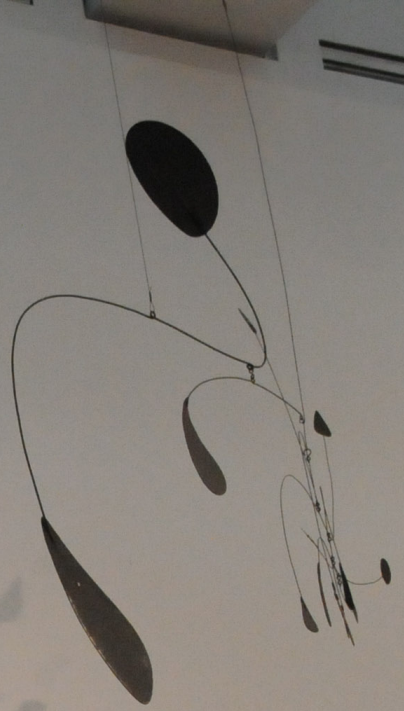
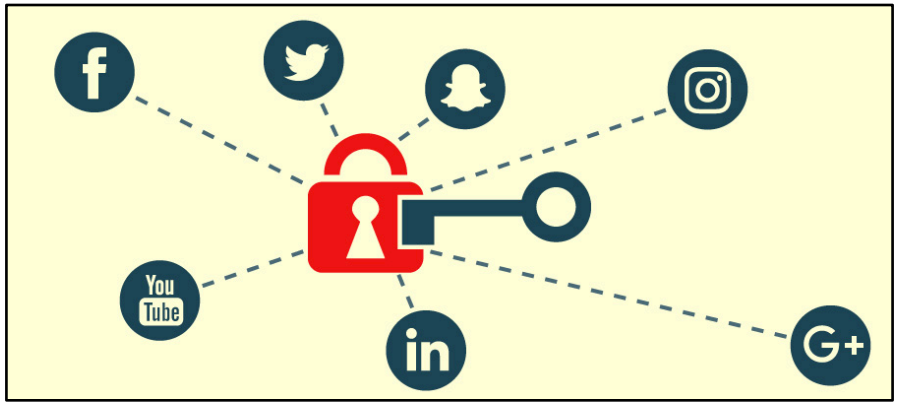
Privacy & Functional Genomics Data



M Gerstein
Yale
(See last slide for more info.)

Slides freely
"tweetable" (via
@MarkGerstein)
& downloadable from
Lectures.GersteinLab.org

Privacy: Does Genomics has similar "Big Data" Dilemma as in the Rest of Society?



- We confront privacy risks every day we access the internet (e.g., social media, e-commerce).
- Sharing & "peer-production" is central to success of many new ventures, with analogous risks to genomics
  - **EG web search**: Large-scale mining essential

Genetic Exceptionalism :

The Genome is very fundamental data, potentially very revealing about one's identity & characteristics

Personal Genomic info. essentially meaningless currently but will it be in 20 yrs? 50 yrs?

Genomic sequence very revealing about one's children. Is true consent possible?

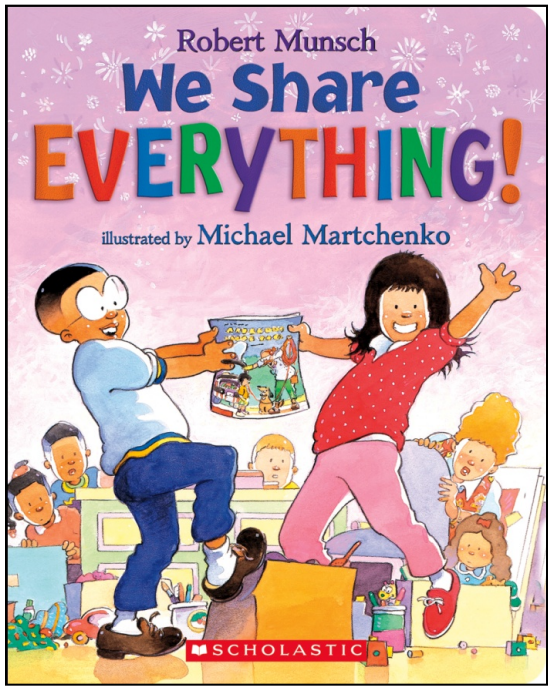
Once put on the web it can't be taken back

Ethically challenged history of genetics

Ownership of the data & what consent means (Hela)

Could your genetic data give rise to a product line?

[Seringhaus & Gerstein ('09), *Hart. Courant* (Jun 5); Greenbaum & Gerstein ('11), *NY Times* (6 Oct), D Greenbaum & M Gerstein ('08). Am J. Bioethics; D Greenbaum & M Gerstein, Hartford Courant, 10 Jul. '08 ; SF Chronicle, 2 Nov. '08; Greenbaum et al. *PLOS CB* ('11) ; Greenbaum & Gerstein ('13), The Scientist; Photos from NY Times, it.wisc.edu]

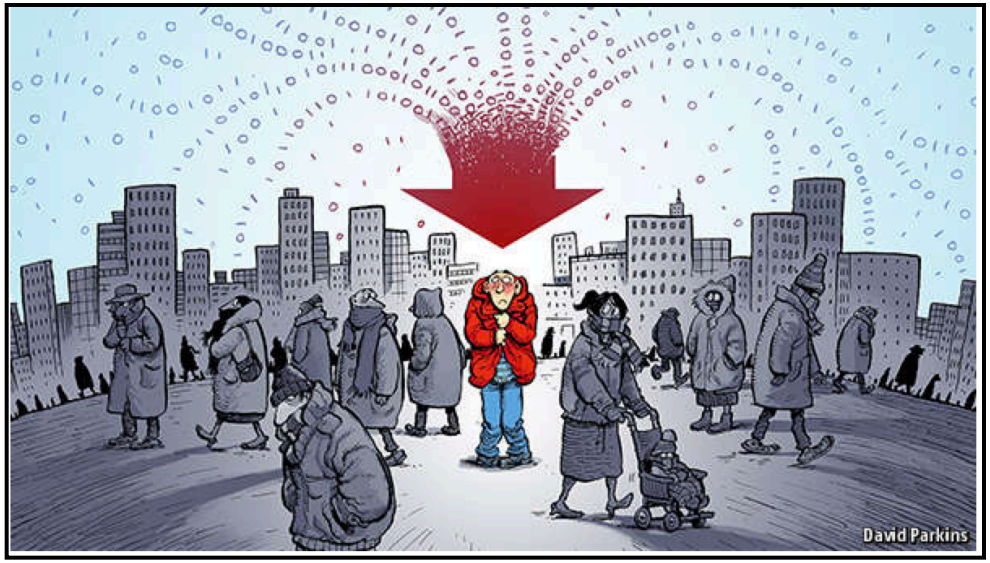


The Dilemma

- The individual (harmed?) v the collective (benefits)
  - But do sick patients care about their privacy?
- How to balance risks v rewards
  – Quantification

The Other Side of the Coin for Genomics: Why we should share

- Sharing helps **speed research**
  - Large-scale mining of this information is important for medical research
  - Statistical power
  - Privacy is cumbersome, particularly for big data



[Economist, 15 Aug '15]

[Yale Law Roundtable ('10). Comp. in Sci. & Eng. 12:8; D Greenbaum & M Gerstein ('09). Am. J. Bioethics; D Greenbaum & M Gerstein ('10). SF Chronicle, May 2, Page E-4; Greenbaum et al. *PLOS CB* ('11)]

Privacy & Functional Genomics

- Intro. to Genomic Privacy
  - The **dilemma**: The genome as fundamental, inherited info that's very private v. need for large-scale sharing & mining for med. research
- Privacy & Functional Genomics Data
  - **2-sided nature** of this data presents particularly tricky privacy issues
  - Overview of **types of the leakage**, from obvious to subtle
- Obvious leakage #1: **reads**
  - How much leakage can we expect?
    - Quantification with available data & real-world environmental samples
  - Using **pBAM** file format to remove obvious large-scale leakage
  - Using **FANCY** to assess the privacy leakage before release of the data
- Subtle Leakage #2: **eQTLs**
  - Quantifying & removing further variant info from expression levels w/ ICI & predictability.
  - Instantiating a practical linking attack w/ noisy quasi-identifiers
- Subtle Leakage #3: **Signal Profiles**
  - Manifest appreciable leakage from large & small deletions.
  - Linking attacks possible but additional complication of SV discovery in addition to genotyping

Privacy & Functional Genomics

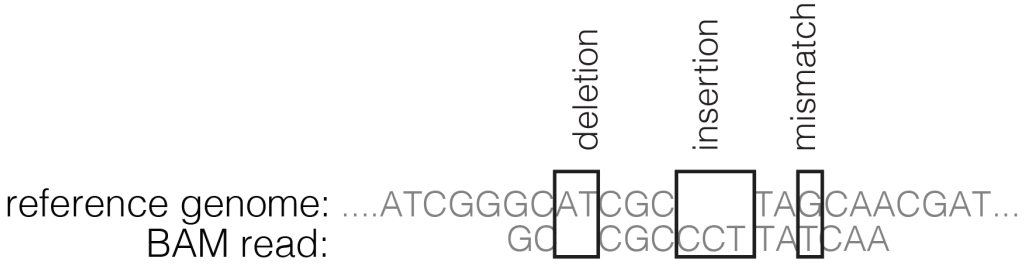
- Intro. to Genomic Privacy
  - The **dilemma**: The genome as fundamental, inherited info that's very private v. need for large-scale sharing & mining for med. research
- Privacy & Functional Genomics Data
  - **2-sided nature** of this data presents particularly tricky privacy issues
  - Overview of **types of the leakage**, from obvious to subtle
- Obvious leakage #1: **reads**
  - How much leakage can we expect?
    - Quantification with available data & real-world environmental samples
  - Using **pBAM** file format to remove obvious large-scale leakage
  - Using **FANCY** to assess the privacy leakage before release of the data
- Subtle Leakage #2: **eQTLs**
  - Quantifying & removing further variant info from expression levels w/ ICI & predictability.
  - Instantiating a practical linking attack w/ noisy quasi-identifiers
- Subtle Leakage #3: **Signal Profiles**
  - Manifest appreciable leakage from large & small deletions.
  - Linking attacks possible but additional complication of SV discovery in addition to genotyping

Peculiarities of Functional Genomics Reads - 1

2-sided nature of functional genomics data: Analysis can be very General/Public or Individual/Private

- **General quantifications** related to overall aspects of a condition – ie gene activity as a function of:
  - Developmental stage, Evolutionary relationships, Cell-type, Disease

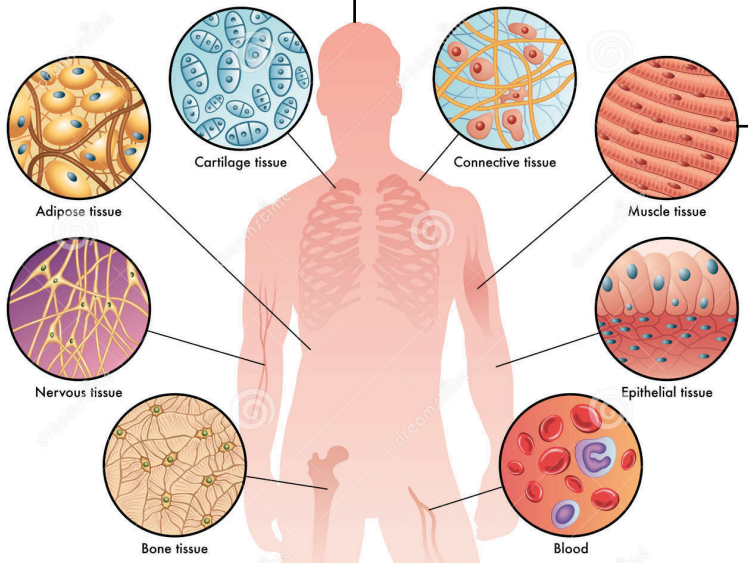
- **Above are not tied to an individual's genotype. However, data is derived from individuals & tagged with their genotypes**



- (Note, a few calculations aim to use explicitly genotype to derive general relations related to sequence variation & gene expression - eg allelic activity)

Peculiarities of Functional Genomics Reads - 2

Amount of data will soon surge those from DNA sequencing different and a new problem



DNA-seq to obtain **SNPs, indels, SVs** – **single BAM**
Aim: finding genotypes, must be protected

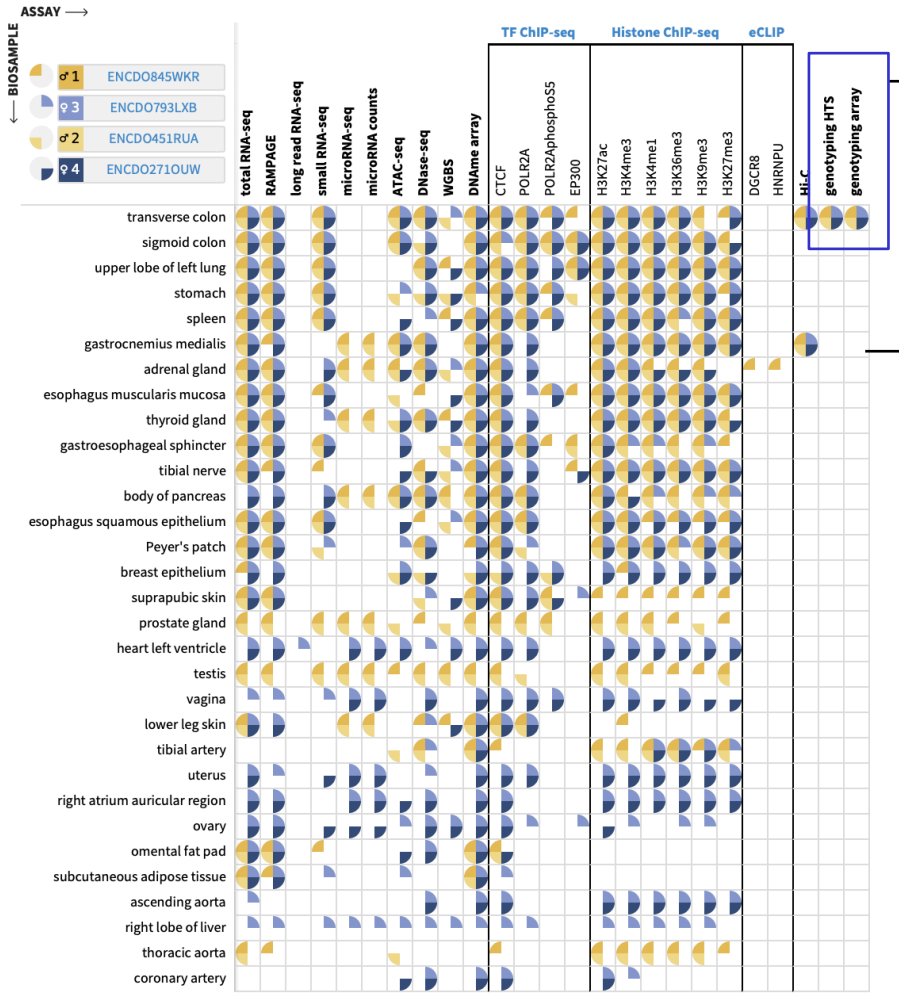
Multiple functional genomics assays **- a BAM for each assay and tissue per individual**
Aim: activity in nucleus in disease and health

- **RNA-Seq**
- **ChIP-Seq**
  - **(TFs, HMs, …)**
- **ATAC-Seq**
- **Dnase**
- **Hi-C**
- **single-cell assays**
  - can be ~0.5 million cells
  - each cell x transcriptome
  - more data than ever before

Figure: adopted from https://www.tes.com/lessons/Q0YL_OHrNtTVGg/biology-2-topic-2-tissues-organ-systems-and-homeostasis

Peculiarities of Functional Genomics Reads - 2

Amount of data will soon surge
those from DNA sequencing
different and a new problem



DNA-seq to obtain **SNPs, indels, SVs** – **single BAM**
Aim: finding genotypes, must be protected

Multiple functional genomics assays **- a BAM for each assay and tissue per individual**
Aim: activity in nucleus in disease and health

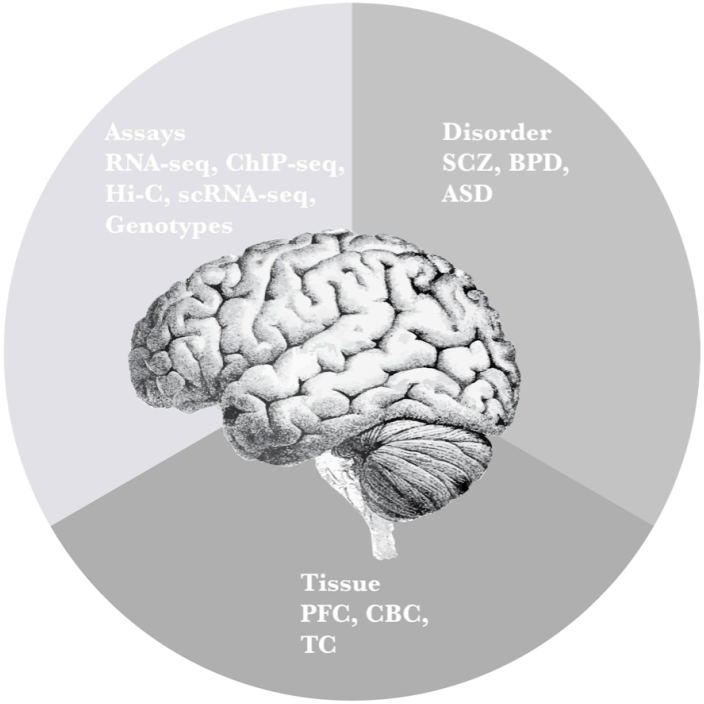
RNA-Seq
ChIP-Seq
(TFs, HMs, …)
ATAC-Seq
Dnase
Hi-C
single-cell assays
can be ~0.5 million cells
each cell x transcriptome
more data than ever before

