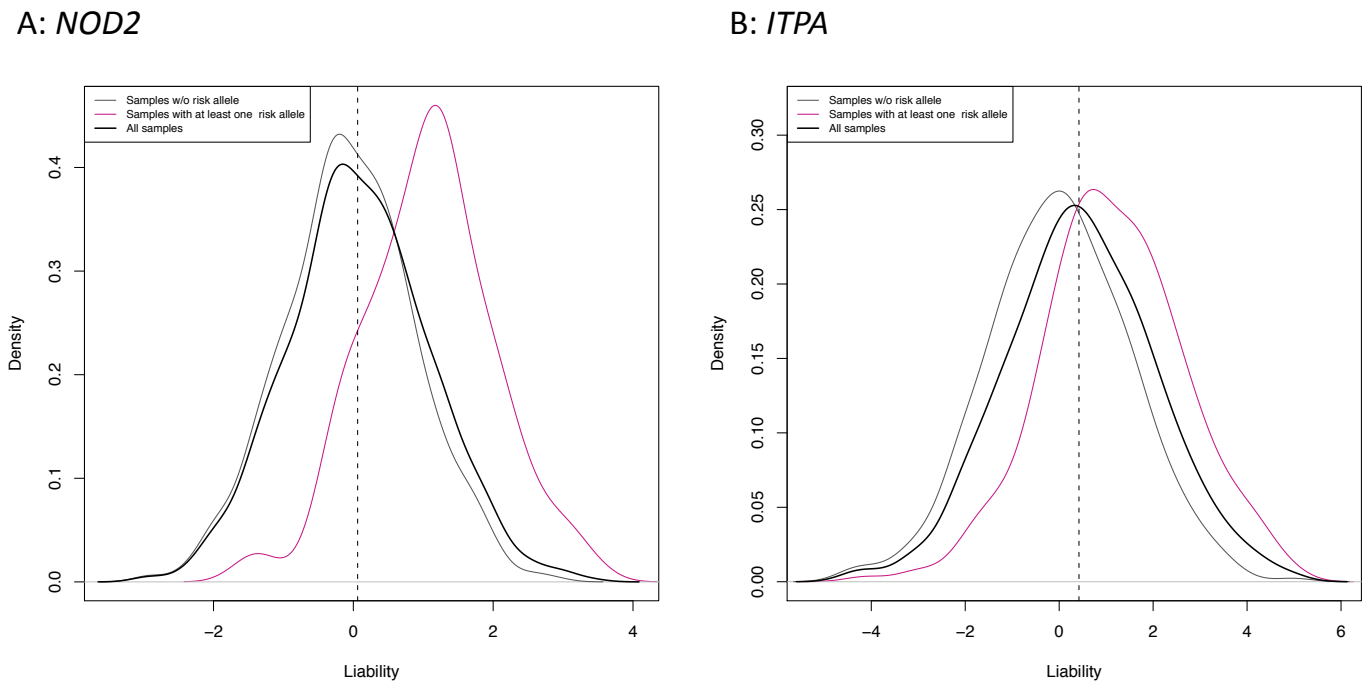

Figure S 3: Distributions of the simulated liability values for exome sequencing samples. Dashed line marks median of the distribution of all samples, which was used to classify samples to cases and controls. (A): *NOD2* data, where 12% of samples were risk allele carriers. (B): *ITPA* data, where 39% of samples were risk allele carriers.