User Manual

ExMEGA v.1.0 & Data Analysis



90.2

<목 차>

1.	ANCOM and Definitions of key terminologies3
2.	ExMEGA Download and Setup4
3.	Differential Abundance (DA)6
4.	Diversity Analysis (ExMEGA GraphicPlus)10

1. Definitions of key terminologies

ANCOM

Metagenome 의 DA(Differential Abundance) 분석은 t-test 와 ANOVA 와 같이 많이 사용하는 기존 통계방법은 적합하지 않습니다. ANCOM(analysis of composition of microbiomes)은 Metagenome 을 분석하는 검증된 통계적 방법입니다. ANCOM 은 CLR(central log-ratio) 방법을 사용하여 Relative abundance 로 Absolute abundance 를 추정할 수 있습니다. ANCOM 을 효과적으로 이용하기 위해서는 그룹 간에 차이가 있으며 서로 연관된 샘플들이어야 합니다. ANCOM 의 CLR 을 통해 평균(0)을 기준으로 거리가 먼 ASV(Amplicon Sequence Variant)를 대략적으로 파악할 수 있습니다. W 는 통계적 수치이며, 값이 높을수록 그룹간 유의미한 ASV 및 있습니다. 분류군을 선별할 수 SigW(Significant W)는 ANCOM 결과에 따른 임계값(threshold)입니다. 기본적으로 설정된 임계값을 바탕으로 TRUE/FALSE 로 유의미한 ASV 를 선별할 수도 있지만, 임계값을 낮춰서 더 많은 ASV를 선별되도록 선택할 수 있습니다. ANCOM은 통계적 유의성을 바탕으로 Taxon 을 선별할 수 있을 뿐만 아니라, 그룹 3 개 이상에 대한 통계적 비교도 가능한 장점이 있습니다.

기타 용어

ASV(Amplicon Sequence Variant)는 dada2 방법으로 생성되며, 기존 OTU(유사도를 기반으로 하는 서열 클러스터링 방법)보다 뛰어난 해상도를 보여줍니다.

Relative abundance 를 통한 분석만으론 Metagenome 생태계를 비교하기는 어렵습니다. 최근에는 다양한 통계방법(ANCOM 등)을 이용하여 Absolute abundance 를 추정하고 있습니다.

용어	정의
OTU	운영 분류 단위:97% 유사성을 갖는 DNA 서열 그룹.
ASV	PCR 증폭 및 시퀀싱 중에 생성된 가짜 시퀀스를 제거한 후 고처리량 마커 유전자 분석에서 회수된 개별 DNA 시퀀스.
Absolute abundance	생태계의 단위 부피에서 관찰할 수 없는 실제 분류군의 풍부함.
Relative abundance	표본에 있는 모든 분류군의 합에 대한 특징표에서 관찰된 분류군의 비율.0과 1 사이

Reference

https://www.zymoresearch.com/blogs/blog/microbiome-informatics-otu-vs-asv a Callahan, Benjamin J., Paul J. McMurdie, and Susan P. Holmes.

2. ExMEGA Download and Setup

(㈜이바이오젠은 Metagenome data 를 엑셀 기반에서 쉽게 분석할 수 있도록 분석 결과 보고 시 ExMEGA (Excel based MetaGenome Analysis) tool 과 ExMEGA Graphic Plus 를 함께 제공한다. ExMEGA 분석 툴은 ㈜이바이오젠이 연구자들이 Metagenome 데이터를 보다 쉽게 다루고 원하는 데이터를 쉽게 얻을 수 있도록 사용자 편의를 최대한 반영한 분석 툴이고 엑셀 프로그램 안에서 다양한 분석을 직관적으로 수행할 수 있도록 개발되었다. ExMEGA 분석 툴은 사용자들의 요구사항을 지속적으로 반영하여 데이터 분석과 엑셀 사용에 익숙하지 못한 연구자들도 쉽게 사용이 가능하도록 계속 업데이트 될 예정이다.

이바이오젠에서 제공하는 Metagenome 를 열기 전에 다운로드 폴더에서 함께 제공한 ExMEGA_v(버전)_Installer.zip 파일의 압축을 풀고, setup.exe 를 실행하면 분석 툴이 설치된다(그림 1-1 A). 만약 설치가 되지 않을 경우, 압축을 푼 파일에 있는 ExMEGA 폴더를 컴퓨터의 로컬 C 드라이브 아래로 복사+붙여넣기 하면 ExMEGA Graphic Plus 프로그램이 설치 완료된다(그림 1-1 B).