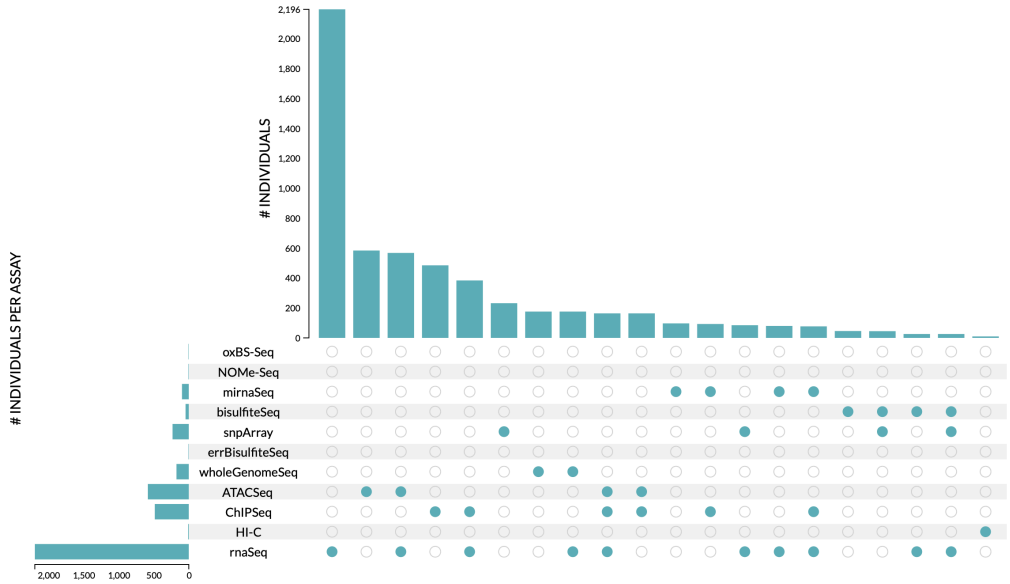
Figure: adopted from https://www.tes.com/lessons/Q0YL_OHrNtTVGg/biology-2-topic-2-tissues-organ-systems-and-homeostasis

Peculiarities of Functional Genomics Reads - 3

A new source of privacy leakage

inferring phenotypes



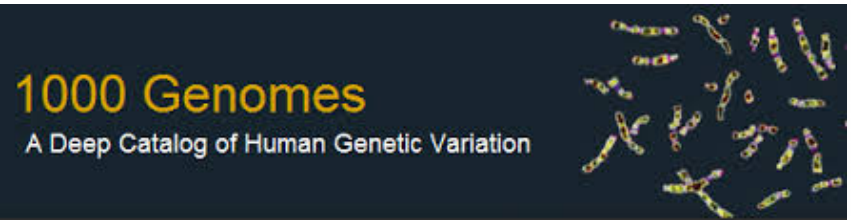


If you find out an individual is in a functional genomics cohort, you also find out potentially sensitive phenotypes

Figures: PsychENCODE consortium

Representative Functional Genomics, Genotype, eQTL Datasets

- Genotypes are available from the 1000 Genomes Project
- mRNA sequencing for 462 individuals from gEUVADIS and ENCODE
  - Publicly available quantification for protein coding genes
- Functional genomics data (ChIP-Seq, RNA-Seq, Hi-C) available from ENCODE
- Approximately 3,000 cis-eQTL (FDR<0.05)



Data Reduction in RNA-Seq: an Overview

Successive steps of Data Reduction

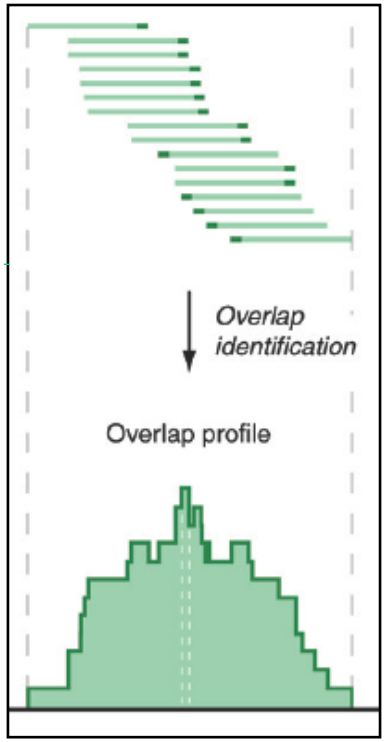
Fastq sequence files
~5-10 GB

```
ATACAAGCAAGTATAAGTTCGTATGCCGTCTT
GGAGGCTGGAGTTGGGGACGTATGCGGCATAG
TACCGATCGAGTCGACTGTAAACGTAGGCATA
ATTCTGACTGGTGTCATGCTGATGTACTTAAA
```

Index-building + Alignment to reference genome

BAM files
~1-2-fold reduction

Conversion to signal track by overlapping reads

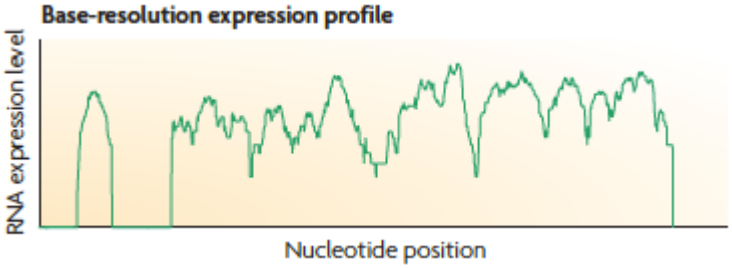


Reads => Signal

BigWig files
~25-fold reduction

Mapping to genes

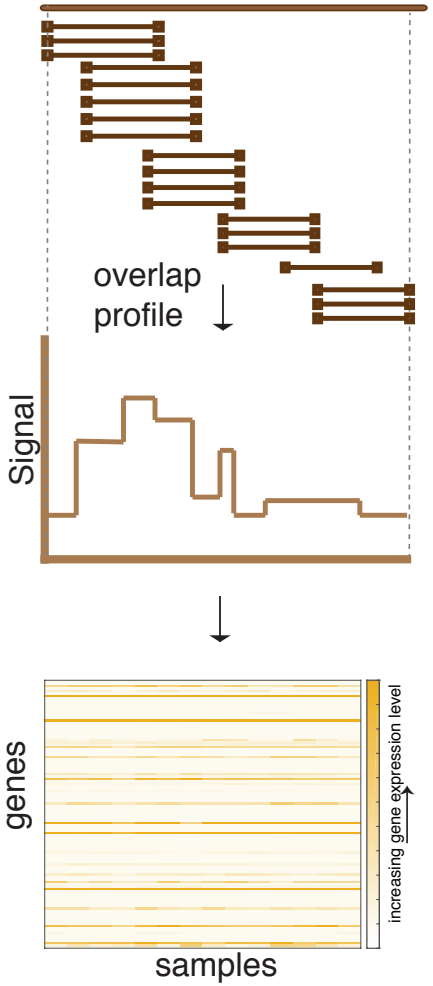
Gene/Transcript expression matrix
~20-fold reduction



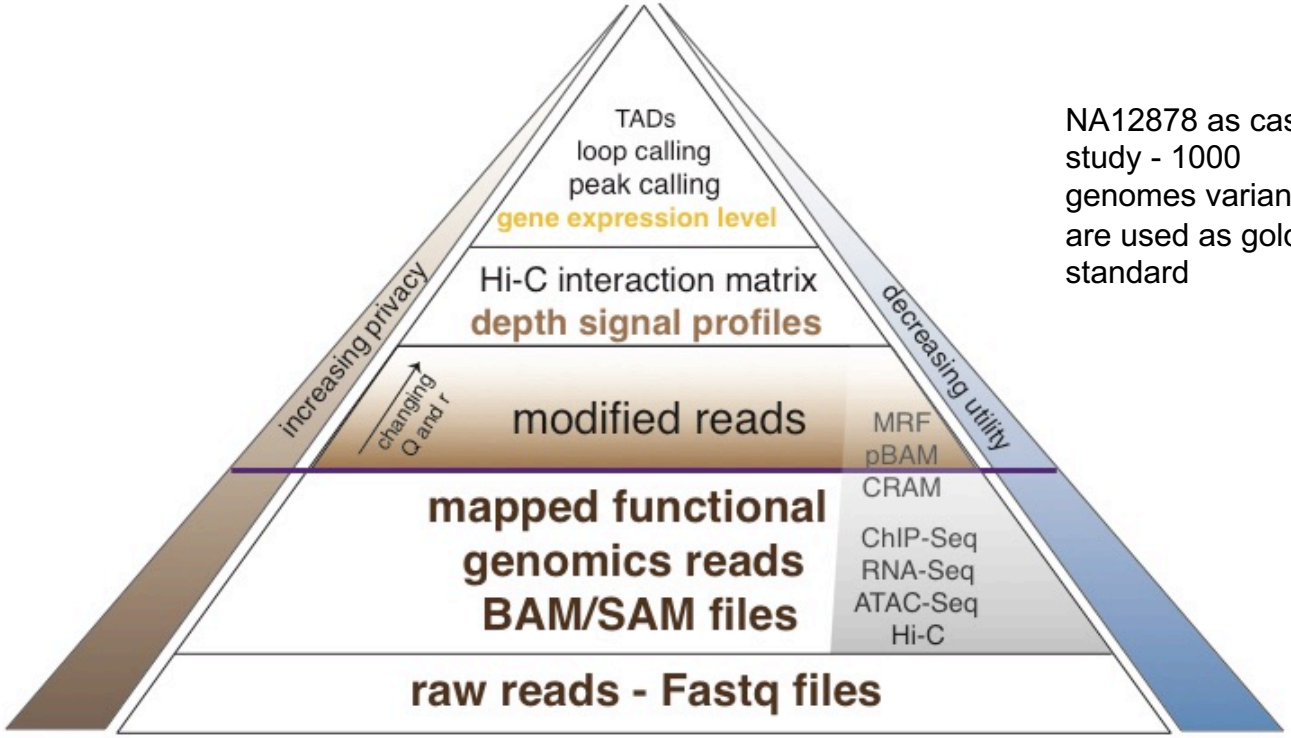
Quantitative information from RNA-seq signal:
average signals at exon level (RPKMs)

[*NAT. REV.* 10: 57; *PLOS CB* 4:e1000158; *PNAS* 4:107: 5254]

Functional genomics data comes with a great deal of sequencing; We can quantify amount of leakage at every step of the data summarization process.



[*Gursoy et al,* Cell, in press]

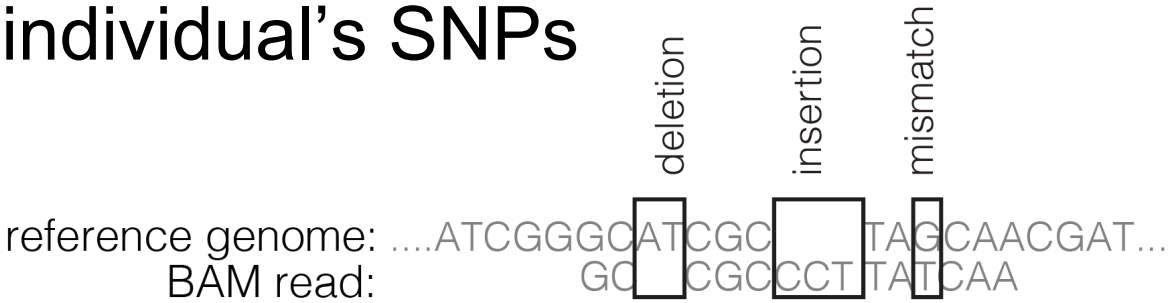


NA12878 as case study - 1000 genomes variants are used as gold standard

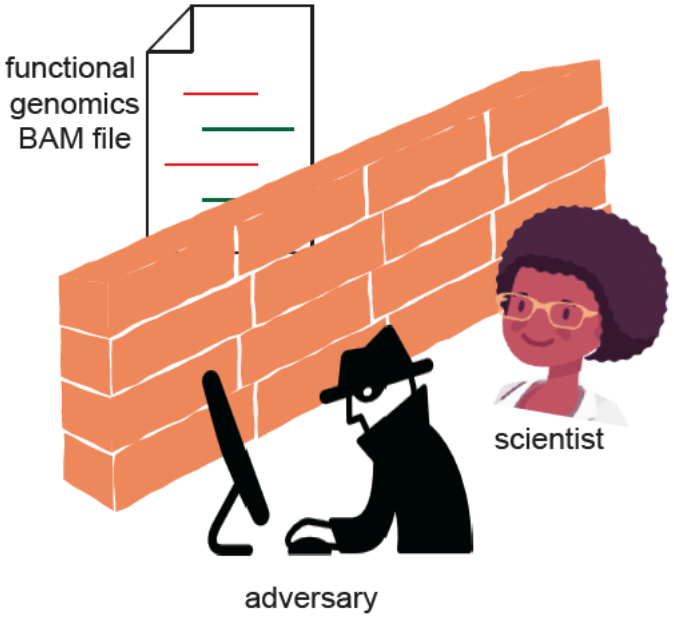
Leakage Source	Leaking Variants	# of potential variants	Average leakage per variant (bits)	Maximum leakage per variant (bits)	# of accessible variants	Total leakage (bits)
Raw reads	Exonic variants	2,682,417	0.10 ± 0.28	9.88 ± 2.12	246,893	24,689
Modified reads Q = {indels}	Exonic SNVs	2,607,969	0.09 ± 0.27	9.95 ± 2.02	231,031	207,92
Modified reads Q = {mismatches}	Exonic indels	51,408	0.33 ± 0.47	7.64 ± 2.42	15,862	5234
Signal profiles	Exonic deletions	48,019	0.29 ± 0.45	7.97 ± 2.42	1,067	298
Gene expression quantification	eQTLs	3,175	1.19 ± 0.36	4.00 ± 1.92	158	188

Functional Genomics Reads

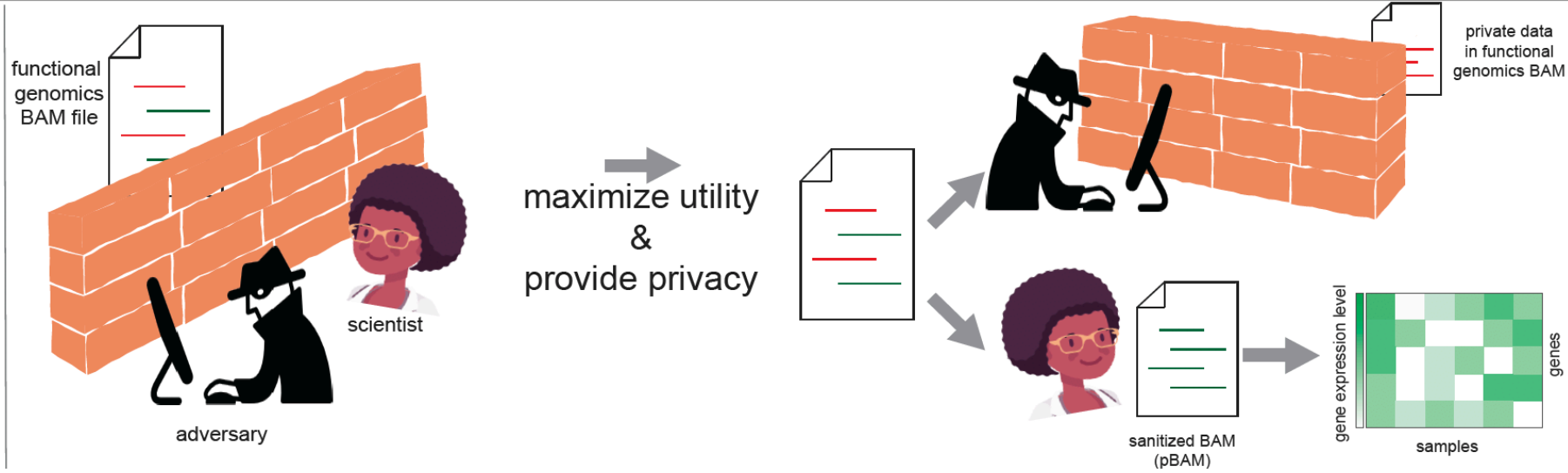
- Usually disseminated in the form of ”BAM” files
- Contain individual’s SNPs



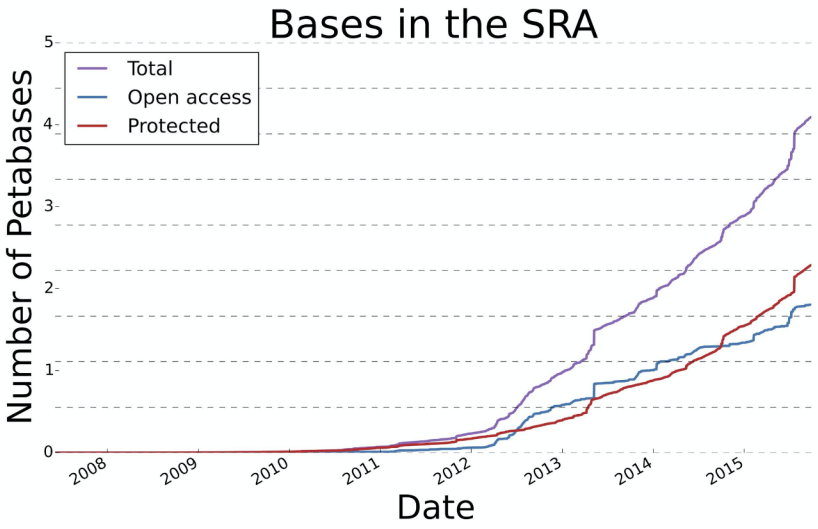
- Traditional approach to protect privacy: dbGAP, EGA, …
- Protects the data from bad actors, but also from scientific community



GOAL



[Gursoy et al., Cell, in press]



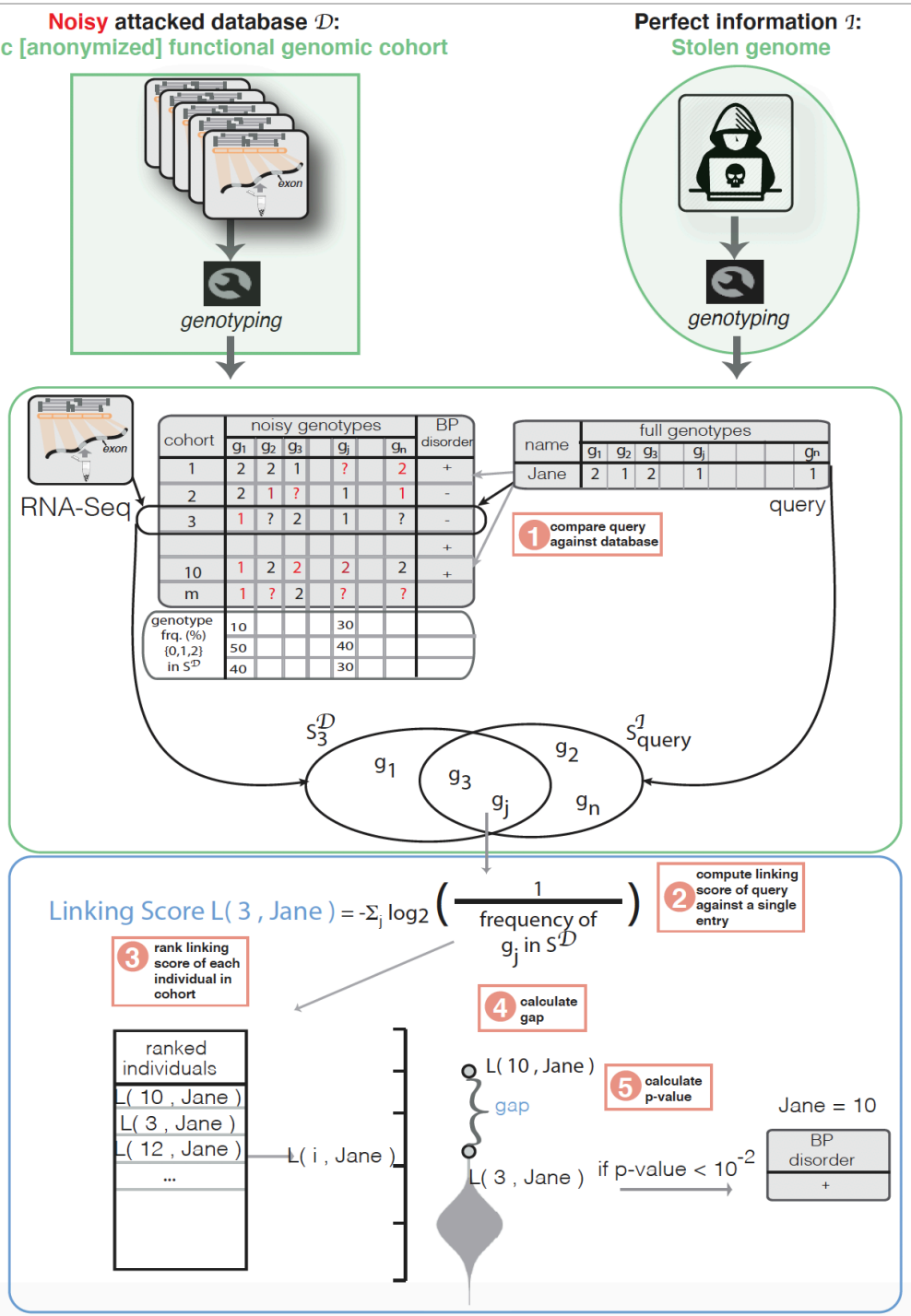
[Muir et al., Genome Biology, 2016]

