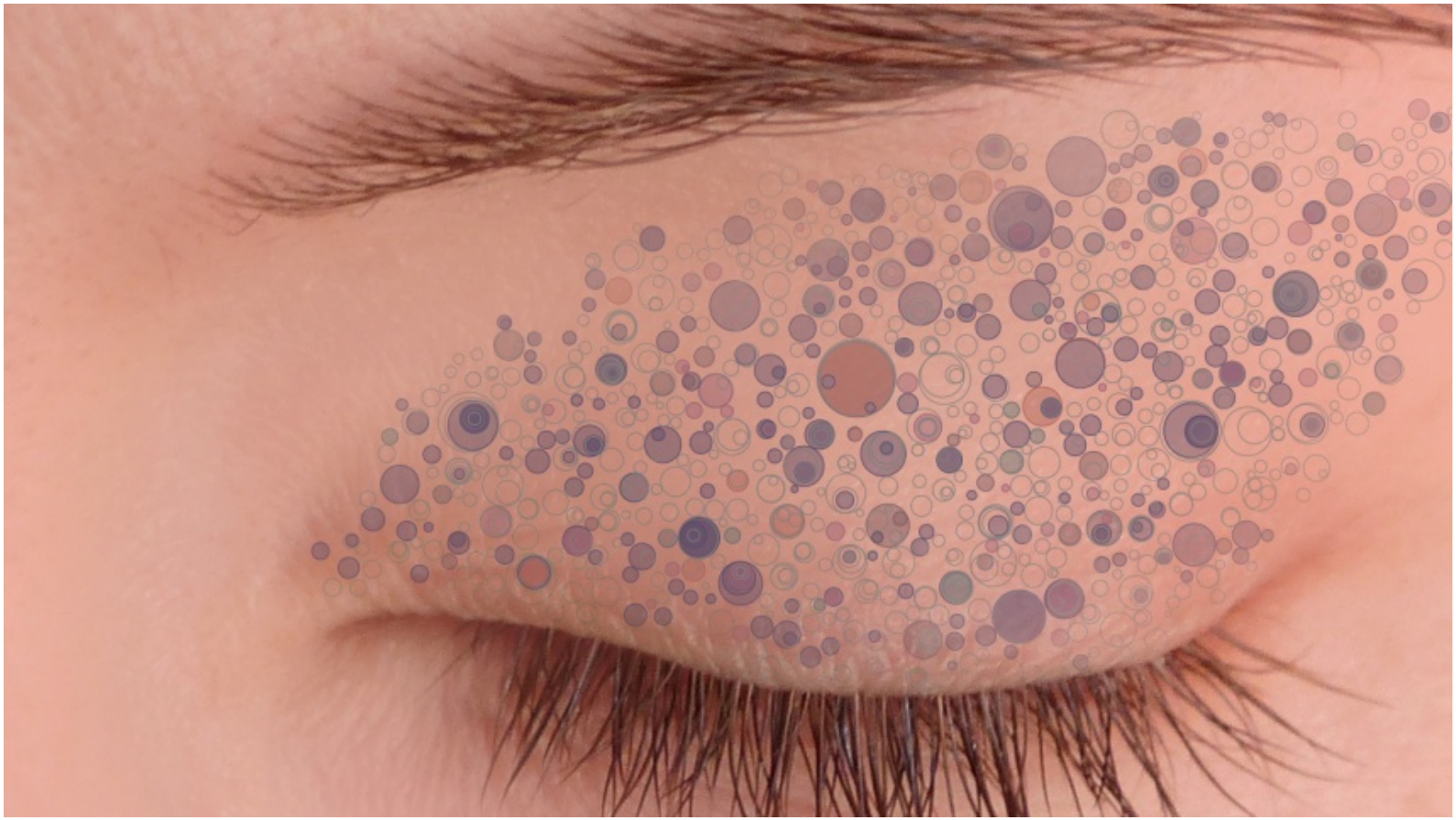


Figure: https://www.sanger.ac.uk/science/programmes/cancer-genetics-and-genomics

Somatic mutation in healthy skin cells

But skin has high UV exposure so is expected to have a higher mutational burden

Somatic mutation in healthy skin cells

But skin has high UV exposure so is expected to have a higher mutational burden

How does this experiment replicate in tissues with a lower rate mutations per cell?

