Hierarchical Clustering and PCA Plots in Skyline Daily

January 8, 2021 Nicholas Shulman

"Cluster" button on the Document Grid



- A "Cluster" button appears on the Document Grid (and related windows) in the newest release of Skyline-Daily.
- When you push that button, Skyline will choose a set of numeric columns to perform hierarchical clustering on
- The clustering works best if some numeric columns have been pivoted by replicate

What is hierarchical clustering?

- Finds which points are closest to each other, and then finds which points or groups of points are closest to each other.
- In this case, A&E are close to each other, and B&D are close to each other. After that, C is closer to B&D than it is to A&E.



What hierarchical clustering looks like in the Document Grid

- Rows get reordered so that the dendrogram can be drawn without the lines crossing
- Values in "Text" column get colored arbitrarily
- Values in "X" column get colored based on how far they are from zero





Group Comparisons

• The Group Comparison grid in Skyline-Daily now has an additional report available "Clustered".

Healthy/Diseased:Grid			
Volcano Plot Bar Graph			
Reports - 🞲 - 👫 - 🕅		1	of 12
✓ default		e	Fo
Clustered		IPWAI TE	13
Customize Report		EVSVACK	1.0
Manage Reports		DAGAA.	. 0.7
NID 600000	GSVM		10

• Also, pushing the colorful cluster button automatically switches the report from "default" to "Clustered.

Clustered Group Comparison Report

- The "Clustered" comparison report displays the replicate abundances in separate columns.
- Hierarchical clustering is performed on the replicate columns, as well as the peptide or protein rows.

										_
										_
	Protein	Peptide	Diseased Fold Change	Diseased Adjusted P-Value	H_159_REP1 Abundance	H_159_REP2 Abundance	H_146_REP1 Abundance	H_146_REP2 Abundance	H_160_REP1 Abundance	H_16 Abun
	<u>s</u>	AFGLSSPR	1.97 (95% CI:1.3	0.0111	1.8478E+0	1.5568E+0	2.5389E+0	2.1798E+0	1.7837E+0	1.341
	NP 620183	LGPLVEQGR	3.68 (95% CI:1.9	0.0037	8.0821E-1	3.2571E-1	1.3095E+0	4.4697E-1	3.3654E-1	2.172
T	<u>s</u>	HLNGFSVPR	2.91 (95% CI:1.4	0.0212	9.2903E-2	1.2024E-2	5.3223E-1	6.2910E-2	2.5264E-2	1.022
	XP_001067936	SPOGLGASTAE	1.81 (95% CI:1.0	0.0550	4.7749E-3	5.8507E-3	2.8985E-3	2.9671E-3	3.1524E-3	2.858
	NP 665722	AEQGAYLGPLP	2.02 (95% CI:1.0	0.0560	2.0482E-2	2.4298E-2	1.3541E-2	1.2914E-2	1.3276E-2	1.417
	NP 665722	AISYLISGYOR	1.9 (95% CI:1.17	0.0318	5.8533E-1	6.0935E-1	3.1228E-1	3.4984E-1	3.8154E-1	4.241
	NP_001008724	GLIDEANQDFT	1.47 (95% CI:0.9	0.1418	1.2167E-2	1.5085E-2	7.9583E-3	8.0401E-3	6.1805E-3	6.375
14	NP_001008724	SQLQEGPPEWK	1.47 (95% CI:1.0	0.0550	1.1370E+0	1.0727E+0	8.3636E-1	7.0021E-1	7.9246E-1	8.0406
	NP 001008724	FGSLTSNFK	1.47 (95% CI:1.0	0.0482	2.1322E+0	2.2287E+0	1.4692E+0	1.4715E+0	1.5530E+0	1.5888
	NP 602308	GSYNLQDLLAQ	1.26 (95% CI:0.7	0.3850			4.8459E-3		4.3424E-3	
	XP_001057320	SGYTFOLLR	1.31 (95% CI:0.9	0.1866	1.5309E-1	1.4194E-1	1.2671E-1	1.2331E-1	1.0386E-1	1.266
	<u>XP_001057320</u>	EENGDFASFR	1.25 (95% CI:0.8	0.2370	4.7771E-2	4.6964E-2	4.8843E-2	4.7385E-2	3.8908E-2	3.773
	NP 036629	CSLPRPWALTF	1.35 (95% CI:0.8	0.2186			6.0454E-3		5.3985E-3	
1 1	NP 001012027	TSDQIHFFFAK	1.3 (95% CI:0.84	0.2615	7.0041E-3	9.1546E-3	6.9477E-3	6.9862E-3	5.1374E-3	6.890
	NP 001012027	LOPLDFK	1.42 (95% CI:0.9	0.1657	3.2153E-1	3.2655E-1	2.6015E-1	2.4869E-1	2.3815E-1	2.166
	NP_037001	TACVLPAPAGP	1.22 (95% CI:0.9	0.1533	1.0983E-2	1.3596E-2	9.8888E-3	1.1332E-2	1.0416E-2	1.046
	NP 062016	ATGVETTLQPLR	2.28 (95% CI:1.1	0.0505	1.6361E-2	1.5013E-2	7.4137E-3	8.5737E-3	6.7845E-3	8.137
	NP 062016	WLPSTSPVTGYR	1.98 (95% CI:1.1	0.0497	3.5215E-2	3.5632E-2	2.1170E-2	1.9682E-2	2.0043E-2	1.998
	NP 001012027	SQLPGIIAEGR	1.8 (95% CI:1.12	0.0392	9.6843E-3	8.8378E-3	8.5799E-3	9.3365E-3	6.5883E-3	7.672
	NP 113692	ATETQGVNLLF	2.01 (95% CI:1.2	0.0287	7.0882E-2	6.8975E-2	4.7786E-2	4.7447E-2	4.6707E-2	4.744
	NP 036869	LNHQMEGLAF	1.95 (95% CI:1.0	0.0604	1.1459E-2	1.2637E-2	6.9926E-3	8.0969E-3	8.1634E-3	8.987
	NP_001007697	NLGVVVAPHALR	2.1 (95% CI:1.18	0.0336	2.0991E-2	2.1025E-2	1.5042E-2	1.2492E-2	1.3491E-2	1.419
	NP 445770	CNADPGI SALL	3 44 (95% CI-1 4	0.0292	0.0242E-2	1 25655.2	5 0017E-2	5 204CE-2	6 72765.2	7 270

Displaying the heat map graph

 There is a dropdown menu item from the cluster button which brings up a heat map graph.



