**HGGA, Volume *5***

## Supplemental information

# A systematic assessment of the impact of rare canonical splice site variants on splicing using functional and *in silico* methods

**Rachel Y. Oh, Ali AlMail, David Cheerie, George Guirguis, Huayun Hou, Kyoko E. Yuki, Bushra Haque, Bhooma Thiruvahindrapuram, Christian R. Marshall, Roberto Mendoza-Londono, Adam Shlien, Lianna G. Kyriakopoulou, Susan Walker, James J. Dowling, Michael D. Wilson, and Gregory Costain**

**Figure S1.** CSSV location by RNA-seq frameshift/NMD or non-frameshift/no NMD outcome.

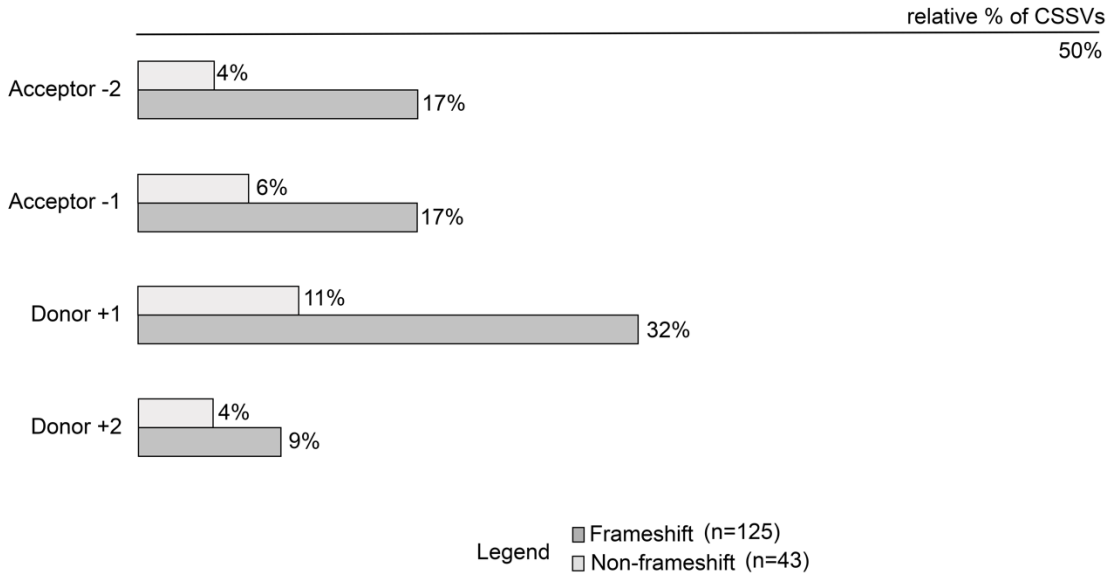

**Figure S2.** Specific nucleotide substitution of CSSVs by RNA-seq frameshift/NMD or non-frameshift/no NMD outcome.

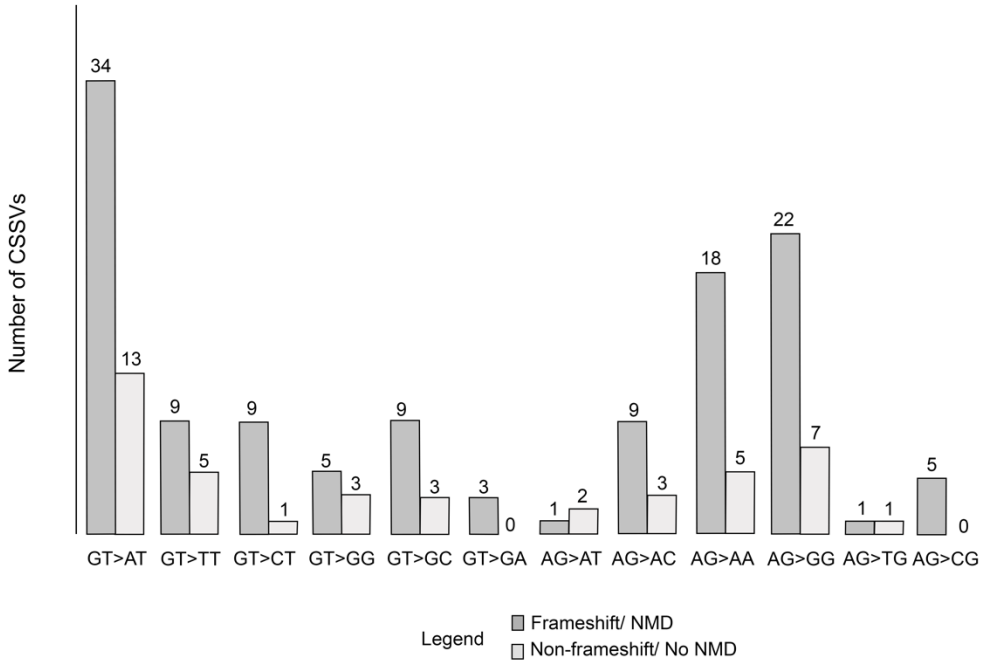

**Figure S3.** CSSVs that showed only WT splicing and comparable read depth to controls by RNA-seq (outcome category B, see main text for details).