설치가 완료되고 ExMEGA format 의 엑셀 데이터를 열면 자동으로 ExMEGA 분석 툴이 구동된다. 참고로 ExMEGA 설치 전에 실행 중인 엑셀 파일이 있으면 종료시킨 후 다시 실행해야 ExMEGA 를 사용할 수 있다.

PC > 다운로드 > 13. Programs > ExMEGA	_v1.0.4_Installer	~ C	ExMEGA_v1.0.4_Installe
이름	수정한 날짜	유형	크기
http://www.com/analysis	2022-07-06 오전 10:42	파일 폴더	
Exmega	2022-07-06 오전 10:43	파일 폴더	
ExMEGASetup.application	2022-07-05 오전 9:56	Application Manifest	2KB
🖫 placer.bat	2022-07-06 오전 10:45	Windows 배치 파일	1KB
📚 setup.exe	2022-07-05 오전 9:56	응용 프로그램	507KB

그림 1-1 A. ExMEGA set up

ЧРС → OS (C:)			~	С
이름	수정한 날짜	유형	크기	
https://www.apps	2021-08-05 오전 11:28	파일 폴더		
늘 Dell	2022-04-05 오후 1:31	파일 폴더		
Drivers	2021-08-06 오전 3:58	파일 폴더		
Exmega	2022-07-04 오후 3:46	파일 폴더		
📁 ilogen	2022-01-05 오후 5:03	파일 폴더		
📁 langpacks	2020-05-06 오후 10:27	파일 폴더		
📁 PerfLogs	2021-06-05 오후 9:10	파일 폴더		
늘 Program Files	2022-07-01 오전 8:54	파일 폴더		
📁 Program Files (x86)	2022-04-05 오후 1:31	파일 폴더		

그림 1-1 B. ExMEGA Graphic Plus installation

EXMEGA format 의 엑셀 파일을 열면, 왼쪽에 Taxonomy 창과 가운데에 Metagenome data, 오른쪽에 DA(Differential Abundance) Analysis 창이 실행된다(그림 1-2). Taxonomy 창에서는 기본 설정된 Taxonomy 정보가 있고, 원하는 Taxon 들만 선택하여 데이터를 필터링할 수 있다. DA Analysis 창 에서는 Relative Abundance, W 값을 선택하여 DA 선별을 쉽게 할 수 있고 DA 를 각 Taxon 별로 그래프를 작성할 수 있다. DA 분석 창에서 Pie Chart 뿐만 아니라, Volcano Plot, Venn Diagram 을 직접 그릴 수 있고 선별된 ASV(Amplicon Sequence Variant) 대상으로 Bar plot, Krona Chart, PCoA, Clustering heatmap 을 그릴 수 있다.

