

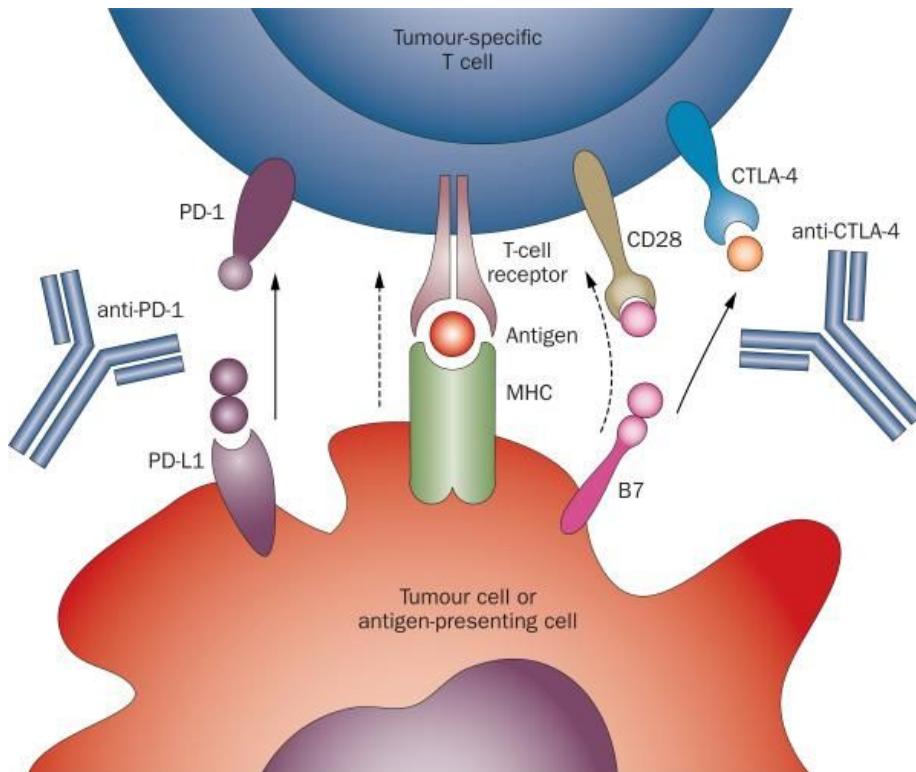
PQE Presentation

Vinay Viswanadham
September 5, 2018

OUTLINE

- Introduction to topic and review of papers (q#1)
- Validation of the gene signature (q#2 and q#3)
- ERV gene expression and immune infiltration (q#5)
- Review of ERV identification methods and proposal (q#4)
- Immunotherapy response and clinical applicability (q#6 and q#7)

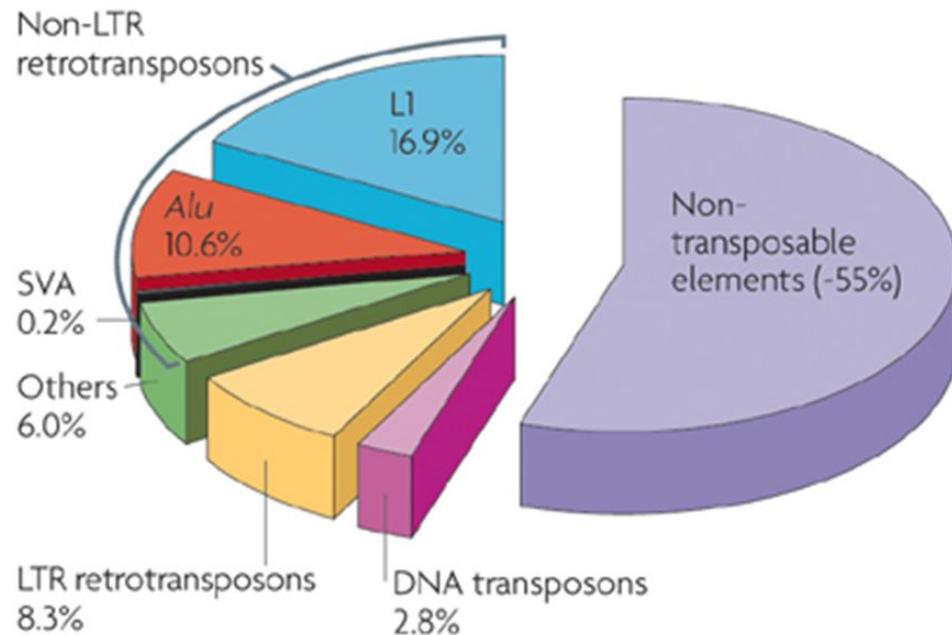
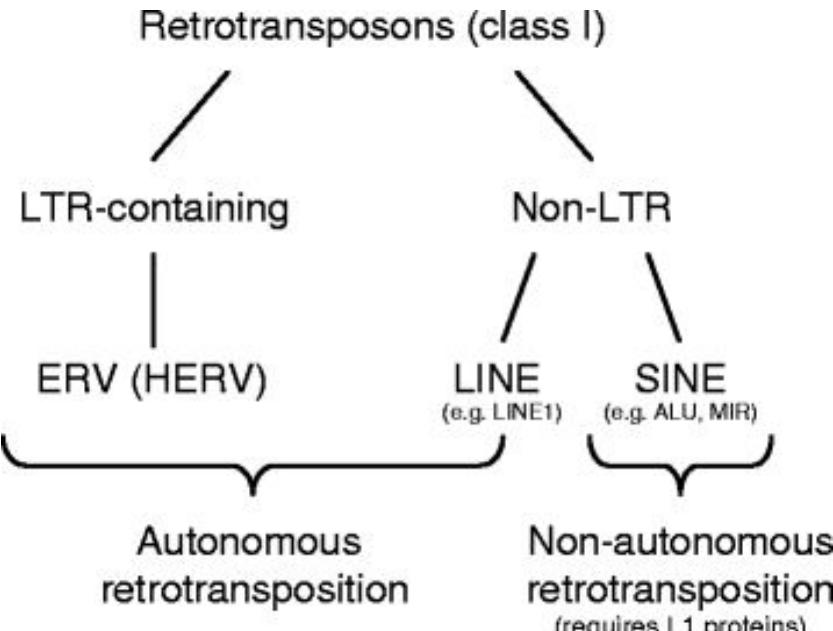
Tumor-immune interactions



- Nucleated cells present internal proteins (“endogenous antigens”) on **Major Histocompatibility Complex I (MHC)**; TCR-MHC interactions provide **primary signals to T-cells**
- Cytokines and chemokines provide **secondary signals** to sustain interactions
- Checkpoint blockade currently targets **PD-1 and CTLA-4**
 - **PD-1:** on CD8 T-cells (Pembrolizumab, Nivolumab)
 - **CTLA-4:** on CD4 T-cells and activated CD8 T-cells (Ipilimumab)

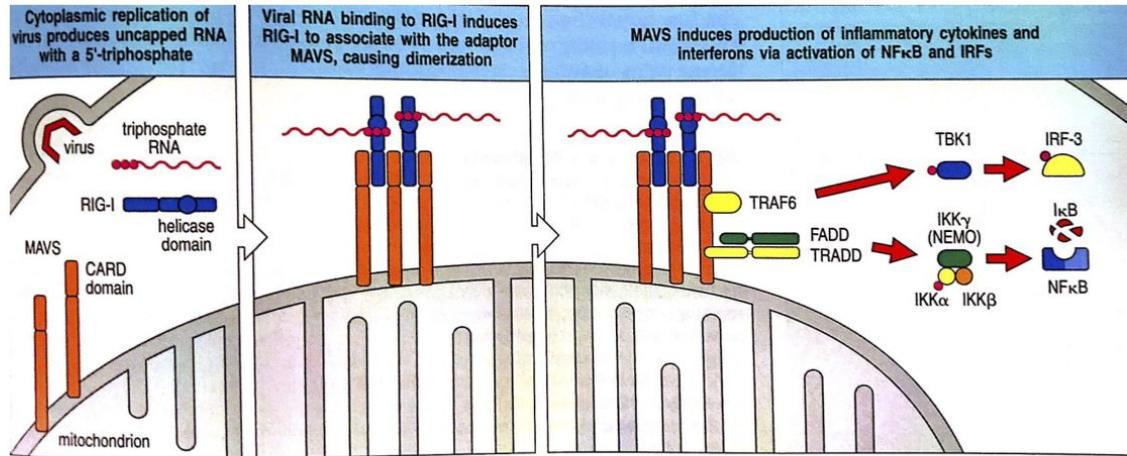
Drake et al Nat Rev Clin Oncol 2014

Endogenous Retroviruses (ERVs)



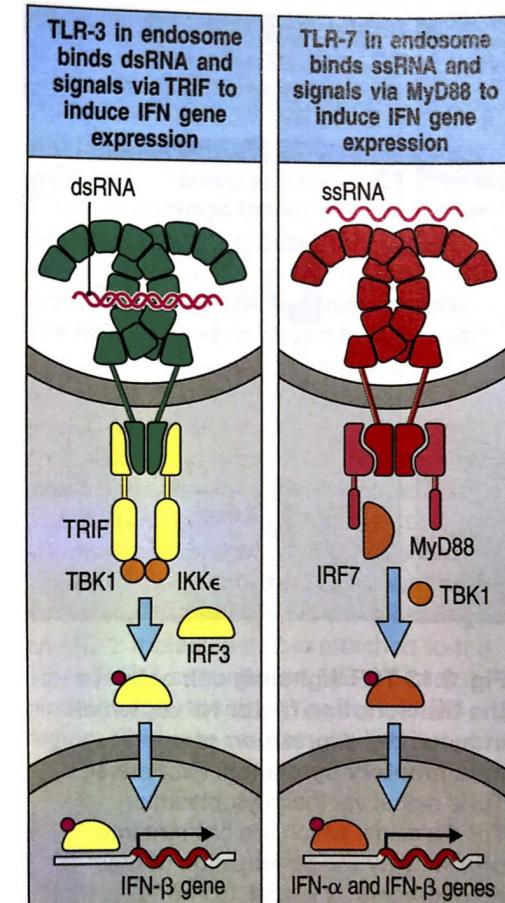
Wolff et al Cell Comm Sig 2017

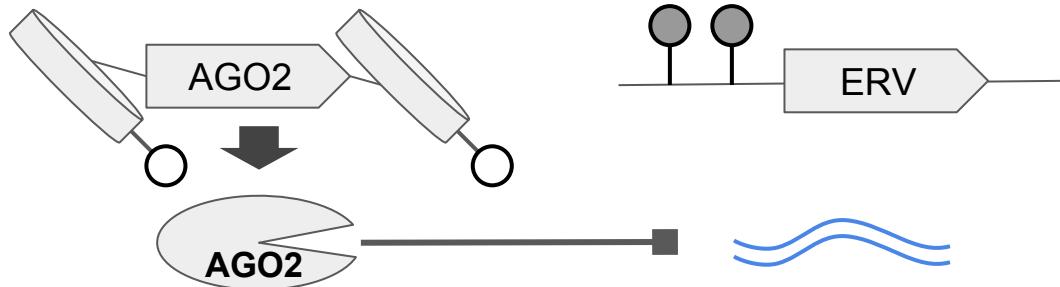
Innate immunity against viruses



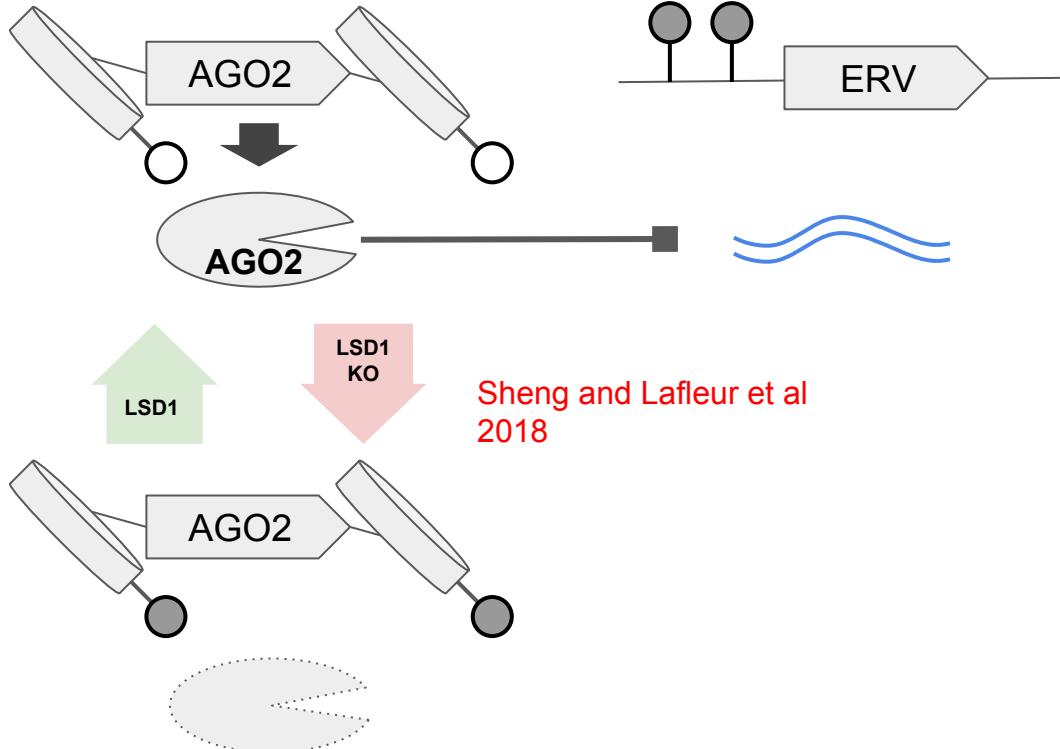
- TLR3 and TLR7 endosomal receptors sense viral dsRNA or ssRNA contained in vesicles and set off the **interferon response**
- RIG-I senses cytosolic dsRNA and orchestrates **interferon response**

