Next-Generation Clustered Heat Maps (NG-CHM)

Bradley Broom and John Weinstein, Department of Bioinformatics and Computational Biology, UT MD Anderson Cancer Center

A visualization tool for the dynamic exploration of large, omic data sets.

Visualization Features

• Easy-to-use navigation
• Flexible zoomed views for inspecting small details in large data sets
• Generate publication-quality graphics
• Scalable
• Highly interactive
Dynamic “Link-outs”

- Can link-out from selected **rows, columns, or matrix elements**
- To other content specific resources,
  - E.g. NCBI, Genecards, Cancer Digital Slide Archive
- Content specific
  - Associate ‘types’ with row and column labels
  - ‘Link-out’ database indexed by type
  - Only matching link-outs included in a specific NG-CHM
- Or to dynamic analyses
  - E.g. Box Plot, Points Plot
- Five major types of resource links
An Exploratory Environment with five major types of resource links
Creating NG-CHMs

Four “builder” levels:

– GUI
  • New non-quantitative user
  • Experienced non-quantitative user

– R
  • Quantitative user

– Specification files
  • System developer
RECENT DEVELOPMENTS
New architecture NG-CHMs

• Previous NG-CHM system architecture
  – Was designed six years ago
  – Relied heavily on back-end server for tile rendering
  – Did not exploit recent advances in web browsers
  – Had performance and scaling issues

• New NG-CHM system architecture
  – Complete re-design to take advantage of new browser capabilities (e.g. WebGL)
  – “Tile Generation” performed entirely in browser
  – Far more responsive user experience
New architecture NG-CHMs

• “Standalone” server-less mode possible
  – Nearly-all NG-CHM functionality is present
    • Cannot “link-out” to server-side dynamic computations
  – Allows NG-CHMs to be:
    • Saved locally (e.g. with other project data)
    • Attached as supplementary material to papers
    • Shared with colleagues via email, web sites, portals, etc.
New architecture NG-CHMs integrated with Galaxy

• Added new architecture NG-CHMs to Galaxy:
  – BigQuery Interface (for ISB-CGC data access)
  – NG-CHM Builder tool
  – NG-CHM Visualization

• Available
  – As a ready-to-run Docker container
  – From the Galaxy Toolshed
  – From Github

• Exposes NG-CHMs to a new user community
Galaxy NG-CHM Builder

MDA Heatmap Generation

Create HeatMap from Dendogram and Matrix (Galaxy Tool Version 5.2.4)

**Input Matrix**

- 2: Mean-Center a Matrix on data 1

**User Defined HeatMap Name**

- HM_name

**Optional HeatMap Description**

- 

**Data Summarization Method**

- average

**Row ordering method**

- Hierarchical Clustering
  - Choices: Hierarchical Clustering, Original Order, Random

**Row Distance Metric**

- Euclidean, binary, manhattan, maximum, canberra, minkowski, or correlation

**Row Clustering Method**

- Ward
  - Choices: 'average' for Average Linkage, 'complete' for Complete Linkage, 'single' for Single Linkage, 'ward', 'mcquitty', 'median', or 'centroid'.

**Column ordering method**

- Hierarchical Clustering
  - Choices: Hierarchical Clustering, Original Order, Random

**Column Distance Metric**

- Euclidean, binary, manhattan, maximum, canberra, minkowski, or correlation

**Column Clustering Method**

- Ward
  - Choices: 'average' for Average Linkage, 'complete' for Complete Linkage, 'single' for Single Linkage, 'ward', 'mcquitty', 'median', or 'centroid'.

**Covariate Bars**

- Insert Covariate Bars

- Execute
Galaxy NG-CHM Visualization Tool
New architecture TCGA NG-CHM Compendium

http://tcga.ngchm.net/
Numerous Improvements to NG-CHM Viewer

• More user options for fine tuning PDF generation
• Improved rendering responsiveness for large NGCHMs
• Ability to create very large heat maps (> 20,000 rows)
  – Contributed dendrogram code to R project (in R 3.4.0)
• Added new features to NG-CHMs:
  – Can add “cuts” (extra blank rows/columns) to heat map
  – Can highlight a small number of top rows/columns in overview pane
• Numerous bug fixes and other improvements
• Created additional tutorial videos
UPCOMING PLANS
Upcoming plans

• Release new architecture NG-CHM manager
• Complete our new architecture NG-CHM compendium
• Further increase viewer responsiveness for very large data matrices
• Develop a tablet-optimized user interface
• Collaborate
  – Add links to/from NG-CHMs from/to other tools
  – Create more advanced tools & data exchange
Credits

MD Anderson Dept of Bioinformatics and Computational Biology
Bradley Broom
John Weinstein
Rehan Akbani
Chris Wakefield
James Melott

In Silico Solutions
Michael Ryan
Robert Brown
Futa Ikeda
Mark Stucky

Hobsons
David Kane