- ×	1	ilter: 153)	~		<u> </u>	Taxonomy	9				α	R			W	, · · · ·	~	
View All Data	2	ID ,1	ASV 🗸	Kingdo -	Phylun -	Class 🚽	Order -	Family -	Genu: -	Specie -	A/B-	A/C -	B/C -	All 🖵	A/B -	A/C-	B/C -	All 🖵	DA Analysis
	3	1	00086a129b3dc9e8641815d44e8791d9	dBacteri	pFirmicu	Clostrid	o_Clostric	f_Clostrid	g_Clostric		0.558031	-0.765062	0.2070312	2.8784931	0	21	43	38	
Taxonomy	- 4	2	003986eaa110161172a7ec41b5d5c817	d_Bacteria							-0.041443	-0.078945	0.1203878	1.0416684	1	24	40	44	Significant Taxon
Phylum	5	3	00408e6774d91032a45f0b773795c810	dBacteria	pFirmicu	Clostrid	o_Lachno	f_Lachnos	g_UC5-1-2		0.2332102	-0.440241	0.2070312	5.3042839	1	25	43	44	Pie Chart
Class	6	4	004323e46d3b6749f80276f5fe7379bf	dBacteria	pProteol	_Alphapr	oRhodos	funcultu	g_uncultu		-0.041443	0.0583816	-0.016939	0.0240837	1	22	34	36	O Materia Dist
• • • • • •	7	5	005c1dbc469b1ba50c1c9d4dc34557fa	dBacteria	pBacterc	Bactero	oBacterc				-0.041443	0.0583816	-0.016939	0.0255373	1	19	35	40	A COLUMN FOR
Orber	8	6	005e0000f4d659ab67404084dd4af50b	dBacteria	p_Firmicu o	Clostrid	o_Lachno:	f_Lachnos	g_Eisenbe		-0.526017	-2.294205	2.8202213	2.910943	0	0	5	8	Venn Diagram
 Family 	9	7	006858cc0d2251544b863550d9d067d0	dBacteria	pBacterc	Bactero	oBacterc				-0.041443	0.0583816	-0.016939	0.0265956	1	19	32	38	
🗢 Genus	10	8	008bb2345da6bffa88ab4245db0d74ba	dBacteria	p_Firmicu (Clostrid	o_Lachno:	f_Lachnos	g_Lachno	s_Lachno	-0.297408	-0.165588	0.4629968	2.3001908	2	24	39	41	GraphicPlus Start
Species	11	9	00acb08eddb1212bdcaef50eccccfd86	dBacteria	p_Firmicu (Clostrid	o_Lachno:	f_Lachnos			-0.865489	-0.165588	1.031077	2.7899426	0	24	28	33	
- sprate	12	10	00ce8cc6ad0083f7a7cef371b71f8e83	dBacteria	p_Firmicu (Bacilli	o_Erysipe	f_Erysipe	gFaecali	sFaecali	-0.39643	-0.165588	0.5620178	1.758645	0	24	33	38	0
	13	11	010ddb7bc606ff747373ce07c940ef6a	dBacteria	pBacterc	Bactero	oBacterc				-0.041443	-0.028262	0.0697046	0.1247591	1	23	37	41	
	14	12	01a26370b5ca9fe7fa78e10095224af8	d_Bacteria	pFirmicu	Clostrid	o_Oscillo:	f_Oscillos	g_UCG-00		-0.041443	0.5785671	-0.537124	1.4787502	1	12	27	26	
	15	13	020f4a9ac9ec34524f8b9a9ec5d6c89b	dBacteria	pBacterc	Bactero	oBacterc	fMuriba	gMuriba	s_Muriba	-0.338506	-0.165588	0.5040942	2.0291622	1	24	36	42	
	16	14	021f18a677c19ad0726e7732c25118d2	dBacteria	pFirmicu	Clostrid	o_Lachno:	f_Lachnos			-0.041443	0.0583816	-0.016939	0.0255373	1	19	35	40	



3. Differential Abundance (DA)

Taxonomy 창 사용 방법

Metagenome 의 Clustering 최소 단위인 ASV(Amplicon Sequence Variant)는 정해진 분류군을 볼 수 있는 직접적인 방법이 아니다. 각 Taxonomy(Phylum, Class, Order, Family, Genus, Species)에서는 원하는 류군만 필터링하여 볼 수 있다. (그림 3-1)

