svviz

`svviz` visualizes high-throughput sequencing data supporting structural variants. The program takes as input one or several bam files with sequencing data, a genome fasta file, and information about a putative structural variant (that has been identified using other methods). From these inputs, `svviz` uses a realignment process to identify reads supporting the reference allele, reads supporting the structural variant (or alternate allele), and reads that are not informative one way or the other (ambiguous).

Reads are then plotted, as in a standard genome browser, but along the sequence of either the alternate allele or the reference allele. The user can thus assess the support for the putative structural variant visually, determine if the breakpoints appear accurate, or estimate a likely genotype for each sample at the given structural variant.

`svviz` differs from existing genome browsers in being able to display arbitrary types of structural variants, such as large insertions, deletions, inversions and translocations. Genome browsers such as IGV are poorly suited to displaying the read data supporting these types of structural variants, which can differ dramatically from the reference genome sequence.

Features

- built-in support for deletions and insertions (including mobile element insertions). `svviz` can easily be extended to analyze translocations and complex variants, but these types are not yet implemented.
- builds reference and alternate allele sequences from genome fasta file and structural variant annotation
- identifies, from the input bam file(s), which reads (both read ends for paired-end sequencing) are likely to be relevant for the given structural variant
- performs Smith-Waterman realignment of all read segments against both alternate and reference allele sequences
- uses alignment score to determine reads supporting reference or alternate allele
- additionally, uses empirical insert-size distribution (rather than mean and stddev) to assign reads as likely derived from alternate or reference allele
- visualizes reads for multiple samples in SVG format (an open-source, web-standard vector graphics format)
- provides a (locally-running) browser-based front-end for inspecting visualizations

Quick-start

1. (OS X only) Ensure that you have a working compiler by following these instructions.
2. Install the latest version of svviz from github using the following terminal command: `sudo pip install git+https://github.com/svviz/svviz.git#svviz`. (The sudo may not be necessary depending on your setup.)
3. Run the following command, which downloads example data and runs it through svviz: `svviz demo`. After several processing steps, a web browser window should open. Click and drag to pan, and zoom using option/alt-scrollwheel.
4. Please report any issues (after making sure they're not explained in the documentation below) using the github issue tracker.