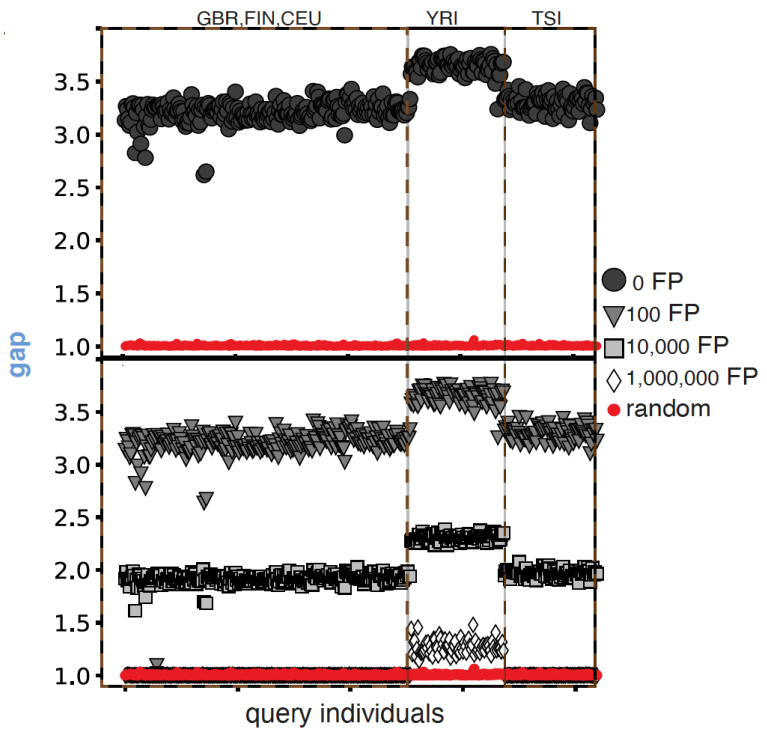
1. **Quantify the amount of leakage in reads**
   - Using perfect reference public datasets
   - Using environmental objects (coffee cups)
     - Under different noise profiles
2. **Develop data-sanitization protocols based on quantifications**
   - Bounds for privacy & utility balance
   - A new mode for sharing data

Privacy & Functional Genomics

- Intro. to Genomic Privacy
  - The **dilemma**: The genome as fundamental, inherited info that's very private v. need for large-scale sharing & mining for med. research
- Privacy & Functional Genomics Data
  - **2-sided nature** of this data presents particularly tricky privacy issues
  - Overview of **types of the leakage**, from obvious to subtle
- Obvious leakage #1: **reads**
  - How much leakage can we expect?
    - Quantification with available data & real-world environmental samples
  - Using **pBAM** file format to remove obvious large-scale leakage
  - Using **FANCY** to assess the privacy leakage before release of the data
- Subtle Leakage #2: **eQTLs**
  - Quantifying & removing further variant info from expression levels w/ ICI & predictability.
  - Instantiating a practical linking attack w/ noisy quasi-identifiers
- Subtle Leakage #3: **Signal Profiles**
  - Manifest appreciable leakage from large & small deletions.
  - Linking attacks possible but additional complication of SV discovery in addition to genotyping

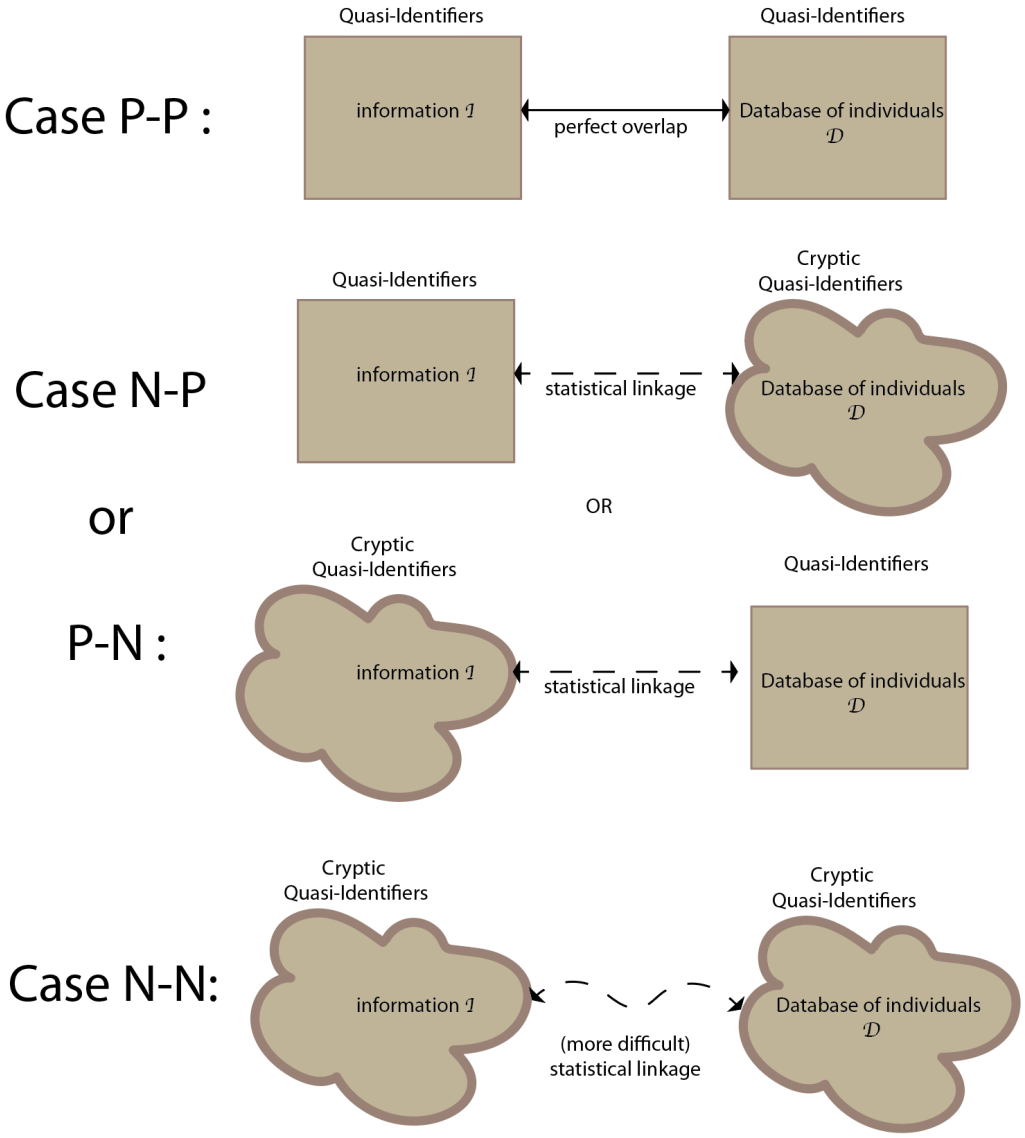
Linking a known individual with perfectly characterized genome to a functional genomics cohort & inference of sensitive phenotypes



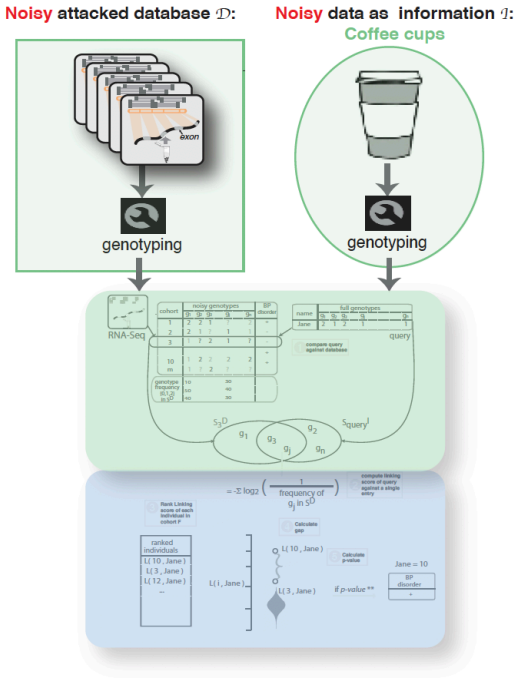


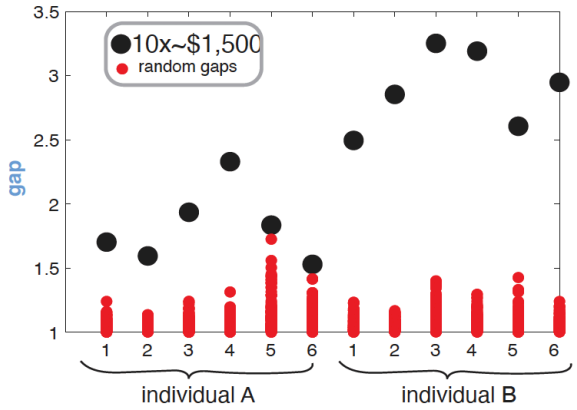
[Gursoy et al., Cell, in press]

Genotypes observed from functional genomics data are noisy: How can we quantify the private information leakage under different noise profiles?

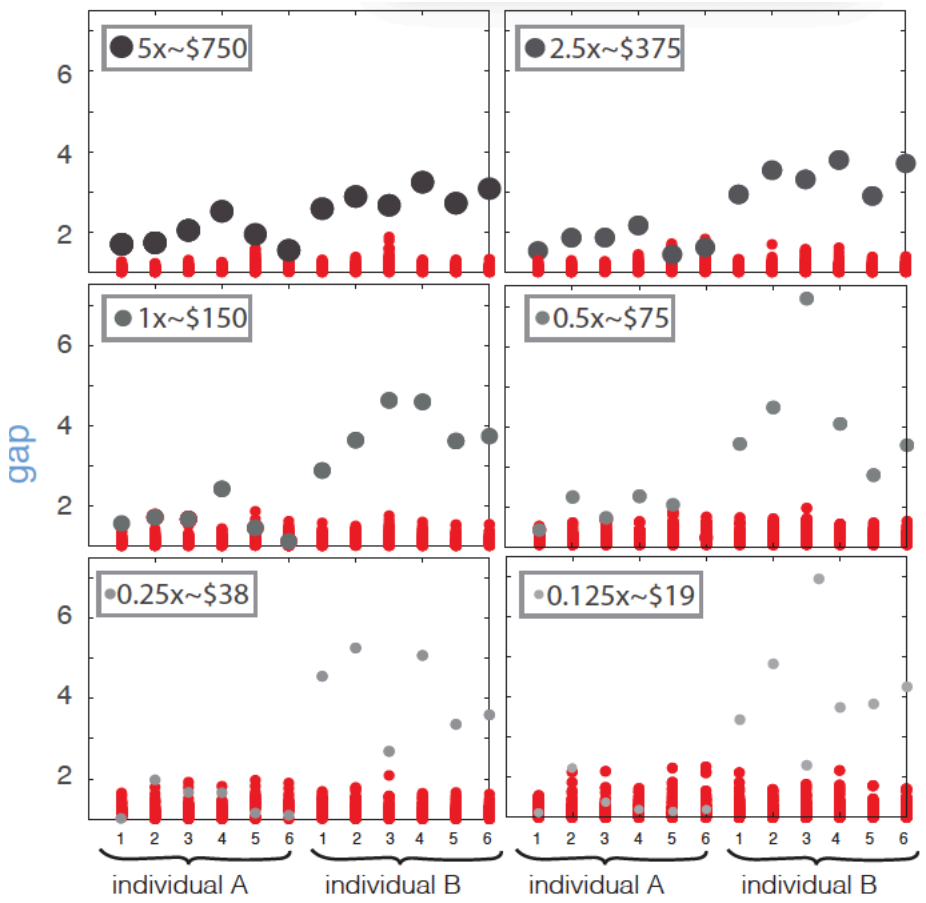


Linking a coffee cup from a known individual to a functional genomics cohort & inference of sensitive phenotypes



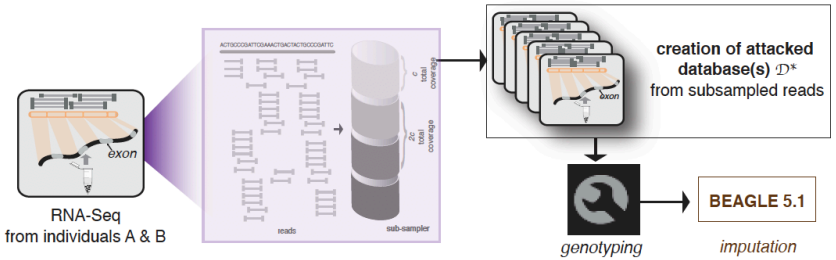


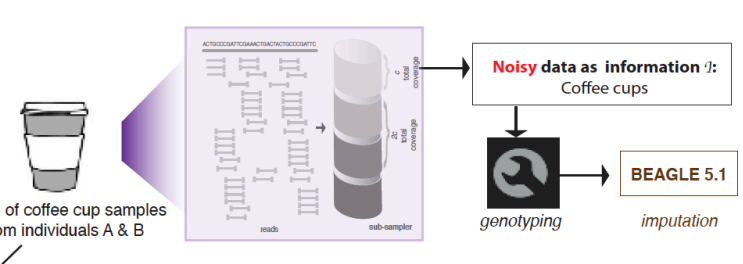
$19 was enough to link the coffee cup to the panel

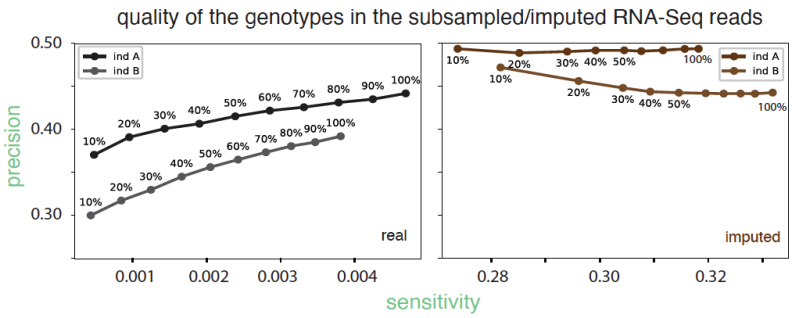


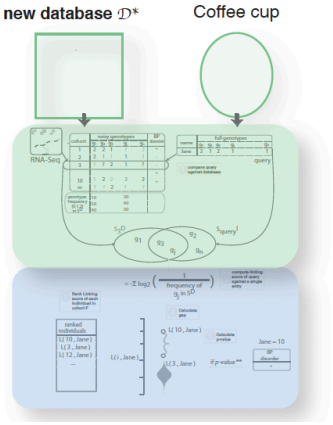
[Gursoy et al., Cell, in press]

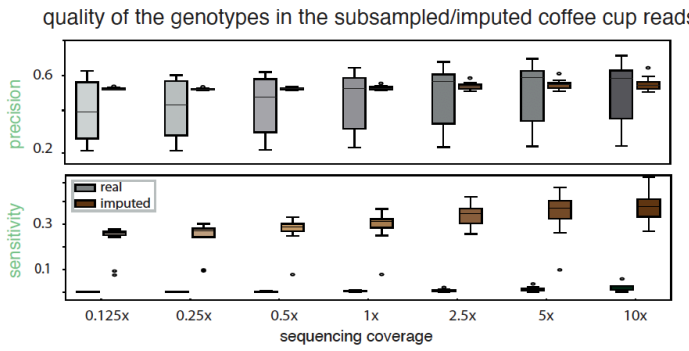
Noise can be changed with more subsampling & linking can be improved by imputation

