



## Data input type formatting guide

**Header titles should be labeled as illustrated in our examples below (ie chrom1, not chromosome1). Additional headers will be ignored by DNA Rchitect.**

### File format

	Required headers
bed	chrom, start, stop
bedpe	chrom1, start1, end1, chrom2, start2, end2, score, samplenumber
bedgraph	chrom, start, stop, value

1. Bedpe format for visualizing arcs from HiC, HiChIP, Promoter Capture HiC, ChIA-PET etc data types.

chrom1	start1	end1	chrom2	start2	end2	name	score	FDR	strand1	strand2	samplenumber
chr6	129208442	129220492	chr6	129231288	129240143	NA	210	0.00	.	.	1
chr2	58151949	5815732	chr2	58158964	58165186	NA	150	0.00	.	.	1
chr7	28371327	28377361	chr7	28377688	28386880	NA	135	0.00	.	.	1
chr7	83649993	83657438	chr7	83659817	83662316	NA	132	0.00	.	.	1
chr9	32707989	32712792	chr9	32714366	32728804	NA	126	0.00	.	.	1
chr2	127310094	127315729	chr2	127316794	127324916	NA	121	0.00	.	.	1

## Required columns and heading titles

2. Bedgraph format for visualizing bp resolution signal from ATAC-seq, ChIP-seq, RNA-seq, or other sequencing data types. Values must be non-negative values.

chrom	start	stop	value
chr1	3010964	3010966	4
chr1	3012331	3012337	4
chr1	3015732	3015736	4
chr1	3015736	3015738	5
chr1	3015738	3015739	6
chr1	3015739	3015745	5

## Required columns and heading titles

3. Bed format for visualizing genomic segments from ATAC-seq, ChIP-seq, or other sequencing data types.

chrom	start	stop
chr2	98666982	98667306
chr17	39847589	39847913
chr3	5860289	5860613
chr12	3109799	3110123
chr9	123461860	123462184
chr9	3000396	3000720

## Required columns and heading titles

## Suggestions for visualization of HiC and other chromatin conformation data using DNA Rchitect:

- Pre-filtering of highly reproducible or significant interactions may be required for proper visualization using DNA Rchitect.
- Large anchors or overlapping anchors may need to be adjusted or merged (can use GenomicRanges R package for this) prior to loading in DNA Rchitect to ensure proper network visualization and interpretation.

## Google Chrome web browser

The screenshot shows the DNA Rchitect web interface. At the top, a dropdown menu lists genomes: mouse\_mm9, mouse\_mm10, human\_Hg19, human\_Hg38, and drosophila\_melanogaster\_6. A button for "Interactive Tutorial" is visible. Below the genome selection, a message says "Genomes supported by web version. Select human Hg19 for tutorial." Three steps are outlined:

- Step 1: Select the types of data you want to analyze**
- Step 2: After browsing for your files, click the button to process the data for plotting**
- Step 3: Make sure your data looks correctly formatted in the tabs below. Then, click on the Plots tab to visualize your data**

On the left, there's a "Select Data Types" section with "HIC" selected. In the center, a "Process Data" button is shown. On the right, a "Go to Plots" button is shown. The bottom section shows "Load Sample Data" with a dropdown menu containing sample names like K562\_NHEK\_GM12878-loops and file separator options (Comma, Semicolon, Tab). A note states: "Data upload selection options available for visualization. Accommodates bed, bedpe, and bedgraph formats for data types like HiC, ChIP-seq, ATAC-seq and RNA-seq. Upload data types from local directory."

Publically available sample HiC and Promoter Capture HiC data

Select **K562-NHEK-GM12878** HiC data and select **Process Data** button for tutorial

The screenshot shows the DNA Rchitect web interface after processing the K562-NHEK-GM12878 HiC data. The "Process Data" button is now active. The "Go to Plots" button is also present. The "Load Sample Data" section shows the selected sample: K562\_NHEK\_GM12878-loops. To the right, a table displays the processed data:

chrom1	start1	end1	chrom2	start2	end2	score	sampenumber
chr10	100180000	100190000	chr10	100410000	100420000	39	1
chr10	101600000	101610000	chr10	101800000	101810000	93	1
chr10	102100000	102105000	chr10	102190000	102195000	59	1
chr10	102100000	102105000	chr10	102265000	102270000	40	1
chr10	102190000	102200000	chr10	102260000	102270000	90	1
chr10	102800000	102810000	chr10	102890000	102900000	78	1

Selecting **Go to Plots** allows you to visualize generate plots tab

The screenshot shows the "Plots" tab configuration. It includes fields for "Number of Samples in HiC Dataset" (set to 1), "HiC Sample 1" (Sample ID), "Chr Number" (chr1), "Start Coordinate" (60853778), and "End Coordinate" (60948071). A "Submit Parameters" button is at the bottom. The "Search By:" section has "Coordinates" selected. The "Select colors for bezier curves" section shows a dropdown menu with color options: Blue, Red, Green, Purple, Orange, Yellow, Brown, and Dk-v.

