

GDdom Manual

1. Introduction

The usefulness of a given genetic marker is determined by its level of polymorphism. Gene diversity (GD), also called polymorphism information content, is a commonly used measure of molecular marker polymorphism for dominant markers such as AFLP, RAPD and multilocus SSRs. We developed a free online computer program, GDdom, which provides easy, quick and accurate calculation of dominant marker GD with the commonly used formula of Roldan-Ruiz et al. (2000): $GD_i = 2f_i (1 - f_i)$, where GD_i is the gene diversity of marker 'i', f_i is the frequency of the amplified allele (band presence), and $1 - f_i$ is the frequency of the null allele. According to this formula, GD for a locus can vary from 0 to 0.5. Results are presented in tabular form for quick interpretation.

2. Data preparing

Open Excel file, rows represent samples and columns represent marker alleles. In first column, write marker number followed by dash (-) and allele size, for example marker1-198. Continue adding columns for all alleles for that marker. Add columns for remaining markers and each of their alleles (e.g. marker2-165, marker2-180) (Figure 1). Score alleles for each dominant marker and individual sample as presence (1), absence (0) and 9 for missing data.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	marker1-198	marker1-253	marker1-260	marker1-270	marker2-165	marker2-180	marker2-192	marker2-205	marker2-214	marker2-224	marker3-220	marker3-235	marker3-242
2	1	1	1	0	1	0	1	0	1	1	1	0	1
3	0	1	1	0	0	1	0	0	1	1	1	1	0
4	1	1	1	9	0	1	0	0	1	1	1	1	0
5	1	1	1	0	0	1	0	0	1	1	1	0	1
6	1	1	0	9	9	9	0	0	1	1	1	0	1
7	1	1	1	1	0	1	0	0	1	1	1	1	1
8	1	1	1	1	1	0	1	0	1	0	1	1	0
9	1	1	0	0	1	1	1	1	1	0	1	0	1
10	1	1	1	0	0	9	0	1	1	9	1	9	1
11	1	1	0	0	1	1	1	0	1	1	1	1	0
12	1	0	1	1	0	1	1	0	1	1	1	0	1
13	0	1	1	1	0	1	1	0	1	1	1	0	0
14	1	1	0	0	1	1	1	1	1	1	1	0	1
15	1	1	1	1	1	1	0	1	1	0	1	0	0
16	1	0	1	1	0	1	1	1	1	1	1	0	1
17	1	1	1	1	0	0	1	0	1	1	1	0	1
18	1	1	1	0	1	0	1	0	1	1	1	0	1
19	1	1	0	0	0	1	0	0	1	0	0	1	1
20	1	1	0	0	0	1	0	0	1	1	1	1	0
21	1	1	0	1	1	0	1	0	1	1	0	1	1
22	1	1	1	1	0	0	1	0	1	1	0	0	1

Figure 1. Excel file showing marker name and allele size in columns. Rows from 2 to 22 represent marker scoring for 21 genotypes which were scored as presence (1), absence (0) or missing data (9).

3. Saving the file

From file menu, the data file should be saved as "CSV (Comma delimited)" or "CSV (MS-DOS)" as illustrated in Figure 2.

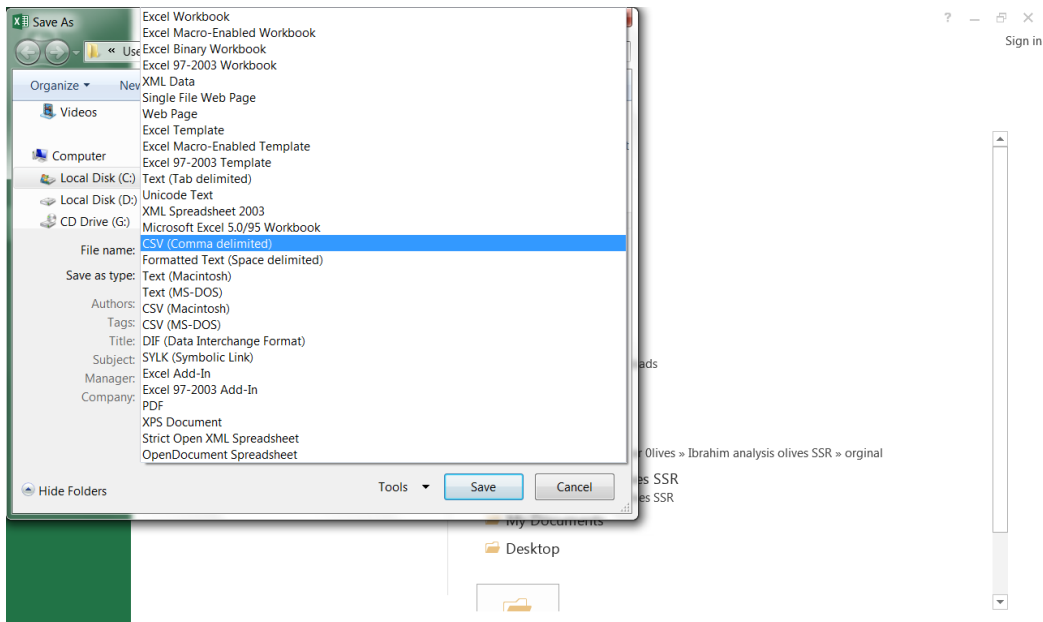


Figure 2. Saving Excel data file as CSV format.

