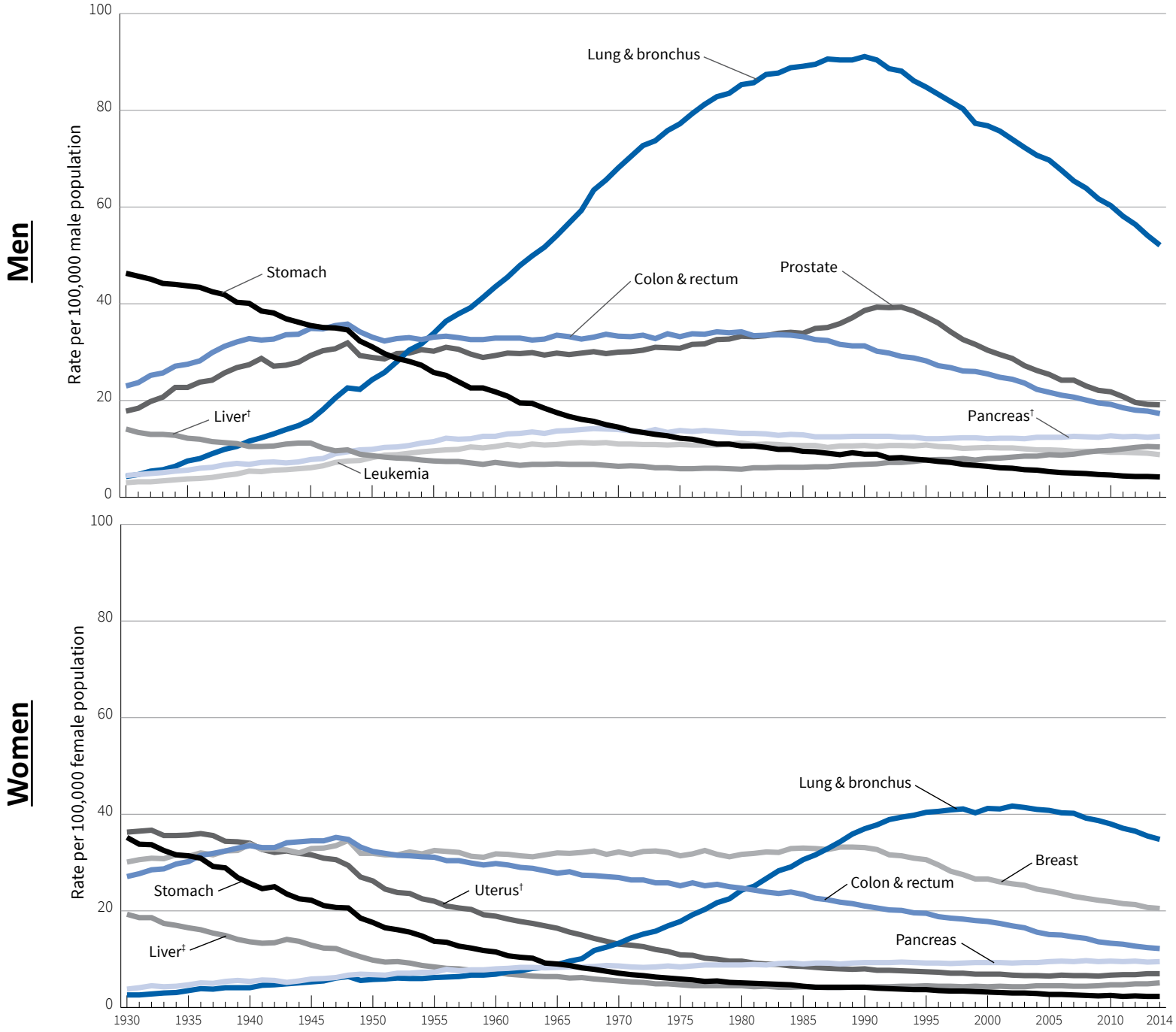
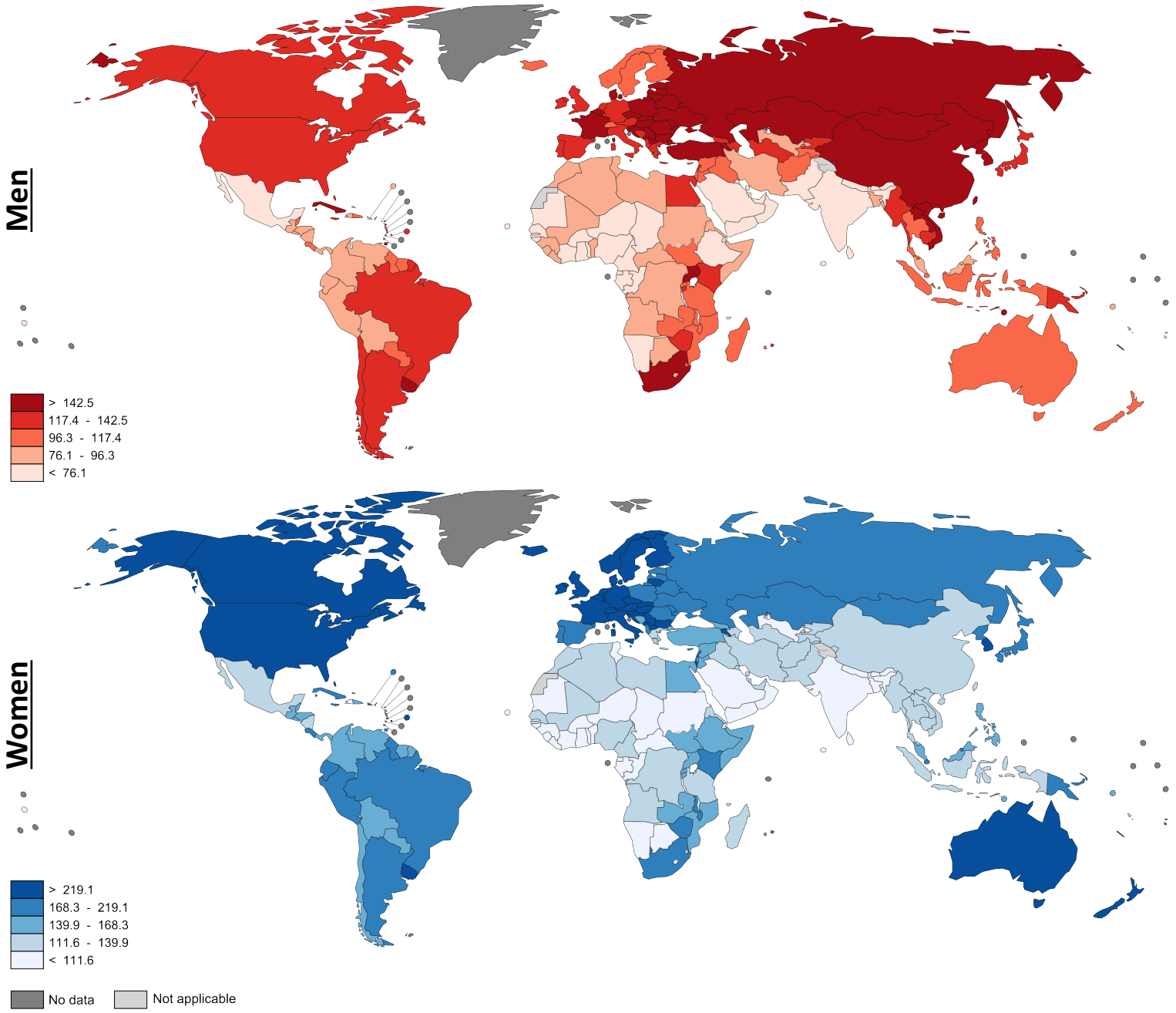
Cancer Genomics
Intro Slides

Epidemiology of cancer-related mortality rates in the United States (1930-2014)



Source: ACS, US Mortality Volumes 1930 to 1959, US Mortality Data 1960 to 2014, National Center for Health Statistics, Centers for Disease Control and Prevention.

