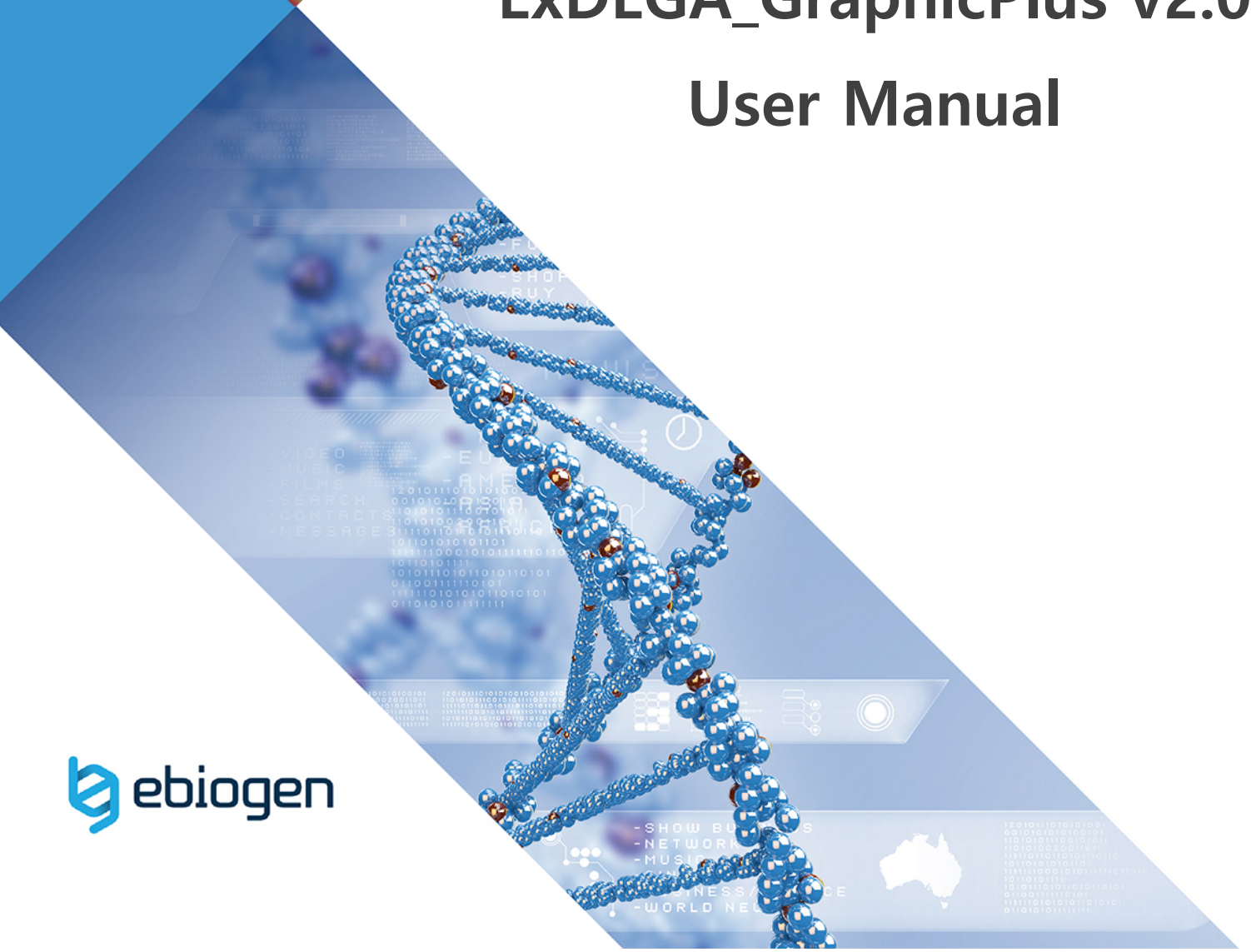


ExDEGA_GraphicPlus v2.0

User Manual



ebiogen

<목 차>

1. Introduction

2. Prerequisites

3. Getting started

- *Execution*
- *Authentication*

4. DAVID Graphic

- *Interface description*

5. Clustering Heatmap

- *Interface description*

6. PCA Plot 2D/3D

- *PCA(Principal Component Analysis)*
- *Interface description*

7. File Format

- *DAVID Graphic*
- *Clustering Heatmap*
- *PCA Plot 2D/3D*

Introduction

㈜이바이오젠에서 개발한 ExDEGA_GraphicPlus v2.0은 생명공학 데이터에 대한 시각화 기능들을 모은 분석 소프트웨어입니다. 연구자가 사용하기에 상당히 간단한 프로그램이며 직관적으로 ep 데이터를 확인 할 수 있습니다. 연구자가 시각화에 필요한 입력 파일을 직접 구성하고 시각화 자료를 편집할 수 있기 때문에 확장성이 높습니다. ExDEGA GraphicPlus v2.0은 컴퓨터 프로그래밍 언어인 Python을 기반으로 작성되었으며 기능을 구현하기 위해 범용적으로 쓰이는 Opensource Package들을 사용하였습니다. ExDEGA GraphicPlus v2.0 User Manual은 해당 프로그램을 사용하기 위해 알아야 할 입력 파일 구성 방법과 기능 사용 방법을 소개 합니다.

Prerequisites

ExDEGA_GraphicPlus v2.0을 이용하기 위해 Python에 대한 지식은 필요하지 않지만 ExDEGA 활용방법, DAVID 분석 방법에 대해서는 최소한의 기본 지식이 있어야합니다.
참고) ExDEGA_UserManual을 통해 위의 내용을 학습하실 수 있습니다.

Getting started

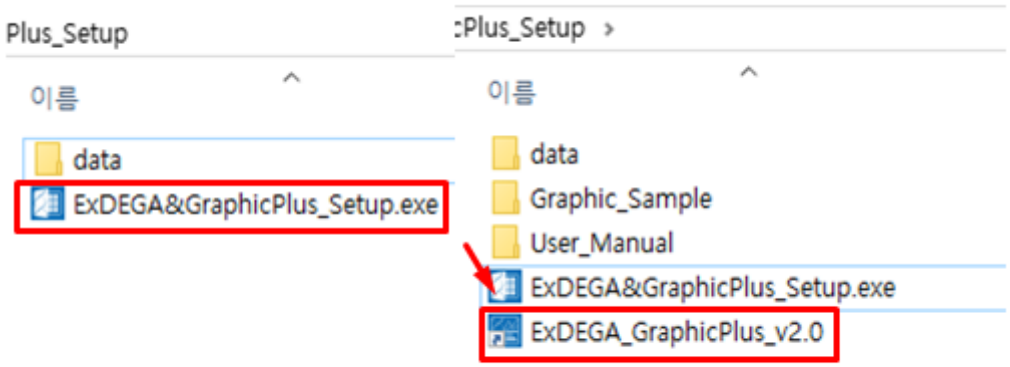
본 챕터에서는 ExDEGA_GraphicPlus의 실행 방법과 사용자 인증 방법에 대해 안내합니다.

Execution

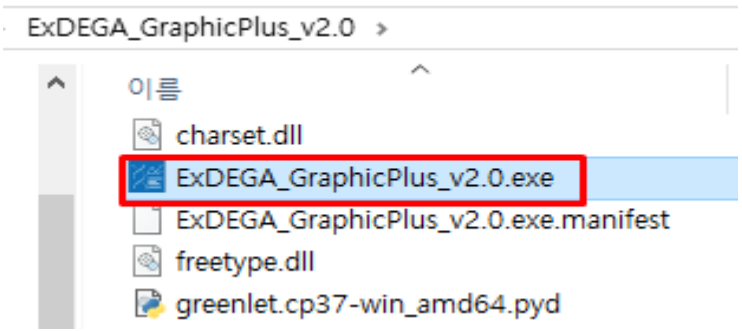
ExDEGA_GraphicPlus는 기본적으로 2가지 방법을 사용하여 실행할 수 있습니다.