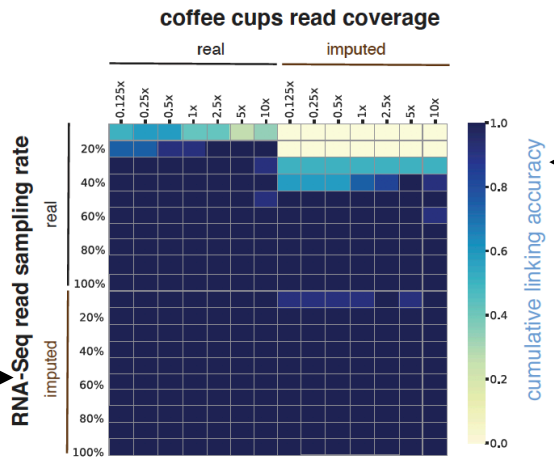










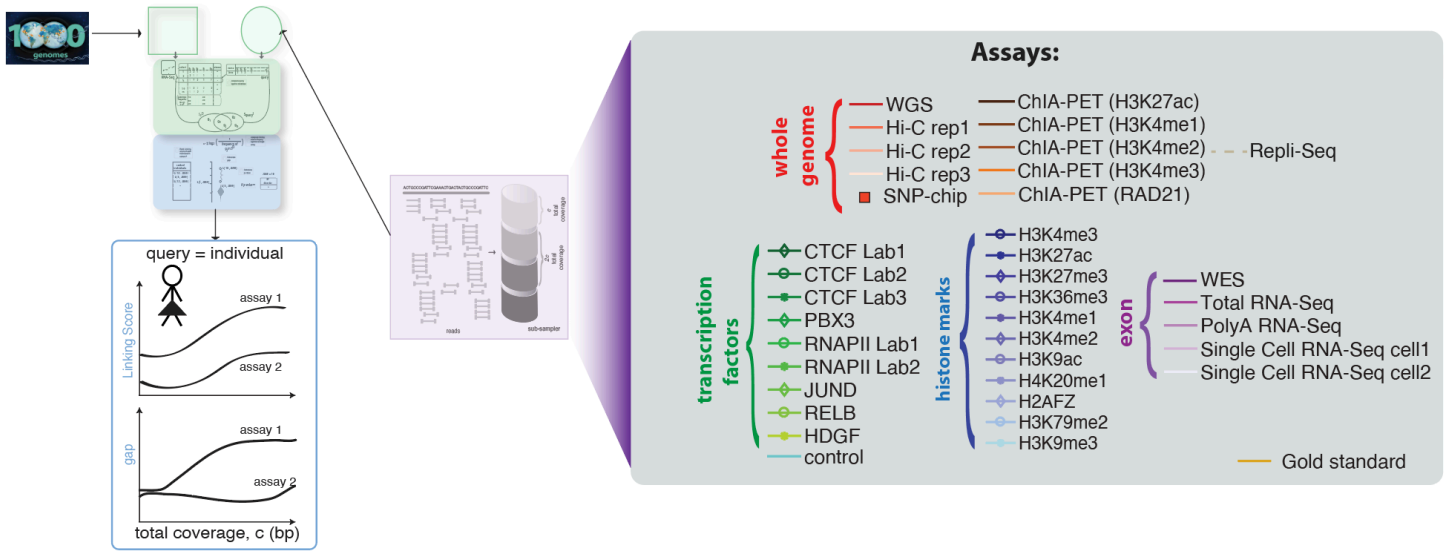


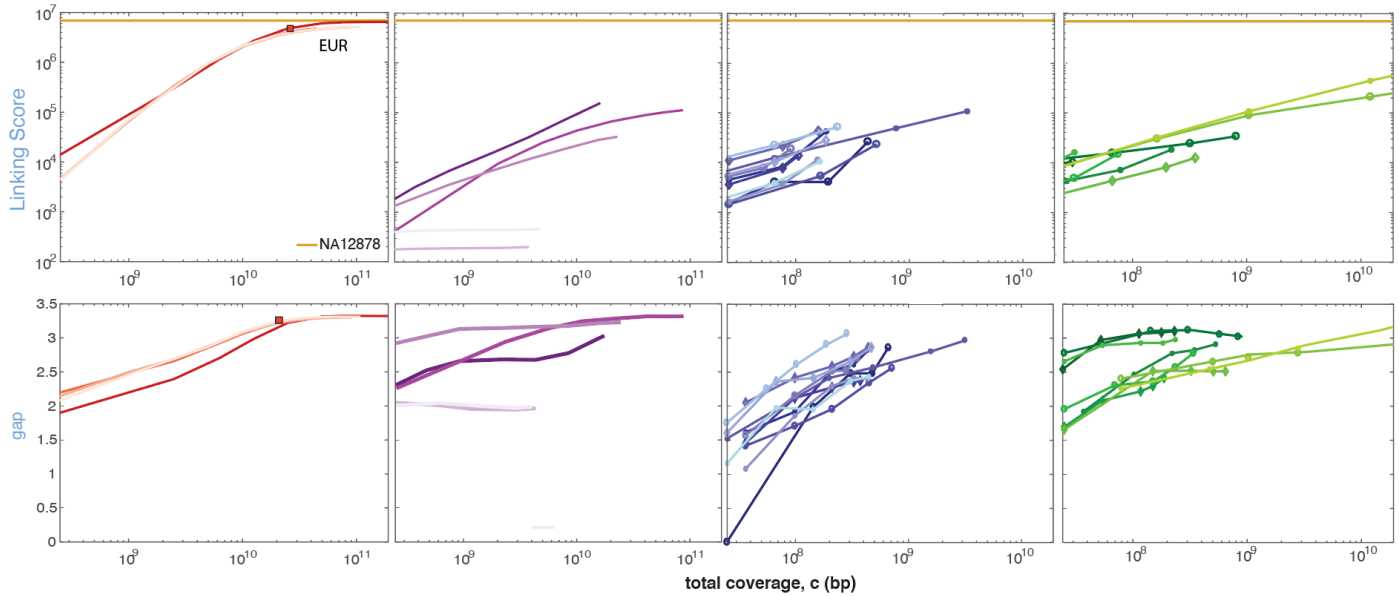
Imputation on RNA-Seq increases linking accuracy

Imputation on coffee-cups decreases linking accuracy

[Gursoy et al., Cell, in press]

Quantifying leakage in different functional genomics assays



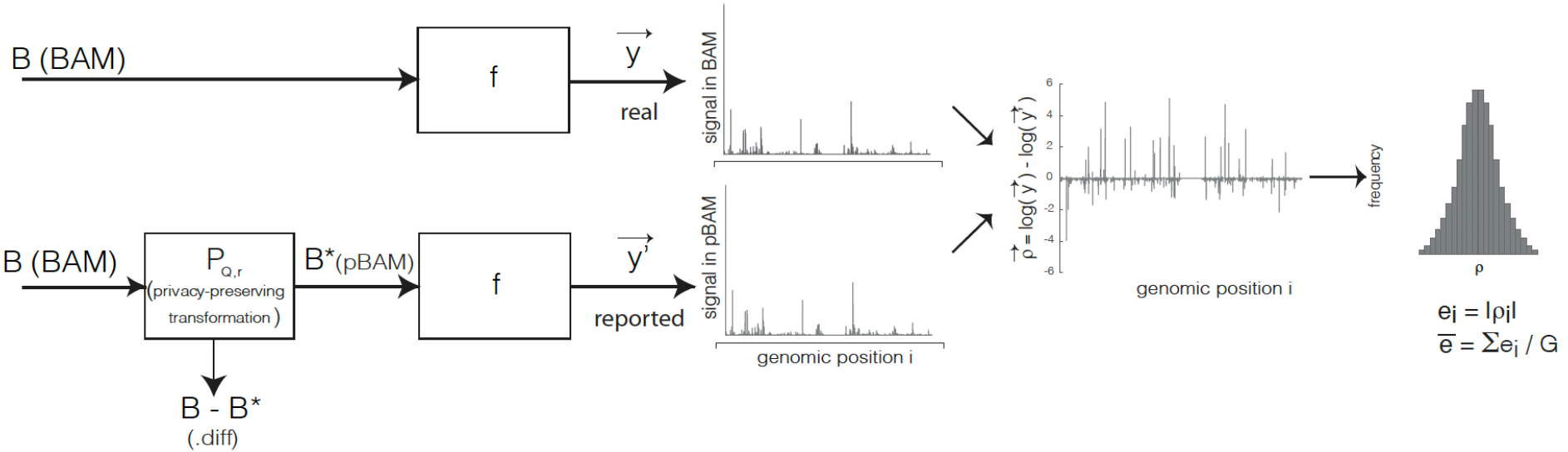


[Gursoy et al., Cell, in press]

Privacy & Functional Genomics

- Intro. to Genomic Privacy
  - The **dilemma**: The genome as fundamental, inherited info that's very private v. need for large-scale sharing & mining for med. research
- Privacy & Functional Genomics Data
  - **2-sided nature** of this data presents particularly tricky privacy issues
  - Overview of **types of the leakage**, from obvious to subtle
- Obvious leakage #1: **reads**
  - How much leakage can we expect?
    - Quantification with available data & real-world environmental samples
  - Using **pBAM** file format to remove obvious large-scale leakage
  - Using **FANCY** to assess the privacy leakage before release of the data
- Subtle Leakage #2: **eQTLs**
  - Quantifying & removing further variant info from expression levels w/ ICI & predictability.
  - Instantiating a practical linking attack w/ noisy quasi-identifiers
- Subtle Leakage #3: **Signal Profiles**
  - Manifest appreciable leakage from large & small deletions.
  - Linking attacks possible but additional complication of SV discovery in addition to genotyping

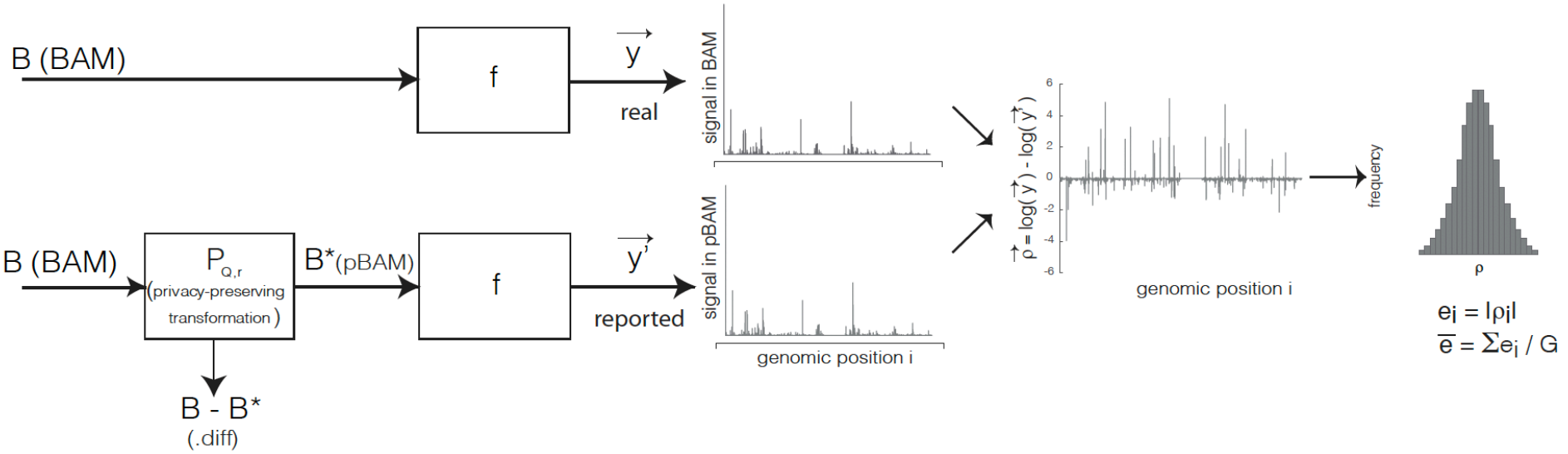
Privacy-preserving Binary Alignment Mapping (pBAM)



- No need to know the sequence of mapped reads to aggregate them
- A manipulation on Binary Alignment Files (BAM)
  - Find leaky fields/tags
  - Generalization
- Goal:
  - Accurate gene/transcript expression quantification
  - Works with the pipelines / SAMtools

[Gursoy et al., Cell, in press]

Privacy-preserving Binary Alignment Mapping (pBAM)



Definitions:

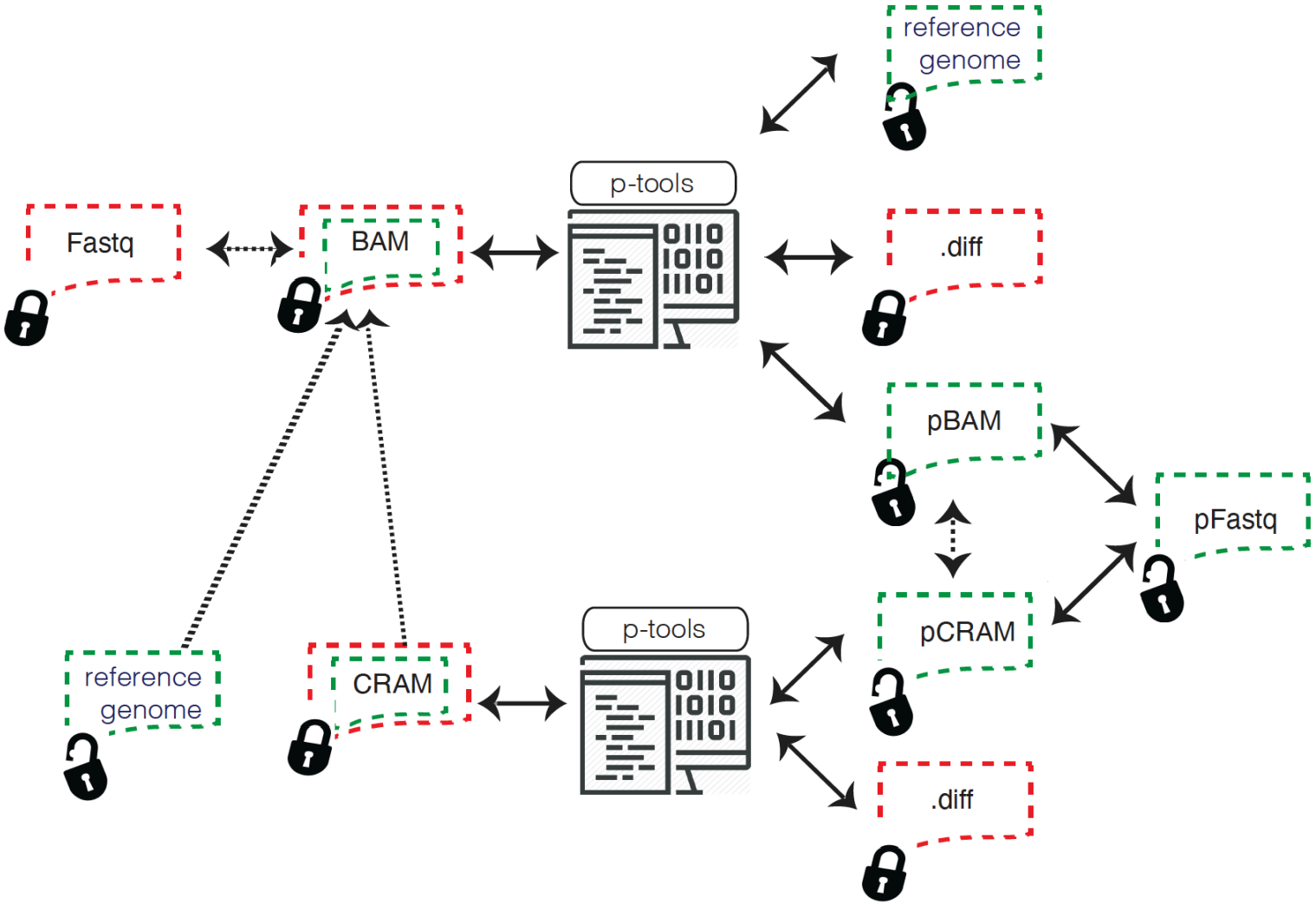
1. Privacy: $B^* = P_{Q,r}(B)$ can be viewed as δ-private with respect to operation Q, if δ = r/t,
r = non-observable (sanitized) variants in pBAM
t = all observed variants in BAM
We can reach 100% privacy when r = t.

2. Utility: $B^* = P_{Q,r}(B)$ can be viewed as having ε -γ -utility with respect to operation Q, if ε = (G − m)/G,
G = total number of units
m = the total number of units with $e_i > \gamma$
Practically, ε is the fraction of the genomic bases affected by the sanitization
& γ can be set based on difference between replicates

Note: Variants can have different effect on privacy based on their rarity; however that will make definitions to be dependent on the composition of attacked database

[Gursoy et al., Cell, in press]

Practical Software for Sanitizing a BAM, creating a pBAM and small ".diff" file



Latest version (pTools v1.0.1): http://privaseq3.gersteinlab.org/download
Development: https://github.com/ENCODE-DCC/ptools
Test files: http://privaseq3.gersteinlab.org/data

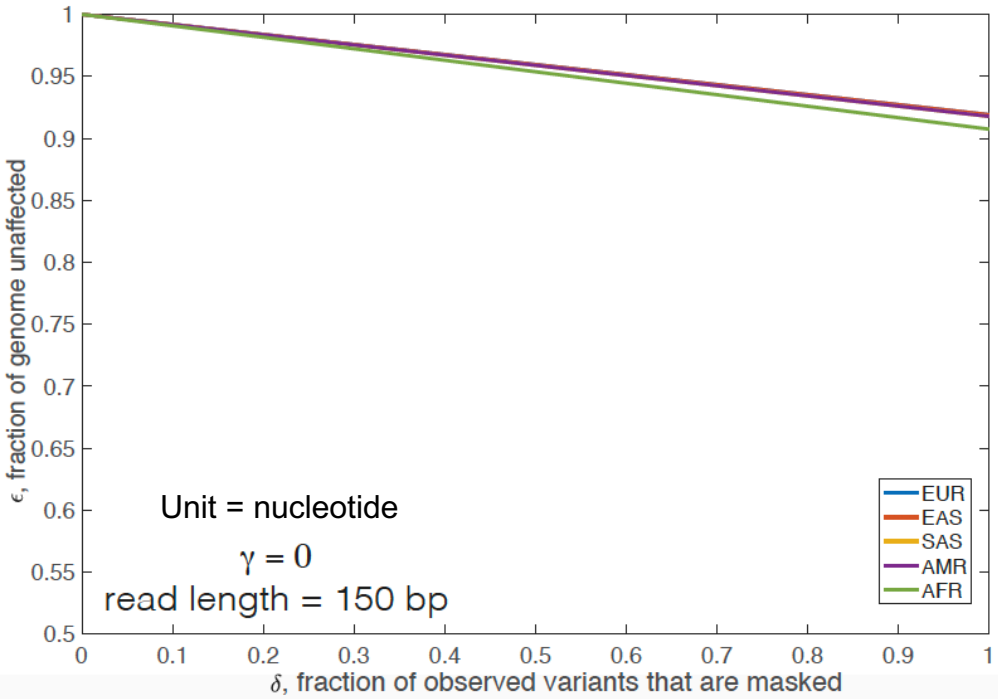
Privacy-preserving Binary Alignment Mapping (pBAM)

(numerical bounds relating privacy and utility)

$$(-\varepsilon \cdot G + G) \leq L_R \cdot (\delta \cdot t - (r_{snp} - r_{del})) + (2 \cdot L_R - 2) \cdot r_{del}$$

- utility parameter
- length of the genome
- read length
- privacy parameter
- number of observable variants
- sanitized variants

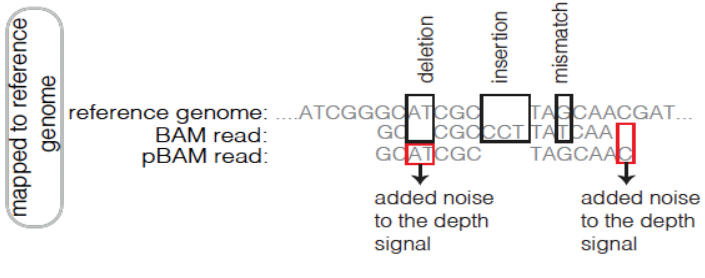
- Here we used average number of genotypes per ancestry as t
- Extreme case as number of observable variants from a functional genomics BAM << t
- Assuming a genomic signal profile, we can see how SNPs & indels maximally affect the profile, giving a numerical bounds relating δ & ε



[Gursoy et al., Cell, in press]

Privacy-preserving Binary Alignment Mapping (pBAM)