Estimated Cancer Incidence Worldwide (2012)



Source: World Health Organization

Goals of the The Cancer Moonshot Initiative

Accelerate progress in cancer, including prevention & screening

- From cutting-edge basic research to wider uptake of standard care

Encourage greater cooperation and collaboration

- Break down silos within and between academia, government, and private sector

Enhance data sharing

- Accessible data that is universally usable
- Annotated patient-level clinical and –omics data

- Adapted from Presidential Memo and Factsheet, 2016

Summary of the 10 Recommendations

A. Establish a network for **direct patient involvement**
B. Create a translational science network devoted to **immunotherapy**
C. Develop ways to overcome **resistance to therapy**
D. Build a national cancer **data ecosystem**
E. Intensify research of the major drivers of **childhood cancer**
F. Minimize cancer treatment's debilitating **side effects**
G. Expand use of proven **prevention and early detection** strategies
H. Mine past patient data to predict future **patient outcomes**
I. Develop a 3D **cancer atlas**
J. Develop new cancer **technologies**

Summary of the 10 Recommendations

A. Establish a network for direct patient involvement

B. Create a translational science network devoted to immunotherapy

C. Develop ways to overcome resistance to therapy

D. Build a national cancer data ecosystem

E. Intensify research of the major drivers of childhood cancer

F. Minimize cancer treatment's debilitating side effects

G. Expand use of proven prevention and early detection strategies

H. Mine past patient data to predict future patient outcomes

I. Develop a 3D cancer atlas

J. Develop new cancer technologies

Intensify research of the major drivers of childhood cancer

- Goal – Accelerate the development of new therapies that target these cancer-causing proteins

- **Fusion oncoproteins**
  - Enhance understanding of molecular and biochemical mechanisms of transformation driven by fusion oncoproteins
  - Develop faithful models
  - Identify key dependencies

Develop a 3D cancer atlas

- Goal – Enable predictive models of tumor progression and response to treatment

- **Generation of human tumor atlases**
  - Adult and pediatric cancers
  - From tumor development through metastasis
  - Immune cell characterization and other cells in the microenvironment
  - Premalignant lesions to create a Pre-Cancer Genome Atlas (PCGA)

Precision Medicine Initiative

"To enable a new era of medicine through research, technology, and policies that empower patients, researchers, and providers to work together toward development of individualized care."

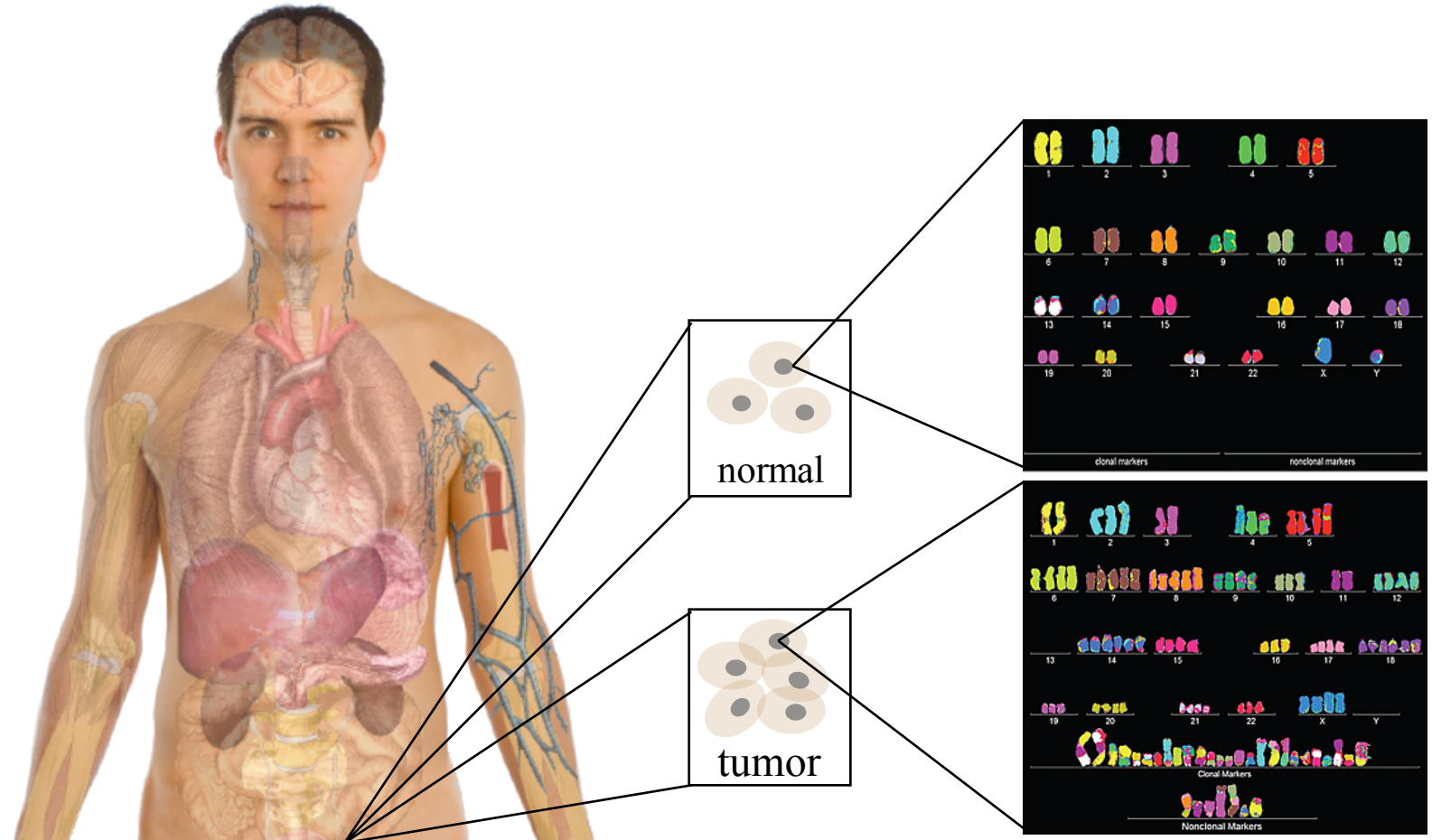
- Barak Obama

THE PRECISION MEDICINE INITIATIVE

$215 million was invested (NIH fiscal year 2016) to accelerate research for selecting the therapies that can be used in more individualized ways. NCI used $70 million of that investment to advance the precision oncology.

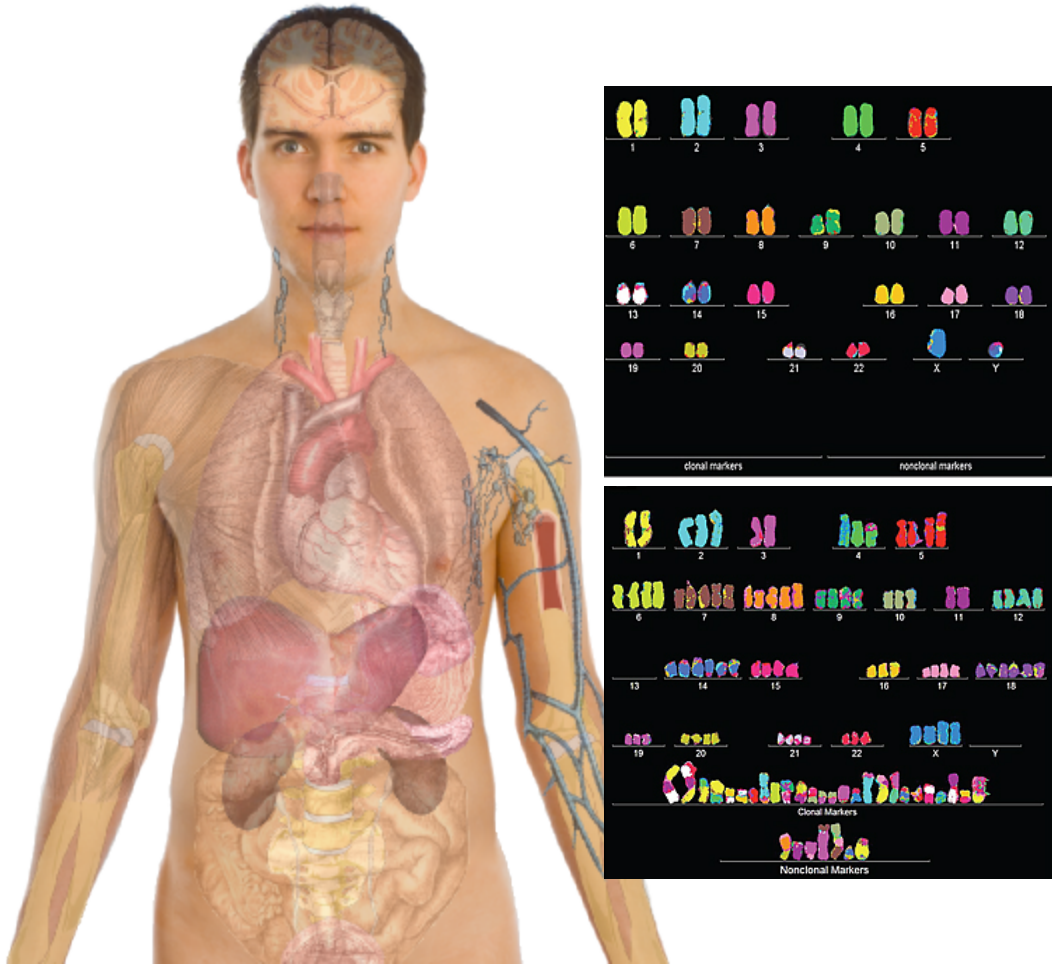
Personal Genomics as a Gateway into Biology

