

# Driver Gene Identification

Federico Abascal  
[fa8@sanger.ac.uk](mailto:fa8@sanger.ac.uk)  
Martincorena's lab  
Wellcome Sanger Institute



# Day 3

- Drivers: theory and detection, *F. Abascal*
  - Exercises with `dndscv` (1 + 1.5 h)
- Invited seminar: *Serena Nik-Zainal* (45 m)
- Plots, *P. Basurto*
  - Introduction and exercises with `maf-tools` (50 m)
- Wrap-up (5-10 m)

# What are drivers and passengers?

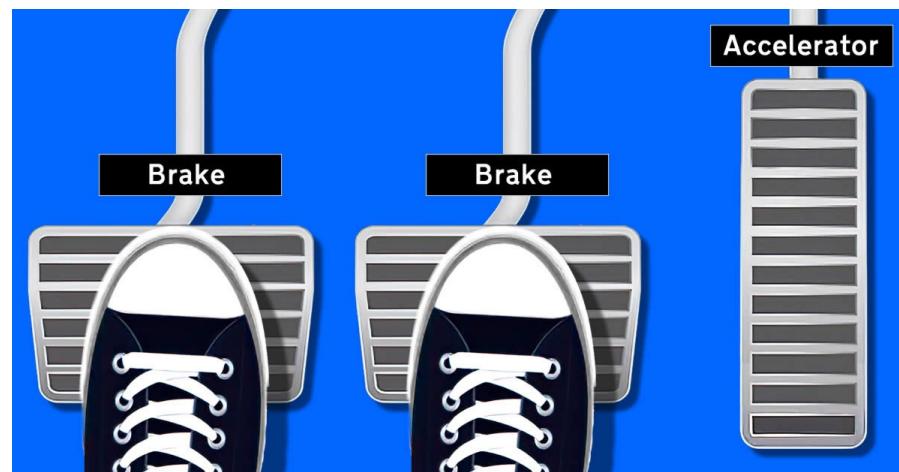
- *American Journal of Traffic and Transportation Engineering*
- Drivers are **causative alterations**
  - Substitutions and small indels – point mutations:
    - Coding alterations: KRAS G12D
    - Regulatory region: *TERT* promoter
  - Structural rearrangements and copy number changes:
    - *BCR-ABL1* in leukemia, *MYC* amplification, long deletions (*TP53*)...
  - Epigenetic alterations:
    - *VHL* expression repression through promoter hypermethylation

# Oncogenes & tumour suppressors



## Loss of function

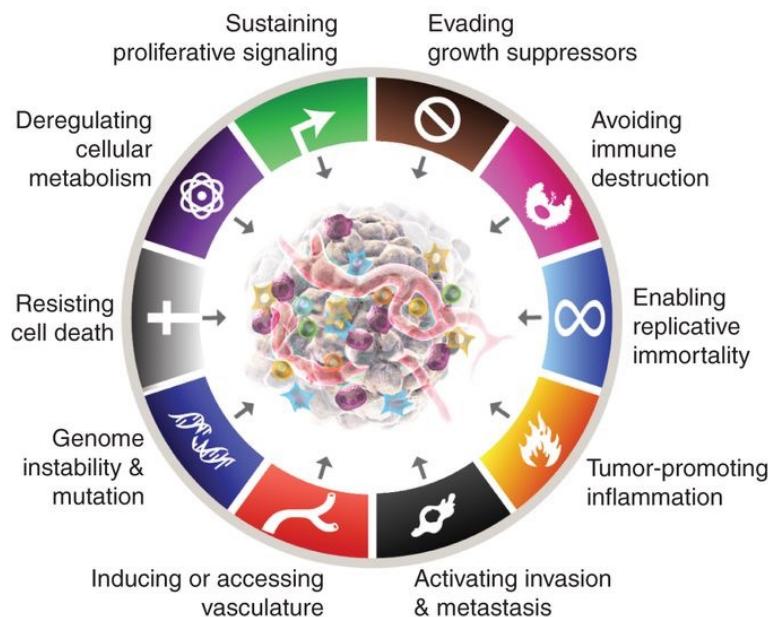
- Double hit
- Missense, splice site, nonsense mutations
- Deletions and insertions
- Loss of loci
  - *TP53*
  - *VHL*



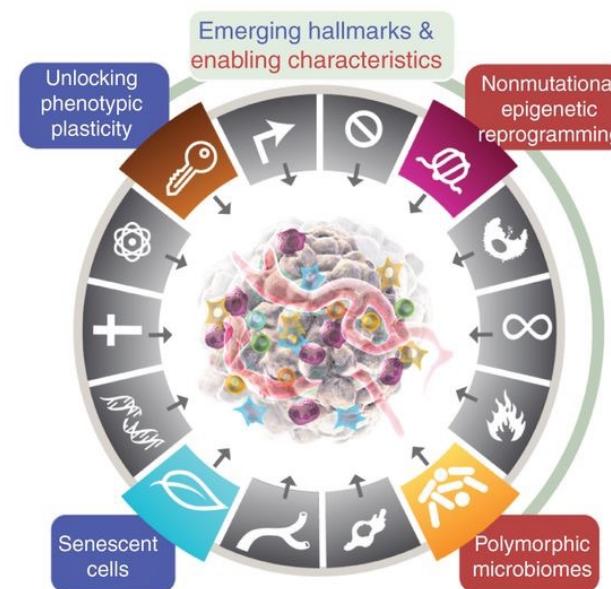
## Gain of function

- Single hit
- Missense mutations
- Copy number gain (amplification)
  - *TERT* promoter
  - *KRAS*
  - *MYC*

# Driver functions – hallmarks of cancer



Hanahan & Weinberg, 2011, *Cell*



Hanahan, 2022, *Cancer Discovery*

# How many mutations are required to develop a tumour?

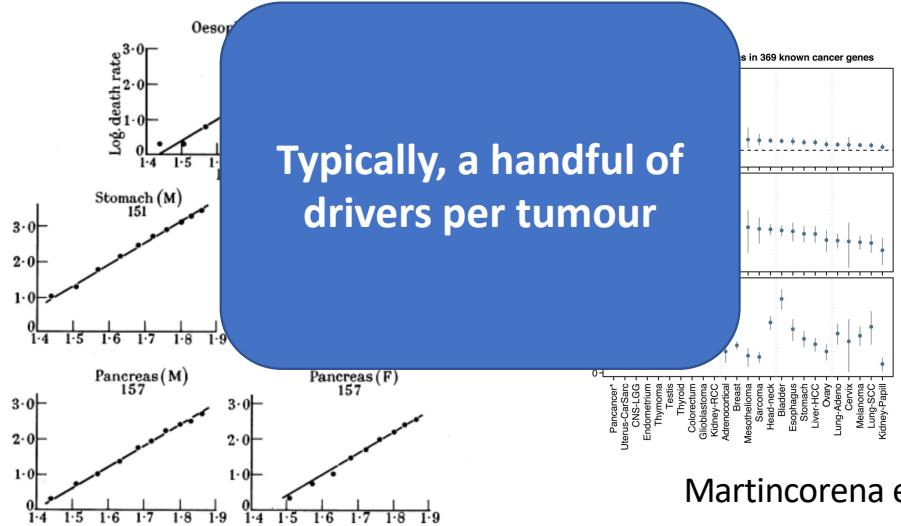
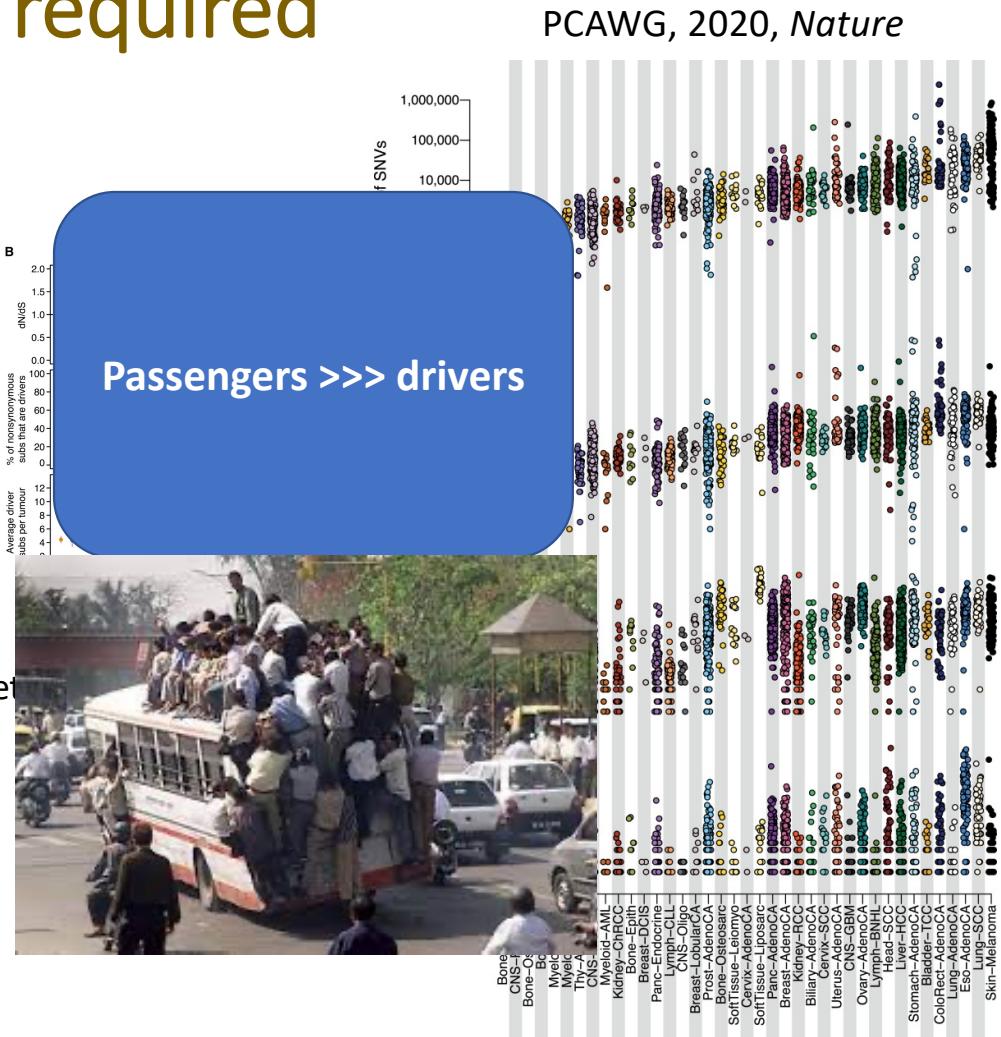