1. ExDEGASetup.exe를 활용하여 ExDEGA_GraphicPlus 설치하여 실행.

2. ExDEGA_GraphicPlus 배포 폴더에서 exe파일을 직접 실행.



실행 방법 1

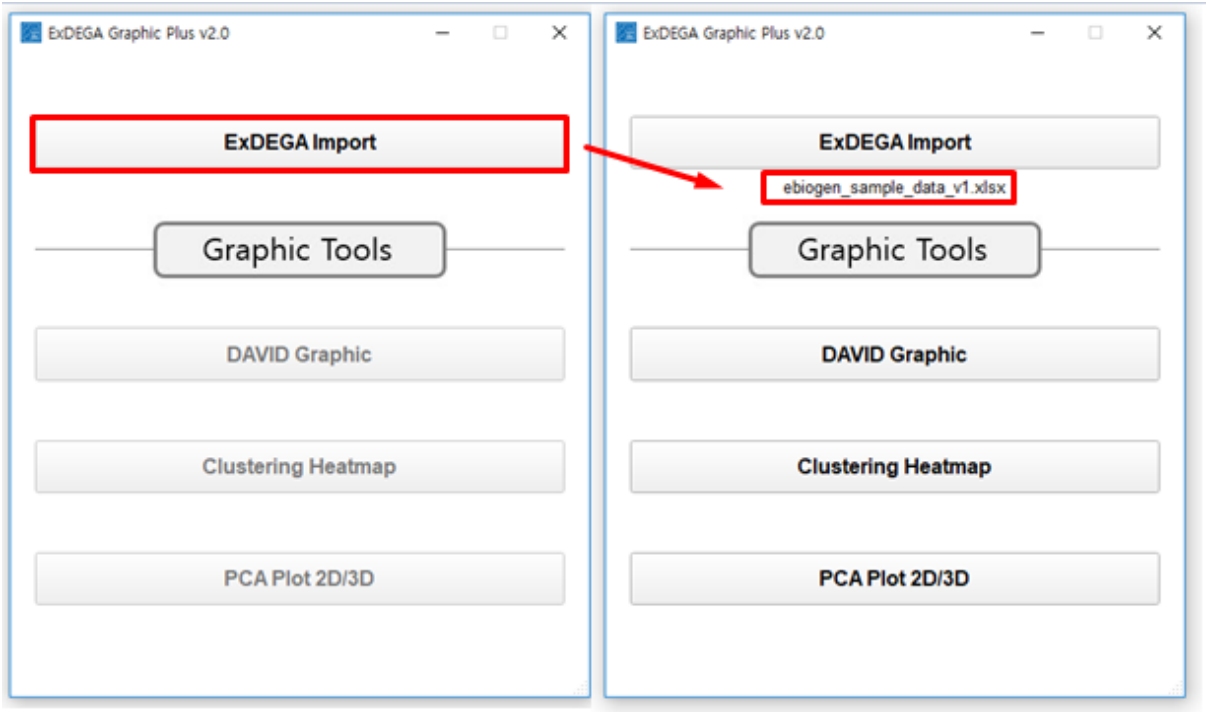


실행 방법 2

Authentication

ExDEGA_GraphicPlus의 모든 기능을 이용하기 위해서는 사용자 인증이 필요합니다. 사용자 인증은 ExDEGA 프로그램이 실행되는 Report파일로 진행되며 인증과 동시에 Report의 데이터는 아티팩트 형식으로 프로그램 내부에 저장됩니다.

참고) Report파일명에 공백이 있을 시 에러가 발생할 수 있습니다.



인증 방법

Function1. DAVID Graphic

DAVID분석은 연구자가 선택한 유전자 리스트에 내포된 생물학적 의미를 알 수 있습니다.

DAVID Graphic은 생물학 기반 분류에 report 데이터를 연동하여 시각화 합니다.

Interface description

A. DAVID output : DAVID 분석 결과 파일을 입력합니다. Here을 통해 자세한 내용을 확인 할 수 있습니다.

B. DAVID input : DAVID 분석에 입력된 Gene Symbol과 그에 상응하는 Fold Change 파일을 입력합니다. Here을 통해 자세한 내용을 확인 할 수 있습니다.

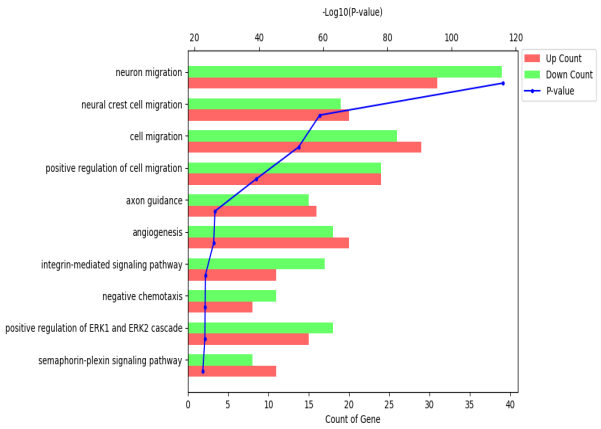
C. Term select : DAVID 분석의 P-value 상위 n개를 분석합니다.

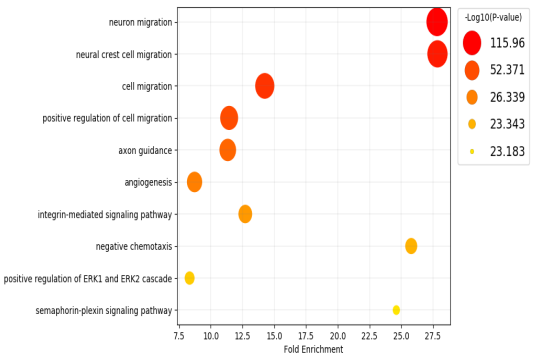
D. Count of Gene : DAVID 분석 결과와 Report 데이터가 연동되어 유전자의 증감을 막대 그래프로 나타냅니다.

E. Fold Enrichment : DAVID 분석 결과와 Report 데이터가 연동되어 유전자 포함 비율을 Bubble Chart로 나타냅니다.

F. Graphic Chart : 설정을 반영하여 해당되는 Chart를 그립니다.

G. Save Graphic Value : 그래프에 사용되는 값들을 Excel 파일 형태로 저장합니다.





Function2. Clustering Heatmap

본 챕터에서는 Heatmap(개별 색상의 직사각형 테이터 행렬)과 Dendrogram(계층적 클러스터링)을 합쳐 Hierarchical Clustering Heatmap을 그리는 방법을 설명합니다. Hierarchical Clustering Heatmap은 연구자가 선택한 유전자의 발현 유사성을 기반으로 Sample의 유사성을 판단할 때 사용합니다.

참고) Hierarchical은 비슷한 군집을 묶어가며 최종적으로 하나의 케이스로 만드는 알고리즘 입니다.

