



Coexpression Tool

User Manual

RegulonDB database

Lucia Pannier

This presentation aims to show the user how to utilize and interpret the coexpression tool.

Summary:

- ➡ Motivation
- ➡ Gene expression and coexpression
- ➡ How is coexpression measured?
- ➡ Interpretation of coexpression values

Copies and Copyright-Notice

User is not entitled to...
update publication...
clear is not entitled to...
from CCG-UNAM.

Citation

User is committed...
update publication...
Susana Gamero-C...
Alberto Santos-Z...
Moreira,
Juan Segura-Salazar, Luis Muñoz-Rascado, Irma Martínez-Flores, Heladia Salgado,
César Bonavides-Martínez, Cel Abreu-Goodger, Carlos Rodríguez-Penagos,
Juan Miranda-Ríos,
Enrique Morett, Enrique Merino, Araceli M. Huerta, Luis Treviño-Quintanilla and Julio Colado-Vides
"RegulonDB (Version 6.0): gene regulation model of Escherichia coli K-12 beyond transcription, active (experimental) annotated promoters and textpresso navigation"
Nucleic Acids Research, 2008, vol 36, D120-D124

Release: 6.4 Date: 2008-09-09

Transcription Factor Matrix and Alignments

The consensus and patser programs were used to create the matrix and alignment

Transcription Factor sites
Total of uniq binding sites

Matrix
A 2 3 1 1 3 2 3 0 0 1 5 2 0 0 1 0 2 2 0 3 2 0 1 1 0 0 0 1 0
C 1 0 0 1 0 1 0 0 1 0 0 1 0 2 0 0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0
G 0 0 1 0 0 0 0 0 1 0 2 0 0 3 1 2 0 0 0 3 0 0 0 2 0 0 2 1 0 1
T 0 0 1 0 0 0 0 2 2 0 0 1 0 0 0 2 0 1 0 1 3 0 2 2 1 2 2 2

AlignmentScore
AAGCAAAGCGCACTGAATACCG
AAAAAATTAAGCGCAAGATTGTGGTTT
CATTACATTGCTGGATAAGAATGTTTAGT9.78

Motivation

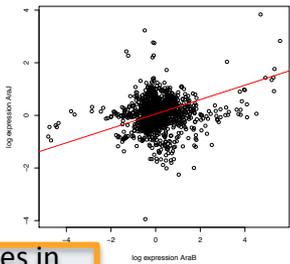
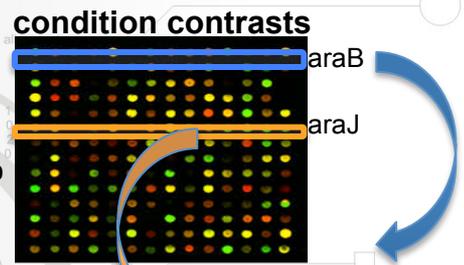
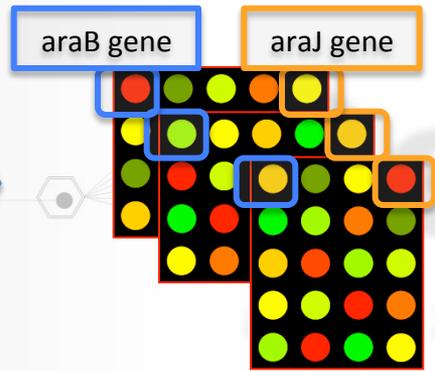
High-throughput assays enable **massive expression measurements** of thousands of genes **simultaneously**. It is possible to measure the similarity of expression between genes to see if they are coexpressed.

Coexpression of genes may indicate that the respective genes are controlled by the same transcriptional regulatory machinery, functionally related or they are members of the same pathway or protein complex.

