

Mutation hotspots are associated with gene expression, signaling pathways, protein domains and drug response

Theo Knijnenburg



William Poole



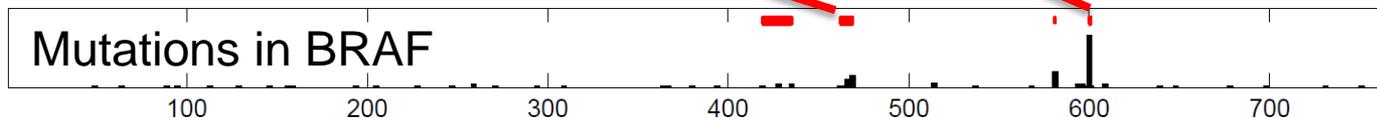
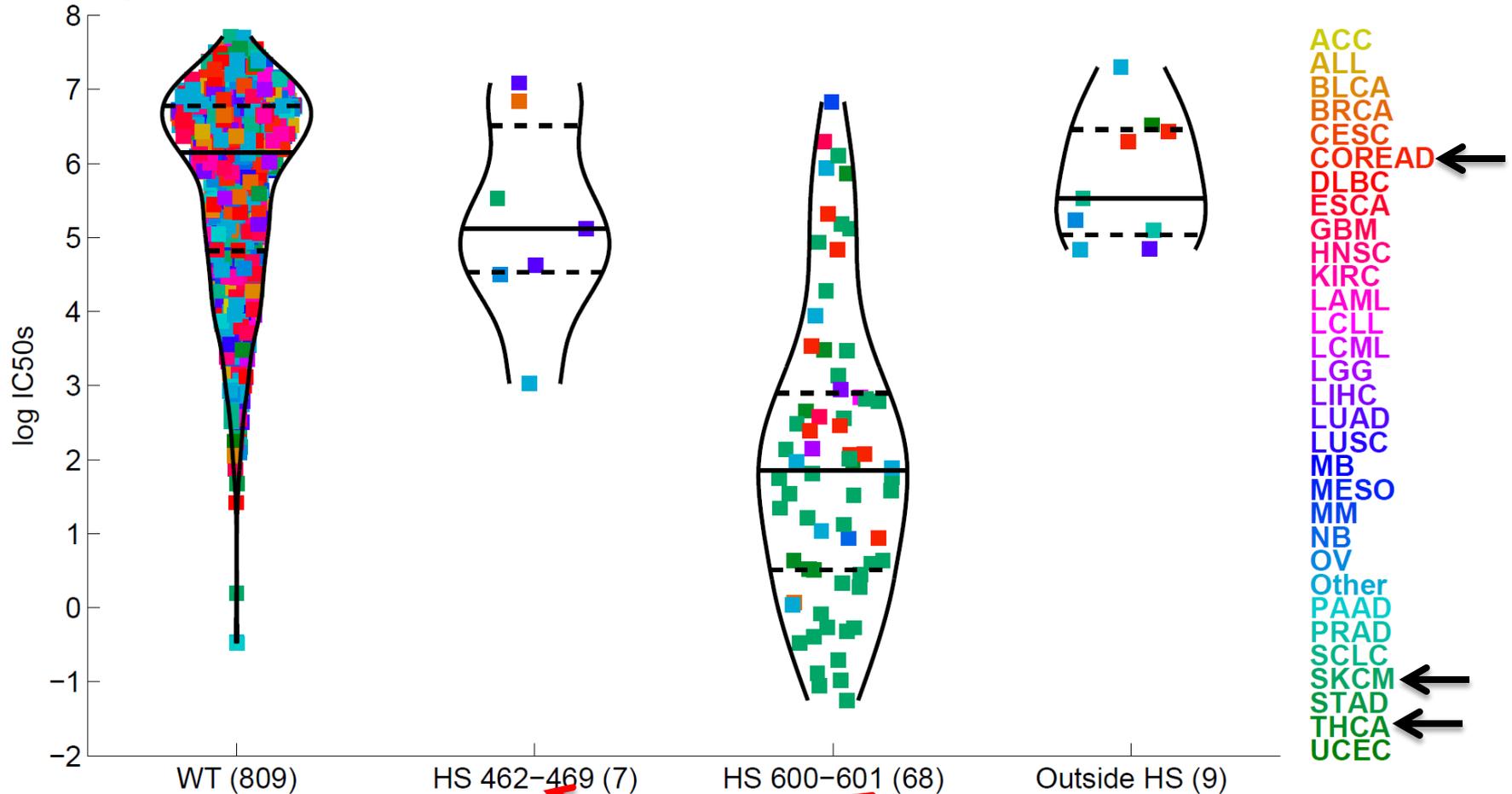
Brady Bernard



Ilya Shmulevich

The location of mutations in a gene matters

Drug response to Vemurafenib (inhibitor of V600E BRAF mutants) in GDSC 1000 cell line panel