Interface description

A. Hierarchical Clustering Input : Clustering Heatmap을 그리기 위해 필요한 데이터 행렬 파일 입니다. 파일 형태와 추출 방법은 Here을 통해 확인 할 수 있습니다.

B. Sample Cluster : Column Dendrogram을 그립니다.

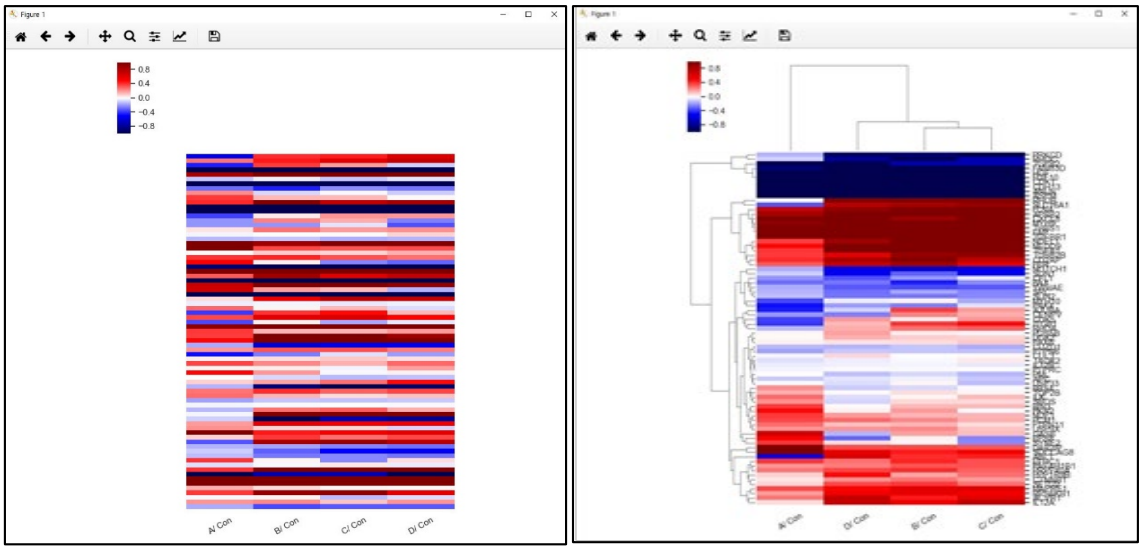
C. Gene Cluster : Raw Dendrogram을 그립니다.

D. Gene Name : Raw에 해당하는 유전자 이름을 표시합니다. 단, 입력한 유전자가 80개 이상일 경우에는 사용할 수 없습니다.

E. Color scheme : Heatmap Theme를 설정합니다.

F. Range : 데이터 표현 범위를 설정합니다. MIN보다 작은 데이터, MAX보다 큰 데이터는 MIN 값과 MAX값으로 합쳐집니다.

G. Clustering Heatmap : 설정 값들을 반영하여 Hierarchical Clustering Heatmap을 그립니다.



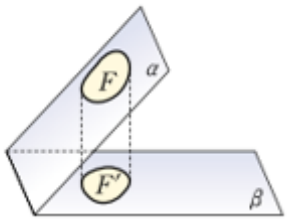
Function3. PCA Plot 2D/3D

본 챕터에서는 PCA 분석과 PCA 2D/3D 를 그리는 방법에 대해 설명합니다. PCA 는 Sample 간의 유사성을 확인할 수 있습니다. Clustering Heatmap 과는 다르게 Sample 에 속해있는 유전자 전체 발현을 기반으로 Sample 간의 유사성을 확인합니다.

PCA(Principal Component Analysis)

PCA Essential Info.

PCA 는 주성분 분석의 준말로 고차원 데이터를 정사영(구조 유지, 차원 감소) 시켜 저차원 데이터로 차원 축소하는 알고리즘입니다.

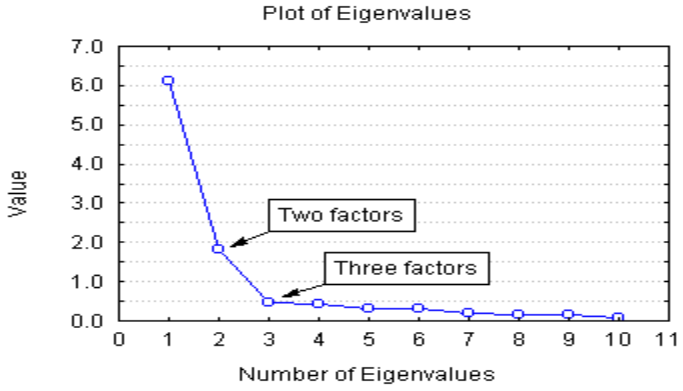


F'은 F 의 β 위로의 정사영

PCA 를 보는 목적은 앞서 설명한 바와 같이 Sample 에 속해있는 전체 유전자 발현을 대상으로 Sample 간의 유사성을 확인입니다. Human RNA-seq 기준으로 각 Sample 에는 약 24,000 개 이상의 유전자가 포함됩니다. 24,000 개의 변수가 생긴다는 말과 동일하기 때문에 Sample 간의 유사성을 파악하기 힘들게 됩니다. PCA 알고리즘을 통해 방향 변화 없이 variance 최대(선형 변환)가 되는 λ 값(eigenvalue)을 계산합니다. λ 값은 sample 의 개수만큼 나오게 되고 주성분 비율은 λ 값의 전체 합에서 해당 λ 값이 차지하는 비율 입니다.

예시 : ⅰ. PC1(60%), PC2(30%) => 주성분 1, 2 로 데이터의 90% 표현

ⅱ. PC1(50%), PC2(30%), PC3(10%) => 주성분 1,2,3 으로 데이터의 90% 표현



PCA additional Info.

다음은 공분산을 이용하여 PCA 를 계산하는 방법을 기술합니다.

2 개의 Sample(열)에서 5 개의 유전자(행)에 발현을 관찰하여 행렬(D)로 만듭니다.

$$D = \begin{bmatrix} 100 & 111 \\ 152 & 45 \\ 19 & 33 \\ 22 & 31 \\ 27 & 10 \end{bmatrix} \epsilon\ R^{5*2}$$

유전자 각각의 변동을 알기 위해 행렬(D)에서 평균 값(기준점)을 뺌으로써 행렬(X) 구할 수 있습니다. (평균 값을 원점으로 데이터가 확장된 정도)

$$X = D - m_{ean}(D) = \begin{bmatrix} 100 & 111 \\ 152 & 45 \\ 19 & 33 \\ 22 & 30 \\ 27 & 10 \end{bmatrix} - \begin{bmatrix} 64 & 46 \\ 64 & 46 \\ 64 & 46 \\ 64 & 46 \\ 64 & 46 \end{bmatrix} = \begin{bmatrix} 36 & 65 \\ 88 & -1 \\ -45 & -13 \\ -42 & -16 \\ -37 & -36 \end{bmatrix}$$

Transpose X 와 X 를 곱함으로써 데이터의 내적 값을 구할 수 있고 (유전자의 개수-1)을 나눔으로 데이터 공분산 행렬을 구할 수 있습니다.

$$X^TX = \begin{bmatrix} 36 & 88 & -45 & -42 & -37 \\ 65 & -1 & -13 & -16 & -36 \end{bmatrix} \begin{bmatrix} 36 & 65 \\ 88 & -1 \\ -45 & -13 \\ -42 & -16 \\ -37 & -36 \end{bmatrix} = \begin{bmatrix} 14198 & 4841 \\ 4841 & 5947 \end{bmatrix}$$

$$\frac{x^Tx}{n-1} = \begin{bmatrix} 14198 & 4841 \\ 4841 & 5947 \end{bmatrix} / 4 = \begin{bmatrix} 3549.5 & 1210.25 \\ 1210.25 & 1486.74 \end{bmatrix}$$

공분산은 공통되게 움직이는 정도를 표현한 것으로 이제 방향(Eigenvalue)를 계산해야 합니다. Eigenvalue(λ)는 nonzero solution vector K (Eigenvector) 가 존재해야 한다는 조건이 있습니다. 따라서, $AK = \lambda K$ 를 만족하여야 하며 $K(A - \lambda I) = 0$ 과 같습니다. 이때 $K(A - \lambda I)$이 역행렬이 존재 한다면 K = 0 이기 때문에 조건에 모순이 됩니다.