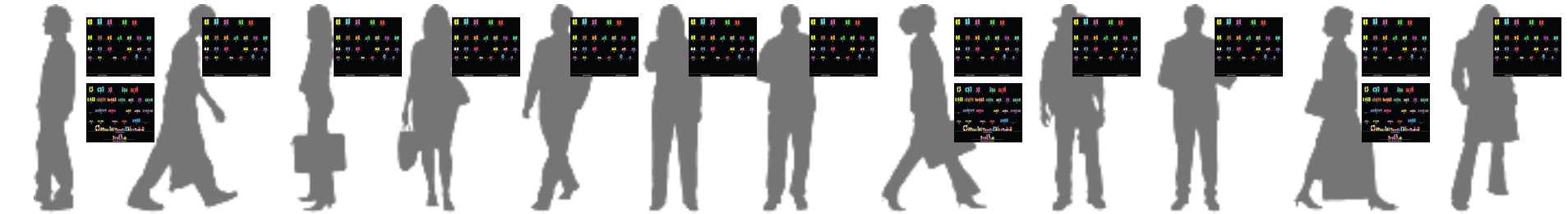
Personal genomes soon will become a commonplace part of medical research & eventually treatment (esp. for cancer). They will provide a primary connection for biological science to the general public.

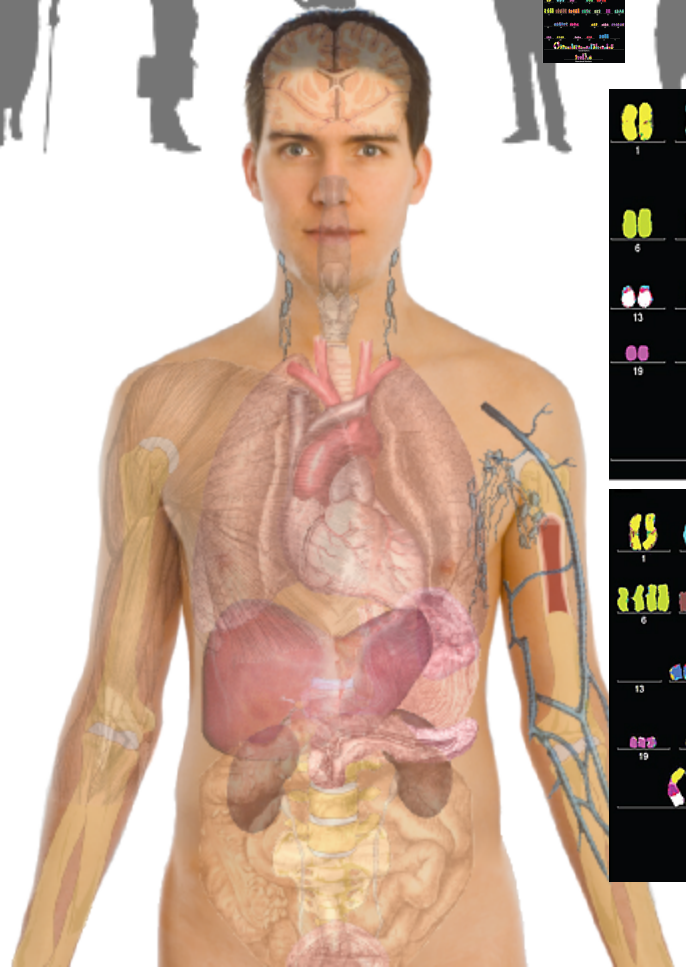


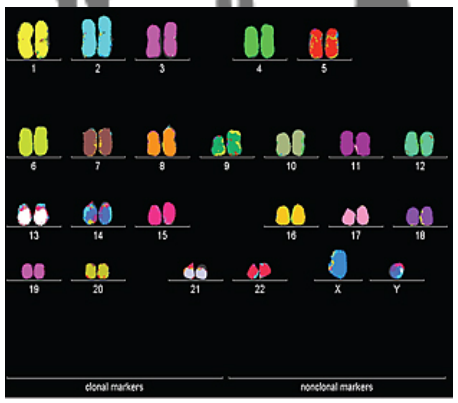
Personal Genomics as a Gateway into Biology

Personal genomes soon will become a commonplace part of medical research & eventually treatment (esp. for cancer). They will provide a primary connection for biological science to the general public.



