

02 RNA-seq 기반 유전자 발현 프로파일링

RNAseq 데이터를 이용해 원하는 대사 경로의 특정 유전자 발현 값의 패턴 분석



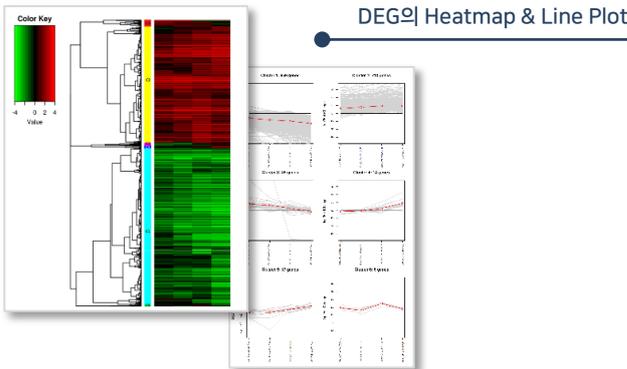
RNAseq 분석을 통해 표준유전체가 완료된 생물뿐만 아니라 진행 중인 생물체를 대상으로 관심 있는 현상에 관여하는 유전자가 무엇인지, 유전자 발현 변화와 패턴을 분석합니다. 시료 준비 과정이나 시퀀싱 과정 중에 발생하는 오차를 최대한 보정하고, 염기서열의 특징에 의해 발생하는 오류 가능성을 최소화하기 위한 다양한 통계 기법을 적용하여 정확도를 높이기 위한 노력을 기울이고 있습니다. 단순 유전자 발현값 측정에 머무르지 않고 목표 형질이나 현상에 관련된 유전자를 선별하기 위해 clustering 기법, GO, KEGG 등 다양한 기법과 문헌자료를 통해 구축한 씨더스 고유 DB를 활용하여 보다 정제된 데이터를 드리 고자 노력하고 있습니다.

RNAseq Gene expression profiling

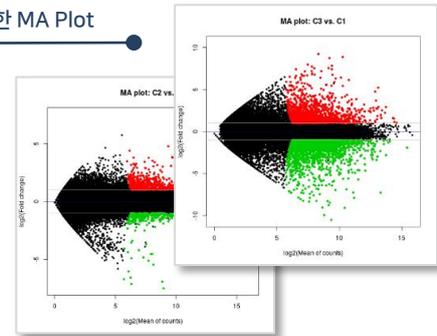
DEG & Clustering 분석

❖ Clustering 분석

- 탐색한 DEG 후보군의 발현 패턴에 따라 군집화 (Clustering) 하여 그래프로 나타냅니다.



유전자 발현값을 표현한 MA Plot



❖ DEG (Differentially Expressed Genes) 선별

- 샘플별 유전자 발현값을 통계적으로 처리하여 대조군과 비교군 샘플 간에 발현이 유의한 DEG 후보군을 선별합니다.

GO & KEGG pathway 분석



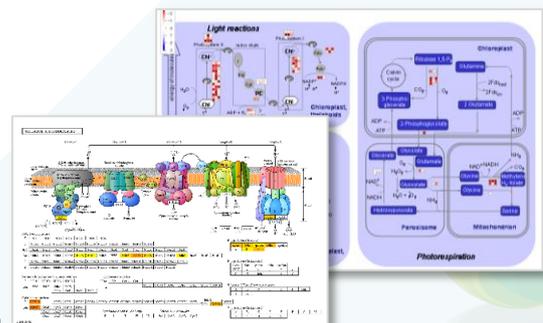
GO 분석 상세 결과 제곱

❖ GO (Gene Ontology) 분석

- 선별된 DEG는 씨더스가 보유한 Gene Ontology DB와 대조하여 Gene annotation을 수행합니다.

❖ KEGG pathway 분석

- KEGG mapper, Mapman 등의 visualization tool을 이용해 목표 Pathway에 관여하는 유전자 발현의 변화를 한 눈에 확인할 수 있습니다.



KEGG pathway visualization



씨더스 고유의 기능별 유전자 DB를 활용한 유전자 선발 사례

병저항성 관련, 스트레스 관련, TF 관련, 식물 호르몬 관련, 지방산 생합성 관련, 개화 관련, Terpenoid 생합성 관련, 과실색 생합성 관련, 광합성 관련 유전자 등 다양한 기능별 유전자군을 문헌자료 통해 정리하여 씨더스 고유 DB로 구축하여 RNAseq 분석에 활용합니다.

