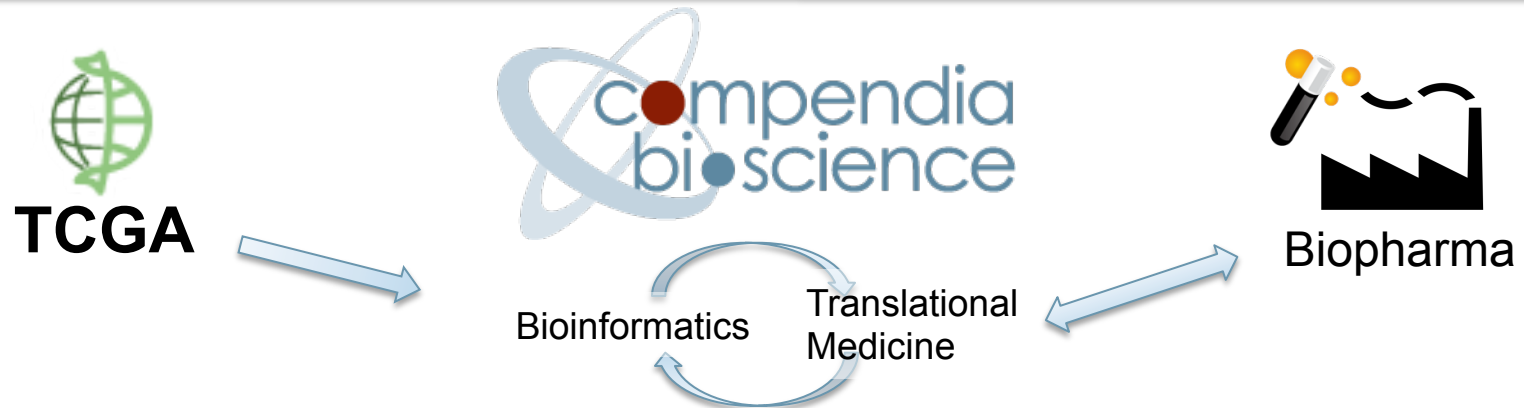


2nd Annual TCGA Symposium - November 28, 2012

Nickolay Khazanov, PhD

Analysis of 3,000 Cancer Exomes to Identify Novel Cancer Drivers and Therapeutic Opportunities

Mission: Cure Cancer With Genomic Data



» **GOALS:**

- » Novel therapeutic targets
- » Right patient population

» **APPROACH:** define comprehensive catalog of driver aberrations in cancer

- » TCGA is critical for this effort!

» **FOCUS: Mutations and Fusions**

Challenges

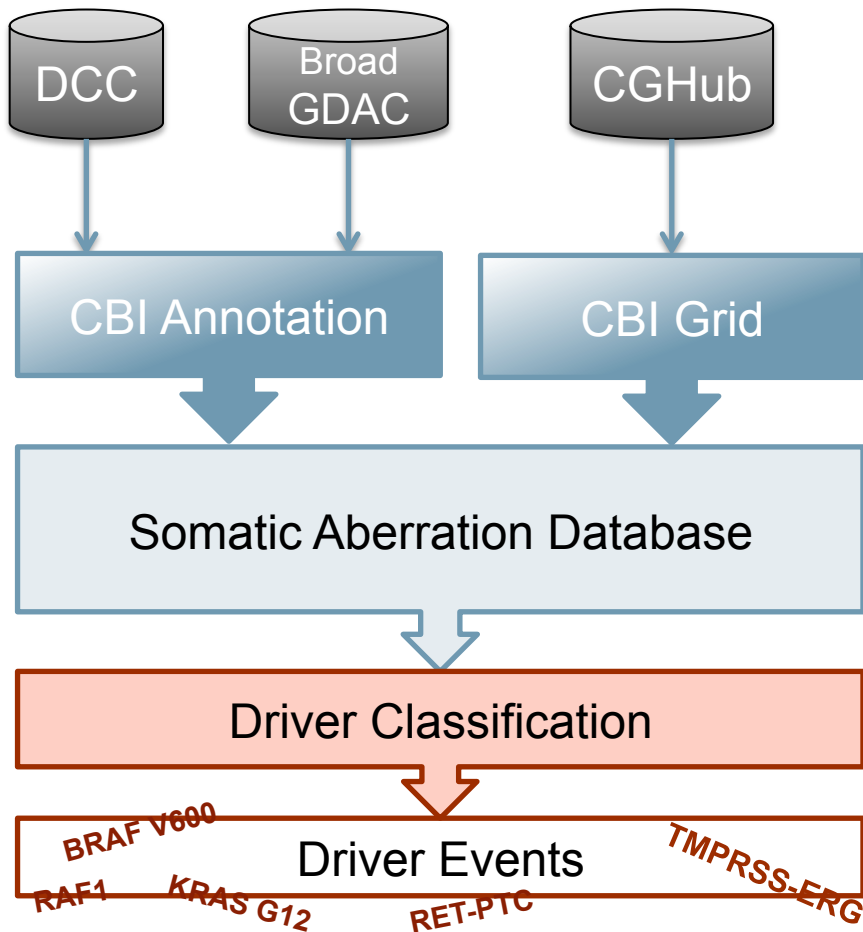
DATA: Large-scale data collection and processing