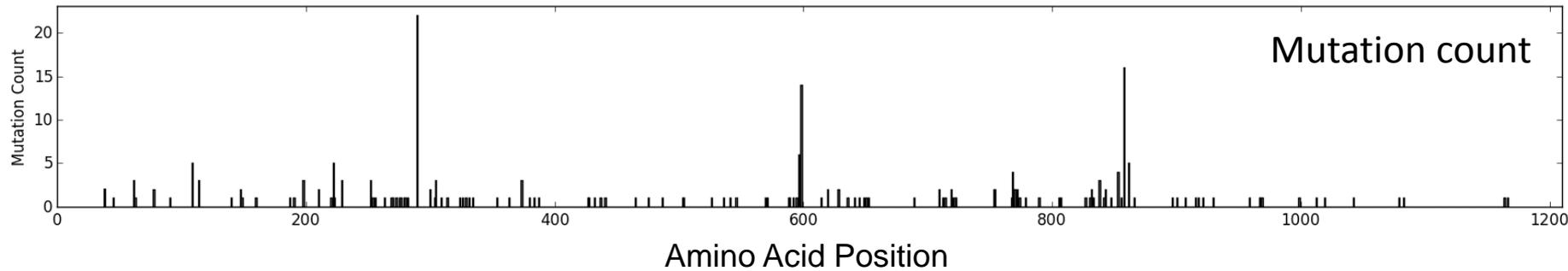
Finding mutation hotspots





Finding mutation hotspots

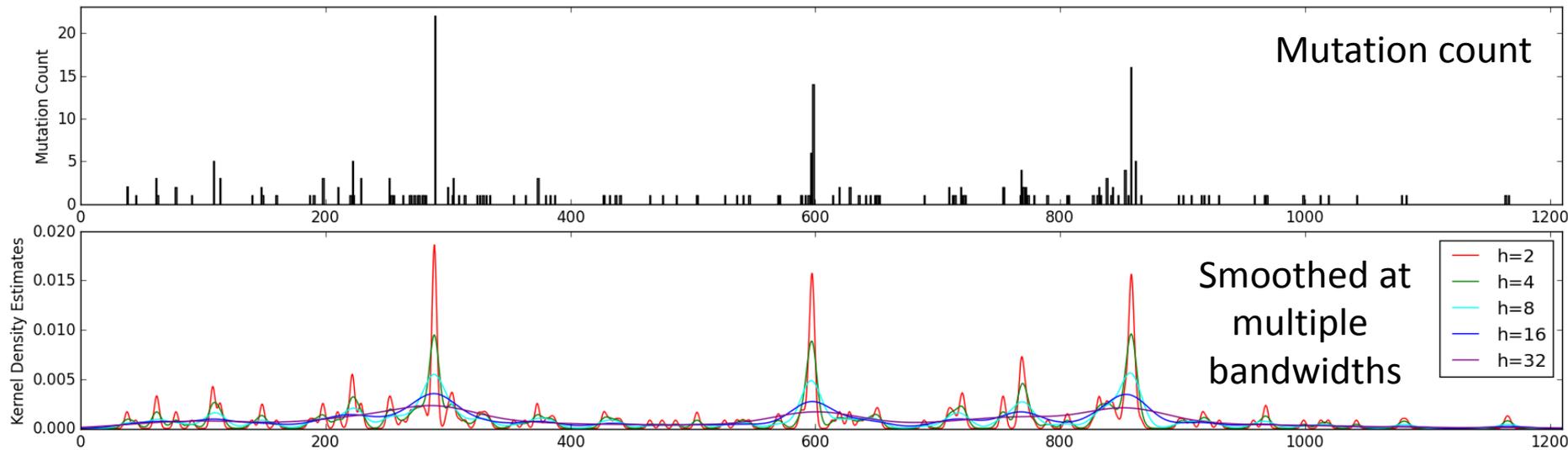
EGFR in PANCAN11





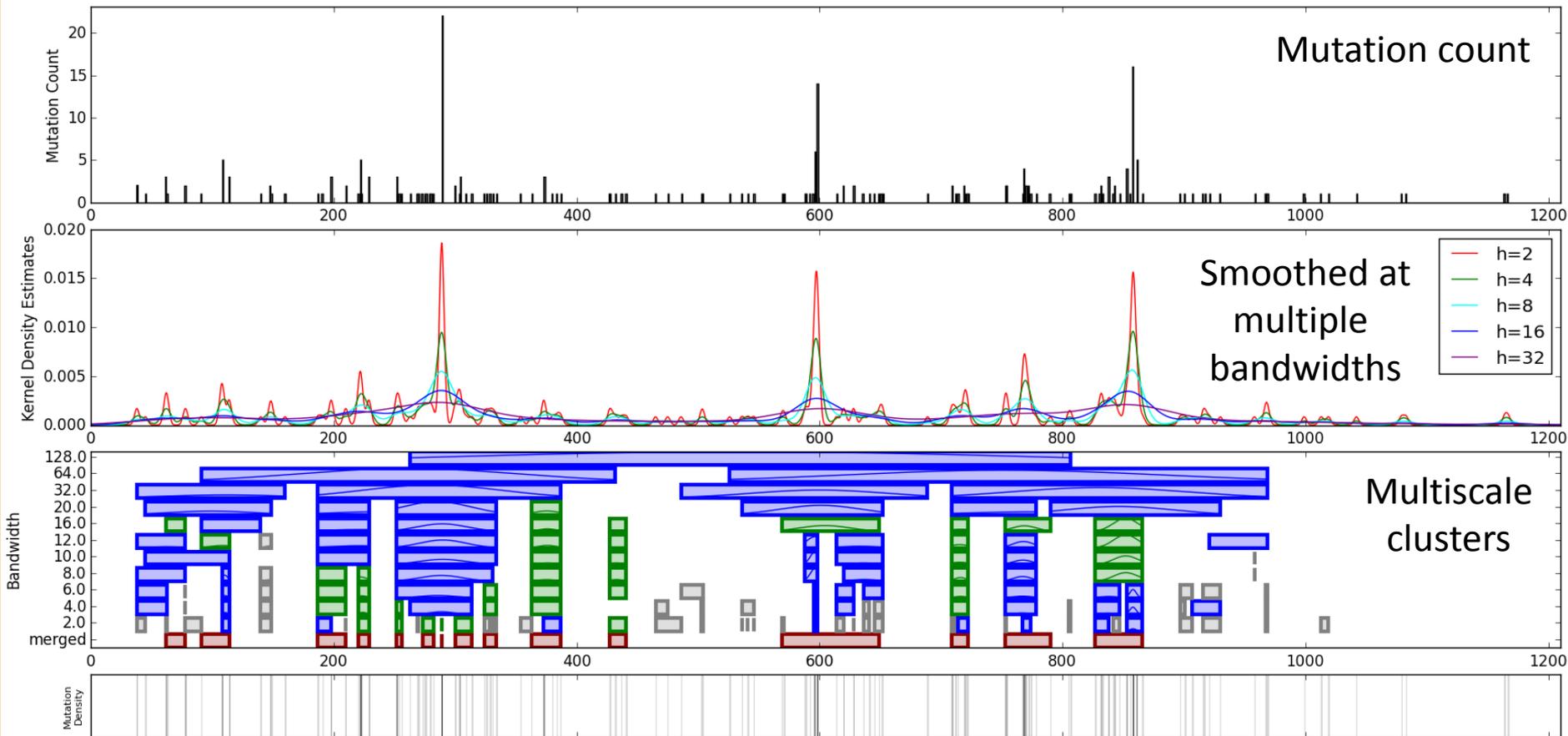
Finding mutation hotspots

EGFR in PANCAN11



Finding mutation hotspots

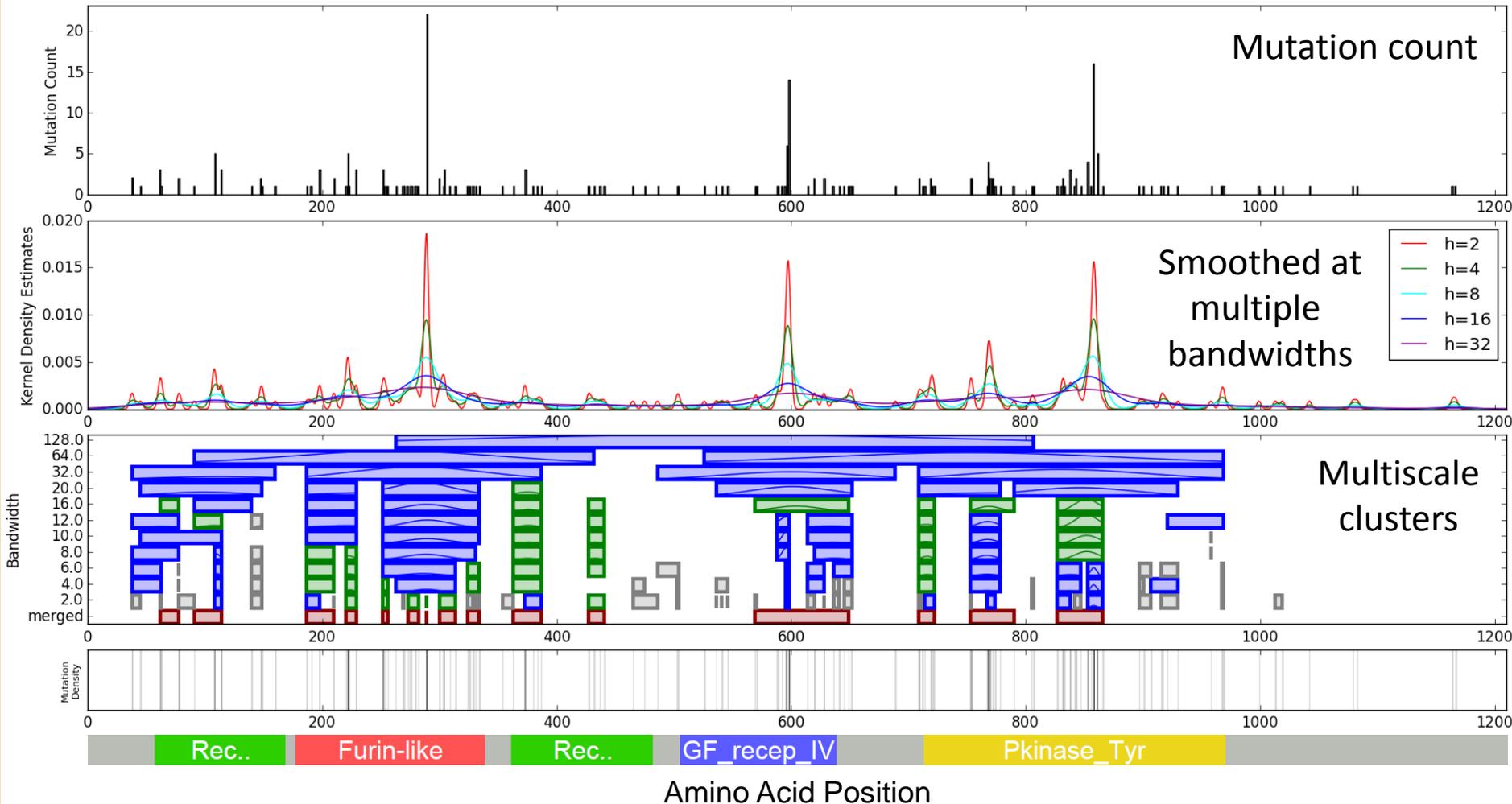
EGFR in PANCAN11





Finding mutation hotspots

EGFR in PANCAN11

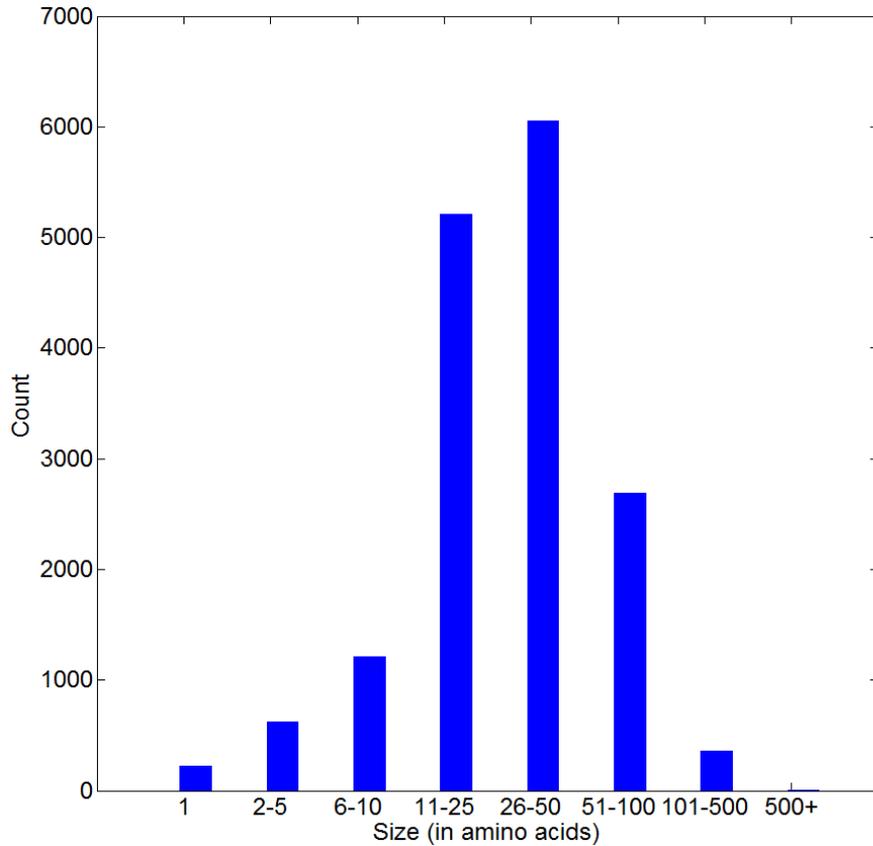


- 75% of the mutations in hotspots are inside protein domains



Mutation hotspots vary substantially in size

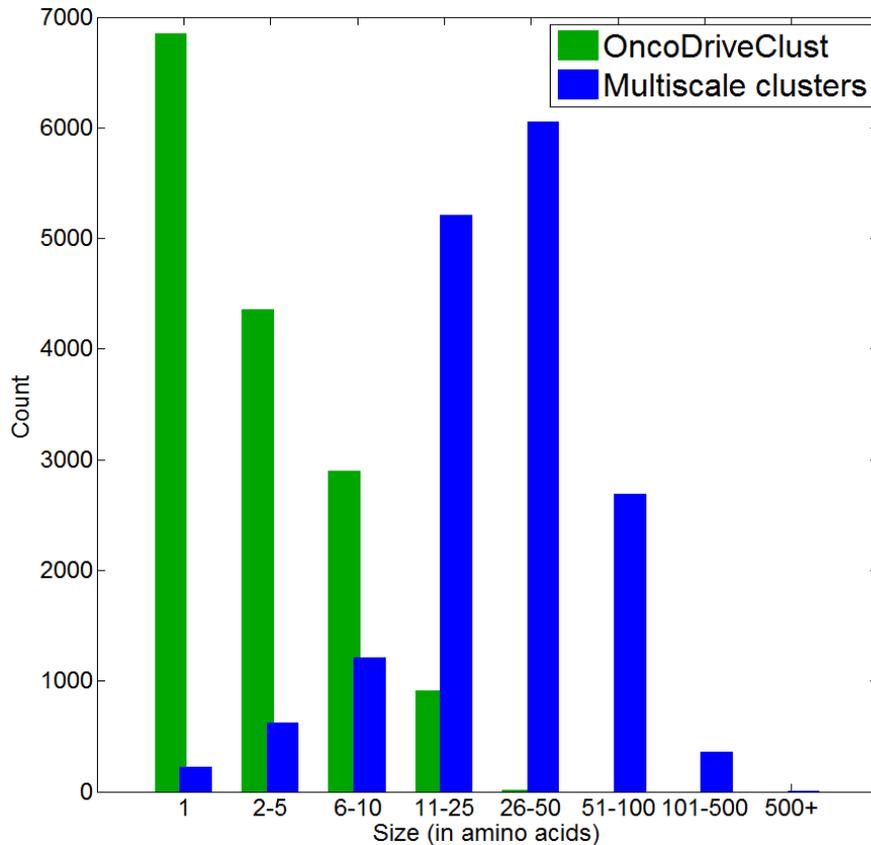
~15,000 hotspots in ~2500 genes in PANCAN11





Mutation hotspots vary substantially in size

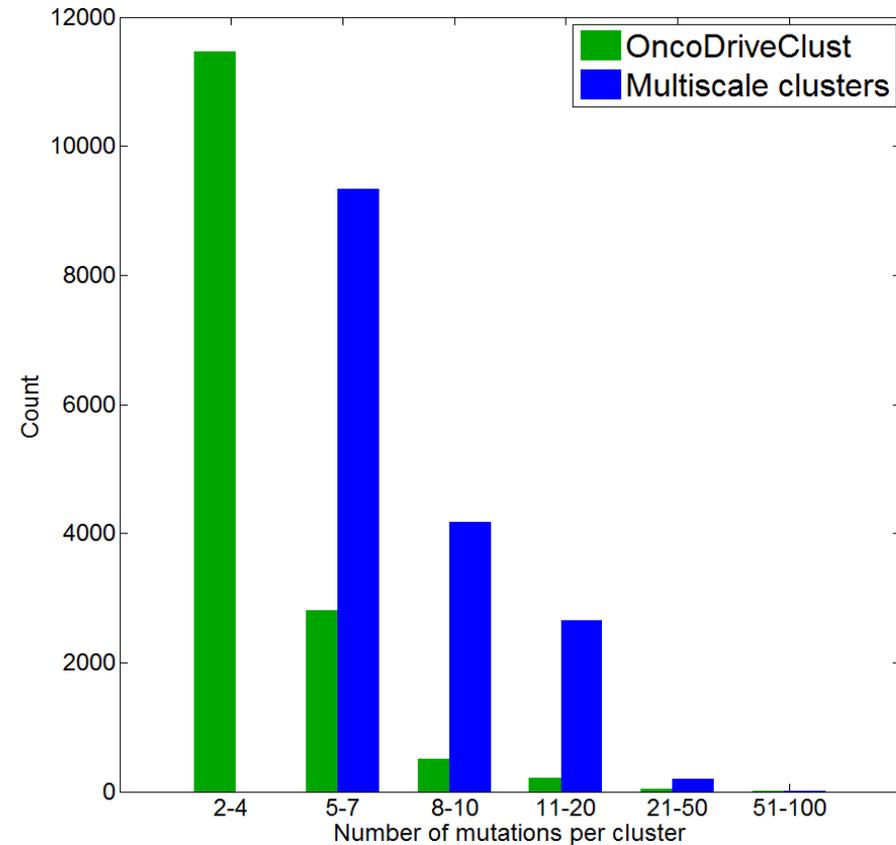
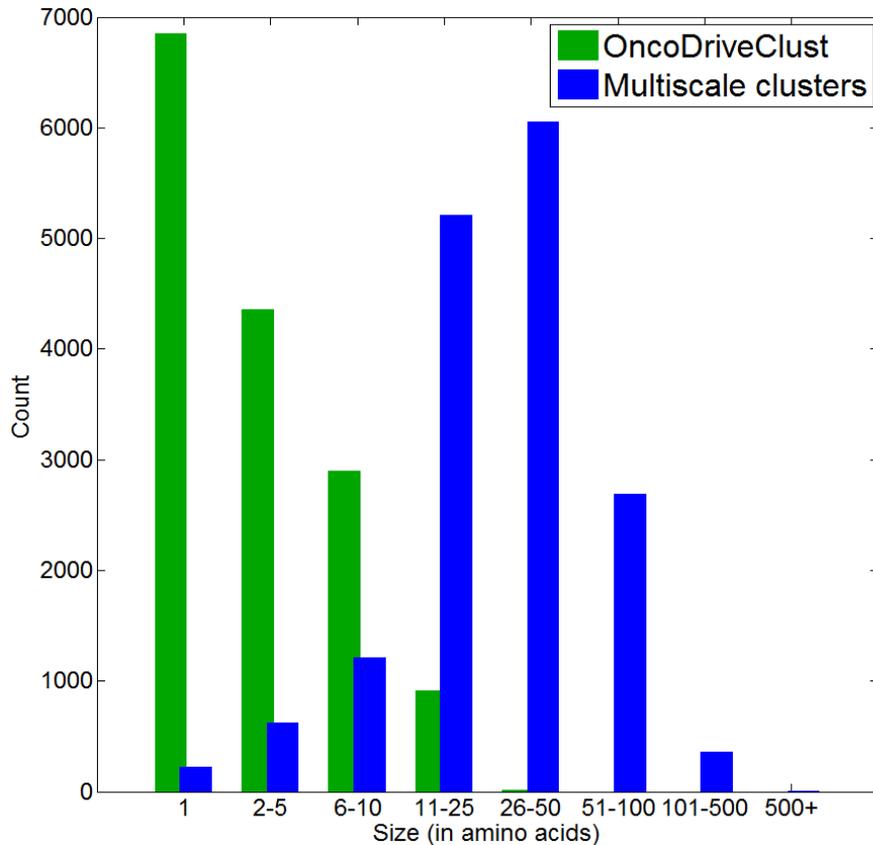
~15,000 hotspots in ~2500 genes in PANCAN11





