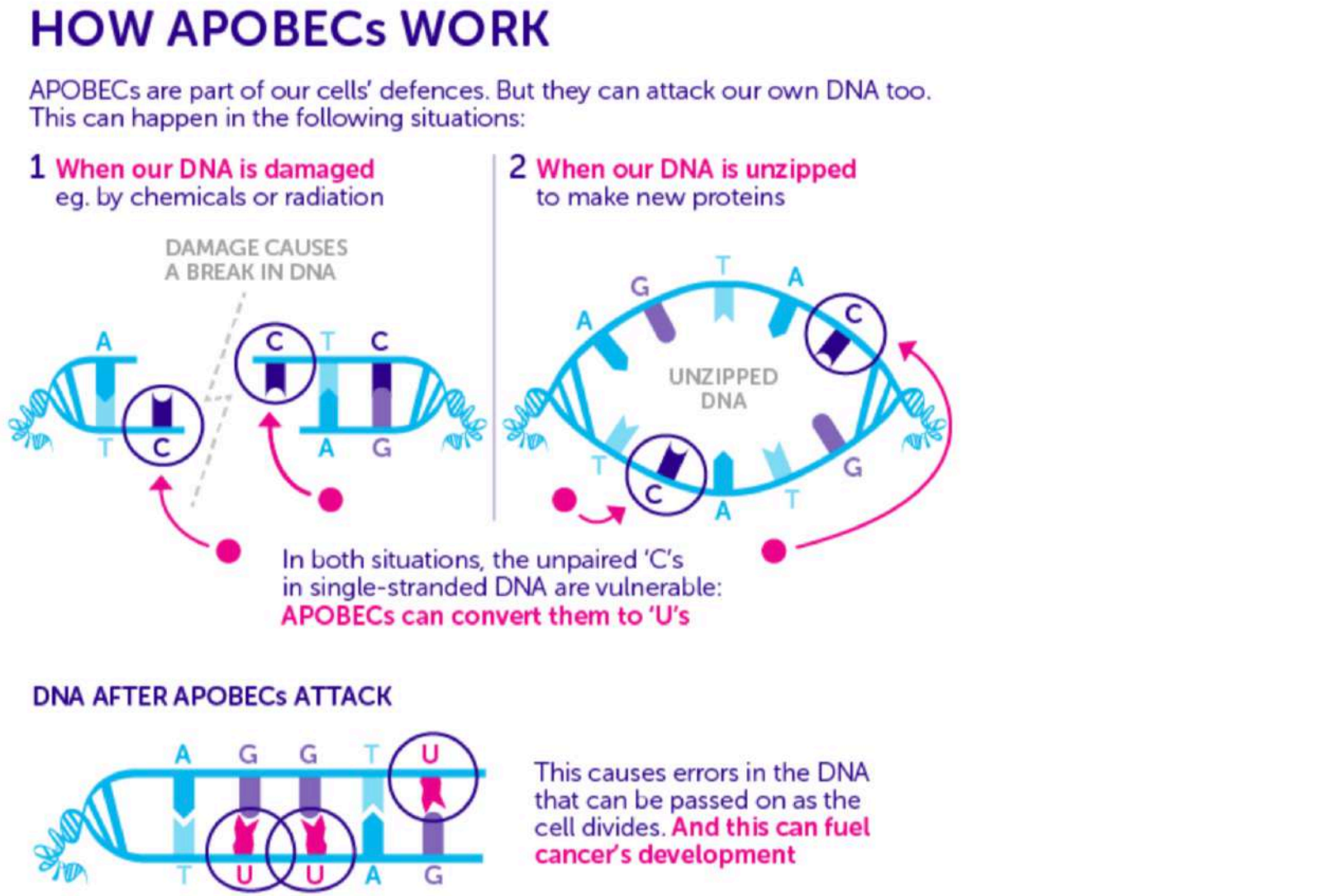
Analysis of 3D Localization of APOBEC-induced Mutations in the Cell Nucleus Using Hi-C Data

Marat Kazanov, Skolkovo Institute of Science and Technology

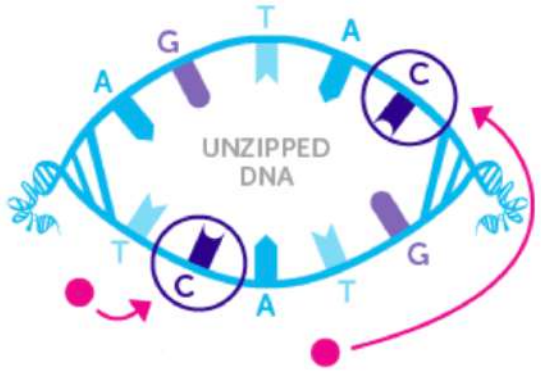
XXIX Symposium on Bioinformatics and Computer-Aided Drug Discovery, 2023

APOBEC family enzymes implicated in cancer mutagenesis



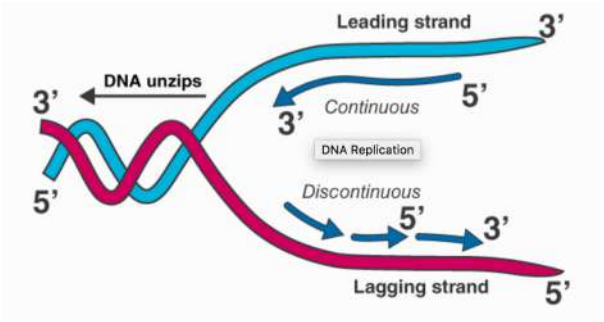
from Cancer Research UK Science Blog, 2015

