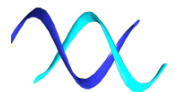


CAGI slide pack guidelines

I agree to my slides being used according to the following terms: _____ *Alin Voskanian and MariceLI Kann* _____

- Authors: please include THIS slide at the beginning of your presentation to indicate that you agree with reuse of your slides in CAGI presentations, according to the terms below, and digitally sign by writing your name above.
- To disseminate the CAGI results, I encourage all CAGI participants to remix these CAGI slides for inclusion in their oral and poster presentations
- CAGI participants may use or publish any content of these slides only in compliance with the CAGI Data Use Agreement
- CAGI participants may not cite or publish any content of these slides except with the written permission of the originating author or in the primary CAGI publications
- CAGI participants must acknowledge the original authors of these slides and CAGI 2013. Include the acknowledgement banner across the bottom
- CAGI participants must include the CAGI credits slide in their presentations
- Slides with a red slash (no remix) may not be included in a presentation unless all attendees are registered CAGI participants who signed the data use agreement
- Slides with a red X over the entire slide may not be used in any circumstances
- Slides with a blue slash (embargoed) may not be used unless the embargo is explicitly lifted
- Slides without slash or X may be tweeted unless flagged with the “No tweet” icon.
- CAGI participants using these slides must explain the “No tweet” icon, which means that the slide content should not be disseminated outside the presenting conference hall. If the presentation venue permits pervasive tweeting, you may not include these slides in your talk
- Note for slide authors: we expect all CAGI participants to abide by these restrictions, and will make best effort to ensure they are followed. However, but bear in mind that we have limited means of enforcing them, and therefore the restrictions cannot be guaranteed.