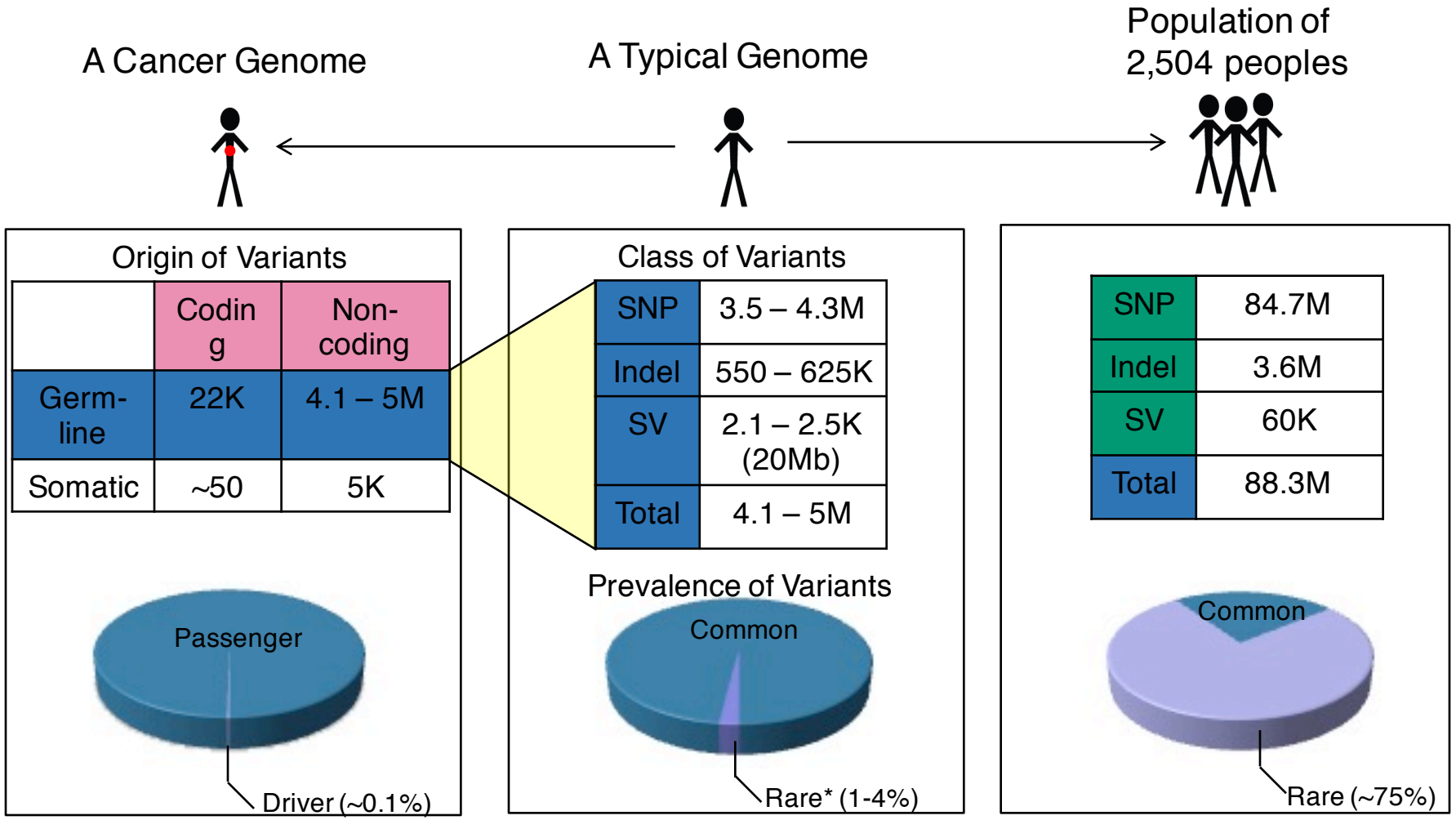





clonal markers
nonclonal markers


Clonal Markers
Nonclonal Markers

Human Genetic Variation

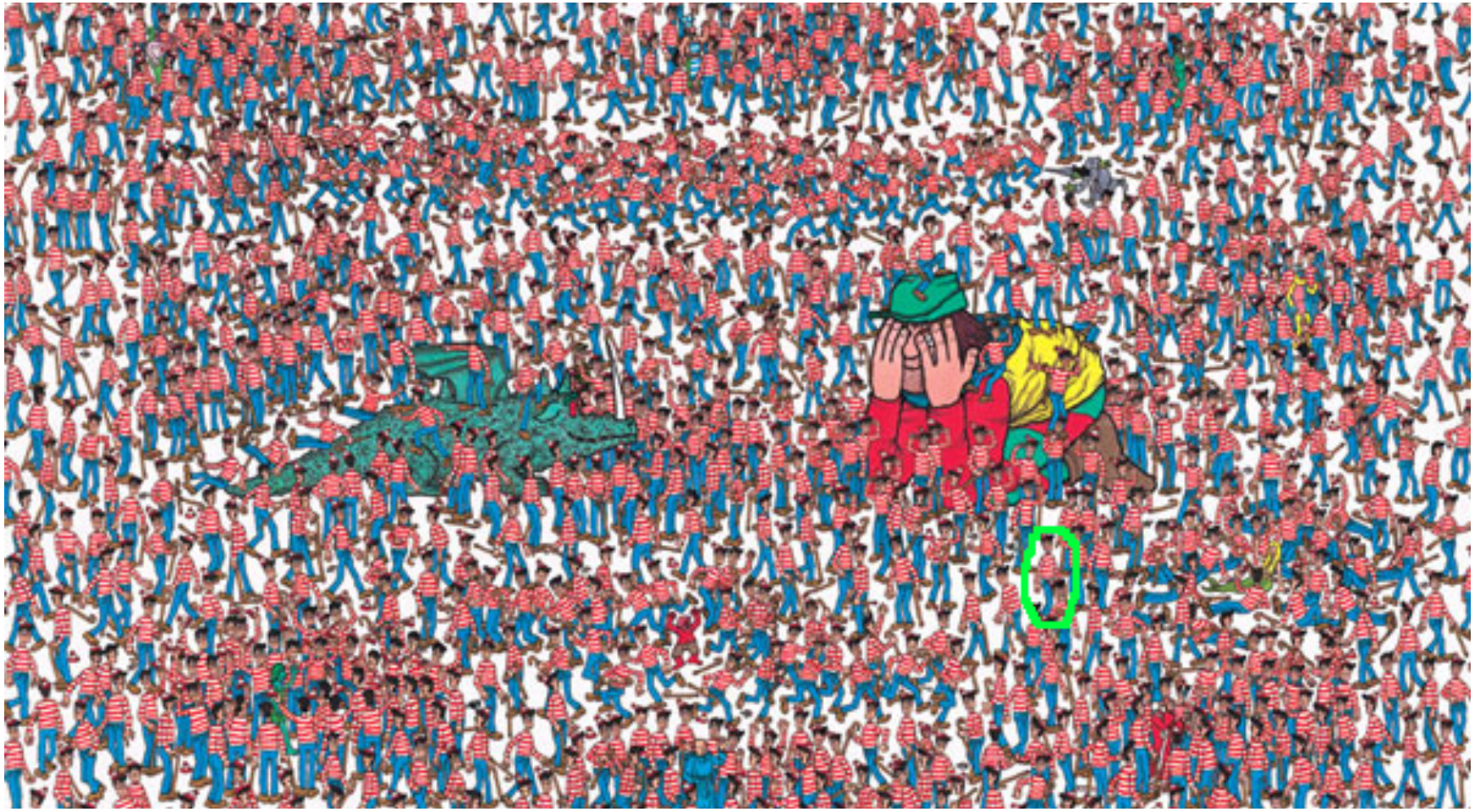


* Variants with allele frequency < 0.5% are considered as rare variants in 1000 genomes project.

The 1000 Genomes Project Consortium, Nature. 2015. 526:68-74
Khurana E. et al. Nat. Rev. Genet. 2016. 17:93-108

Cancer drivers: Significance & identification

(Finding the key mutations in ~3M Germline variants &
~5K Somatic Variants in a Tumor Sample)



Finding Key Variants

Germline

CAN YOU FIND THE PANDA?

- **Common variants**
  - Can be associated with phenotype (ie disease) via a Genome-wide Association Study (GWAS), which tests whether the frequency of alleles differs between cases & controls.
  - Usually their functional effect is weaker.
  - Many are non-coding
  - Issue of LD in identifying the actual causal variant.
- **Rare variants**
  - Associations are usually underpowered due to low frequencies.
  - They often have larger functional impact
  - Can be collapsed in the same element to gain statistical power (burden tests).
  - In some cases, causal variants can be identified through tracing inheritance of Mendelian subtypes of diseases in large families.

McCarthy, M. et al. Nat. Rev. Genet. 2008. 9, 356-369, Zuk, O. et al. PNSA. 2014. Vol. 11, no. 4, MacArthur DG et al. Nature 2014. 508:469-476

Cancer drivers: Significance & identification

Finding Key Variants

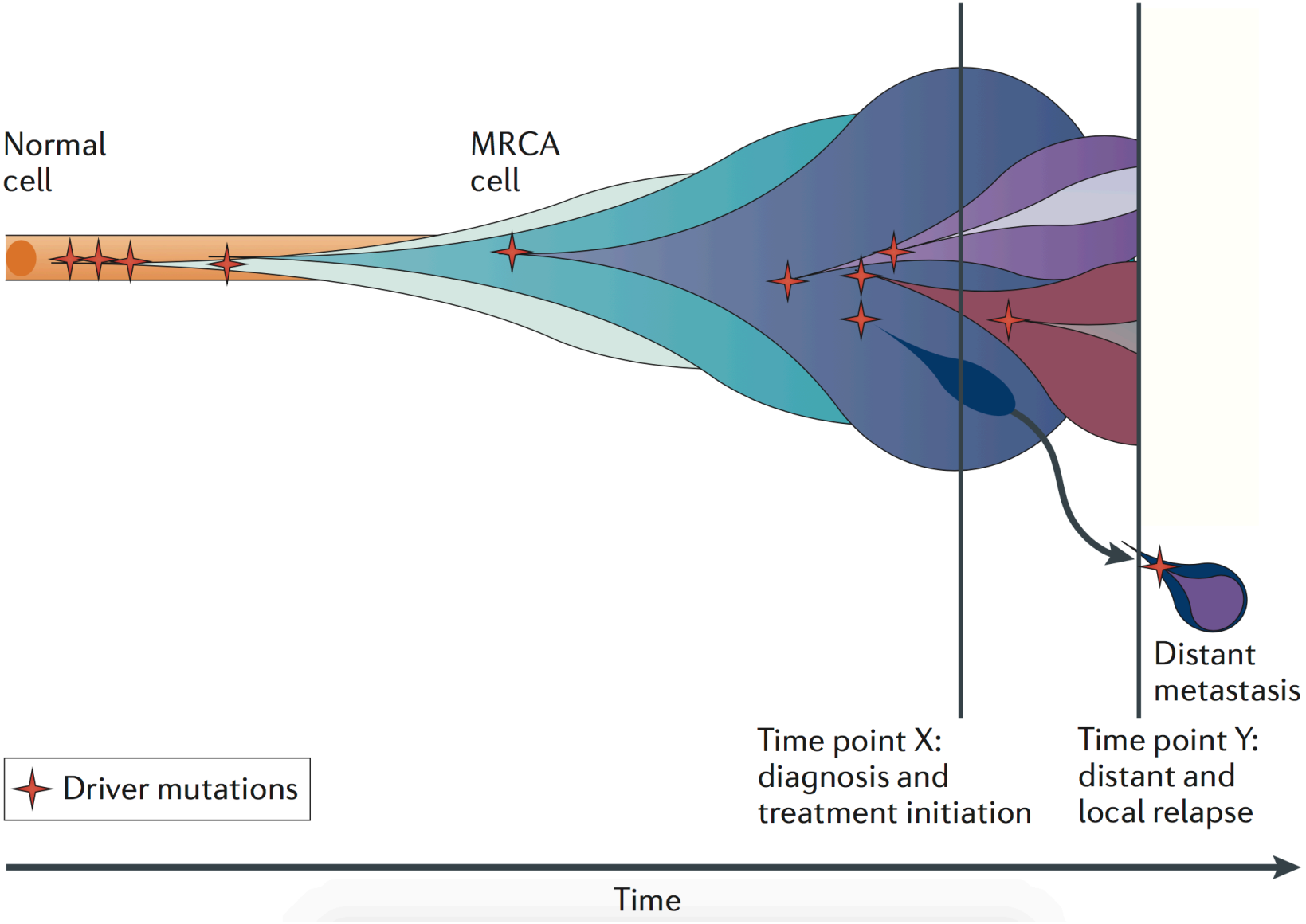
Somatic

CAN YOU FIND THE PANDA?