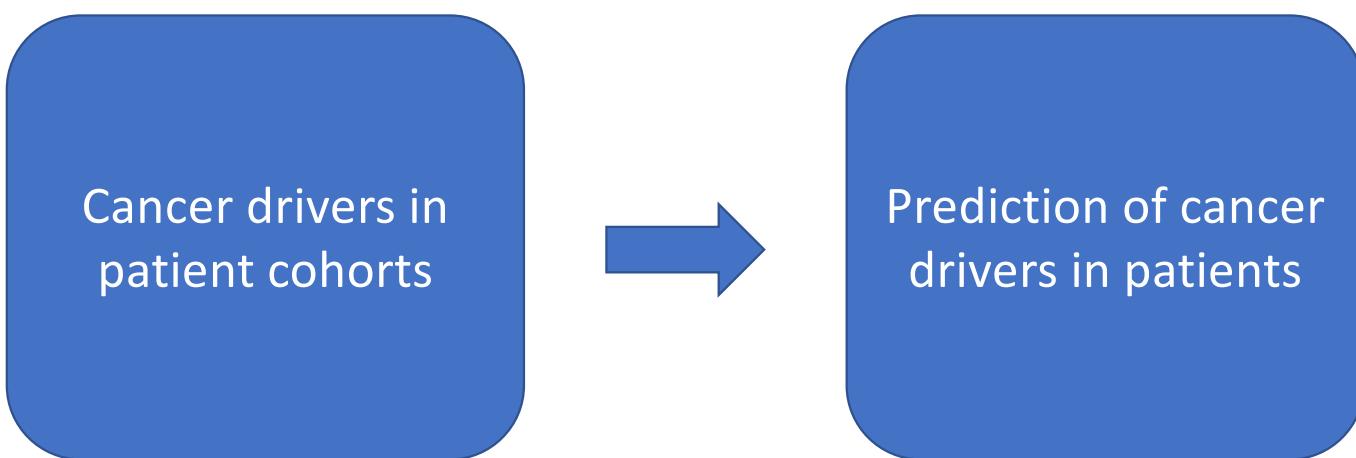
FIG. 1.—Change in mortality with age for cancer of the oesophagus, stomach and pancreas in men and for cancer of the stomach and pancreas in women shown on a double logarithmic scale, that is, the logarithm of the death rate per million persons plotted against the logarithm of the mid-point of the age group. The straight line through the points has been drawn arbitrarily to give the best fit, subject to the gradient being 6 to 1.

Armitage and Doll, 1954,  
*British Journal of Cancer*

Martincorena et al.



# Research on cancer drivers



General Biology  
Cancer Biology  
New drivers  
Cohort characterisation  
Drug research

Prediction of cancer  
drivers in patients

Risk prediction  
Treatment

# Types of point mutations

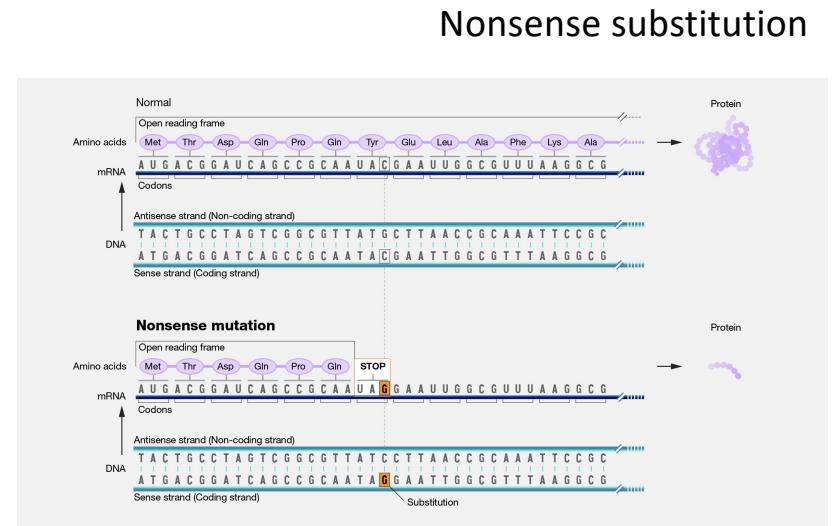
		Second letter					
		U	C	A	G		
First letter	U	UUU } Phe UUC UUA } Leu UUG	UCU } Ser UCC UCA UCG	UAU } Tyr UAC UAA Stop UAG Stop	UGU } Cys UGC UGA Stop UGG Trp	U C A G	
	C	CUU } Leu CUC CUA CUG	CCU } Pro CCC CCA CCG	CAU } His CAC CAA Gln CAG	CGU } CGC CGA Arg CGG	U C A G	
	A	AUU } Ile AUC AUA <b>AUG Met</b>	ACU } Thr ACC ACA ACG	AAU } Asn AAC AAA Lys AAG	AGU } Ser AGC AGA Arg AGG	U C A G	
	G	GUU } Val GUC GUA GUG	GCU } Ala GCC GCA GCG	GAU } Asp GAC GAA Glu GAG	GGU } Gly GGC GGA GGG	U C A G	

Synonymous substitution, e.g. TAT > TAC (Tyr → Tyr)

Missense substitution, e.g. TAT > TGT (Tyr → Cys)