4. Data analysis

Open GDdom program using the link (<http://plantmolgen.iyte.edu.tr/GDdom/>), upload CSV data file and submit. The program summarizes all data in tabular form. The first output table describes the missing values for all samples (Figure 3). The second table shows detailed results for the number of alleles, number of polymorphic alleles, average GD value, standard deviation and standard error (Figure 3). The third table gives summary statistics for the average GD value for each marker (Figure 4). The final table includes maximum, minimum and average GD over all markers (Figure 4).

The screenshot shows the GDdom web application interface. At the top, there is a navigation bar with the URL plantmolgen.iyte.edu.tr/GDdom/index.php. Below the navigation bar, there is a message: "- Choose your file by clicking the button 'Choose file' then click 'Submit'." Below this message, there is an "Upload CSV" section with a "Choose File" button and the text "No file chosen". A "Submit" button is located below the "Upload CSV" section. Below the "Submit" button, the "File Name" is displayed as "GDdom sample.csv".

Below the "File Name" section, there are two tables. The first table is titled "Table1: Missing Values' table: assuming that the (9) value in the input table is a missing value." and contains the following data:

Row Number	Missing Data	% Missing Data
3	1	1.9231
5	3	5.7692
9	3	5.7692

The second table is titled "Table 2: Frequency and GD statistics for marker alleles." and contains the following data:

Marker	Number of 1s	Number of 0s	Number of 9s	Total no. of samples	Samples without 9	Fp	Fa	GD	Number of polymorphic fragments/total no. fragments (%)	Average GD value	Standard deviation	Standard error
marker1-198	19	2	0	21	21	0.9048	0.0952	0.1723	4/4 (100)	0.3219	0.1598	0.0754
marker1-253	19	2	0	21	21	0.9048	0.0952	0.1723				
marker1-260	14	7	0	21	21	0.6667	0.3333	0.4444				
marker1-270	9	10	2	21	19	0.4737	0.5263	0.4986				
marker2-165	8	12	1	21	20	0.4	0.6	0.48	5/6 (83.3)	0.3475	0.1667	0.0681
marker2-180	13	6	2	21	19	0.6842	0.3158	0.4321				
marker2-192	12	9	0	21	21	0.5714	0.4286	0.4898				

Figure 3. Output from GDdom program. Tables 1 and 2 describe the missing values for all samples and detailed results for the number of alleles, number of polymorphic alleles, average GD value, standard deviation and standard error.

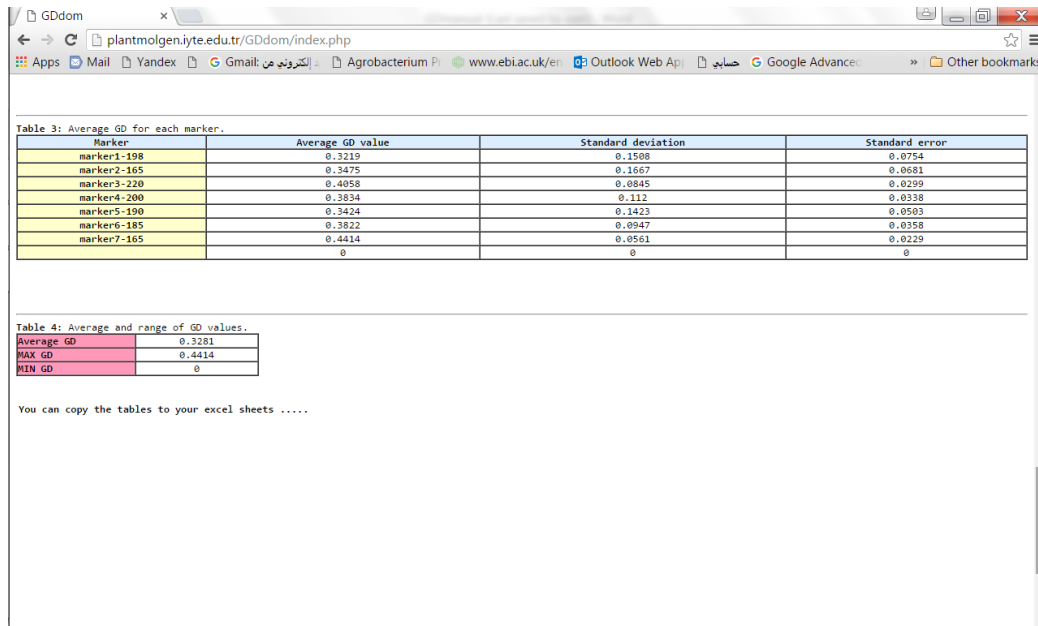


Figure 4. Summary statistics for the average GD value for each marker in Table 3. Table 4 includes maximum, minimum and average GD over all markers.