HETEROGENEITY: Different formats and analysis methods across data types and data centers

SPEED: Accommodating a rapidly growing dataset to get latest-and-greatest data hot off the TCGA “press”

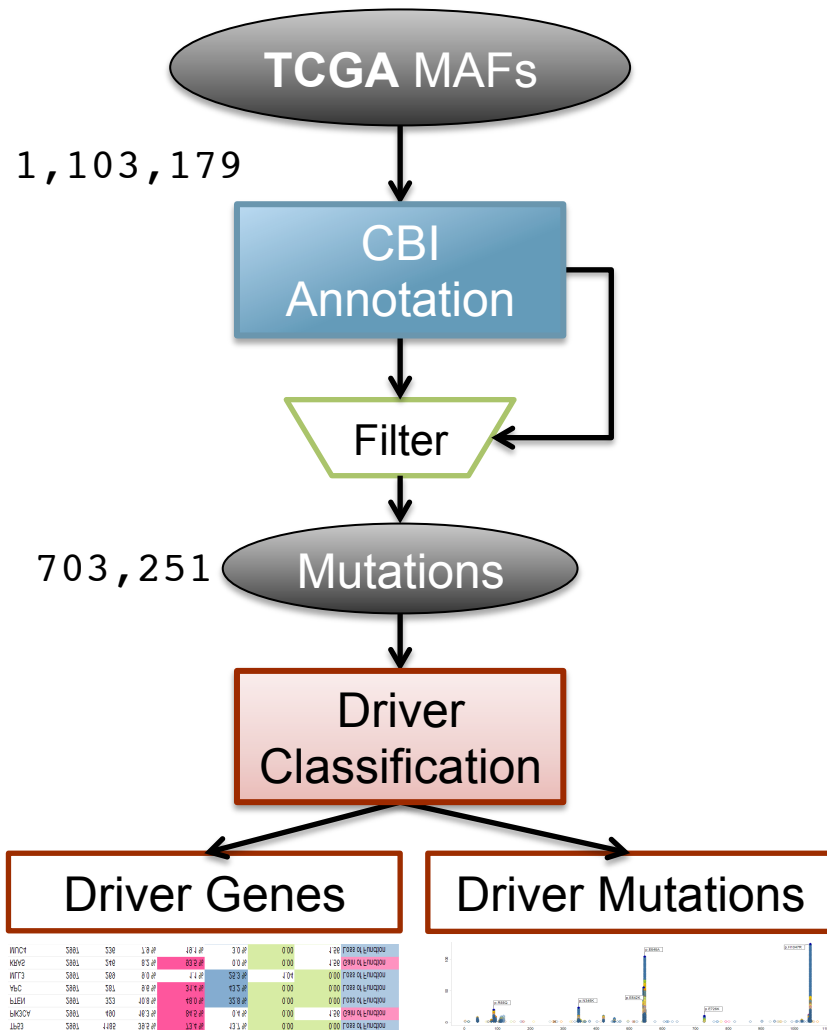
METHODS: Identifying true driver events from immature and evolving methods