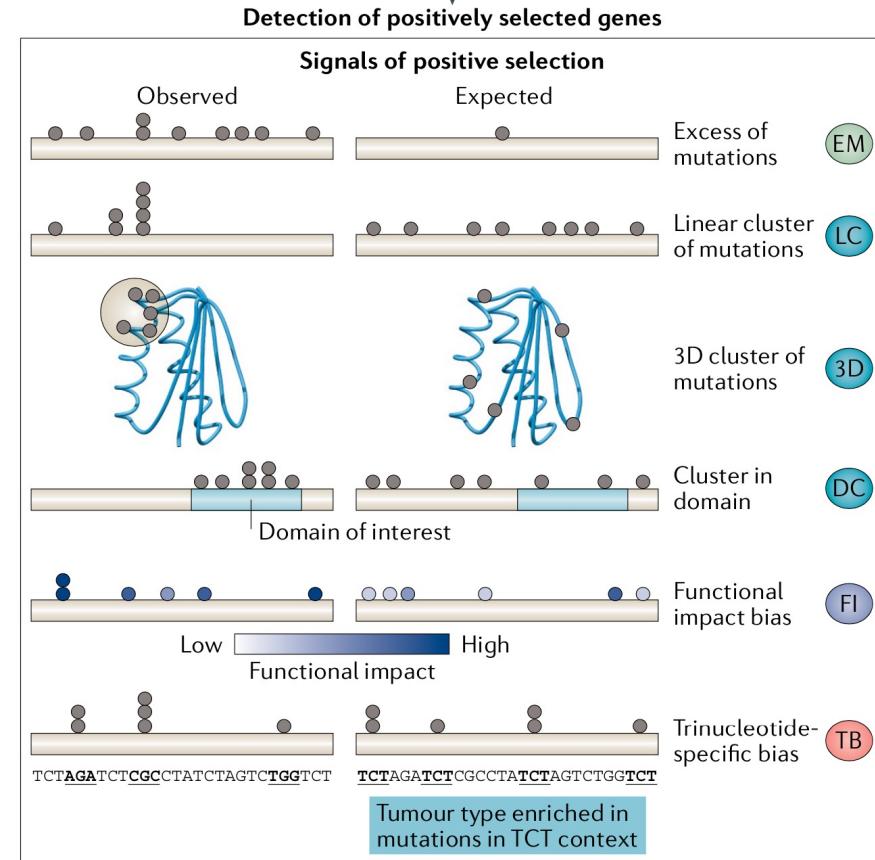
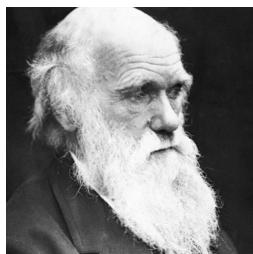
Nonsense substitution, e.g. TAT > TAA (Tyr → Stop\*)

Indels: insertions/deletions, in frame vs out of frame



# How to find drivers in cohorts?

- Like finding needles in a haystack
- Recurrence – signature of positive selection
- Key – properly modelling the mutational process: null ("expected") model

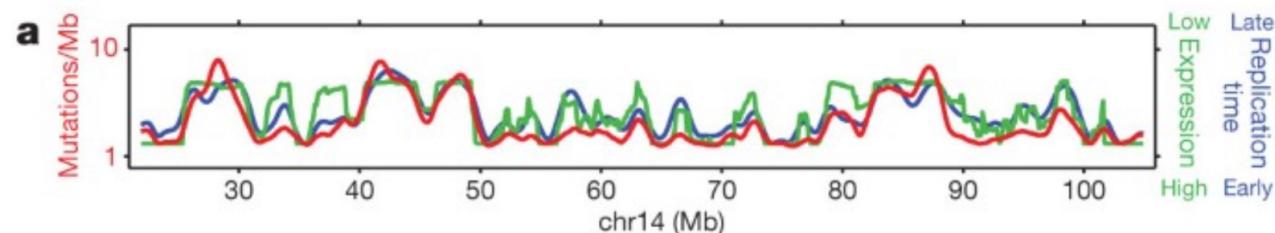
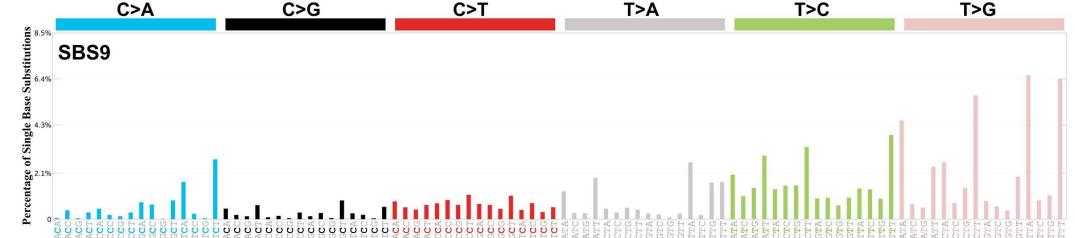
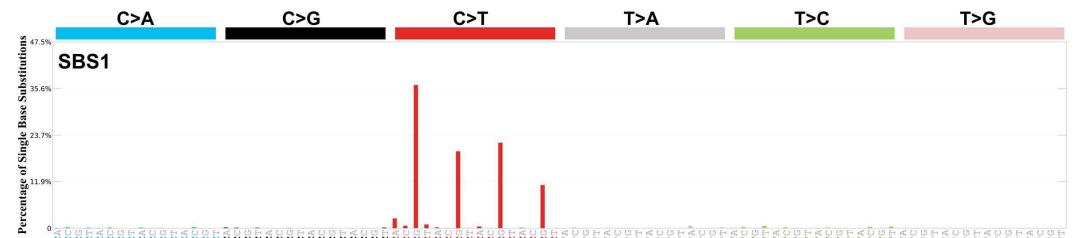


Martínez-Jiménez et al, 2020, *Nature Reviews Cancer*

# Modelling the mutational process

- Null model
  - Difficult for:
    - structural variants
      - Gistic [https://www.genepattern.org/modules/docs/GISTIC\\_2.0.html](https://www.genepattern.org/modules/docs/GISTIC_2.0.html)
      - Gene fusions, etc
    - epigenetic alterations
  - Better understood for point mutations
    - substitutions
    - small indels

*Most substitution mutational processes can be described using the trinucleotide context*



Lawrence et al, 2013, *Nature*

# Selection for substitutions in coding regions

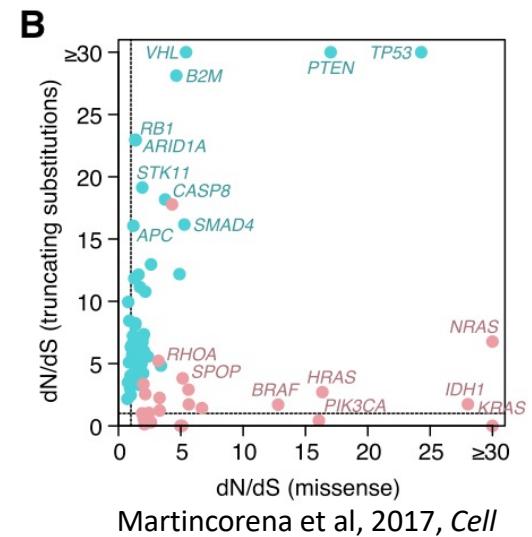
- dN/dS: a formal observed/expected test
  - Rate of non-synonymous substitutions divided by the rate of synonymous substitutions (=obs /exp under null model)
  - $dN/dS < 1$ ? **Negative selection**
  - $dN/dS = 1$ ? **Neutral evolution**
  - $dN/dS > 1$ ? **Positive selection!**
- **dndscv** <https://github.com/im3sanger/dndscv>
- Non-synonymous substitutions:
  - Missense (e.g. Leu → Pro) *(OGs and TSGs)*
  - Nonsense (e.g. Ser → Stop)
  - Splice sites

*Truncating (TSGs)*

		Second letter					
		U	C	A	G		
First letter	U	UUU } Phe UUC UCC UUA } Leu UUG	UCU } Ser UCC UCA UCG	UAU } Tyr UAC UAA Stop UAG Stop	UGU } Cys UGC UGA Stop UGG Trp	U C A G	
	C	CUU } Leu CUC CUA CUG	CCU } Pro CCC CCA CCG	CAU } His CAC CAA } Gln CAG	CGU } Arg CGC CGA CGG	U C A G	
A	A	AUU } Ile AUC AUA AUG Met	ACU } Thr ACC ACA ACG	AAU } Asn AAC AAA } Lys AAG	AGU } Ser AGC AGA } Arg AGG	U C A G	
G	G	GUU } Val GUC GUA GUG	GCU } Ala GCC GCA GCG	GAU } Asp GAC GAA } Glu GAG	GGU } Gly GGC GGA GGG	U C A G	

# Selection for point mutations in coding regions

- Oncogenes vs tumour suppressors:
  - OG: missense
  - TSG: nonsense, splice, indels, missense
- Coefficient of selection (dN/dS) interpretation
  - dN/dS of 1: all obs non-syn are exp (neutral)
  - dN/dS of 2: 50% of non-syn selected
  - dN/dS of 10: 90% of non-syn selected
  - dN/dS of 100: 99% of non-syn selected



# dndscv detecting selection

- R package dndscv (Martincorena et al, *Cell* 2017)
  - dN/dS model – coef. of selection
  - Sophisticate mutational model:
    - Trinuc frequencies
    - Covariates for regional variation + data at hand
    - Overdispersion – negative binomial (equivalent of Poisson distribution)
  - User friendly
  - <https://github.com/im3sanger/dndscv>
- Selection estimates at:
  - Gene level (or domains)
  - Global
  - Sites & codons
  - Not only for cancer