			1						- 1				
- × -				A	В	Ц	C	D	E			H	
			1	Filter: 661		4			Ta	xonomy			
	View All Data	\sim	2	ID 🛒	ASV	-	Kingdo -	Phylui -	Class -	Orde 👻	Famil [®] -	Genu 👻	Specie -
	View All Data	4.1	3	1	321592f084359113f23e3e38237c8093	_	dBacteria	p_Campilo	Campyl o	Campyl	fHelicob	g_Helicob	sHelicob
Taxonor	my		4	2	8373a6f34bdfb5996c716e2c6e764ce2	_	dBacteria	pBacterc	Bactero o	Bacterc	fRikenell	gAlistipe	sAlistipe
· ·	,		5	3	fa387857a369df7328f595391386555d	_	dBacteri;	pBacterc	Bactero o	Bacterc	fRikenell	gAlistipe	
Y Phyl	um		6	4	8d0a3ca24ccba95ef79fb0fbbd91709f	6	dBacteria	p_Campilo	Campyl o	Campyl	fHelicob	g_Helicob	s_Helicob
Class	s			5	e01914f46e26f1aaed24bd00c9c442b8	6	dBacteria	pBacterc	Bactero o	Bactero	fMuribac	gMuriba	sMuriba
Check	TaxonName	1	8	6	c84137e57f8283dbc56df43da161dd1c		dBacteri	pBacterc	Bactero o	Bacterc	fPrevote		
			10	8	82a938698291b68cecfcb9d46457229b	,	dBacteria	pBacterc	Bactero o	Bactero	fBactero	gBactero	
	cCampyiobacteria		12	10	d52e4e06b7fc88a48f5464555c0973f3	,	dBacteria	pBacterc	Bactero o	Bactero	fPrevote		
v	cBacteroidia		13	11	e28749be518256d40762e358d5f48821	,	dBacteria	pBacterc	Bactero o	Bactero	fPrevote	g_Allopre	
	cVampirivibrionia	-	14	12	edaebfecd88812ac54235907a571a16d		dBacteria	pBacterc	Bactero o	Bactero	fMuribac	gMuriba	suncultu
	cVerrucomicrobiae	-	16	14	3bfbf0a80abd58e7c6a16da6dec33154	,	dBacteria	Bacterc	Bacteroo	Bactero	fRikenell	g_Alistipe	s_Alistipe
	cClostridia	-	17	15	af5fe3a72c0bbba02162e60150c31887		d Bacteria	Bacterc	Bacteroo	Bacterc	f Bactero	g Bactero	
	cDeferribacteres	-	19	17	43ad07aa6643c9dd6d7fd103956c27d0		d Bacteria	Bacterc	Bacteroo	Bacterc	f Muribac	g Muriba	s unident
	cBacilli	-	20	18	ecb1f8a7d07e9f1e081f4f5faf2327de		d Bacteria	Bacterc	Bactero o	Bacterc	 f Muribac	g Muriba	s uncultu
	cAlphaproteobacteria	-	21	19	5eafb639b86bf607f38a817820e982a9		d Bacteria	Bacterc	Bactero o	Bacterc	 f Muribac	g Muriba	s uncultu
	cGammaproteobacteri	6	22	20	4dda3a74f2bf9a7d1a5be15fbf8fbc64		d Bacteria	Bacterc	Bacteroo	Bacterc	f Bactero	g Bactero	
	cSaccharimonadia	-	23	21	d60740cdfee329165f0fbbc482494635	-	d Bacteria	Bacterc	Bactero o	Bacterc	 f Rikenell	g Alistipe	
	cDesulfovibrionia	-	24	22	84b1653758e1a862e7b599d66960273c	-	d Bacteria	Bacterc	Bactero o	Bacterc	 f Muribac	g Muriba	s Muriba
	cActinobacteria	-	25	23	34e8c93392eff3a8245f395589dbfcc2		d Bacteria	Bactero	Bacteroo	Bactero	f Rikenell	ø Alistine	
	cCoriobacteriia		26	24	7a8c6354aaa1d4c6f0cb647e9fd91356	Ť	d Bacteria	Bactero	Bacteroo	Bactero	f Muribac	g Muriba	s uncultu
	cNegativicutes		27	25	3af92fc1a7a277a4a8322b86c7149c7a	Ť	d Bacteriu	Bactero	Bacteroo	Bactero	f_Muribac	g Muriba	s uncultu
	cIncertae_Sedis		28	26	08f0f10726o00/a36f5f5d0fb628cc3a	f	d Bactoriu	Bactero	Bacteroo	Bactero	f Muribac	8Wumba	uncultu
<	>		20	28	42404400260254450151500100260058	f	dDacterii	pDactero	Dacter00	Dactero	fNumbac	· · · · · · · · · · · · · · · · · · · ·	• • • • • • • • • • • • • • • • • • • •
			- 50	20	1424001032000263015131035280816/5	1	u Bacteria	p Bacterd	Bacteroo	pactero	i iviuribac	g iviuriba	(s iviuriba)

그림 3-1. Taxonomy selection

가장 왼쪽 상단에 'View All Data' 버튼을 누르면 필터를 모두 해제하여 다시 전체 결과를 볼 수 있다.