APOBEC implicated in cancer mutagenesis



Some single-stranded DNA sources

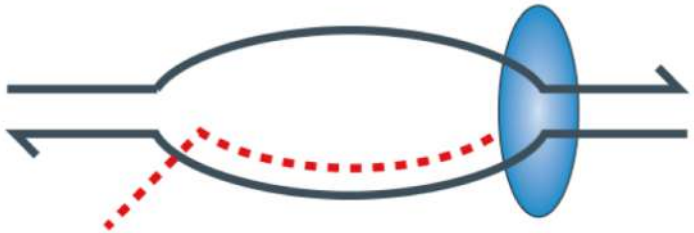
DNA replication



M.D. Kazanov et al., Cell Reports, 2016

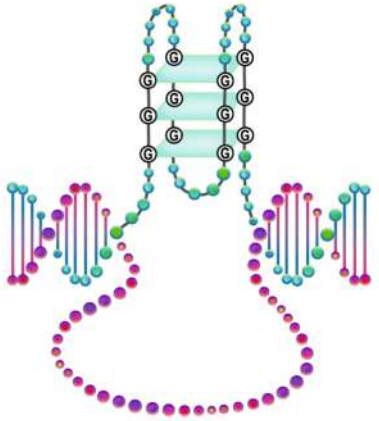
V.B. Seplyarskiy et al., Genome Res., 2016

Transcription (R-loops)



