

The Genetic Landscape of the Cell

[Supplementary Online Material \(SOM\)](#)

Supplementary data files

General information

The SGA genetic interaction dataset is composed of 1711 queries crossed to 3885 array strains. Of 1711 queries, 1377 are deletion mutants of non-essential genes and 334 are essential gene alleles (214 temperature-sensitive and 120 DAmP alleles). TS and DAmP queries are indicated by "_tsq" and "_damp" suffixes, respectively. For a subset of essential genes, several different TS alleles were tested, as indicated by a unique number following the "_tsq" suffix. The set of array strains originally contained 4293 non-essential deletion mutants, but 408 mutants were removed for quality control reasons or due to incompatibility with SGA technology. Approximately 645,000 individual gene pairs were filtered from the 1711 x 3885 genetic interaction matrix for the same reasons. The resulting dataset contains ~6 million double mutants.

Supplementary data file S1. Raw data file

[sgadata_costanzo2009_rawdata_101120.txt.gz](#) - 160MB

Update (Jan. 28, 2010). Inconsistencies between single and double mutant fitnesses reported for 94 query strains and 2 array strains were corrected. The raw data file was updated with the correct single and double mutant fitness values for these strains. None of the SGA scores reported in the original data file were affected by these changes.

Update (Nov. 20, 2010). The *clc1*Δ screen was removed because an allele discrepancy was noticed.

The file contains the complete SGA genetic interaction dataset in a tab-delimited format with 13 columns:

- Query ORF
- Query gene name
- Array ORF
- Array gene name
- Genetic interaction score (ϵ)
- Standard deviation
- p-value

- Query single mutant fitness (SMF)
- Query SMF standard deviation
- Array SMF
- Array SMF standard deviation
- Double mutant fitness
- Double mutant fitness standard deviation

Supplementary data file S2. Raw data matrix

[sgadata_costanzo2009_rawdata_matrix_101120.txt.gz](#) - 19MB

Update (Nov. 20, 2010). The *clc1Δ* screen was removed because an allele discrepancy was noticed.

The file contains the complete SGA genetic interaction matrix in [Java Treeview](#) format.

Supplementary data file S3. Dataset at lenient cutoff

[sgadata_costanzo2009_lenientCutoff_101120.txt.gz](#) - 12MB

Update (Nov. 20, 2010). The *clc1Δ* screen was removed because an allele discrepancy was noticed.

The file contains the SGA genetic interaction dataset with a lenient cutoff applied (p-value < 0.05). Reciprocal interactions (AB vs BA) were processed as follows: if AB and BA show opposite interaction signs (AB is positive and BA is negative, or viceversa), both pairs were removed; if AB and BA show the same interaction sign (both positive or both negative), the interaction with the lowest p-value was retained and both pairs are reported with that interaction. The file is provided in a tab-delimited format with 7 columns:

- Query ORF
- Query gene name
- Array ORF
- Array gene name
- Genetic interaction score (ϵ)
- Standard deviation
- p-value

Supplementary data file S4. Dataset at intermediate cutoff

[sgadata_costanzo2009_intermediateCutoff_101120.txt.gz](#) - 3.1MB

Update (Nov. 20, 2010). The *clc1Δ* screen was removed because an allele discrepancy was noticed.

The file contains the SGA genetic interaction dataset with an intermediate cutoff applied ($|\epsilon| > 0.08$, p-value < 0.05). Reciprocal interactions (AB vs BA) were processed as follows: if AB and BA show opposite interaction signs (AB is positive and BA is negative, or viceversa), both pairs were removed; if AB and BA show the same interaction sign (both positive or both negative), the interaction with the lowest p-value was retained and both pairs are reported with

that interaction. The file is provided in a tab-delimited format with 7 columns:

- Query ORF
- Query gene name
- Array ORF
- Array gene name
- Genetic interaction score (ϵ)
- Standard deviation
- p-value

Supplementary data file S5. Dataset at stringent cutoff

[sgadata_costanzo2009_stringentCutoff_101120.txt.gz](#) - 1.3MB

Update (Nov. 20, 2010). The *clc1* Δ screen was removed because an allele discrepancy was noticed.

The file contains the SGA genetic interaction dataset with a stringent cutoff applied ($\epsilon < -0.12$, p-value < 0.05 or $\epsilon > 0.16$, p-value < 0.05). Reciprocal interactions (AB vs BA) were processed as follows: if AB and BA show opposite interaction signs (AB is positive and BA is negative, or viceversa), both pairs were removed; if AB and BA show the same interaction sign (both positive or both negative), the interaction with the lowest p-value was retained and both pairs are reported with that interaction. The file is provided in a tab-delimited format with 7 columns:

- Query ORF
- Query gene name
- Array ORF
- Array gene name
- Genetic interaction score (ϵ)
- Standard deviation
- p-value

Supplementary data file S6. Biological process annotations

[bioprocess_annotations_costanzo2009.xls](#) - 800KB

Supplementary data file S7. Chemical genomics data

[chemgenomic_data_costanzo2009.xls](#) - 3.5MB

The file contains the fitness defect scores for 4933 non-essential homozygous and 1107 essential heterozygous deletion mutants in the presence of 11 chemical treatments. Each row represents a gene deletion strain, given by its ORF name. Each column represents a treatment experiment.

Supplementary data file S8. The cell map poster

[poster_costanzo2009.pdf](#) - 16MB (updated on Apr. 1, 2010)

Supplementary data file S9. List of queries

[sgadata_costanzo2009_query_list_101120.txt](#) - 26KB

Update (Nov. 20, 2010). The *clc1* Δ screen was removed because an allele discrepancy was noticed.

Supplementary data file S10. List of arrays

[sgadata_costanzo2009_array_list.txt](#) - 53KB