Mutation hotspots vary substantially in size

~15,000 hotspots in ~2500 genes in PANCAN11



- 84% of the **OncoDrive** clusters overlap with the **Multiscale** clusters



Are these mutation hotspots functional?

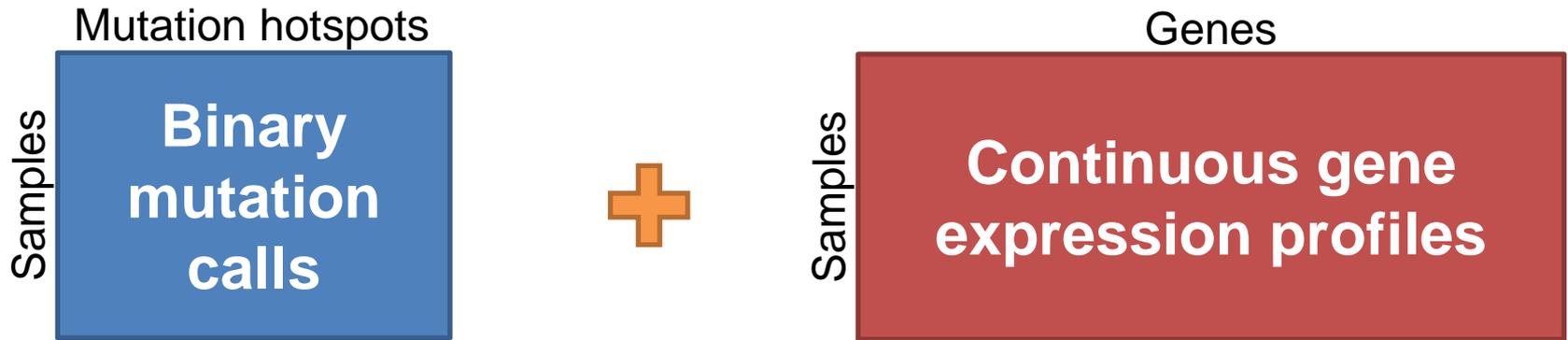


Are these mutation hotspots functional?

Identify statistical associations with:

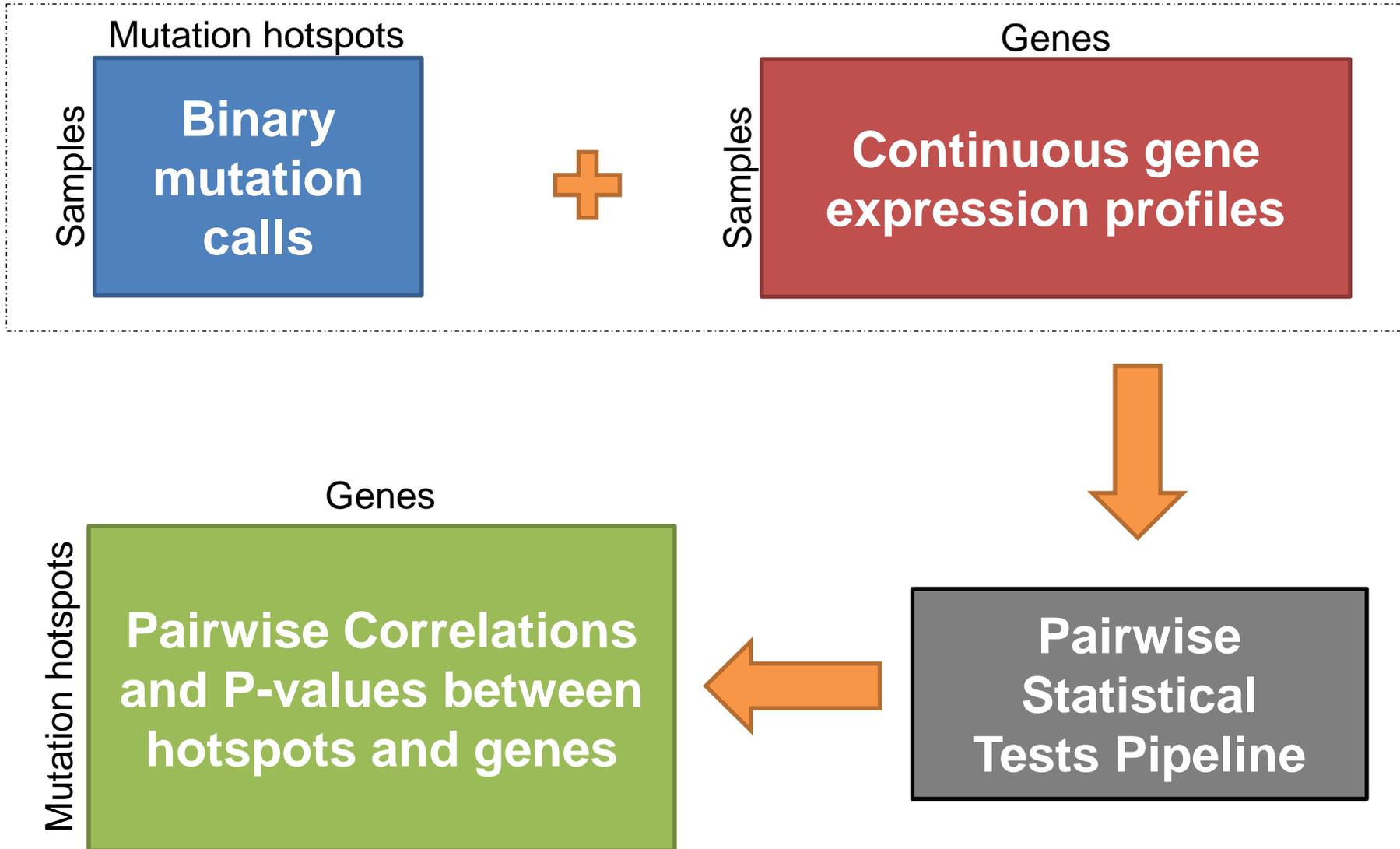
- Protein domains
- Gene expression
- Signaling pathways
- Drug response

