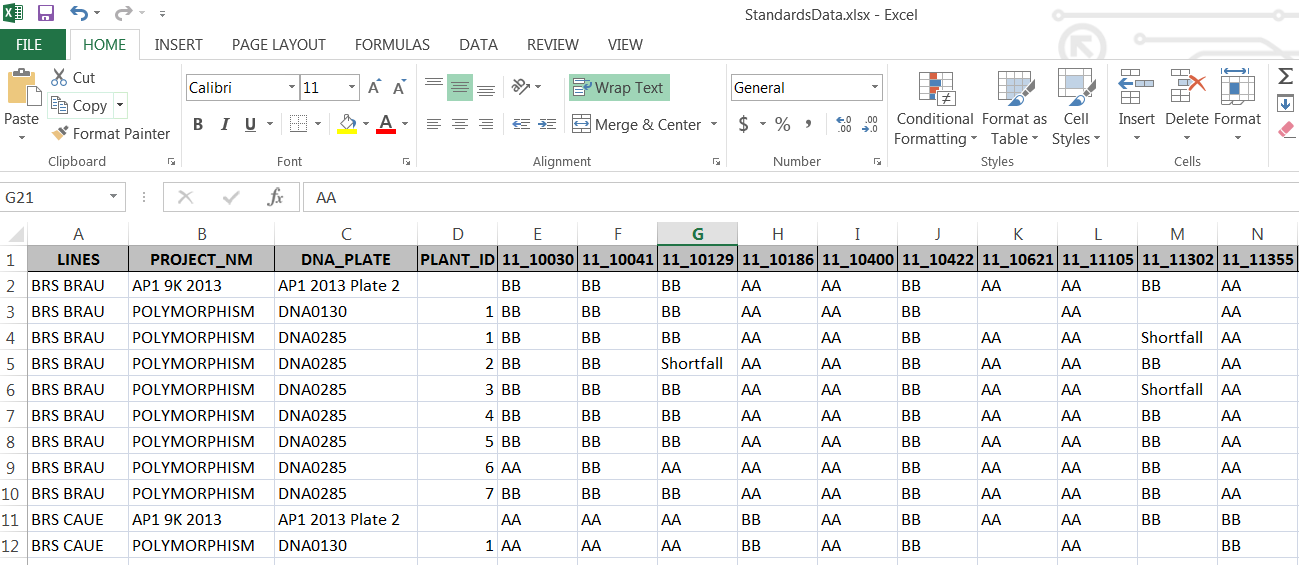
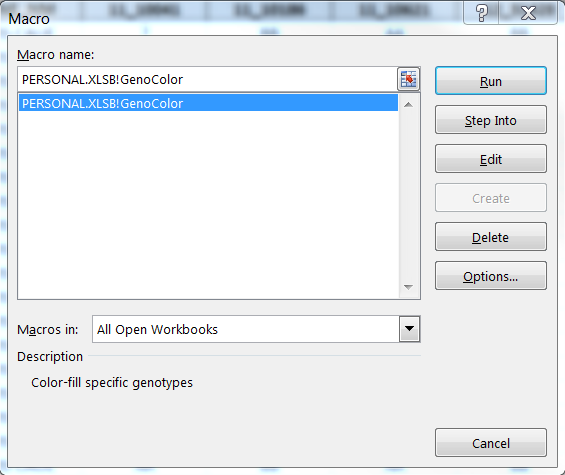
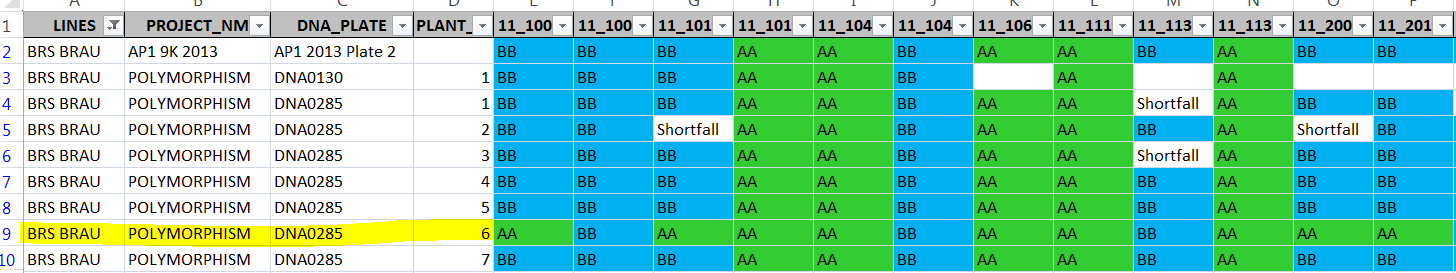
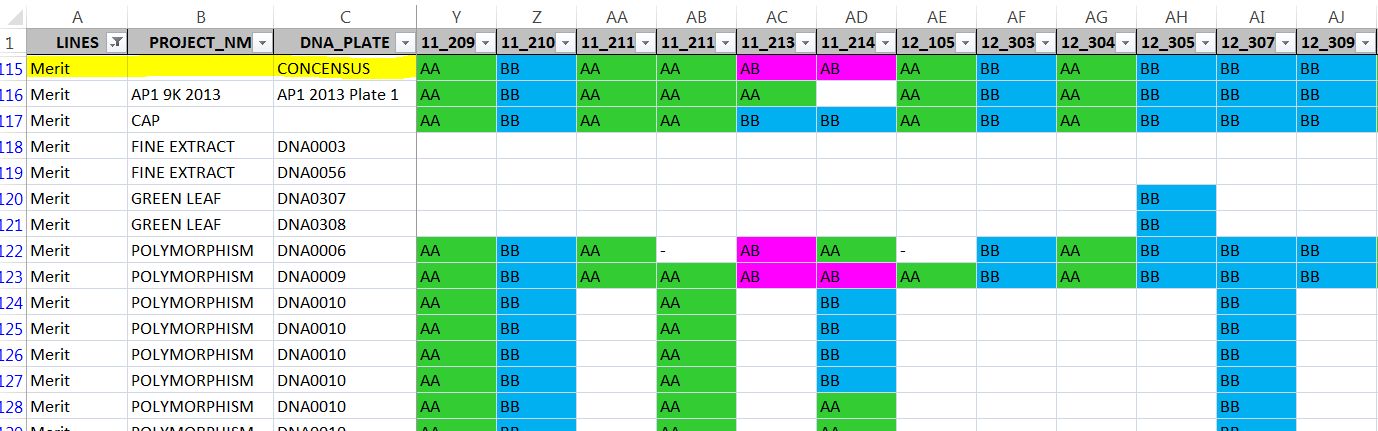
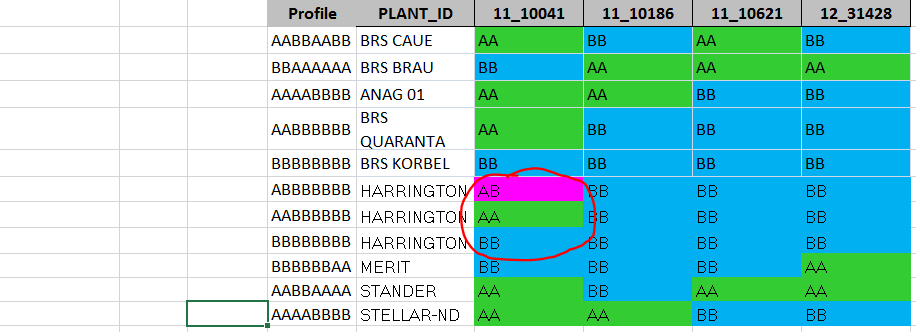
**SOP for creation of Line Profiles key for Purity Analysis**