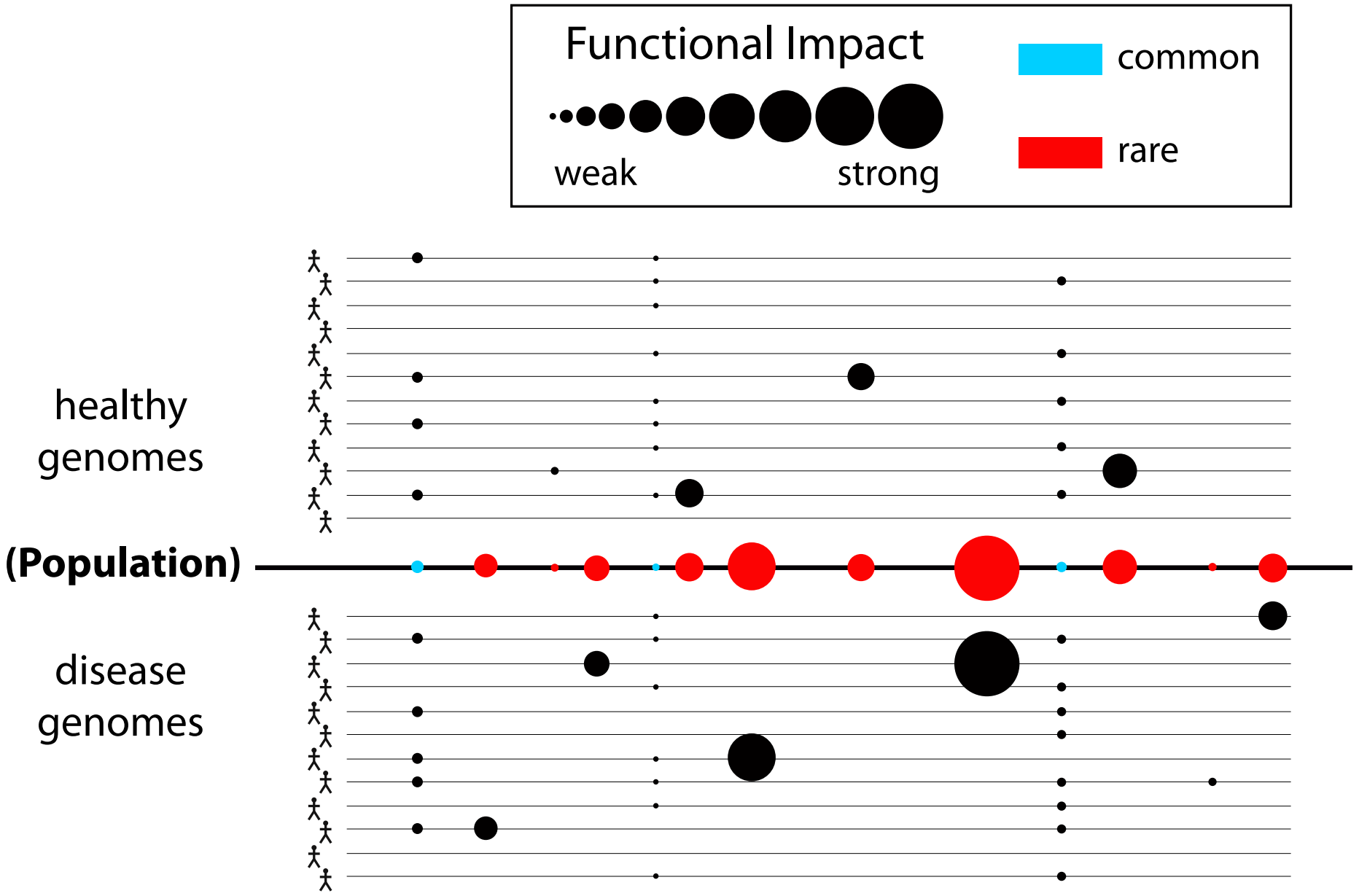
- **Overall**
  - Often these can be conceptualized as very rare variants
  - A challenge to identify somatic mutations contributing to cancer is to find driver mutations & distinguish them from passengers.
- **Drivers**
  - Driver mutation is a mutation that directly or indirectly confers a selective growth advantage to the cell in which it occurs.
  - A typical tumor contains 2-8 drivers; the remaining mutations are passengers.
- **Passengers**
  - Conceptually, a passenger mutation has no direct or indirect effect on the selective growth advantage of the cell in which it occurred.

Vogelstein B. Science 2013. 339(6127):1546-1558

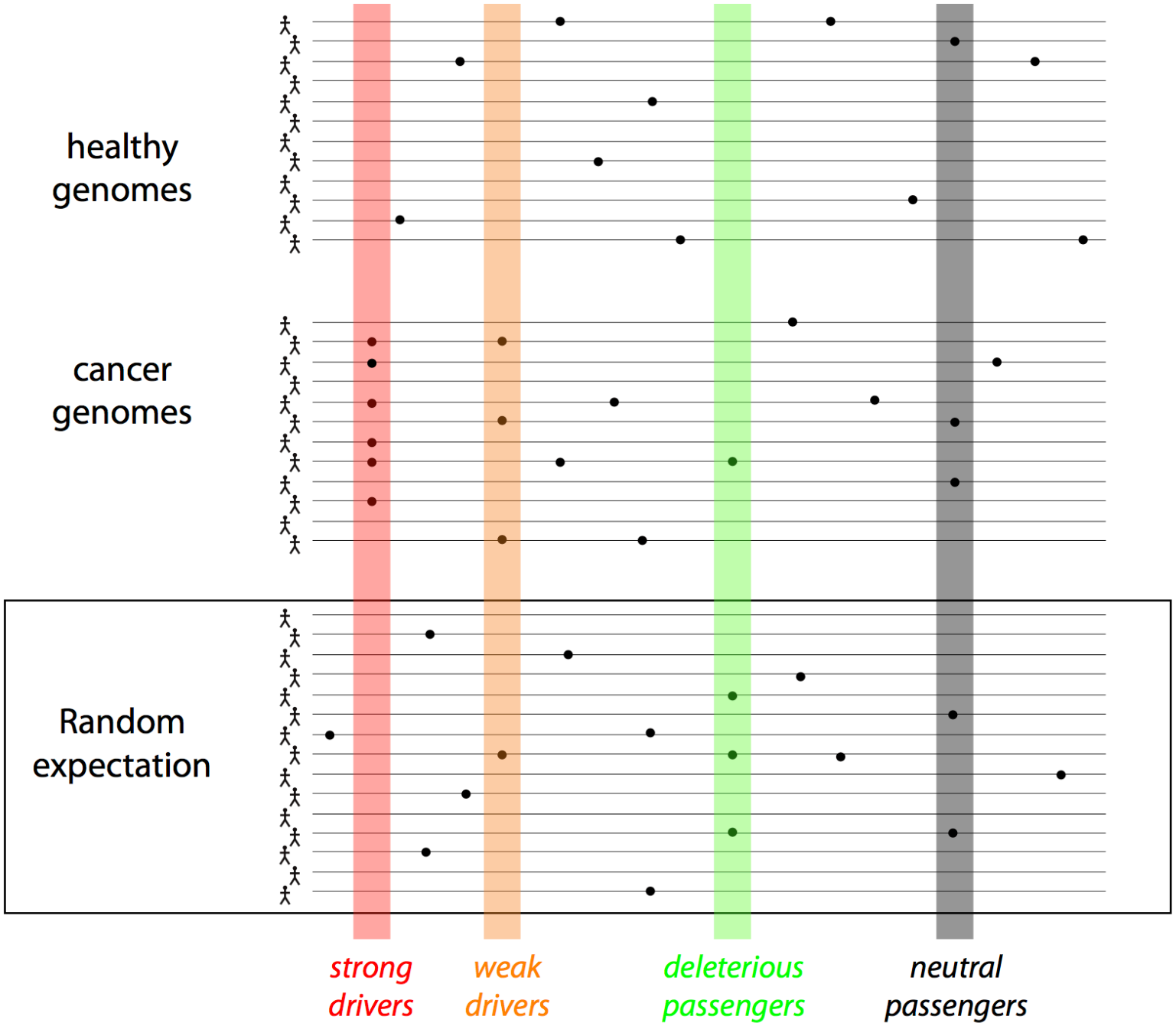
Cancer drivers: Significance & identification



