Edinburgh Genome Foundry

Sabotage report

This report highlights the parts which significantly impair the success of assemblies they belong to.

Significant parts

Part	p-value	Effect
part_C	0.004	+42%
part_S	0.059	+34%

0.72
F1 SCORE

These parts have been selected because they tend to appear in assemblies that fail more, and be absent of assemblies that succeed.

The p-value (from an ANOVA test) indicates how statistically significant this pattern is for each selected part. Smaller p-values mean greater significance.

The "Effect" indicates how impactful the part is, i.e. how much the part belonging to a assembly increases this assembly's chances of failure, in average.

The F1 score (between 0 and 1, 1 being best) score indicates how accurately one could predict the success or failure of a assembly, simply knowing which of these parts are in the assembly. A low score (below 0.5) indicates that this report hasn't found any significant effect of a assembly's parts on its success rate.

Were these parts appear

The next table shows the different assemblies, from the less to the most successful in average, and which failure-associated parts each assembly contains.

The "Mystery" column, when featuring two stars or more, indicates assemblies with a success rate that is not well explained by the presence of these parts. This means either that the assembly was particularly successful while containing failure-associated parts, or on the contrary that the assembly was unsuccessful despite containing no such part. In both cases, this indicates that the assembly's success was mainly driven by external factors.

Assembly	Failure Rate (%)	part_C	part_S	Mystery
assembly_11	100		✔	*
assembly_1	87	✔		
assembly_5	87	✔		
assembly_7	62		✔	
assembly_9	62	✔		
assembly_12	62			*
assembly_3	50		✔	
assembly_4	37			
assembly_10	37			
assembly_6	25			
assembly_8	25			
assembly_2	12			