Approach



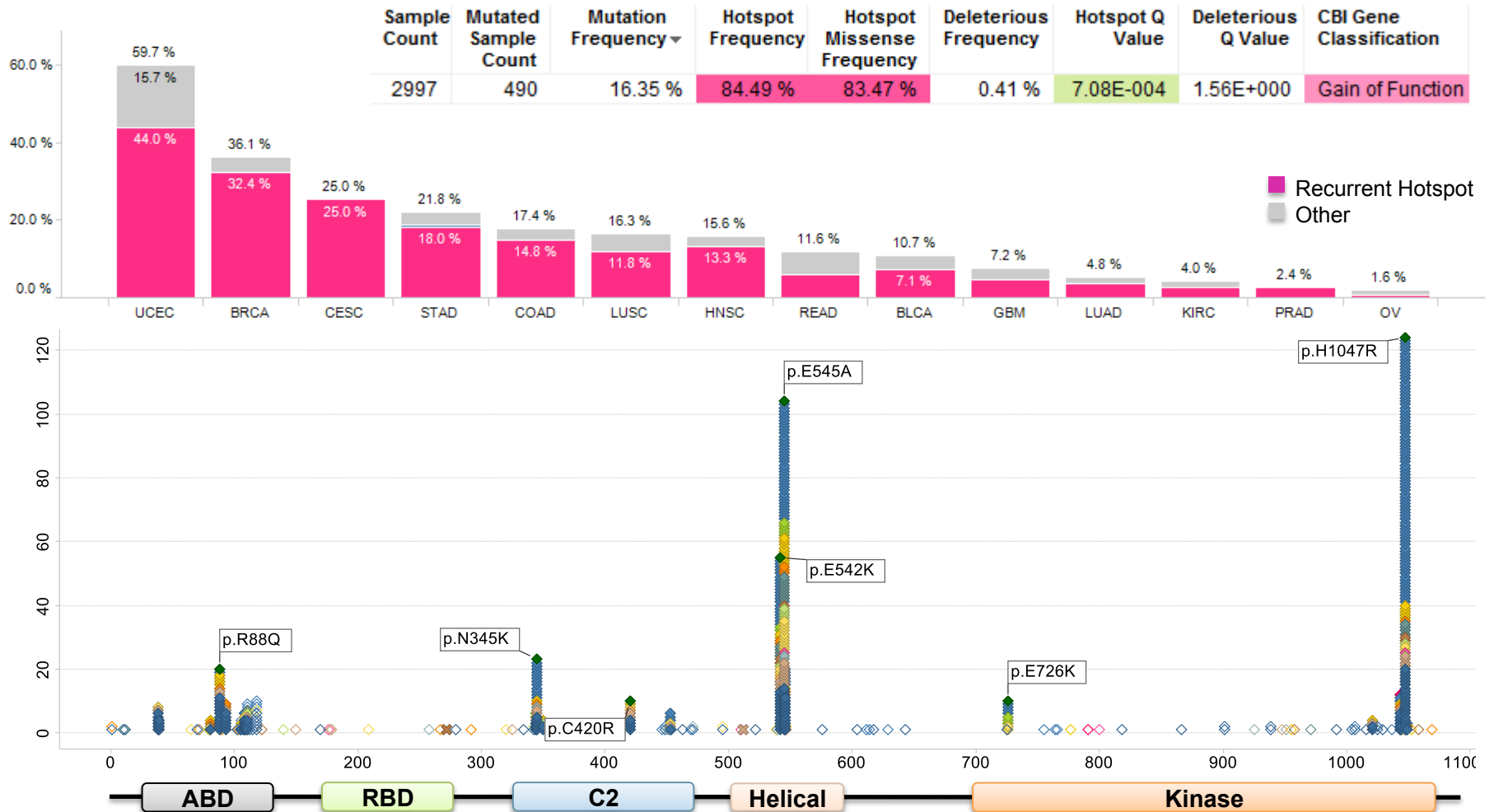
- » Quarterly data surveys
- » Method adaptation
- » HPC grid processing
- » Data standardization
- » Pan-cancer analysis
- » Gold-standard trained methods