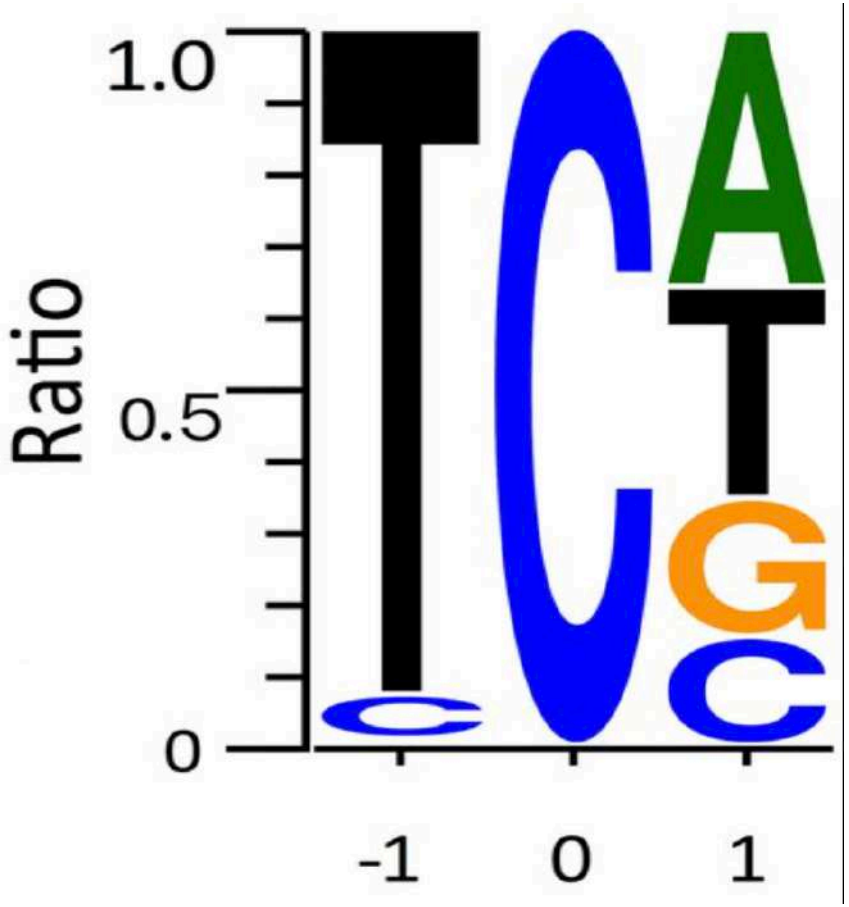
A.A. Chervova et al., NAR Cancer, 2021

Non-canonical DNA structures

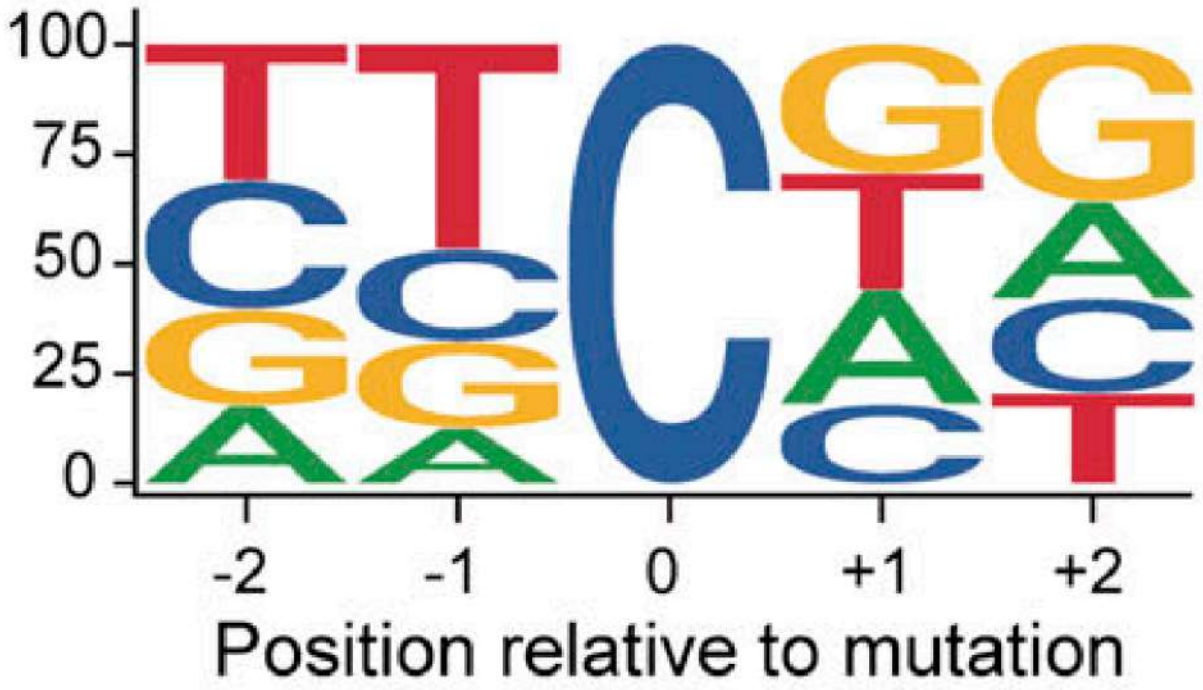


G.V.Ponomarev et al., iScience, 2022

APOBEC specificity

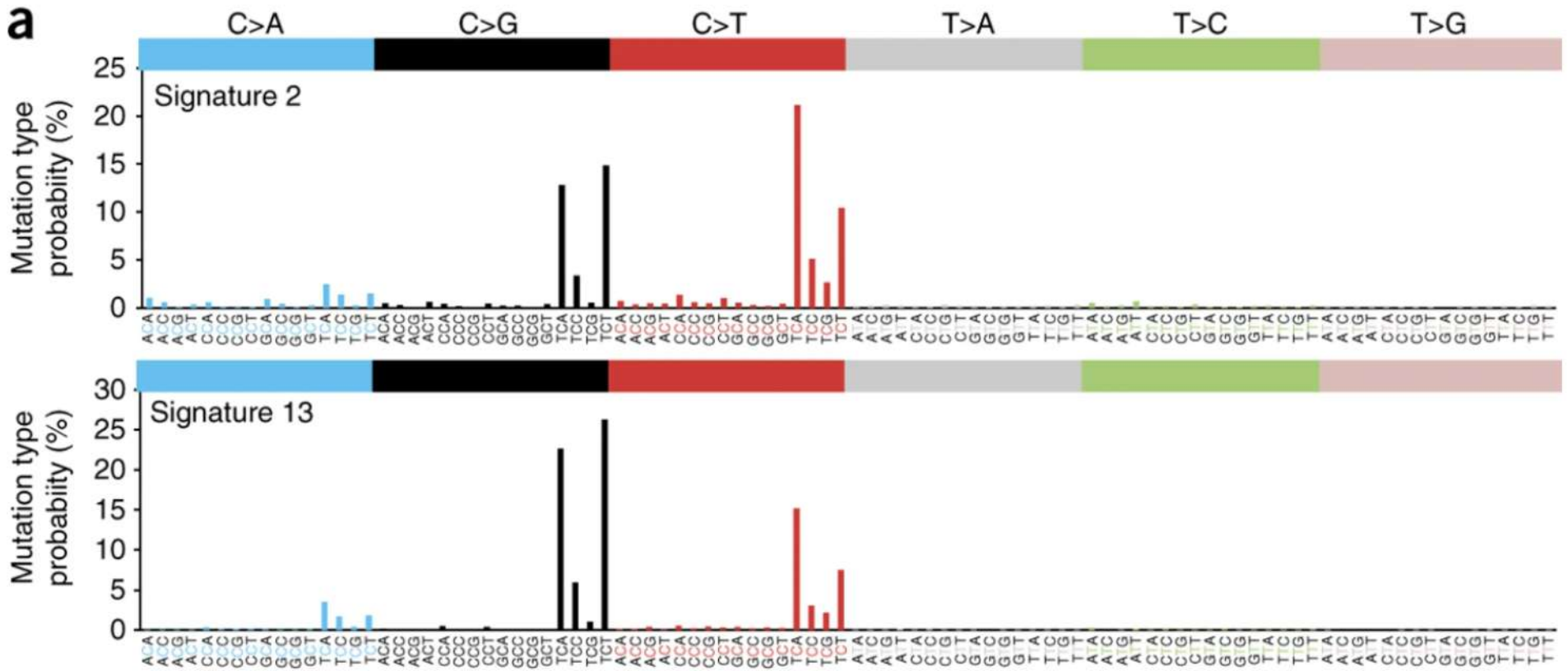


from Sui et al., PNAS, 2020



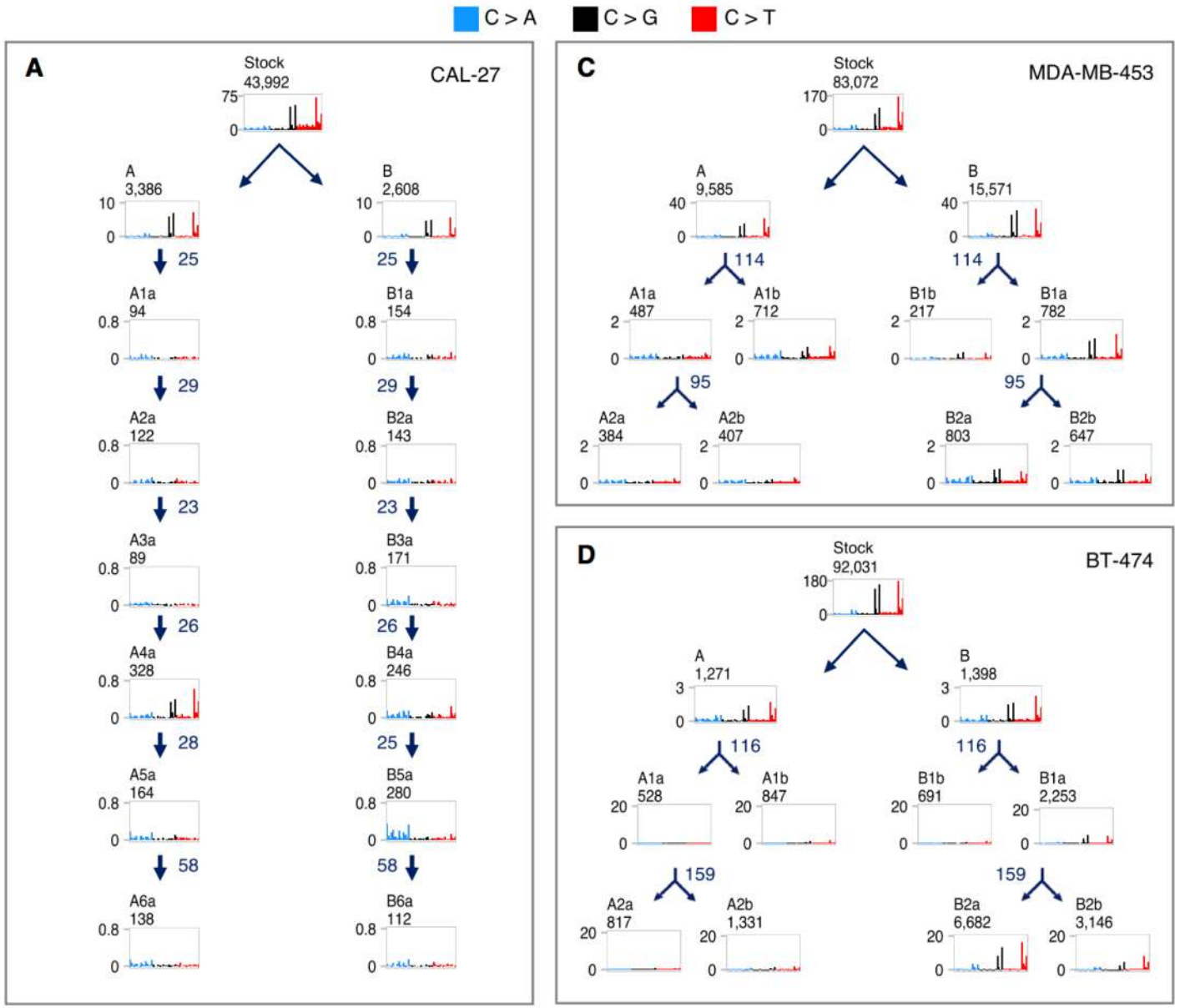
from Law et al, JEM, 2020

APOBEC mutational signature



from Nik-Zainal et al., Nature Genetics, 2014

APOBEC mutagenesis occurs in episodic bursts



from Petljak et al., Cell, 2023

APOBEC associated with drug resistance to anticancer therapy

Article

Therapy-induced APOBEC3A drives evolution of persistent cancer cells

https://doi.org/10.1038/s41586-023-06303-1

Received: 28 September 2020

Accepted: 8 June 2023

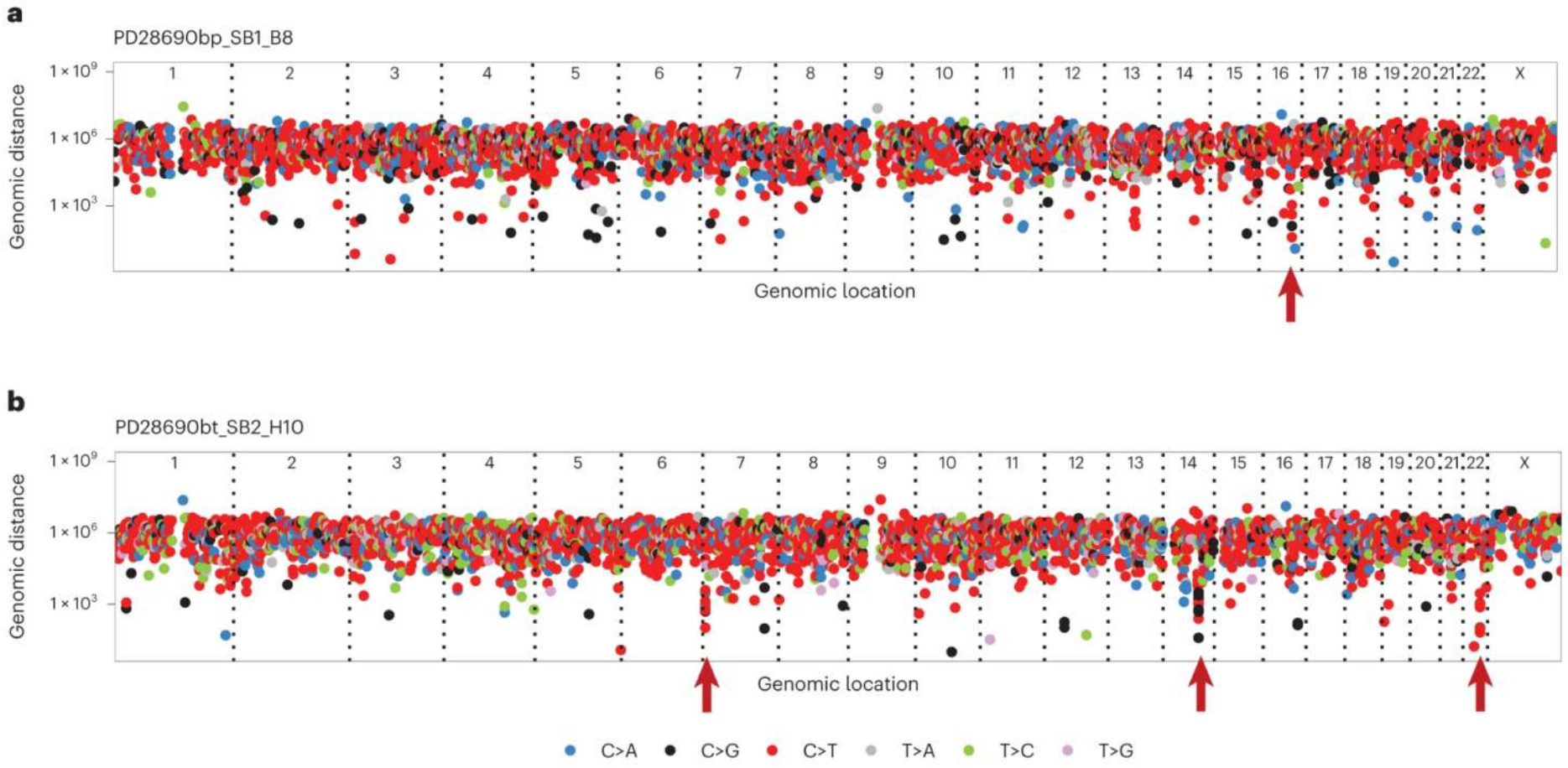
Published online: 05 July 2023

Check for updates