Annotations provide instructions:

- "Select number of samples (3 for tutorial)" points to the "Number of Samples in HiC Dataset" field.
- "Select colors for 3 samples" points to the "Select colors for bezier curves" dropdown menu.

At the bottom, navigation tabs include Network, Hubs, Degree Distribution, Groups, and Node Degree.

Plots

Number of Samples in HiC Dataset

Search By:  
 Coordinates  
 Genes

Number of samples

Sample Labeling  
 HiC Sample 2: NHEK  
 HiC Sample 3: GM12878  
 Select colors for bezier curves: Blue Red Gray

Type gene name: (backspace to clear)  


MYCL chr1:40361098-40367928  
 MYCN chr2:16080686-16087129  
 MYCBP chr1:39328636-39347289  
 MYC chr8:128747680-128753674  
 MYCNUN chr2:16060521-16076139

Type gene name: (backspace to clear)  


Genomic interval to left of gene (bases)  


Genomic interval to right of gene (bases)

Submit Parameters

Select the gene MYC. Set Genomic intervals left and right of gene to 1Mb (1000000) for tutorial

[Network](#)
[Hubs](#)
[Degree Distribution](#)
[Groups](#)
[Node Degree](#)

Selecting **Submit Parameters** allows you to visualize your data

HiC contacts for 3 human cell lines near MYC locus

Submit Parameters

Download Bezier curves as PDF

Download Bezier curves as SVG

Interactive network of visualized loops

Nodes may be selected and toggled

Network Hubs Degree Distribution Groups Node Degree

Download as HTML

Download Network as XGMML

Download All as PDE

SVG, PDF, and network files can be readily downloaded or exported into other tools

Various **network analysis** can be readily viewed and downloaded for this genomic locus

Hubs

chr8:12875000-128600000  
 chr8:12875000-128755000  
 chr8:129325000-129350000  
 chr8:12965000-129675000  
 chr8:12875000-128760000  
 chr8:128730000-128740000  
 chr8:12856000-128570000  
 chr8:12789000-127900000  
 chr8:12818000-128200000  
 chr8:12820000-128230000  
 chr8:12818000-128250000  
 chr8:12874000-128750000  
 chr8:12800000-128050000  
 chr8:12919000-129200000  
 chr8:12830000-128325000

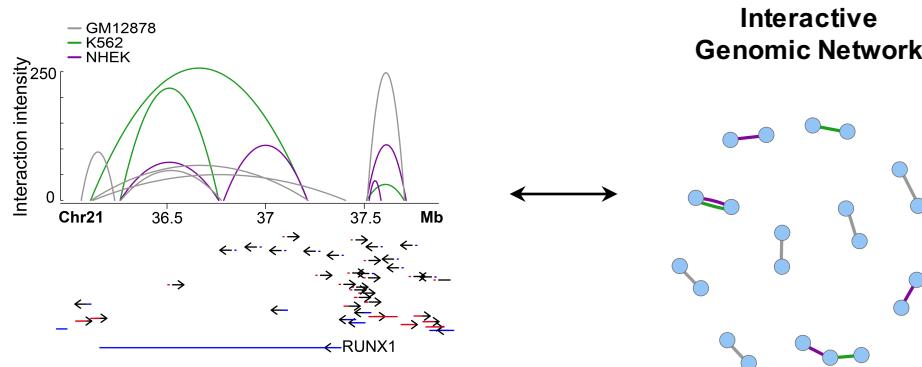
Community Detection

Node Degree Density Plot

## Network analysis Guide

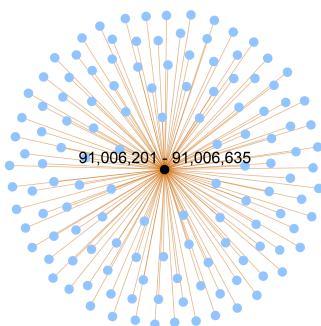
This guide details information on how network analysis can be useful for biological interpretation of HiC data using DNA Rchitect