Janeway Immunobiology 9th edition





Common findings



Common findings

Q#1

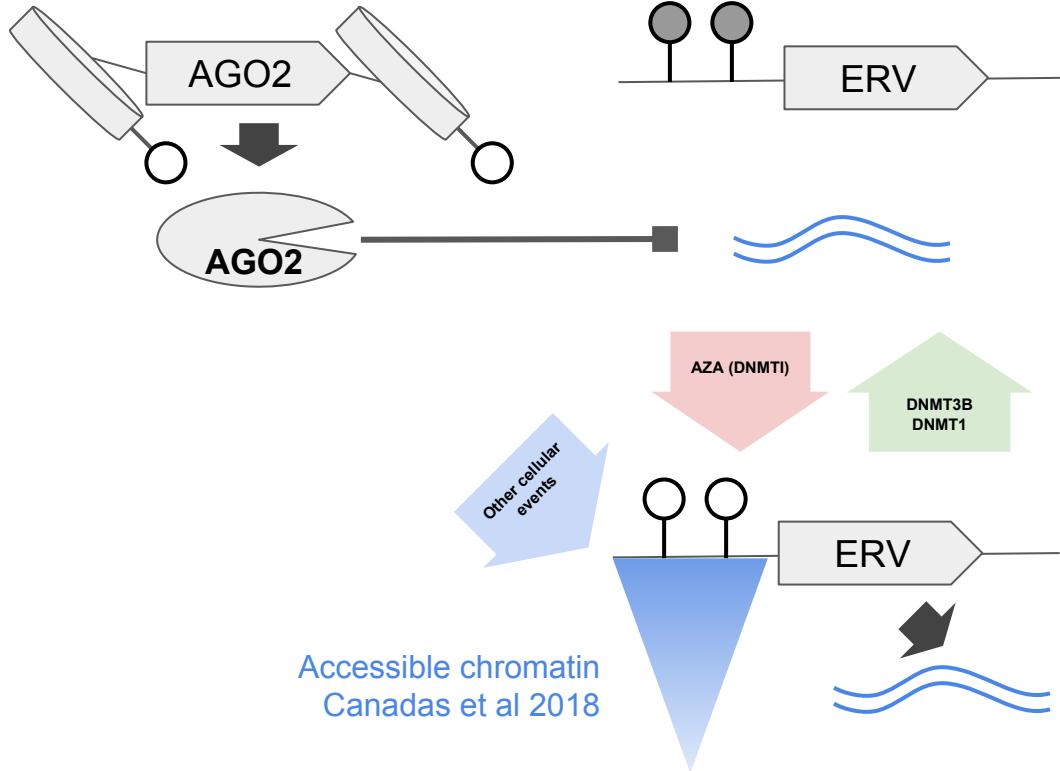
Q#2 + Q#3

Q#5

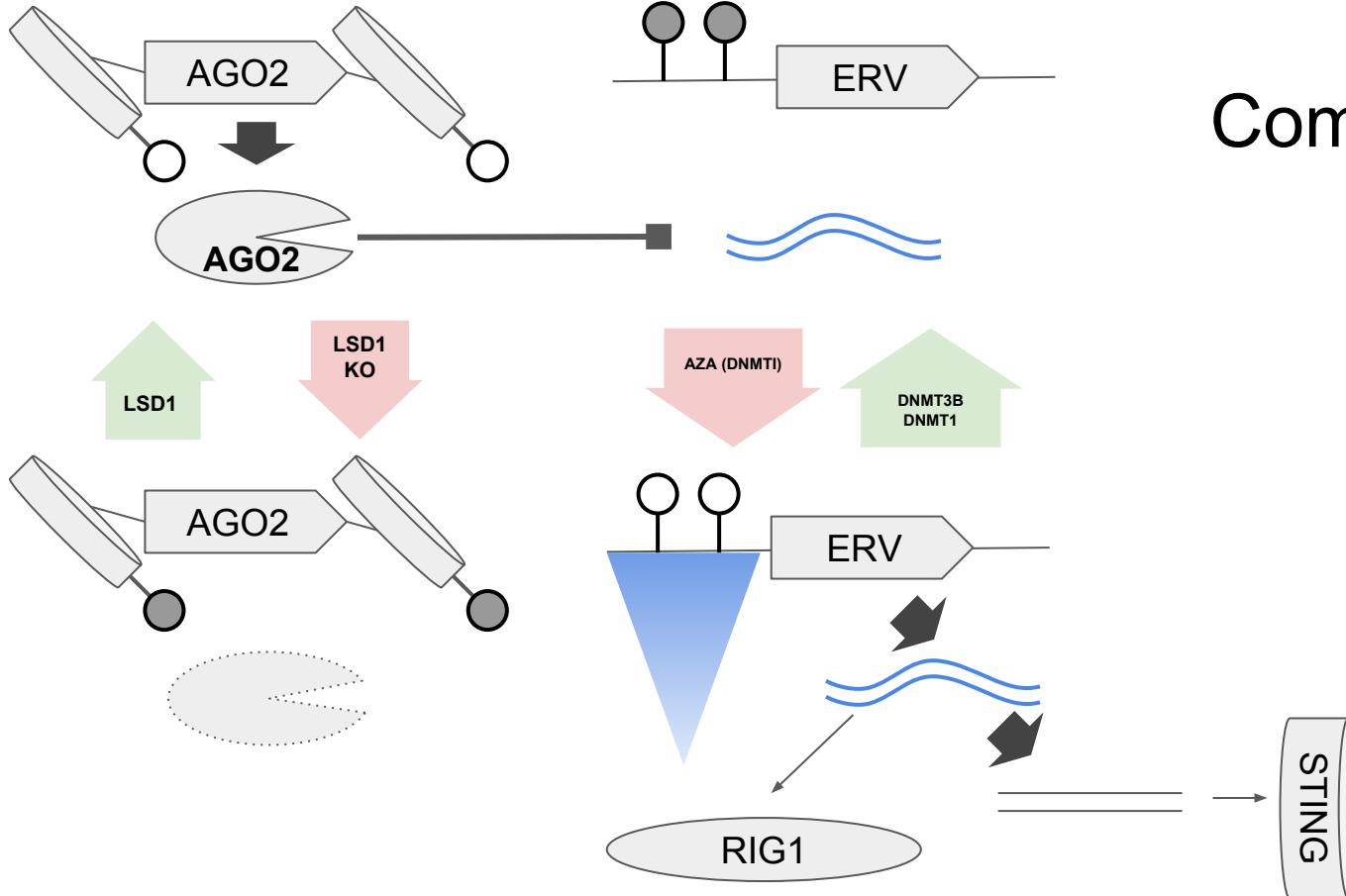
Q#4

Q#6 + Q#7

Common findings



Li et al 2014
Chiappinelli et al 2015



Common findings

Q#1

Q#2 + Q#3

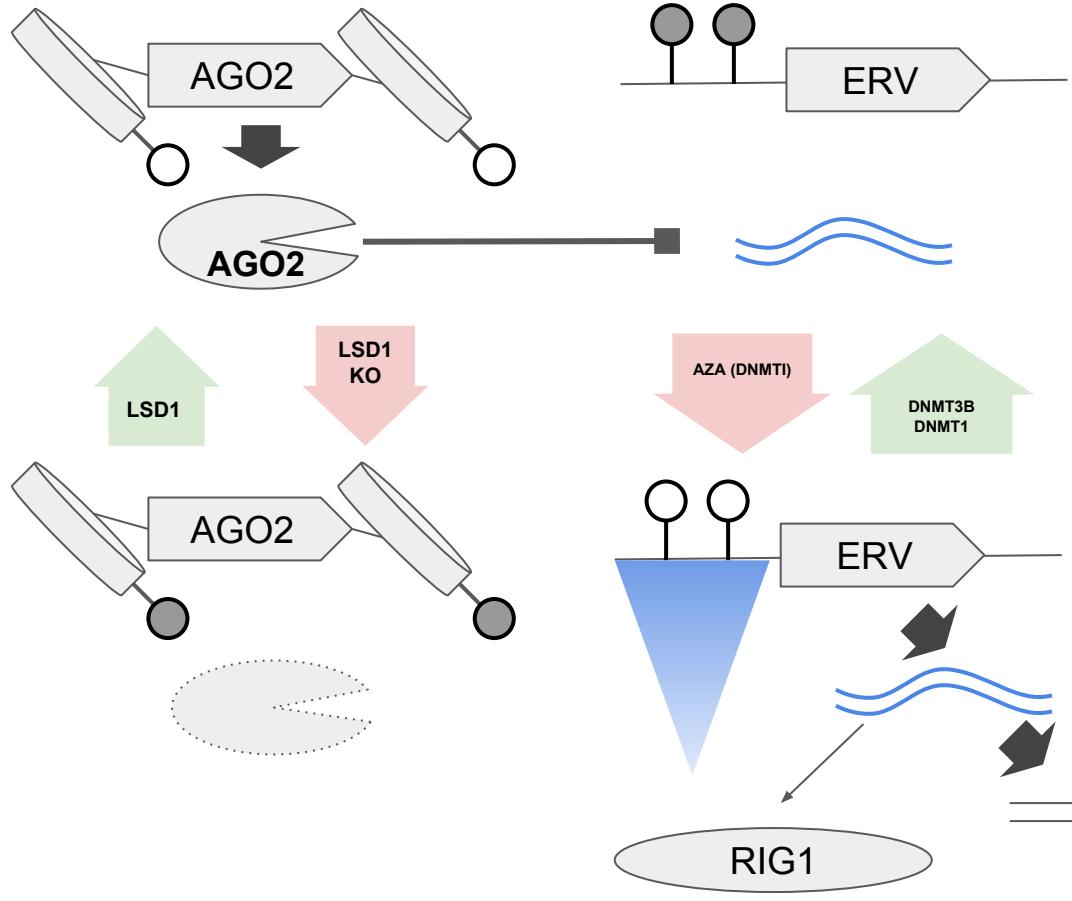
Q#5

Q#4

Q#6 + Q#7

Common findings

- STAT1-driven IFN-beta response genes
- Additional ERVs (SPARCS)
- Inflammatory cytokine secretion (and immune cell recruitment)



Q#1

Q#2 + Q#3

Q#5

Q#4

Q#6 + Q#7

Summary of papers

Chiappinelli et al 2015

- Inhibiting DNA methyltransferases in an EOC line activates ERV expression
- A subset of AIMs (from Li et al 2014) is upregulated as part of the interferon response against ERVs, which segregate tumors by immune responsiveness and correlated with anti-CTLA4

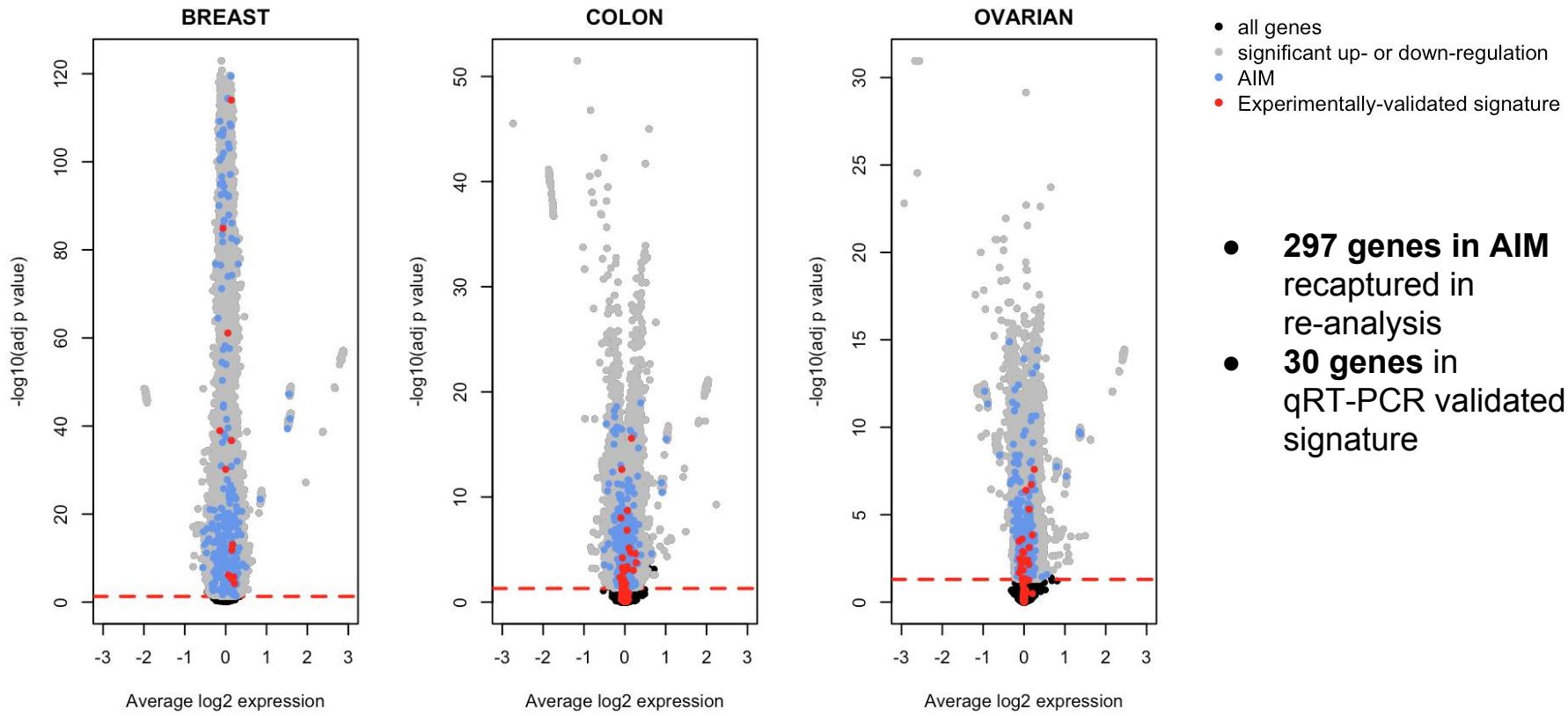
Canadas et al 2018

- A subset of genes interferon response in small lung cancer cell lines contain 3' UTR antisense ERVs that are activated upon MET activation
- The innate immune response feeds-forward, promotes immune cell infiltration into the tumors, and supports anti-PD1 blockade

