BioJS-HGV Viewer: A BioJS component for visualzing protein variants

Saket K Choudhary*

Leyla J Garcia, Andrew Nightingale†

skchoudh@usc.edu

Abstract

Genomic studies have resulted in catalogs of genetic variants in humans. Studying the pattern of damaging and non-damaging variants can not only help understand evolution but can also potentially improve human health by identifying the key driver elements.

We present BioJS-HGV Viewer, a BioJS component to represent and visualize genetic variants pooled from various sources. The component presents information at different levels allowing the end user to study the pattern of variations in detail in a user friendly manner.

The code for BioJS-HGV Viewer is available at:
`https://github.com/saketkc/biojs-genetic-variation-viewer`.

A demo is available at: `http://saketkc.github.io/biojs`

I. Introduction

With the advent of next-generation sequencing technologies, it has been possible to profile genomes in large numbers. One of the chief outcomes of such projects has been catalog of genetic variants such as dbSNP[1] and COSMIC[2]. These catalogs contain publicly accessible sets of genetic variants found in humans which can be utilized to study evolutionary relationships and disease specific variations. COSMIC database is a curated set of somatic mutations as observed in cancer samples. The number of such variations are huge. dbSNP 129 had reportedly more than 14 million unique variants [3]. The availability of data at such a large scale makes the analysis challenging.

Any exploratory attempt at making sense of the variation data would involve visualizing the variants across the genome to determine specific sites, if any where the mutations are more frequent or are absent completely. BioJS-HGV Viewer is a BioJS [4] component developed to visualize genetic variants in a comprehensive manner.BioJS is an open source project providing various components to visualize biological data. These components use javascript for rendering visualization. The visualizations are web based and hence are absolutely platform independent.

II. Methods

The functionality provided by BioJS-HGV Viewer has two parts:

- Overview
- Detailed or Zoomed View
  The architecutre of this component is designed to shandle both DNA and protein variants. The current implementation makes use of protein variants. These variant sites have been generated by an

*Molecular and Computational Biology, University of Southern California, Los Angeles, California, USA
†European Bioinformatics Institute EMBL-EBI, Hinxton, Cambridge, CB10 1SD, UK

un-published webservice made available through EBI. This service has an indexed database of protein variants as reported in the COSMIC and UniProt[**?**] database and is made available as a JSON[**?**] file.

I. Overview Mode

In the default mode the viewer reprents variant

III. Results

IV. Discussion

I. Subsection One

II. Subsection Two

References

[1] E. M. Smigielski, K. Sirotkin, M. Ward, and S. T. Sherry, "dbSNP: a database of single nucleotide polymorphisms.," *Nucleic acids research*, vol. 28, pp. 352–5, Jan. 2000.

[2] S. a. Forbes, N. Bindal, S. Bamford, C. Cole, C. Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, J. W. Teague, P. J. Campbell, M. R. Stratton, and P. A. Futreal, "COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer.," *Nucleic acids research*, vol. 39, pp. D945–50, Jan. 2011.

[3] "dbsnp data statistics." `http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm`. Accessed: 2014-10-04.

[4] M. Corpas, "The BioJS article collection of open source components for biological data visualisation," *F1000Research*, vol. 56, pp. 6–8, Feb. 2014.

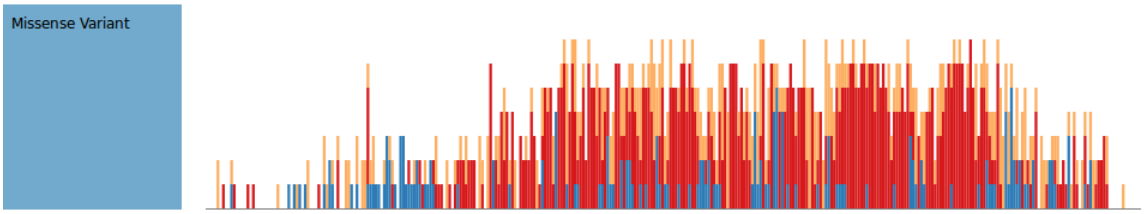

Figure 1: *'Overview' of genetic variants as shown in by HGV viewer*

Figure 2: *'Detailed view' of genetic variants. The SIFT/Polyphen scores and associated information with the mutations is rendered using tooltips(not shown here)*