It seems like an easy thing to do the single snp wise simple linear regression indicated in the video. However, can more be done? E.g multi-dimensional regression over sets of snps or classification or association rule mining. I think the analysis can be reduce to searching for patterns in the described matrix below.

In the Gatti data, there are ~400 transactions-ROWS (individuals) and ~7500 items_COLUMNS (alleles/snp). Population is also partitioned by sex. Each allele (SNP) is one of six types one for each distinct pair of {A,C,G,T}. The each type has one of three values (1=XX, 2=XY, 3=YY. The associated two phenotypes, white blood counts (WBC) and neutrophils (NEUT) are real values. I haven't checked what the phenotype distribution looks like, but it could be possible find a binary threshold for clearly high or low levels. Then label the threshold 0=above threshold, 10=below. If one adds this column to each column in the snp matrix, we have a 1,2,3, 11, 12, 13 valued matrix. Then set all two digit values to 0. This give a 0,1,2,3 matrix that can be analyzed for non-zero entries.