Sheng and Lafleur et al 2018

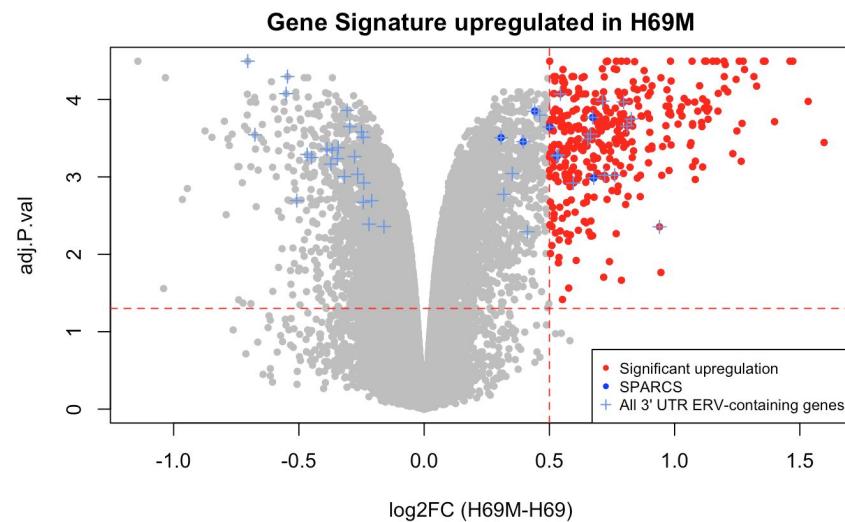
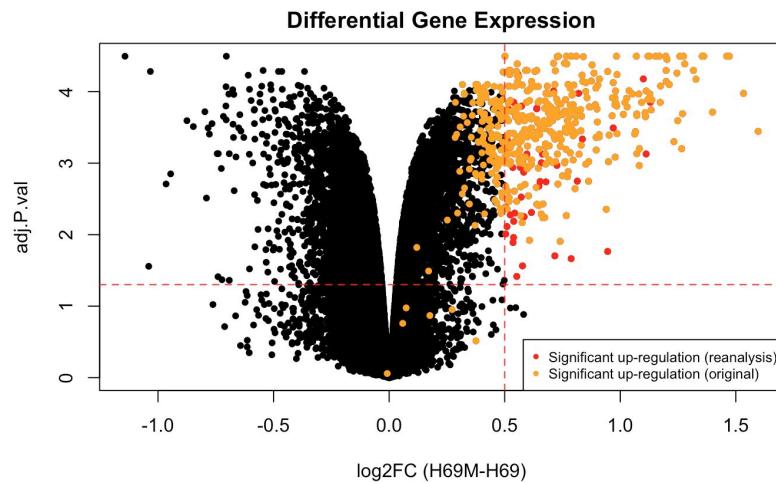
- Loss of LSD1 inhibition prevents activation of RISC components, which target and degrade dsRNA from cytosolic ERVs that are sensed by dsRNA sensors
- Interferon pathway activation combined with anti-PD1 therapy promotes CD8+ T-cell infiltration into tumors

Reanalysis of Chiappinelli et al 2015 and Li et al 2014



- **297 genes in AIM** recaptured in re-analysis
- **30 genes in qRT-PCR validated signature**

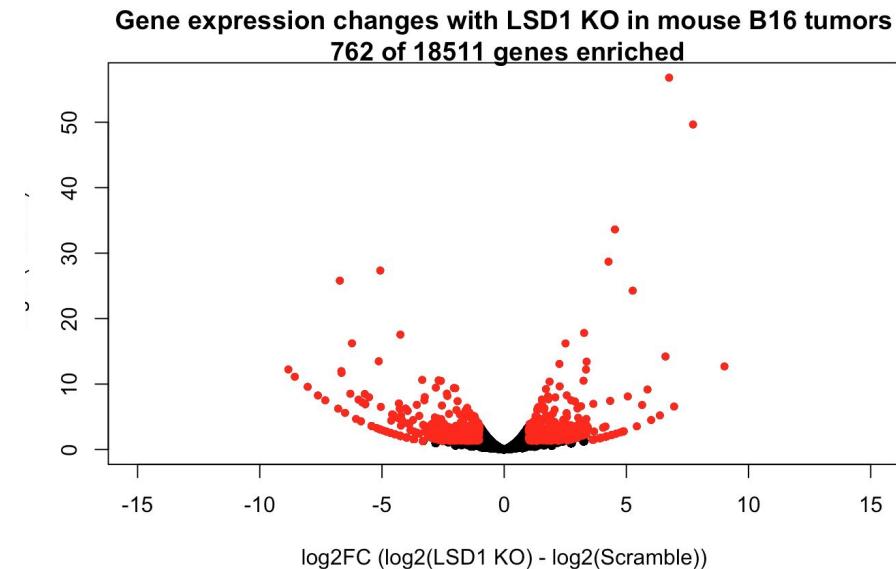
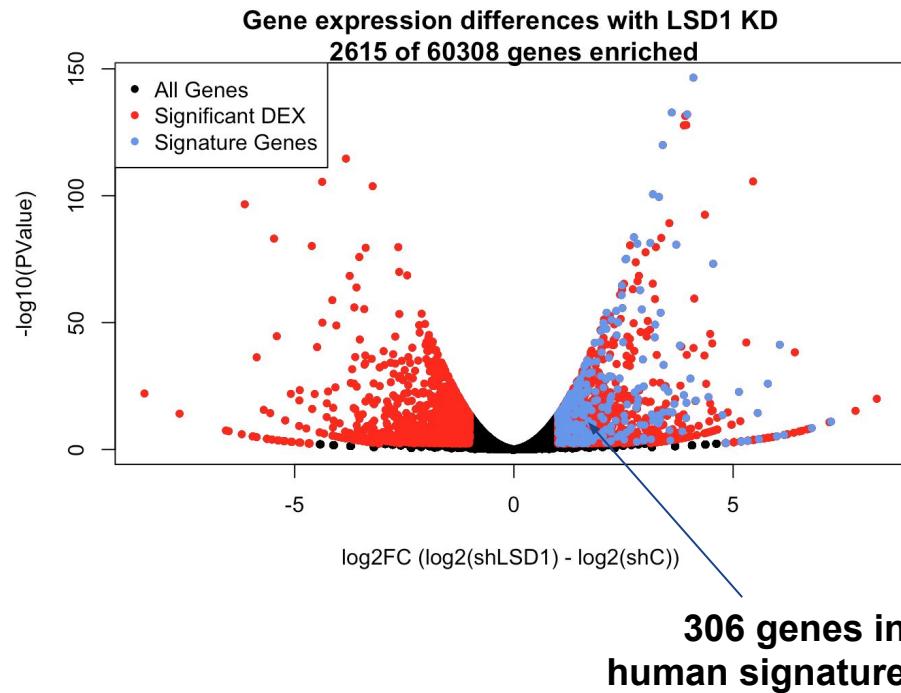
Reanalysis of Canadas et al 2014 and Canadas et al 2018



306 genes in signature

Orange: Canadas et al 2018's significantly-expressed set
Red: Significantly-expressed set from re-analysis
Blue: 3'UTR ERV genes (+ = sense; o = antisense)

Reanalysis of Sheng and LaFleur 2018



Q#1

Q#2 + Q#3

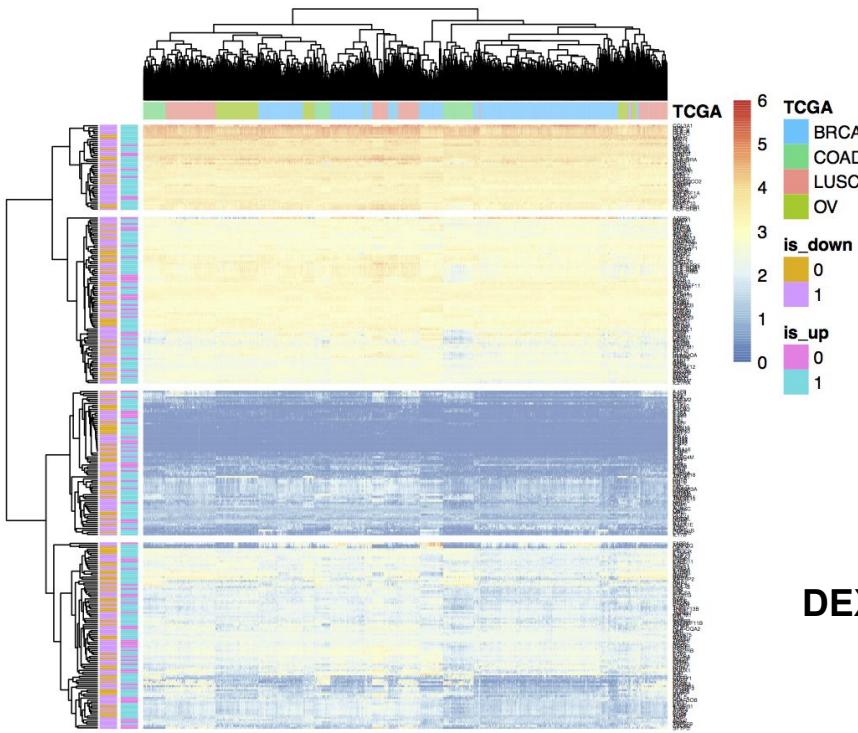
Q#5

Q#4

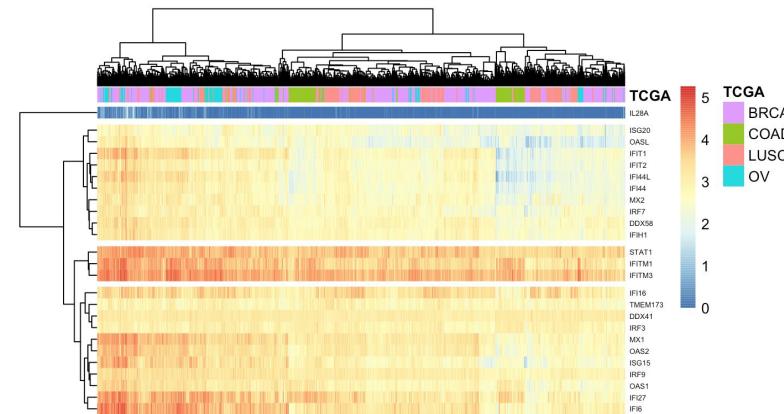
Q#6 + Q#7

Chiappinelli DEX (AIM) and signature genes in TCGA

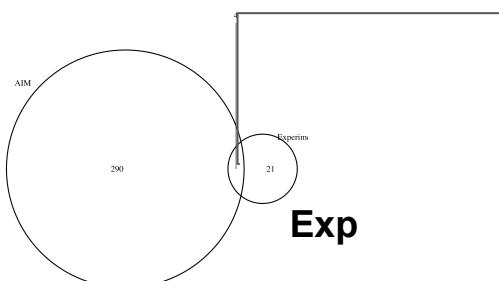
DEX genes from Li et al 2014 and Chiappinelli et al 2015
in 2399 TCGA samples



Experimentally-validated signature from Chiappinelli et al 2015
in 2399 TCGA samples



DEX



**IL28A
IRF7
IRF9
STAT1**

Q#1

Q#2 + Q#3

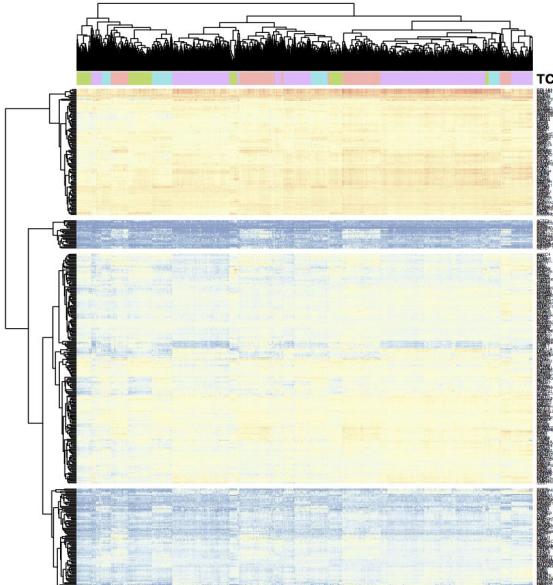
Q#5

Q#4

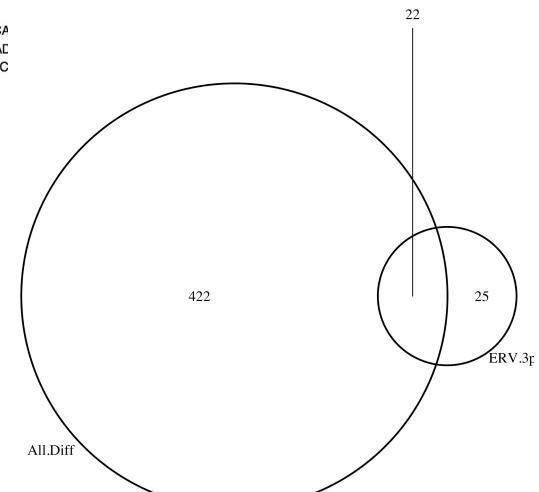
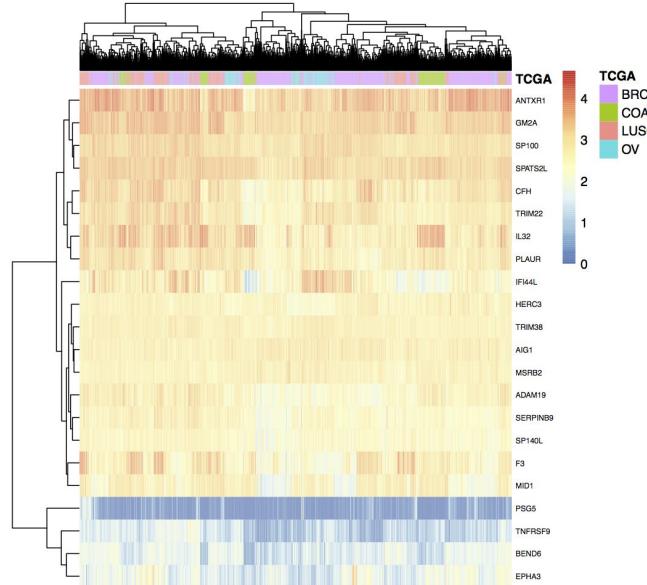
Q#6 + Q#7

Canadas DEX and SPARCS genes in TCGA

DEX genes from Canadas et al 2014 and Canadas et al 2018
in 2399 TCGA samples



3' UTR ERV-containing DEX genes from Canadas et al 2014 and
Canadas et al 2018 in 2399 TCGA samples



