

# iGenomicViewer: R package for visualization of high dimension genomic data

Daniel P. Gaile<sup>1,2,3,\*</sup>, Lori A. Shepherd<sup>1,2,3,\*</sup>, Andrew E. Bruno<sup>4</sup>, David Gold<sup>1,2,3</sup>, Song Liu<sup>1,2,3</sup>, Carl D. Morrison<sup>6</sup>, Norma J. Nowak<sup>2,5</sup>, Lara E. Sucheston<sup>1,2,3</sup>, and Jeffrey C. Miecznikowski<sup>1,2,3,†</sup>

<sup>1</sup>*Department of Biostatistics, University at Buffalo, Buffalo, NY 14214-3000* <sup>2</sup>*New York State Center of Excellence in Bioinformatics and Life Sciences, Buffalo, NY* <sup>3</sup>*Department of Biostatistics, Roswell Park Cancer Institute, Buffalo, NY* <sup>4</sup>*Center for Computational Research, University at Buffalo, Buffalo, NY* <sup>5</sup>*Department of Biochemistry, University at Buffalo, Buffalo, NY* <sup>6</sup>*Department of Pathology, Roswell Park Cancer Institute, Buffalo, NY*

---

---

### 3. Background

The analysis of high dimensioned data, particularly those generated by advanced biotechnologies, is not guaranteed to provide small dimension results. Each component of a set of 'results' may have assay and sample specific attributes which are of potential interest to the collaborative research team investigating a scientific question. In addition, because current science necessitates multi-disciplinary and multi-institutional research teams, it is imperative results be communicated clearly and efficiently via a platform independent vehicle that allows for a custom viewing experience.

The iGenomicViewer R application provides interactive heatmap plots with tool-tip content and customizable annotation for genomic data. The software provides interactive heatmap representations of assay value data matrices with matrix columns corresponding to assayed samples, matrix rows corresponding to particular probes (genes, SNPs, etc), matrix entries corresponding to observed assay values, and tool-tip data displays that are customizable. An accompanied annotation representation with tool-tip capabilities may be customized for display. The tool-tip displays enhance data integration by allowing inclusion of links and images from other sources. Surveys of software for the analysis of genomic data such as microarrays are provided in 2 and 4. Other software packages currently do not provide all the functionalities of iGenomicViewer. It combines and extends useful functionality of R functions and packages such as image and sendplot (<http://www.r-project.org>), while also remaining Bioconductor compatible (<http://www.bioconductor.org>). It allows for use of marray,

---

\*equal contributions

†to whom correspondence should be addressed

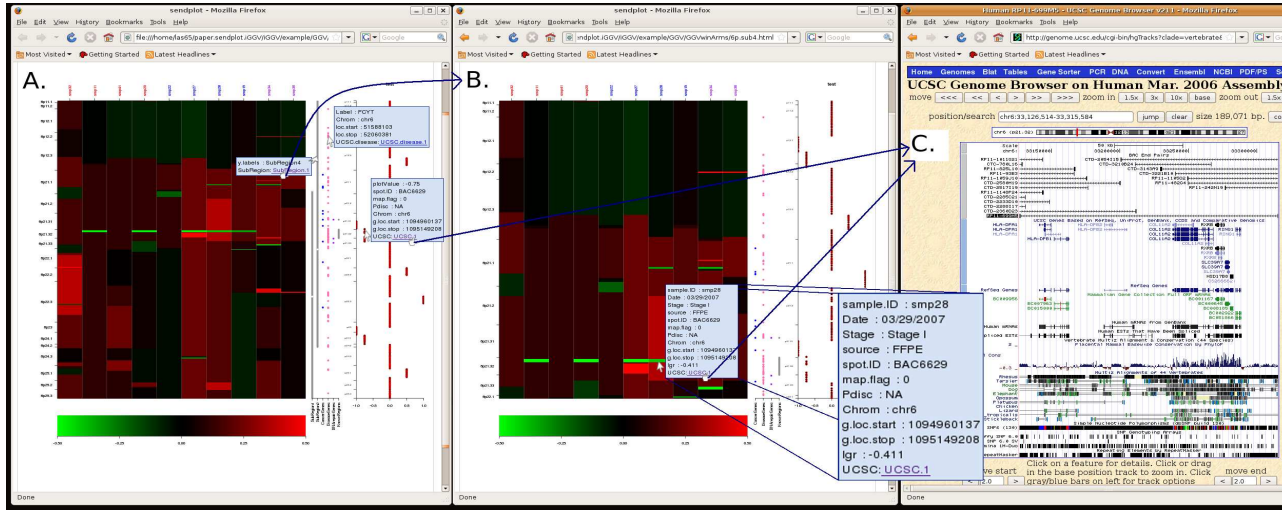


Figure 1: interactive display for array comparative genomic hybridization data. (A) A heatmap image spanning a chromosome arm for entries in a data matrix. The rows correspond to particular BAC aCGH assays, columns correspond to particular assayed clinical samples, and the entries, color coded according to ranges in the data matrix. Accompanying the heatmap is a plot of customized annotations, here showing disease, cancer, and dna repair genes, as well as a customized plot of statistical values (p-values, linear order statistics). The heatmap is divided into smaller subregions which may be viewed to examine heatmap patterns more closely and to account for high dimension data. (B) The subregion has the same interactive as the larger heatmap, and also shows details for each data matrix entry based on user specifications. Only one tool-tip display may show at any one time, but the capability of multiple interactive plots is depicted. It is possible to include links to outside data sources and images. (C) This example shows tool-tips with links to the UCSC genome browser.