Copies and Copyright-Notice
 Use is not intended to change or erase data sets of the RegulonDB database. Furthermore, RegulonDB partly written consent from CCG-UNAM.

colombos

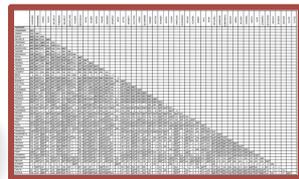
thousands of microarrays ,
 thousands of conditions for
Escherichia coli K-12



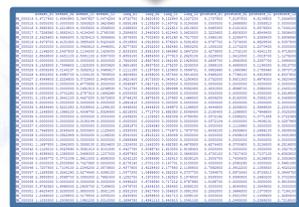
araB, araJ measures in all conditions

Measure of similarity of gene expression by pairs of genes, in all conditions for all genes

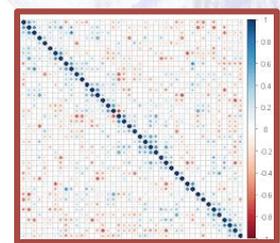
Transformations



Spearman Correlation Coefficient (SCC) matrix for all genes vs all genes



Transformed SCC into Ranked SCC matrix



Spearman Correlation Rank (SCR) matrix

What is coexpression ?

Similarity of gene expression

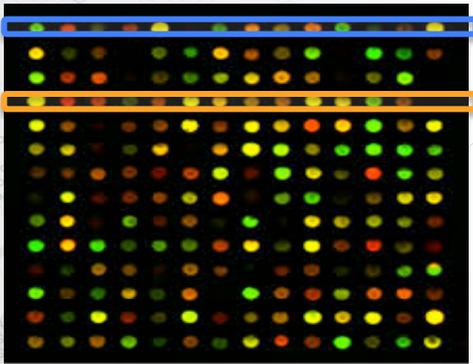
Copies and Copyright-Notice
User is not entitled to change or erase data sets of the RegulonDB database or to eliminate copyright notices from RegulonDB. Furthermore, user is not entitled to expand RegulonDB or to integrate RegulonDB partly or as a whole into other databank systems, without prior written consent from CCG-UNAM.

Citation
User is committed to cite properly the use of RegulonDB. The current most update publication concerning RegulonDB is:

Siciliano-García-Castro, Verónica Jiménez-Jasinto, Martín Perilla-Gil, Alberto Santos-Zavaleta, Mónica I. Peñafoza-Spindola, Bruno Contreras-Moreira, Juan Segura-Salazar, Luis Muñoz-Ramos, Irma Martínez-Flores, Hilaria Salgado, César Bonavides-Martínez, Cel Abreo, Woodger, Carlos Rodríguez-Pérez, Juan Miranda-Ríos, Enrique Morett, Enrique Merino, Araceli and Julio Colado-Vides
"RegulonDB (Version 6.0): gene regulation in *Escherichia coli* K-12 beyond transcription, active (experimental) annotated promoters and textpresso navigation"
Nucleic Acids Res
Release: 6.4 Database