Evaluations of CHEK2 predictions for Hispanic Women

Alin Voskanian

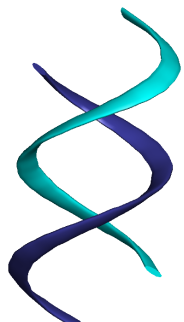
Ph.D Student

Dr. Maricel G. Kann

Associate Professor

Dept. of Biological Sciences

University of Maryland, Baltimore County



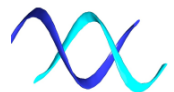
AAG	GCC	AAU	AAC
TTC	CGG	TTA	TTG



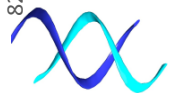
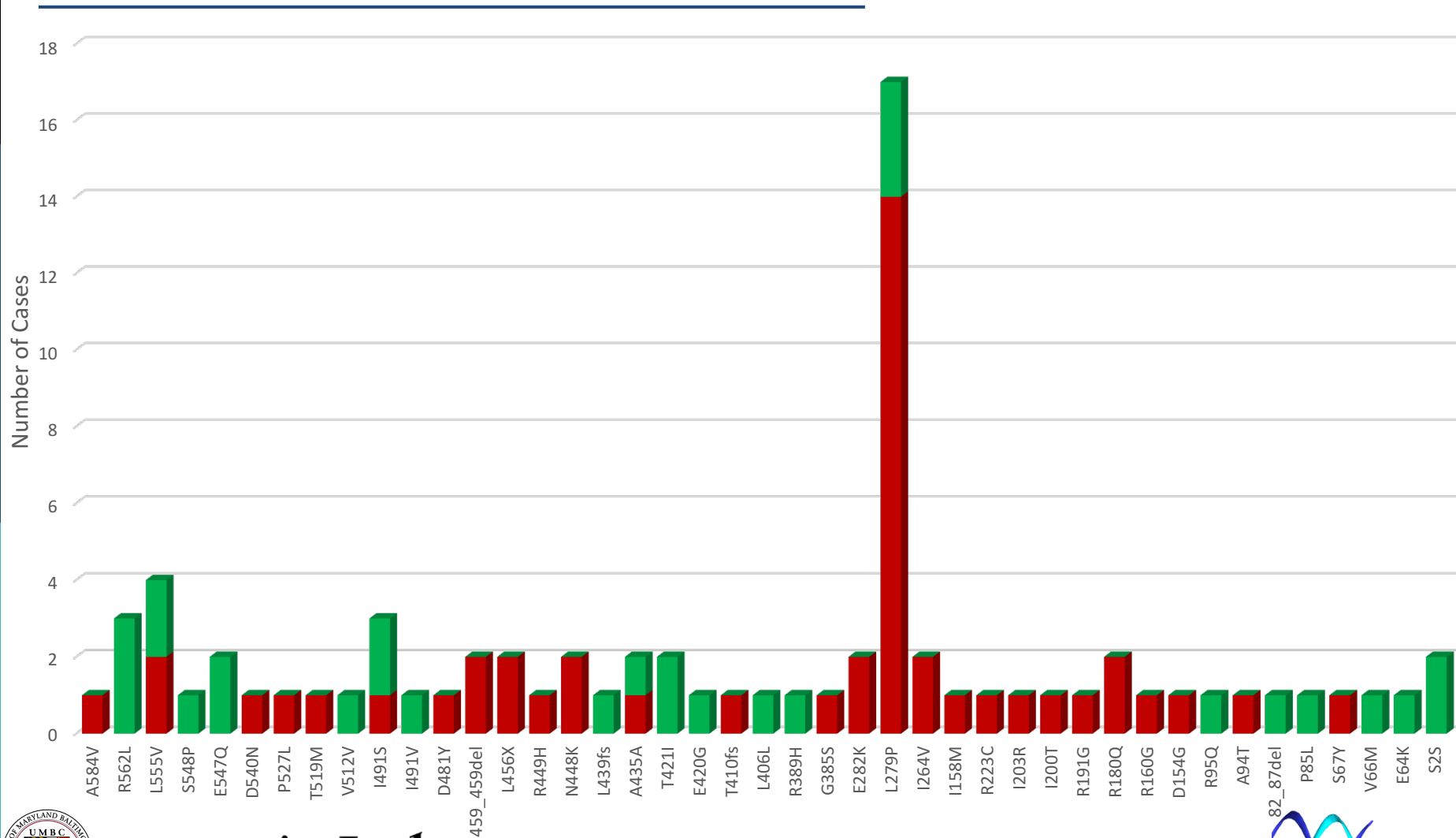
Kann's Lab

Analysis of CHEK2 case-control mutation screening data

- Latina population GWAS
- Based on a dataset of:
 - 2385 control
 - 7342 case
- Total exonic mutations: 43
 - **Exonic non synonymous SNV: 34**
 - Exonic synonymous SNV: 4
 - Exonic stop-gain: 1
 - Exonic non-frameshift deletions: 2
 - Exonic frameshit deletions: 2
- Total UTR =10



Case Control Distribution by Position



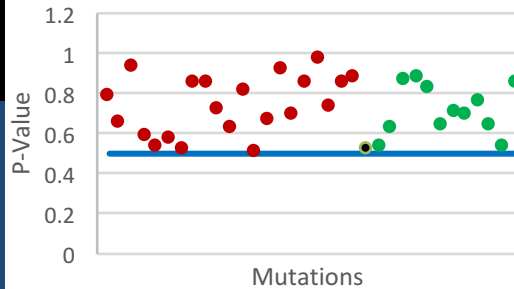
Submission Overview

- 18 total Submissions

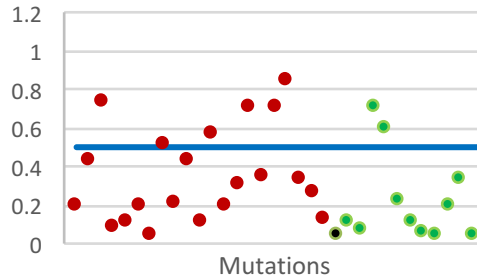
Method		Submissions					
1	1						
2	2						
3	3.1	3.2	3.3	3.4	3.5		
4	4						
5	5.1	5.2					
6	6						
7	7						
8	8.1	8.2	8.3	8.4	8.5	8.6	

