

SOFTWARE

Open Access



A flexible tool to plot a genomic map for single nucleotide polymorphisms

Fuquan Zhang

Abstract

Background: Most genetic association studies use single-nucleotide polymorphisms (SNPs) as the research targets. However, resources to visualize the genomic map of candidate SNPs in a programming manner are limited. We have previously created an R package, *mapsnp* v0.1, to plot the genomic map for a panel of SNPs within a genomic region of interest. It failed to work under the latest version of *Gviz* package.

Results: We updated the *mapsnp* package to keep up with the latest package environment and improved its functionality by adding more parameters to fine tune plotting outputs.

Conclusions: The *mapsnp* package is a flexible software to visualize genomic map for SNPs, involving the relative chromosome location and the transcripts in the region.

Keywords: R package, *Mapsnp*, SNP map

Background

Single-nucleotide polymorphisms (SNPs) are the most common type of genetic variation among people. SNPs are used for estimating predisposition to disease.

Visualizing genomic map relevant to SNPs may inform the reader intuitively. Genome browsers are common tools to show a SNP's genomic information, including NCBI genome browsers [1], UCSC [2], and Ensembl Genome Browser [3]. These browsers offer retrieval resources and serve as reference datasets for individual SNPs or genes. UCSC and Ensembl Browser offers a set of annotation 'tracks' for a genomic region. However, they are not programmatically accessible and have limited plotting options to render users' data. With these tools, it is not possible to produce a map for a specific set of SNPs.

The R language [4] is a widely used language and software environment for statistical computing and graphics. Several programming tools have been developed under R environment for visualizing genomic data, including *GenomeGraphs* [5], *ggbio* [6], and *Gviz* [7]. Within these packages, individual types of genomic features or data are represented by separate tracks, and there are constructor functions to coordinate

and plot these tracks. However, none of these packages provide a method specified to plot genomic information for a panel of user-supplied SNPs.

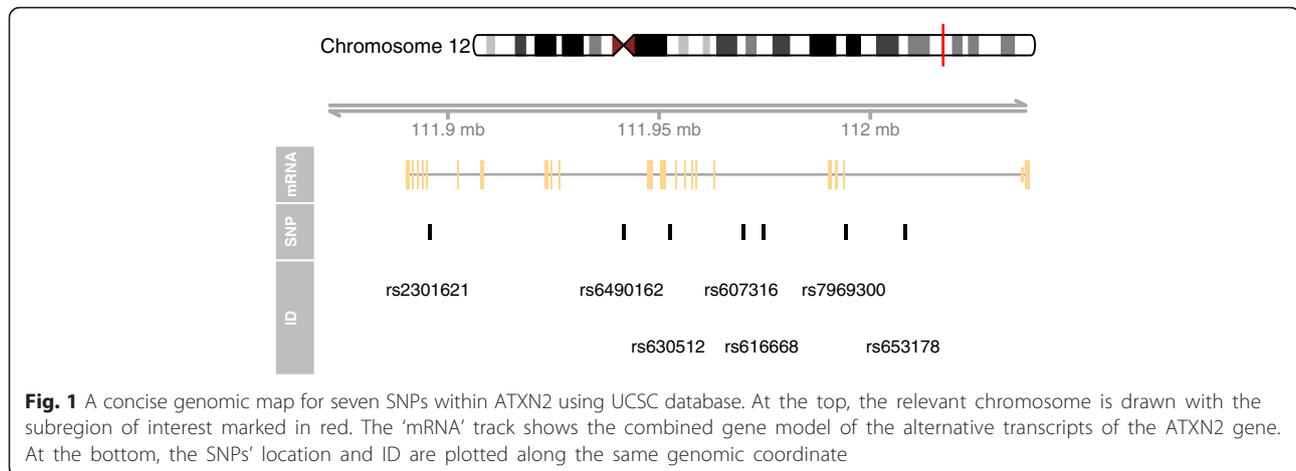
To fulfill this need, we have developed *mapsnp* v0.1 [8] to plot genomic maps for SNPs. It works under R v2.15 and *Gviz* v1.2.1. As the upgrading of R, the *Gviz* package was also updated, which deprecated our *mapsnp* package. To keep up with the latest R environment and the *Gviz* package, we created *mapsnp* v0.2.

Implementation

The *mapsnp* package leverages the *Gviz* system [7] to plot a genomic map for SNPs. A SNP map includes five tracks, an ideogram track for a chromosome, an axis track for genomic coordinates, a transcript track for relevant transcripts, a SNP location track, and a SNP label track annotating their ID symbol.

The *mapsnp* v0.2 package contains one function, 'msb'. The function has three mandate parameters, 'M', 'start', and 'end'. Parameter 'M' is a data frame consist of three columns, including chromosome, SNP ID, and SNP genomic location. The 'start' and 'end' parameter define the range of a highlighting region, typically a gene region where the SNPs are located to. For transcript track, the 'msb' function utilizes *Homo Sapiens* data from UCSC build hg19 based on the

Correspondence: zhangfq@njmu.edu.cn
Wuxi Mental Health Center, Nanjing Medical University, 156 Qianrong Road,
Wuxi, Jiangsu Province 214151, China



knownGene table, implemented in the 'TxDb.Hsapiens.UCSC.hg19.knownGene' package [9].