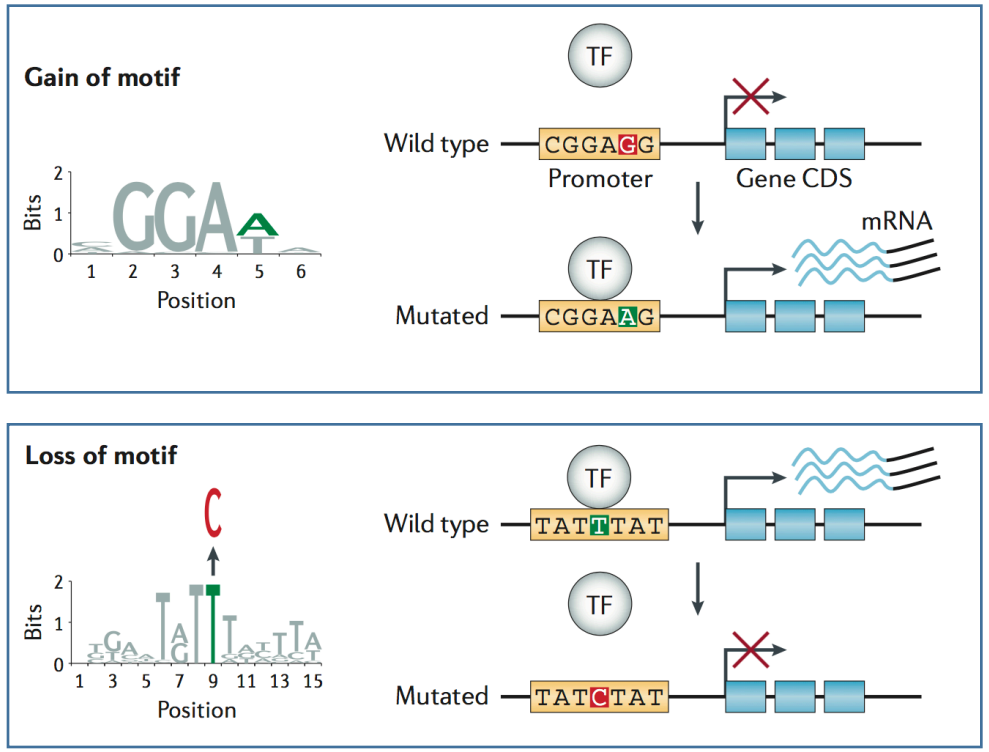
Yates et al, NRG (2012)

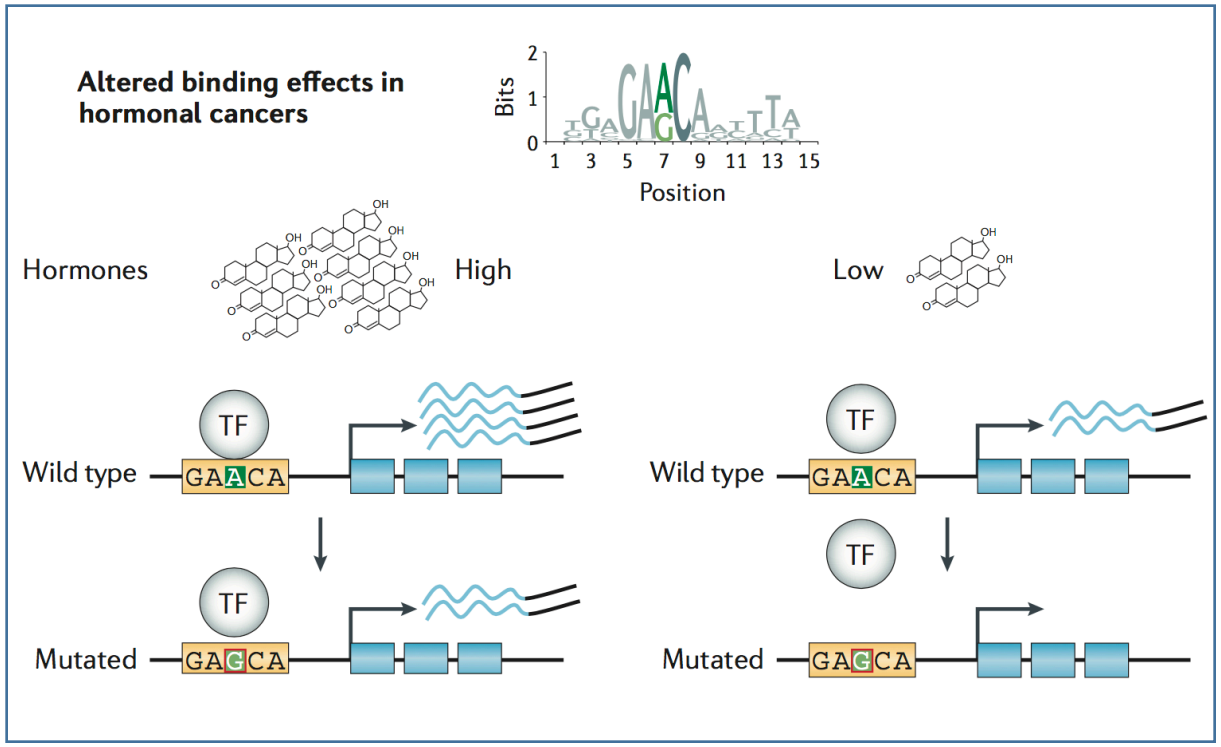

Functional Impact
weak
strong
common
rare
healthy genomes
(Population)
disease genomes