Somatic mutant clones colonize the human esophagus with age

Science

RESEARCH ARTICLES

Cite as: I. Martincorena *et al.*, *Science*
10.1126/science.aau3879 (2018).

Somatic mutant clones colonize the human esophagus with age

Iñigo Martincorena[1*†], Joanna C. Fowler[1*], Agnieszka Wabik[1], Andrew R. J. Lawson[1], Federico Abascal[1], Michael W. J. Hall[1,2], Alex Cagan[1], Kasumi Murai[1], Krishnaa Mahbubani[3], Michael R. Stratton[1], Rebecca C. Fitzgerald[2], Penny A. Handford[4], Peter J. Campbell[1,5], Kourosh Saeb-Parsy[3], Philip H. Jones[1†]

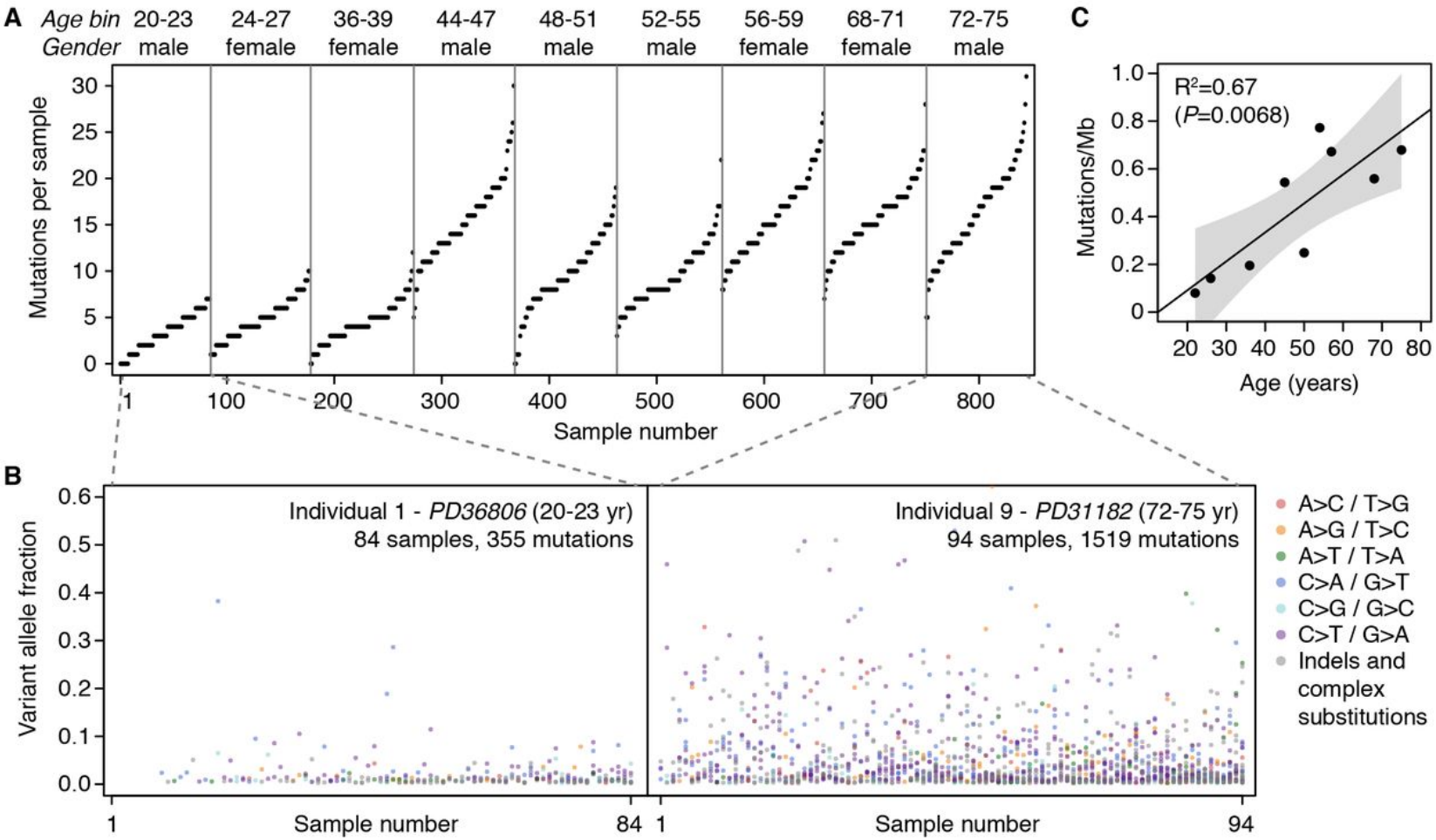
'These findings have implications for our understanding of cancer and ageing'