Hideko Isozaki[1,2]✉, Ramin Sakhtemani[1,3], Ammal Abbasi[1], Naveed Nikpour[1], Marcello Stanzione[1], Sunwoo Oh[4], Adam Langenbucher[1], Susanna Monroe[1], Wenjia Su[1], Heidie Frisco Cabanos[1,2], Faria M. Siddiqui[1], Nicole Phan[1], Pégah Jalili[4], Daria Timonina[1], Samantha Bilton[1], Maria Gomez-Caraballo[1], Hannah L. Archibald[1], Varuna Nangia[1], Kristin Dionne[1], Amanda Riley[1], Matthew Lawlor[1], Mandeep Kaur Banwait[1], Rosemary G. Cobb[1], Lee Zou[1,2,5], Nicholas J. Dyson[1,2], Christopher J. Ott[1,2,3], Cyril Benes[1,2], Gad Getz[1,3,6,7], Chang S. Chan[8], Alice T. Shaw[1,2], Justin F. Gainor[1,2], Jessica J. Lin[1,2], Lecia V. Sequist[1,2], Zofia Piotrowska[1,2], Beow Y. Yeap[1,2], Jeffrey A. Engelman[1,2], Jake June-Koo Lee[2], Yosef E. Maruvka[9], Rémi Buisson[4,10], Michael S. Lawrence[1,2,3,11]✉ & Aaron N. Hata[1,2,11]✉

Acquired drug resistance to anticancer targeted therapies remains an unsolved clinical problem. Although many drivers of acquired drug resistance have been identified[1–4], the underlying molecular mechanisms shaping tumour evolution during treatment are incompletely understood. Genomic profiling of patient tumours has implicated apolipoprotein B messenger RNA editing catalytic polypeptide-like (APOBEC) cytidine deaminases in tumour evolution; however, their role during therapy and the development of acquired drug resistance is undefined. Here we report that lung cancer targeted therapies commonly used in the clinic can induce cytidine deaminase APOBEC3A (A3A), leading to sustained mutagenesis in drug-tolerant cancer cells persisting during therapy. Therapy-induced A3A promotes the formation of double-strand DNA breaks, increasing genomic instability in drug-tolerant persisters. Deletion of A3A reduces APOBEC mutations and structural variations in persister cells and delays the development of drug resistance. APOBEC mutational signatures

Clusters of APOBEC mutations (kataegis)