Mutations: Result Summary

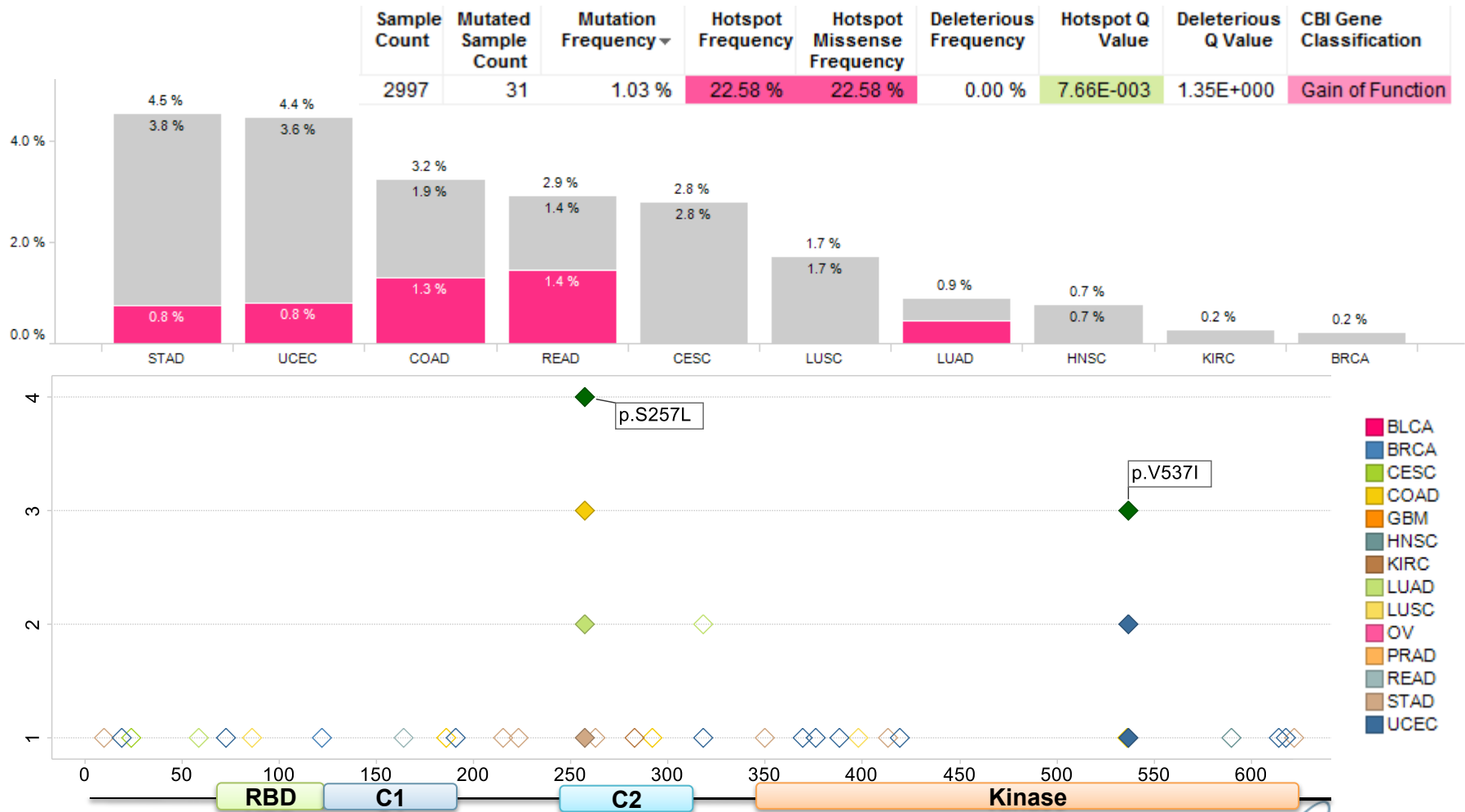


- » **2,998** samples from **15** diseases (Jul, 2012)
- » CBI Annotation
 - » Variant Classification
 - » Variant Position
- » Gene Classification
 - » Recurrent Hotspots
 - » Deleterious
 - » Statistical significance
 - » Gold-standards
- » Pan-cancer
 - » **107 “Gain of Function”** genes (e.g. Ras family)
 - » **120 “Loss of Function”** genes (e.g. PTEN, APC)
- » Within-cancer also calculated