Somatic mutant clones colonize the human esophagus with age

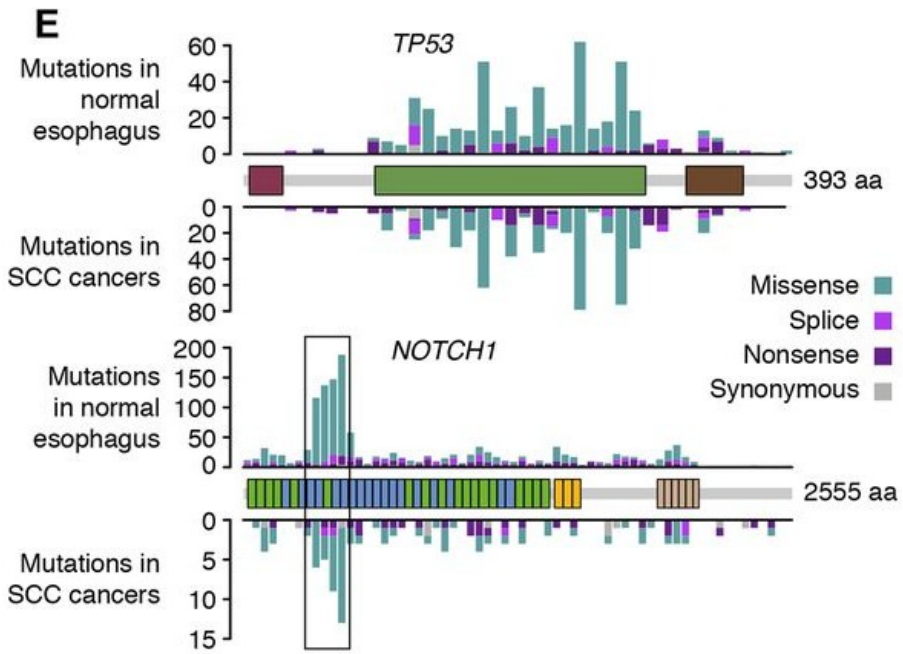
Experimental overview

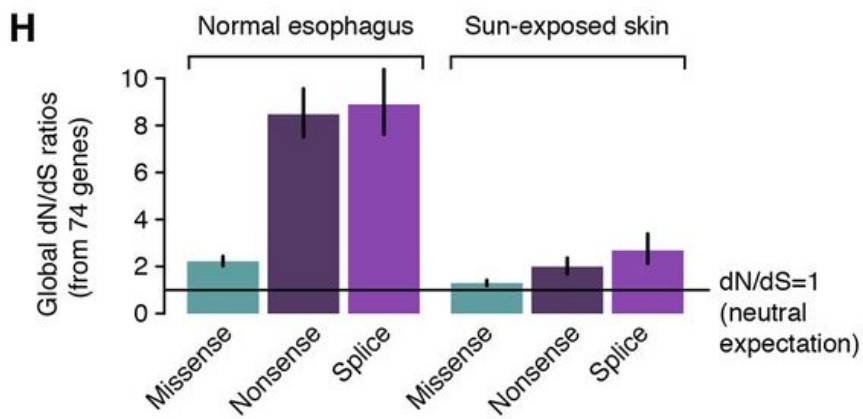
- 844 samples from 9 deceased donors (Age=20-75)
- Ultra-deep sequencing of 74 cancer genes with an median depth of $\approx 870x$
- Modified Shearwater algorithm for detection of mutants with median frequency of 1.6% with $\frac{1}{3}$ of mutations below 1%
- 21 samples dominated by large clones selected for WGS median $\approx 37x$