Significant Taxon

오른편의 DA Analysis 부분에서 "Significant Taxon" 창은 그룹별로 비교한 결과에서 유의하게 발현 차이가 나는 분류군을 필터링 할 수 있도록 만들어 놓은 것이다. Relative Abundance(상대 풍부도) 와 W(통계 유의성)값을 이용해 필터링하여 각 샘플 및 그룹에 대한 조건이 적용된 결과를 확인할 수 있다. W 값을 설정하면 설정한 값 이상을 만족하는 Taxon 만 필터링 되어 통계적 유의성을 만족하는 분류군에 대한 분석이 가능하다. Relative Abundance 도 설정한 값 이상의 Taxon(ASV)만 필터링하여 낮은 분류군들은 제거하여 분석할 수 있다.

A	В	N	0	P	Q	R	S	Т	U	V	w	Х	Y	Z	AA	AB	1			
Filter: 1	7		w				Sig	ş₩		Average of	Relative Ab	oundance	Average	of Feature	Count					
ID ,	T ASV 🗸	A/B 🗸	A/C 💷	B/C ↓	All 👻	A/B 🗸	A/C -	B/C -	All 👻	C 🖵	Α 👻	B 👻	С 🚽	Α 🚽	В 🚽	A1 👻		DA Analysis		
454	4dda3a74f2bf9a7d1a5be15fbf8fbc64	1	1418	1534	1531	FALSE	TRUE	TRUE	TRUE	2.97%	0.00%	0.00%	750.875	0	0	0.00%				
792	889bfa9e582f8e8b58e7dfbf4f7b66a7	1	1374	1439	39	FALSE	TRUE	TRUE	FALSE	0.00%	0.00%	0.00%	0.5	0	0	0.00%		 Significant Taxon 		
69	0b2f98063db272dc65373defc3dd0907	1	1360	1432	1466	FALSE	TRUE	TRUE	TRUE	0.00%	0.62%	0.36%	0	162.5	82.142857	0.67%		Delative Abundance(%)		
883	98cd478dcbe17c474061139e7505cf0f	0	1302	21	24	FALSE	TRUE	FALSE	FALSE	0.01%	0.00%	0.00%	2.625	0	0	0.00%		Relative Abundance(%	<u></u>	
1047	b4dca0048bafb098b206531033c09819	1	1269	1446	41	FALSE	TRUE	TRUE	FALSE	0.00%	0.00%	0.00%	0.25	0	0	0.00%		0.00	-	
554	5e329e61daa94379223d221363003615	1	1200	1412	1448	FALSE	FALSE	TRUE	TRUE	1.01%	0.00%	0.00%	256	0	0	0.00%		-W		
1180	c6dc8737919138e03a6536d8bba231e6	1	1185	1408	39	FALSE	FALSE	TRUE	FALSE	0.00%	0.00%	0.00%	0.5	0	0	0.00%		1000	÷	
1137	c0bc5dde25d896d271f66c9cbe3d6d62	1	1170	1396	37	FALSE	FALSE	TRUE	FALSE	0.00%	0.01%	0.00%	0	3.5	1	0.00%		Comela (Control		
1254	d29ed019f8f94bdeb975627ccbd98d0a	1	1162	1406	45	FALSE	FALSE	TRUE	FALSE	0.00%	0.00%	0.00%	0.25	0	0	0.00%		Sample / Control		
942	a382f914b74187cc027862e9232206b2	1	1156	1394	40	FALSE	FALSE	TRUE	FALSE	0.00%	0.00%	0.00%	0	0	0.8571429	0.00%		Name Name	_	
1182	c6e4e41126581c5cf596b4d23c8218ce	1	1133	1319	22	FALSE	FALSE	FALSE	FALSE	1.87%	0.54%	0.51%	473.125	142.5	116	0.03%		A/6	- 1	
141	194f6fa35dfc9b0650eca906507a486b	1	1130	1321	1338	FALSE	FALSE	FALSE	FALSE	0.30%	0.00%	0.00%	74.75	0	0	0.00%		V A/C		
450	4cb800503906a9e601d50141060e8a49	1	1122	1307	1340	FALSE	FALSE	FALSE	FALSE	0.32%	0.00%	0.00%	80.25	0	0	0.00%		B/C		
1426	ef83a8dcd040566537b1e812378fc048	1	1122	1343	41	FALSE	FALSE	FALSE	FALSE	0.00%	0.00%	0.00%	0.375	0	0	0.00%		All		
1471	f6e2f231f56930f016a3edfd8706a60c	1	1082	1261	94	FALSE	FALSE	FALSE	FALSE	0.13%	0.19%	0.00%	32.25	49	0	0.12%				
635	6bdec4eb36e22f5346fd4b79d3b4636d	1	1020	1329	41	FALSE	FALSE	FALSE	FALSE	0.00%	0.00%	0.00%	0.25	0	0	0.00%				
518	58c9269c2280f20bb254b4ebf0d2bd1a	1041	1017	43	1192	TRUE	FALSE	FALSE	FALSE	0.00%	0.09%	0.00%	0	24.5	0	0.14%			_	