# dndscv gene level

- How to run:
  - dndscv function
  - panel vs no panel
  - Hypermutators
- Quick tutorial:  
<http://htmlpreview.github.io/?http://github.com/im3sanger/dndscv/blob/master/vignettes/dNdScv.html>
- Outputs:

```
dout = dndscv(muts)
names(dout)
[1] "globaldnds"      "sel_cv"           "sel_loc"          "annotmuts"       "genemuts"
[6] "geneindels"       "mle_submodel"    "exclsamples"     "exclmuts"        "nbreg"
[11] "nbregind"         "poissmodel"      "wrongmuts"       "N"                "L"
```

# dndscv example: TCGA bladder carcinoma

gene_name	n_syn	n_mis	n_non	n_spl	n_ind	wmis_cv	wnon_cv	wspl_cv	wind_cv	qglobal_cv
TP53	5	143	35	4	24	41.0	97.7	97.7	366.9	0
PIK3CA	3	80	0	0	1	14.5	0.0	0.0	6.1	0
FGFR3	6	53	1	0	2	12.6	2.6	2.6	13.7	0
ARID1A	7	34	45	6	30	2.7	38.1	38.1	65.7	0
KDM6A	6	21	33	9	34	2.5	33.6	33.6	136.9	0
KMT2D	9	39	50	10	29	1.7	23.6	23.6	26.2	0
RB1	0	8	34	14	21	2.3	91.4	91.4	138.5	0
STAG2	8	12	22	4	14	1.3	21.8	21.8	64.8	0
ELF3	0	29	1	1	20	17.1	10.1	10.1	138.9	0
CDKN2A.p16	0	12	3	2	5	16.2	104.4	104.4	188.4	0
CDKN1A	1	5	6	0	23	5.6	51.2	51.2	389.1	0
TSC1	3	9	12	5	6	1.4	25.6	25.6	27.1	1.86E-13
ZFP36L1	0	7	3	0	18	4.1	29.5	29.5	85.4	1.54E-12
FBXW7	0	19	10	1	2	7.5	35.3	35.3	16.8	5.32E-10
EP300	2	40	16	1	7	4.5	14.4	14.4	14.1	2.90E-09

...

**dndscv** tutorials:

<https://drive.google.com/drive/folders/1-M4ee1sdzqZCyEwEnFNHFDZs2LkZA0dr?usp=sharing>

<https://github.com/im3sanger/dndscv>

Other tools:

- **MutSigCV:** <https://www.genepattern.org/modules/docs/MutSigCV>
- **IntOGen:** <https://www.intogen.org/search>

$w = dN/dS$  (coef. of selection)

$(w-1)/w$  = fraction mutations selected

*TP53* wmis = 41.0, 97.5% selected

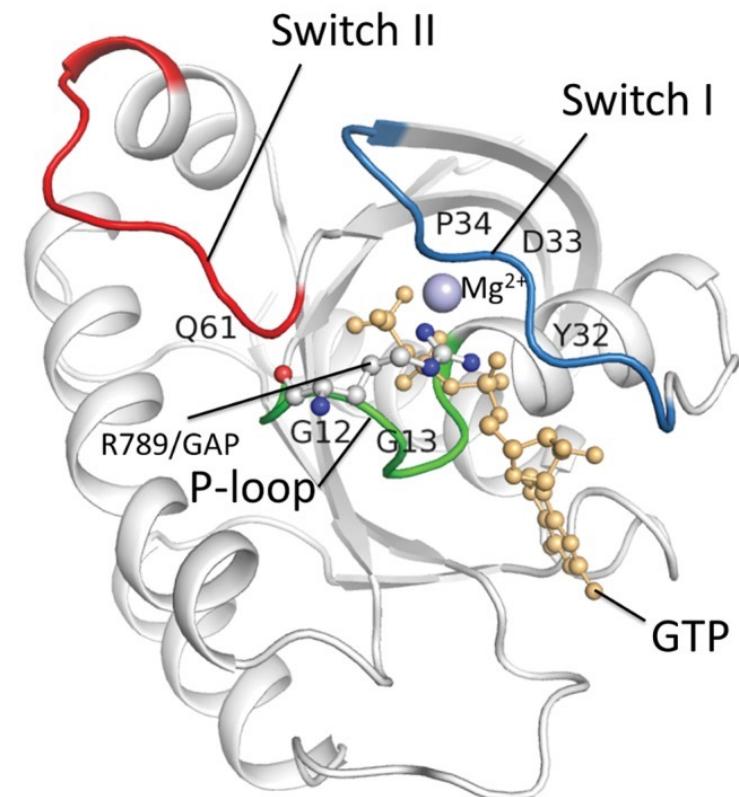
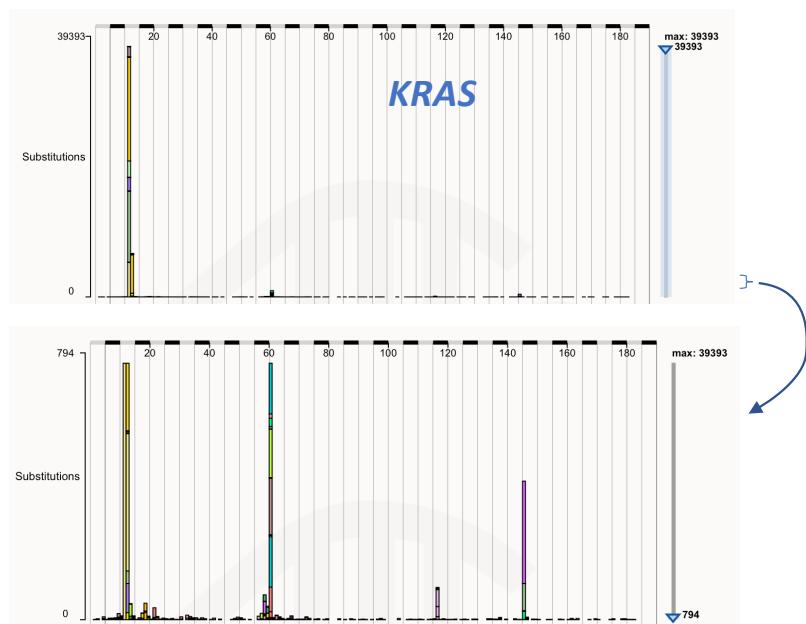
*EP300* wmis = 4.5, 77.8% selected

Why are coefficients of selection important?

- Personalised medicine implications
- Hypermutators lower dN/dS

# dndscv detecting selection at sites/codons

- Recurrence - hotspots
- COSMIC:



Chen et al, 2013, *PLoS one*

# Hotspot analysis with dndscv: sitednds and codondnds

Site dN/dS on TCGA data

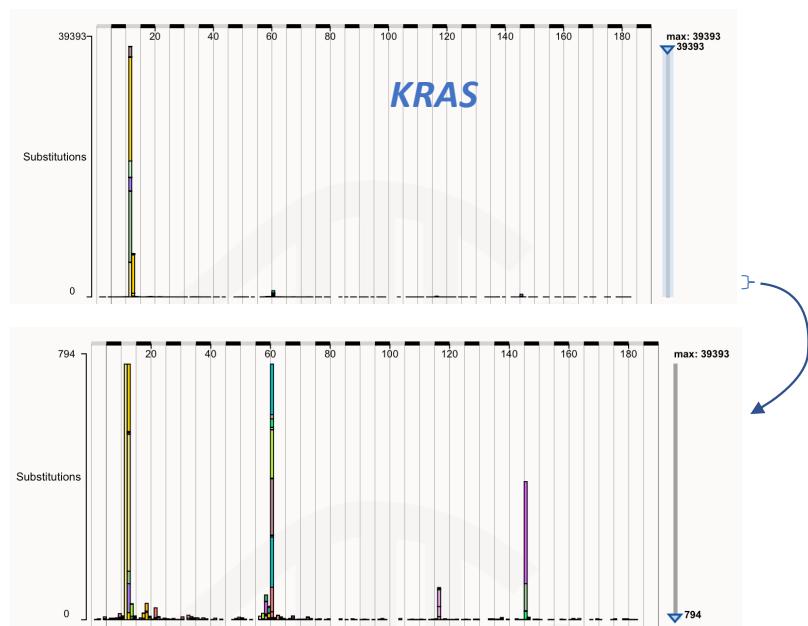
chr	pos	ref	mut	gene	aachange	impact	ref3_cod	mut3_cod	freq	mu	dnds	pval	qval
7	140453136	A	T	BRAF	V600E	Missense	GTG	GAG	408	0.002	176526.1	0	0
2	209113112	C	T	IDH1	R132H	Missense	CGT	CAT	353	0.079	4469.6	0	0
3	178952085	A	G	PIK3CA	H1047R	Missense	CAT	CGT	161	0.007	22252.8	0	0
12	25398284	C	T	KRAS	G12D	Missense	GGT	GAT	114	0.008	14578.7	6.21E-241	1.62E-233
3	178936091	G	A	PIK3CA	E545K	Missense	TGA	TAA	150	0.029	5112.8	4.94E-232	1.03E-224
12	25398284	C	A	KRAS	G12V	Missense	GGT	GTT	95	0.007	13634.6	1.26E-205	2.19E-198
1	115256529	T	C	NRAS	Q61R	Missense	CAA	CGA	72	0.003	25980.9	1.02E-184	1.51E-177
3	178936082	G	A	PIK3CA	E542K	Missense	TGA	TAA	95	0.029	3238.1	3.17E-147	4.13E-140
17	7578406	C	T	TP53	R175H	Missense	CGC	CAC	108	0.056	1938.8	2.16E-138	2.50E-131
12	25398285	C	A	KRAS	G12C	Missense	TGG	TTG	62	0.007	8299.3	1.36E-132	1.42E-125
17	7577538	C	T	TP53	R248Q	Missense	CGG	CAG	92	0.045	2025.9	9.10E-126	8.62E-119
17	7577121	G	A	TP53	R273C	Missense	GCG	GTG	90	0.056	1596.8	5.13E-115	4.46E-108
1	115256530	G	T	NRAS	Q61K	Missense	ACA	AAA	48	0.004	11185.0	2.62E-114	2.10E-107
17	7578190	T	C	TP53	Y220C	Missense	TAT	TGT	49	0.015	3290.9	2.00E-90	1.49E-83
4	1803568	C	G	FGFR3	S249C	Missense	TCC	TGC	35	0.004	8415.1	5.05E-84	3.51E-77
17	7577094	G	A	TP53	R282W	Missense	CCG	CTG	60	0.049	1216.8	3.12E-80	1.91E-73
17	7577539	G	A	TP53	R248W	Missense	CCG	CTG	60	0.049	1216.8	3.12E-80	1.91E-73
11	533874	T	C	HRAS	Q61R	Missense	CAG	CGG	29	0.002	14650.1	4.55E-79	2.63E-72
17	7577120	C	T	TP53	R273H	Missense	CGT	CAT	69	0.083	832.5	1.53E-77	8.36E-71
12	25398285	C	G	KRAS	G12R	Missense	TGG	TCG	28	0.002	14496.8	1.15E-76	5.99E-70
12	25398281	C	T	KRAS	G13D	Missense	GGC	GAC	34	0.008	4430.7	1.19E-72	5.90E-66
17	7578394	T	C	TP53	H179R	Missense	CAT	CGT	32	0.008	3808.1	3.61E-67	1.71E-60
3	178952085	A	T	PIK3CA	H1047L	Missense	CAT	CTT	26	0.003	8447.0	5.41E-66	2.45E-59

.... 565 with q<0.1

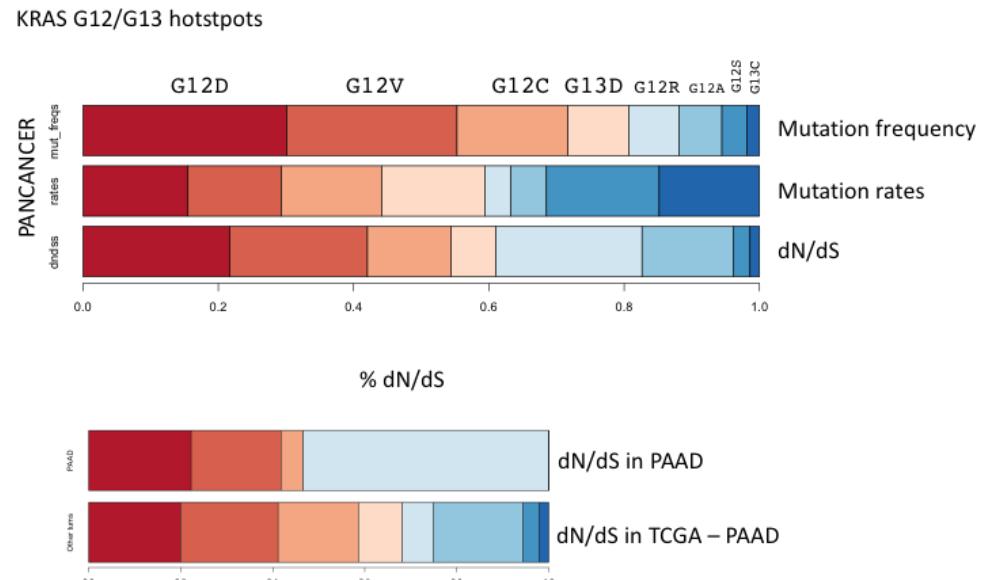
Quick tutorial: <https://rdrr.io/github/im3sanger/dndscv/f/vignettes/sitednds.Rmd>

# Hotspots: sitednds and codondnds

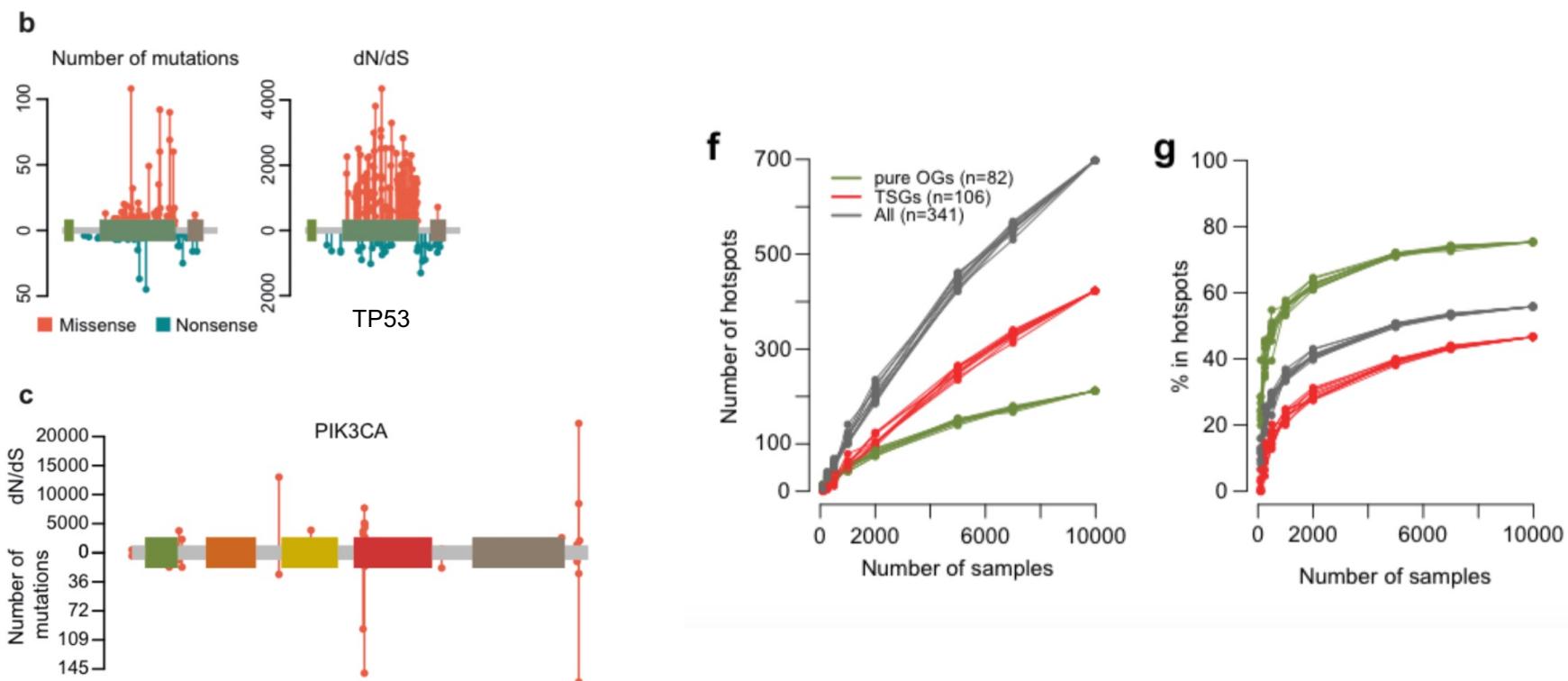
- Recurrence
- COSMIC:



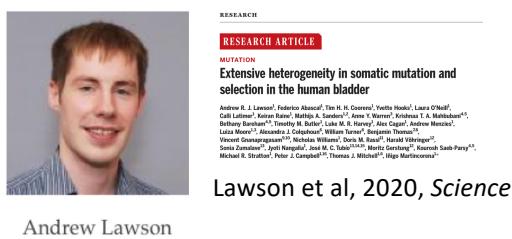
*KRAS hotspots vary from tumour to tumour*