Somatic mutant clones colonize the human esophagus with age

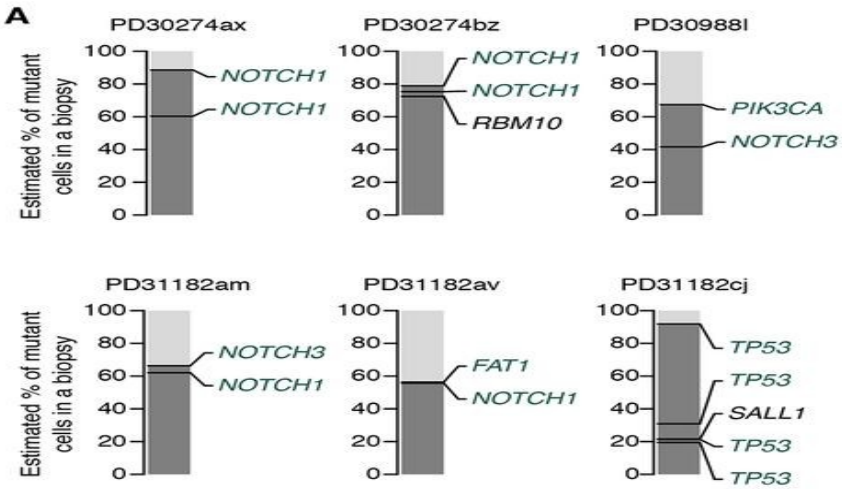


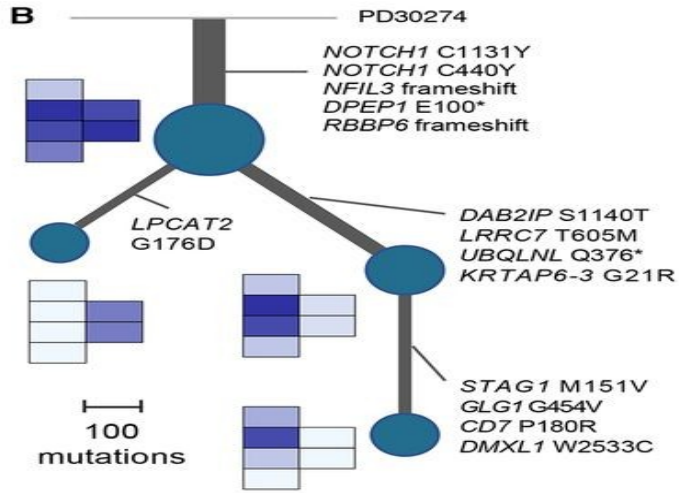
Somatic mutant clones colonize the human esophagus with age