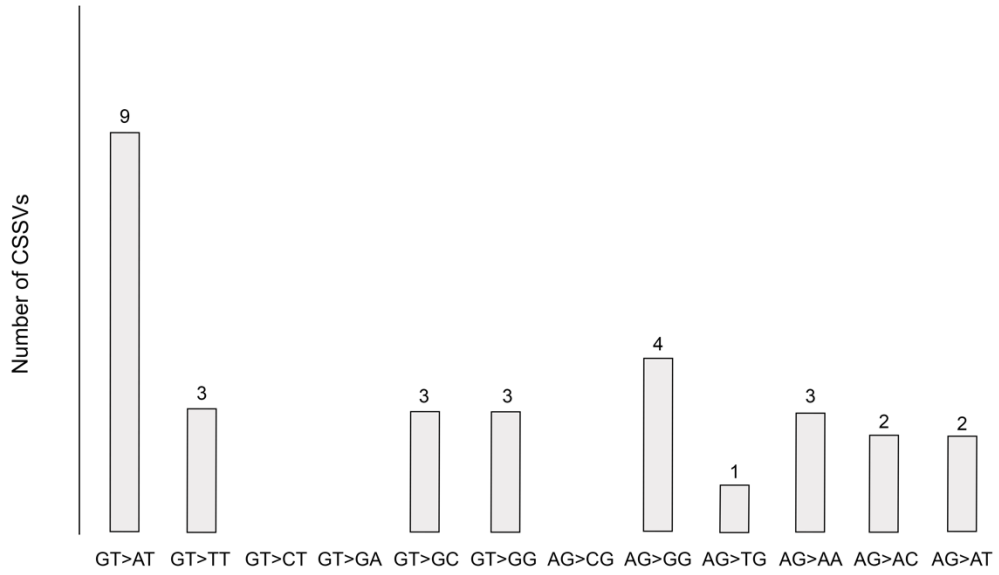

**Figure S4.** Scatterplot of gnomAD v2.1 genome allele frequencies for the 168 CSSVs considered in this study.

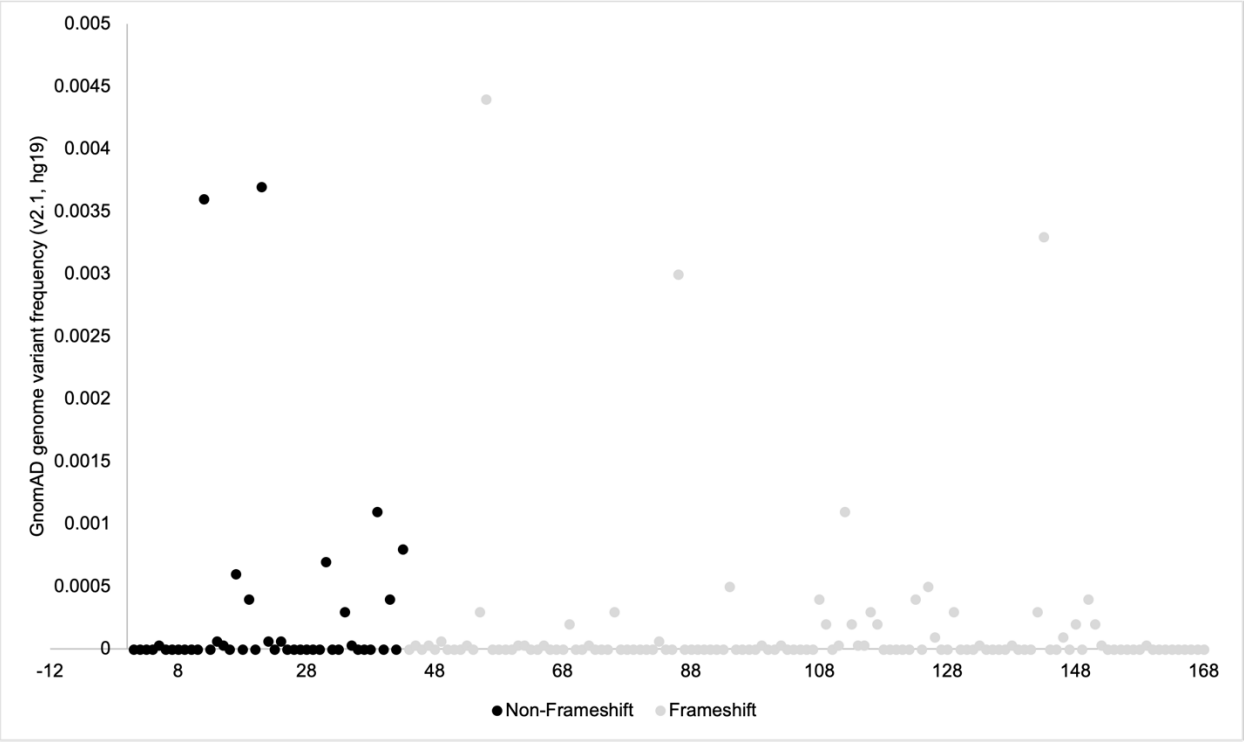


Figure legend: Each column on the x-axis represents a single variant. Within each of the non-frameshift/no NMD variant category (left side of the figure; dark circles) and the frameshift/NMD variant category (right side of the figure; gray circles), variants were ordered alphabetically by gene symbol. There was no significant difference in allele frequencies between the variants that were and were not resulting in a frameshift/NMD per RNA-seq (Mann Whitney U-test p-value =0.5287).

**Table S1.** Demographics of the study cohort.

| | **Individuals (N=112)** |
|---|---|
| **Sex** | **n (%)** |
| **Sex** | |
| Male | 57 (50.9%) |
| Female | 55 (49.1%) |
| **Age at RNA-seq Requisition** | |
| <1 | 12 (10.7%) |
| 1 to 5 | 34 (30.4%) |
| 6 to 10 | 25 (22.3%) |
| 11 to 17 | 22 (19.6%) |
| >19 | 19 (17.0%) |
| **Ancestry** | |
| East Asian | 4 (3.6%) |
| African | 2 (1.8%) |
| Latino/a | 3 (2.7%) |
| Ashkenazi Jewish | 5 (4.5%) |
| European | 55 (49.1%) |
| South Asian | 19 (17.0%) |
| Middle Eastern | 7 (6.3%) |
| Other/ Mixed | 17 (15.2%) |
| **Molecular diagnosis** | |
| Yes, not via CSSV in blood expressed genes | 40 (35.7%) |
| Yes, via CSSV in blood expressed gene | 2 (1.8%) |
| No | 70 (62.5%) |
| **Relationship to others in cohort** | |
| Proband | 93 (83.0%) |
| Parent | 16 (14.3%) |
| Affected Sibling | 3 (2.7%) |

**Table S2.** Please see attached .xls file for an annotated list of the 168 CSSVs studied in this report, including splicing outcomes using RNA-seq and *in silico* predictions.