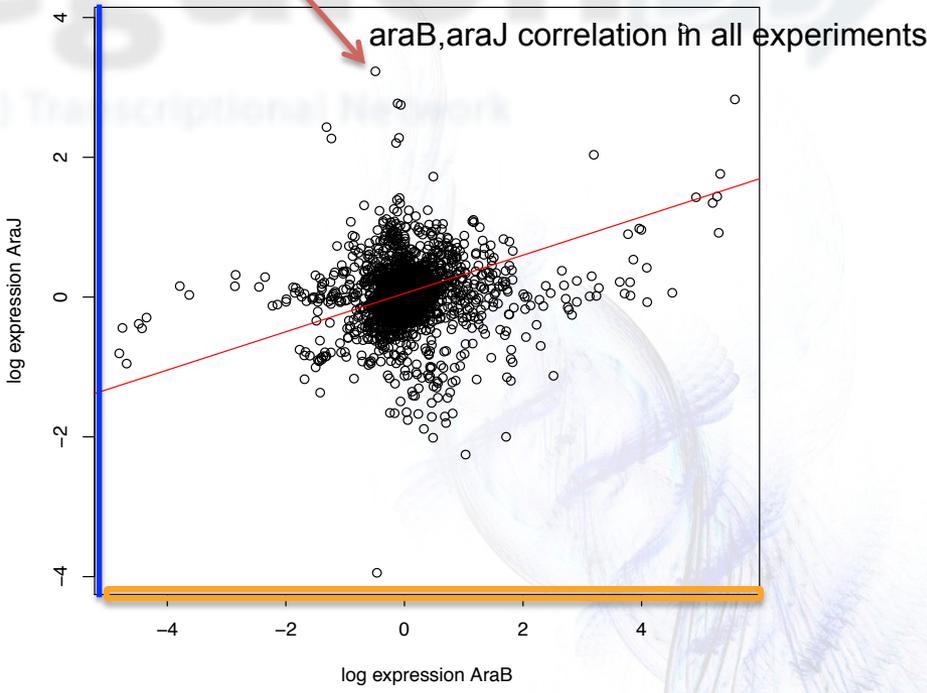


2470 condition contrasts



araB
araJ

One circle represents expression of two genes in one condition contrast
= test versus control of one experiment



Transcript Facto
The con... and
Transcript Facto
Total of uniq bindin
Matrix
A 2
C 1
G 0
T 0
AlignmentScore
AAGCAAAGCCCA
AAAAAATTAAG
CATTACATTGCTG

pairwise correlation as a measure of similarity

Copies and Copyright-Notice

User is not entitled to change or erase data sets of the RegulonDB database or to eliminate copyright notices from RegulonDB. Furthermore, user is not entitled to expand RegulonDB or to integrate RegulonDB partly or as a whole into other databank systems, without prior written consent from CCG-UNAM.

Citation

User is committed to cite properly the use of RegulonDB. The current most update publication concerning RegulonDB is:
Socorro Gama-Castro, Verónica Jiménez-Jacinto, Martín Peralta-Gil, Alberto Santos-Zavaleta, Monica I. Peñalosa-Spindola, Bruno Contreras-Moreira, Juan Segura-Salazar, Luis Muñoz-Rascado, Irma Martínez-Flores, Heladia Salgado, César Bonavides-Martínez, Cel Almey-Goodner, Carlos Rodríguez-Penagos, Juan Miranda-Ries, Enrique Morett, Enrique Martínez-Morales, and Julio Colado-Vides. "RegulonDB (Version 6.0): gene regulation model of Escherichia coli K-12 beyond transcription, active (experimental) promoters and textpresso navigation" Nucleic Acids Research, 2008, vol 36, pp 1531-1536. DOI:10.1093/nar/gkn124
Release: 6.4 Date: 10-AUG-09

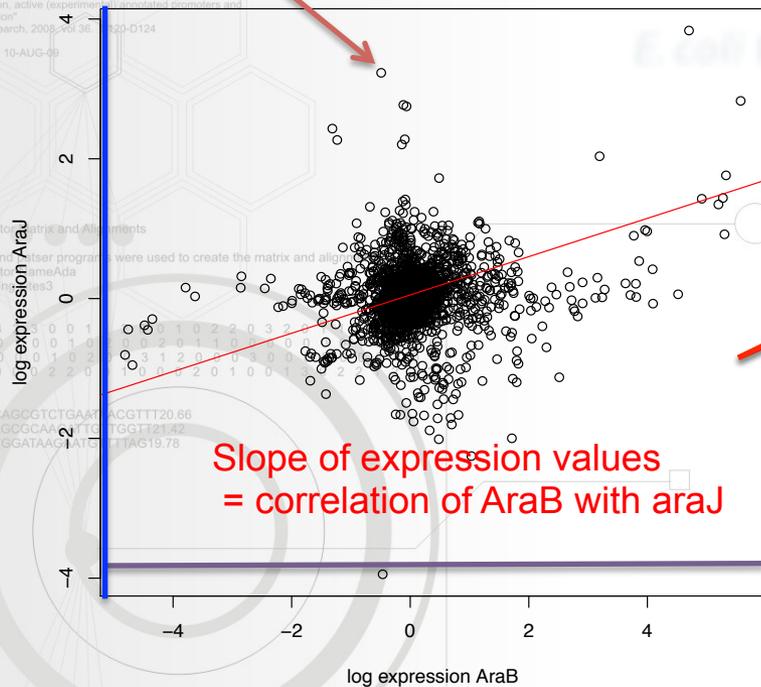
Transcription Factor Matrix and Alignments