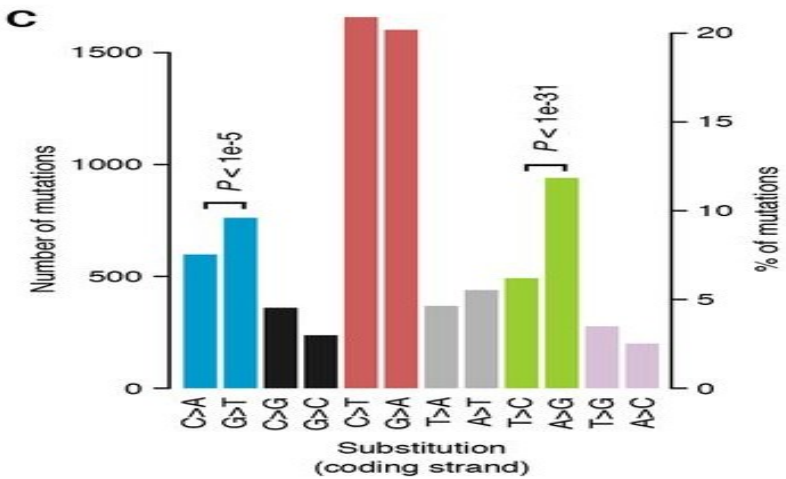


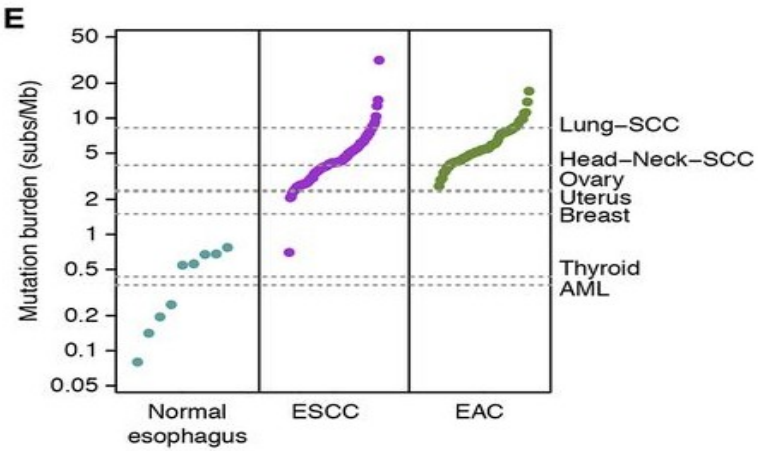


Somatic mutant clones colonize the human esophagus with age

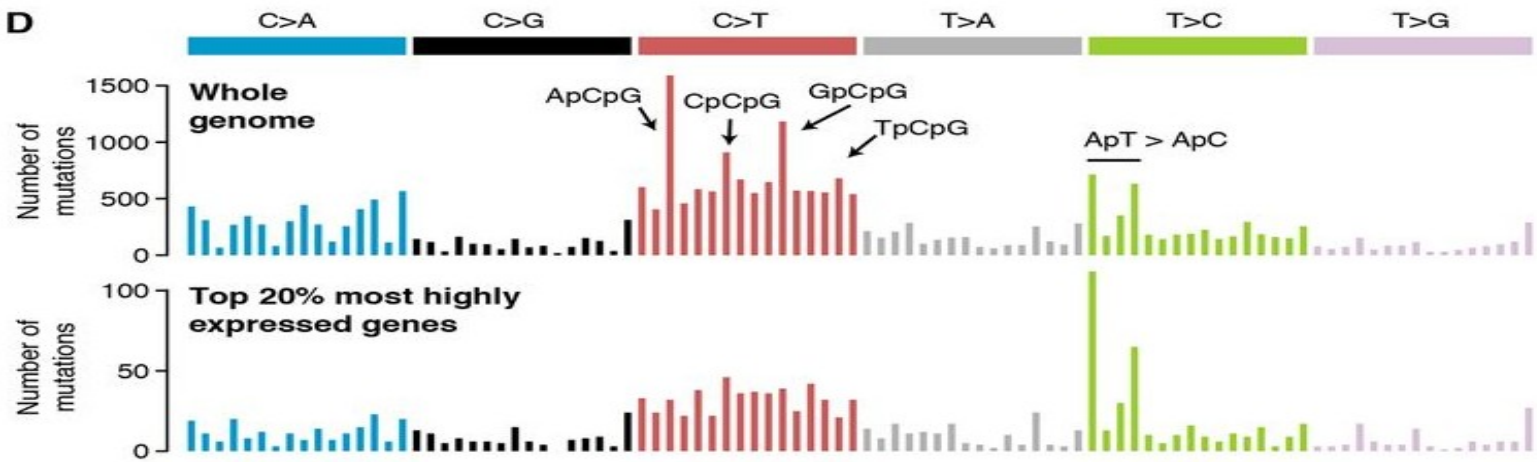


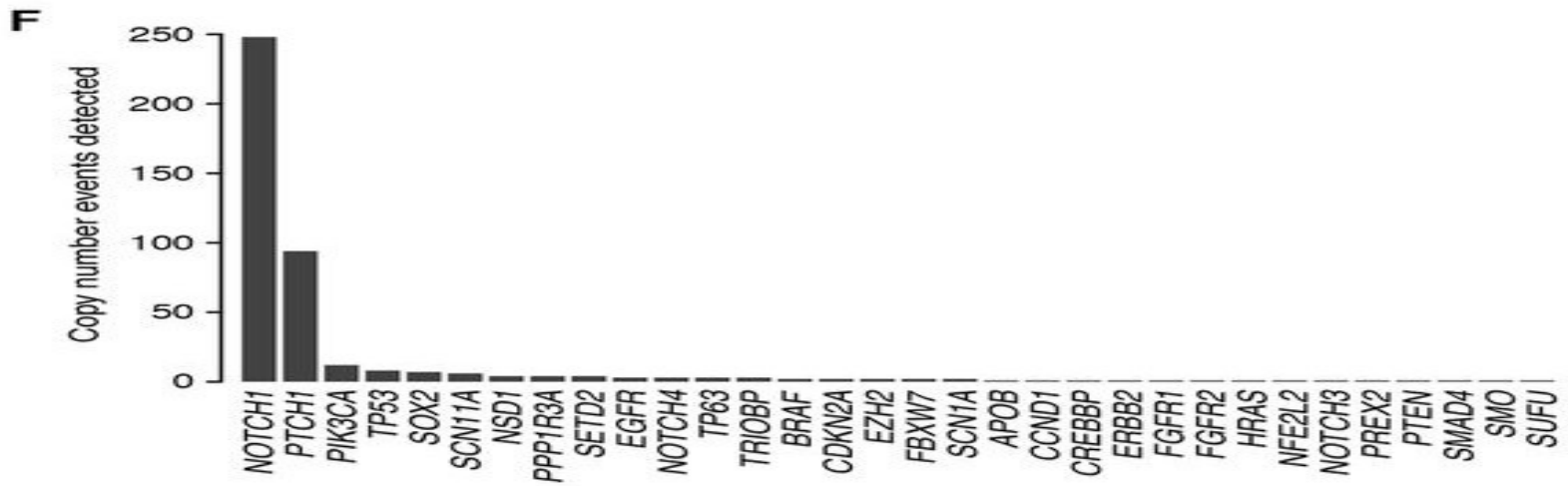




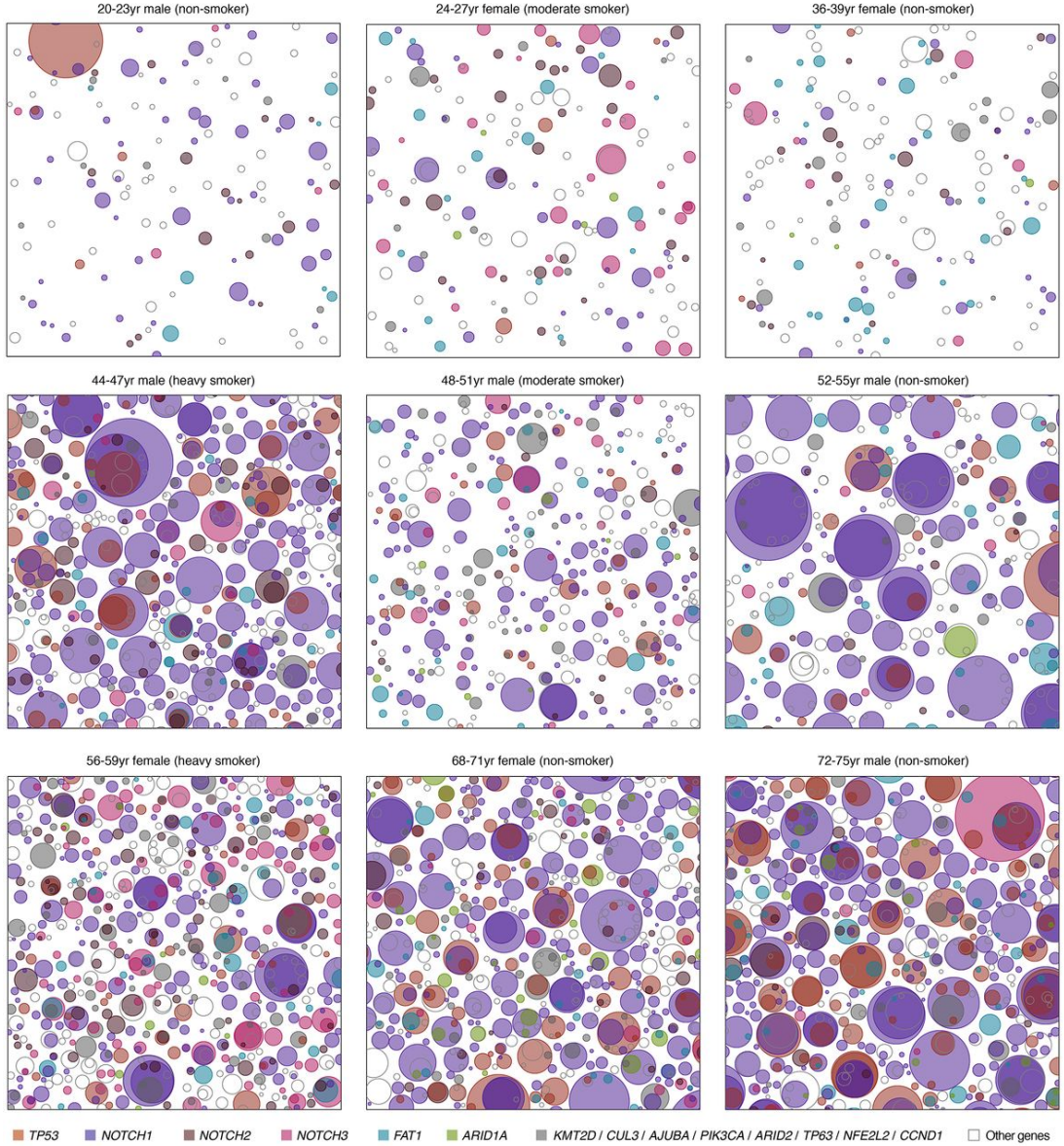


Somatic mutant clones colonize the human esophagus with age