$\therefore\ \det(A - \lambda I) = 0$

위의 예제를 공식에 따라 대입해보면 다음과 같습니다.

$$\det\left(\begin{bmatrix} 3549.5 - \lambda & 1210.25 \\ 1210.25 & 1486.74 - \lambda \end{bmatrix}\right) = 0 \rightarrow (3549.5 - \lambda)(1486.74 - \lambda) - 1210.25^2 = 0$$

$\therefore\ \boldsymbol{\lambda_1 = 275.101,\ \lambda_2 = -5311.341}$

Interface description

A. Default Data(Single Normalized Data) : 초기화면에서 입력한 Report 파일의 Normalized Data 가 자동으로 입력됩니다. Array Data 의 경우 해당 옵션을 이용하실 수 없습니다.

B. Group Data : Group 별로 색상을 동일하게 PCA 를 그릴 수 있습니다. 파일 형태와 추출 방법은 Here 을 통해 확인 할 수 있습니다.

C. Filtered Data : 유전자를 필터링하여 PCA 를 그릴 수 있습니다. 파일 형태와 추출 방법은 Here 을 통해 확인 할 수 있습니다.

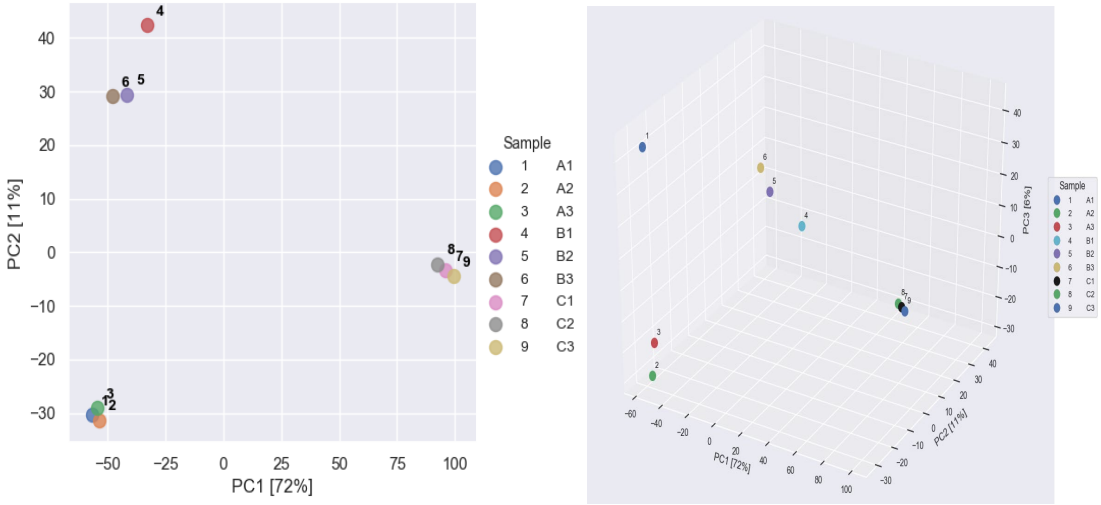
D. Denote Index : Sample 의 인덱스를 표시합니다.

E. PCA 2D : 설정 값들을 반영하여 PCA 를 진행하고 결과를 2 차원 평면에 그립니다.

F. PCA 3D : 설정 값들을 반영하여 PCA 를 진행하고 결과를 3 차원 공간에 그립니다.

G. Save Eigenvalues : 계산된 Eigenvalue 의 결과를 Excel 파일 형식으로 저장합니다.

H. Save PC Values : 각 주성분이 전체 데이터를 얼마만큼 설명할 수 있는 지 분산 비율로 계산되어 Excel 파일 형식으로 저장합니다.



File Format (sample download)

DAVID Graphic

DAVID Output

DAVID 분석의 결과 파일로 Term, Count, P-value, Genes, Pop Hits, Fold Enrichment 의 열이 필수적으로 포함되어 있어야 합니다.

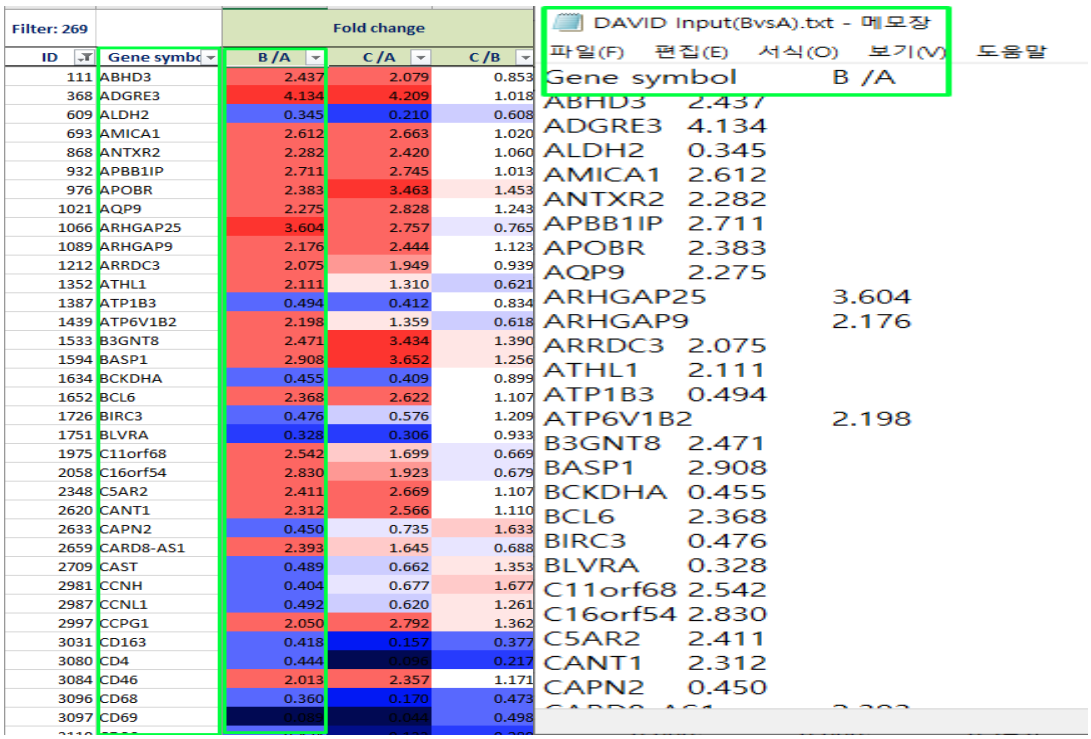
참고) DAVID 분석 진행 방법은 'ExDEGA UserManual'을 참고해주세요.