The consensus and matrix programs were used to create the matrix and alignments. Transcription Factor name: AraJ
Total of unique binding sites: 3

Matrix
A 2 3 1 1 3 1 1 0 0 1
C 1 0 0 1 0 0 1 0 0 1
G 0 0 1 0 0 0 1 0 0 2
T 0 0 1 1 0 0 2 2 2 0

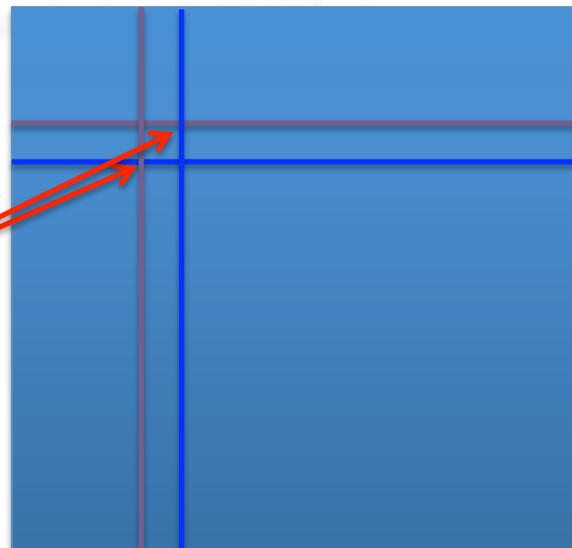
AlignmentScore
AAGCAAAGCGCAGCGTCTGAATACGTTT20.66
AAAAAATTAAAGCGCAATTTGTGGTT21.42
CATTACATTGCTGGATAACATTTTAGT19.78