Somatic mutant clones colonize the human esophagus with age



Somatic mutant clones colonize the human esophagus with age

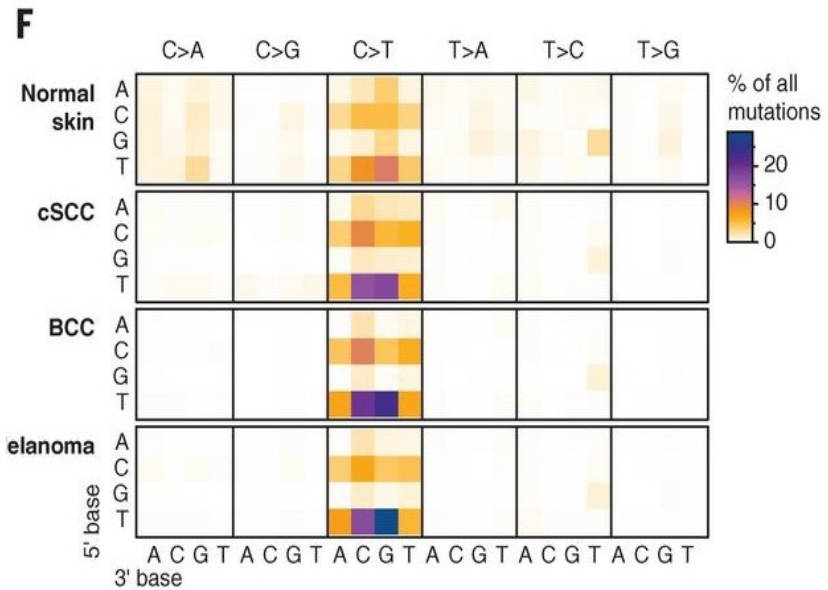
Recap

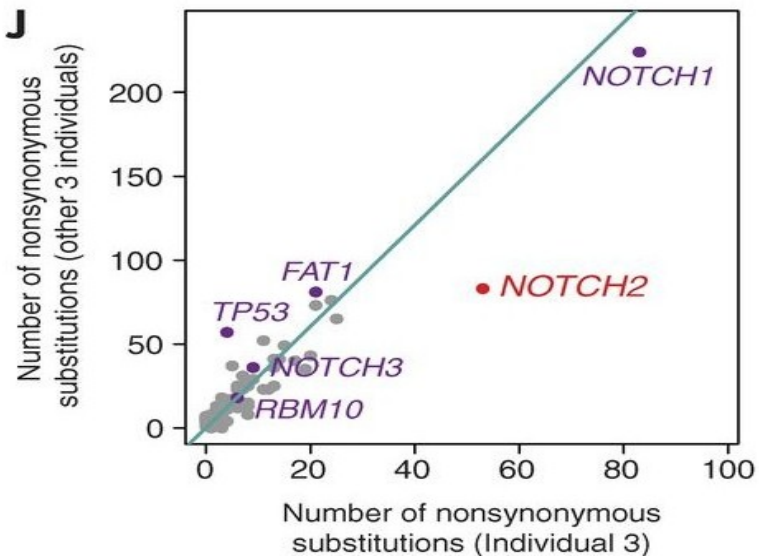
- Deep sequencing of esophageal tissue uncovered a hidden world of mutation and selection as we age