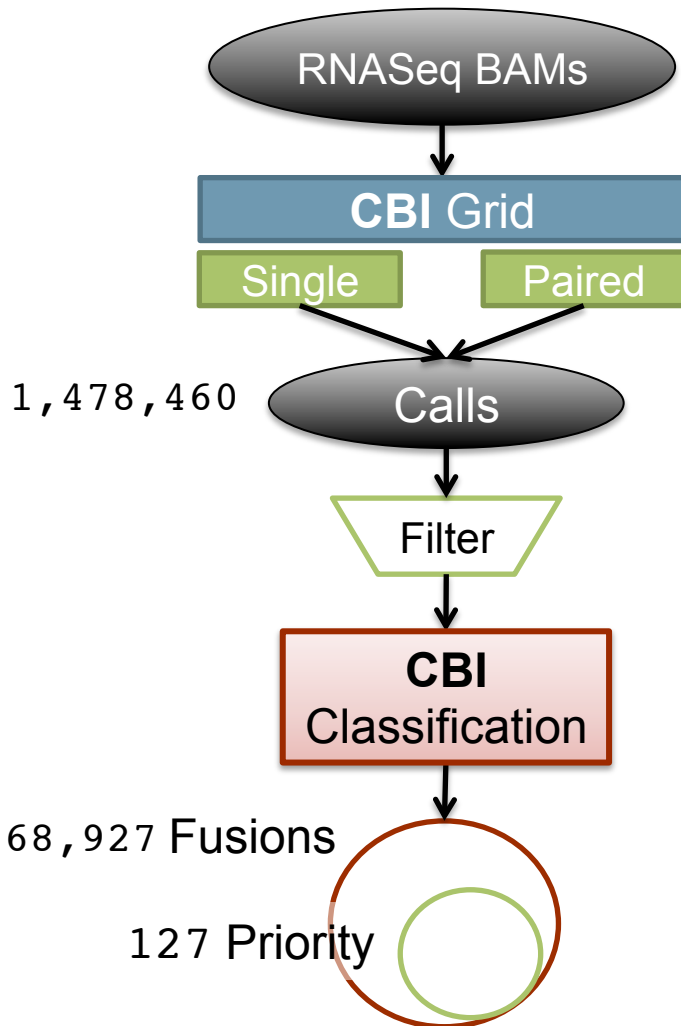
Mutations Example #1 – PIK3CA



Mutations Example #2 – RAF1 (CRAF)