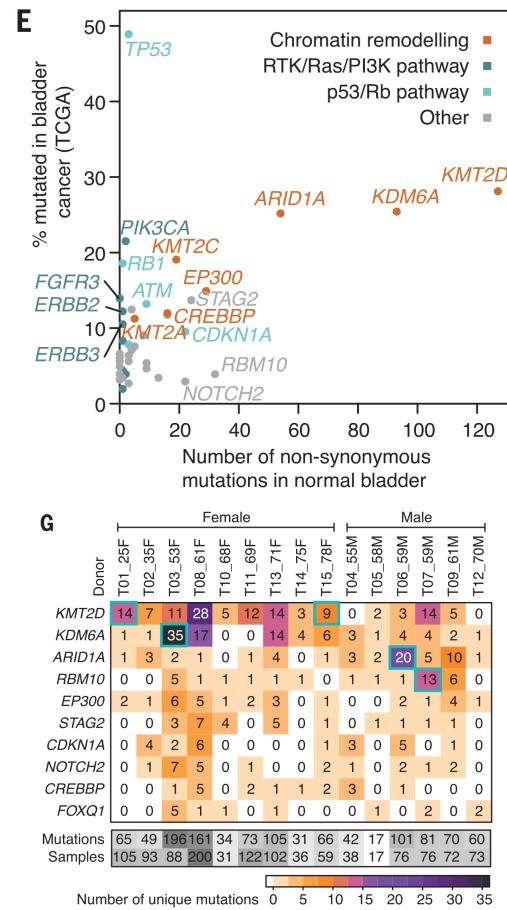
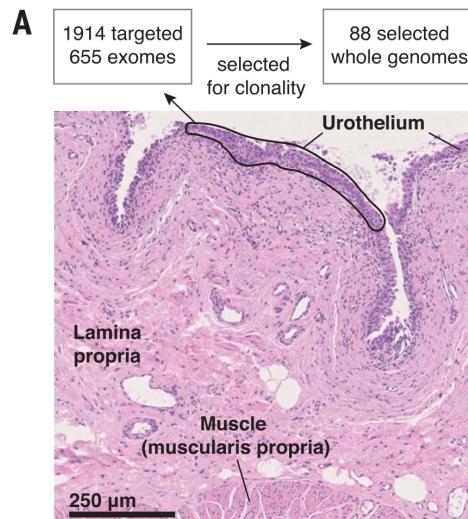
# Hotspots: best for OGs



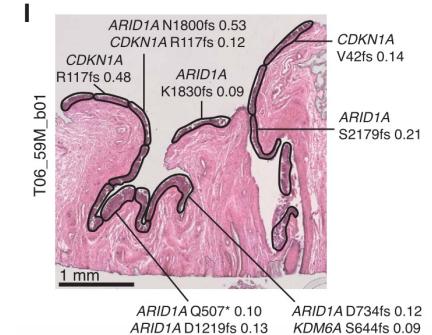
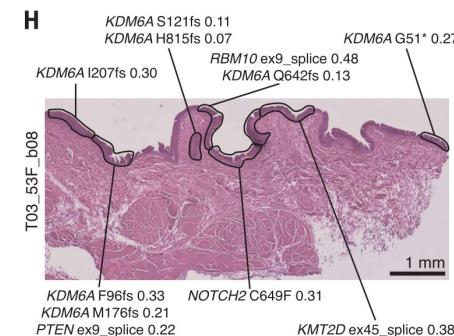
# dndscv example: drivers in healthy bladder



Andrew Lawson



- *Chromatin modifiers!*
  - *TP53 diagnostic?*
  - *Differences between donors*



# Regulatory regions and noncoding genes

- Protein coding: 1% of the genome
  - PCAWG: 2,658 whole genomes (TCGA exomes)
    - microRNAs, lncRNAs, tRNAs...
    - Promoters, 3' and 5' UTRs, enhancers
  - Large set of driver candidates
    - lncRNAs: *NEAT1* & *MALAT1*...
    - UTRs/promoters: *WDR74*...
    - tRNAs
    - small RNAs
    - micro RNAs

## Article

## Analyses of non-coding somatic drivers in 2,658 cancer whole genomes

Received: 19 January 2018  
Accepted: 2 December 2019  
Published online: 5 February 2020  
Open access

Ester Rhynhart<sup>1,2\*</sup>, Marten Muller Nielsen<sup>3</sup>, Fabrizio Abascal<sup>4</sup>, Juliano H. Azevedo<sup>5</sup>,  
Ofer Shepp<sup>6</sup>, Grace Tsao<sup>7</sup>, Henrik Mørk<sup>8</sup>, Juliano M. Ferreira<sup>9</sup>, Renzo Inderli<sup>10</sup>, Ziao Li<sup>11</sup>,  
Lars Malmström<sup>12</sup>, Shimrit Shani<sup>13</sup>, Anders Larsson<sup>14</sup>, Carl Hermann<sup>15</sup>, Yoffi<sup>16</sup>, Maruska<sup>17</sup>,  
Clyne Shyan<sup>18</sup>, Samukwana B. Amat<sup>19</sup>, Pranav Bandopadhyay<sup>20</sup>, Johanna Bert<sup>21</sup>,  
Kathleen M. Brinkmann<sup>22</sup>, Michael C. Burman<sup>23</sup>, Daniel J. Cane<sup>24</sup>, Michael C. Chappell<sup>25</sup>,  
Calvin Weng Yiu Chan<sup>26</sup>, David Craft<sup>27</sup>, Qiyanan Dhingra<sup>28</sup>, Kew Diamond<sup>29</sup>,  
Nuno A. Fonseca<sup>30</sup>, Abdellatif M. Perez<sup>31</sup>, Qianyan Ding<sup>32</sup>, Mark J. Hamerling<sup>33</sup>,  
Nishanth K. Jayaram<sup>34</sup>, Suresh K. Jayaram<sup>35</sup>, S. K. Jayaram<sup>36</sup>, S. K. Jayaram<sup>37</sup>,  
Andres Kalman<sup>38</sup>, Abdulrahman Kahloun<sup>39</sup>, Vassilios Karayannidis<sup>40</sup>, Kuan Kurnia<sup>41</sup>,  
Sudheendra Kumar<sup>42</sup>, Donghong Lee<sup>43</sup>, Kyoung Van Lee<sup>44</sup>, Jun Konowacki<sup>45</sup>, Kuan Kurnia<sup>46</sup>,  
Luisa Lanza<sup>47</sup>, Ming Tang<sup>48</sup>, Ming Tang<sup>49</sup>, Ming Tang<sup>50</sup>, Ming Tang<sup>51</sup>, Ming Tang<sup>52</sup>,  
Steven E. Schuchman<sup>53</sup>, Nikita Sidropoulos<sup>54</sup>, Liya Szwarcberg<sup>55</sup>, Nasim-Sinnot-Armstrong<sup>56</sup>,  
Clif Steward<sup>57</sup>, David Tambrone<sup>58</sup>, Jose M. C. Tavares<sup>59</sup>, Wei Wang<sup>60</sup>,  
Lin Xia<sup>61</sup>, Xiangyu Xie<sup>62</sup>, Xiangyu Xie<sup>63</sup>, Xiangyu Xie<sup>64</sup>,  
Cheng-Zhang Zhang<sup>65</sup>, Jing Zhang<sup>66</sup>, Jamie A. Haber<sup>67</sup>, Asper Hobolth<sup>68</sup>,  
Marcello J. J. da Cunha<sup>69</sup>, Alessio Kalisz<sup>70</sup>, Michael J. Lawlor<sup>71</sup>, Jennifer M. Matting<sup>72</sup>,  
Hidenori Miyake<sup>73</sup>, Bejanapoor Rajabi<sup>74</sup>, Rui Wang<sup>75</sup>, Chris Yiu<sup>76</sup>,  
Lincoln D. Stein<sup>77</sup>, Joshua M. Stutts<sup>78</sup>, Tatsuhiko Suzuki<sup>79</sup>, David A. Wheeler<sup>80</sup>,  
Peter J. Campbell<sup>81</sup>, Nicola Loper<sup>82</sup>, Peter O’Donovan<sup>83</sup>,  
Peter J. Campbell<sup>84</sup>, Nicola Loper<sup>85</sup>,  
PCAWD Working Group<sup>86</sup>, PCAWD Structural Variation Working Group<sup>87</sup>, Joachim Weischenfeldt<sup>88</sup>,  
Bansal et al.<sup>89</sup>, Brem et al.<sup>90</sup>, Brem et al.<sup>91</sup>, Jakob Søe Pedersen<sup>92</sup>,  
Gert Gutz<sup>93,94</sup>, and CCW Consortium<sup>95</sup>