P(case) Distributions

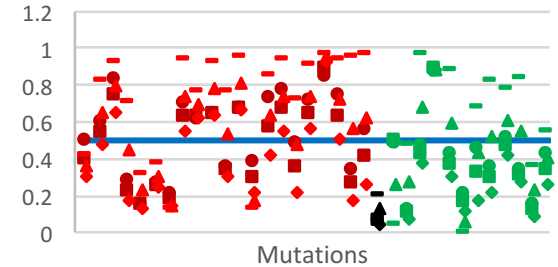
Submission 1 P-value Distribution



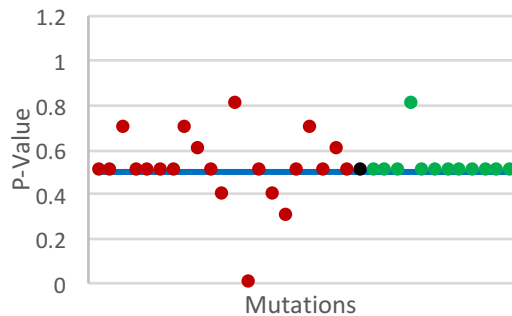
Submission 2 P-value Distribution



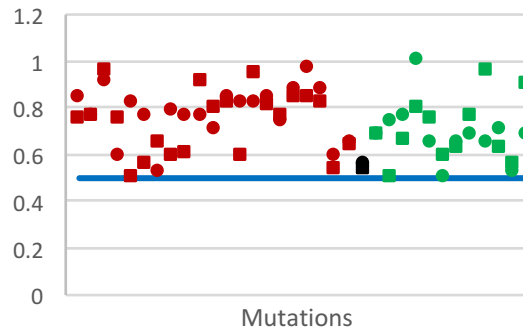
Submission 3 P-value Distribution



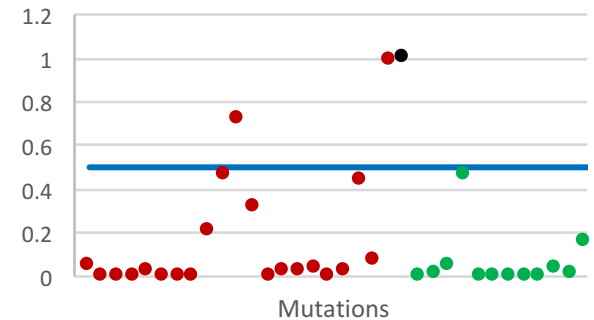
Submission 4 P-value Distribution



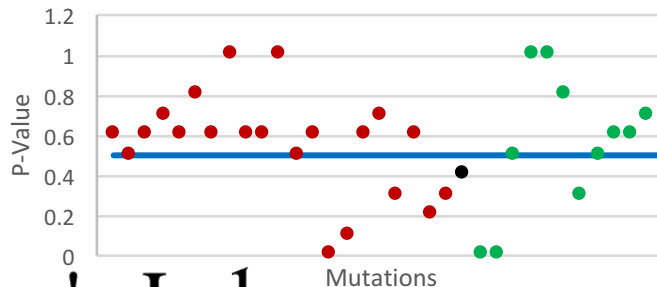
Submission 5 P-value Distribution



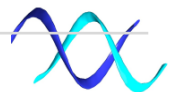
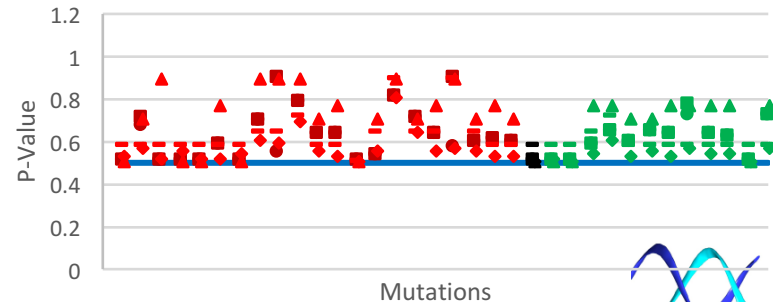
Submission 6 P-value Distribution