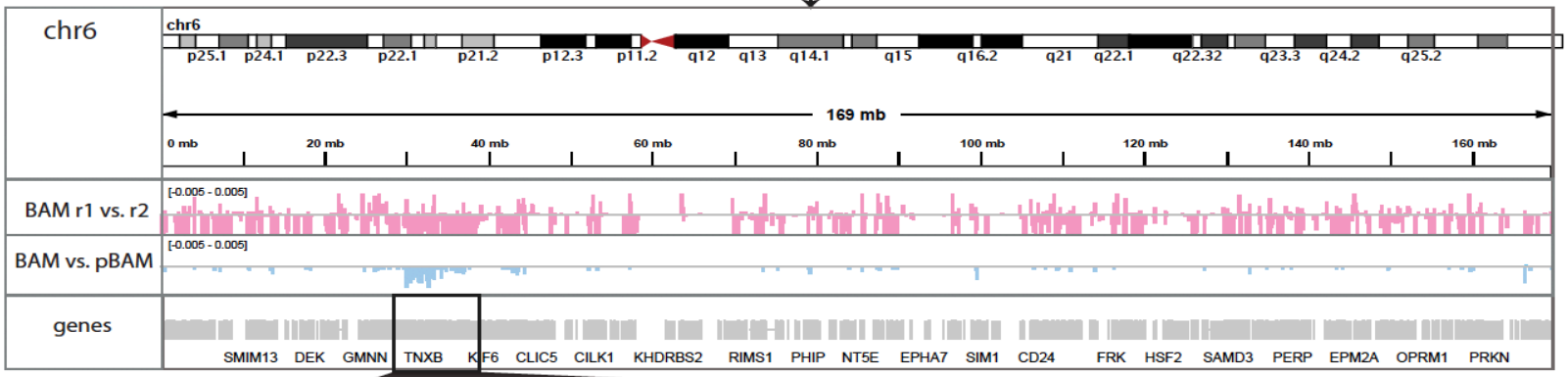
(empirical utility observations)
NA12878 RNA-Seq BAM files

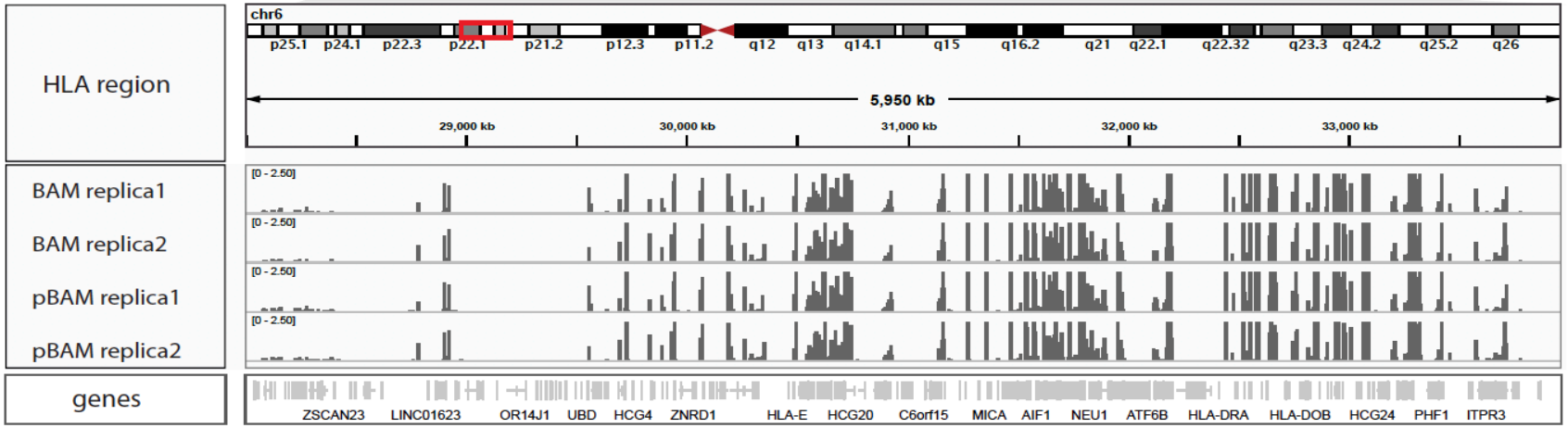


Variation between replicates
Variation between BAM-pBAM

[Gursoy et al., Cell, in press]

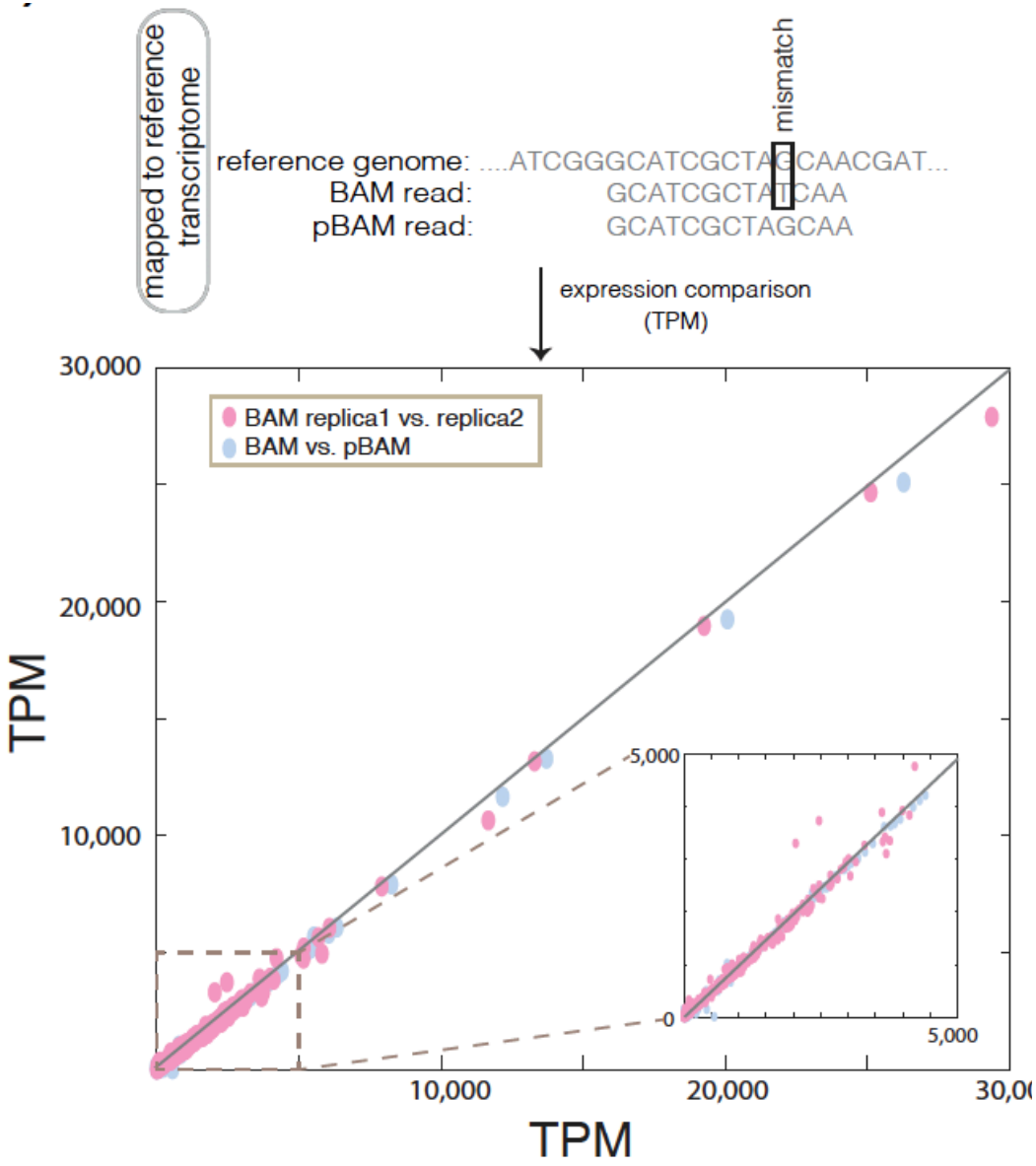
signal comparison (BigWig)





Privacy-preserving Binary Alignment Mapping (pBAM)

(empirical utility observations)
NA12878 RNA-Seq BAM files



Gene level

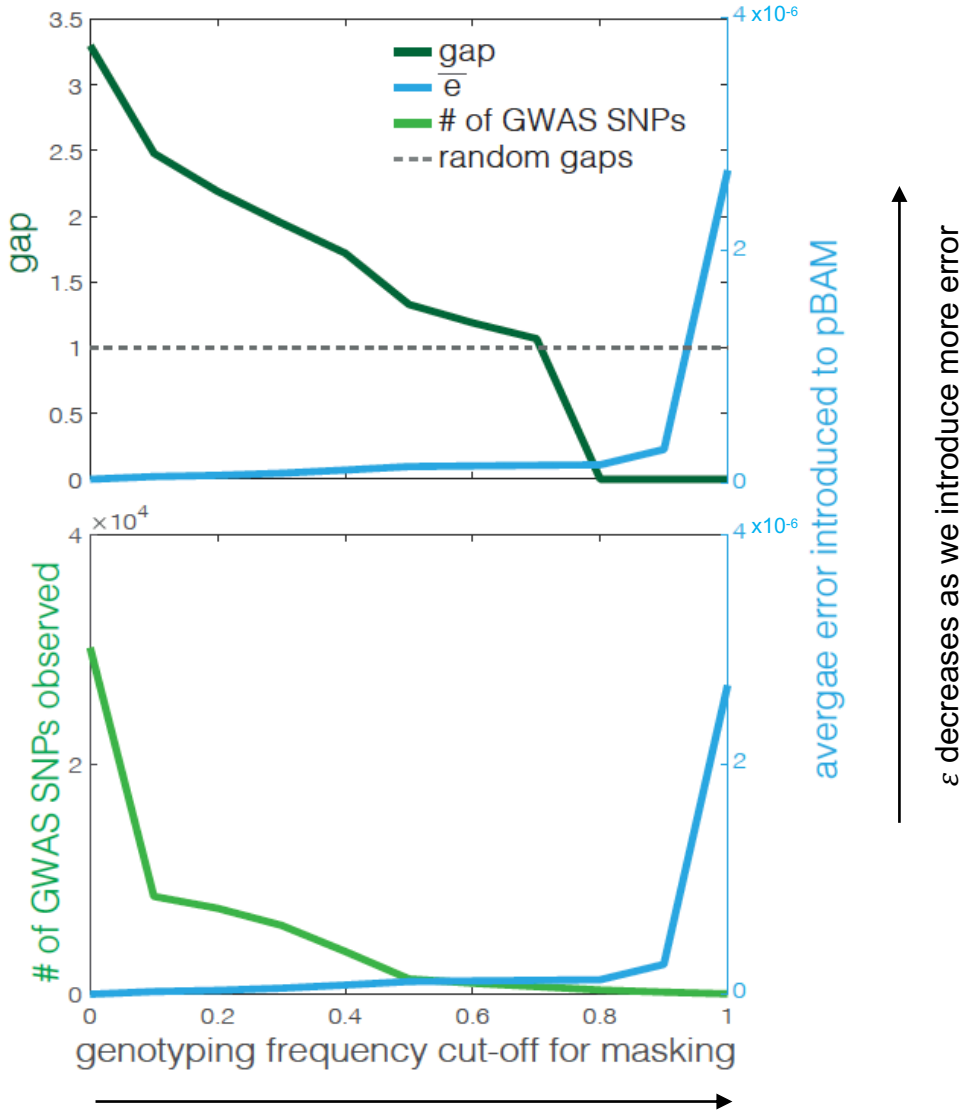
Variation between replicates
Variation between BAM-pBAM

[Gursoy et al., Cell, in press]

Privacy-preserving Binary Alignment Mapping (pBAM)

(grounded in privacy and utility)

- Unit = nucleotide (signal track)
- NA12878 RNA-Seq data
- Test the privacy for each level of masking
- Measure the error introduced



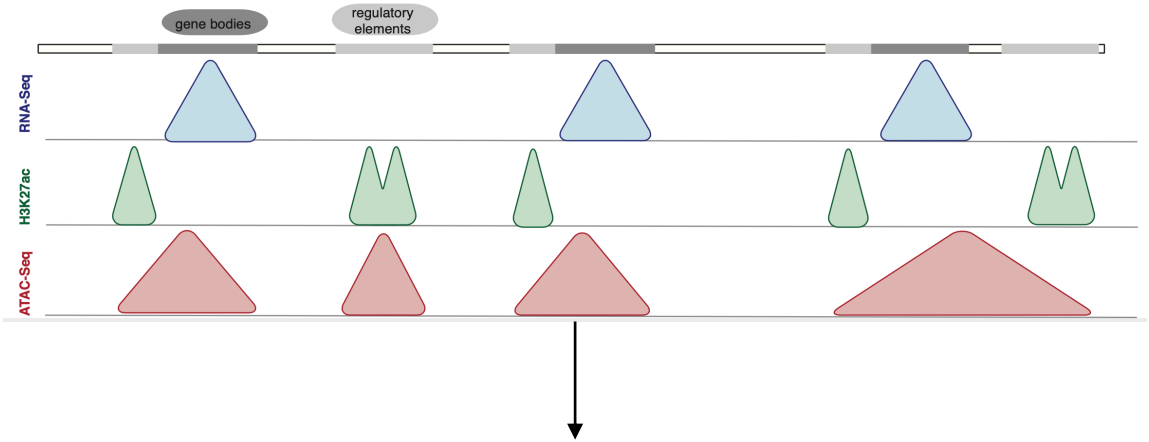
[Gursoy et al., Cell, in press]

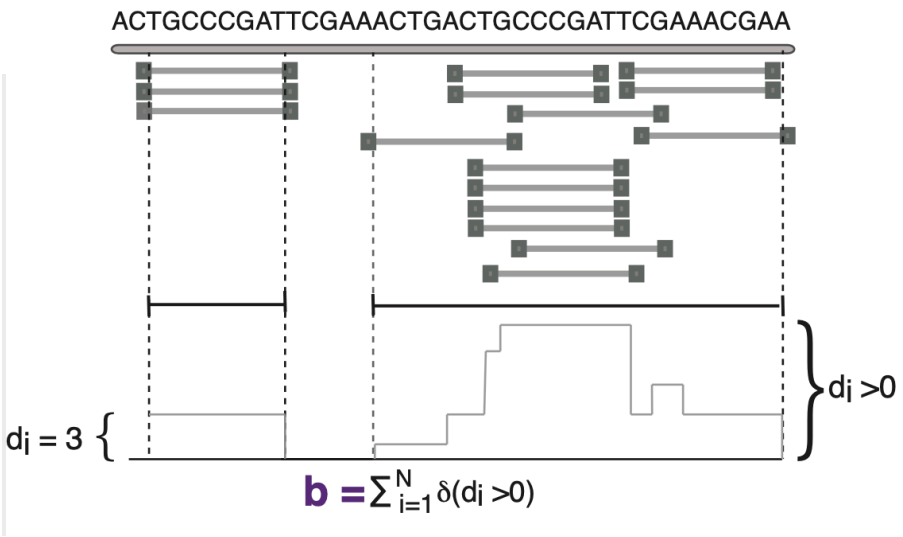
Privacy & Functional Genomics

- Intro. to Genomic Privacy
  - The **dilemma**: The genome as fundamental, inherited info that's very private v. need for large-scale sharing & mining for med. research
- Privacy & Functional Genomics Data
  - **2-sided nature** of this data presents particularly tricky privacy issues
  - Overview of **types of the leakage**, from obvious to subtle
- Obvious leakage #1: **reads**
  - How much leakage can we expect?
    - Quantification with available data & real-world environmental samples
  - Using **pBAM** file format to remove obvious large-scale leakage
  - Using **FANCY** to assess the privacy leakage before release of the data
- Subtle Leakage #2: **eQTLs**
  - Quantifying & removing further variant info from expression levels w/ ICI & predictability.
  - Instantiating a practical linking attack w/ noisy quasi-identifiers
- Subtle Leakage #3: **Signal Profiles**
  - Manifest appreciable leakage from large & small deletions.
  - Linking attacks possible but additional complication of SV discovery in addition to genotyping

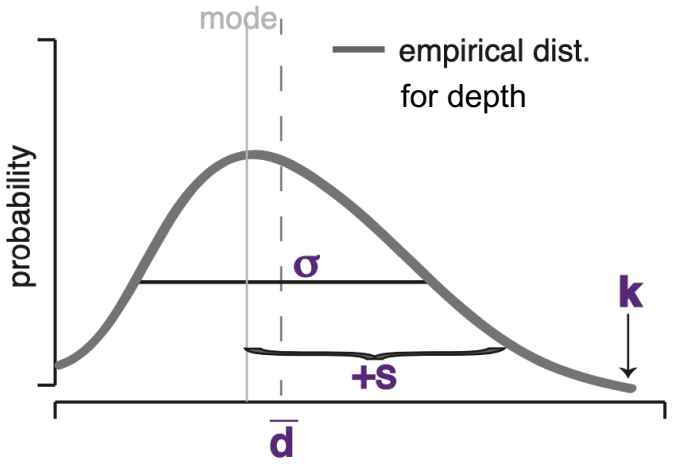
FANCY: Fast Estimation of Privacy Risk in Functional Genomics Data

Can we predict the amount of leakage before the release of the data without the need for genotyping?