TF 관련 유전자 set 추출

Symbol	Ref. Count	Transcripts Count	Symbol	Ref. Count	Transcripts Count	Symbol	Ref. Count	Transcripts Count	Symbol	Ref. Count	Transcripts Count
AP2	45	75 (22)	ERF	185	92 (39)	M-type	114	4 (4)	SRS	7	6 (5)
ARF	33	209 (18)	FAR1	4	3 (2)	MIKC	53	32 (16)	STAT	1	1 (1)
ARR-B	21	32 (8)	G2-like	85	67 (33)	MYB	126	97 (63)	TALE	32	63 (23)
B3	88	64 (29)	GATA	50	46 (19)	MYB_related	107	75 (43)	TCP	43	29 (17)
B8R-BPC	12	12 (6)	GRAS	71	262 (40)	NAC	131	115 (49)	Trithelix	39	36 (24)
BES1	12	29 (11)	GRF	12	22 (8)	NF-X1	2	2 (1)	VOZ	4	7 (2)
C2H2	136	157 (62)	GeBP	14	11 (7)	NF-YA	28	20 (9)	VOX	8	5 (4)
C3H	68	126 (41)	HB-PHD	3	3 (2)	NF-YB	31	9 (6)	WRKY	125	131 (49)
CAMTA	0	0 (0)	HB-other	12	13 (9)	NF-YC	12	9 (5)	Whirly	4	2 (2)
CO-like	16	29 (11)	HD-ZIP	77	175 (85)	NZZ/SPL	0	0 (0)	YABBY	12	5 (4)
CPF	4	6 (4)	HRT-like	2	2 (1)	Nin-like	27	95 (15)	ZF-HD	19	10 (7)
DBB	16	14 (11)	HSF	41	44 (21)	RAW	1	2 (2)	bHLH	206	176 (89)
Dof	40	43 (23)	LBD	47	28 (22)	S1F8-like	7	0 (0)	bZIP	95	157 (46)
E2F/DP	20	38 (8)	LPF	1							
EIL	14	43 (7)	LSD	10							

❖ 사례 1. TF에 해당하는 DEG Annotation

- TF와 관련된 유전자만을 선별하기 위해 씨더스가 확보한 총 53개 Transcription Factors DB에 연결된 유전자 (DEG)와 유전자 발현 값을 정리하였습니다.

Gene id	Reference id	E-value	Identity	TF gene id	TF symbol	Log2FoldChange	C1.1	C1.2	C1.3	C2.1	C2.2	C2.3	C3.1	C3.2	C3.3	C2 vs C1	Log2FC
Subseeder2510299380001	PGSC0002DMF400053152	4.00E-160	99.16	PGSC0002DMF400053042	ERF	11.924	11.919	11.918	11.919	11.241	12.291	9.950	10.004	10.342	10.342	0.00	-0.02
Subseeder2510299380006	PGSC0002DMF400053426	1.20E-127	100	PGSC0002DMF400053244	LSD	7.662	7.658	7.658	7.658	7.658	7.658	7.658	7.658	7.658	7.658	0.00	0.00
Subseeder2510299380002	PGSC0002DMF40005496	0	100	PGSC0002DMF400051114	NAC	5.234	5.233	5.233	5.233	5.451	6.747	1.947	2.216	2.449	2.449	0.27	0.27
Subseeder2510299380003	PGSC0002DMF40005496	0	100	PGSC0002DMF400051112	NAC	5.234	5.233	5.233	5.233	5.451	6.747	1.947	2.216	2.449	2.449	0.27	0.27
Subseeder2510299380004	PGSC0002DMF40005496	0	55.45	PGSC0002DMF400049229	NAC	5.234	5.233	5.233	5.233	5.451	6.747	1.947	2.216	2.449	2.449	0.27	0.27
Subseeder2510299380005	PGSC0002DMF40005496	0	100	PGSC0002DMF400051113	NAC	5.234	5.233	5.233	5.233	5.451	6.747	1.947	2.216	2.449	2.449	0.27	0.27
Subseeder2510299380008	PGSC0002DMF400053589	0	100	PGSC0002DMF400022848	ERF1H	4.440	4.439	4.439	4.439	2.817	3.237	2.111	2.465	2.317	2.317	0.51	-0.51
Subseeder2510299380009	PGSC0002DMF400053700	0	100	PGSC0002DMF400004804	Tobacco	4.367	4.327	4.313	4.345	3.857	4.522	1.349	1.295	2.555	2.555	0.40	-0.40
Subseeder2510299380010	PGSC0002DMF400053452	0	71.57	PGSC0002DMF400050996	Non-like	4.002	4.001	4.001	4.001	2.242	2.242	4.002	3.888	3.314	3.314	-0.42	-0.42
Subseeder2510299380008	PGSC0002DMF400053452	0	63.26	PGSC0002DMF400024895	Non-like	4.002	4.002	4.002	4.002	2.242	2.242	4.120	3.629	3.375	3.375	-0.42	-0.42