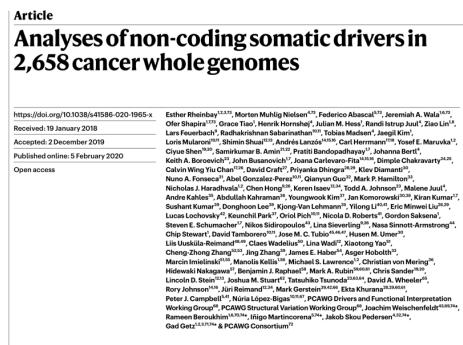
Rheinbay et al, 2020, *Nature*

# Regulatory and noncoding genes

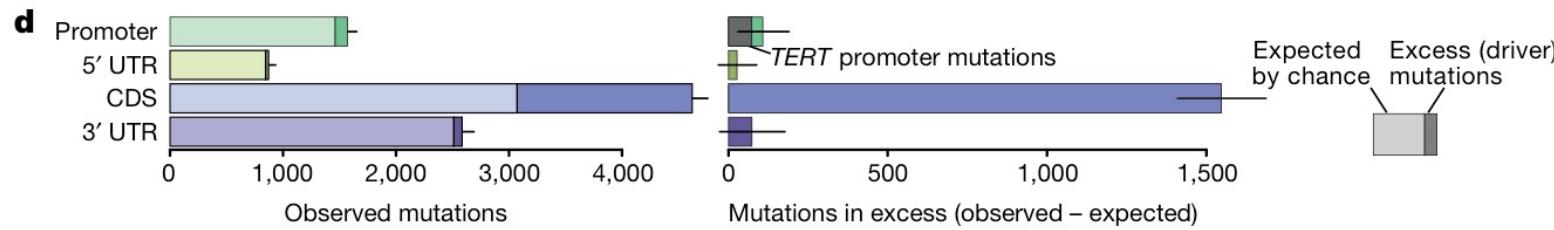
- Preliminary set of candidates
  - lncRNAs: NEAT1 & MALAT1...
  - UTRs/promoters: WDR74, etc...
  - tRNAs
  - small RNAs
  - micro RNAs



# Regulatory regions of cancer genes: promoters and UTRs



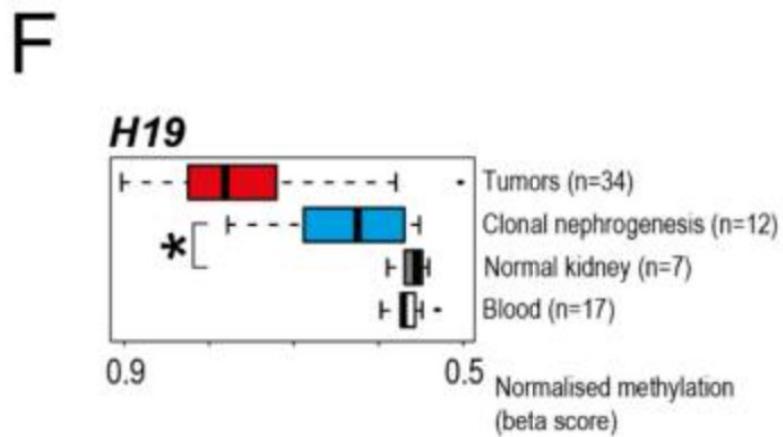
- No significant hits in cancer genes  
(except for *TERT* promoter, *TP53* 5'UTR and *TOB1* 3'UTR)



- Not much beyond *TERT* promoter

# Epigenetic drivers

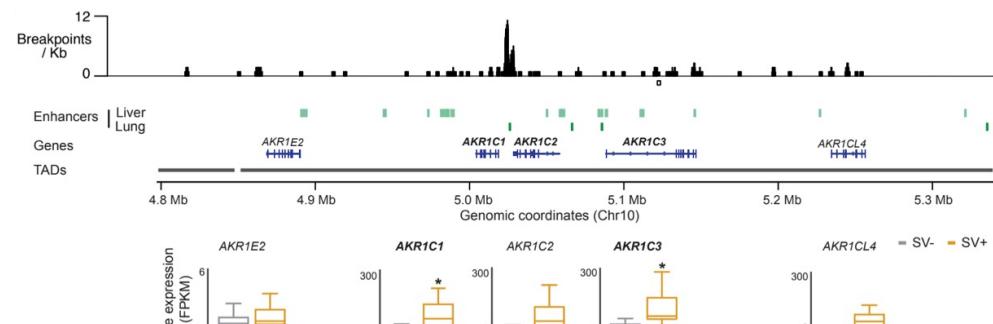
- Chromatin modifiers
- Plasticity: Δ individuals, cell types, environment → null model?
- Wilms tumour: *H19* hypermethylation



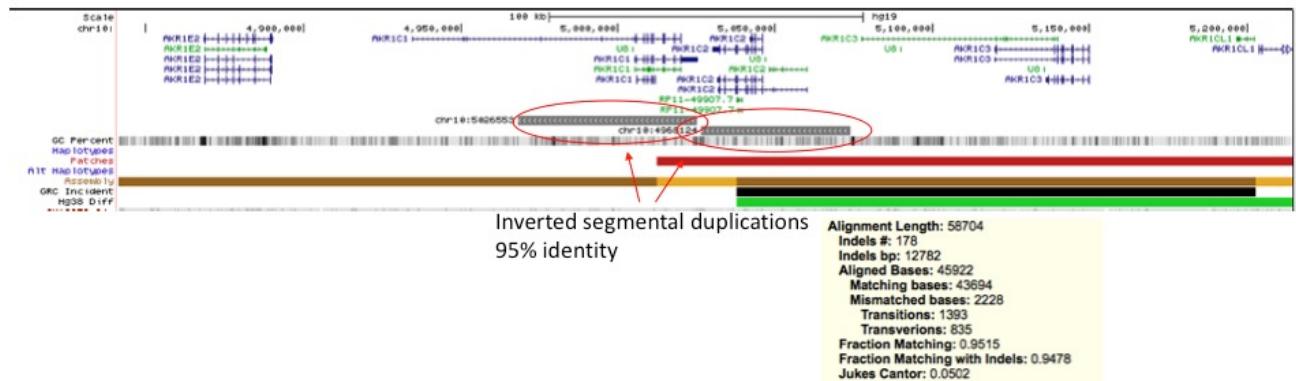
Coorens et al, 2019, *Science*

# Structural drivers

- Null model? Fragile sites *why?*
- Gene fusions, enhancer hitch-hiking, copy number gain/loses, etc
- AKR1C locus

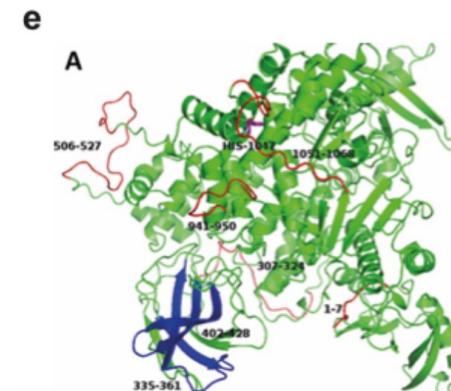
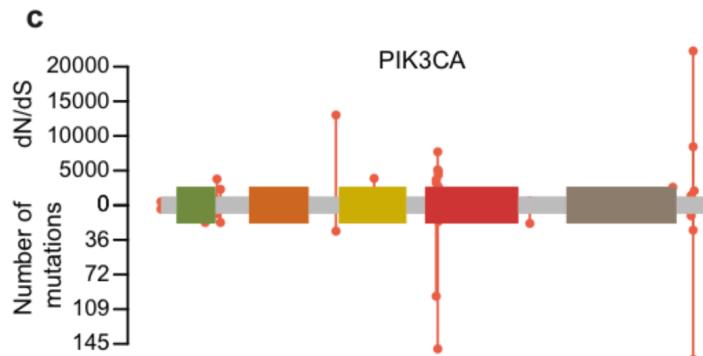


Rheinbay et al, 2020, *Nature*



# Take home messages

- OG and TSG behave very differently – no B&W in Biology
- Passengers >>> drivers (hypermutators particularly problematic)
- Recurrence = positive selection (obs>exp)
- Most drivers are protein-coding (+TERT) – 1% of the genome!
- Not all non-synonymous mutations are drivers (dN/dS)
- Structural and epigenetic alterations can be drivers too



Start codon loss in OG?