Category	Term	Count	%	PValue	Genes	List Total	Pop Hits	Pop Total	Fold Enrichment	Bonferron	Benjamini	FDR
GOTERM_BP_DIRECT	GO:0055114~oxidation-reduction process	36	16.43836	2.70E-14	D3YZE4, P5	203	676	18082	4.743580027	3.05E-11	3.05E-11	4.33E-11
GOTERM_BP_DIRECT	GO:0006631~fatty acid metabolic process	19	8.675799	1.50E-13	P04117, P5	203	156	18082	10.84874321	1.70E-10	8.48E-11	2.40E-10
GOTERM_BP_DIRECT	GO:0006635~fatty acid beta-oxidation	12	5.479452	1.43E-12	Q9DC50, C	203	44	18082	24.29287953	1.62E-09	5.40E-10	2.30E-09
GOTERM_BP_DIRECT	GO:0008152~metabolic process	28	12.78539	2.10E-12	P19157, P5	203	463	18082	5.386758025	2.37E-09	5.93E-10	3.36E-09
GOTERM_BP_DIRECT	GO:0006629~lipid metabolic process	24	10.9589	1.96E-09	P51174, Q9	203	459	18082	4.657458386	2.22E-06	4.43E-07	3.14E-06
GOTERM_BP_DIRECT	GO:0006810~transport	48	21.91781	3.28E-08	P04117, P2	203	1822	18082	2.346622831	3.71E-05	6.18E-06	5.26E-05
GOTERM_BP_DIRECT	GO:0006637~acyl-CoA metabolic process	7	3.196347	1.05E-06	Q8VCT4, Q	203	31	18082	20.1134594	0.001191	1.70E-04	0.001689
GOTERM_BP_DIRECT	GO:0070527~platelet aggregation	7	3.196347	3.70E-06	Q9Z1Q5, P	203	38	18082	16.40834846	0.004178	5.23E-04	0.005935
GOTERM_BP_DIRECT	GO:0006754~ATP biosynthetic process	5	2.283105	1.13E-04	Q03265, D:	203	23	18082	19.36388948	0.120415	0.014155	0.181743
GOTERM_BP_DIRECT	GO:0015671~oxygen transport	4	1.826484	1.56E-04	P02089, P0	203	10	18082	35.62955665	0.161483	0.017458	0.249389
GOTERM_BP_DIRECT	GO:0006749~glutathione metabolic process	6	2.739726	2.12E-04	P48774, P1	203	49	18082	10.90700714	0.213205	0.021563	0.339389
GOTERM_BP_DIRECT	GO:0051791~medium-chain fatty acid metabol	3	1.369863	3.70E-04	Q9DC50, C	203	3	18082	89.07389163	0.342104	0.034291	0.591879

DAVID Input

DAVID 분석에 입력한 Gene symbol 과 상응하는 fold change 로 구성하여 만듭니다. fold change 는 비교 조합 중 하나만 선택할 수 있습니다. 파일은 메모장과 엑셀 두가지 방법으로 제작 할 수 있습니다.

1. 메모장으로 만들기



Filter: 269		Fold change		
ID	Gene symbo	B /A	C /A	C /B
111	ABHD3	2.437	2.079	0.853
368	ADGRE3	4.134	4.209	1.018
609	ALDH2	0.345	0.210	0.608
693	AMICA1	2.612	2.663	1.020
868	ANTXR2	2.282	2.420	1.060
932	APBB1IP	2.711	2.745	1.013
976	APOBR	2.383	3.463	1.453
1021	AQP9	2.275	2.828	1.243
1066	ARHGAP25	3.604	2.757	0.765
1089	ARHGAP9	2.176	2.444	1.123
1212	ARRDC3	2.075	1.949	0.939
1352	ATHL1	2.111	1.310	0.621
1387	ATP1B3	0.494	0.412	0.834
1439	ATP6V1B2	2.198	1.359	0.618
1533	B3GNT8	2.471	3.434	1.390
1594	BASP1	2.908	3.652	1.256
1634	BCKDHA	0.455	0.409	0.899
1652	BCL6	2.368	2.622	1.107
1726	BIRC3	0.476	0.576	1.209
1751	BLVRA	0.328	0.306	0.933
1975	C11orf68	2.542	1.699	0.669
2058	C16orf54	2.830	1.923	0.679
2348	C5AR2	2.411	2.669	1.107
2620	CANT1	2.312	2.566	1.110
2633	CAPN2	0.450	0.735	1.633
2659	CARD8-AS1	2.393	1.645	0.688
2709	CAST	0.489	0.662	1.353
2981	CCNH	0.404	0.677	1.677
2987	CCNL1	0.492	0.620	1.261
2997	CCPG1	2.050	2.792	1.362
3031	CD163	0.418	0.157	0.377
3080	CD4	0.444	0.096	0.217
3084	CD46	2.013	2.357	1.171
3096	CD68	0.360	0.170	0.473
3097	CD69	0.089	0.044	0.498

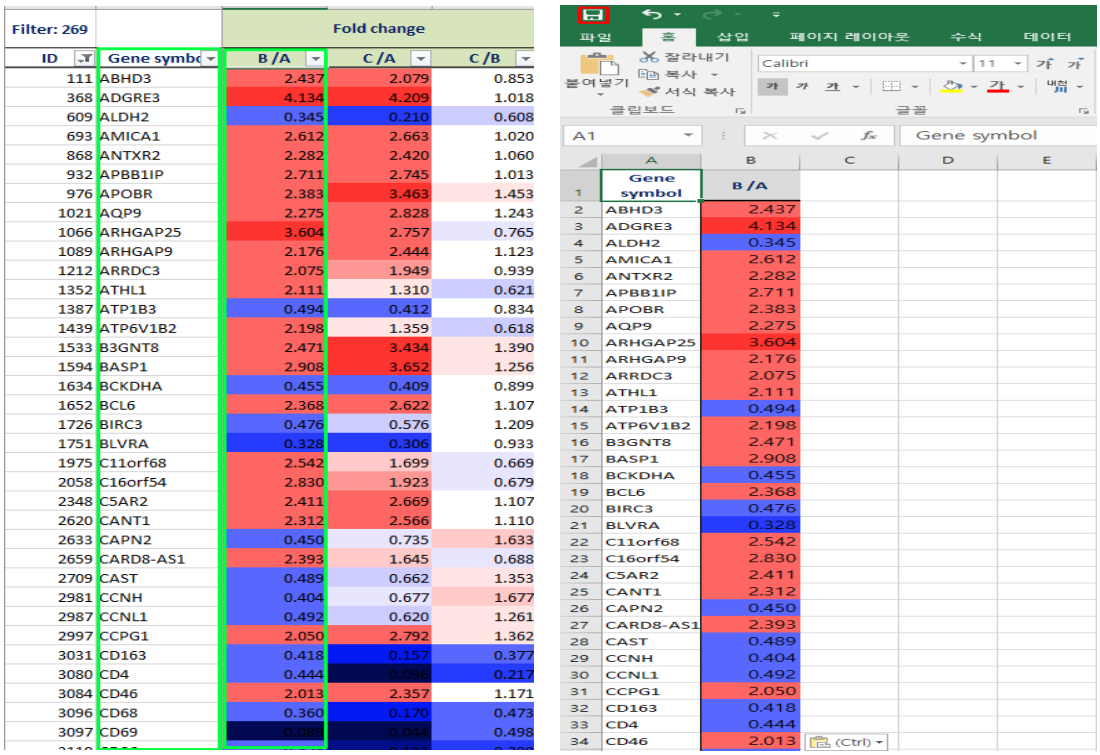
DAVID Input(BvsA).txt - 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말

Gene symbol	B /A
ABHD3	2.437
ADGRE3	4.134
ALDH2	0.345
AMICA1	2.612
ANTXR2	2.282
APBB1IP	2.711
APOBR	2.383
AQP9	2.275
ARHGAP25	3.604
ARHGAP9	2.176
ARRDC3	2.075
ATHL1	2.111
ATP1B3	0.494
ATP6V1B2	2.198
B3GNT8	2.471
BASP1	2.908
BCKDHA	0.455
BCL6	2.368
BIRC3	0.476
BLVRA	0.328
C11orf68	2.542
C16orf54	2.830
C5AR2	2.411
CANT1	2.312
CAPN2	0.450

Copy (Report File) → Paste (Note)