1 condition contrast

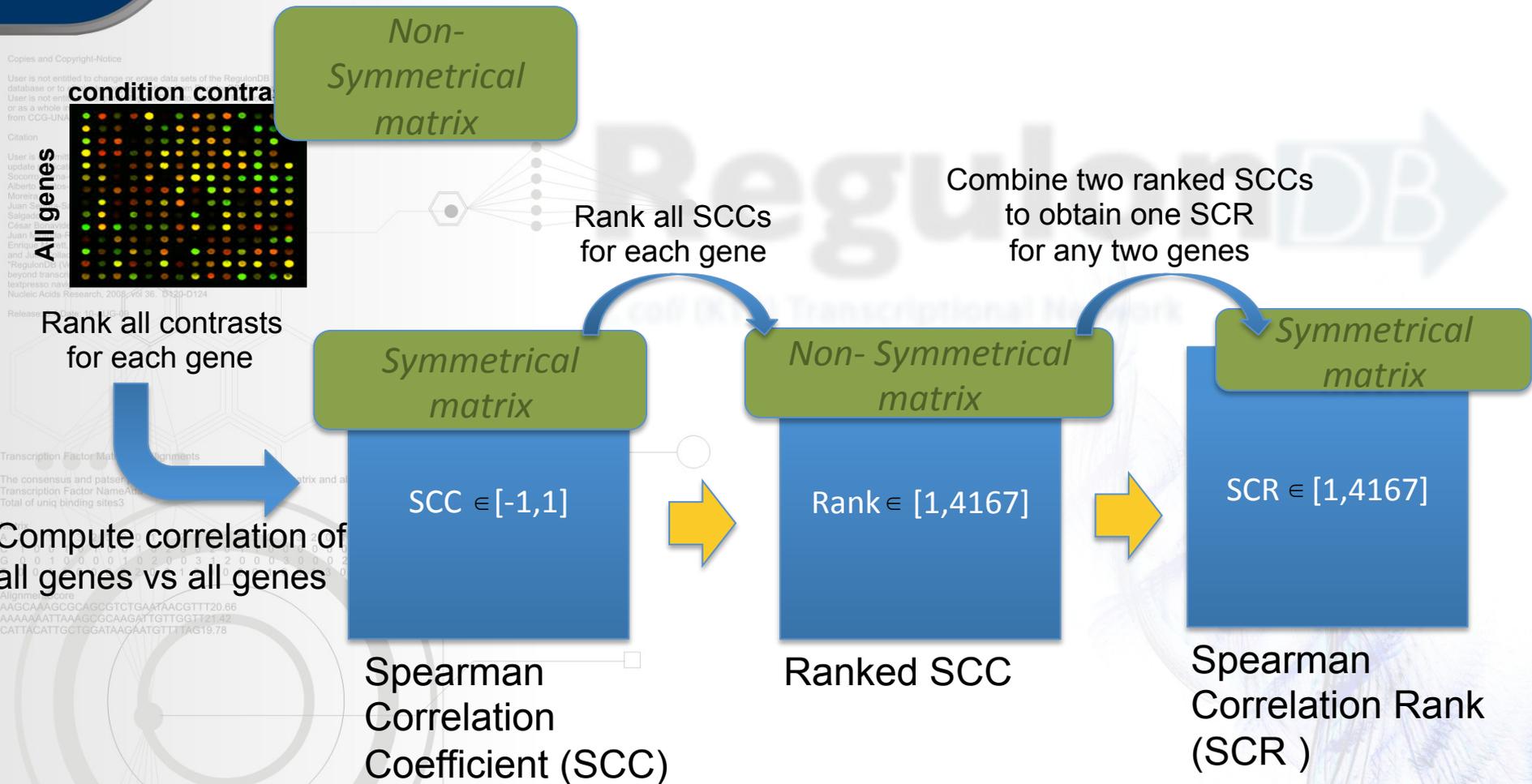


Correlation Matrix
(all genes vs all genes)

AraB AraJ



Transformations



$$SCR(A, B) = \sqrt{rankedSCC_{A(B)} rankedSCC_{B(A)}}$$

Ranking all SCC for a gene

Copies and Copyright-Notice

User is not entitled to change or erase data sets of the RegulonDB database or to eliminate copyright notices from RegulonDB. Furthermore, user is not entitled to expand RegulonDB or to integrate RegulonDB partly or as a whole into other databank systems, without prior written consent from CCG-UNAM.

Citation

User is committed to cite properly the use of RegulonDB. The current most update publication concerning RegulonDB is:

Socorro Gama-Castro, Verónica Jiménez-Jacinto, Martín Peralta-Gil, Alberto Santos-Zavaleta, Mónica I. Peñalosa-Spindola, Bruno Contreras-Moreira, Juan Segura-Salazar, Luis Pacheco-Recas, Irma Martínez-Pérez, Higinio Salgado, César Bonavides-Martín, Juan Miramanda-Ríos, Enrique Morett, Enrique Merino, Araceli M. Huerta, Luis Treviño-Quintanilla and Julio Colado-Vides
"RegulonDB (Version 6.0): gene regulation model of Escherichia coli K-12 beyond transcription, active (experimental) annotated promoters and textpresso navigation"
Nucleic Acids Research 2008, vol 36, D120-D124