Statistical association between hotspots and gene expression

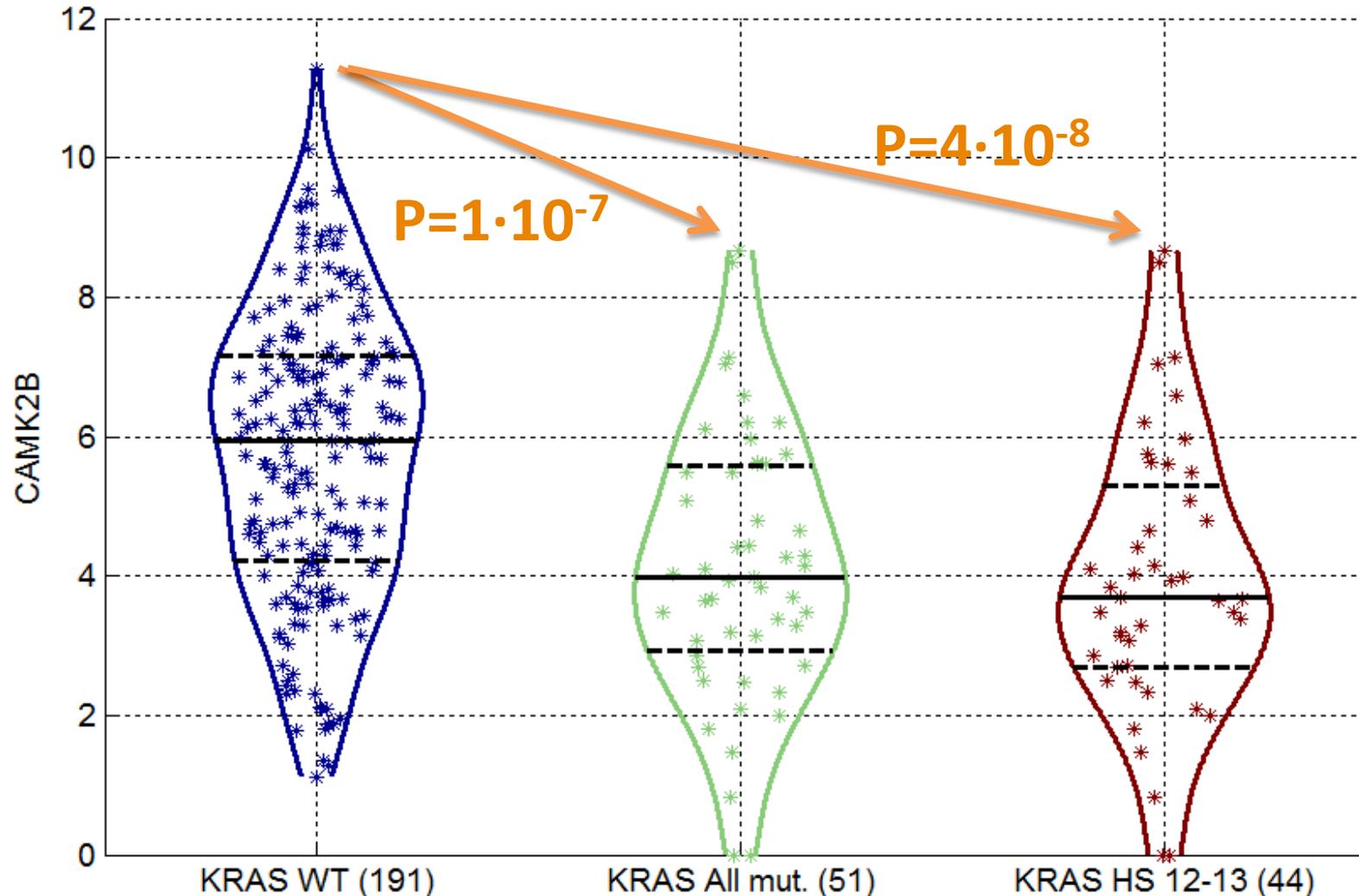




Statistical association between hotspots and gene expression

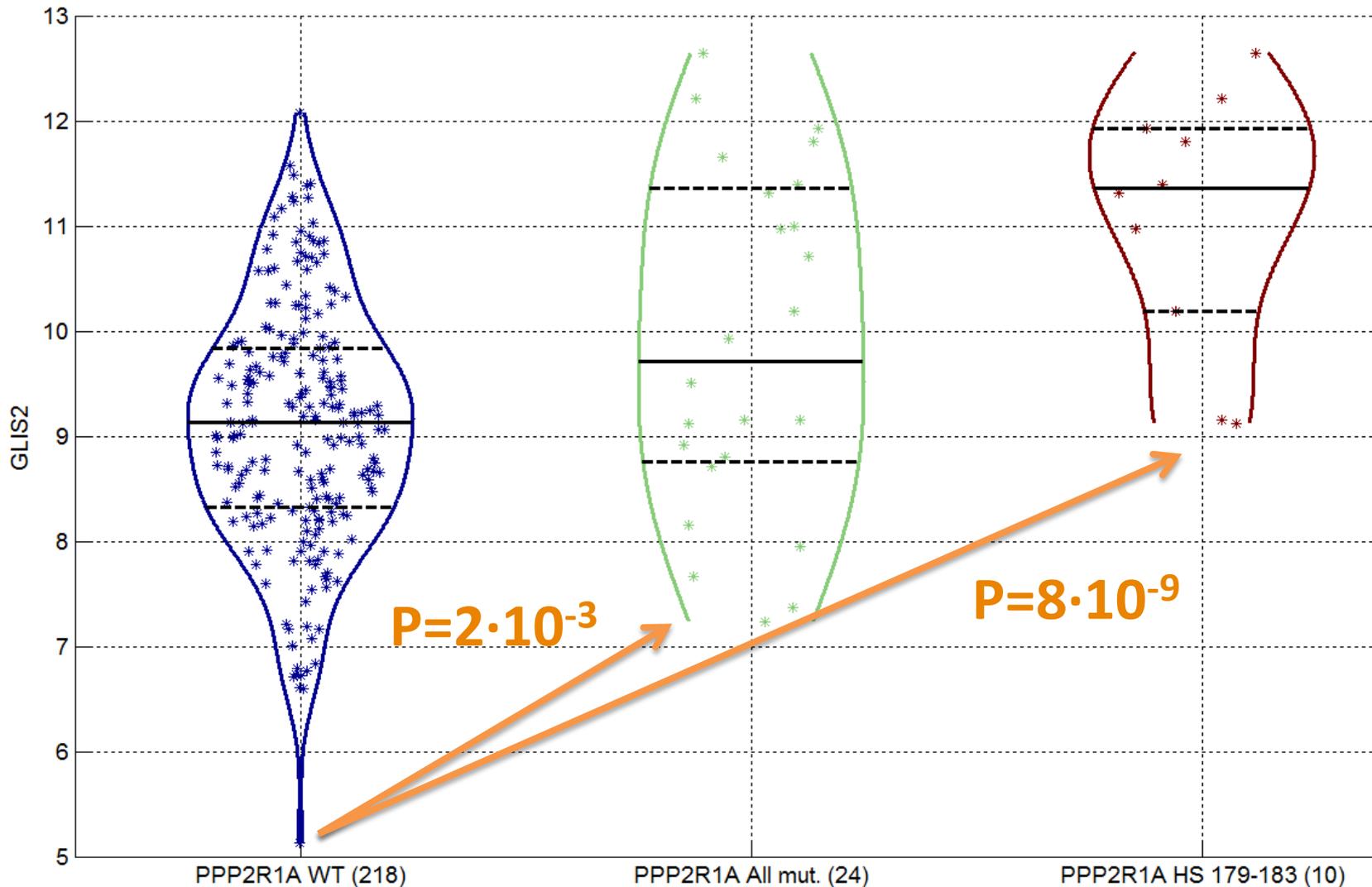


Statistical association between hotspots and gene expression in UCEC





Statistical association between hotspots and gene expression in UCEC



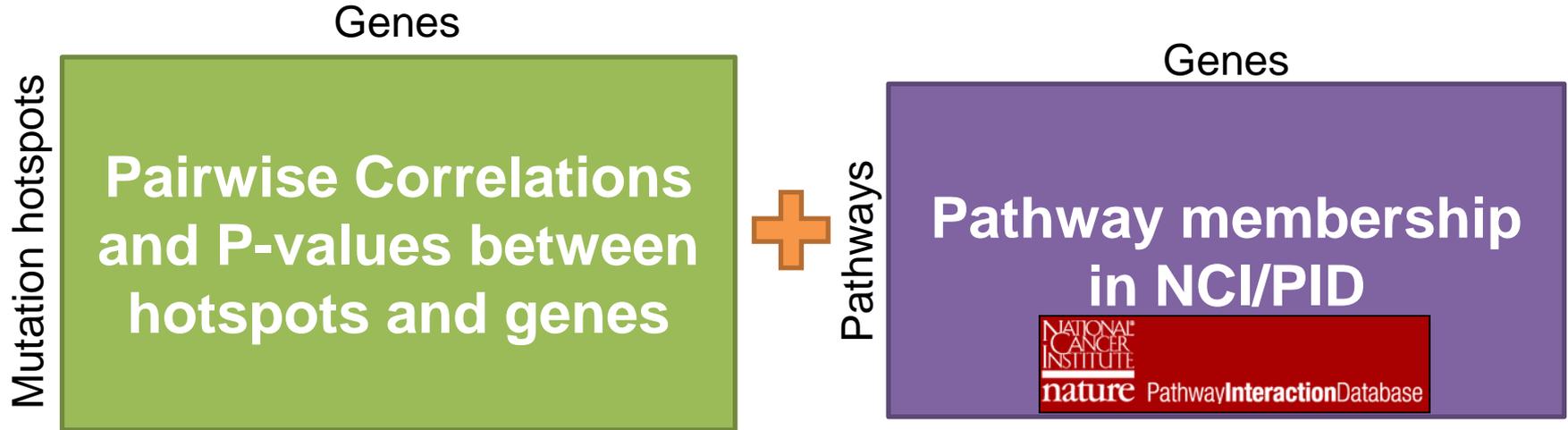


Statistical association between hotspots and gene expression in UCEC

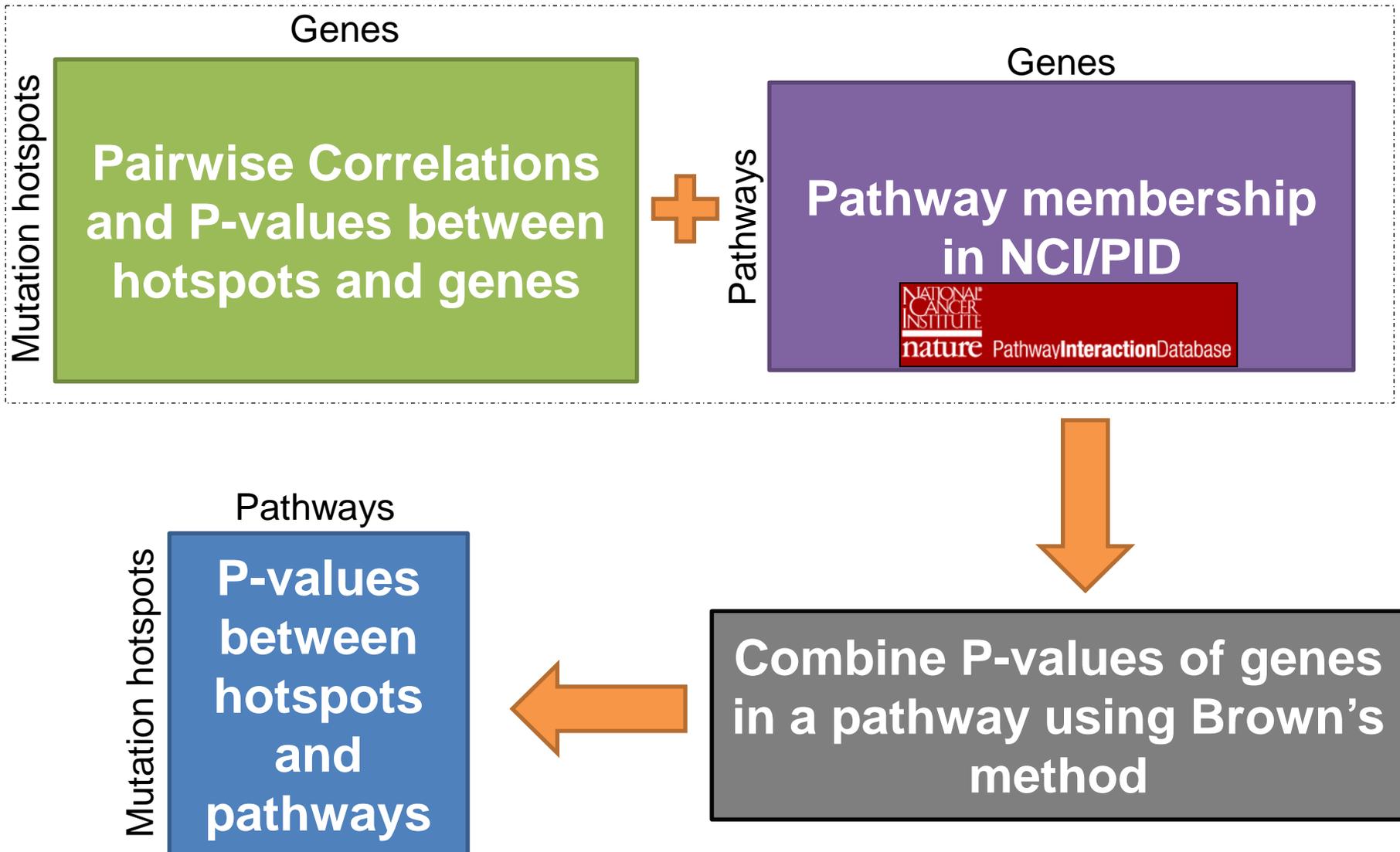
Type	# Significant associations
Only in hotspot (hotspot $P < 10^{-4}$, all mutations $P > 10^{-2}$)	60
Only in gene (all mutations $P < 10^{-4}$, hotspot $P > 10^{-2}$)	1350
Both in gene and hotspot (all mutations $P < 10^{-4}$, hotspot $P < 10^{-4}$)	1360