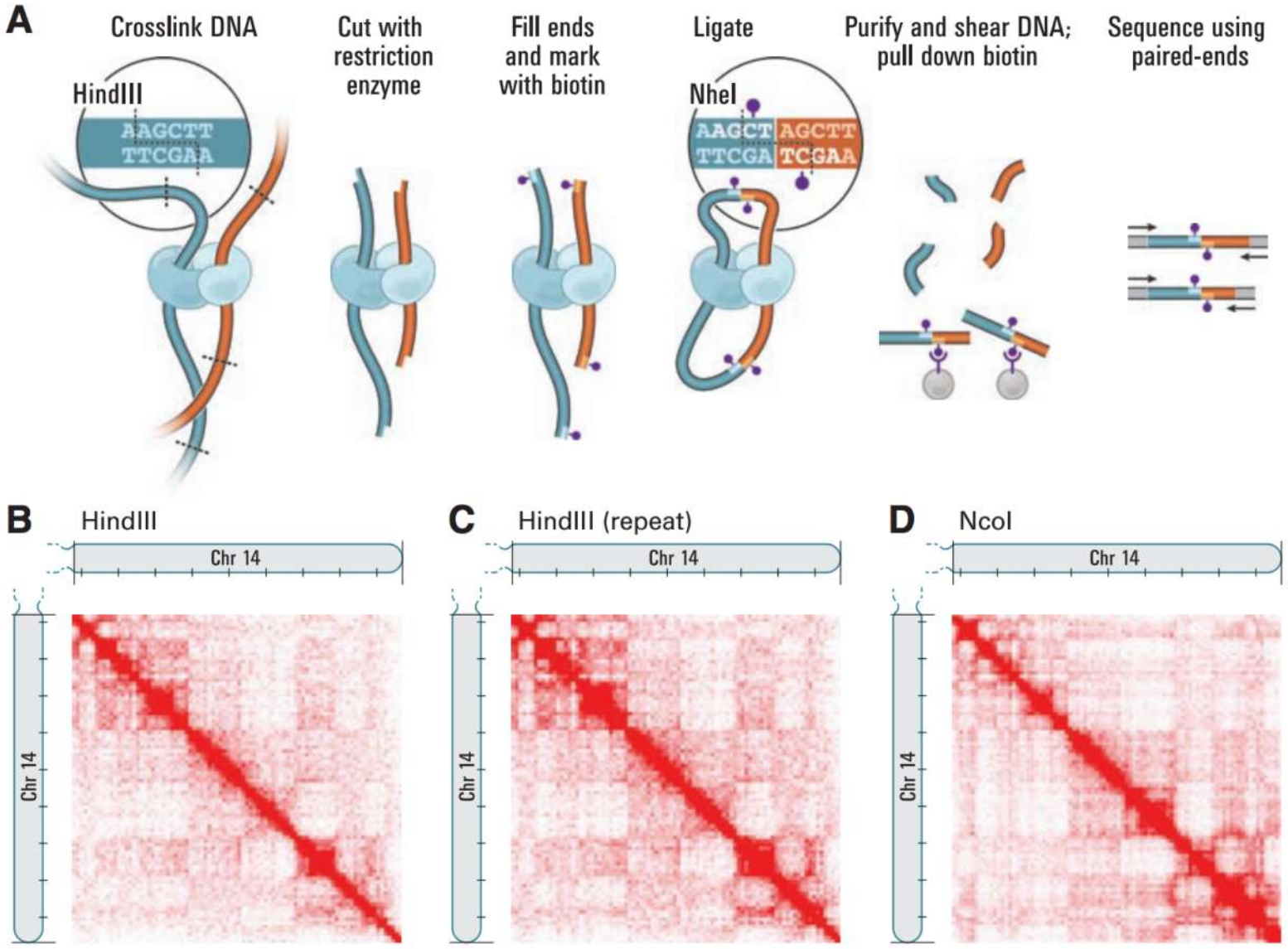


a,**b**, Two examples of crypts with kataegis. Types of SBSs are indicated by the color code below. Red arrowheads show the location of kataegis. **a**, Rainfall plot of a crypt from PD28690 showing a mutation cluster on chromosome 16. **b**, Rainfall plot of a crypt from PD28690 showing mutation clusters on chromosomes 7, 14 and 22.

Project aim

- Analyze 3D distribution of the APOBEC-induced mutations inside the cell nucleus using 3D chromatin models and mutational data from cancer samples

What is Hi-C data?



from Lieberman-Aiden et al, Science, 2012

Methods

- Data: Somatic mutations 3D chromatin organization

- APOBEC-induced mutation density:

$$D = \frac{Number\ of\ C \rightarrow T\ \&\ C \rightarrow G\ mutations\ in\ TCN\ context}{Number\ of\ TCN\ motifs}$$

- Estimation of APOBEC activity in a cancer sample:

$$APOBEC\ Enrichment = \frac{D}{Number\ of\ C \rightarrow T\ \&\ C \rightarrow G\ mutaions\ /\ Number\ of\ C}$$

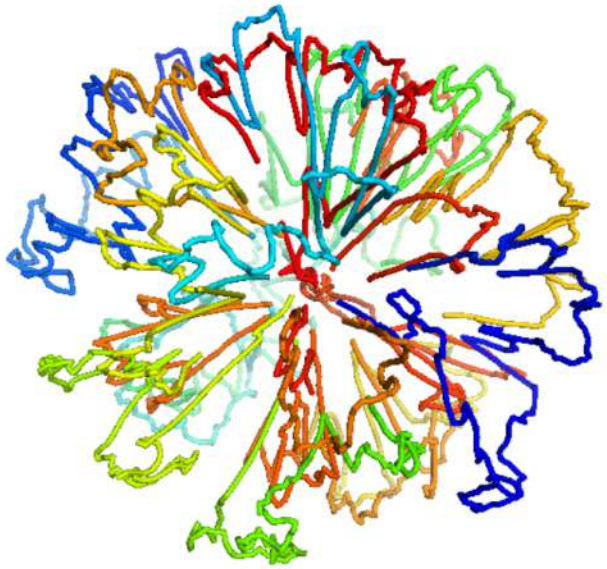
(Roberts, Gordenin, et al. 2012-2021)

Methods

- Mutational data from PCAWG project: six cancer types known for APOBEC activity were considered:
  - breast invasive carcinoma (BRCA),
  - lung adenocarcinoma (LUAD),
  - lung squamous cell carcinoma (LUSC),
  - head and neck squamous cell carcinoma (HNSC),
  - cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), bladder urothelial carcinoma (BLCA),
  - esophageal carcinoma (ESCA).
- Samples with high APOBEC activity were selected using enrichment metric (APOBEC enrichment > 2.0)
- 1D mutational clusters of APOBEC-induced mutations were detected and excluded from the analysis

3D genome structural models from Hi-C data

GSDB : Genome Structure Database



GSDB is a database of three-dimensional (3D) chromosome and genome structures reconstructed from Hi-C data by multiple 3D modeling tools in the field. In recent years, numerous Hi-C datasets have been generated and many chromosome/genome structure construction algorithms have been developed. However, there is no public repository of 3D chromosome and genome structures available for the community to use.

GSDB aims to fill this gap by providing a comprehensive repository of 3D chromosome and genome structures reconstructed from for Hi-C data. We hope this database will enable the exploration of the dynamic architectures of chromosomes and genomes for biomedical research.

Over 50,000 structures from 12 start-of-the-art Hi-C data structure prediction algorithms for 32 Hi-C datasets each containing varying resolutions.

Get Started

Methods

- 3D model allows to calculate distance between pair of mutations:

$$d = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}$$

- We estimated tendency of mutations to form clusters in 3D using Hopkins statistic:

Let X be the set of n data points.

Generate a random sample $\widetilde{X}$ of $m \ll n$ data points sampled without replacement from X.

Generate a set Y of m uniformly randomly distributed data points.