Release: 6.4 Date: 10-AG-09

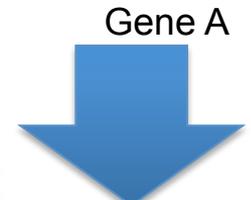
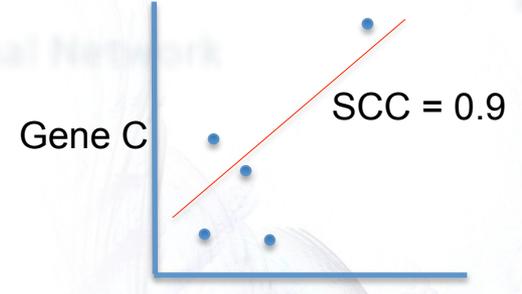
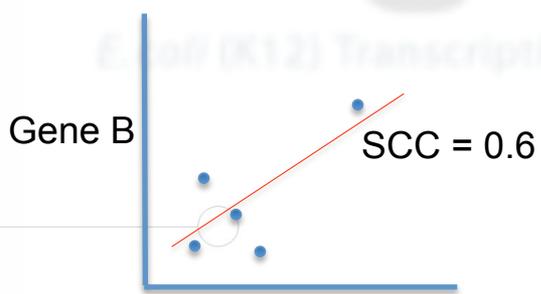
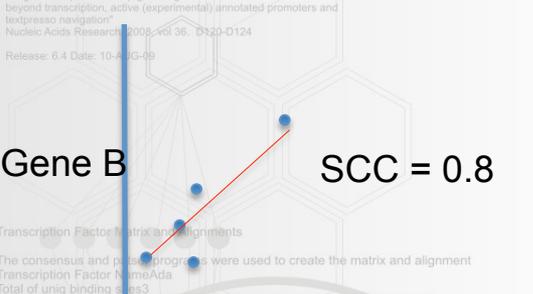
Transcription Factor Matrix and Alignments

The consensus and position programs were used to create the matrix and alignment
Transcription Factor NameAda
Total of uniq binding sites3

Matrix
A 2 3 1 1 3 2 3 0 0 1 1 2 2 2 0 1 1 0 0 0 1 0
C 1 0 0 1 0 1 0 0 1 0 2 0 0 0 0 0 0 0 0 1 0 0 0 0
G 0 0 1 0 0 0 0 1 0 2 0 0 3 1 2 0 0 0 3 0 0 0 2 0 1 0 1
T 0 0 1 1 0 0 0 2 2 0 0 1 0 0 0 2 0 1 0 0 1 3 0 2 2 1 2 2 2

AlignmentScore
AAGCAAAGCGCAGCGTCTGAATAACGTTT20.66
AAAAAATTAAGCGCAAGATTGTTGGTT21.42
CATTACATTGCTGGATAAAGATGTTTAGT19.78

Gene expression measurements of genes A, B and C in 5 experiments (conditions)



Non-Symmetrical matrix

Spearman Correlation coefficient (SCC)	Gene A	Gene B	Gene C
Gene A	1	0.6	0.9
Gene B	0.6	1	0.8
Gene C	0.9	0.8	1

Obtaining SCR from two ranked SCC

Copies and Copyright-Notice
User is not entitled to change or erase data sets of the RegulonDB database or to eliminate copyright notices from RegulonDB. Furthermore, user is not entitled to expand RegulonDB or to integrate RegulonDB partly or as a whole into other databank systems, without prior written consent from CCG-UNAM.

Citation
User is committed to cite properly the use of RegulonDB. The current most update publication concerning RegulonDB is:
Socorro Gama-Castro, Verónica Jiménez-Jasinto, Martín Perilla-Gil, Alberto Santos-Zavaleta, Monica I. Peñalosa-Spindola, Bruno Contreras-Moreira, Juan Segura-Salazar, Luis Muñoz-Rascado, Irma Martínez-Flores, Heladia Salgado, César Encarnación-Martínez, Cal Abreu-Grobois, Carlos Rodríguez-Penagos

Correlation coefficient	Gene A	Gene B	Gene C
Gene A	1	0.6	0.9
Gene B	0.6	1	0.8
Gene C	0.9	0.8	1

Example :
Gene B's **highest** correlation ($r = 0.8$) is with gene C so **rank = 1**
Gene C's **second highest** correlation ($r=0.8$) is with gene B so **rank = 2**

Symmetrical matrix

Non-Symmetrical matrix

Correlation Rank	Gene A	Gene B	Gene C
Gene A	--	2	1
Gene B	2	--	1
Gene C	1	2	--

Two genes B,C have two ranks (1) rank of gene B for gene C (2) rank of gene C for gene B

Obtaining SCR from two ranked SCC

Copies and Copyright-Notice
User is not entitled to change or erase data sets of the RegulonDB database or to eliminate copyright notices from RegulonDB. Furthermore, user is not entitled to expand RegulonDB or to integrate RegulonDB partly or as a whole into other databank systems, without prior written consent from CCG-UNAM.