Q#1

Q#2 + Q#3

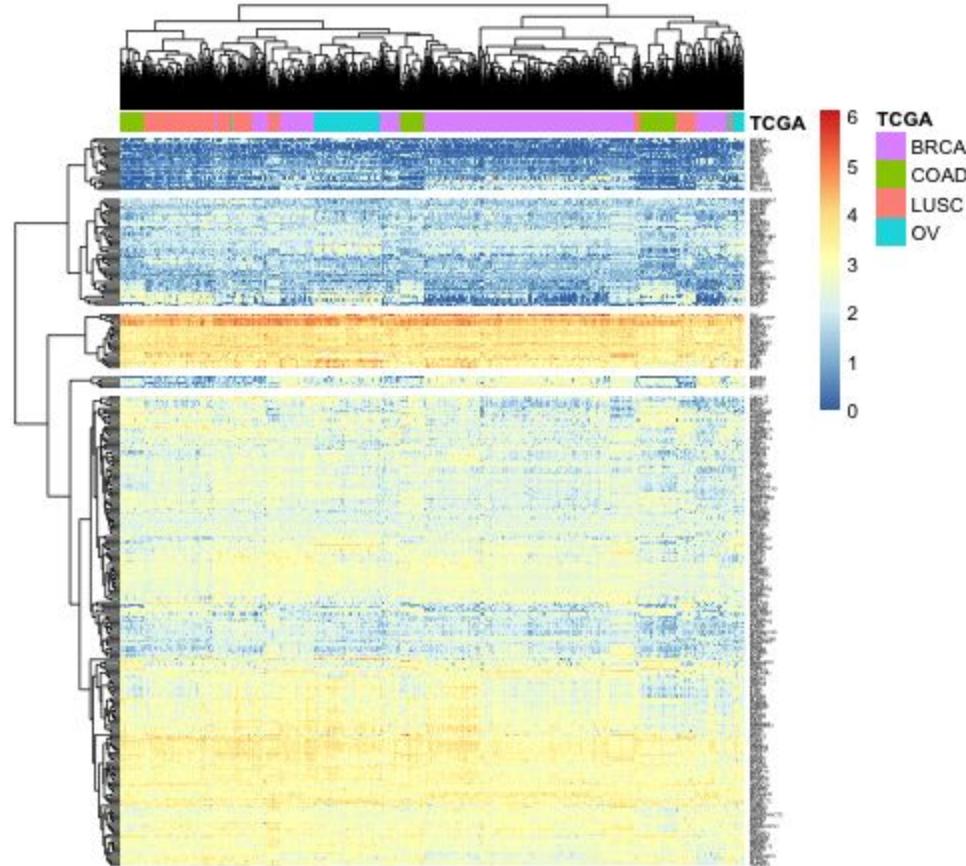
Q#5

Q#4

Q#6 + Q#7

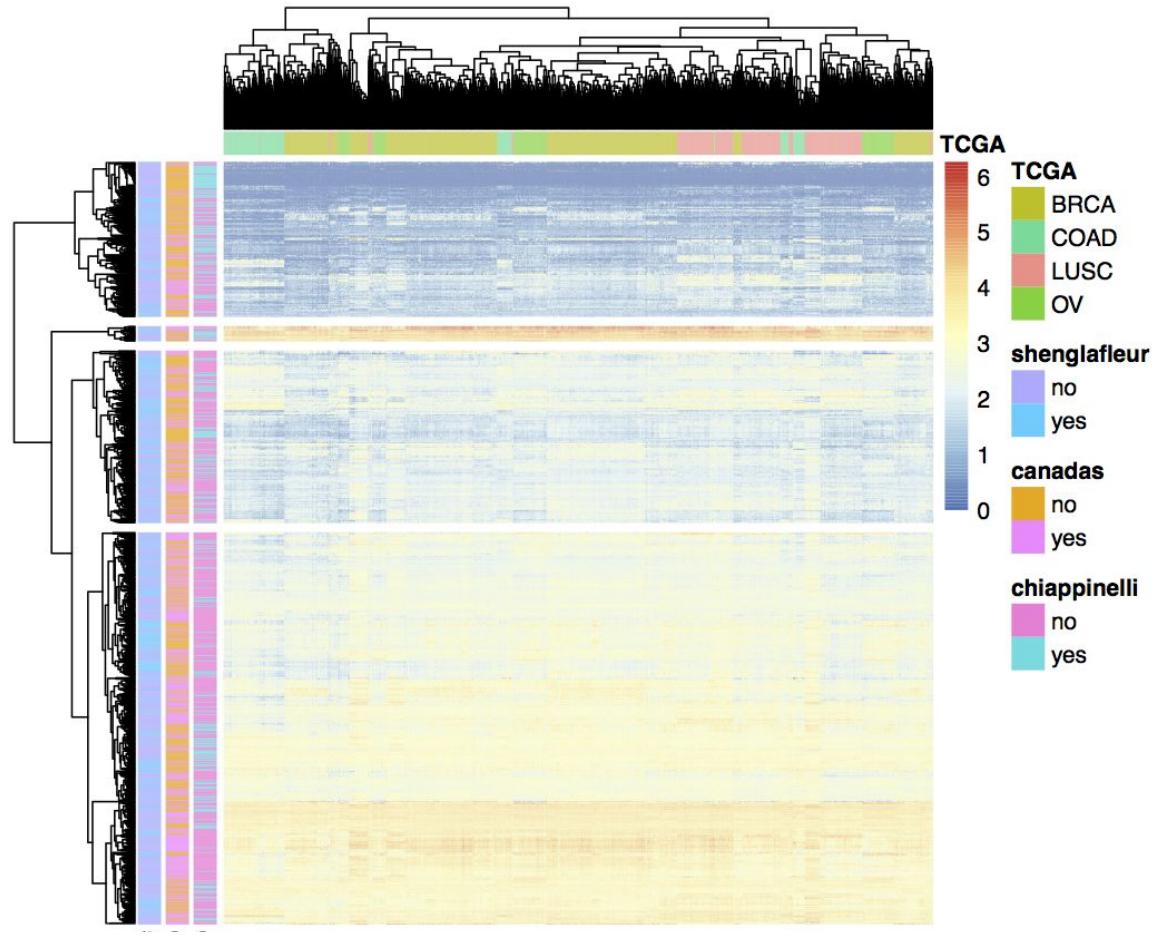
Sheng and LaFleur (human) signature in TCGA

DEX genes from Sheng and LaFleur et al 2018
in 2399 TCGA samples



All DEX genes found in at least one dataset
in 2399 TCGA samples

Unified



Q#1

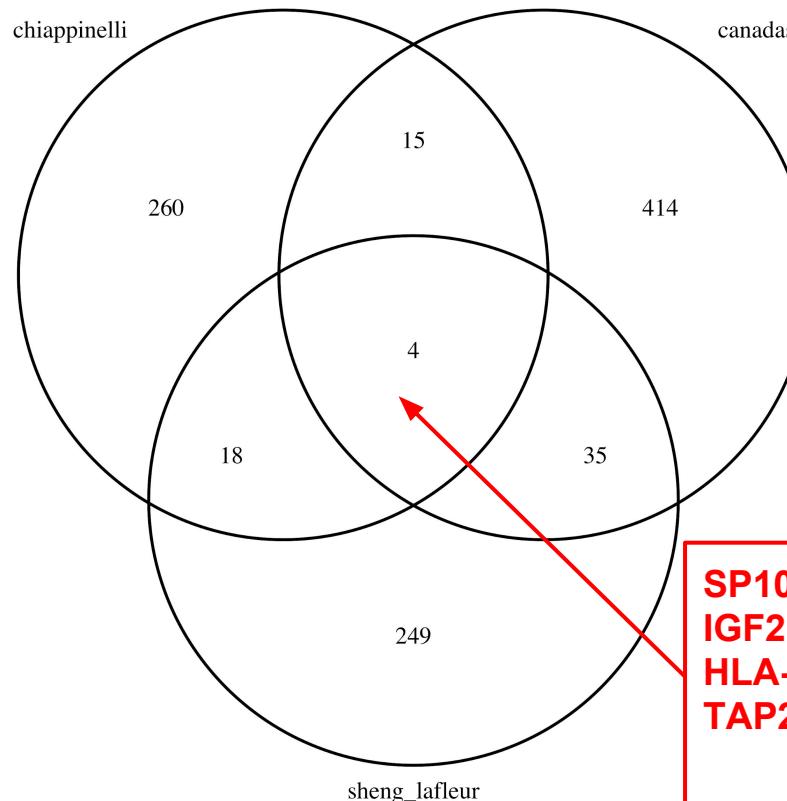
Q#2 + Q#3

Q#5

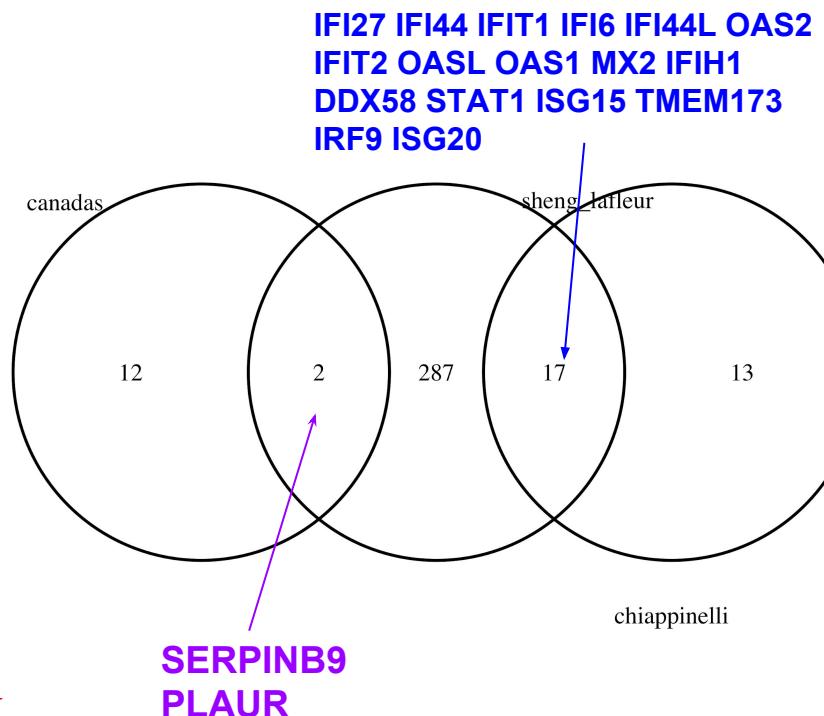
Q#4

Q#6 + Q#7

Intersection between all DEX genes



Intersection between paper-selected signatures



Q#1

Q#2 + Q#3

Q#5

Q#4

Q#6 + Q#7

Microarrays and RNA-seq have different measurement distributions

Two-channel microarrays present **ratio of fluorescent emissions for a particular probe**, which are normally distributed → *continuous distributions*

- a. Chiappinelli et al 2015
- b. Li et al 2014
- c. Canadas et al 2014
- d. Canadas et al 2018

RNA-seq provides **counts (integer or real-number) of reads for detected transcripts** → can be treated as a *discrete distribution*

- a. Sheng and LaFleur et al 2018

When using a standard gene set (Type-I interferon signature from MSigDB), how would the number of intersections vary if the experimentally-derived set were from a microarray vs from a sequencing experiment

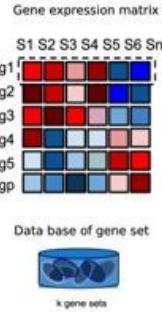
Approaches

Hänzelmann et al BMC Bioinf 2013

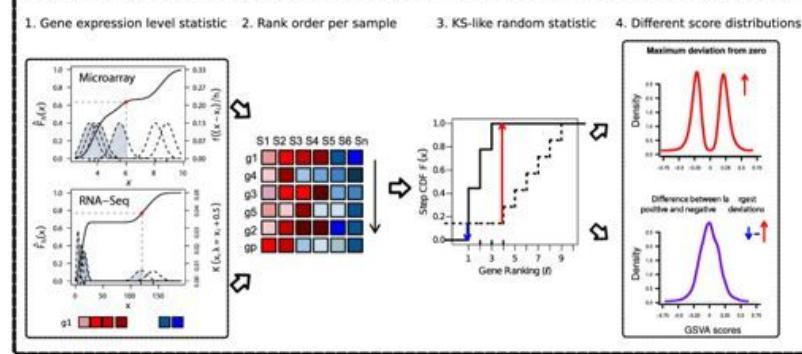
GSVA

- derive a KDE as an estimate for the underlying gene expression distribution
- Non-parametric test of enrichment of highly-expressed genes across gene sets
- Alternative to ssgSEA, gProfileR, and standard GSEA

Input



GSVA Algorithm



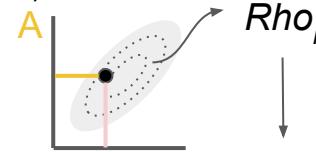
Output



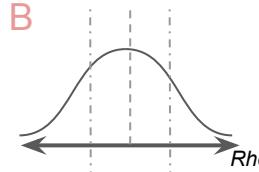
Bootstrap estimate of gene signature correlation

- Empirically estimate how correlated are the median expression values of genes in a signature for a specific sample