그림 3-2. Significant Taxon

예를 들어 A / C 를 기준으로 W 값을 1000 으로 설정하여 총 17 개의 ASV 가 필터링 된 것을 확인할 수 있다.

Pie Chart

Pie Chart 를 활용하여 각 단일 샘플 및 그룹에 대한 6 가지 Taxon 별 분포 비율을 알 수 있다. ASV 단위가 아닌 실제 Taxon 별로 합산한 상대적 비율을 Pie Chart 로 확인할 수 있다.



Phylum_Taxon

그림 3-3. Pie Chart

그림 3-3 은 단일 샘플에 대한 Phylum 에 대한 예시 파이 차트이다. 샘플 및 그룹에 대한 필터링이 적용된 풍부도를 알 수 있으며, Chart 내 taxa 를 클릭하면 필터링 된 결과만 엑셀창에서 확인할 수 있다. 단, 한 샘플만 선택해야한다.

Volcano Plot

Volcano Plot을 이용하여 선별된 ASV를 기준으로 시각화 가능하다.

"Graph View"를 클릭하면 모든 그룹에 대한 Plot 이 자동 생성된다. 생성된 그래프는 마우스 스크롤로 크기 조절할 수 있다. 원하는 임계값(SigW)을 기준으로 위는 빨강, 아래는 초록색으로 표시가 가능하다.



그림 3-4. Volcano Plot

그림 3-4 는 'B / C' 그룹비교를 예시로 한 Volcano plot 이다. Sig W 값을 500 을 기준으로 유의성 있는 ASV 를 선별할 수 있다. 오른쪽 W 값을 수정하여 임계값(가로선)에 따른 원하는 그래프 제작이 가능하다.

Venn Diagram

원하는 샘플 및 그룹에서 ASV 에 대한 벤다이어그램을 제작할 수 있다. 최대 4 개까지 비교 가능하며, 선별 기준은 Relative Abundance 만으로 필터링할 수 있다.



그림 3-5. Venn Diagram

그림 3-5 는 A, B, C 에 대해 Relative Abundance 1%로 선별하여 제작된 그림이다. 각 그룹별로 1% 이상인 ASV 행만 선별하여 공통된 region 과 그렇지 않은 region 에 대한 벤다이어그램이 제작된다. 특정 region 마우스 오른쪽 버튼으로 'Region All'을 클릭하면 해당 region 의 ASV 행만 선별하여 Excel 창에서 보여준다.

4. Graphic Plus

DA Analysis 부분에서 "GraphicPlus Start" 창을 펼치면 작동한다. Metagenome 분석 시각화에 대표적으로 사용되는 Bar Plot, Krona Chart, PCoA, Clustering heatmap 을 제작할 수 있다.

ExMEGA Graphic Plus v1.0.0	?	×
ExMEGA_Report_V2.3.xlsx		
Bar Plot Krona Chart PCoA Clustering Heatmap		
Bar Plot Data Input		_
Single Data		
W Value Select Filtering Data Column 0 ● Relative Abundance ○ A / B 0,00 % ● Average of RA ○ A / C 0,00 % ●	1	
Single Data ColumnGroup Data ColumnA1CA2AA3BB1B2		
Select All Data Column Select All Data Column	umn	

그림 4-0. Graphic Plus

모든 그래프는 단일 샘플 비교(Single Data)와, 그룹 비교(Group Data)가 모두 가능하도록 설계되어 있다. W value 를 적용할 비교조합을 선택하여야 하며, All Column 은 모든 그룹에 대한 비교로, 기본적으로 제공하는 Data Column 이다. 만약 그룹 분석이 아닌 데이터나 통계분석을 원하지 않을 경우 All Column 을 체크하여 W value 를 0으로 설정하면 Relative Abundance 만으로 분석할 수 있다. 'A / B' Data Column 과 같이 특정 비교조합을 선택하면 A 와 B 에 속하지 않은 샘플들은 통계적 유의성을 가지지 못하기 때문에 A 와 B 그룹에 속한 샘플들만 분석하는 것이 좋다.