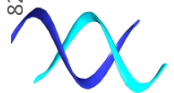
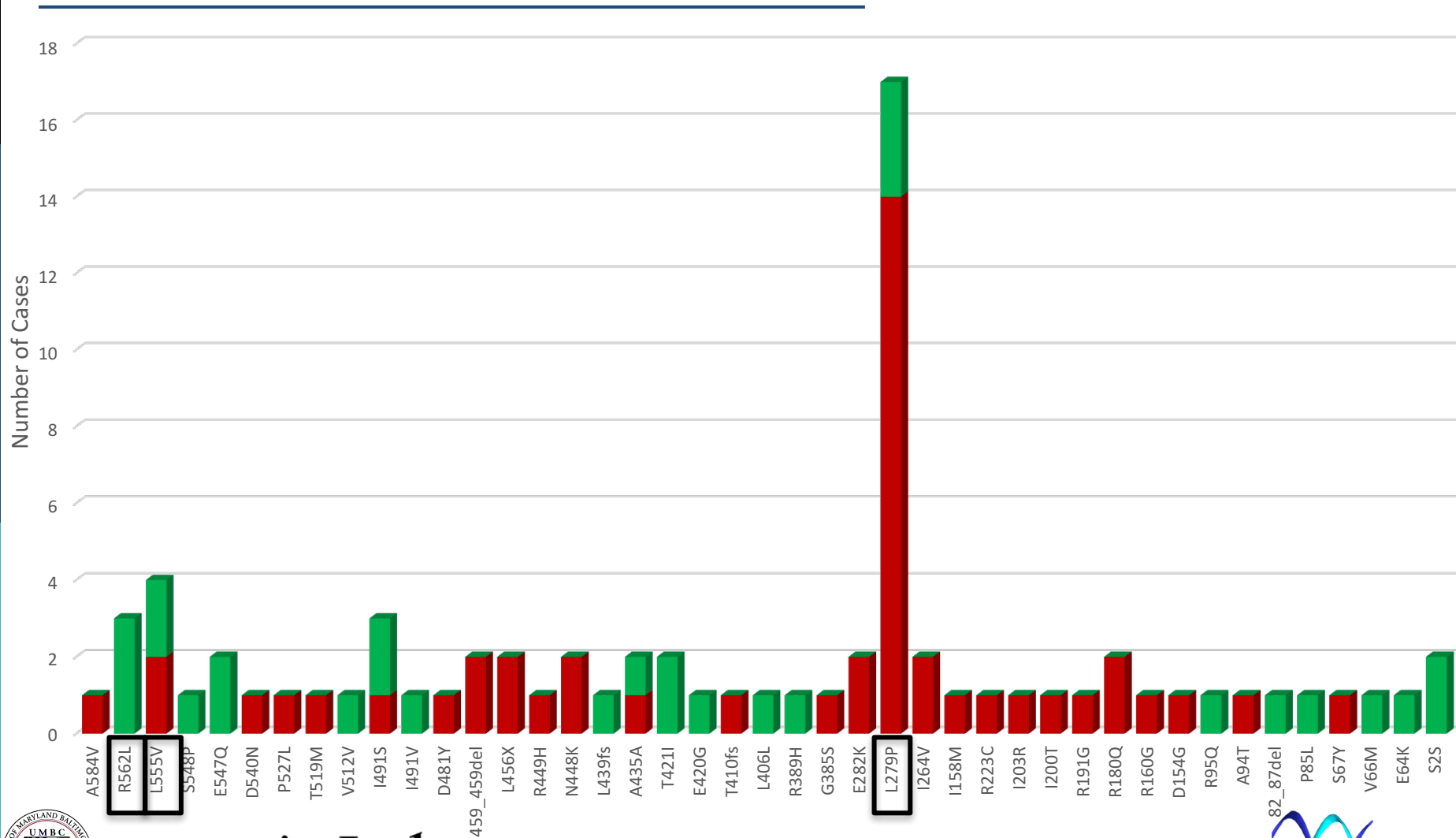
Submission 7 P-value Distribution