식물 호르몬 관련 유전자 set 추출

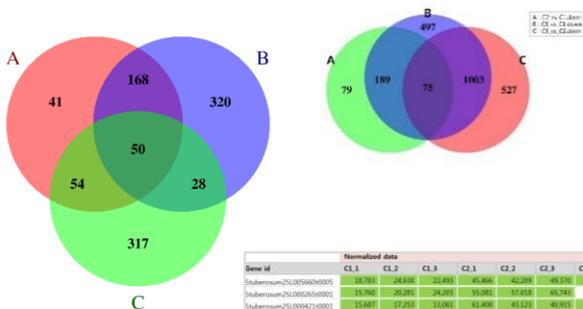
❖ 사례 2. Plant Hormone에 해당하는 DEG 선별

- ABA, Auxin, Cytokinin, Ethylene, GA 등 9개의 식물 호르몬에 해당하는 유전자(DEG)와 유전자 발현 값을 정리하였습니다.

Plant hormone	분석에 사용된 plant hormone gene 개수	Plant hormone 과 맥친된 transcripts 개수
ABA(abscisic acid)	56	329 (144)
Auxin	106	608 (168)
Brassinosteroid	33	747 (329)
Cytokinin	61	238 (61)
Ethylene	41	263 (98)
Gibberellin	34	55 (34)
Jasmonic acid	46	132 (70)
Salicylic acid	61	236 (163)
		2 (2)

Gene id	Reference id	Category	E-value	Identity	SI	Plant Hormone Symbol	Plant Hormone Gene C1.1	C1.2	C1.3	C2.1	C2.2	C2.3	C3.1	C3.2	C3.3	C2 vs C1	Log2FC	
Subseeder2510297780002	PGSC0002DMF400032964	ABA_Signaling	8.00E-142	45.47	AT3G21750	UGT75B1	UGP-GLUCYLIC TRANS	8.961	10.246	10.430	11.424	14.684	15.762	1.512	1.298	1.362	0.3	0.3
Subseeder2510297780003	PGSC0002DMF400032964	ABA_Signaling	1.00E-139	48.23	AT3G21750	UGT75B1	UGP-GLUCYLIC TRANS	8.961	10.246	10.430	11.424	14.684	15.762	1.512	1.298	1.362	0.3	0.3
Subseeder2510297780004	PGSC0002DMF400032964	ABA_Signaling	1.00E-132	44.22	AT4G03280	UGT75B1	UGP-GLUCYLIC TRANS	8.961	10.246	10.430	11.424	14.684	15.762	1.512	1.298	1.362	0.3	0.3
Subseeder2510297780005	PGSC0002DMF400032964	ABA_Signaling	4.00E-129	44.63	AT3G21750	UGT75B1	UGP-GLUCYLIC TRANS	8.961	10.246	10.430	11.424	14.684	15.762	1.512	1.298	1.362	0.3	0.3
Subseeder2510297780006	PGSC0002DMF400032964	ABA_Signaling	1.00E-122	42.37	AT1G08850	ACT2/NMT2	ABA-IMPORTING TRAN	8.929	13.980	14.807	12.261	12.1	13.381	1.68	1.46	1.55	0.08	0.08
Subseeder2510297780007	PGSC0002DMF400032964	ABA_Signaling	0.00E-100	74.64	AT3G04860	ABC2/MMP1	ATP-BINDING CASSETT	10.508	10.674	10.674	10.792	10.674	10.674	10.674	10.674	10.674	0.0	0.0
Subseeder2510297780008	PGSC0002DMF400032964	ABA_Signaling	2.00E-173	48.54	AT5G06930	ABC22/MRC25	ATP-BINDING CASSETT	10.411	10.249	10.249	10.171	9.627	9.491	9.54	9.67	9.49	-0.5	-0.5
Subseeder2510297780009	PGSC0002DMF400032964	ABA_Signaling	2.00E-166	48.12	AT1G08260	ABC22/MRC25	ATP-BINDING CASSETT	10.411	10.249	10.249	10.171	9.627	9.491	9.54	9.67	9.49	-0.5	-0.5
Subseeder2510297780010	PGSC0002DMF400032964	ABA_Signaling	4.00E-175	75.4	AT4G03950	SNRQ2/ED371	SNRQ-RELATED PROTE	8.918	11.142	11.473	14.418	14.445	14.415	14.1	13.91	13.45	5.09	-0.5
Subseeder2510297780011	PGSC0002DMF400032964	ABA_Signaling	1.00E-170	73.83	AT5G06880	SNRQ2.3	SUCROSE NONFERMENT	8.918	11.142	11.473	14.418	14.445	14.415	14.1	13.91	13.45	5.09	-0.5
Subseeder2510297780012	PGSC0002DMF400032964	ABA_Signaling	1.00E-165	71.7	AT1G03950	SNRQ2.2	SUCROSE NONFERMENT	8.918	11.142	11.473	14.418	14.445	14.415	14.1	13.91	13.45	5.09	-0.5
Subseeder2510297780013	PGSC0002DMF400032964	ABA_Signaling	2.00E-151	41.54	AT4G03950	SNRQ2/ED371	SNRQ-RELATED PROTE	8.918	11.142	11.473	14.418	14.445	14.415	14.1	13.91	13.45	5.09	-0.5