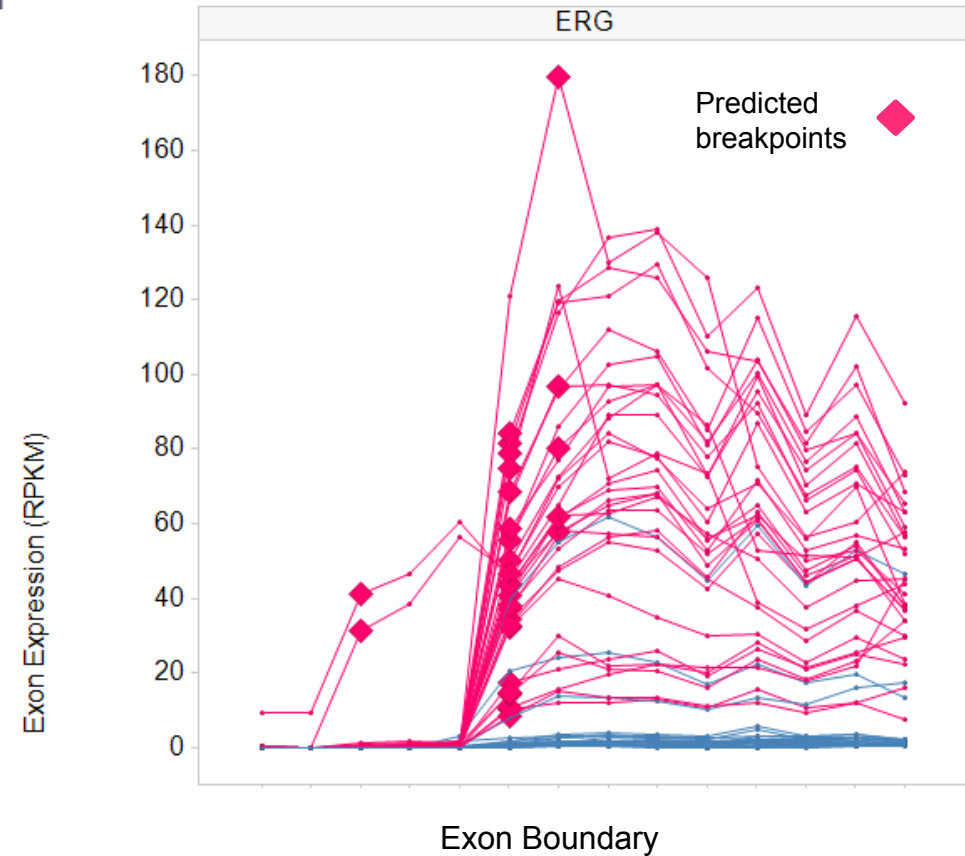
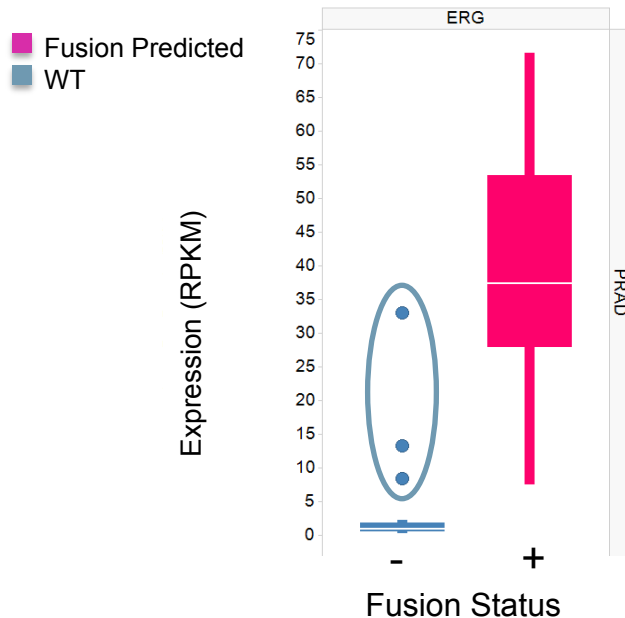
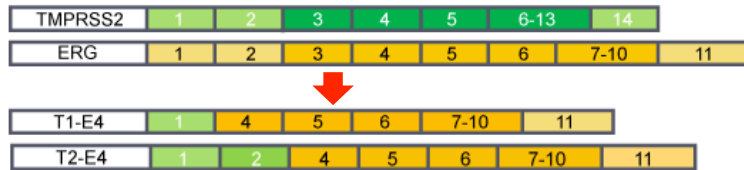
Fusions: Method Summary



- » **1,475** samples from **6** diseases
- » Adapted leading software for batch parallel processing
- » Annotate and filter against CBI gene model
- » Classify
 - » Dual-caller validation
 - » Recurrence
 - » Evidence-based scheme developed using gold-standard positives*
 - » Gene partners