AnnotatedDataFrame, and data.frame objects (3) making iGenomicViewer capable of handling a variety of formats such as Affymetrix, Agilent, illumina, and SNP data.

## 4. Description

### 4.1. Interactive Tooltips

The iGenomicViewer is designed as a genomic specific utilization of sendplot. The sendplot functions generate an HTML file with embedded Javascript code to create an interactive web page for viewing an image graphic generated from R. Using meta-data provided by R about the image, a pixel coordinate mapping (image map) is generated using special HTML markup which links various areas of the image to customized Javascript code for rendering tool-tip windows. The Javascript code detects when the users mouse enters a specified area of the image and displays a tool-tip window containing the variable names and values defined for that specific location. The wz.tooltip.js Javascript library ([http://www.walterzorn.com/tooltip/tooltip\\_e.htm](http://www.walterzorn.com/tooltip/tooltip_e.htm)) is embedded within the HTML page and used for rendering the tool-tips.

Tool-tip content is categorized as either x-specific, y-specific, or point specific (e.g., xy-specific) and convenient mechanisms exist to attach user-specified attributes to the interactive plot. For example, in Figure 1 the tool-tip contains x-specific information (i.e., spot.ID, loc.start, Chrom, and Fine.Band), y-specific information (i.e., sample.ID and sex), and point specific information (i.e., log2.ratios.fitted and LGR). The flexibility of adding user customized tooltip data is extended by allowing for inclusions of hyperlinks and images, as well as allowing for the interactivity to exist for multiple graphical representations.

#### 4.2. Platform Independent

The iGenomicViewer is platform independent with respect to biological and technological platforms, thus allowing for broader compatibility. Any genomic data output can be incorporated into the application by creating a mapping object, which contains chromosome information, genomic location information, and identifiers for probes (genes, SNPs, etc). It therefore supports a variety of platforms including but not limited to Affymetrix, Agilent, CGH, and SNP data. The files generated as output are generic and can be viewed by any machine regardless of operating system utilizing a supported web browser.

#### 4.3. Example

In the example below, three of iGenomicViewer's functions are introduced. Data from the R library aCGH is used to show additional compatibility with other applications. This is achieved by creating unique mapping objects for a dataset using one of the mappingObj functions; here mappingObjDF for data.frames is utilized. The annotation object provided with the iGenomicViewer package is loaded. Taking an object-oriented approach, an object is initialized through the initGGV method, and then plotted through makeGGV.

```
library(iGenomicViewer)
library(aCGH)
data(colorectal); data(annObj)
mapObj = mappingObjDF(df=colorectal$clones.info,
                      spot.ID=1, chrom=3, chrom.levels=1:23,
                      locBy="within", loc=4)
GGV = initGGV(vls=colorectal$log2.ratios,
              mapObj=mapObj, annObj=annObj,
              x.labels = colorectal$phenotype[,1:3],
              xy.labels = list(vls=colorectal$log2.ratios))
makeGGV(GGV=GGV)
```

A manual with more detailed examples and descriptions can be downloaded from the iGenomicViewer website.

## 5. Conclusions

The iGenomicViewer allows for quick and convenient generation of customizable interactive, portable, platform-independent plots of genomic data.

Although it is applied to CGH data above, it can be applied to a variety of different platforms.

*Funding:* This work was supported by grants from the National Cancer Institute (CA 27469, to the GOG Administrative Office; CA 37517, to the GOG Statistical Office; CA 113882, to Norma J. Nowak; and CA 16056-27-RPCI, Cancer Center Support Grant to Roswell Park Cancer Institute).

*Conflict of Interest:* none declared

## References

- [1] Albertson, Donna G, Pinkel, Daniel (2003). Genomic microarrays in human genetic disease and cancer, *Genomic microarrays in human genetic disease and cancer*, **12** Spec No 2: 142-152.
- [2] Hibbs, M.A., Dirksen, N.C., Li, K., Troyanskaya, O.G. (2005). Visualization methods for statistical analysis of microarray clusters. *BMC Bioinformatics*, **6**: 115.
- [3] Gentleman RC, Carey VJ et al.. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biololgy*. 2004; 5:R80. doi: 10.1186/gb-2004-5-10-r80
- [4] Prasad, T.V., Ahson, S.I. (2006). Visualization of microarray gene expression data. *Bioinformatics*, **4**: 141-145.