Cancer drivers: Significance & identification: frequency-based approaches



Cancer drivers: Significance & identification: functional annotations

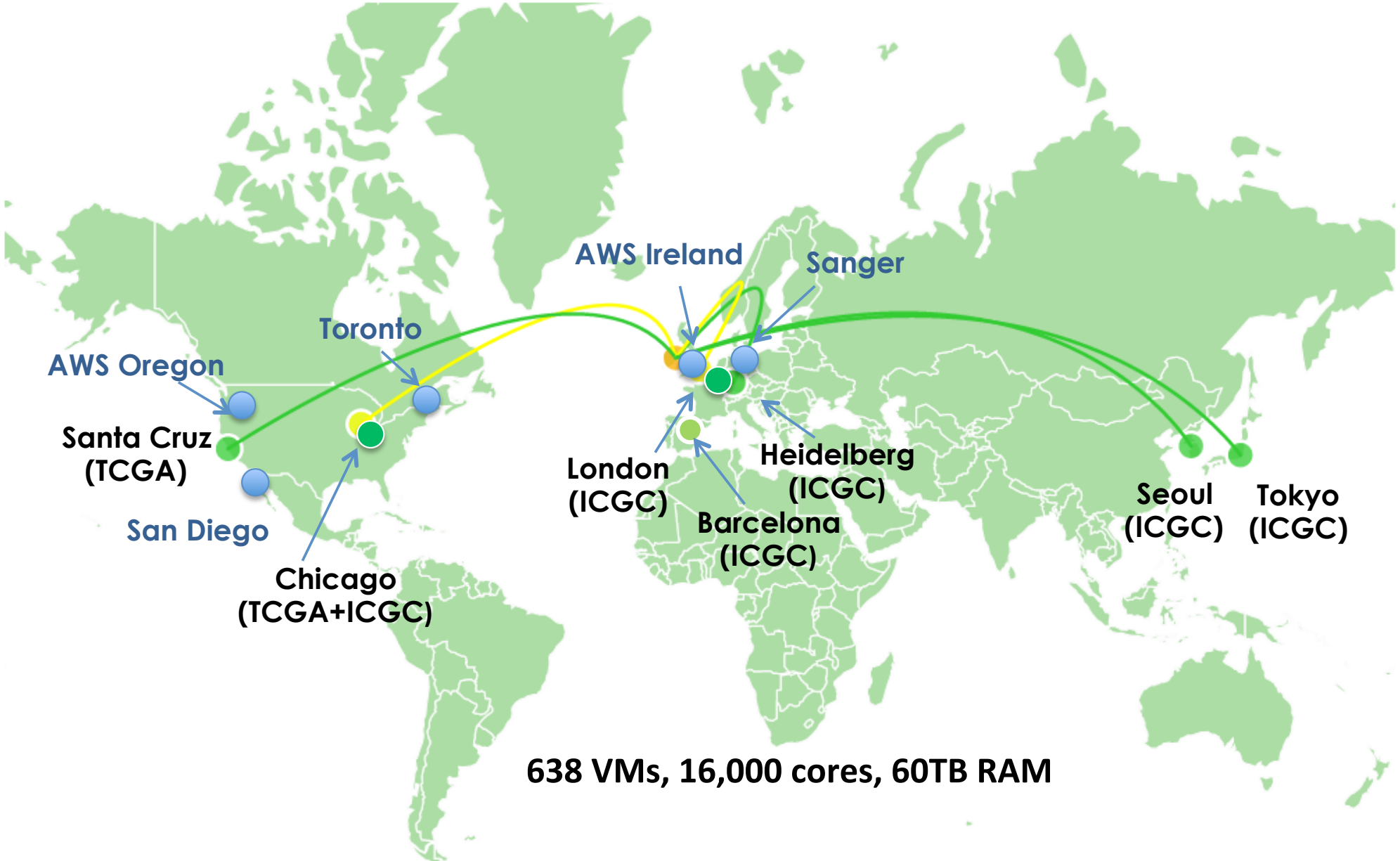




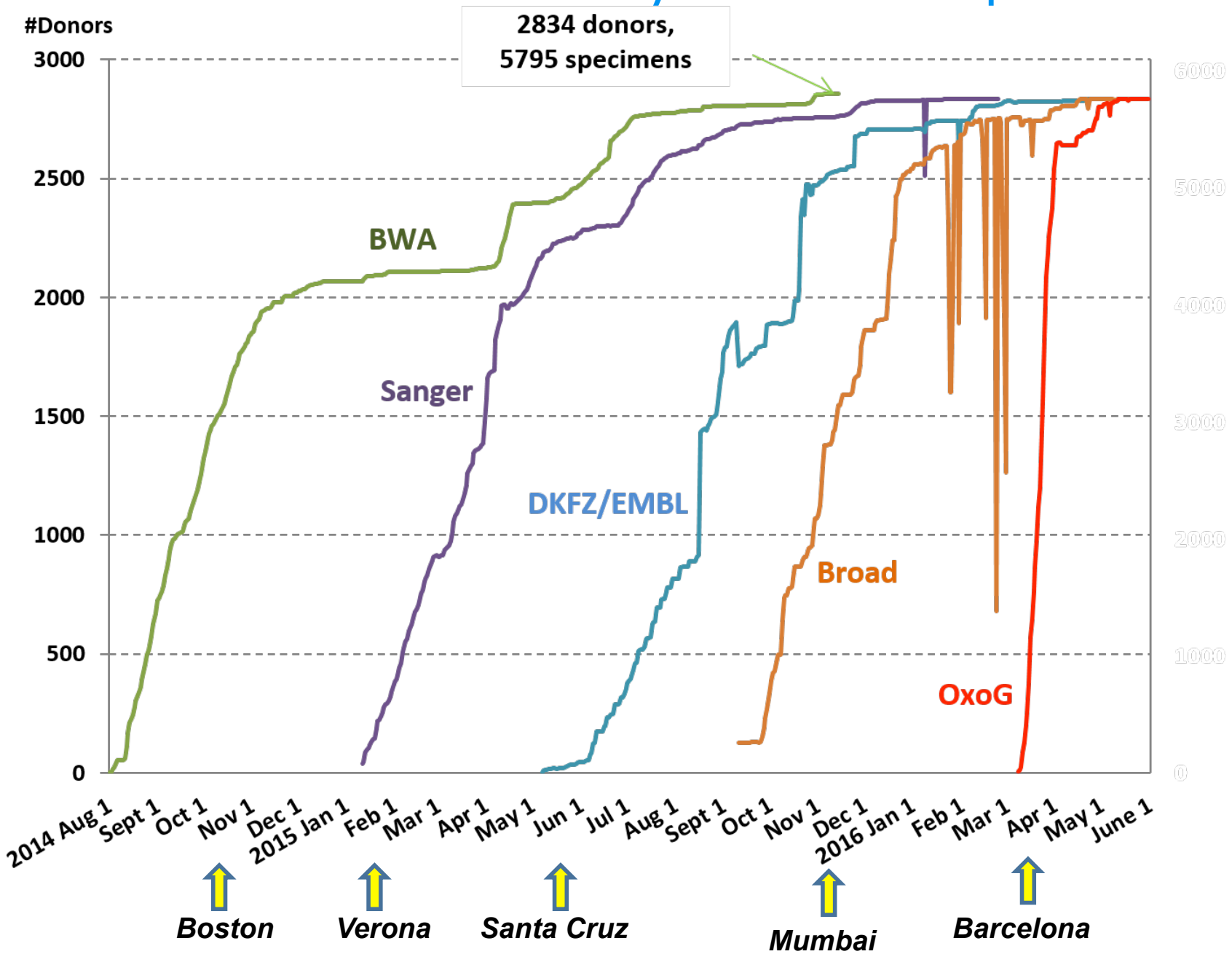
Khurana et al, NRG (2016)