$\text{median}(A1, A2, A3, A4 \dots)$



$\text{median}(B1, B2, B3, B4 \dots)$

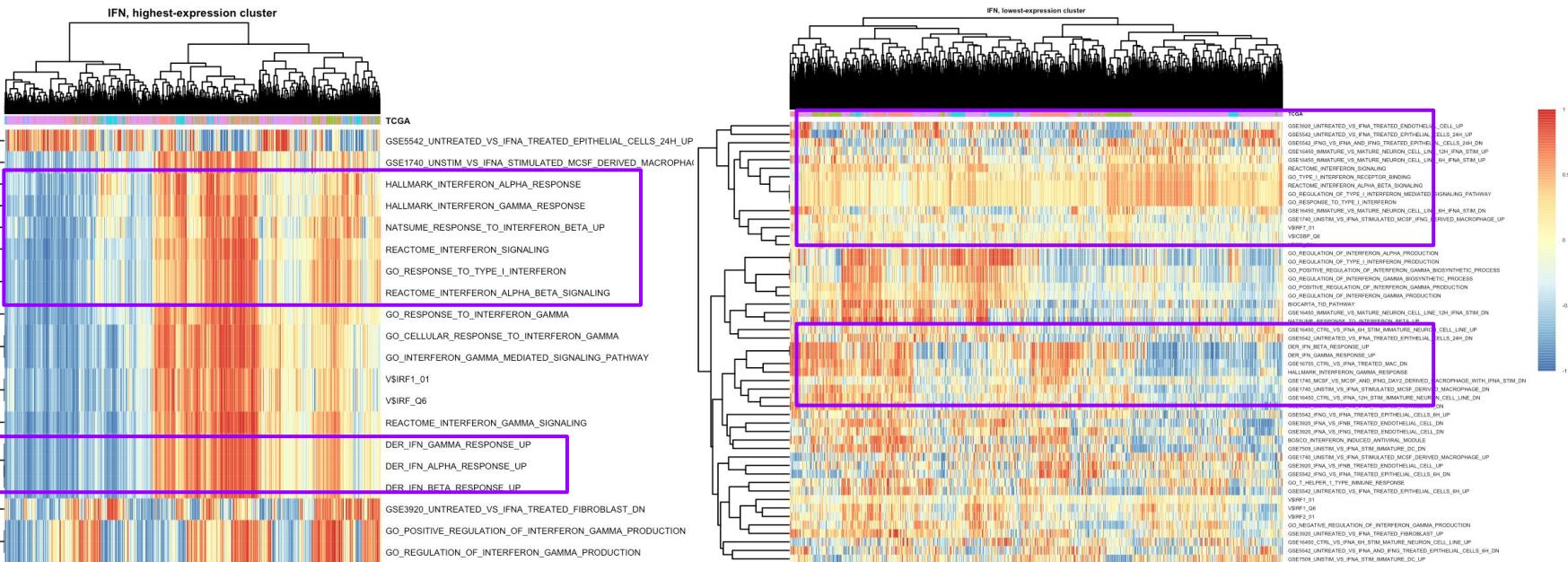


95% CI for correlation
between A and B

Rho_i

Rho

Enrichment analysis for IFN signatures in TCGA by GSVA



Highest-expressed TCGA cluster 20/120 gene sets enriched

Sparsest expressed TCGA cluster 52/120 gene sets enriched

Q#1

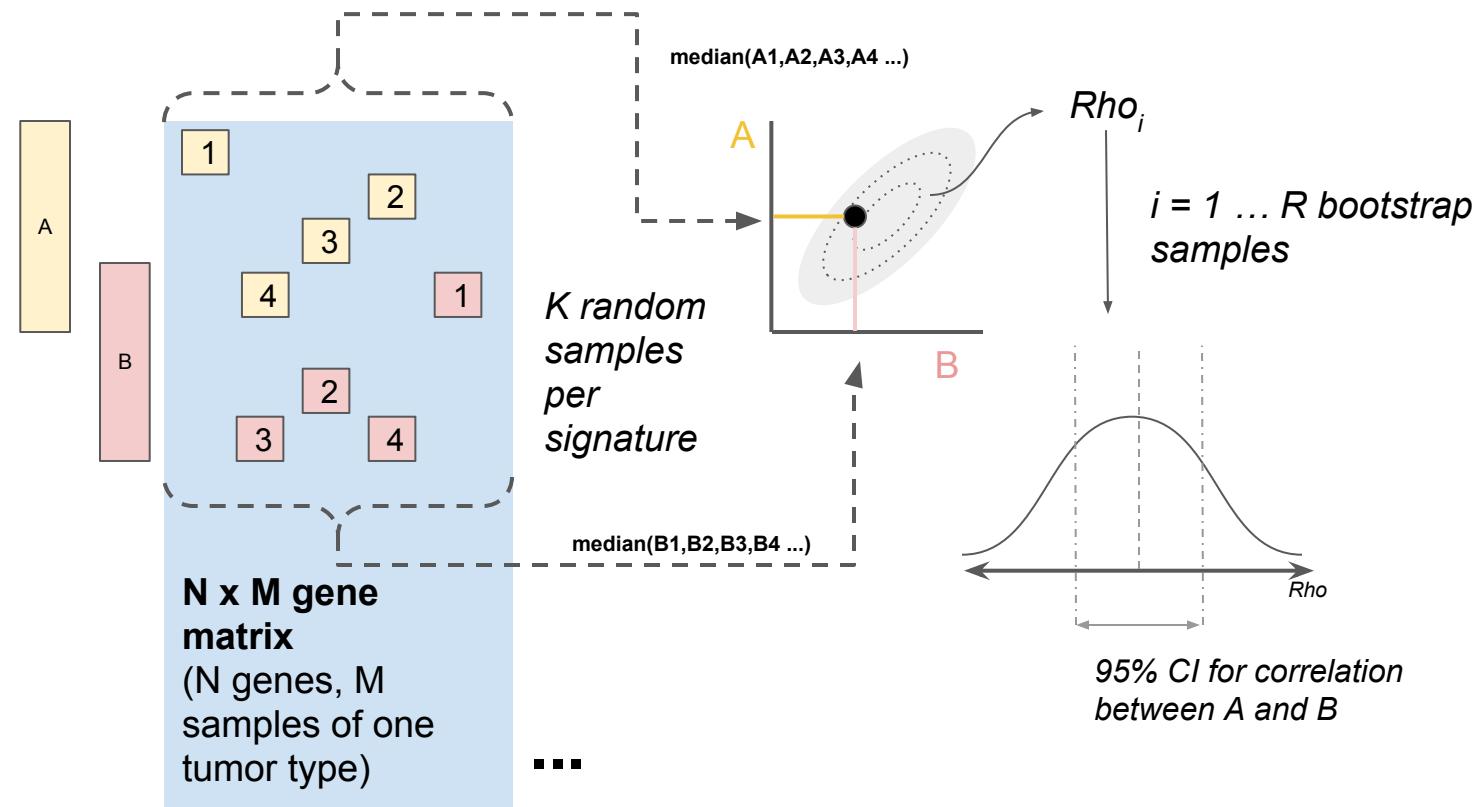
Q#2 + Q#3

Q#5

Q#4

Q#6 + Q#7

Estimating correlation in signature expression



Q#1

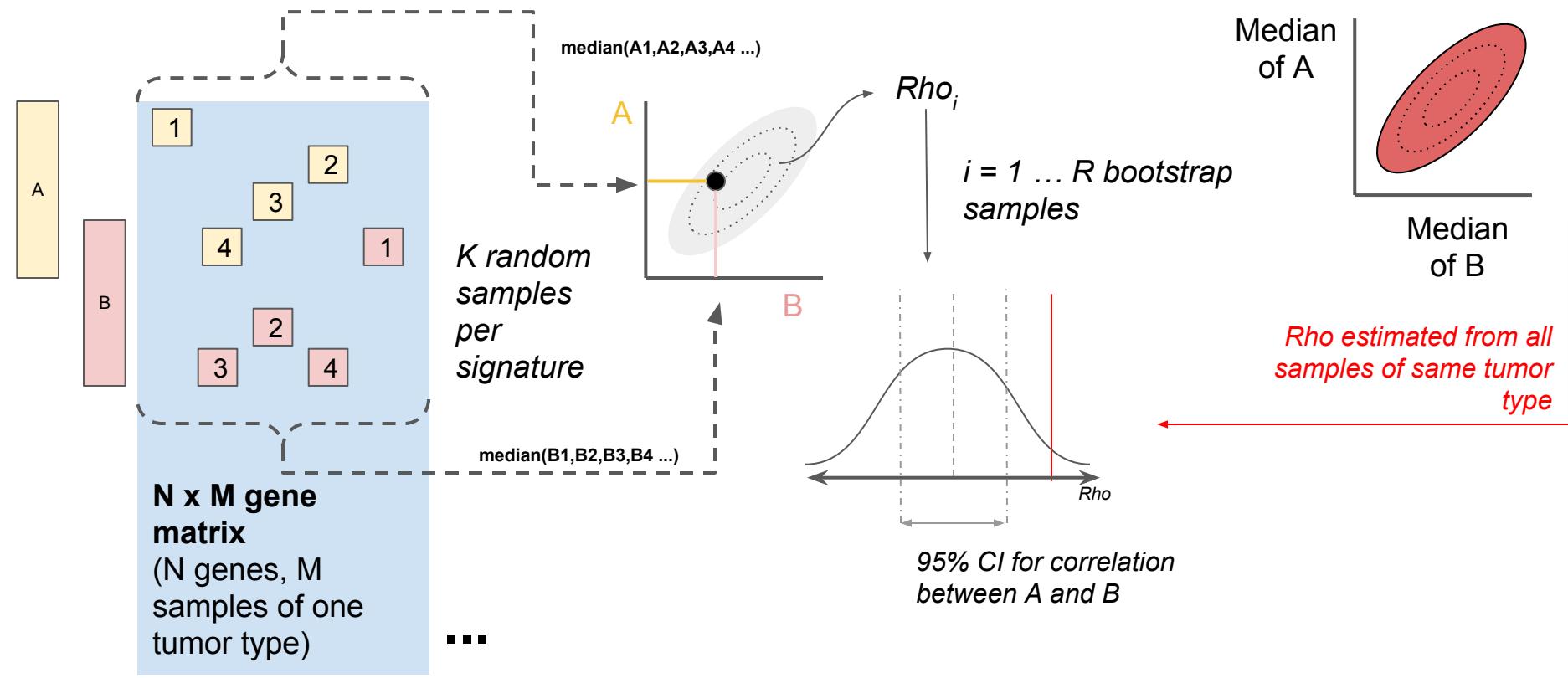
Q#2 + Q#3

Q#5

Q#4

Q#6 + Q#7

Estimating correlation in signature expression



Q#1

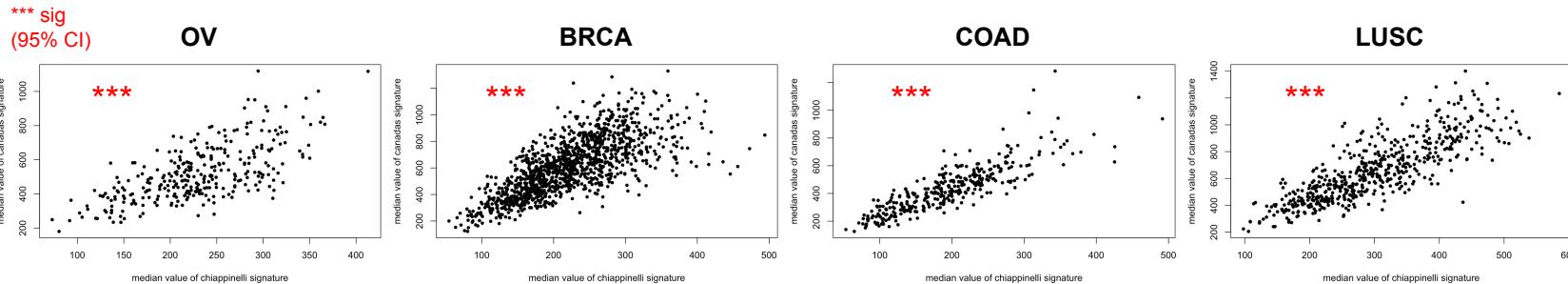
Q#2 + Q#3

Q#5

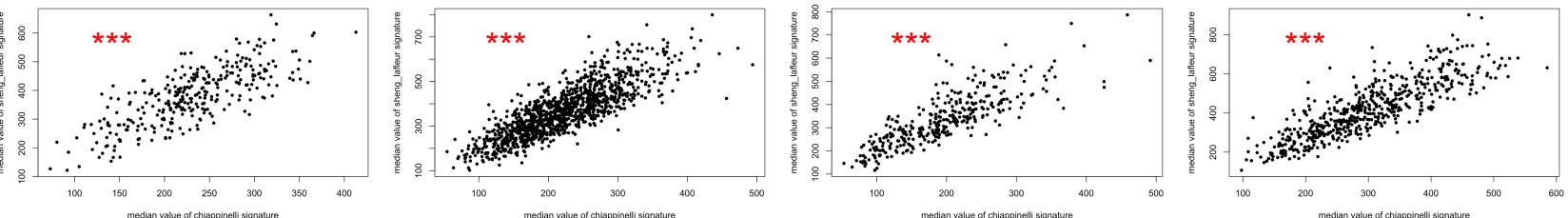
Q#4

Q#6 + Q#7

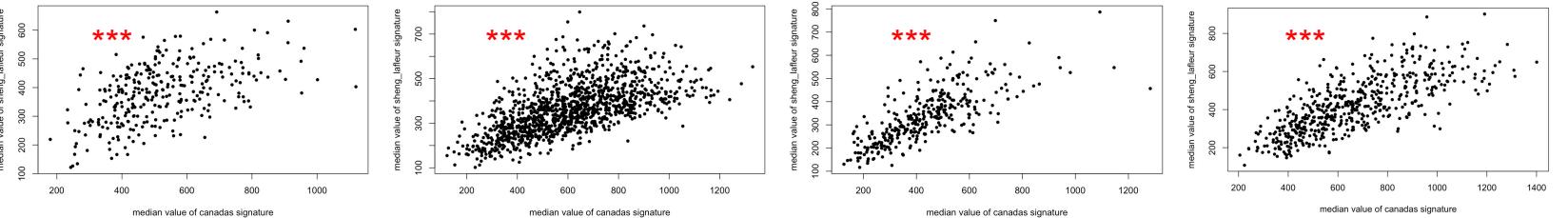
Estimating correlation in median log10 signature expression in TCGA