Bar Plot

원하는 샘플 및 그룹에 대한 조건을 만족하는 Bar Plot을 제작한다.



그림 4-1. Bar Plot

Bar Width 를 통해 Bar 크기를 조절할 수 있으며, Taxonomic 탭에서 모든 Taxonomy(총 7 가지 분류군)에 대한 그래프를 모두 확인할 수 있다. Bar Plot 을 필터링 없이 제작하면 낮은 Taxon 들로 시각화에 방해될 수 있다. 보통 Relative Abundance 값을 0.1~1%로 설정하여 낮은 분류군을 필터링하는 방법으로 많이 사용한다.

Krona Chart

KronaTools를 사용한 Krona chart는 확대/축소가 가능한 multi-layered pie chart이다. 이바이오젠에 선 필터링 조건을 적용한 Krona Chart를 매크로 방법으로 제작한다. 여러 샘플을 선택할 경우 Chart를 그리는데 시간이 오래 소요될 수 있다. 사용 방법과 자세한 설명은 해당 매뉴얼 페이지 에 자세히 설명되어 있다. <u>https://github.com/marbl/Krona/wiki/Browsing-Krona-charts</u>



그림 4-2. Krona Chart

PCoA

PCoA(Principal Coordinates Analysis)는 Metagenome 에 사용되는 대표적인 Beta diversity(샘플 간 분석) 분석으로 샘플간 유사성을 거리로 계산하는 차원 축소 개념이다. PCoA 차원은 2D/3D Type 중 선택하며, Metric type 은 distance matrix 알고리즘으로 대중적으로 사용하는 braycurits 와 jaccard 방법 중 선택하여 사용할 수 있다. Marker Type 을 Group 으로 설정하면 각 샘플에 대한 그룹 정보도 확인할 수 있다. "Save Value Options" 탭에서는 PCoA 에 활용한 각 수치 값들을 파일로 저장할 수 있다. PCoA 도 다른 그래프와 마찬가지로 필터링을 적용하여 유의미한 Taxon 분석할 수 있다.

Bar Plot Krona Chart PCoA	Clustering Heatmap						
PCoA Data Input							
<u>Data Type</u>	Metric Type						
🗿 <u>Single</u> 🔘 <u>Group</u>	• <u>braycurits</u> • jaccard						
PCoA Type	Marker Type						
O 2D ○ 3D	Single O Group						
Sample Cluster	Coloct Filtering Date Colump						
W Value							
V 💌 Relative Abundance							
	Ğ₿∕Č						
Average of RA	🗆 All						
0,00 % 🖨							
Single Data Column	Group Data Column						
□ A1							
🗆 A2 🔰	A A						
□ A3	B						
Select All Data Column	Select All Data Column						
Savo Valuo Ontions							
Save PC Values							
 Save Eigen Values 	Save Values						
Save Distance Values							
)						
Draw	PCoA Plot						

그림 4-3-A. PCoA



그림 4-3-B는 2D로 분석한 샘플들 간의 PCoA 결과이다. 각 샘플들은 A / B / C 샘플들끼리 유사함을 알 수 있다.

Clustering Heatmap

Clustering Heatmap 은 ASV 마다 Relative Abundance(%) 비교를 시각화 할 수 있다. Relative Abundance 값은 0.05~0.5%까지 추천하며, 너무 높은 값을 필터링하면 많은 ASV 가 제거되어 그래프가 제작되지 않는다. W 값을 기준으로 필터링하는 방법도 있다 (**그림 4-4-B).** Volcano plot 등을 이용하여 일정 기준 W 값에 대한 ASV 들을 확인하면, 해당 ASV 만으로 풍부도를 시각화 할 수 있다.



그림 4-4-A. Clustering Heatmap



그림 4-4-B. Clustering Heatmap2