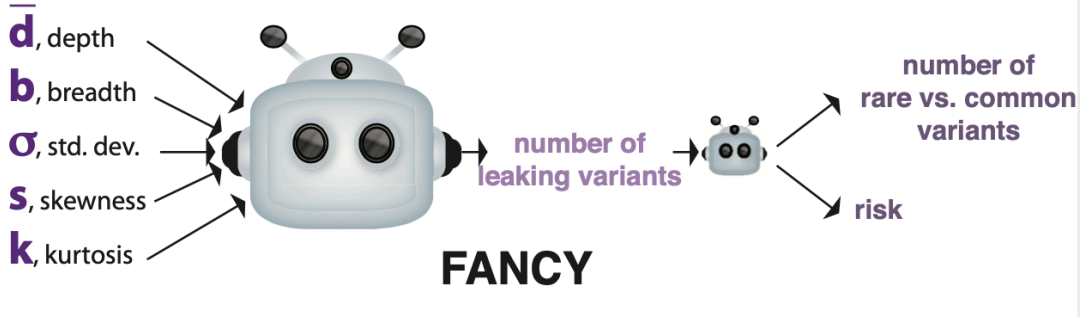
depth & **b**readth

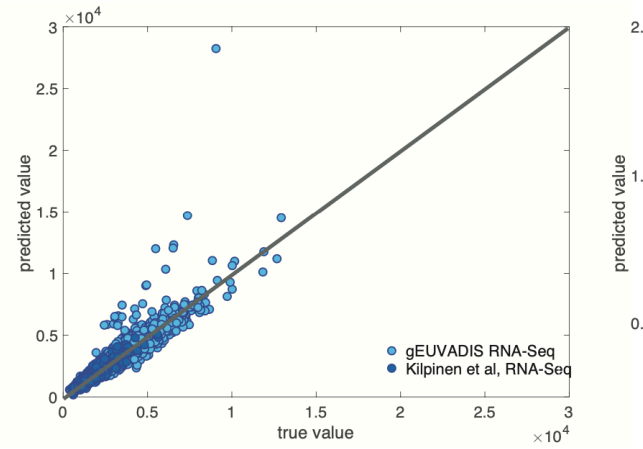


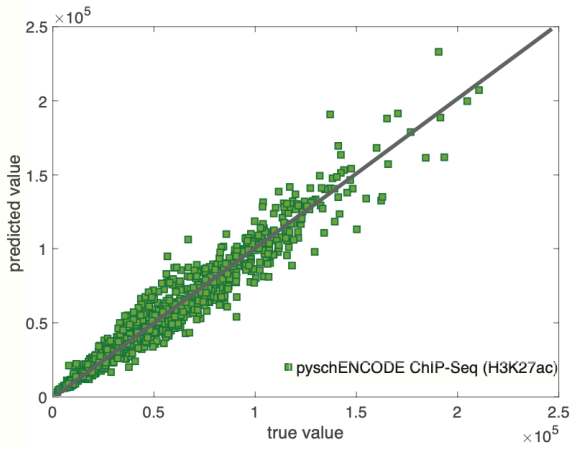
[Gursoy et al., Bioinformatics, 2020]

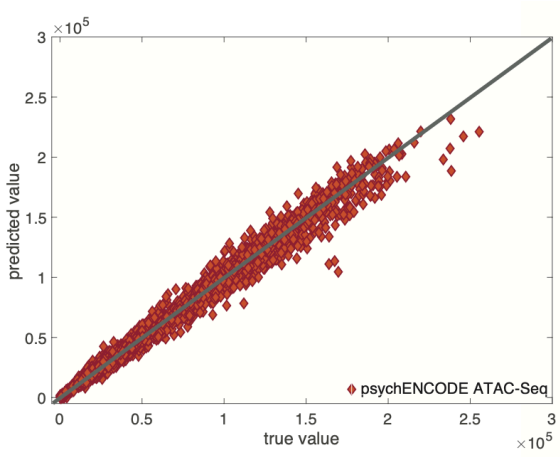
FANCY: Fast Estimation of Privacy Risk in Functional Genomics Data

Can we predict the amount of leakage before the release of the data without the need for genotyping?





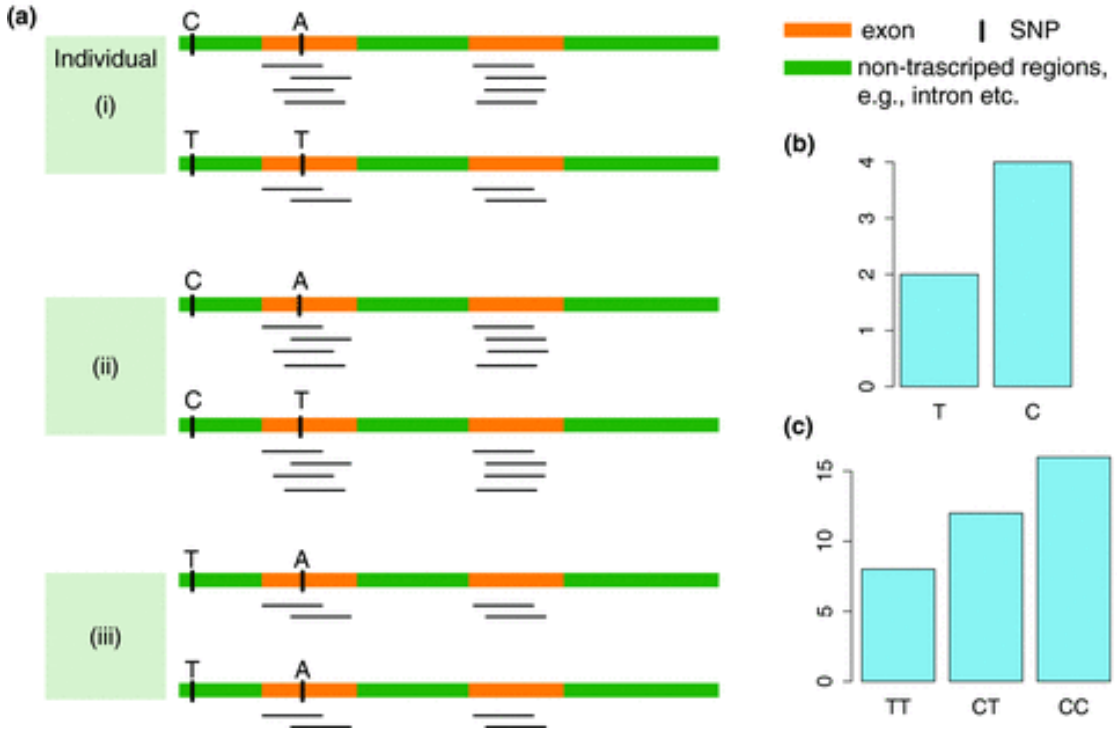




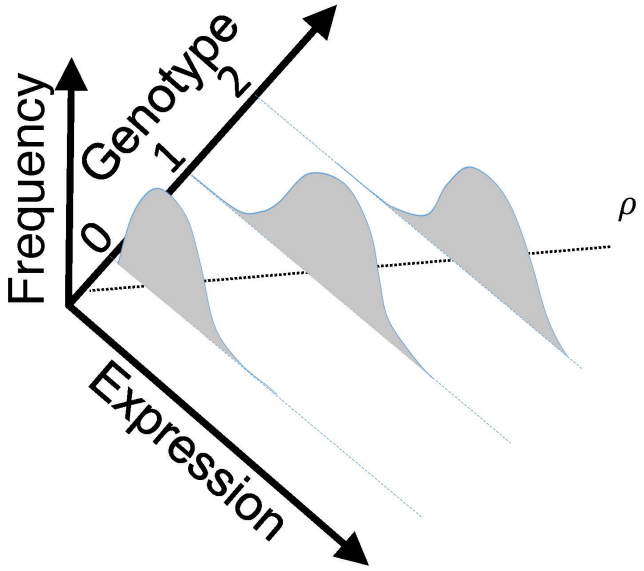
[Gursoy et al., Bioinformatics, 2020]

Privacy & Functional Genomics

- Intro. to Genomic Privacy
  - The **dilemma**: The genome as fundamental, inherited info that's very private v. need for large-scale sharing & mining for med. research
- Privacy & Functional Genomics Data
  - **2-sided nature** of this data presents particularly tricky privacy issues
  - Overview of **types of the leakage**, from obvious to subtle
- Obvious leakage #1: **reads**
  - How much leakage can we expect?
    - Quantification with available data & real-world environmental samples
  - Using **pBAM** file format to remove obvious large-scale leakage
  - Using **FANCY** to assess the privacy leakage before release of the data
- Subtle Leakage #2: **eQTLs**
  - Quantifying & removing further variant info from expression levels w/ ICI & predictability.
  - Instantiating a practical linking attack w/ noisy quasi-identifiers
- Subtle Leakage #3: **Signal Profiles**
  - Manifest appreciable leakage from large & small deletions.
  - Linking attacks possible but additional complication of SV discovery in addition to genotyping



[*Biometrics 68(1) 1–11*]

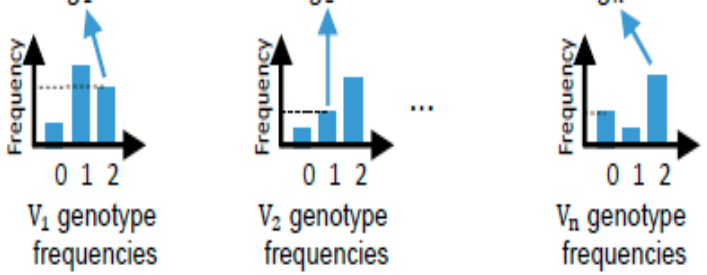


eQTL Mapping Using RNA-Seq Data

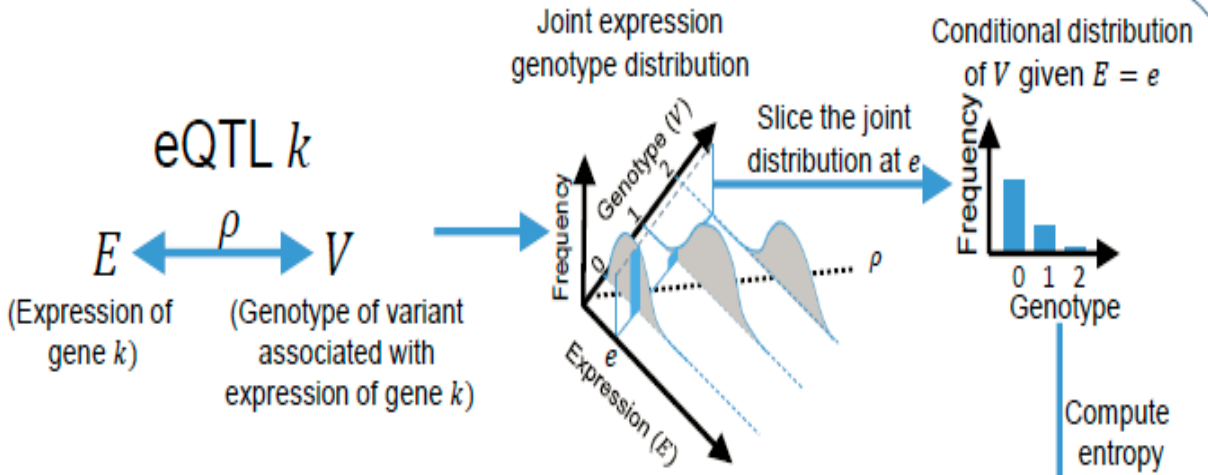
- eQTLs are genomic loci that contribute to variation in mRNA expression levels
- eQTLs provide insights on transcription regulation, and the molecular basis of phenotypic outcomes
- eQTL mapping can be done with RNA-Seq data

Information Content and Predictability

$$ICI\left(\begin{array}{l}\text{Individual has variant}\\ \text{genotypes } g_1, g_2, \ldots, g_n\\ \text{for variants } V_1, V_1, \ldots, V_n\end{array}\right) = \log\left(\frac{1}{\text{Frequency of } V_1 \text{ genotype } g_1 = 2}\right) + \log\left(\frac{1}{\text{Frequency of } V_2 \text{ genotype } g_2 = 1}\right) + \ldots + \log\left(\frac{1}{\text{Frequency of } V_n \text{ genotype } g_n = 2}\right)$$



- Naive measure of information (no LD, distant correlations, pop. struc., &c)
- Higher frequency: Lower ICI
- Additive for multiple variants

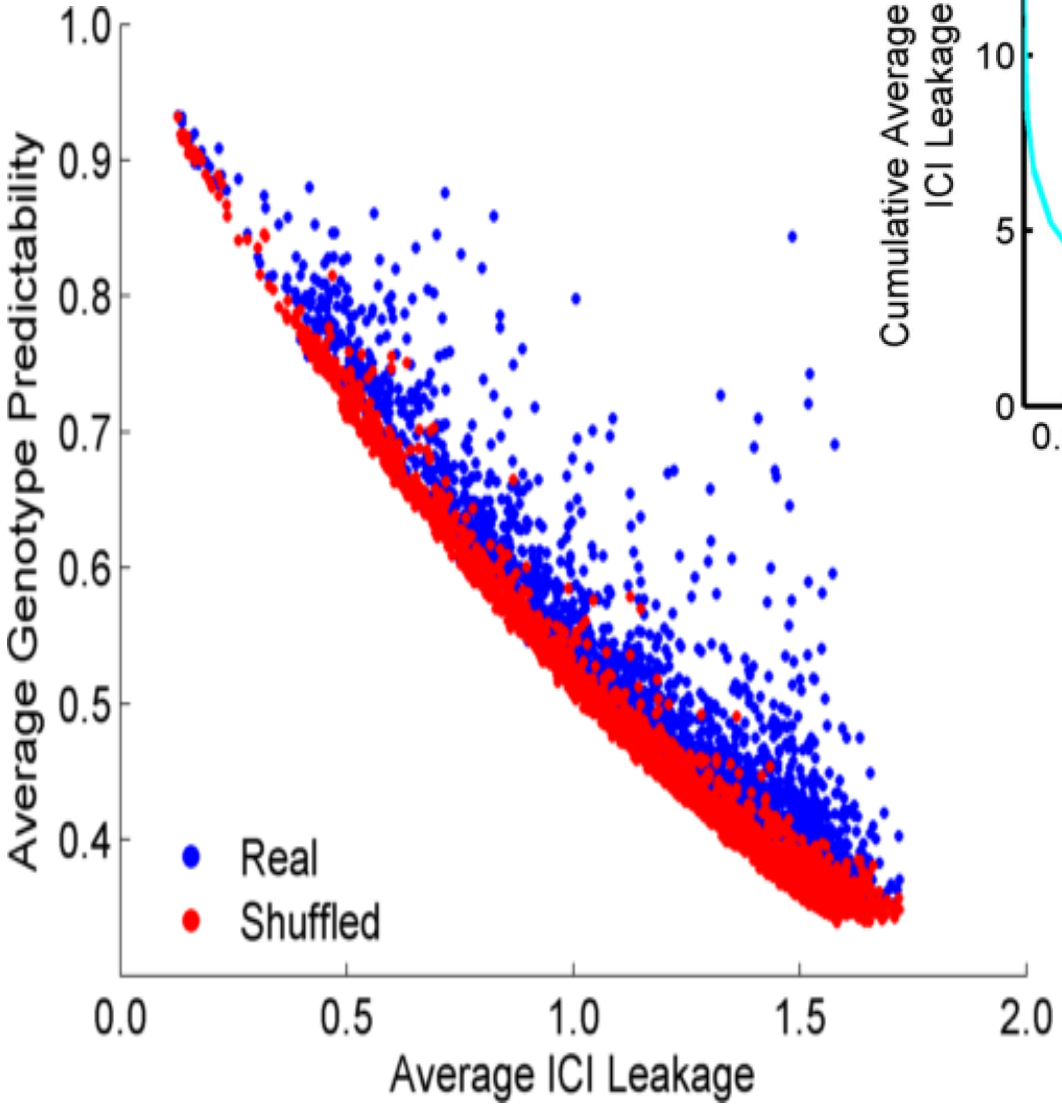


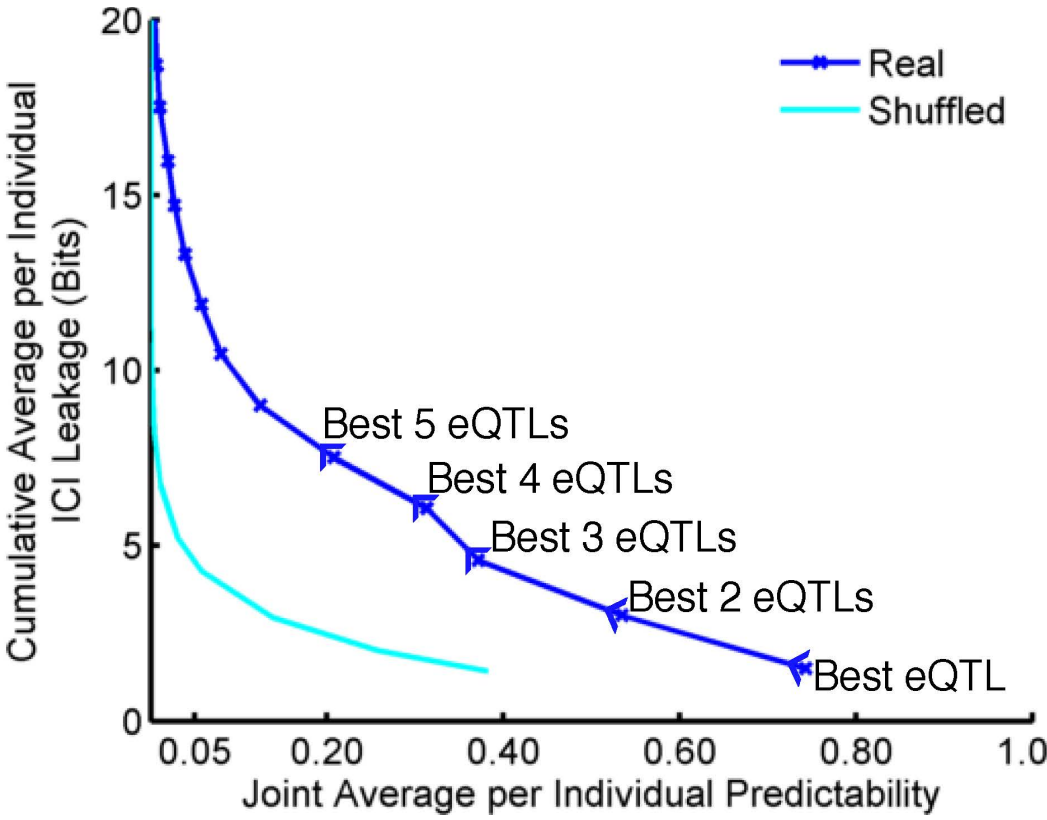
$$\pi\left(V \mid E = e\right) = e^{-H\left(V \mid E = e\right)}$$

- Condition specific entropy
- Higher cond. entropy: Lower predictability
- Additive for multiple eQTLs

[Harmanci et al. Nat. Meth. 2016]