# Prediction of driver mutations in a given patient

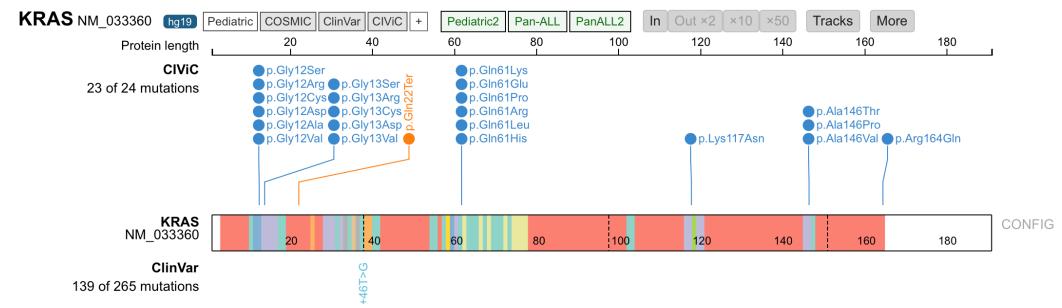
**Table 1.** Molecular targets for personalised cancer therapies

Cancer type	Cellular target	Targeted agent	Class of agent
Colorectal <sup>16–18</sup>	KRAS	Cetuximab	Monoclonal antibody against EGFR
Breast <sup>19,20</sup>	HER2	Trastuzumab	Monoclonal antibody against HER2/ Neu (EGFR2)
Chronic myeloid leukaemia <sup>21,22</sup>	BCR-ABL fusion protein	Imatinib	Receptor tyrosine kinase inhibitor
Gastrointestinal stromal tumours <sup>23,24</sup>	c-KIT	Imatinib	Receptor tyrosine kinase inhibitor
Non-small-cell lung cancer <sup>25–28</sup>	EGFR	Erlotinib and gefitinib	Receptor tyrosine kinase inhibitor
Non-small-cell lung cancer <sup>29,30</sup>	EML4-ALK fusion protein	Crizotinib	Receptor tyrosine kinase inhibitor
Metastatic malignant melanoma <sup>31,32</sup>	BRAF V600E	Vemurafenib	B-raf/MEK/ERK pathway inhibitor
Ovarian, breast and prostate can- cer (under investigation) <sup>33,34</sup>	BRCA1, BRCA2	Olaparib	Poly(ADP-ribose) polymerase (PARP) inhibitor

**Abbreviations:** APC: adenomatous polyposis coli; CML: chronic myeloid leukaemia; CRC: colorectal cancer; EGFR: epidermal growth factor receptor; EML4-ALK: echinoderm microtubule-associated protein-like 4—anaplastic lymphoma kinase fusion gene; FAP: familial adenomatous polyposis coli; GIST: gastrointestinal stromal tumour; NICE: National Institute for Health and Care Excellence; NSCLC: non-small cell lung cancer; PARP: poly(ADP-ribose) polymerase; TK: tyrosine kinase; TKI: tyrosine kinase inhibitor

# Clinical practice

- Top genes: most mutations are drivers (*TP53*, *KRAS*, *BRAF*...)
  - Careful with hypermutators
- Hotspots: most drivers
- OGs and TSGs behave differently
- Useful tools:
  - CiVIC: <https://civicdb.org/home>
  - Cancer Genome Interpreter: <https://www.cancergenomeinterpreter.org/home>



# Cancer Genome Interpreter example



Tamborero et al, 2018,  
*Genome Medicine*  
Muiños et al, 2021,  
*Nature*  
**BoostDM:** CiVIC,  
recurrence,  
conservation, ...

This analysis will be removed after 6 months

**Mutations**   **CNAs**

Show entries with:  Mutations identified as drivers  Mutations with oncogenic annotations  Other mutations

Sample ID	Gene	Protein Change	Oncogenicity	Mutation	Consequence	Oncogenic annotation
TCGA-AG-3999	KRAS	G12S				2 216
TCGA-AG-3999	TP53	R213*				
TCGA-AA-A00D	TP53					
TCGA-AA-A00D	PIK3CA					
TCGA-AA-A00D	APC					
TCGA-AA-A00D	BRAF					
TCGA-AG-3999	APC					
TCGA-AA-A00D	APC					
TCGA-AG-3999	BCI					
TCGA-AG-3999	PTEN	M KRAS (G12S)	KRAS (12,13)		Cetuximab (EGFR mAb inhibitor)	Colorectal adenocarcinoma
TCGA-AG-3999	PTEN	M KRAS (G12S)	KRAS (12,13,59,61,117,146)		Panitumumab (EGFR mAb inhibitor)	Colorectal adenocarcinoma
TCGA-AG-3999	UBR3	M KRAS (G12S)	KRAS oncogenic mutation		Panitumumab (EGFR mAb inhibitor)	Colorectal adenocarcinoma
TCGA-AG-3999	PNL	M KRAS (G12S)	KRAS oncogenic mutation		Cetuximab (EGFR mAb inhibitor)	Colorectal adenocarcinoma
TCGA-AA-A00D	C10	M KRAS (G12S)	KRAS (D119N,G12F,F156L,G60R,F28I)		Panitumumab + Cetuximab	Colorectal adenocarcinoma
TCGA-AG-3999	DDIT3	M KRAS (G12S)	KRAS oncogenic mutation		Trastuzumab + Lapatinib (ERBB2 mAb inhibitor + EGFR mAb inhibitor)	Colorectal adenocarcinoma
TCGA-AG-3999	TRF	M KRAS (G12S)	KRAS (A146T,G13D,G12C,,,A146P,Q66E)		Cetuximab	Colorectal adenocarcinoma
TCGA-AG-3999	KCNQ1	M KRAS (G12S)	KRAS oncogenic mutation		Bevacizumab	Colorectal adenocarcinoma
TCGA-AA-A00D	LRF	M KRAS (G12S)	KRAS oncogenic mutation		Cetuximab (EGFR mAb inhibitor)	Colorectal adenocarcinoma
TCGA-AA-A00D	APC	M BRAF (V600E)	BRAF (V600E)		Cetuximab (EGFR mAb inhibitor)	Colorectal adenocarcinoma
play a menu 3999	CPT	TCGA-AA-A00D	M PIK3CA (H1047L)	PIK3CA oncogenic mutation	Cetuximab (EGFR mAb inhibitor)	Colorectal adenocarcinoma
		TCGA-AA-A00D	M BRAF (V600E)	BRAF (V600E)	Panitumumab (EGFR mAb inhibitor)	Colorectal adenocarcinoma
		TCGA-AA-A00D	M BRAF (V600E)	BRAF (.,G469A,V600E,D594G)	Cetuximab	Colorectal adenocarcinoma
		TCGA-AA-A00D	M BRAF (V600E)	BRAF (V600E,D594G)	Cetuximab + Panitumumab	Colorectal adenocarcinoma
		TCGA-AA-A00D	M BRAF (V600E)	BRAF (V600E)	Panitumumab + Cetuximab	Colorectal adenocarcinoma
		TCGA-AA-A00D	M BRAF (V600E)	BRAF (V600E)	Fluorouracil	Colorectal adenocarcinoma
		TCGA-AA-A00D	M BRAF (V600E)	BRAF (V600E)	Bevacizumab	Colorectal adenocarcinoma
		TCGA-AA-A00D	M BRAF (V600E)	BRAF (V600.,G596R)	Vemurafenib	Colorectal adenocarcinoma
		TCGA-AA-A00D	M PIK3CA (H1047L)	PIK3CA (E542K,E545K,H1047R,,)	Cetuximab	Colorectal adenocarcinoma

DETAILS ⓘ

en: (cc) BY-NC

## Recommended readings and software

- Martincorena, Iñigo, et al. "Universal patterns of selection in cancer and somatic tissues." *Cell* 171.5 (2017): 1029-1041.
- Rheinbay, Esther, et al. "Analyses of non-coding somatic drivers in 2,658 cancer whole genomes." *Nature* 578.7793 (2020): 102-111.
- Lawson, Andrew RJ, et al. "Extensive heterogeneity in somatic mutation and selection in the human bladder." *Science* 370.6512 (2020): 75-82
- Tamborero, David, et al. "Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations." *Genome medicine* 10.1 (2018): 1-8
- Gonzalez-Perez, Abel, et al. "IntOGen-mutations identifies cancer drivers across tumor types." *Nature methods* 10.11 (2013): 1081-1082.
- Software: [dndscv](#), [maf-tools](#), [IntOGen](#), [Cancer Genome Interpreter](#)

# Thank you!

- Q&A
- [fa8@sanger.ac.uk](mailto:fa8@sanger.ac.uk)