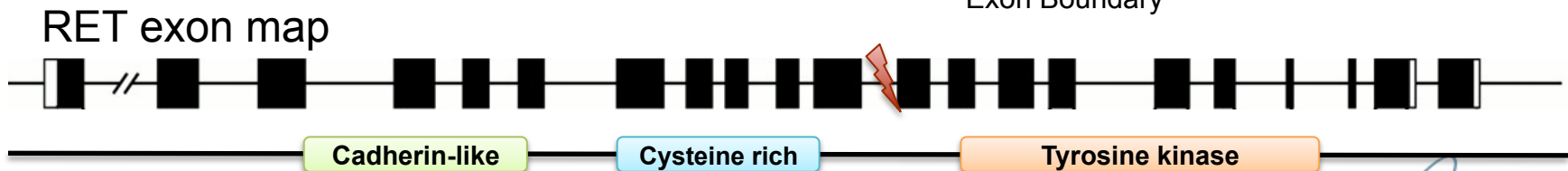
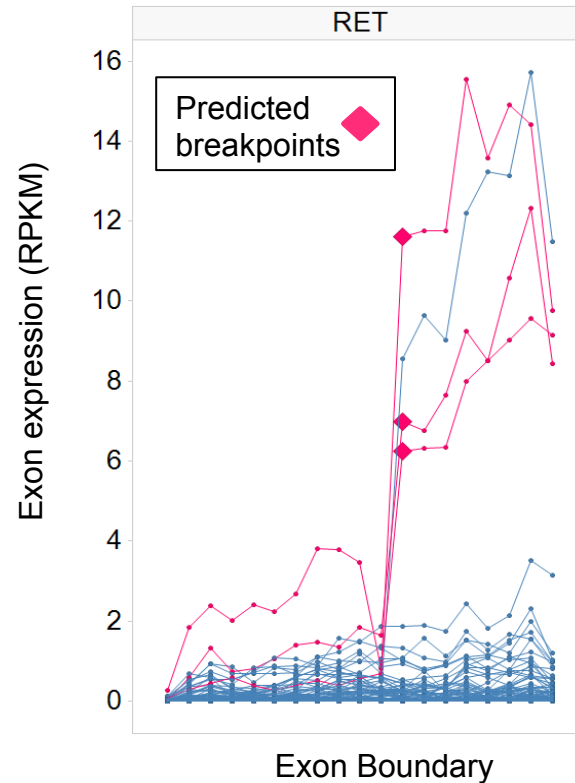
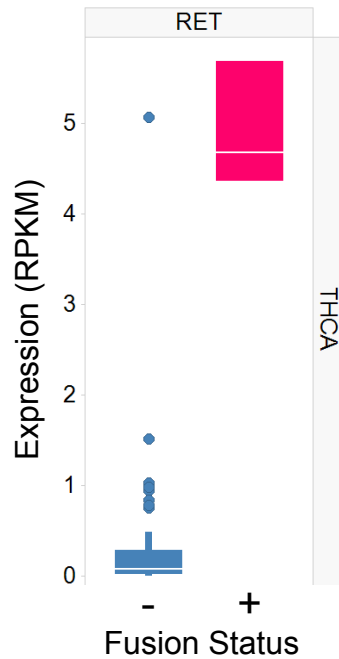
Fusions Example #1 – TMPRSS2-ERG

» 58% frequency in PRAD (30/53)



Fusions Example #2 – RET Fusions

- » Multiple RET partners across 3 diseases – 14 samples total



Future Direction

- » **TCGA:** important resource for foreseeable future
- » **Integrative analyses:**
 - » Gene, Pathway-level summarization
 - » Co-occurrence, mutual exclusivity, clinical subtype association, and outcome analysis
 - » Workflows for target discovery and target exploration
- » **More cancer types:** beyond TCGA
- » **Model systems:** map drivers
- » **Continued partnership with biopharma** to answer their most pressing questions

Acknowledgements

Compendia:

- » Dan Rhodes
- » Seth Sadis
- » Peter Wyngaard
- » Paul Williams
- » Supra Gajjala
- » Armand Bankhead

Research Informatics

Translational Medicine

Scientific Content

- » *Poster #7*

TCGA:

- » Kenna Shaw

Broad:

- » Michael Noble
- » Kristian Cibulskis

UC Santa Cruz:

- » Mark Diekhans
- » Chris Wilks

TopHat:

- » Daehwan Kim