### Genomic visualization of HiC contacts



### Interactive network:

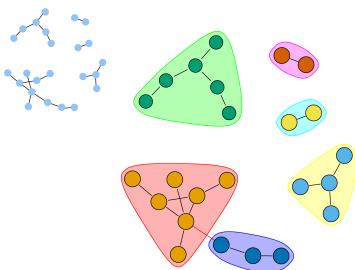
- Useful for selecting interacting nodes (ie. DNA elements such as enhancers/promoters) which are interacting across any genomic window or gene loci.
- Quick selection of DNA nodes that have high numbers of interactions (ie. Super enhancer-like function). See hub network analysis panel for additional information.



### Promoter-capture interactive network example:

- We demonstrate visualization and use of promoter-capture HiC data to illustrate a one-to-many view of DNA contacts. The central node is the promoter with edges drawn to all interacting DNA elements captured using this 3C-technology.
- Similarly, 3C, 4C, 5C contacts which may also include one-to-many type contacts can be visualized using DNA Rchitect.

### Community detection



### Community detection:

- Analysis that takes a 'birds-eye view' of HiC contacts and resolves clusters of HiC communities.
- Can be useful in the interrogation of enhancer-enhancer, enhancer-promoter, and promoter-promoter contacts both within and across clusters.
- An example use of community detection can be seen in:  
Babaei S, Mahfouz A, Hulsman M, Lelieveldt BPF, de Ridder J, et al. (2015) Hi-C Chromatin Interaction Networks Predict Co-expression in the Mouse Cortex. PLOS Computational Biology 11(5): e1004221. <https://doi.org/10.1371/journal.pcbi.1004221>

Upload File    Plots

**Step 1:** Select the types of data you want to analyze, then browse for your files  
**Select Data Types**

**Step 2:** After browsing for your files, click the button to process the data for plotting

**Step 3:** Make sure your data looks correctly formatted in the tabs below. Then, click on the Plots tab to visualize your data

**Select ATAC data format**  
 Bed  
 Bedgraph

**Choose a TXT, CSV, or BG File**  
 No file selected

Header  
**Separator**  
 Comma  
 Semicolon  
 Tab  
**Quote**  
 None  
 Double Quote  
 Single Quote

**Data type is ATAC-seq with options for Bed or Bedgraph format. Same for ChIP-seq data.**

**Step 1:** Select the types of data you want to analyze, then browse for your files  
**Select Data Types**

**Step 2:** After browsing for your files, click the button to process the data for plotting

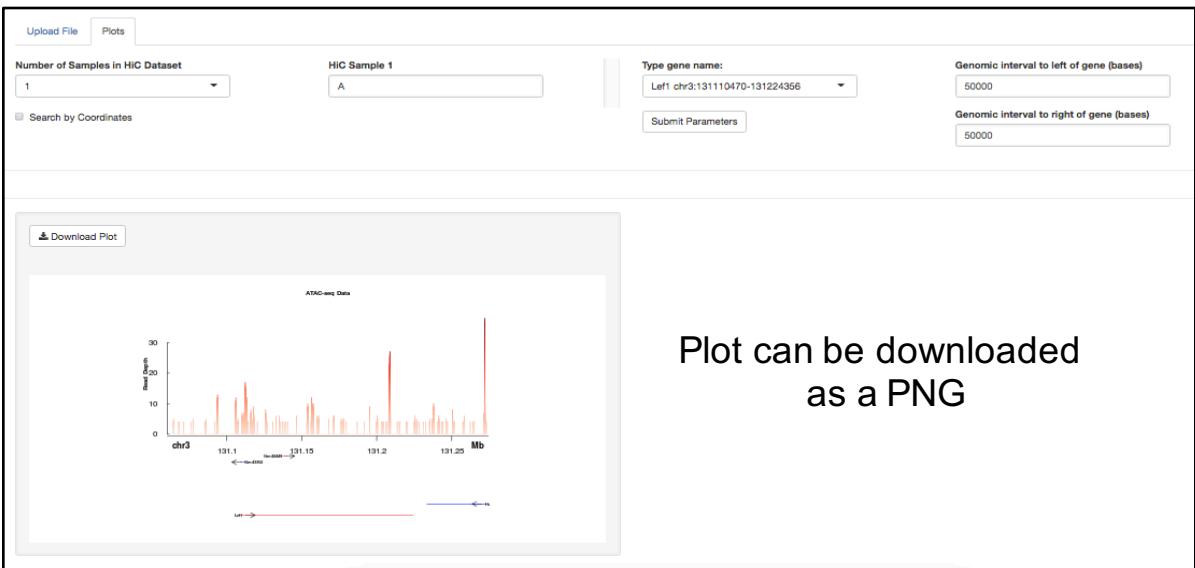
**Select ATAC data format**  
 Bedgraph

**Choose a TXT, CSV, or BG File**  
 ATAC.bg

Header  
**Separator**  
 Comma  
 Semicolon  
 Tab  
**Quote**  
 None  
 Double Quote  
 Single Quote