Statistical association between hotspots and signaling pathways

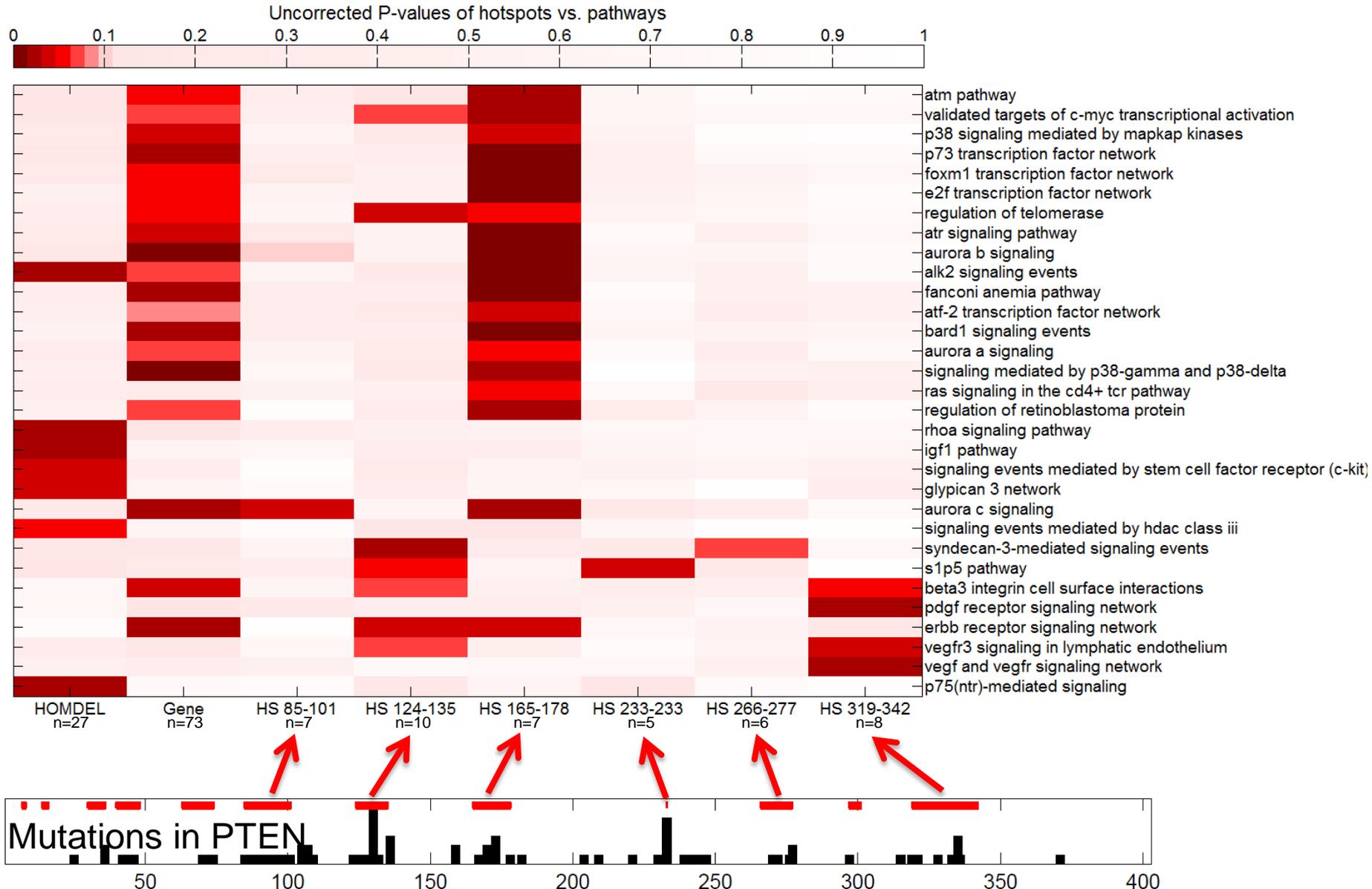


Statistical association between hotspots and signaling pathways

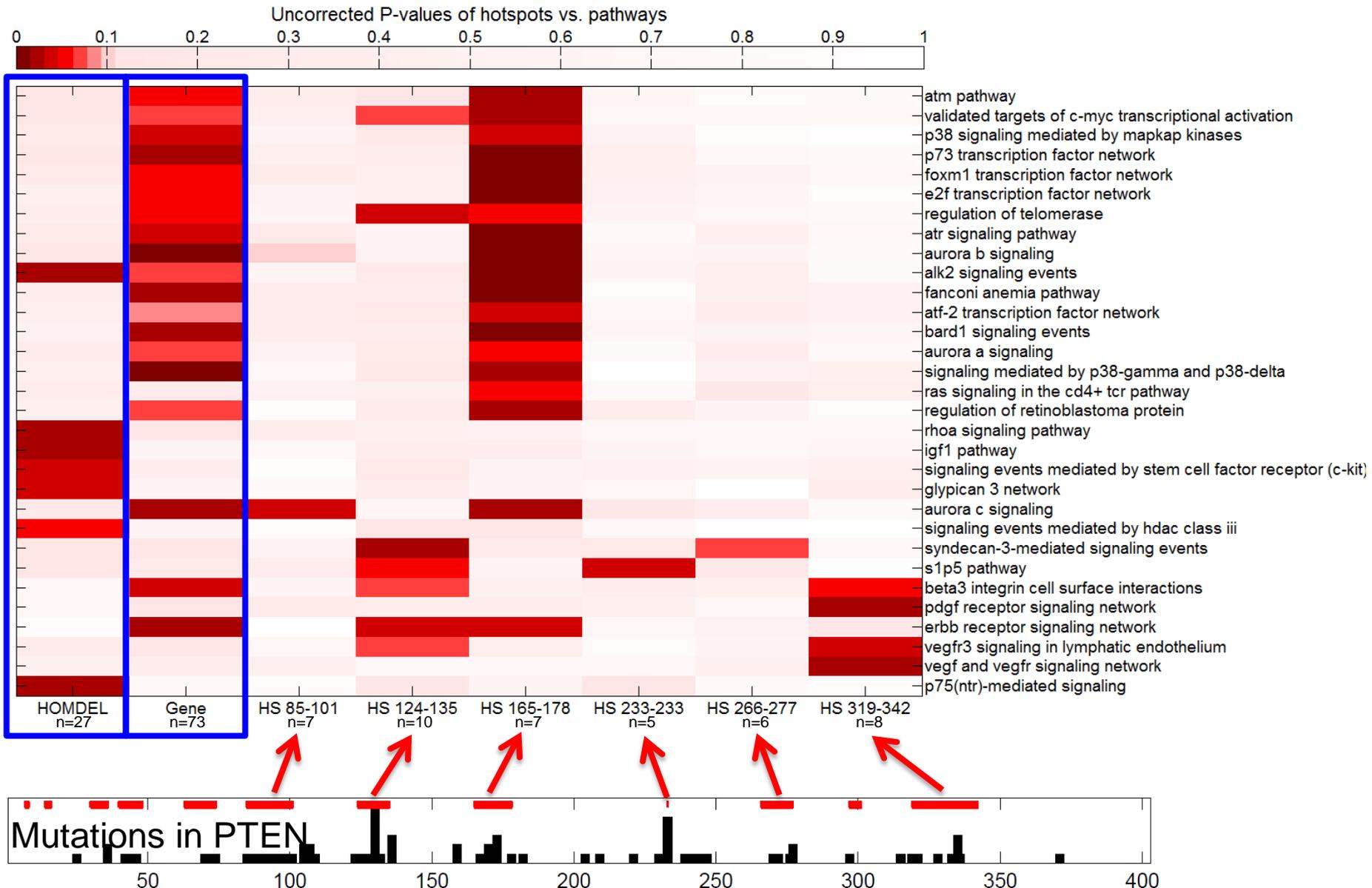




Statistical association between PTEN hotspots and signaling pathways in GBM

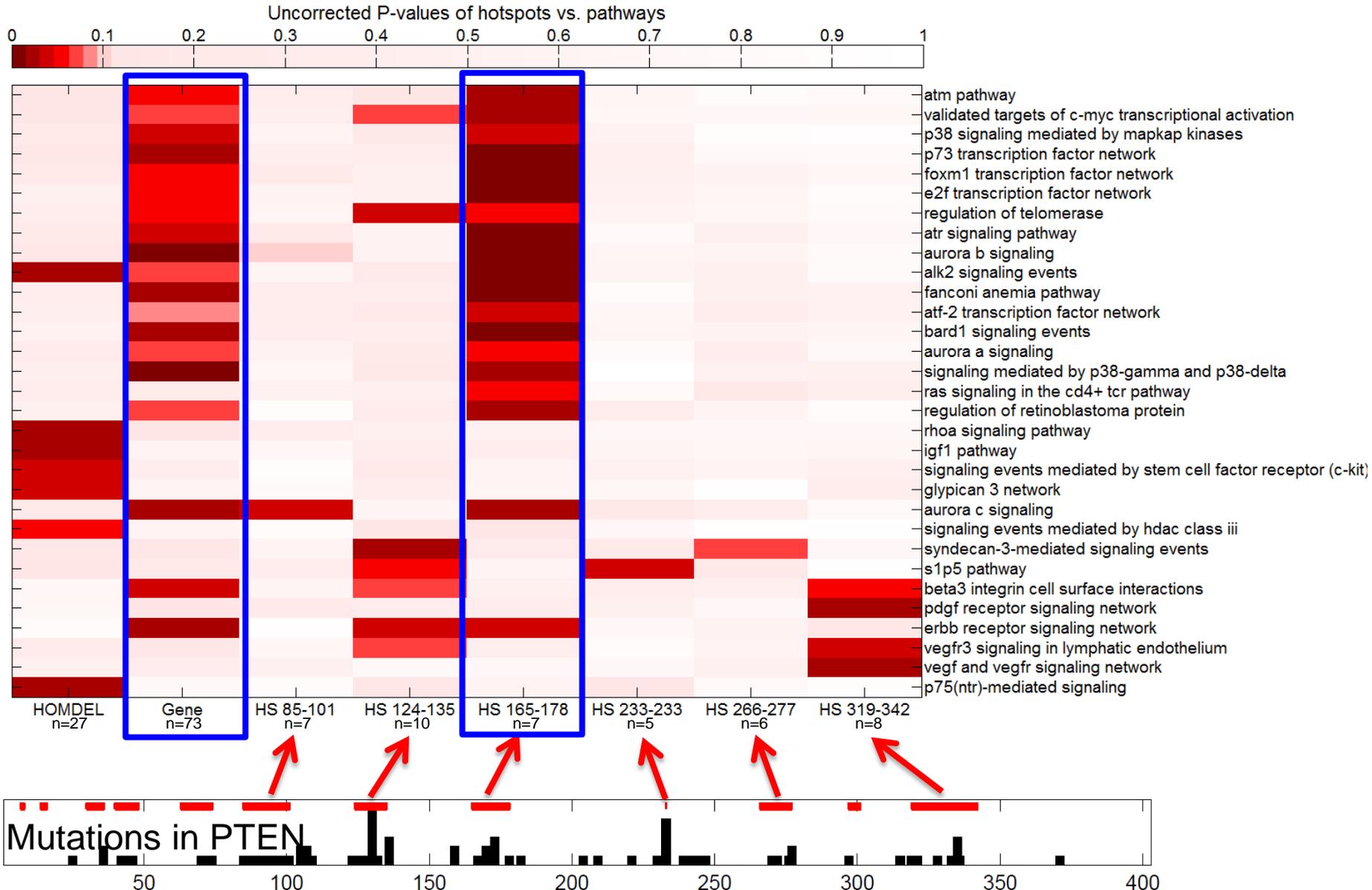


Statistical association between PTEN hotspots and signaling pathways in GBM



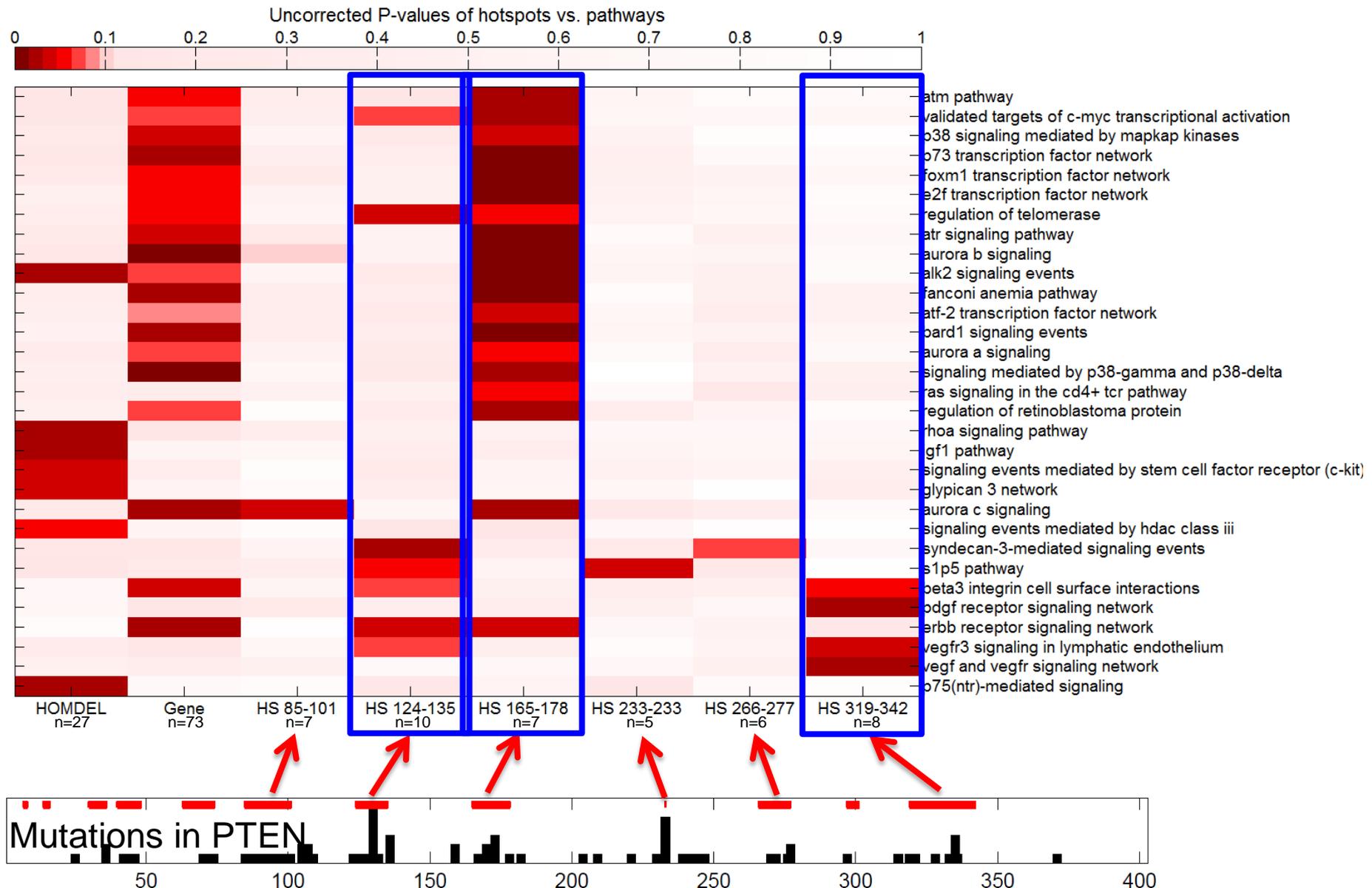


Statistical association between PTEN hotspots and signaling pathways in GBM