There are dozens of other parameters, which fine-tune other track properties, such as color, size, track name, annotation text, and so on. The detailed usage of the package is described in Additional file 1.

Results and discussion

To illustrate the use of `mapsnp` v0.2, we show an example for 'msb' on the built-in dataset involving seven candidate SNPs within the ATXN2 gene (Fig. 1) [10]. The genomic range of this gene is from 111950277 to 112036294 base pair.

```
> library (mapsnp)
> library (TxDb.Hsapiens.UCSC.hg19.knownGene)
> data (snp)
> msb (M = snp, start = 111950277, end = 112036294)
```

Compared with its precursor, `mapsnp` v0.2 offers more parameters to fine-tune the output map. See Additional file 2 for several examples using alternative plotting options.

SNPs occur normally throughout a person's genome. They can act as biological markers, helping scientists locate genes that are associated with disease. Visualization of SNPs facilitates exploration and discovery by revealing genomic patterns of variations [6]. `mapsnp` provides an easy-to-use method to visualize genomics annotations for a group of SNPs. The output maps deliver views of SNP locations, genomic regions, summary views of splicing patterns, and genome-wide overviews with karyogram. The package is especially useful for most candidate gene studies by exploring genomic features for relevant SNPs.

Widely-used visualization tools are implemented in the form of a genome browser. Comparison of our tool with UCSC Genome Browser and Ensembl has been described previously [11]. Of note is that the package can handle only one chromosome at a time.

Also, users need an established internet connection to fetch data from UCSC.

Conclusion

The `mapsnp` package provides a simple and flexible function to plot genomic maps for a set of SNPs.

Availability and requirements

Project name: `mapsnp`

Project home page: <https://sourceforge.net/projects/mapsnp/files/>

Operating system(s): Platform independent.

Programming language: R platform.

Other requirements: Internet connection.

License: GPL (≥ 3)

Additional files

Additional file 1: `mapsnp` manual. (PDF 90 kb)

Additional file 2: Plotting examples. Examples to plot the map with other options. (PDF 193 kb)

Competing interests

The author declares that he has no competing interests.

Author's contributions

FZ designed the software package and wrote the manuscript.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (81471364).

Received: 26 February 2015 Accepted: 29 March 2016

Published online: 02 April 2016

References

1. Coordinators NR. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 2014;42(Database issue): D7–17.
2. Rosenbloom KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, Dreszer TR, Fujita PA, Guruvadoo L, Haeussler M, Harte RA, Heitner S, Hickey G, Hinrichs AS, Hubley R, Karolchik D, Learned K, Lee BT, Li CH, Miga KH, Nguyen N, Paten B, Raney BJ, Smit AF, Speir ML, Zweig AS, Haussler D, Kuhn

- RM, Kent WJ. The UCSC genome browser database: 2015 update. *Nucleic Acids Res.* 2015;43(Database issue):D670–81.
3. Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, Gil L, Giron CG, Gordon L, Hourlier T, Hunt S, Johnson N, Juettemann T, Kahari AK, Keenan S, Kulesha E, Martin FJ, Maurel T, McLaren WM, Murphy DN, Nag R, Overduin B, Pignatelli M, Pritchard B, Pritchard E, Riat HS et al. Ensembl 2014. *Nucleic Acids Res.* 2014;42(Database issue):D749–55.
 4. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.
 5. Durinck S, Bullard J. GenomeGraphs: Plotting genomic information from Ensembl. 2014, R package version 1.22.0
 6. Yin T, Cook D, Lawrence M. ggbio: an R package for extending the grammar of graphics for genomic data. *Genome Biol.* 2012;13(8):R77.
 7. Hahne F, Durinck S, Ivanek R, Mueller A, Lianoglou S, Tan G. Gviz: Plotting data and annotation information along genomic coordinates. 2014, R package version 1.6.0.
 8. mapsnp. [https://github.com/csuzfq/mapsnp_pkg].
 9. Carlson M. TxDb.Hsapiens.UCSC.hg19.knownGene: Annotation package for TranscriptDb object(s). 2014, R package version 2.10.1.
 10. Zhang F, Wang G, Shugart YY, Xu Y, Liu C, Wang L, Lu T, Yan H, Ruan Y, Cheng Z, Tian L, Jin C, Yuan J, Wang Z, Zhu W, Cao L, Liu Y, Yue W, Zhang D. Association analysis of a functional variant in ATXN2 with schizophrenia. *Neurosci Lett.* 2014;562:24–7.
 11. Zhang F, Xu Y, Cao H, Jin C, Cheng Z, Wang G, Shugart YY. Mapsnp: an R package to plot a genomic map for single nucleotide polymorphisms. *PLoS One.* 2015;10(4):e0123609.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

