



Figure S2. Gene expression in CD4⁺ T cells at diagnosis of clinical T1D in subjects in the placebo and nasal insulin treatment arms of INIT II. Multidimensional scaling (MDS) plots of RNA-seq log2 counts per million (log-CPM) values in the discovery cohort, with samples coloured by time-point and treatment arm in INIT II. The log-CPM values were corrected for age and batch variables. (B) Heat map of genes differentially expressed in CD4⁺ T cells between progressors and non-progressors over the period of observation. Samples are colour coded by progression status, time-point and treatment group. The Z-score underneath the heatmap represents the number of standard deviation units a specific datapoint is from the mean (mean is derived from the entire set of genes expressed in CD4⁺ T cells). Samples and genes were arranged according to “ward.D2” clustering method with the function coolmap in the limma package. Genes with an asterisk are cytotoxicity-related genes while those bolded are EOMES target genes.