PCAWG
Intro

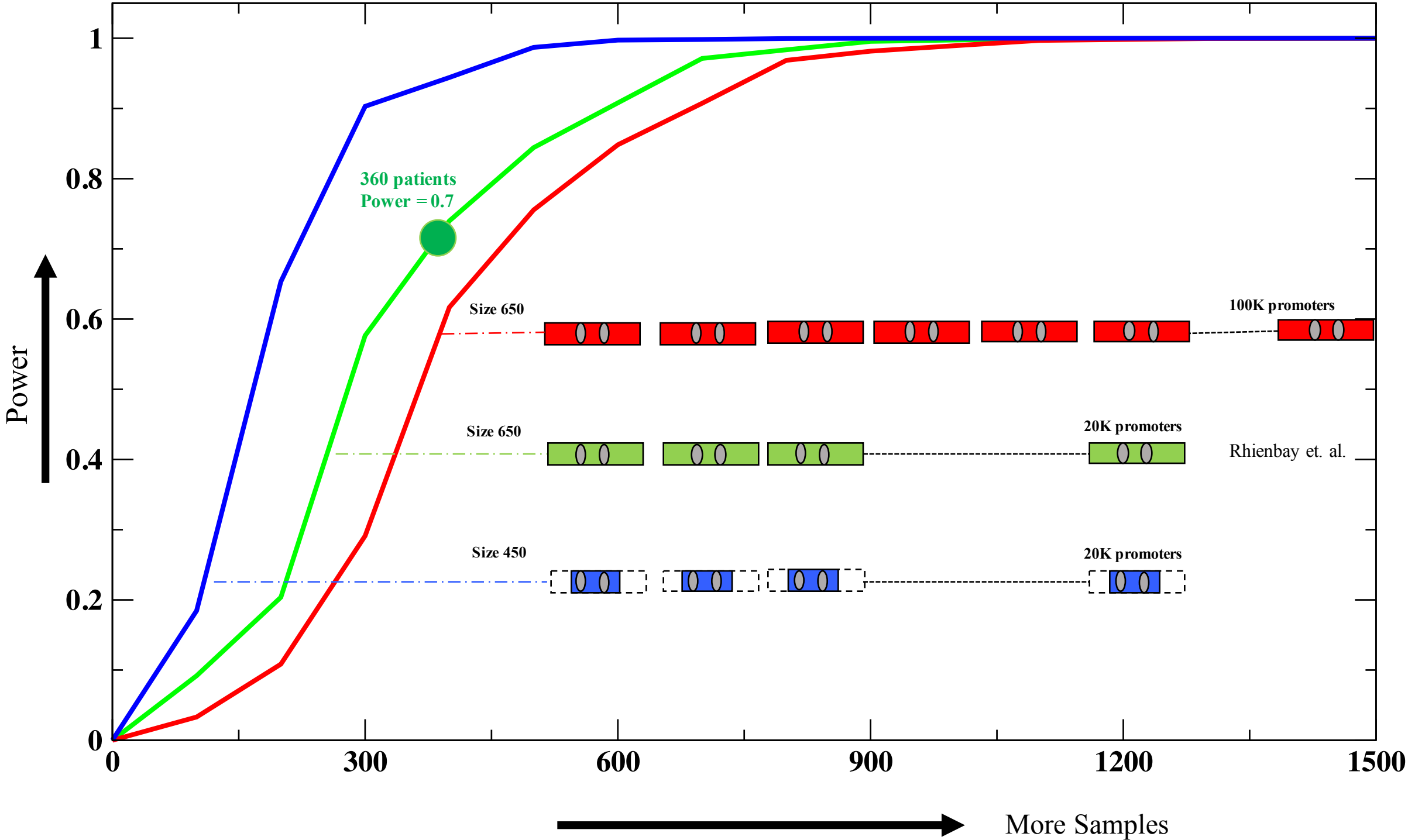
PCAWG Data Processing



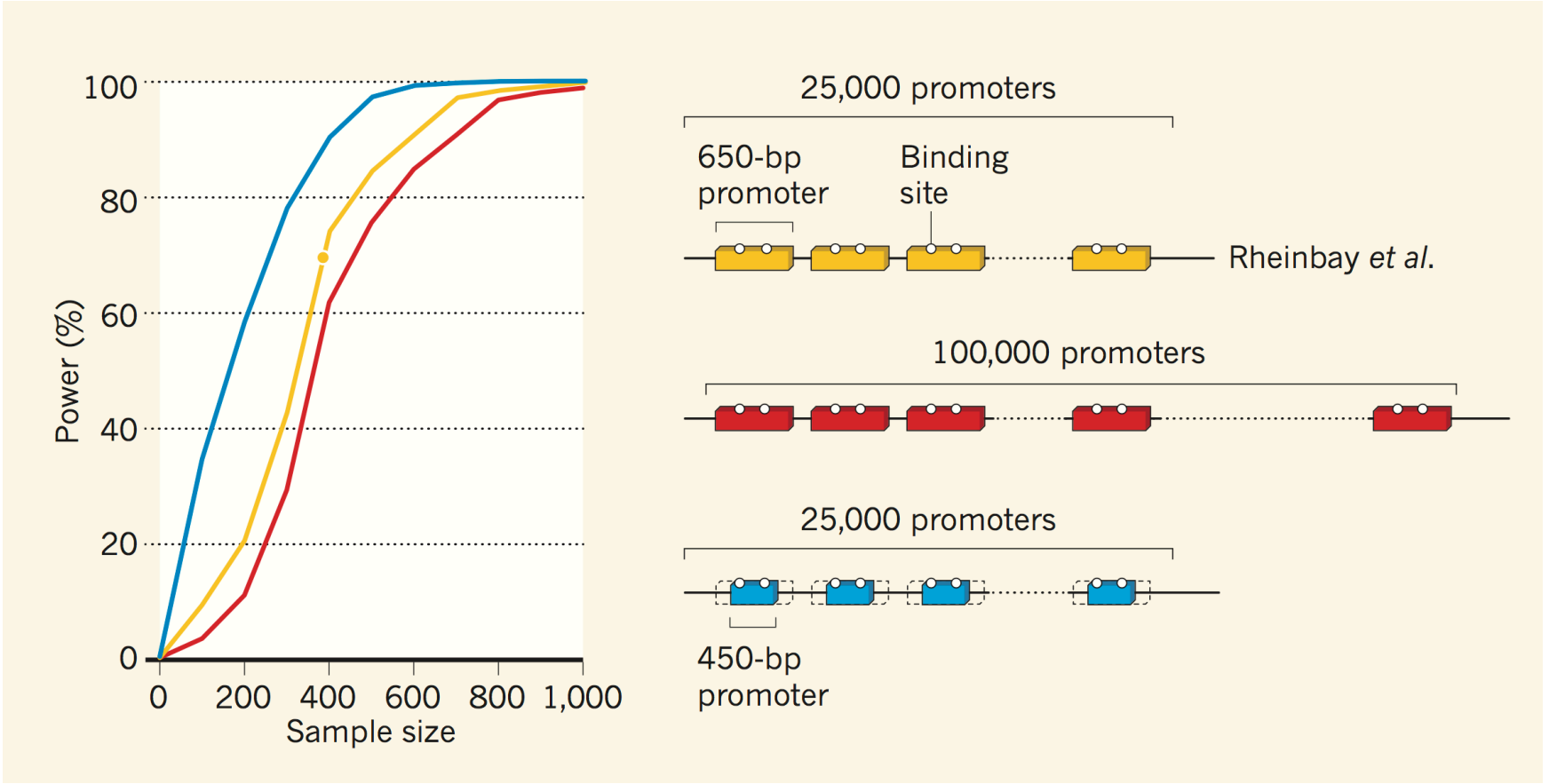
PCAWG Core Analyses Completed



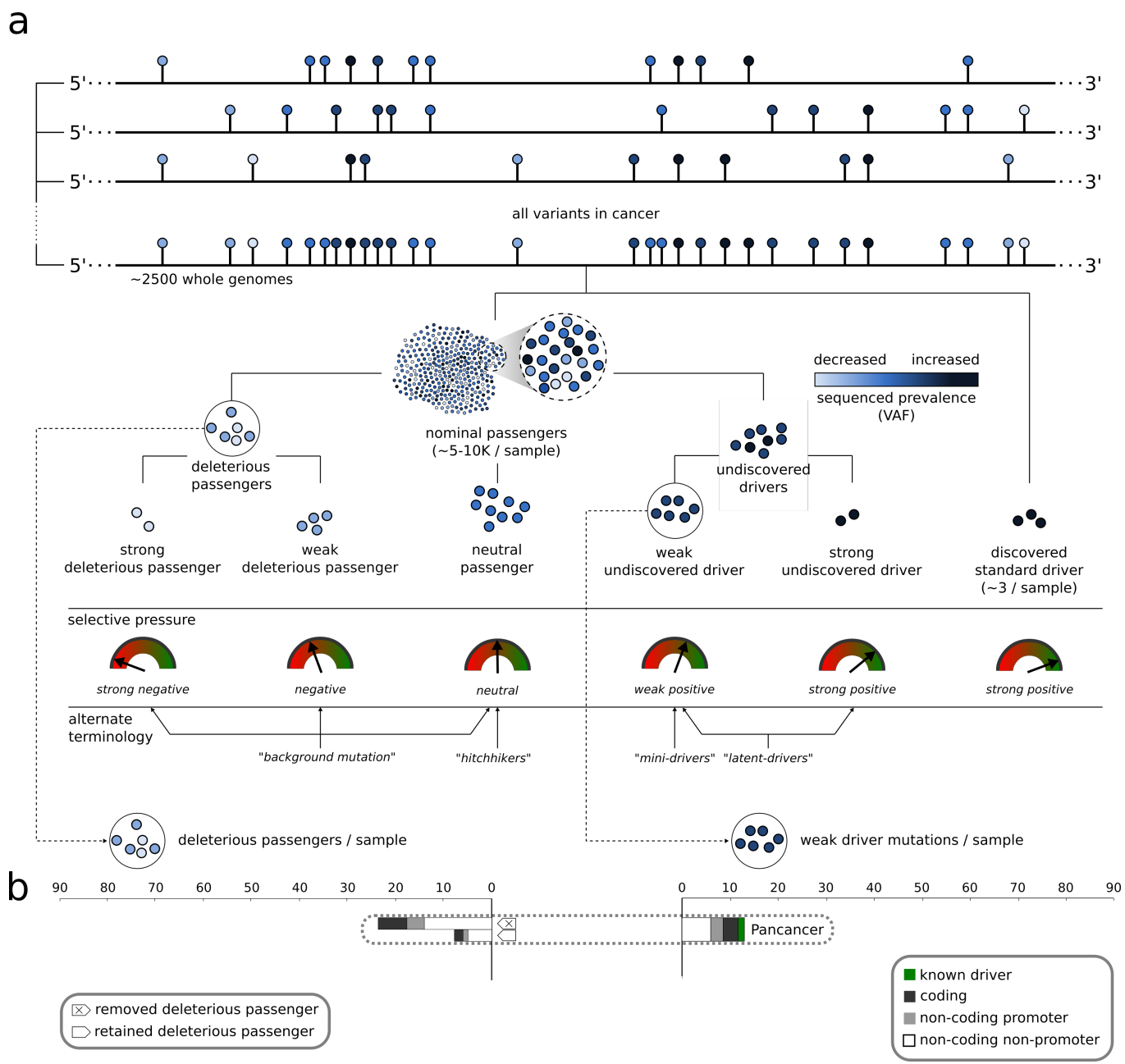
Potential Paper E
Intro (next few slides)


1
0.8
0.6
0.4
0.2
0
Power
0
300
600
900
1200
1500
360 patients
Power = 0.7
Size 650
100K promoters
Size 650
20K promoters
Rhienbay et. al.
Size 450
20K promoters
More Samples

Improving discovery of cancer-driving mutations in the non-coding genome



S. Kumar & M. Gerstein, Nature (2017)


a
5'
3'
all variants in cancer
~2500 whole genomes
decreased
increased
sequenced prevalence (VAF)
nominal passengers (~5-10K / sample)
deleterious passengers
undiscovered drivers
strong deleterious passenger
weak deleterious passenger
neutral passenger
weak undiscovered driver
strong undiscovered driver
discovered standard driver (~3 / sample)
selective pressure
strong negative
negative
neutral
weak positive
strong positive
strong positive
alternate terminology
"background mutation"
"hitchhikers"
"mini-drivers"
"latent-drivers"
deleterious passengers / sample
weak driver mutations / sample
b
90 80 70 60 50 40 30 20 10 0
0 10 20 30 40 50 60 70 80 90
Pancancer
removed deleterious passenger
retained deleterious passenger
known driver
coding
non-coding promoter
non-coding non-promoter