1. Export data from DNA genotyping database for all lines known to be grown in region as well as plate standards (See StandardsData.xlsx). Data will contain some missing values as well as Shortfall or Over values which are considered FAILS and the data is not used. 
2. Open TemplatesMacro.xlsm file, then in StandardsData.xlsx run the TemplatesMacro.xlsm!GenoColor macro to assign genotype colors to cells with conditional formatting.
3. Filter by line name and visually scan each marker to determine uniform calls for each line. If an entry differs from the majority of calls across multiple markers, remove or ignore this entry as it is likely a contaminant. 
4. Create a new row for each Line variety with the term CONCENSUS in the DNA\_PLATE column. If all the genotypes are consistent across entries(rows) for a line for each marker, copy these values into the CONCENSUS row. If more than 10% of calls are divergent between the entries (ie, some are AA and BB and even AB) then enter an AB into the cell for this marker in the CONCENSUS row. 
5. Data can then be filtered to just the CONCENSUS row for each LINE(Variety). In order to accommodate all possible genotye patterns for a given set of markers. The AB genotypes must be represented as all possible variations. This can become challenging if there is variation across multiple markers. 
6. Select only CONCENSUS rows and matching Markers for use in Purity Analysis.

Suggested Improvements:

It would be good to have this as an automated process that generates not only the final collection of CONCENSUS rows but also a report providing information for each LINE/Marker such as:

# of datapoints available  
Ratio of A:B  
Variance  
Whether any data was removed due to variance of genotypes at multiple marker locations (see step3)