C.1.D.iii. ***Power (y-axis) to detect percent fold changes (x-axis) in transcript levels for 250 asthmatics in the discovery stage***. Here we assumed an additive model in the number of variant alleles for the expression levels, and considered a two-sided hypothesis test for the slope in the regression model under false discovery rate control (5% **left,** 10% **right**), for 4 assumed MAFs (5%, 10%, 20%, 50%) under HWE (**Fig. 5**). The proportion of true eQTLs was assumed to be 10% (**solid line**) and 1% (**dashed line**) to reflect cis- and trans-eQTLs, respectively. Assumed within-group variability for the power calculations was delineated from a previous array-based gene expression study of 396 subjects of the same population, using the median between sample variability observed. Since no technical replicates were generated, the variability observed in the array-based gene expression data represents biological variability plus technical variability. In our experience, biological variability tends to be larger than technical variability for both array-based gene expression and RNA-seq. In addition, while the range of measured transcript levels can be larger in RNA-seq due to a commonly observed "ceiling effect" in fluorescence based gene expression data, the overall technical variability is usually lower for RNA-seq data[45, 117, 118]. Thus, we believe the above power calculations are accurate or even slightly conservative, and show that even modest fold changes in transcription levels can be reliably detected.

Fig. 5. Power calculations SA1: eQTL on 250 asthmatics.