X: Chiap
Y: Can



X: Chiap
Y: SF



X: Can
Y: SF

Q#1

Q#2 + Q#3

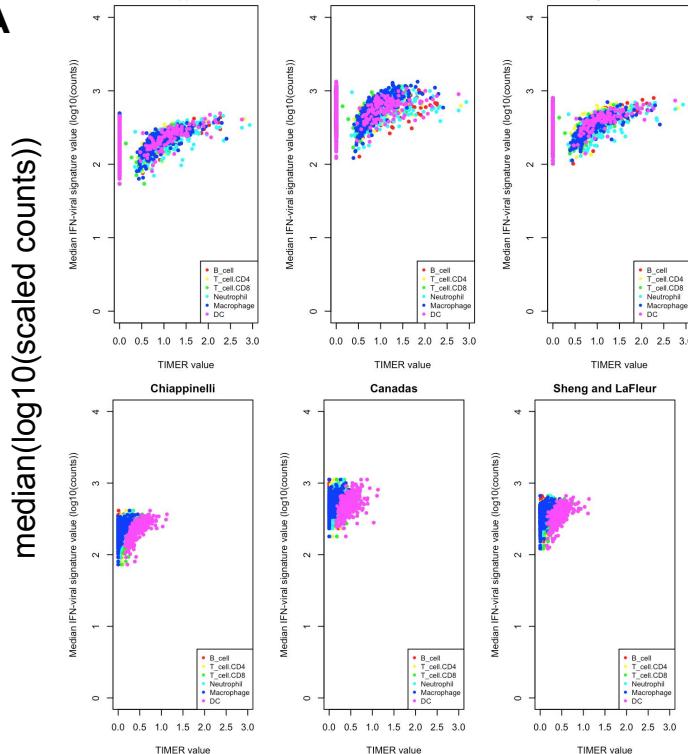
Q#5

Q#4

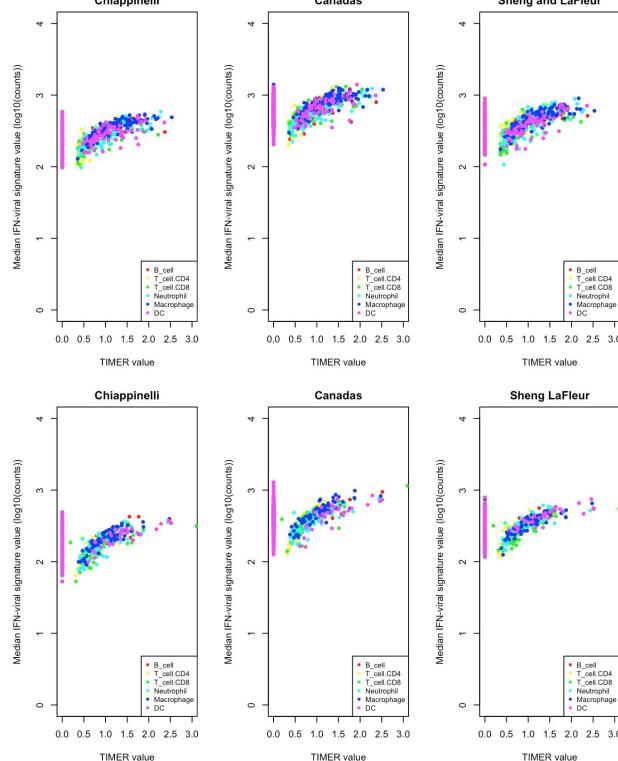
Q#6 + Q#7

Correlations of TIMER with IFN-viral signatures

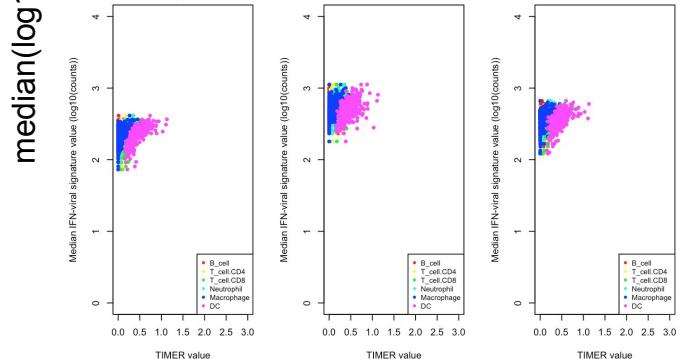
BRCA



LUSC

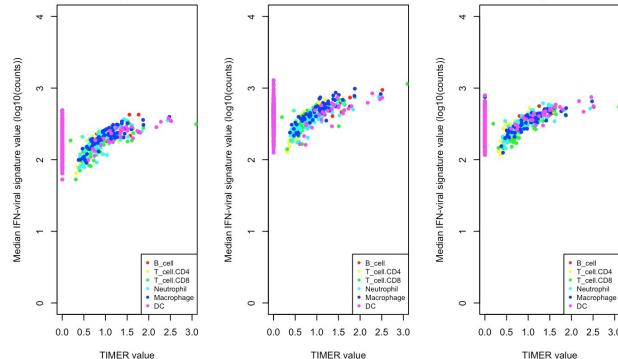


OV



- B_cell
- T_cell.CD4
- T_cell.CD8
- Neutrophil
- Macrophage
- DC

COAD



TIMER estimated abundance

Q#1

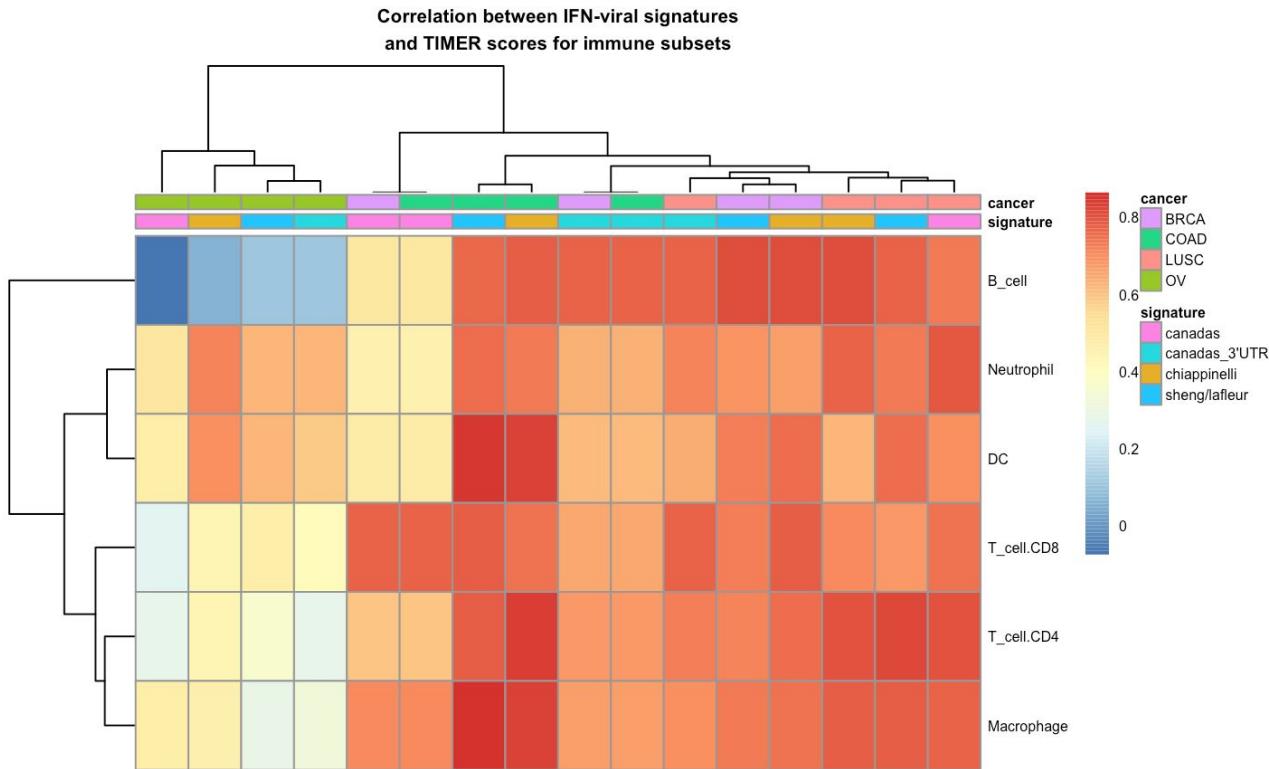
Q#2 + Q#3

Q#5

Q#4

Q#6 + Q#7

Correlations of TIMER with IFN-viral signatures



Q#1

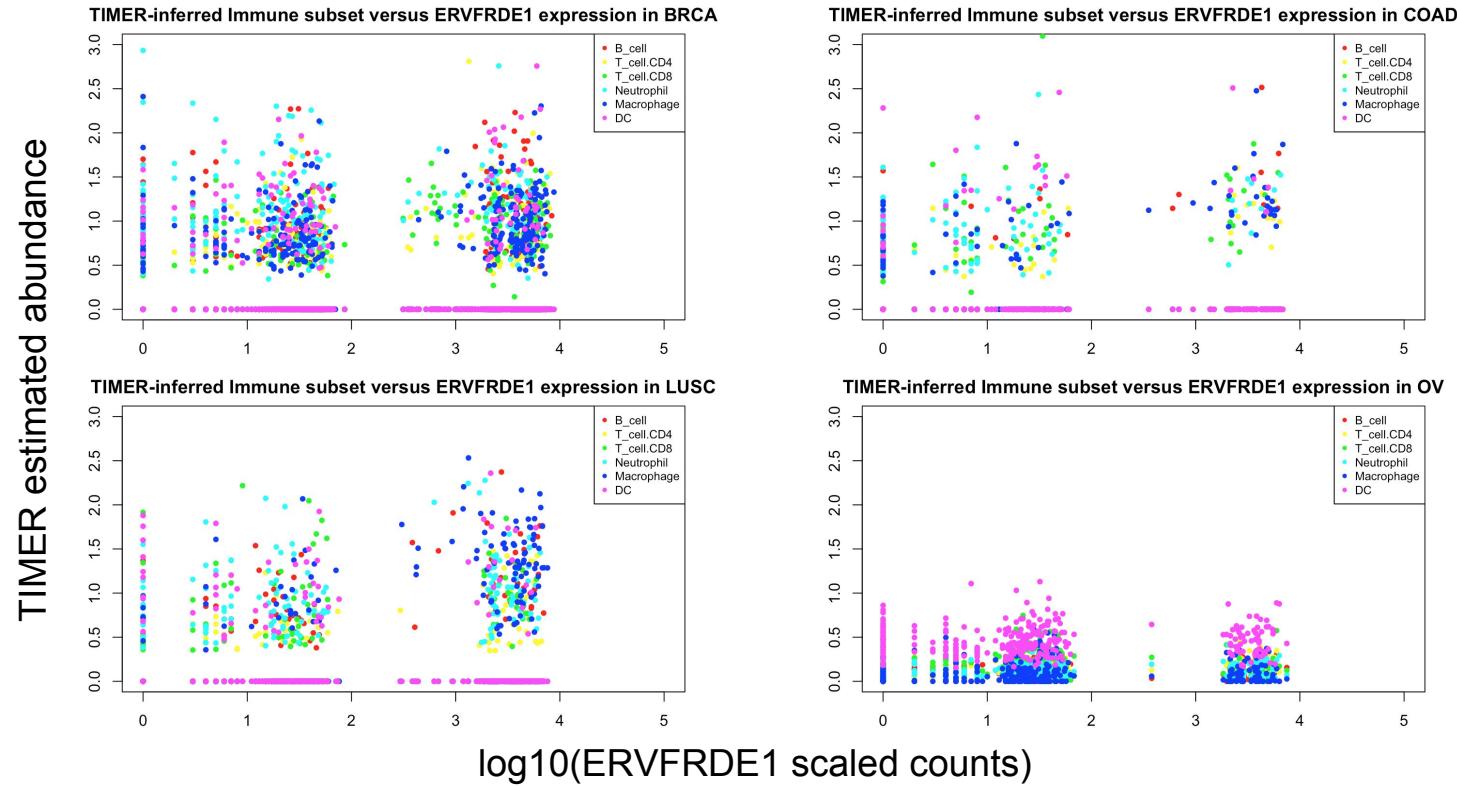
Q#2 + Q#3

Q#5

Q#4

Q#6 + Q#7

Correlation of ERVFRDE1 and IFN-viral signatures



Q#1

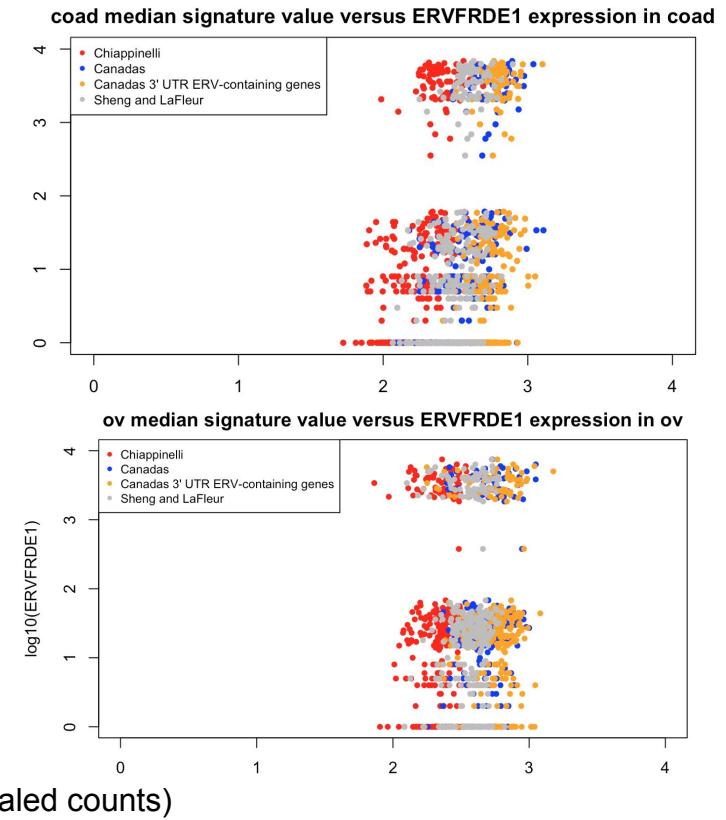
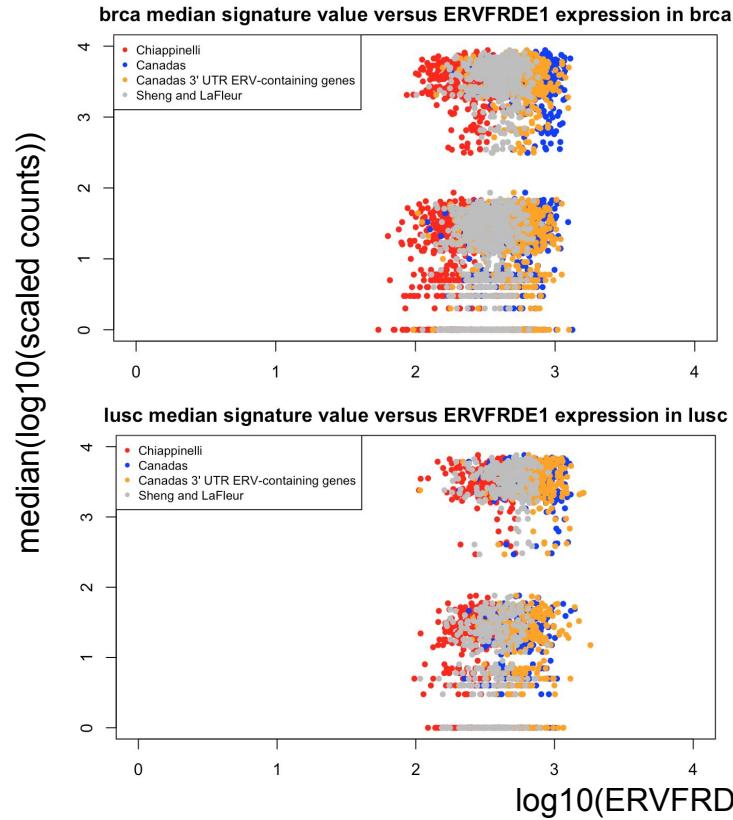
Q#2 + Q#3