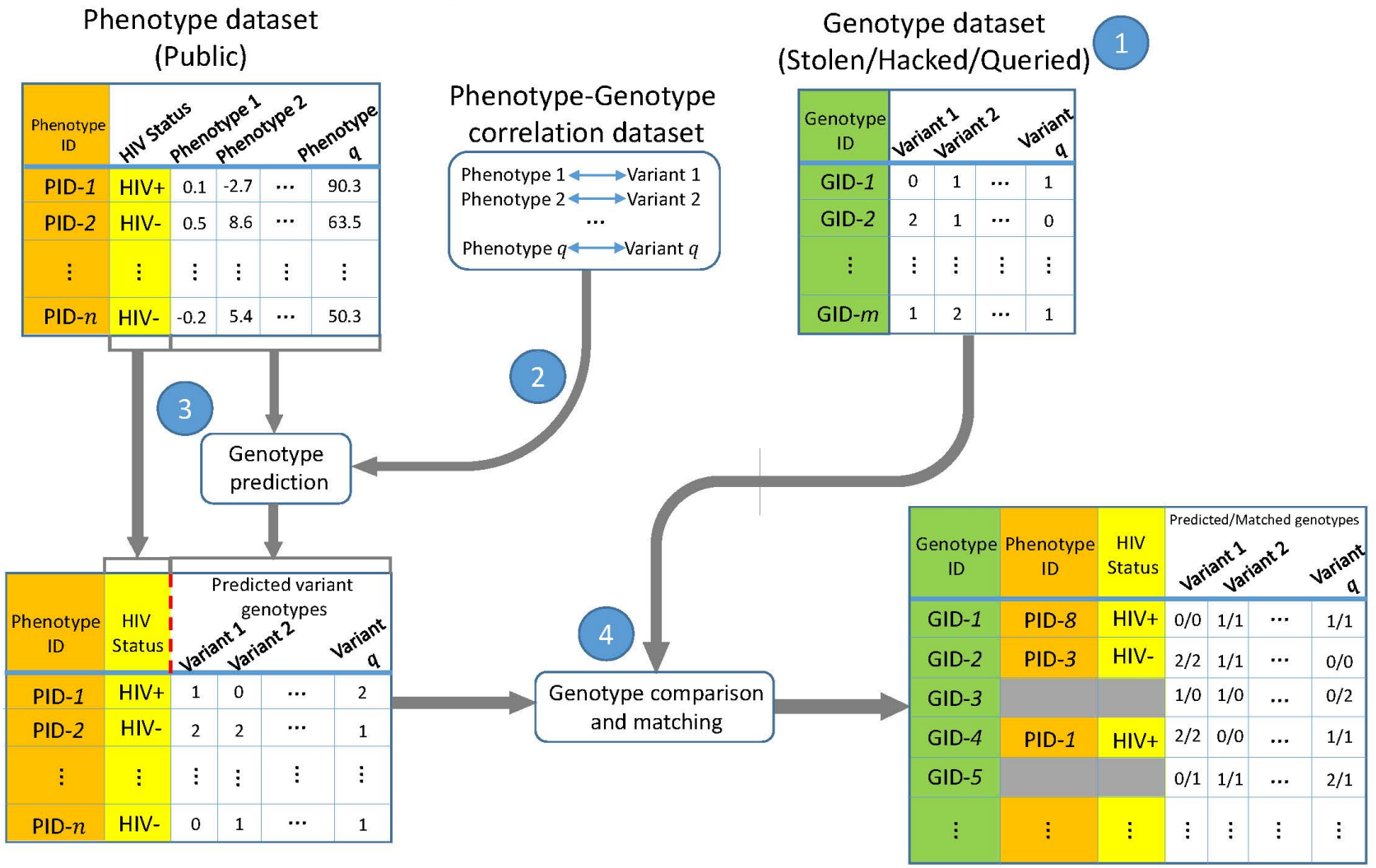
[Harmanciet al. Nat. Meth. (‘16]





ICI Leakage versus Genotype Predictability

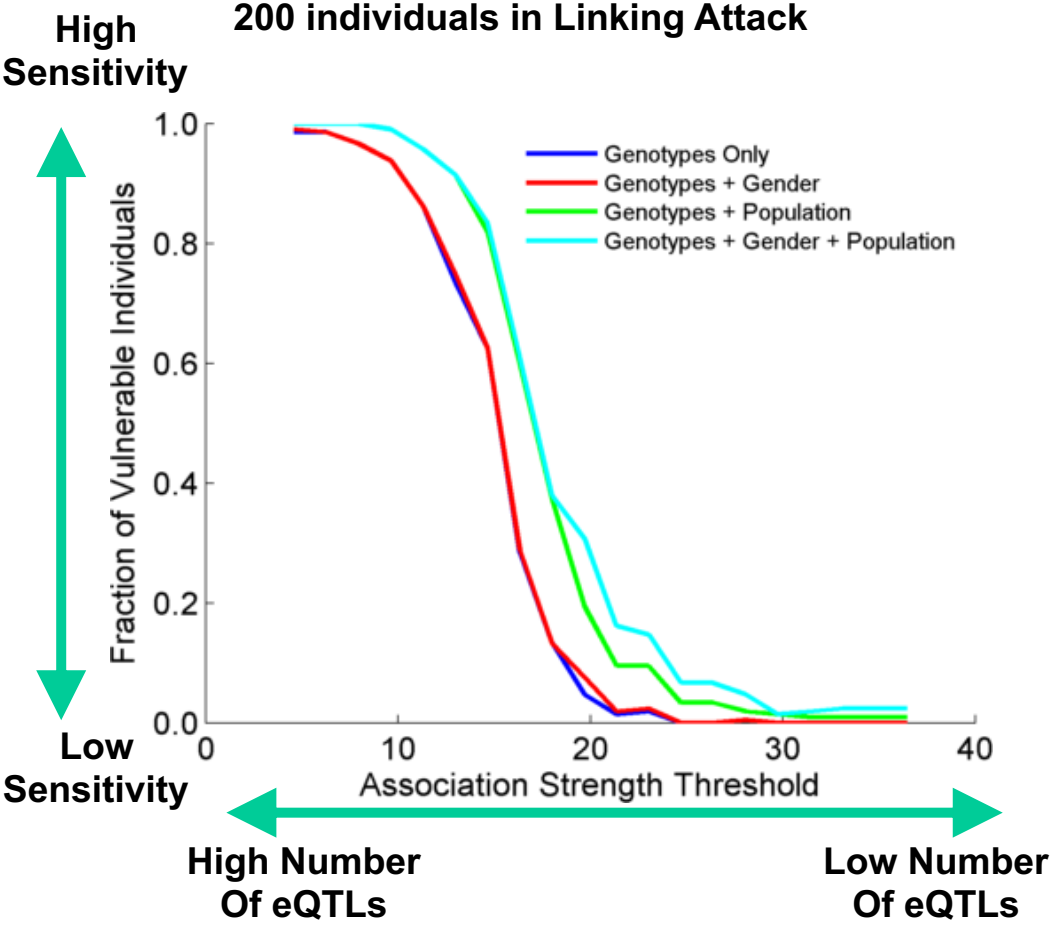
Linking Attack Scenario



[Harmanciet al. Nat. Meth. (‘16)]

Success in Linking Attack with Extremity based Genotype Prediction

200 individuals eQTL Discovery
200 individuals in Linking Attack

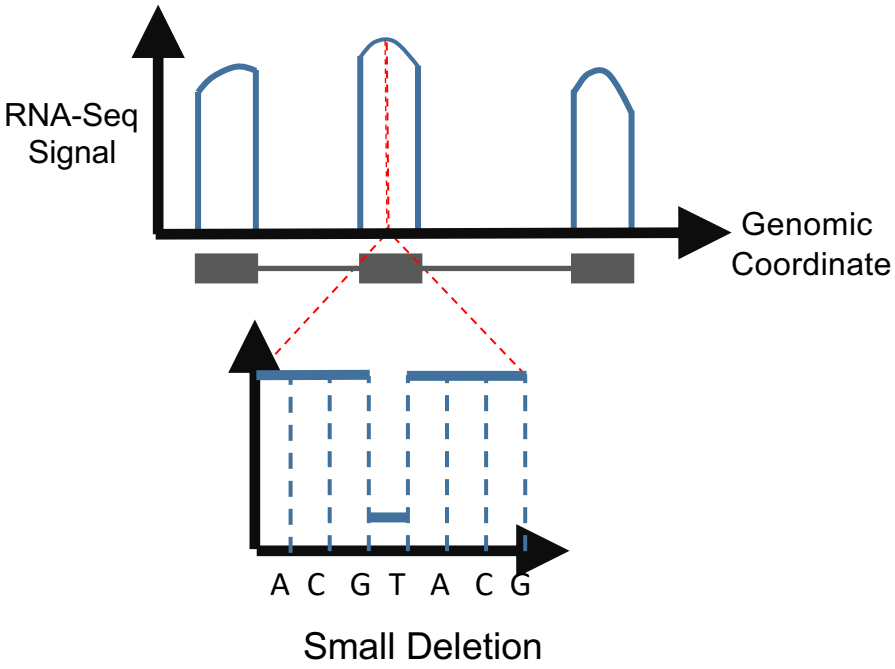


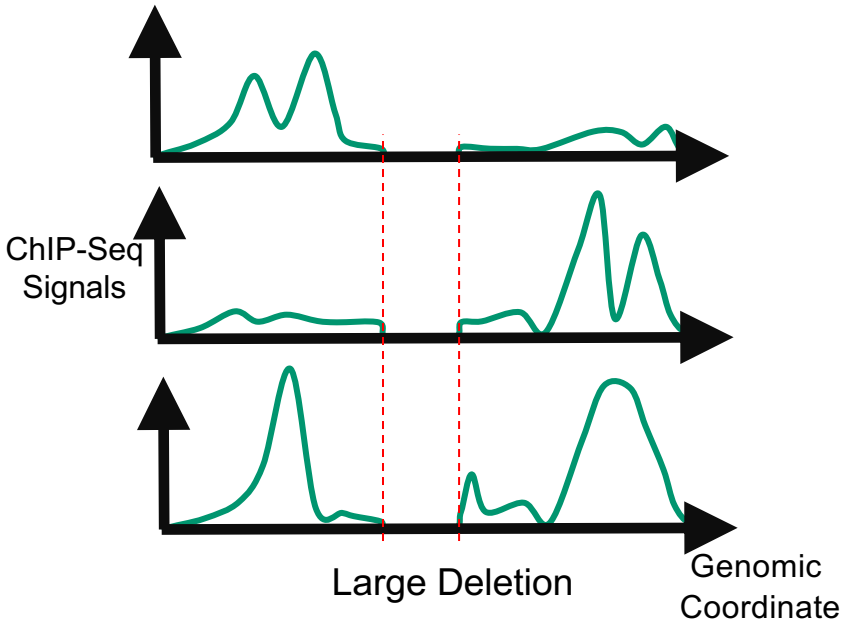
[Harmanci et al. *Nat. Meth.* (16)]

Privacy & Functional Genomics

- Intro. to Genomic Privacy
  - The **dilemma**: The genome as fundamental, inherited info that's very private v. need for large-scale sharing & mining for med. research
- Privacy & Functional Genomics Data
  - **2-sided nature** of this data presents particularly tricky privacy issues
  - Overview of **types of the leakage**, from obvious to subtle
- Obvious leakage #1: **reads**
  - How much leakage can we expect?
    - Quantification with available data & real-world environmental samples
  - Using **pBAM** file format to remove obvious large-scale leakage
  - Using **FANCY** to assess the privacy leakage before release of the data
- Subtle Leakage #2: **eQTLs**
  - Quantifying & removing further variant info from expression levels w/ ICI & predictability.
  - Instantiating a practical linking attack w/ noisy quasi-identifiers
- Subtle Leakage #3: **Signal Profiles**
  - Manifest appreciable leakage from large & small deletions.
  - Linking attacks possible but additional complication of SV discovery in addition to genotyping

Detection & Genotyping of small & large SV deletions from signal profiles

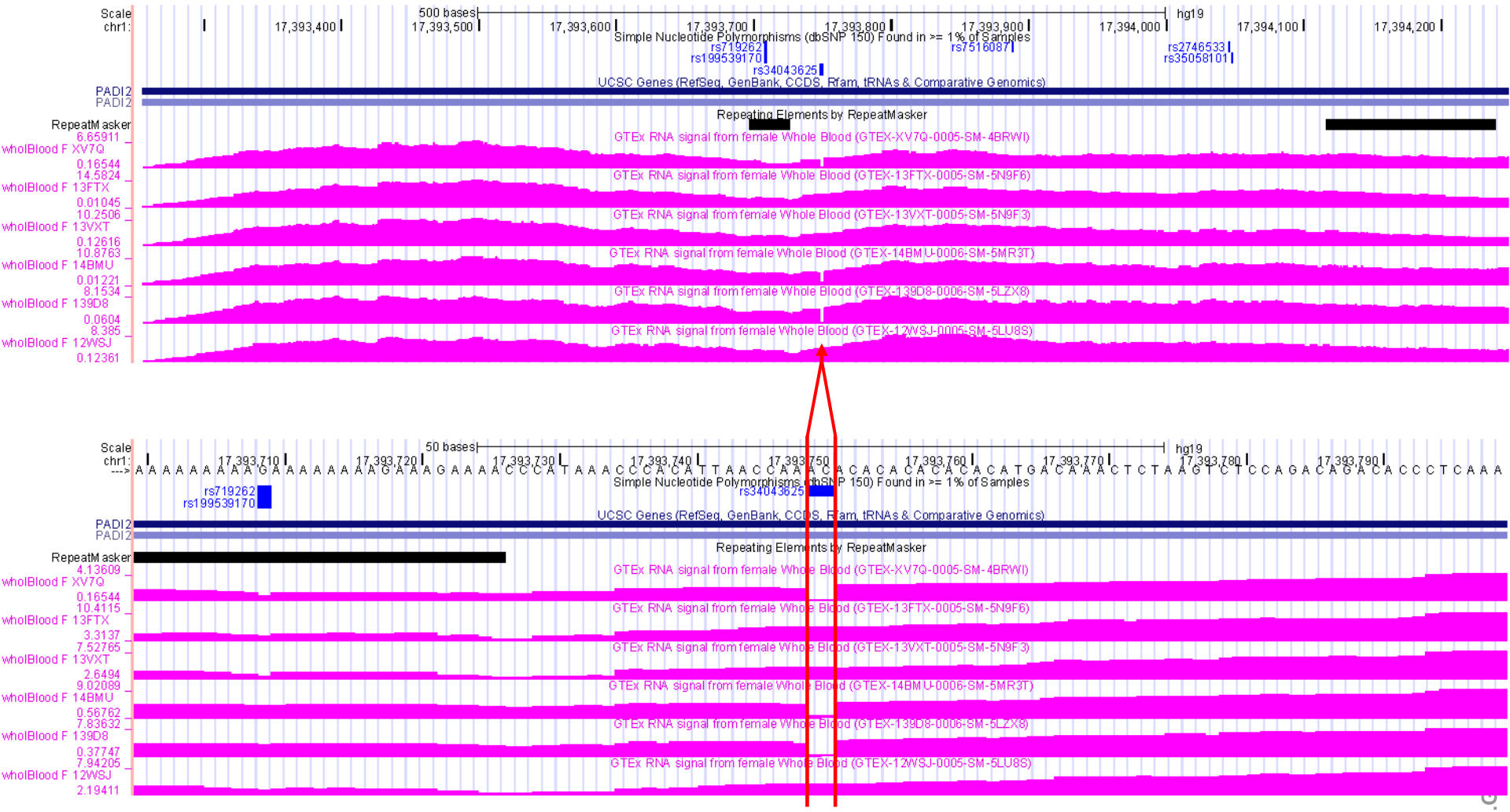




RNA-seq also shows large deletions

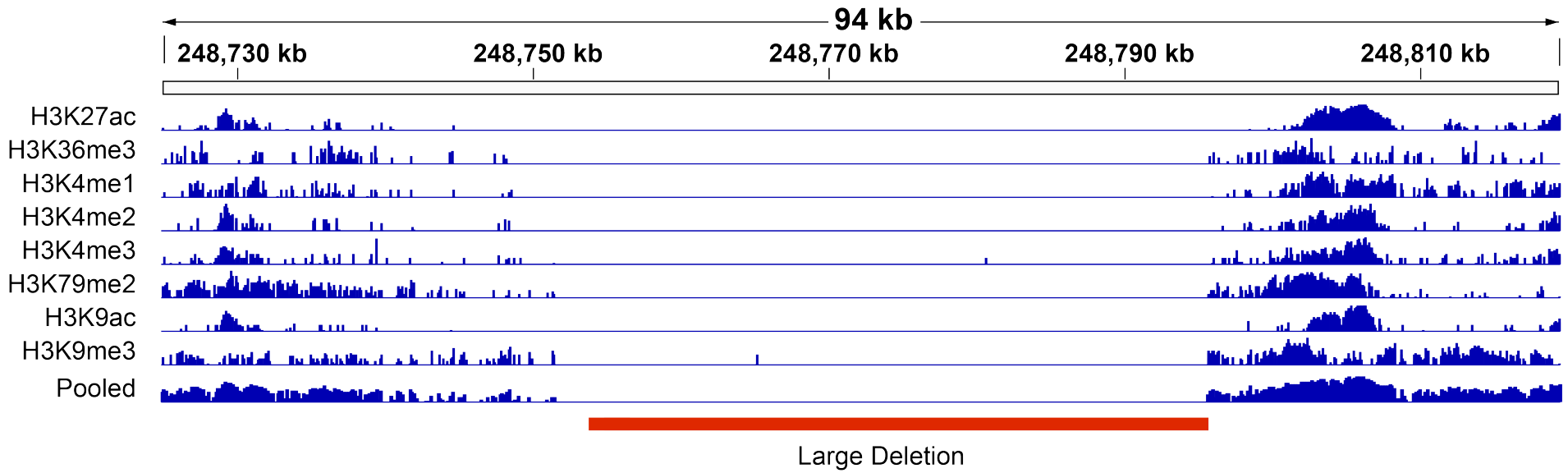
[Harmanci & Gerstein, *Nat. Comm.* ('18)]

Example of Small Deletion Evident in Signal Profile



[Harmanci & Gerstein, *Nat. Comm.* ('18)]

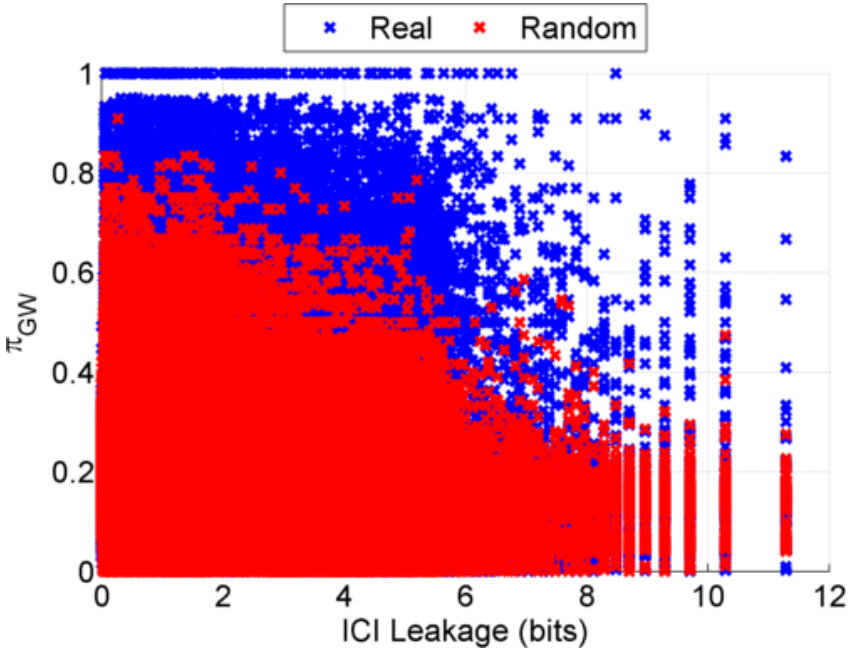
Example of Large Deletion Evident in Signal Profile