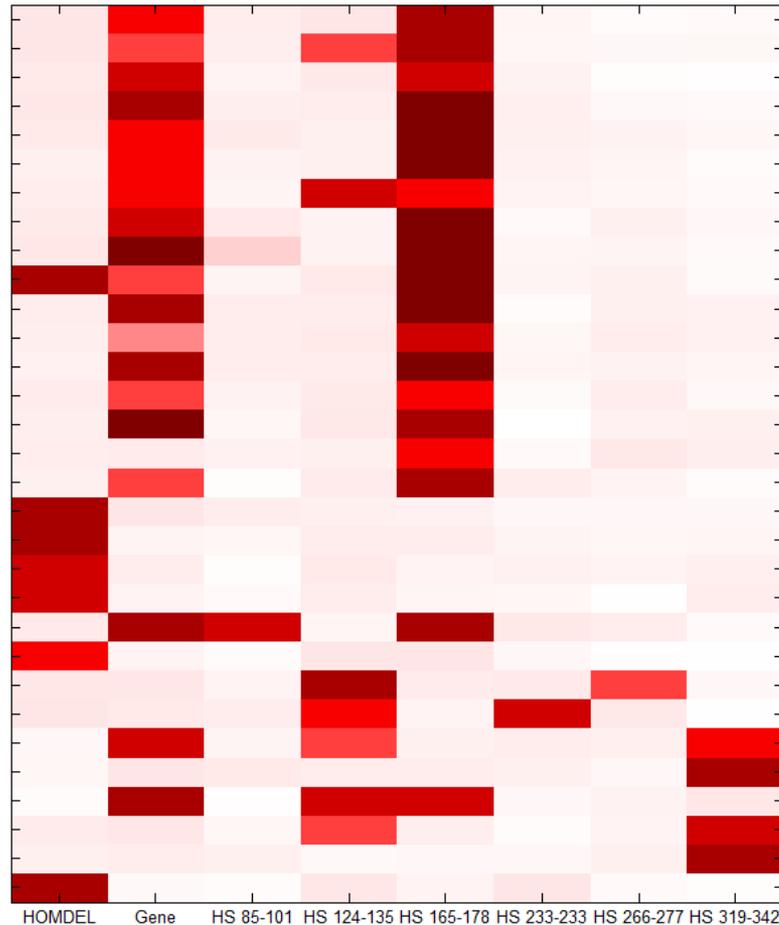




Statistical association between PTEN hotspots and signaling pathways in GBM



Statistical association between PTEN hotspots and signaling pathways in GBM



WPD Loop

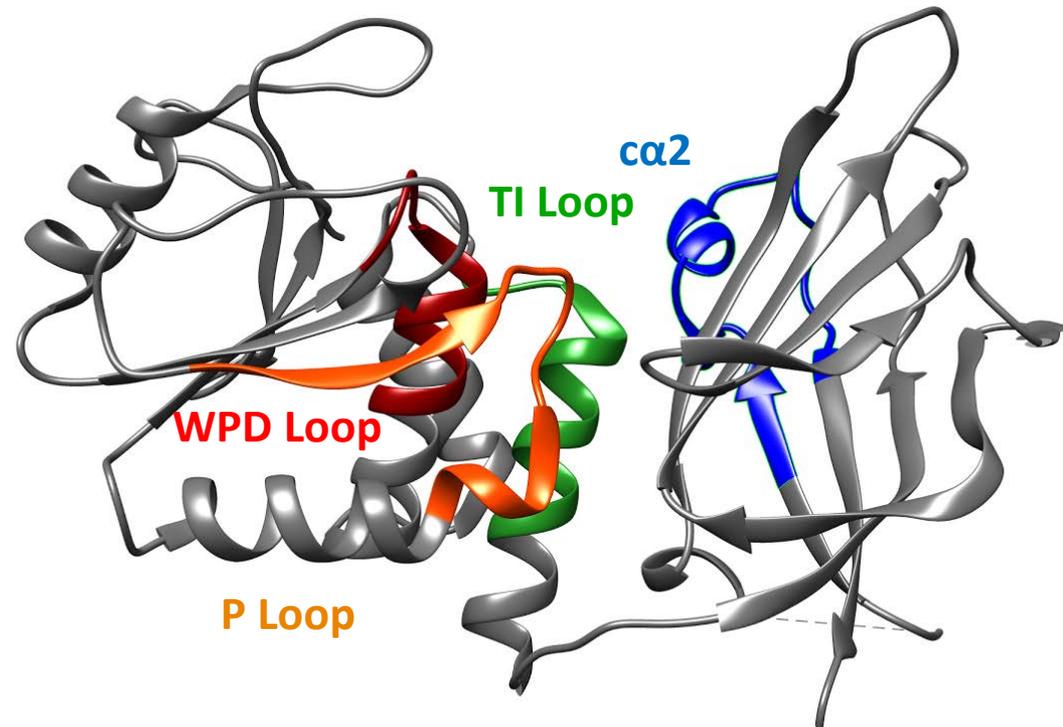
P Loop

TI Loop

$\alpha 2$

Phosphatase domain

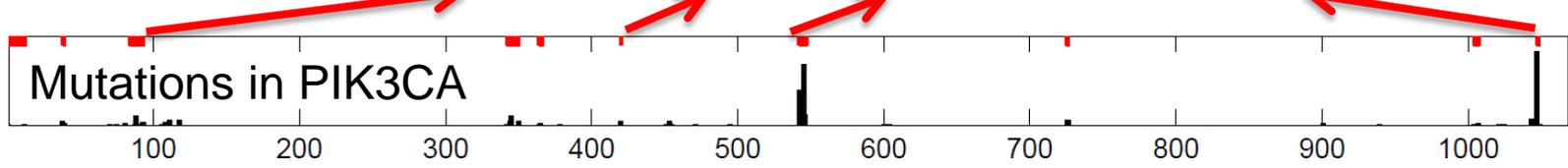
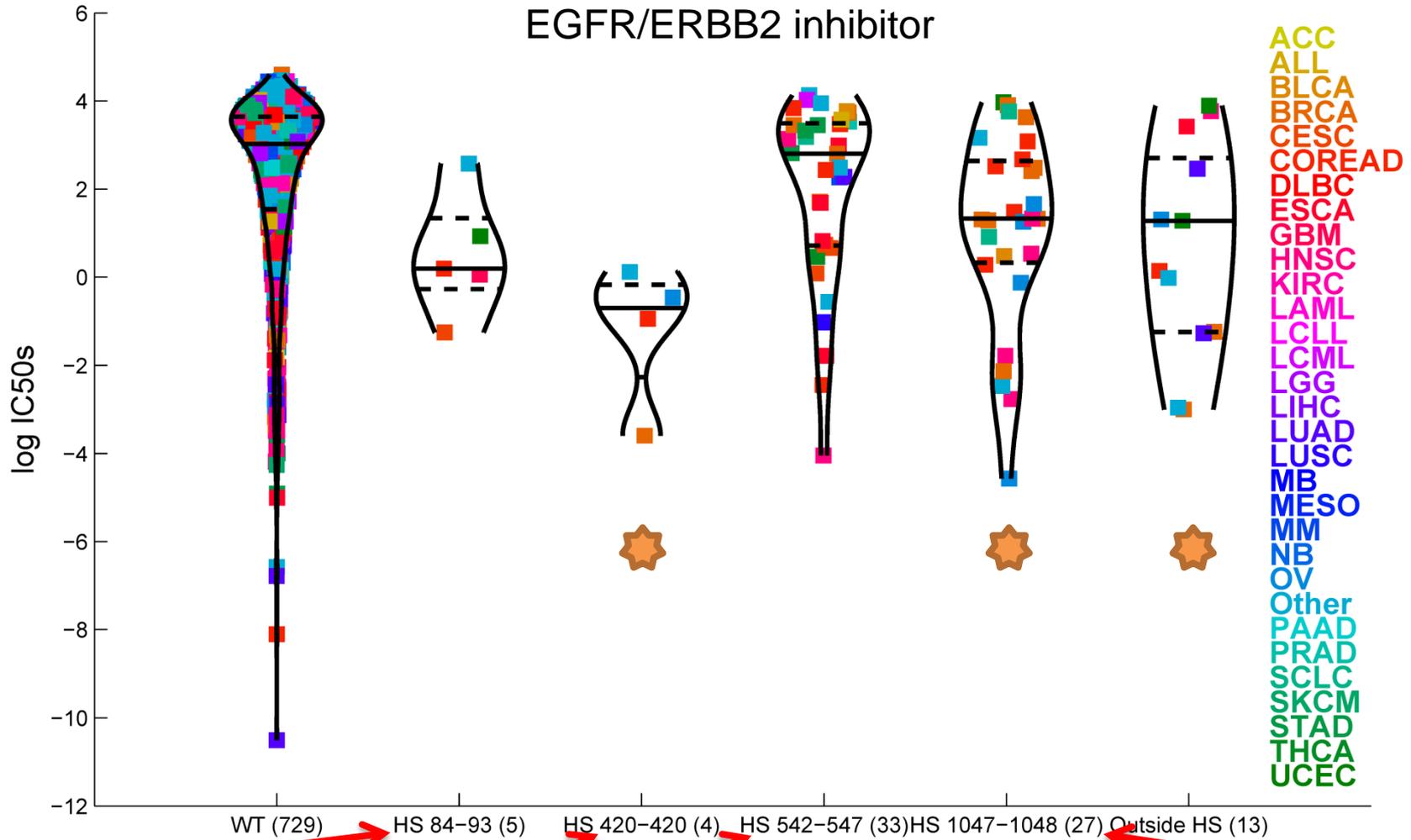
C2 domain



Lee et al. Crystal Structure of the PTEN Tumor Suppressor: Implications for Its Phosphoinositide Phosphatase Activity and Membrane Association. Cell, 1999.