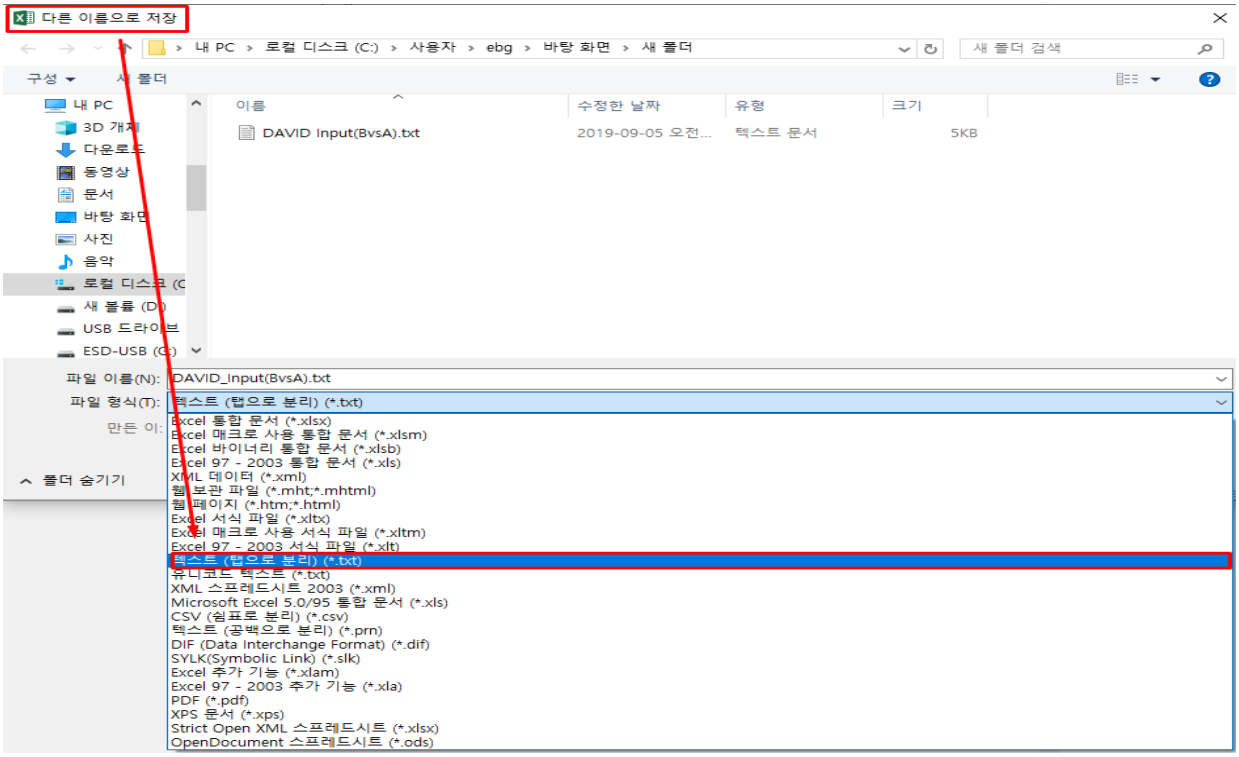
2. Excel 파일로 만들기



Filter: 269		Fold change		
ID	Gene symbo	B /A	C /A	C /B
111	ABHD3	2.437	2.079	0.853
368	ADGRE3	4.134	4.209	1.018
609	ALDH2	0.345	0.210	0.608
693	AMICA1	2.612	2.663	1.020
868	ANTXR2	2.282	2.420	1.060
932	APBB1IP	2.711	2.745	1.013
976	APOBR	2.383	3.463	1.453
1021	AQP9	2.275	2.828	1.243
1066	ARHGAP25	3.604	2.757	0.765
1089	ARHGAP9	2.176	2.444	1.123
1212	ARRDC3	2.075	1.949	0.939
1352	ATHL1	2.111	1.310	0.621
1387	ATP1B3	0.494	0.412	0.834
1439	ATP6V1B2	2.198	1.359	0.618
1533	B3GNT8	2.471	3.434	1.390
1594	BASP1	2.908	3.652	1.256
1634	BCKDHA	0.455	0.409	0.899
1652	BCL6	2.368	2.622	1.107
1726	BIRC3	0.476	0.576	1.209
1751	BLVRA	0.328	0.306	0.933
1975	C11orf68	2.542	1.699	0.669
2058	C16orf54	2.830	1.923	0.679
2348	C5AR2	2.411	2.669	1.107
2620	CANT1	2.312	2.566	1.110
2633	CAPN2	0.450	0.735	1.633
2659	CARD8-AS1	2.393	1.645	0.688
2709	CAST	0.489	0.662	1.353
2981	CCNH	0.404	0.677	1.677
2987	CCNL1	0.492	0.620	1.261
2997	CCPG1	2.050	2.792	1.362
3031	CD163	0.418	0.157	0.377
3080	CD4	0.444	0.096	0.217
3084	CD46	2.013	2.357	1.171
3096	CD68	0.360	0.170	0.473
3097	CD69	0.085	0.044	0.498

	A	B
1	Gene symbol	B /A
2	ABHD3	2.437
3	ADGRE3	4.134
4	ALDH2	0.345
5	AMICA1	2.612
6	ANTXR2	2.282
7	APBB1IP	2.711
8	APOBR	2.383
9	AQP9	2.275
10	ARHGAP25	3.604
11	ARHGAP9	2.176
12	ARRDC3	2.075
13	ATHL1	2.111
14	ATP1B3	0.494
15	ATP6V1B2	2.198
16	B3GNT8	2.471
17	BASP1	2.908
18	BCKDHA	0.455
19	BCL6	2.368
20	BIRC3	0.476
21	BLVRA	0.328
22	C11orf68	2.542
23	C16orf54	2.830
24	C5AR2	2.411
25	CANT1	2.312
26	CAPN2	0.450
27	CARD8-AS1	2.393
28	CAST	0.489
29	CCNH	0.404
30	CCNL1	0.492
31	CCPG1	2.050
32	CD163	0.418
33	CD4	0.444
34	CD46	2.013

Copy (Report File) → Paste (Excel)