Q#5

Q#4

Q#6 + Q#7

ERVFRDE1 and expression of different signatures



External studies

- TCGA-wide analyses of ERV transcripts (using controlled-access data) show that ERV expression is well-correlated with positive outcomes for patients with urothelial cancer (Solovyok et al Cell Reports 2018) and modestly so with CTLA-4 and PD-1 expression levels within tumors

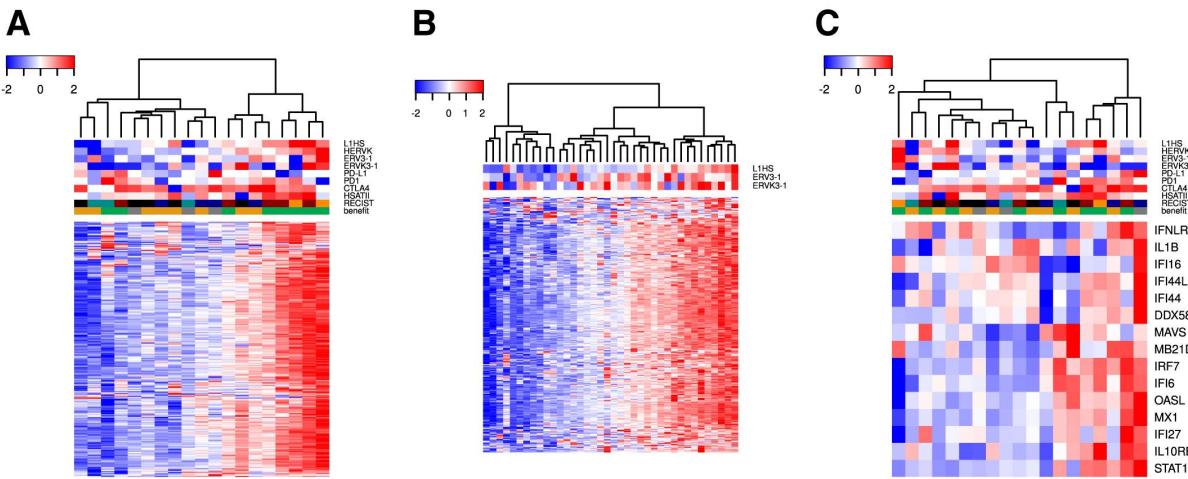


Figure 3 from Solovyok et al.
(a) ERV clustering for
urothelial cancer patient cohort
from Snyder et al PLoS 2017
(b) TCGA-wide clustering of
ERVs

Q#1

Q#2 + Q#3

Q#5

Q#4

Q#6 + Q#7

Current methods to directly identify ERV expression

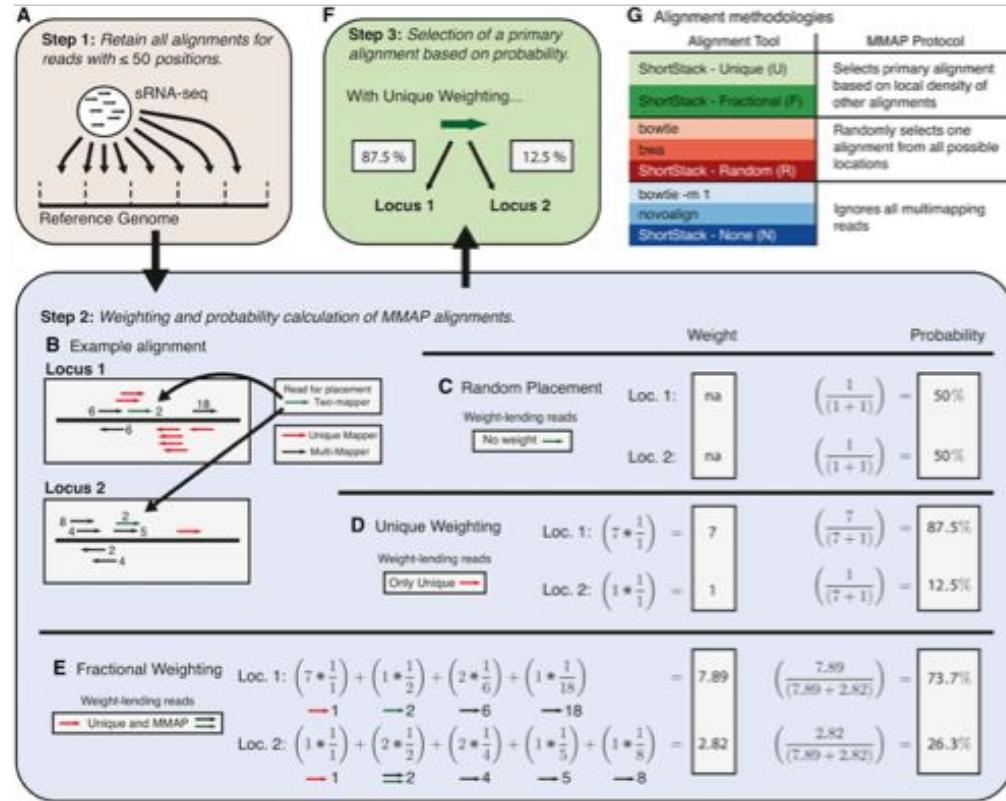
1. **Direct quantification of ERVs:** map raw RNA-seq reads to RepeatMasker annotations and normalize as in typical RNA-seq (*current method*)
 - a. Rooney et al Cell 2015
 - b. Sheng and LaFleur et al Cell 2018
2. **Map reads around boundaries:** identify reads that map across the boundaries of ERVs and the surrounding gene
 - a. Day et al Genome Biology 2013
3. **Specialized technologies** for capturing and selective amplification

Problems with ERV identification

1. Repetitive elements often “soak” up a large number of reads and are often excluded from conventional RNA-seq analysis (eg: TCGA). **Extreme right-tail of the count distribution**
2. Multiple copies of repetitive elements **can drive high rates of multiple mapping (50-80% MMAP driven by rRNA in RepeatMasker track from which ERV estimates are derived)** that confounds ERV estimation

Proposed Approach for multiple mapping: Adopt ShortStack

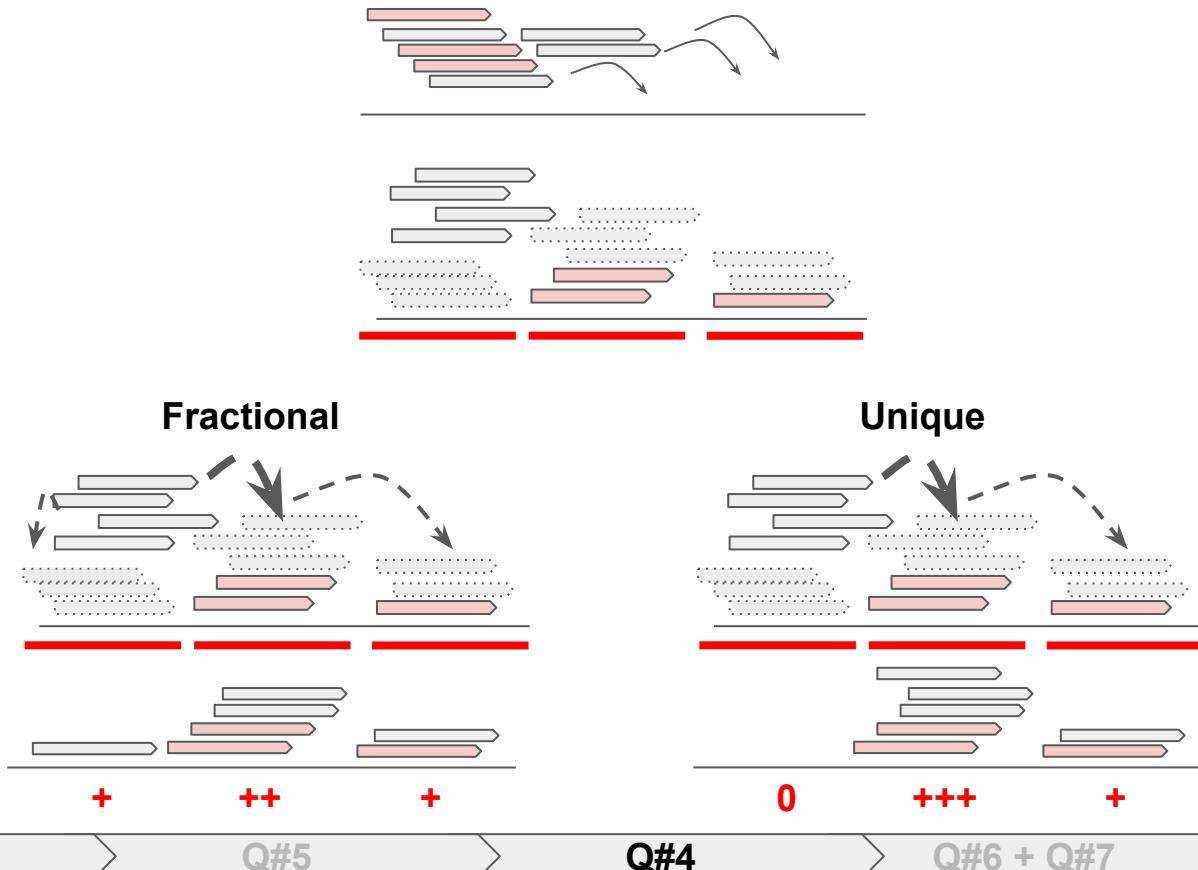
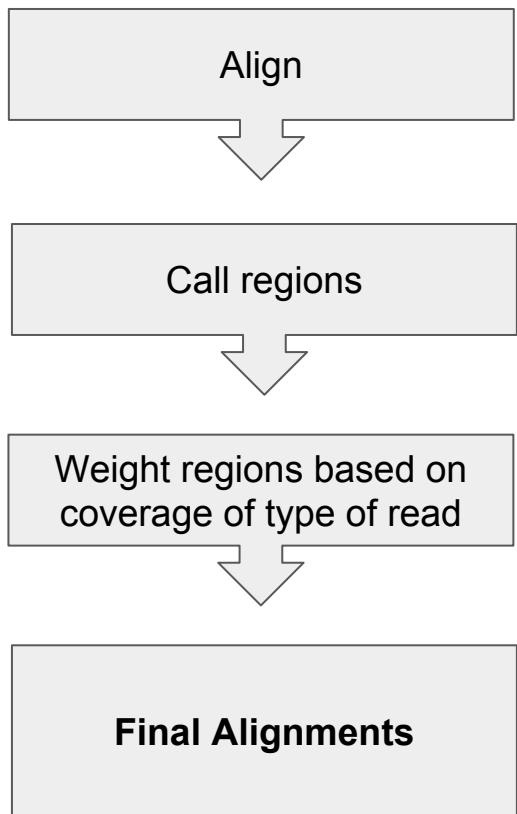
Use the local coverage of unique-mapping reads at a site to weight the site's contribution to the read



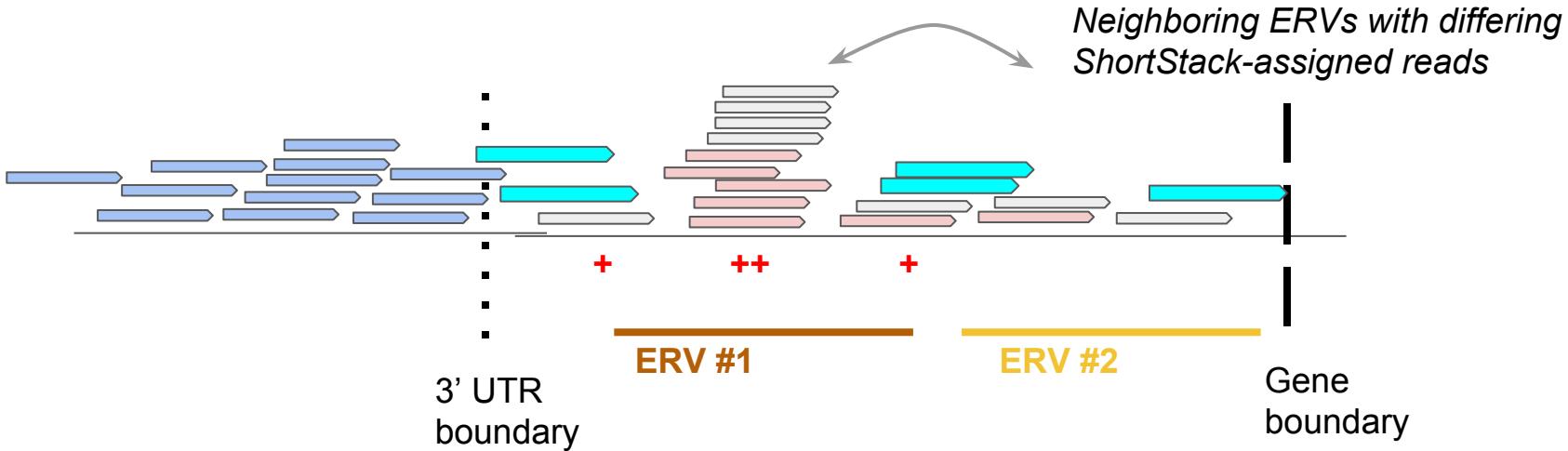
Johnson et al G3 2016

A schematic depiction

→ = Uniquely-mapping reads
→ = multiple-mapping (MMAP) reads
→ = possible alignment for MMAP read