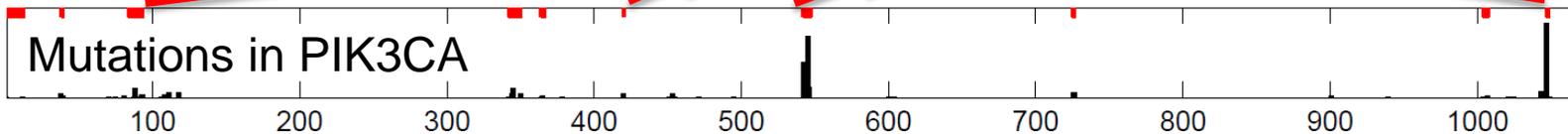
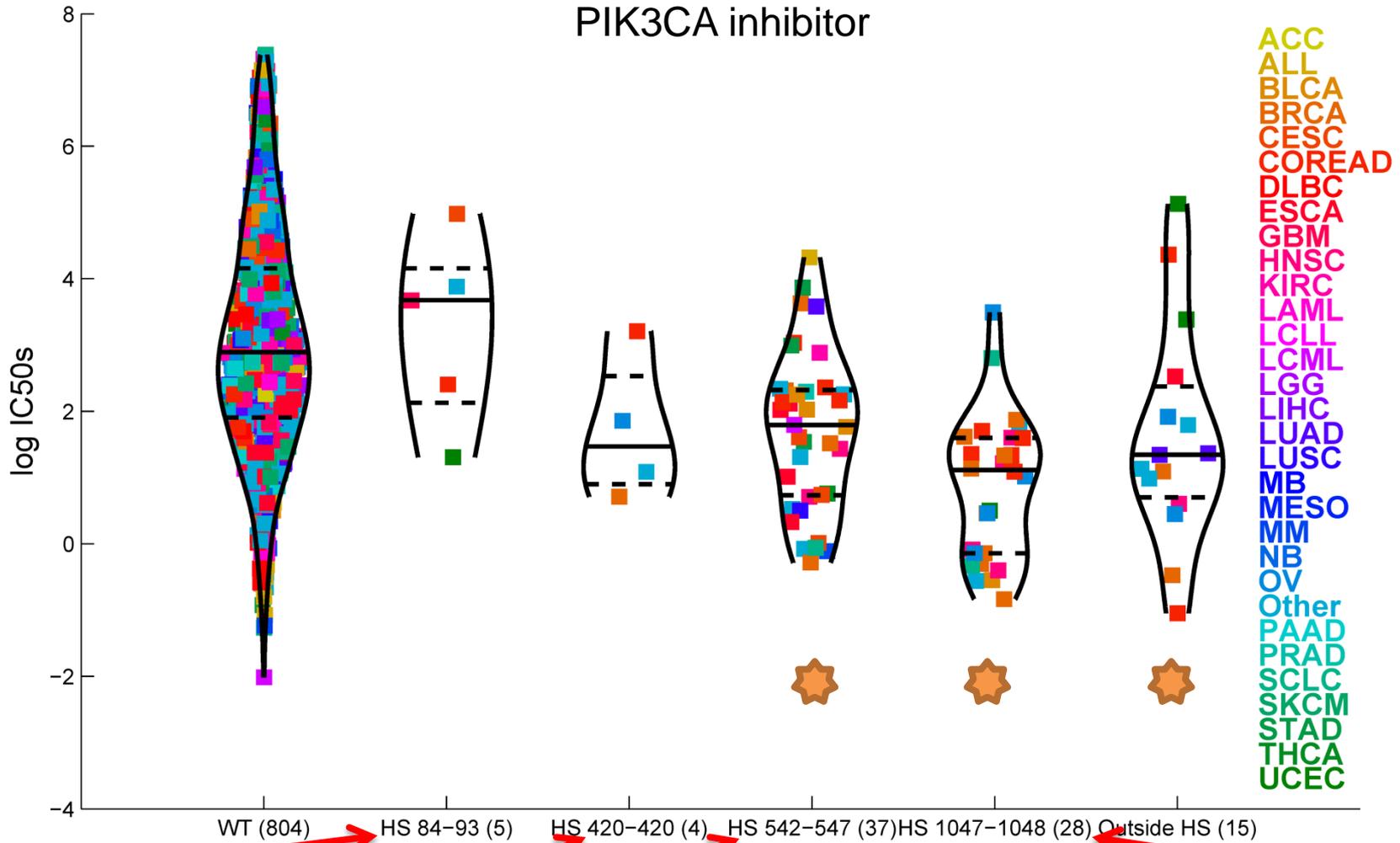


Drug response depends on mutation hotspots





Drug response depends on mutation hotspots





Summary

- We have developed a novel multiscale clustering algorithm to robustly identify mutation hotspots in genes
- We uncovered statistical associations between many of these mutation hotspots and gene expression, signaling pathways and drug response



Future work

- From PANCAN11 to the PANCAN Atlas
- Integrate the mutation hotspots into Regulome Explorer (RE)
- Make code and hotspots available



Acknowledgements

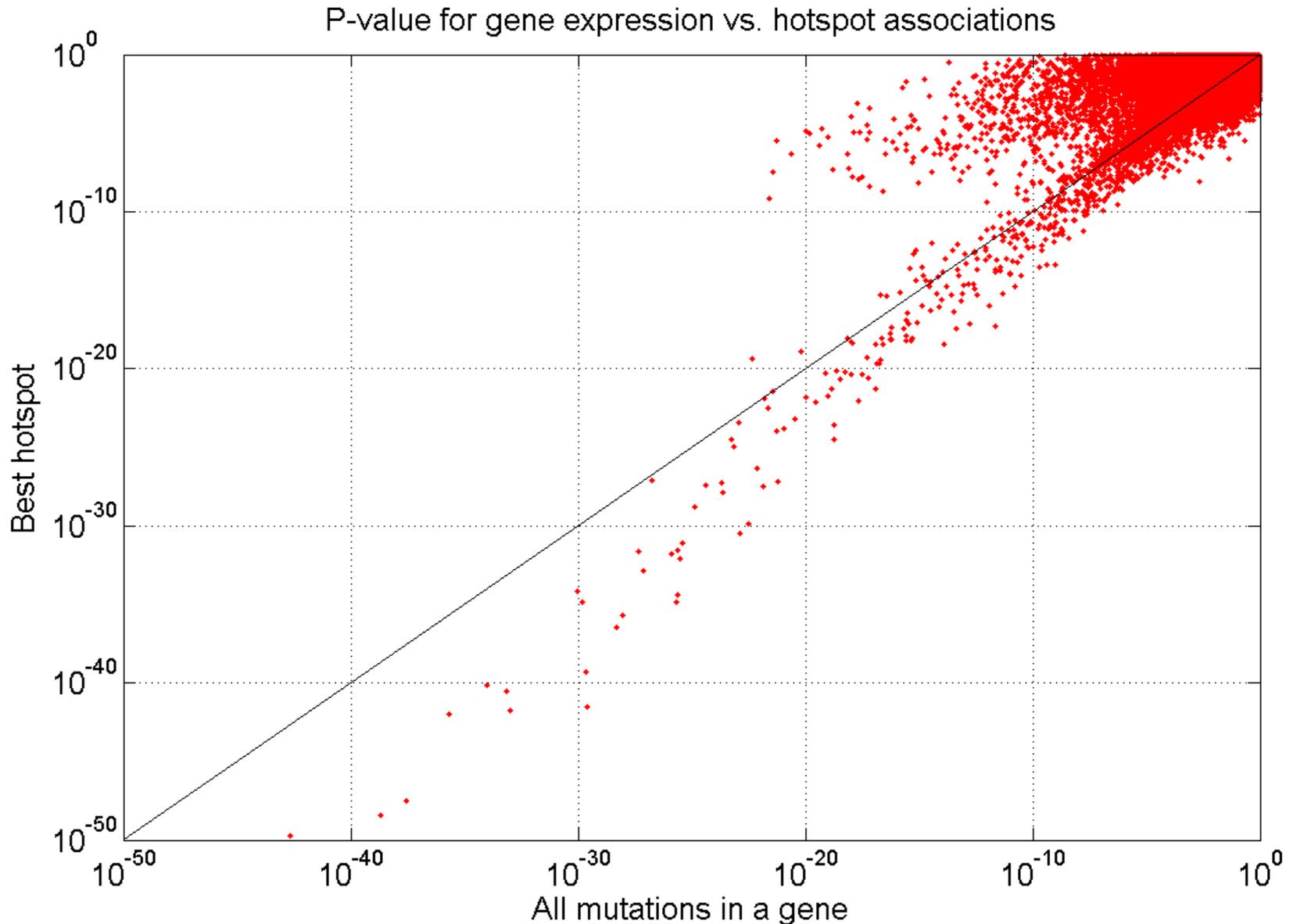
- **William Poole**
- Brady Bernard
- Ilya Shmulevich
- Sheila Reynolds
- Vesteynn Thorsson
- TCGA Genome Data Analysis Center (GDAC)
for Systems Analysis of the Cancer Regulome
(CSACR)