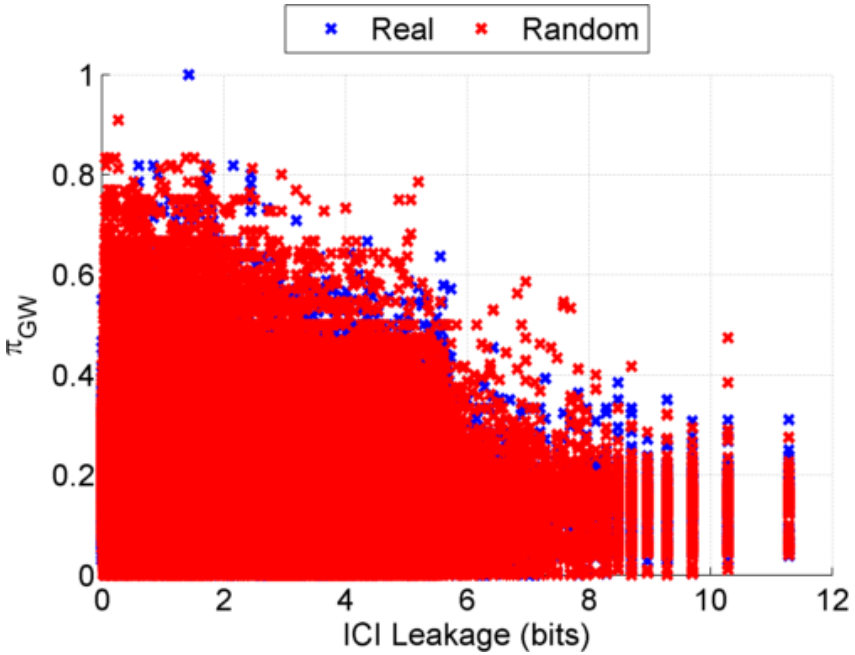
[Harmanci & Gerstein, *Nat. Comm.* ('18)]

Information Leakage from SV Deletions

a) Before Anonymization



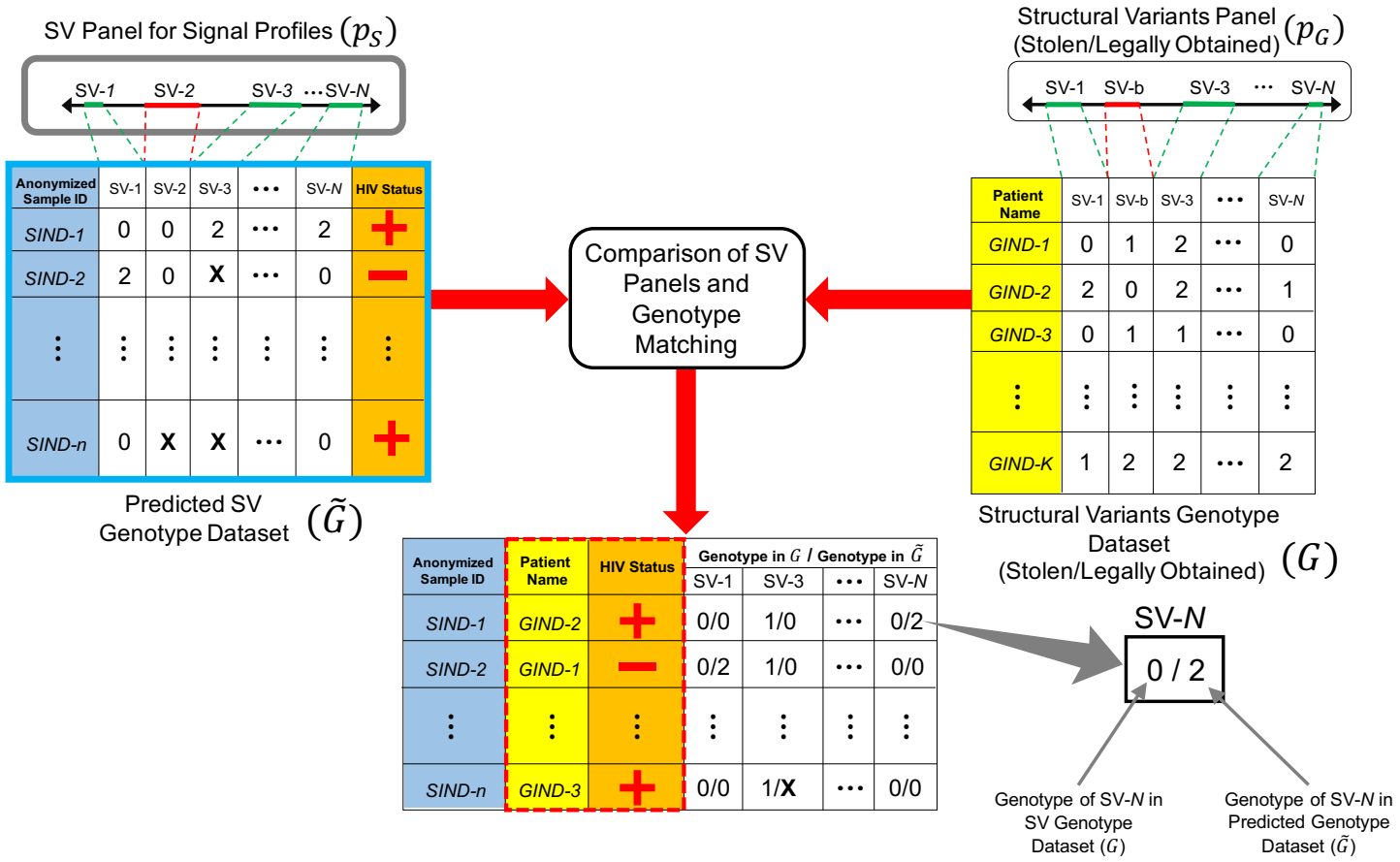
b) After Anonymization



Simple anonymization procedure (filling in deletion by value at endpoints) has dramatic effect

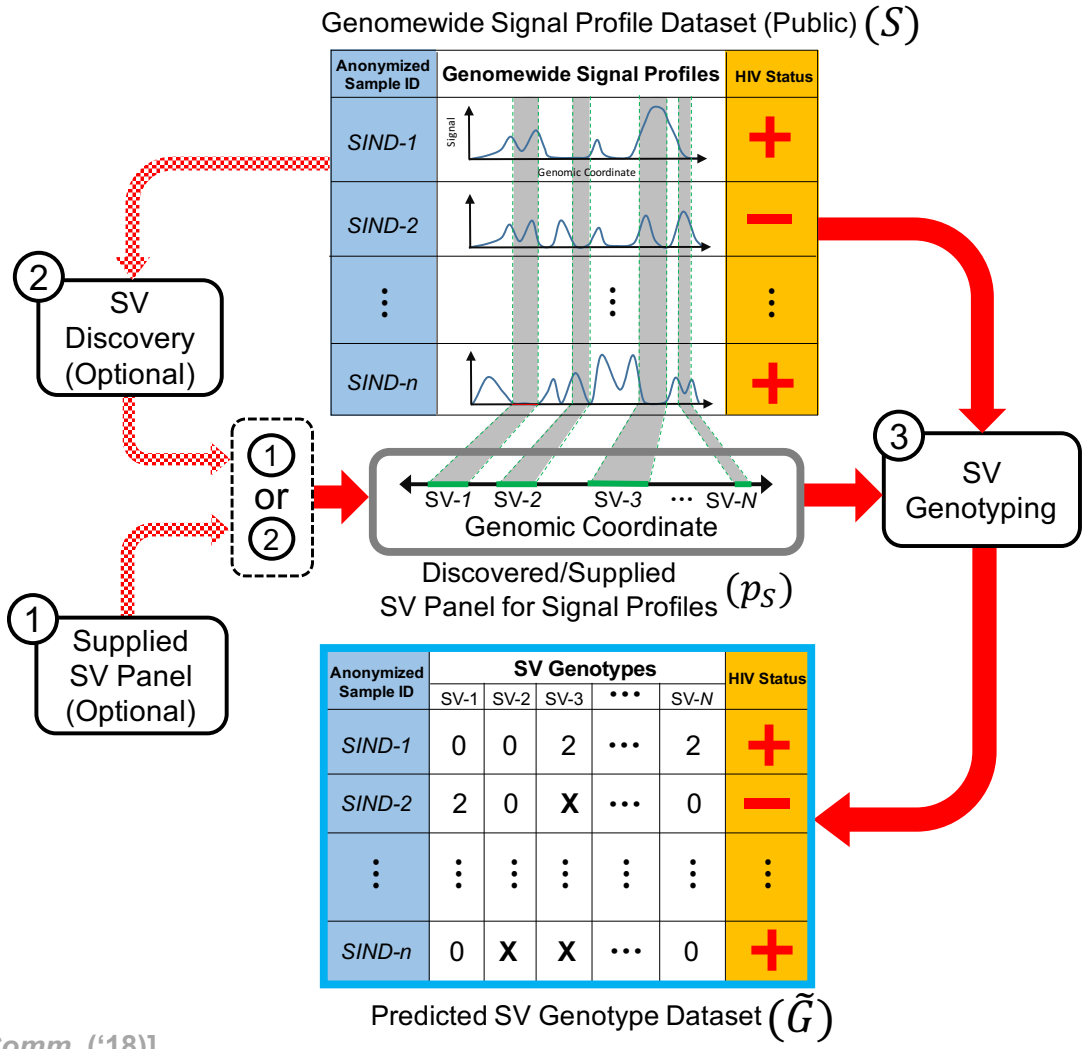
[Harmanci & Gerstein, *Nat. Comm.* ('18)]

Another type of Linking Attack: Linking based on SV Genotyping



[Harmanci & Gerstein, *Nat. Comm.* ('18)]

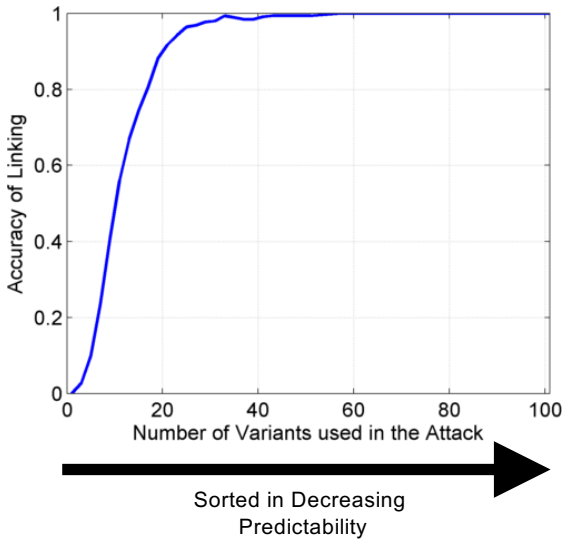
Another type of Linking Attack: First Doing SV Genotyping



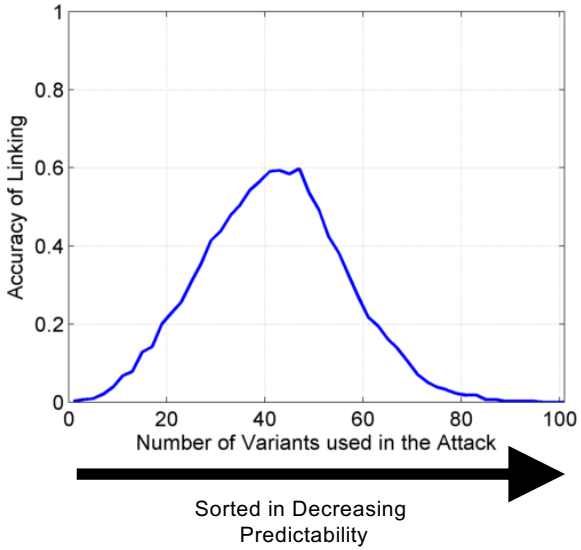
[Harmanci & Gerstein, *Nat. Comm.* ('18)]

Linking Attack Based on SV Deletions in gEUVADIS Dataset

c) Genotyping
(1kG MAF>0.01)



d) Discovery + Genotyping



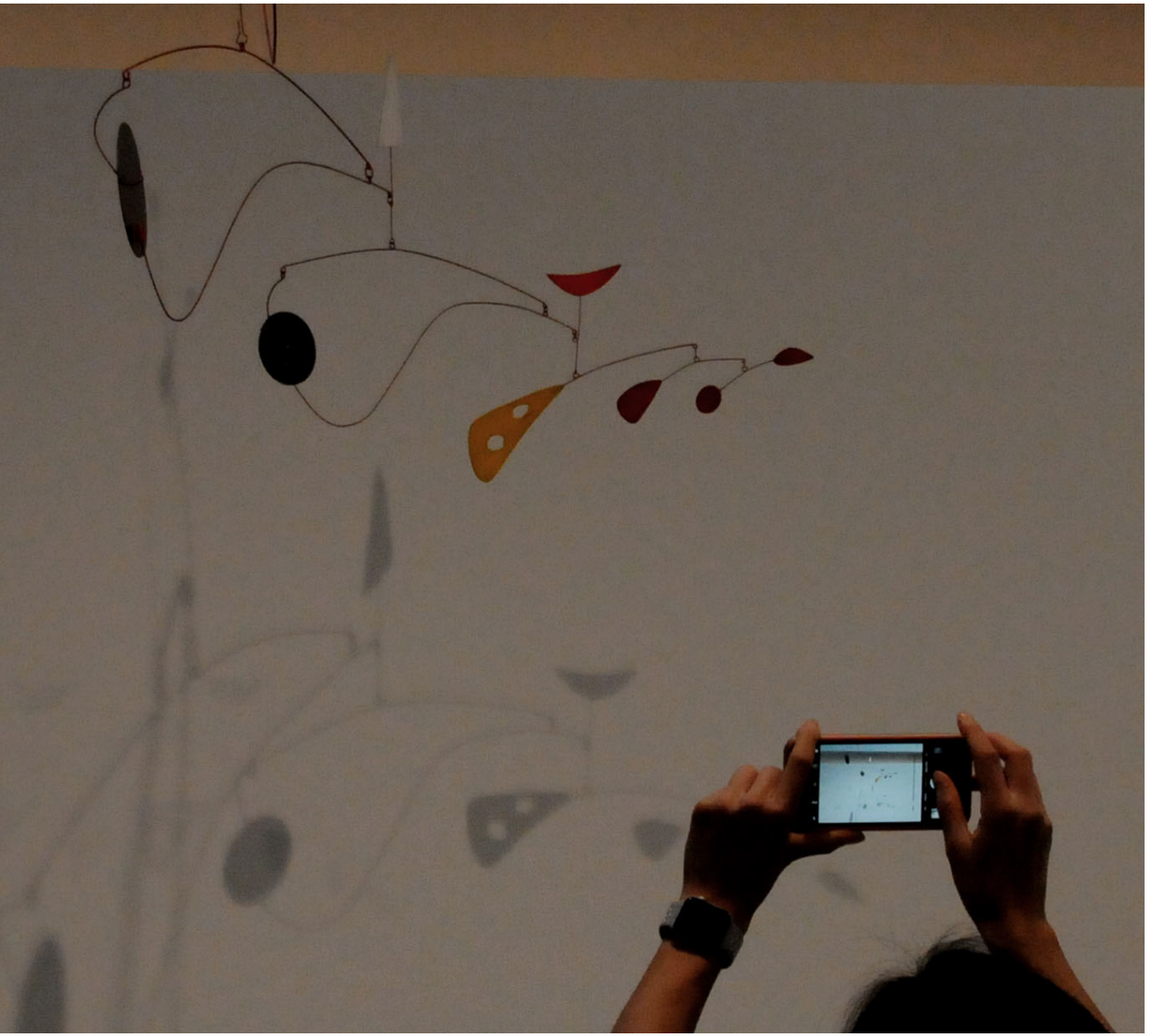
[Harmanci & Gerstein, *Nat. Comm.* ('18)]

Privacy & Functional Genomics

- Intro. to Genomic Privacy
  - The **dilemma**: The genome as fundamental, inherited info that's very private v. need for large-scale sharing & mining for med. research
- Privacy & Functional Genomics Data
  - **2-sided nature** of this data presents particularly tricky privacy issues
  - Overview of **types of the leakage**, from obvious to subtle
- Obvious leakage #1: **reads**
  - How much leakage can we expect?
    - Quantification with available data & real-world environmental samples
  - Using **pBAM** file format to remove obvious large-scale leakage
  - Using **FANCY** to assess the privacy leakage before release of the data
- Subtle Leakage #2: **eQTLs**
  - Quantifying & removing further variant info from expression levels w/ ICI & predictability.
  - Instantiating a practical linking attack w/ noisy quasi-identifiers
- Subtle Leakage #3: **Signal Profiles**
  - Manifest appreciable leakage from large & small deletions.
  - Linking attacks possible but additional complication of SV discovery in addition to genotyping

Privacy & Functional Genomics

- Intro. to Genomic Privacy
  - The **dilemma**: The genome as fundamental, inherited info that's very private v. need for large-scale sharing & mining for med. research
- Privacy & Functional Genomics Data
  - **2-sided nature** of this data presents particularly tricky privacy issues
  - Overview of **types of the leakage**, from obvious to subtle
- Obvious leakage #1: **reads**
  - How much leakage can we expect?
    - Quantification with available data & real-world environmental samples
  - Using **pBAM** file format to remove obvious large-scale leakage
  - Using **FANCY** to assess the privacy leakage before release of the data
- Subtle Leakage #2: **eQTLs**
  - Quantifying & removing further variant info from expression levels w/ ICI & predictability.
  - Instantiating a practical linking attack w/ noisy quasi-identifiers
- Subtle Leakage #3: **Signal Profiles**
  - Manifest appreciable leakage from large & small deletions.
  - Linking attacks possible but additional complication of SV discovery in addition to genotyping



Acknowledgements

G Gürsoy, A Harmanci, D Greenbaum, P Emani, C Brannon, S Strattan, O Jolanki, M Cherry, A Miranker, F Navarro

papers.gersteinlab.org/subject/**privacy**

PrivaSeq3.gersteinlab.org

FANCY.gersteinlab.org

PrivaSeq.gersteinlab.org

PrivaSig.gersteinlab.org

Also:

JOBS.gersteinlab.org

Extra

Info about content in this slide pack

No Conflicts

Unless explicitly listed below,
there are no conflicts of interest relevant to the material in this talk

General PERMISSIONS

- This Presentation is copyright Mark Gerstein, Yale University, 2019.
- Please read permissions statement at
  sites.gersteinlab.org/Permissions
- Basically, feel free to use slides & images in the talk with PROPER acknowledgement (via citation to relevant papers or website link). Paper references in the talk were mostly from Papers.GersteinLab.org.

PHOTOS & IMAGES

For thoughts on the source and permissions of many of the photos and clipped images in this presentation see streams.gerstein.info . In particular, many of the images have particular EXIF tags, such as kwpotppt , that can be easily queried from flickr, viz: flickr.com/photos/mbgmbg/tags/kwpotppt