Citation
User is committed to cite properly the use of RegulonDB. The current most update publication concerning RegulonDB is:
Socorro Gama-Castro, Verónica Jiménez-Jasinto, Martín Perilla-Gil, Alberto Santos-Zavaleta, Monica I. Peñalosa-Spindola, Bruno Contreras-Moreira, Juan Segura-Salazar, Luis Muñoz-Rascado, Irma Martínez-Flores, Heladia Salgado, César Bonavides-Martínez, Cel Abreu-Goodger, Carlos Rodríguez-Penagos, Juan Miranda-Ríos, Enrique Morán, Enrique Marino, Araceli M. Huerta, Luis Treviño-Quintanilla and Julio C. "RegulonDB beyond tra textpresso Nucleic Ac Release: 6



Non-Symmetrical matrix

Ranked SCC	Gene A	Gene B	Gene C
Gene A	--	2	1
Gene B	2	--	1
Gene C	1	2	--

Example :
Two genes B,C have two ranks
(1) rank of gene B for gene C
(2) rank of gene C for gene B

Geometric mean

Symmetrical matrix

SCR	Gene A	Gene B	Gene C
Gene A	--	2	1.4
Gene B	2	--	1.4
Gene C	1.4	1.4	--

Two genes B,C have one mutual rank (1.4) = the geometrical mean of the two ranks

Obtaining SCR from two ranked SCC

Example

Copies and Copyright-Notice
User is not entitled to change or erase data sets of the RegulonDB database or to eliminate copyright notices from RegulonDB. Furthermore, user is not entitled to reproduce or to disseminate, in any form, the whole or as a whole into any other database or to use the information contained therein for any purpose other than that for which it was originally intended.
Citation
User is committed to update publication information in the RegulonDB database. Please contact Alberto Santos-Zeigler, Juan Segura-Salas, Saiguel, Cesar Bonavides, Juan Miranda-Rios, Enrique Morett, Enrique and Julio Collado-Chavez, and the RegulonDB team beyond transcriptome textpresso navigat Nucleic Acids Res
Release: 6.4 Date

Correlation Rank	Gene A	Gene B	Gene C
Gene A	--	2	1
Gene B	2	--	1
Gene C	1	2	--

Non-Symmetrical matrix

$$SCR(A, B) = \sqrt{rankedSCC_{A(B)} rankedSCC_{B(A)}}$$

Why geometric mean ?

The geometric mean is closer to the smallest argument than to the largest argument.

→ the mutual rank tends to be closer to the “best” rank than to the “worst” rank

(see Pannier L et al, *in prep.* for more detailed explanation)

Symmetrical matrix

Mutual Correlation Rank	Gene A	Gene B	Gene C
Gene A	--	2	1.4
Gene B	--	--	1.4
Gene C	--	--	--

EVALUATION OF THE SCR

RANGE : 1 until 4167 (number of genes -1)

INTERPRETATION : If two genes have $SCR=1$, it means that these two genes have the highest similarity with each other compared to the similarities they have with all other genes

If two genes have $SCR=4167$, it means these two genes have the lowest similarity with each other compared to the similarities with all other genes

Copies and Copyright-Notice

User is not entitled to change or erase data sets of the RegulonDB database or to eliminate copyright notices from RegulonDB. Furthermore, user is not entitled to expand RegulonDB or to integrate RegulonDB partly or as a whole into other databank systems, without prior written consent from CCG-UNAM.

Citation

User is committed to cite properly the use of RegulonDB. The current most update publication concerning RegulonDB is:

Socorro Gama-Castro, Verónica Jiménez-Jasinto, Martín Peralta-Gil, Alberto Santos-Zavaleta, Monica I. Peñalzo-Spindola, Bruno Contreras-Moreira,

Juan Segura-

Salgado,

César Bonavito,

Juan Miranda,

Enrique Morea,

and Julio Coll.

"RegulonDB 10.0

beyond trans-

criptpresso na-

Nucleic Acids

Research

Release: 6.4.0

2013

10.1186/1745-2918-6-4

doi:10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4

10.1186/1745-2918-6-4