Poster: 57

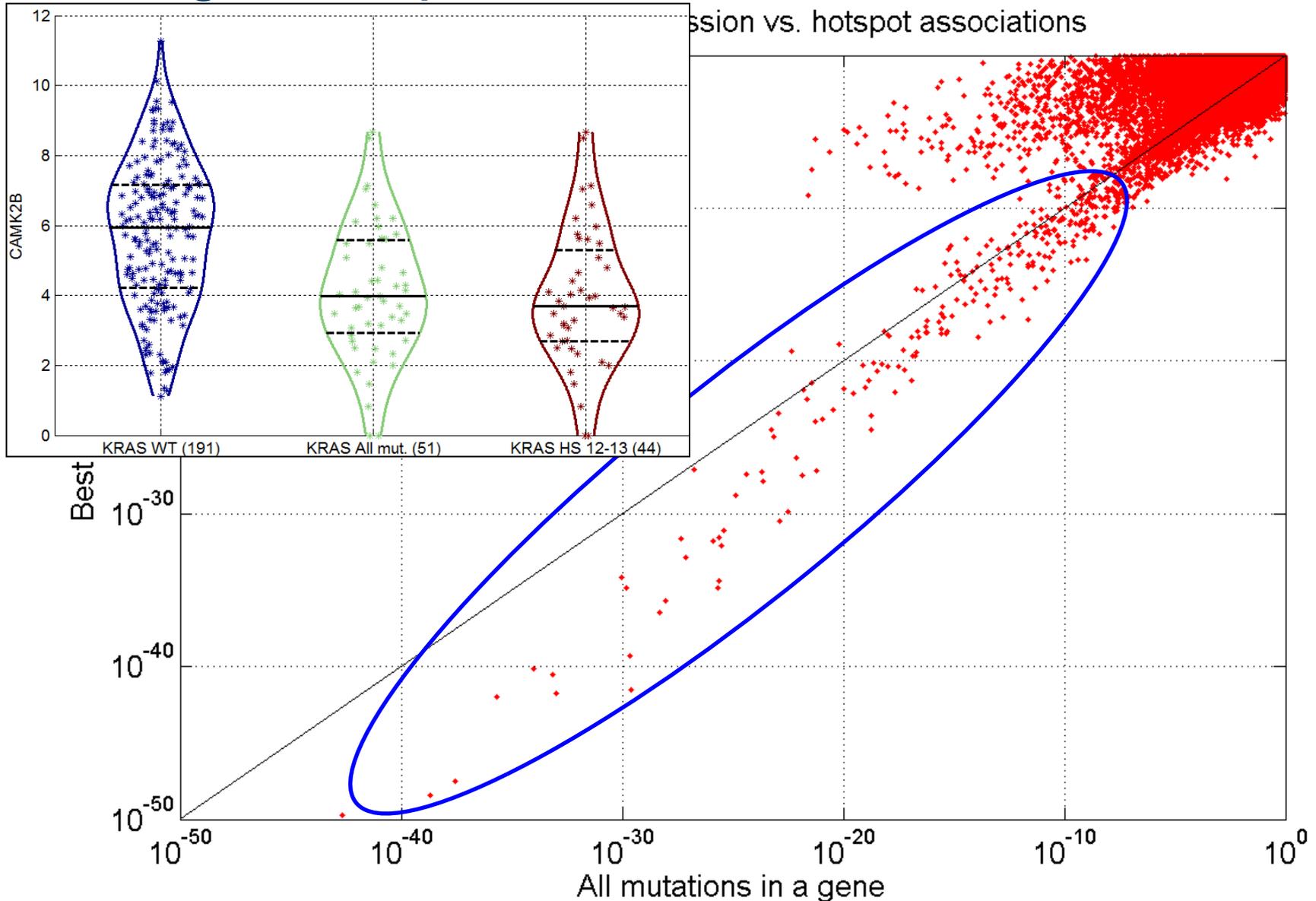


Statistical association between hotspots and gene expression in UCEC



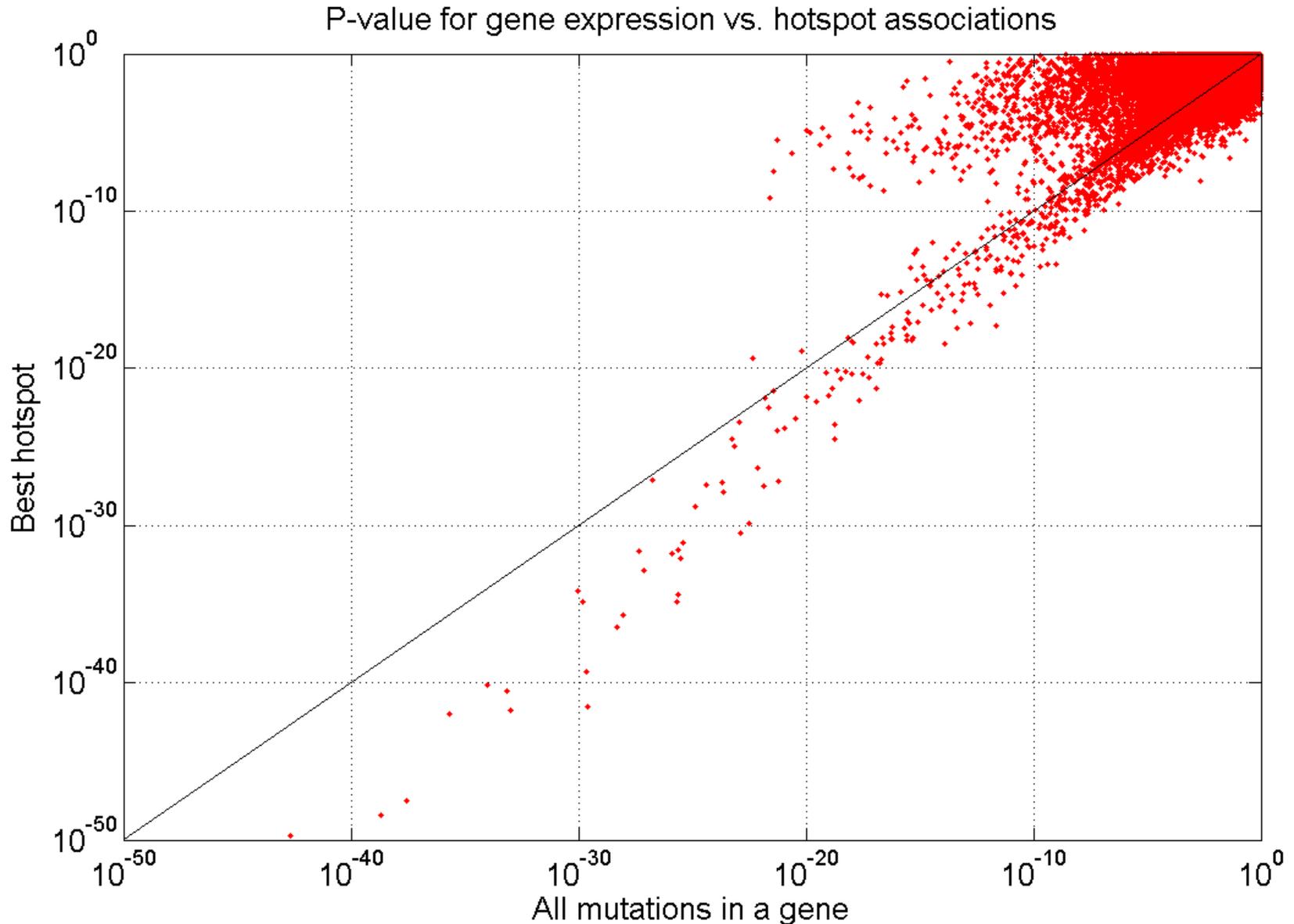


Statistical association between hotspots and gene expression in UCEC



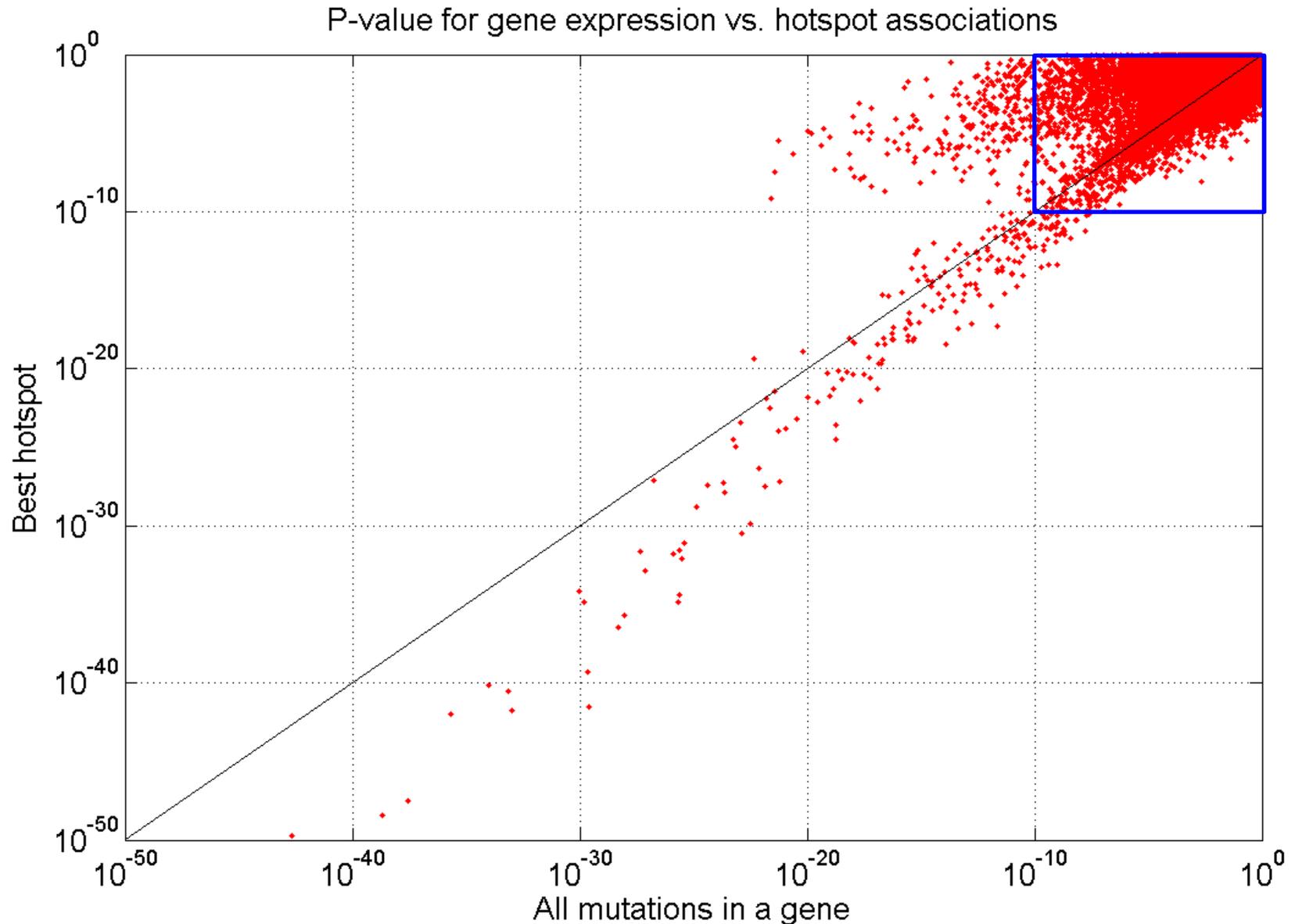


Statistical association between hotspots and gene expression in UCEC





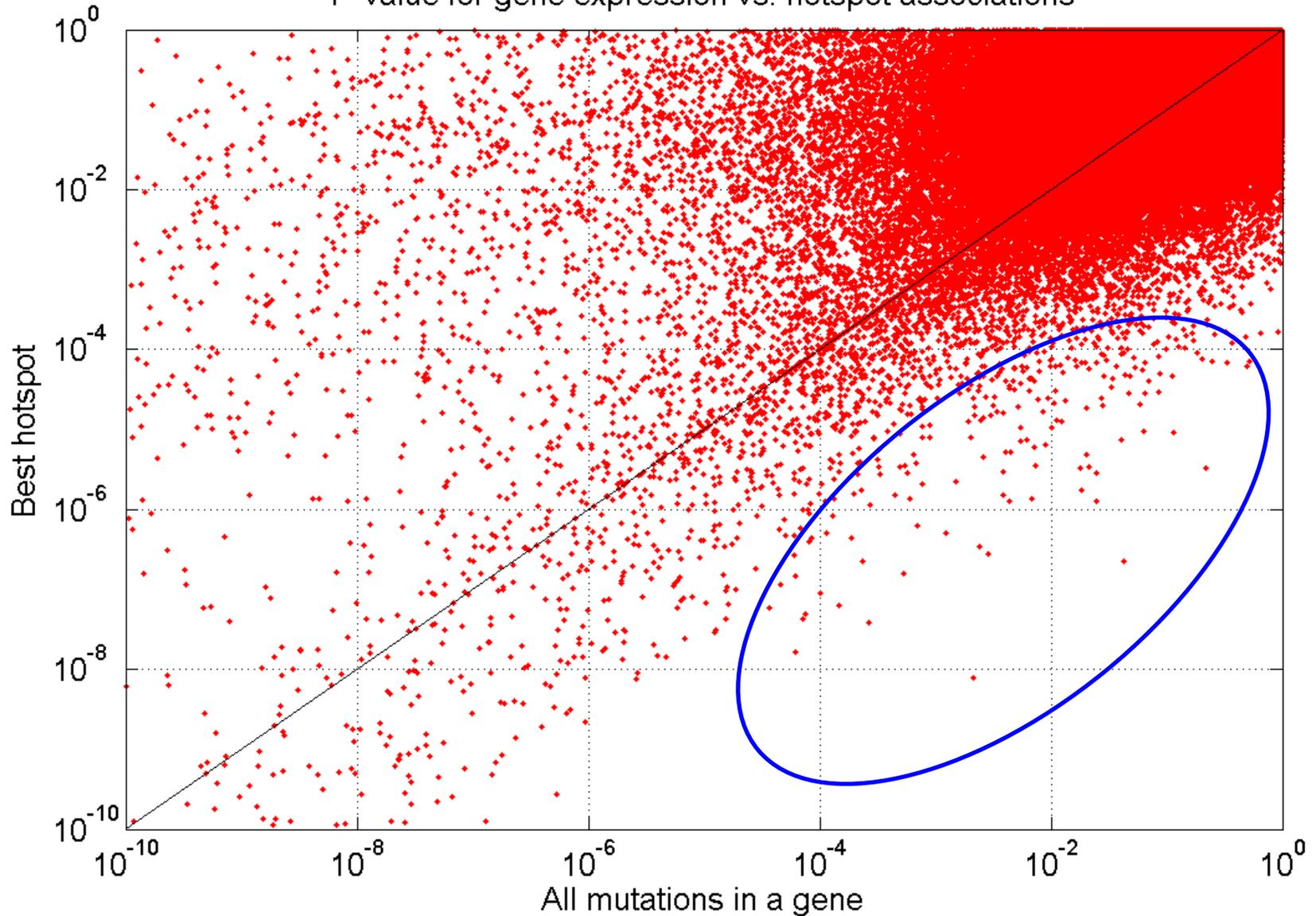
Statistical association between hotspots and gene expression in UCEC





Statistical association between hotspots and gene expression in UCEC

P-value for gene expression vs. hotspot associations





Statistical association between hotspots and gene expression in UCEC