목표 형질 관련 유전자군 선별



❖ 사례 3. 벤 다이어그램을 통한 공통/특이 유전자 선별

- DEG를 대상으로 특정 조건에서 공통적으로, 혹은 단독으로 발현하는 유전자들을 구분하고 벤 다이어그램과 같은 시각적인 정보로 나타내어 효과적으로 확인할 수 있습니다.

Gene id	Reference id	Category	E-value	Identity	SI	Annotation	NR	Unigene	Protein ID	Description	Species				
Subseeder2510005660005	18.781	24.636	22.481	45.466	42.209	49.376	8.34	9.942	8.331	1.06	-1.30	-2.35	0.00E+00	99.29	gkNP_0304.PREDICTED: Camellia sa
Subseeder2510005660001	19.790	26.281	24.261	59.992	57.991	66.741	7.7	9.61	7.64	1.54	0.24	-2.41	0	93.71	gkNP_0063.PREDICTED:110: Solanum tub
Subseeder2510005660002	13.681	17.233	16.261	61.836	60.835	67.783	7.293	22.096	24.938	1.46	0.58	-1.81	0.00E+00	98.73	gkNP_0012.5: Solanum tub
Subseeder2510005660003	30.791	7.87	6.133	25.536	24.490	18.472	1.96	6.47	8.381	1.34	-0.25	-1.89	0.00E+00	100	gkNP_0063.PREDICTED: Solanum tub
Subseeder2510005660004	8.98	9.933	9.44	17.992	18.897	20.173	11.234	11.429	14.979	1.14	0.76	-0.38	1.00E-130	71.51	gkNP_0350.PREDICTED: Solanum pa
Subseeder2510005660008	4.709	8.219	8.049	20.233	21.391	14.790	10.970	10.947	13.847	1.51	0.17	-0.95	0.00E+00	96.5	gkNP_0012.5: Solanum tub
Subseeder2510005660005	4.604	9.969	9.329	14.864	16.298	19.849	10.849	20.054	14.478	1.21	0.95	-0.29	1.00E-20	107	gkNP_0012.5: Solanum tub
Subseeder2510005660001	4.280	4.764	3.918	10.972	9.688	11.694	3.71	4.130	2.674	1.29	-0.09	-1.88	0.00E+00	100	gkNP_0063.PREDICTED: Solanum tub
Subseeder2510005660002	4.174	3.394	3.854	7.627	5.117	7.920	1.209	3.764	4.865	1.19	0.34	-0.88	0.00E+00	100	gkNP_0063.PREDICTED: Solanum tub
Subseeder2510005660001	3.818	2.964	3.562	12.139	9.278	10.427	4.832	2.54	4.811	1.39	0.27	-1.52	0.00E+00	100	gkNP_0063.PREDICTED: Solanum tub
Subseeder2510005660002	3.796	4.401	4.711	13.190	13.832	8.91	4.61	4.866	4.866	1.51	0.68	-0.89	0.00E+00	100	gkNP_0063.PREDICTED: Solanum tub
Subseeder2510005660007	3.694	4.454	2.377	9.780	6.040	9.631	4.489	3.690	4.466	1.32	0.26	-1.05	0.00E+00	99.45	gkNP_0012.5: Solanum tub



Contact Us

042-710-4035 | www.seeders.co.kr | 대전시 유성구 테크노1로 11-3

문의 사항은 위 연락처로 연락 주시기 바랍니다. For further details of our Technology, please contact SEEDERS.