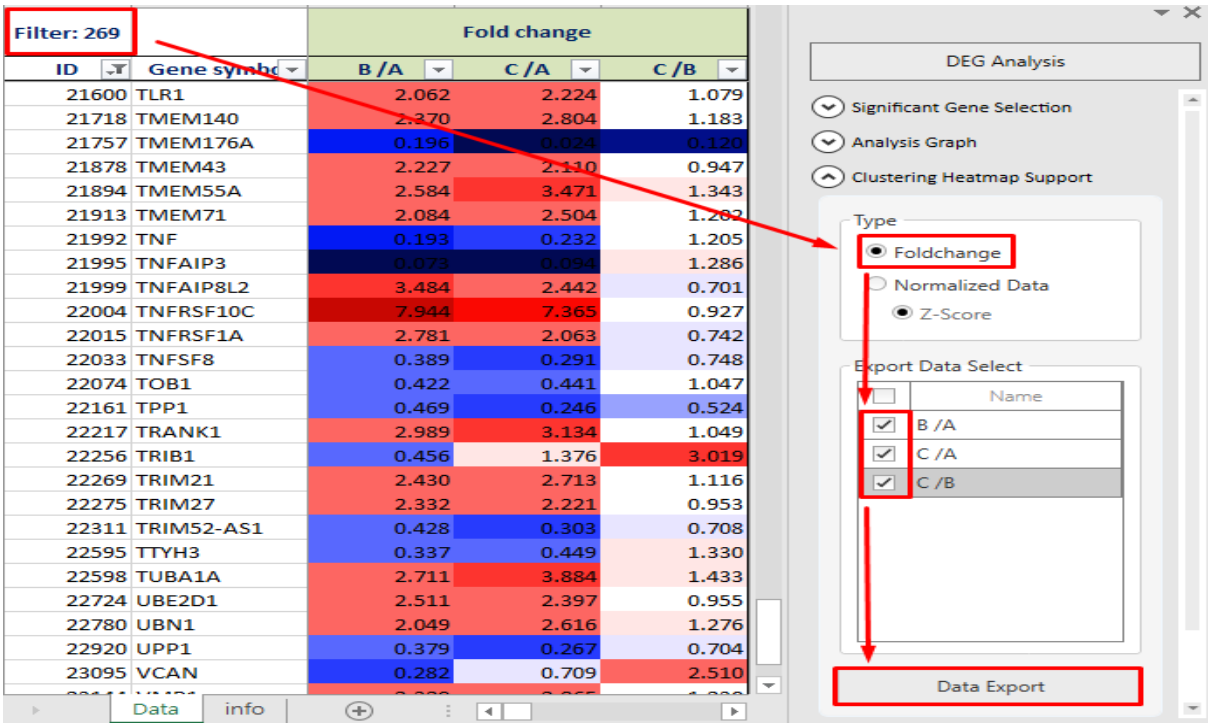


Separated by Tab

Clustering Heatmap

Hierarchical Clustering Input

Clustering Heatmap 은 연구자가 선택한 유전자의 상관 관계를 나타냅니다. 따라서, Input File 제작에 앞서 유전자를 선별하는 것이 필요합니다. 유전자 선별은 Gene Ontology 분석, Kegg 분석, DEG 분석 등으로 선별할 수 있습니다. 유전자를 선별한 후 ExDEGA 기능 중 Clustering Heatmap Support 를 통해 Input File 을 제작합니다. 위의 파일은 Gene Symbol 과 상응하는 데이터 열로 구성됩니다. 데이터 열은 비교 조합의 Fold Change 나 각 Sample(Group) Normalized Data 의 Z-score 로 구성되며 비교 조합 또는 Sample(Group)의 개수는 관계 없습니다.



Heatmap(Allcompare_FC).txt - 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말

Gene Symbol	B /A	C /A	C /B
ABHD3	1.28484278864215	1.05564963600944	-0.229193276278856
ADGRE3	2.04741030637742	2.07337604787126	0.0259657102702727
ALDH2	-1.53704893440065	-2.25385030388101	-0.716801237557984
AMICA1	1.38502619744398	1.41309305964294	0.0280669503482263
ANTXR2	1.19047263791619	1.27486038592147	0.0843881542963257
APBB1IP	1.43866864864493	1.45671679922983	0.0180472923397484
APOBR	1.2529266643143	1.79222275950347	0.539295814093037
AQP9	1.18587922828891	1.49957447547321	0.313695136333861

PCA Plot 2D/3D

Group Data

같은 Group 의 Sample 들을 묶어 같은 색상으로 표시하여 PCA 를 표현할 수 있습니다. Input File 은 Excel 에서 작성합니다.

작성 순서

1) Normailized data(log2)항목에서 데이터 복사

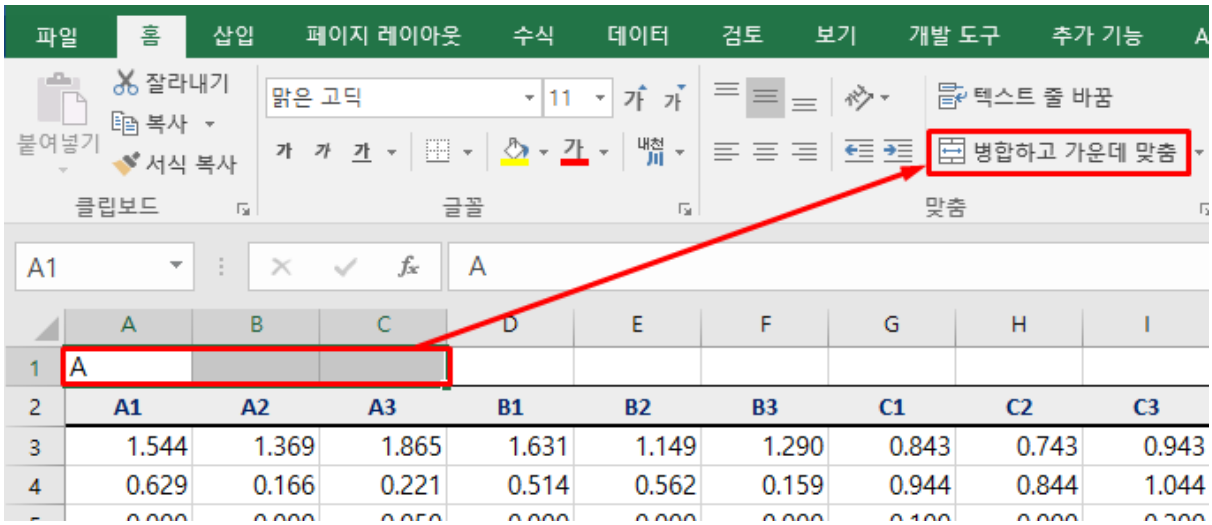
	Normalized data (log2)								
	A1	A2	A3	B1	B2	B3	C1	C2	C3
5	1.544	1.369	1.865	1.631	1.149	1.290	0.843	0.743	0.943
6	0.629	0.166	0.221	0.514	0.562	0.159	0.944	0.844	1.044
2	0.000	0.000	0.050	0.000	0.000	0.000	0.100	0.000	0.200
0	0.231	0.330	0.100	0.642	0.091	0.071	0.328	0.228	0.428
5	0.303	0.799	0.111	1.771	0.917	0.294	0.923	0.823	1.023
2	0.000	0.000	0.063	0.055	0.000	0.088	0.100	0.000	0.200
2	0.000	0.290	0.000	0.000	0.098	0.000	0.100	0.000	0.200
2	0.440	0.002	0.000	0.001	0.001	0.862	0.100	0.000	0.200
4	0.218	0.564	0.000	0.001	0.002	0.000	1.742	1.642	1.842
2	0.000	0.000	0.266	0.000	0.000	0.000	0.100	0.000	0.200
2	0.406	0.001	0.000	0.000	0.000	0.000	0.100	0.000	0.200
6	3.136	2.775	2.732	2.813	3.270	3.323	2.474	2.374	2.574
3	1.510	1.210	0.904	1.276	1.172	0.962	2.241	2.141	2.341
2	0.000	0.000	0.000	0.098	0.000	0.157	0.100	0.000	0.200
2	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200

참고) 전체 데이터 선택 방법은 A1(초록색 네모) 클릭 후 C3(파란색 네모)를 Shift + 클릭(빨간색 화살표) 합니다. A1~ C3 까지 선택 후에 Shift + 아래 방향키(↓)를 입력하면 전체 데이터가 한번에 선택됩니다.

2) 새 통합 문서를 실행 후에 **2 번째** 열에 '붙여넣기'

	A	B	C	D	E	F	G	H	I
1									
2	A1	A2	A3	B1	B2	B3	C1	C2	C3
3	1.544	1.369	1.865	1.631	1.149	1.290	0.843	0.743	0.943
4	0.629	0.166	0.221	0.514	0.562	0.159	0.944	0.844	1.044
5	0.000	0.000	0.050	0.000	0.000	0.000	0.100	0.000	0.200
6	0.231	0.330	0.100	0.642	0.091	0.071	0.328	0.228	0.428
7	0.303	0.799	0.111	1.771	0.917	0.294	0.923	0.823	1.023
8	0.000	0.000	0.063	0.055	0.000	0.088	0.100	0.000	0.200
9	0.000	0.290	0.000	0.000	0.098	0.000	0.100	0.000	0.200
10	0.440	0.002	0.000	0.001	0.001	0.862	0.100	0.000	0.200
11	0.218	0.564	0.000	0.001	0.002	0.000	1.742	1.642	1.842
12	0.000	0.000	0.266	0.000	0.000	0.000	0.100	0.000	0.200
13	0.406	0.001	0.000	0.000	0.000	0.000	0.100	0.000	0.200
14	3.136	2.775	2.732	2.813	3.270	3.323	2.474	2.374	2.574
15	1.510	1.210	0.904	1.276	1.172	0.962	2.241	2.141	2.341
16	0.000	0.000	0.000	0.098	0.000	0.157	0.100	0.000	0.200
17	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200
18	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200
19	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200