Define two distance measures,

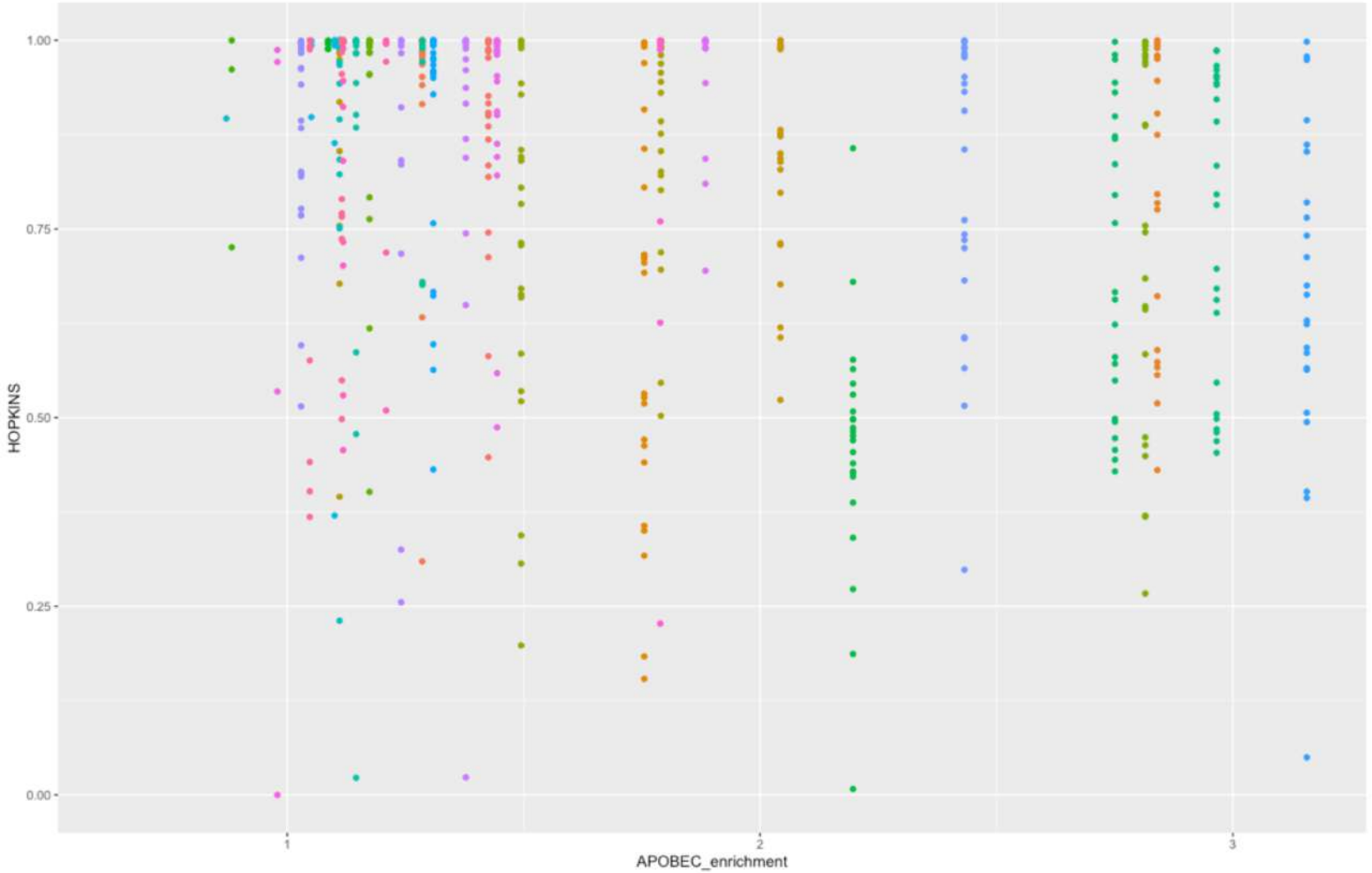
u_i, the minimum distance (given some suitable metric) of $y_i \in Y$ to its nearest neighbour in X, and

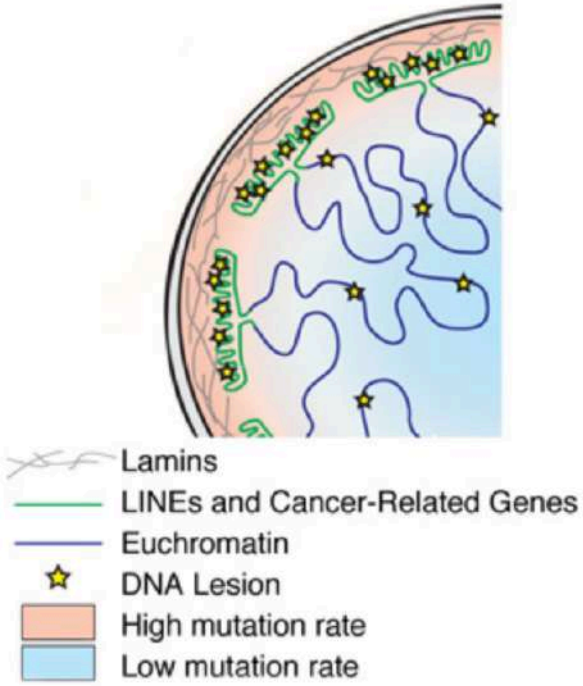
w_i, the minimum distance of $\widetilde{x}_i \in \widetilde{X} \subseteq X$ to its nearest neighbour $x_j \in X$, $\widetilde{x}_i \neq x_j$.

With the above notation, if the data is d dimensional, then the Hopkins statistic is defined as:

$$H = \frac{\sum_{i=1}^{m} u_i^d}{\sum_{i=1}^{m} u_i^d + \sum_{i=1}^{m} w_i^d}$$

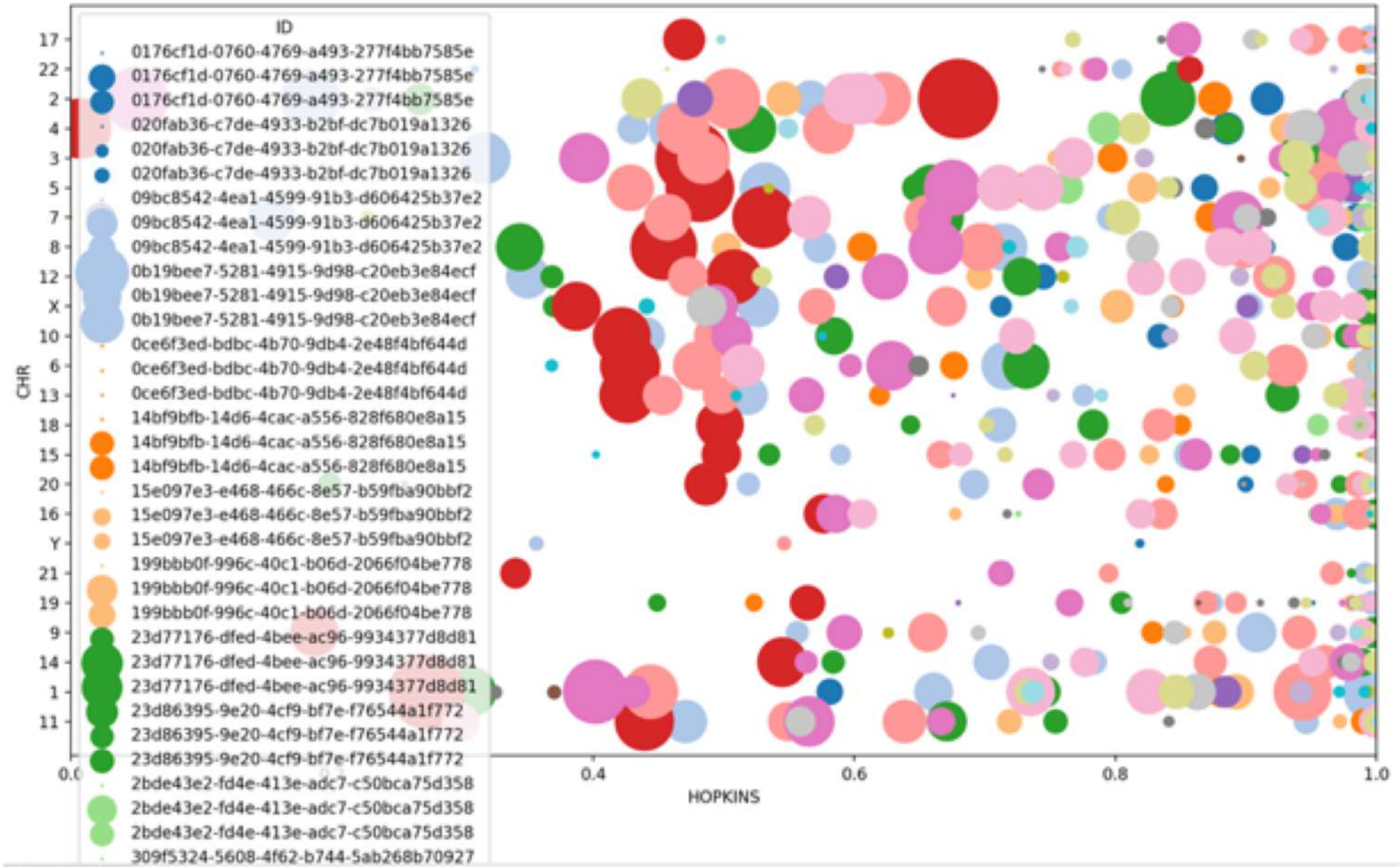
Results





Adapted from P.E. Garcia-Nieto, The EMBO journal 36, 2017

Results


ID
0176cf1d-0760-4769-a493-277f4bb7585e
0176cf1d-0760-4769-a493-277f4bb7585e
0176cf1d-0760-4769-a493-277f4bb7585e
020fab36-c7de-4933-b2bf-dc7b019a1326
020fab36-c7de-4933-b2bf-dc7b019a1326
020fab36-c7de-4933-b2bf-dc7b019a1326
09bc8542-4ea1-4599-91b3-d606425b37e2
09bc8542-4ea1-4599-91b3-d606425b37e2
09bc8542-4ea1-4599-91b3-d606425b37e2
0b19bee7-5281-4915-9d98-c20eb3e84ecf
0b19bee7-5281-4915-9d98-c20eb3e84ecf
0b19bee7-5281-4915-9d98-c20eb3e84ecf
0ce6f3ed-bdbc-4b70-9db4-2e48f4bf644d
0ce6f3ed-bdbc-4b70-9db4-2e48f4bf644d
0ce6f3ed-bdbc-4b70-9db4-2e48f4bf644d
14bf9bfb-14d6-4cac-a556-828f680e8a15
14bf9bfb-14d6-4cac-a556-828f680e8a15
14bf9bfb-14d6-4cac-a556-828f680e8a15
15e097e3-e468-466c-8e57-b59fba90bbf2
15e097e3-e468-466c-8e57-b59fba90bbf2
15e097e3-e468-466c-8e57-b59fba90bbf2
199bbb0f-996c-40c1-b06d-2066f04be778
199bbb0f-996c-40c1-b06d-2066f04be778
199bbb0f-996c-40c1-b06d-2066f04be778
23d77176-dfed-4bee-ac96-9934377d8d81
23d77176-dfed-4bee-ac96-9934377d8d81
23d77176-dfed-4bee-ac96-9934377d8d81
23d86395-9e20-4cf9-bf7e-f76544a1f772
23d86395-9e20-4cf9-bf7e-f76544a1f772
23d86395-9e20-4cf9-bf7e-f76544a1f772
2bde43e2-fd4e-413e-adc7-c50bca75d358
2bde43e2-fd4e-413e-adc7-c50bca75d358
2bde43e2-fd4e-413e-adc7-c50bca75d358
309f5324-5608-4f62-b744-5ab268b70927
CHR
17
22
2
4
3
5
7
8
12
X
10
6
13
18
15
20
16
Y
21
19
9
14
1
11
0.0
0.4
0.6
0.8
1.0
HOPKINS

Acknowledgements

Gennady Ponomarev

Alexandra Denisova