chrom	start	stop	value
chr1	3010964	3010966	4
chr1	3012331	3012337	4
chr1	3015732	3015736	4
chr1	3015736	3015738	5
chr1	3015738	3015739	6
chr1	3015739	3015745	5

ATAC-seq bedgraph upload successful.  
**Process Data** allows one to view data processed and ready for plotting.



Select **sample number**, and either **gene** or **gene coordinates** and press **Submit Parameters**  
For this example ATAC-seq bedgraph file is shown for the *Lef1* gene.

Google Chrome web browser

Upload File Plots

**Step 1:** Select the types of data you want to analyze, then browse for your files  
**Select Data Types**

**Step 2:** After browsing for your files, click the button to process the data for plotting

**Step 3:** Make sure your data looks correctly formatted in the tabs below. Then, click on the Plots tab to visualize your data

Select ChIP data format

Data type selected is ChIP-seq in Bed format

ChIP

Choose a TXT, CSV, or BG File  
 No file selected

**Step 1:** Select the types of data you want to analyze, then browse for your files  
**Select Data Types**

**Step 2:** After browsing for your files, click the button to process the data for plotting

**Step 3:** Make sure your data looks correctly formatted in the tabs below. Then, click on the Plots tab to visualize your data

Select ChIP data format

ChIP

Choose a TXT, CSV, or BG File  
 Peaks.txt

Header  
**Separator**  
 Comma  
 Semicolon  
 Tab

chrom	start	stop
chr2	9866982	98667306
chr17	39847589	39847913
chr3	5860289	5860613
chr12	3109799	3110123
chr9	123461860	123462184
chr9	3000396	3000720

ChIP-seq bed upload successful.  
**Process Data** allows one to view data processed and ready for plotting.

Upload File Plots

Number of Samples in HiC Dataset

HIC Sample 1

Type gene name:

Genomic interval to left of gene (bases)

Search by Coordinates

Submit Parameters

Genomic interval to right of gene (bases)

Plot can be downloaded as a PNG

ChIP-seq Data

Pressing the **Process Data** button allows for visualization of data uploaded. Please ensure that headings follow our **bed** required formatting in order to proceed with plotting.

# Troubleshooting and error handling

1. Error will appear if Bedpe file format is incorrect.

2. Error shown of HiC plotting. This is due to the anchor plotting boundary. This can be resolved by adjusting the genomic window coordinates.

Upload File Plots

Number of Samples in HiC Dataset: 2

HiC Sample 1: A

Type gene name: IL10 chr1:131019844-131024974

Genomic interval to left of gene (bases): 50000

HiC Sample 2: B

Search by Coordinates

Submit Parameters

Genomic interval to right of gene (bases): 50000

[Download Plot](#)

Error: Current genomic window cannot be plotted, probably because an anchor crosses the plot boundary. Adjust genomic window coordinates (zoom in or out) and re-submit

[Download as PNG](#)



3. Error in cytoscape network graph. No graph is shown as this genomic window does not contain any data.

Upload File Plots

Number of Samples in HiC Dataset: 2

HiC Sample 1: A

Type gene name: IL2 chr3:37120522-37125959

Genomic interval to left of gene (bases): 50000

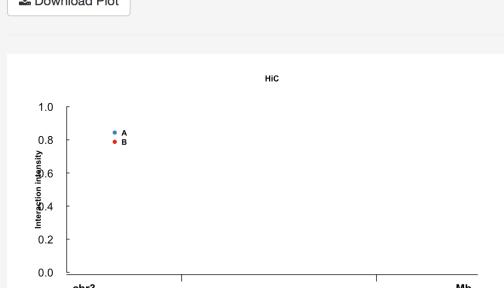
HiC Sample 2: B

Search by Coordinates

Submit Parameters

Genomic interval to right of gene (bases): 50000

[Download Plot](#)



[Download as PNG](#)

Error: The genomic window does not contain any nodes