3) 1 번째 열에는 그룹명 입력 후 병합



	A	B	C	D	E	F	G	H	I
1	A								
2	A1	A2	A3	B1	B2	B3	C1	C2	C3
3	1.544	1.369	1.865	1.631	1.149	1.290	0.843	0.743	0.943
4	0.629	0.166	0.221	0.514	0.562	0.159	0.944	0.844	1.044

4) 완성

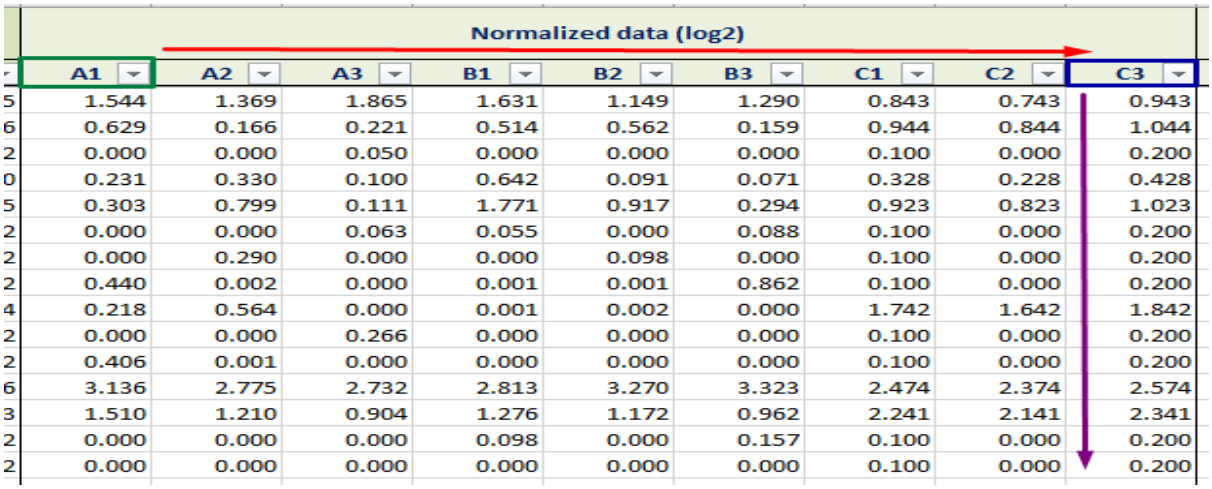
	A	B	C	D	E	F	G	H	I
1	A			B			C		
2	A1	A2	A3	B1	B2	B3	C1	C2	C3
3	1.544	1.369	1.865	1.631	1.149	1.290	0.843	0.743	0.943
4	0.629	0.166	0.221	0.514	0.562	0.159	0.944	0.844	1.044
5	0.000	0.000	0.050	0.000	0.000	0.000	0.100	0.000	0.200
6	0.231	0.330	0.100	0.642	0.091	0.071	0.328	0.228	0.428
7	0.303	0.799	0.111	1.771	0.917	0.294	0.923	0.823	1.023
8	0.000	0.000	0.063	0.055	0.000	0.088	0.100	0.000	0.200
9	0.000	0.290	0.000	0.000	0.098	0.000	0.100	0.000	0.200
10	0.440	0.002	0.000	0.001	0.001	0.862	0.100	0.000	0.200
11	0.218	0.564	0.000	0.001	0.002	0.000	1.742	1.642	1.842
12	0.000	0.000	0.266	0.000	0.000	0.000	0.100	0.000	0.200
13	0.406	0.001	0.000	0.000	0.000	0.000	0.100	0.000	0.200
14	3.136	2.775	2.732	2.813	3.270	3.323	2.474	2.374	2.574
15	1.510	1.210	0.904	1.276	1.172	0.962	2.241	2.141	2.341
16	0.000	0.000	0.000	0.098	0.000	0.157	0.100	0.000	0.200
17	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200
18	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200
19	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200
20	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200

Filtered Data

수정된 데이터를 입력할 수 있는 File Format 입니다.

작성 순서

1) Normailized data(log2)항목에서 데이터 복사



	Normalized data (log2)								
	A1	A2	A3	B1	B2	B3	C1	C2	C3
5	1.544	1.369	1.865	1.631	1.149	1.290	0.843	0.743	0.943
6	0.629	0.166	0.221	0.514	0.562	0.159	0.944	0.844	1.044
2	0.000	0.000	0.050	0.000	0.000	0.000	0.100	0.000	0.200
0	0.231	0.330	0.100	0.642	0.091	0.071	0.328	0.228	0.428
5	0.303	0.799	0.111	1.771	0.917	0.294	0.923	0.823	1.023
2	0.000	0.000	0.063	0.055	0.000	0.088	0.100	0.000	0.200
2	0.000	0.290	0.000	0.000	0.098	0.000	0.100	0.000	0.200
2	0.440	0.002	0.000	0.001	0.001	0.862	0.100	0.000	0.200
4	0.218	0.564	0.000	0.001	0.002	0.000	1.742	1.642	1.842
2	0.000	0.000	0.266	0.000	0.000	0.000	0.100	0.000	0.200
2	0.406	0.001	0.000	0.000	0.000	0.000	0.100	0.000	0.200
6	3.136	2.775	2.732	2.813	3.270	3.323	2.474	2.374	2.574
3	1.510	1.210	0.904	1.276	1.172	0.962	2.241	2.141	2.341
2	0.000	0.000	0.000	0.098	0.000	0.157	0.100	0.000	0.200
2	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200

참고) 전체 데이터 선택 방법은 A1(초록색 네모) 클릭 후 C3(파란색 네모)를 Shift + 클릭(빨간색 화살표) 합니다. A1~ C3 까지 선택 후에 Shift + 아래 방향키(↓)를 입력하면 전체 데이터가 한번에 선택됩니다.

2) 새 통합 문서를 실행 후에 **2 번째** 열에 '붙여넣기' 후 완성

	A	B	C	D	E	F	G	H	I
1	A1	A2	A3	B1	B2	B3	C1	C2	C3
2	1.544	1.369	1.865	1.631	1.149	1.290	0.843	0.743	0.943
3	0.629	0.166	0.221	0.514	0.562	0.159	0.944	0.844	1.044
4	0.000	0.000	0.050	0.000	0.000	0.000	0.100	0.000	0.200
5	0.231	0.330	0.100	0.642	0.091	0.071	0.328	0.228	0.428
6	0.303	0.799	0.111	1.771	0.917	0.294	0.923	0.823	1.023
7	0.000	0.000	0.063	0.055	0.000	0.088	0.100	0.000	0.200
8	0.000	0.290	0.000	0.000	0.098	0.000	0.100	0.000	0.200
9	0.440	0.002	0.000	0.001	0.001	0.862	0.100	0.000	0.200
10	0.218	0.564	0.000	0.001	0.002	0.000	1.742	1.642	1.842
11	0.000	0.000	0.266	0.000	0.000	0.000	0.100	0.000	0.200
12	0.406	0.001	0.000	0.000	0.000	0.000	0.100	0.000	0.200
13	3.136	2.775	2.732	2.813	3.270	3.323	2.474	2.374	2.574
14	1.510	1.210	0.904	1.276	1.172	0.962	2.241	2.141	2.341
15	0.000	0.000	0.000	0.098	0.000	0.157	0.100	0.000	0.200
16	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200
17	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200
18	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200
19	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200
20	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200
21	0.000	0.000	0.000	0.000	0.000	0.000	0.100	0.000	0.200