UTR-ERV junction reads inform relative expression



→ = RNA-seq
read in a
non-ERV
region

→ = ERV-UTR
junction
reads

- Junction reads allow for estimates of relative expression of UTR vs ERV
- Sense and antisense biases can be obtained by considering different-stranded transcripts separately
- Neighboring ERVs in UTRs can provide estimates on ERV copy number (based on common UTR abundance)

Q#1

Q#2 + Q#3

Q#5

Q#4

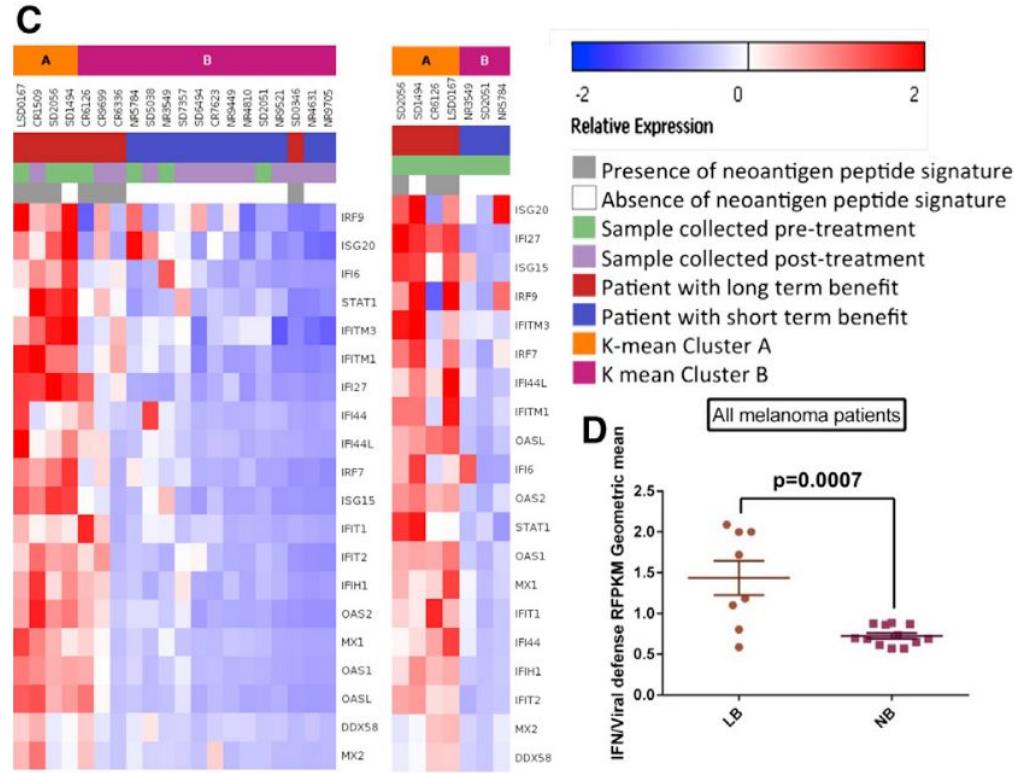
Q#6 + Q#7

Summary of Approach

- Adopt **ShortStack** to select one of the many possible alignments of each multiple-mapping (MMAP read) at ERVS.
- Use **ERV-UTR junction reads** to observe and correct for large-scale discrepancies in ERV-UTR expression
- Return adjusted ERV counts and estimates of ERV copy number based on **neighboring or same-UTR ERVs**

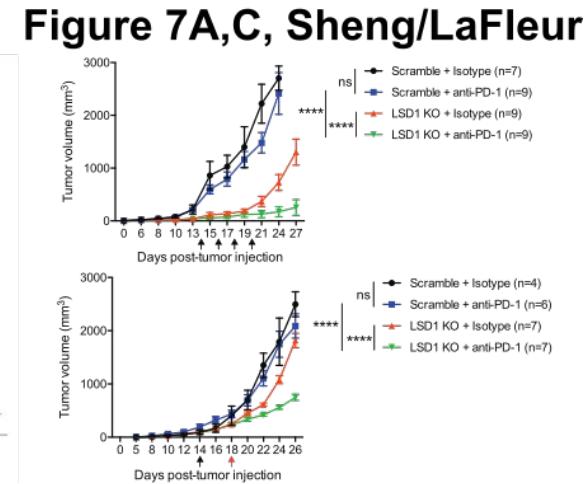
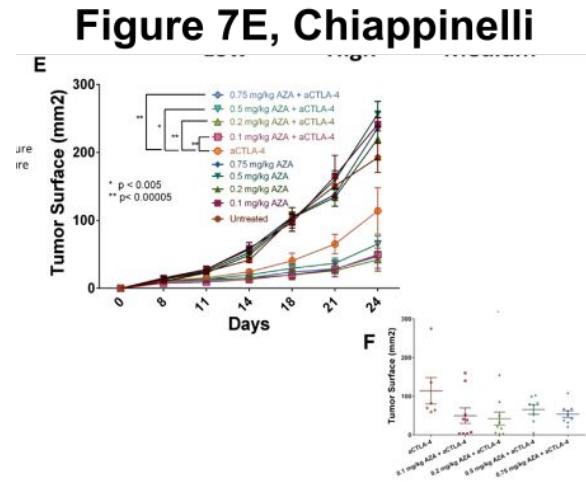
ERV and immunotherapy: human studies

- Patients who show long-term benefit from anti-CTLA4 therapy (Snyder et al 2014) tend to have higher expression of (a subset of) the IFN-ERV signature.
- However, **TP rate of signature diagnostic is ~83% and FP rate is ~17%**



ERV and immunotherapy: Mouse studies

- In mice, epigenetic enzyme inhibition treatment **strengthens the effect of checkpoint blockade therapy**
- ERV activation and inflammation without an immune response can **aggravate tumors and worsen mouse outcomes**
- Epigenetic inhibitors **heightens tumor inflammation.** Checkpoint blockade **expands the capacity of the immune response.**



Clinical applications of ERV and immunotherapy

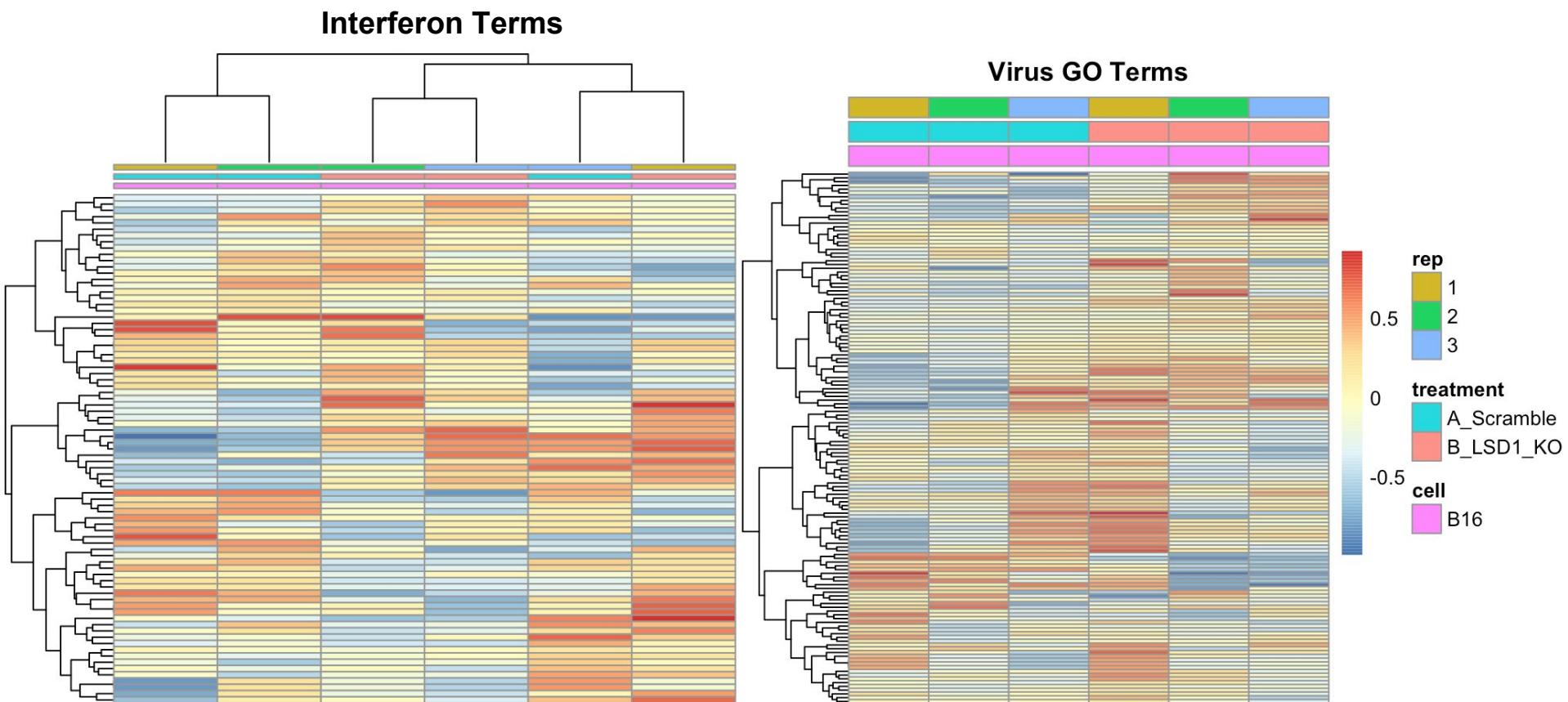
- **Drug delivery and localization** of epigenetic enzyme inhibitors/modulators is critical
- **Tissue-specific subsets of immune cells (CD8+ T-cells, DCs, and macrophages in particular)** can suppress or amplify innate immune activation
- ERV activation is also implicated in **autoimmune disorders** (eg: SLE/lupus)

Summary

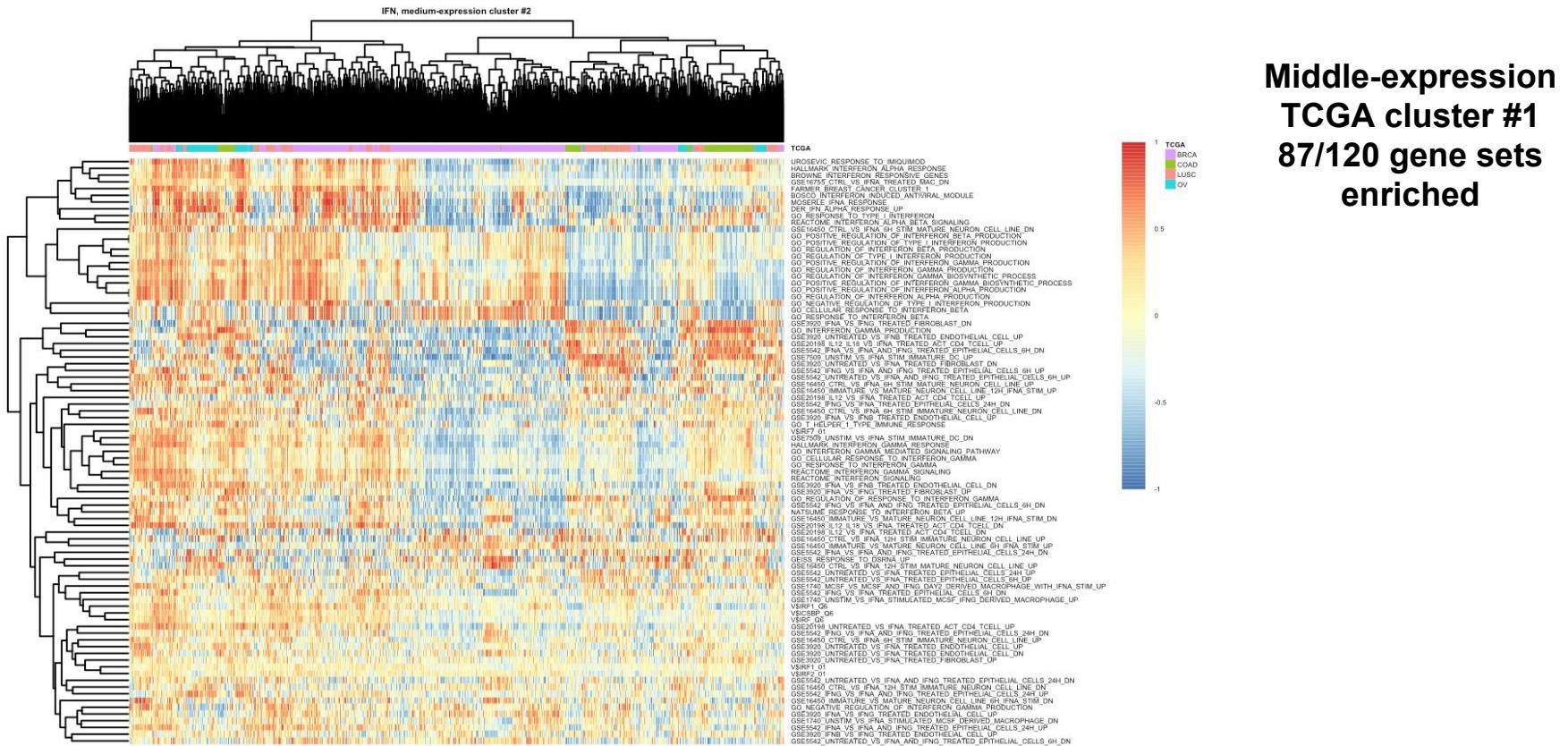
- Different signatures between the three papers may reflect different experimental conditions, although they provide correlated insight into tumor inflammation
- ERV-response pathways and inflammatory gene activation in tumors is structured, with constitutively-active parts and more “sparse” and “variegated” signatures
- ERV-response signatures still require “unification” across platforms and samples to develop a high-TP-low-FP diagnostic

SUPPLEMENTAL SLIDES

Supplemental GSVA, Sheng/LaFleur



Supplemental: Enrichment analysis for IFN signatures in TCGA



Enrichment analysis for virus-related signatures in TCGA