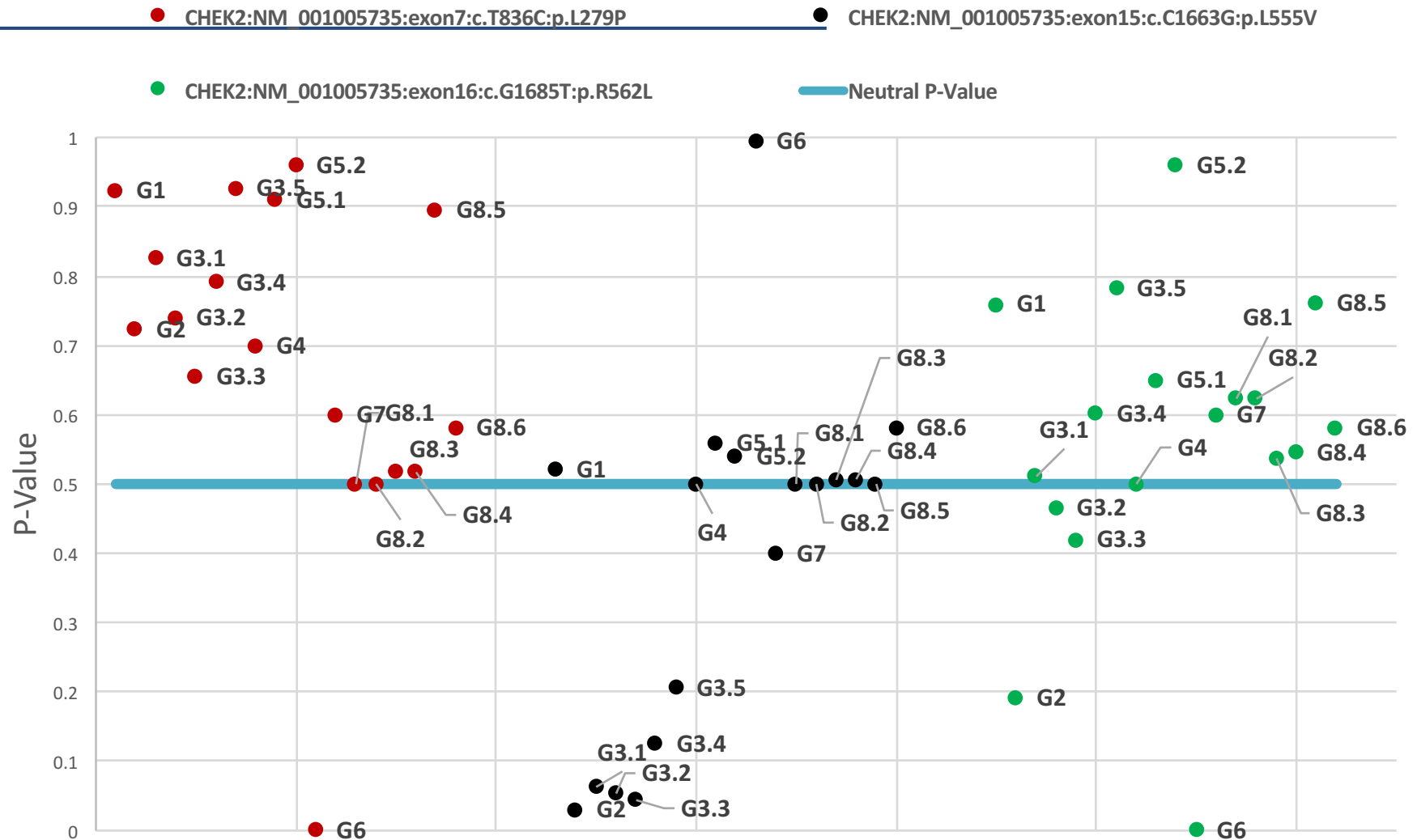
Submission 8 P-value Distribution



Case Control Distribution by Position

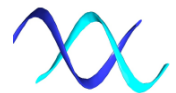


Best Representative Case for Cancerous, Neutral and Protective Mutation



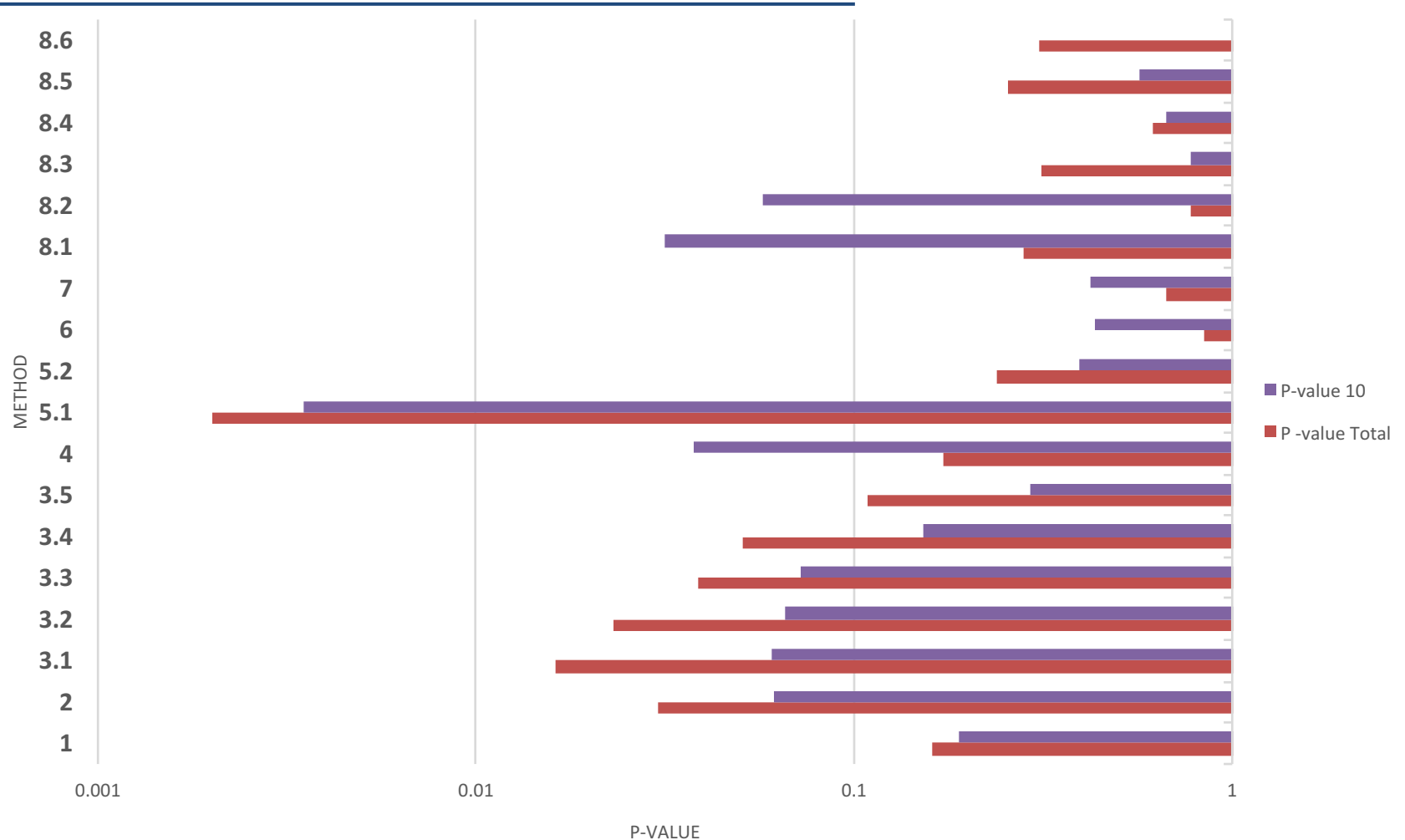
Accuracy Measures

	1	2	3.1	3.2	3.3	3.4	3.5	4	5.1 5.2	6	7	8.1,8.2,8.5	8.3,8.4,8.6
Sensitivity	90.9	27.2	63.6	59.0	45.4	72.7	72.7	Most P values are in neutral range	100	15	72.7	90.4	Most P values are in neutral range
Specificity	41.6	83.3	91.7	91.7	91.7	91.7	50		16.7	92.3	66.7	30	