- Although esophageal tissue mutates at a rate 10x lower than sun-exposed skin, positive selection acting on cancer driving genes is much stronger
- In 2 postage-stamp size biopsies the number of driver mutations was 4000 equivalent to $\approx$ 1000 cancer genomes with 50% occurring in *Notch1* alone
- *Notch1* is mutated in 15% of ESCC yet occurs in 40% of normal cells by middle age

Somatic mutant clones colonize the human esophagus with age

Further...

In the sun-exposed paper 3 of the individuals were of European ancestry while one individual was of South Asian ancestry





Somatic mutant clones colonize the human esophagus with age

Further.....

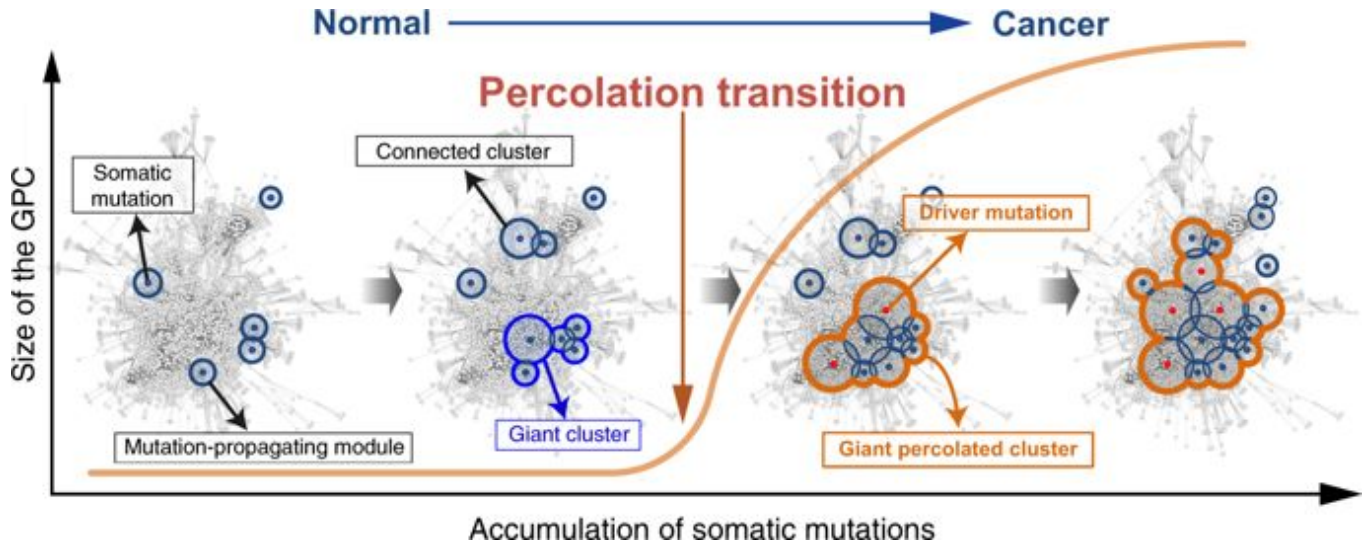


Figure: Shin *et al.* 2017