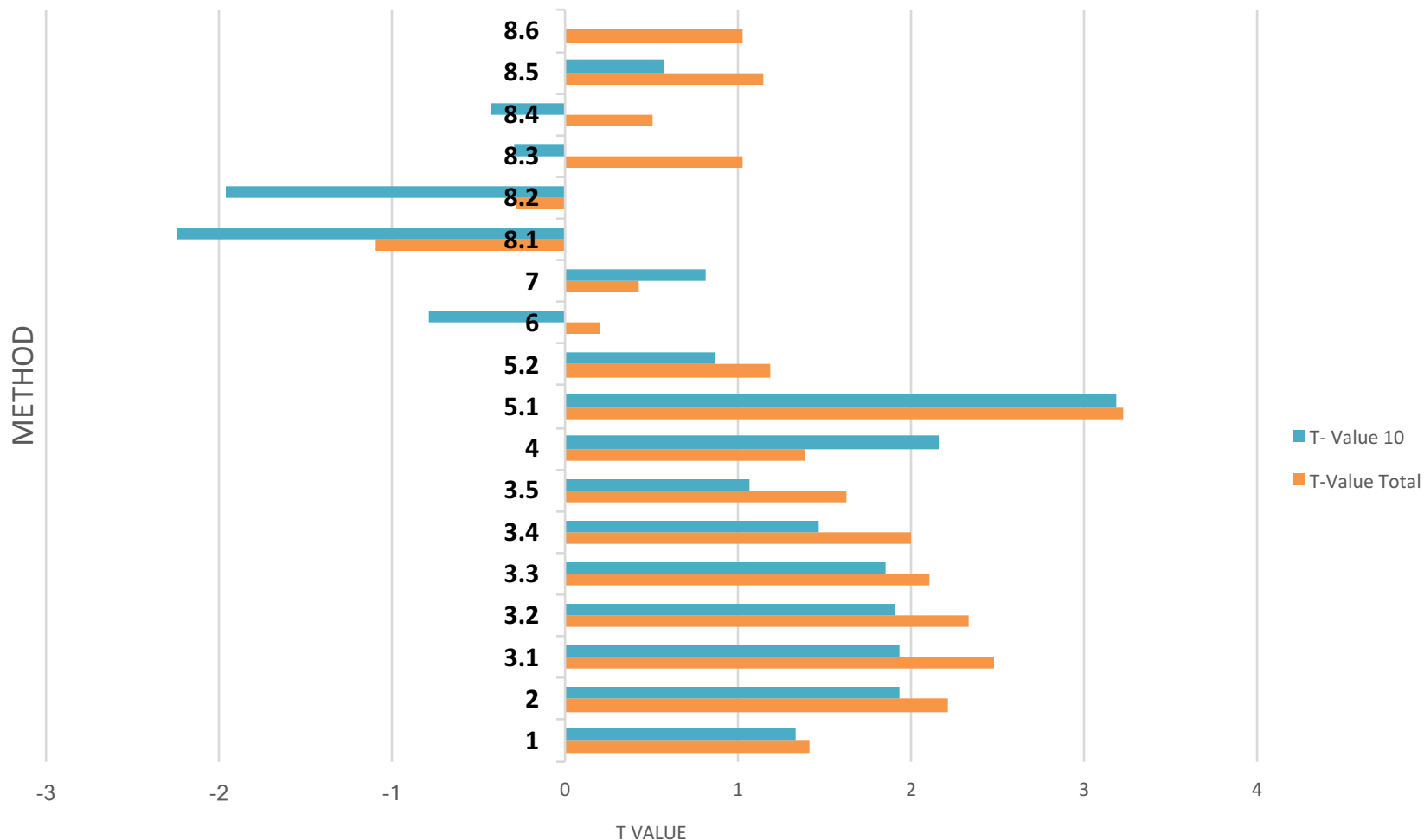
P-values by Method on Log Scale

Group 8.6 is missing values for selected 10 positions, because all their P-values were identical



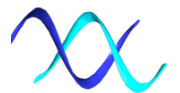
T Value by Method

Group 8.6 is missing values for selected 10 positions, because all their P-values were identical



Conclusions

- Only 1 case present for most positions
- Group 5 submission 1 seems to do best but has many false positives
- Standard deviations were missing for many groups and not really taken into account in this analysis
- Most variants were in cancerous cases thus methods that favored this did better



Acknowledgments

CAGI team.

Gaia Andreoletti

Sean V. Tavtigian

Predictors



Kann Lab



Kann's Lab

