Visualizing Structural Genomic Variants

CPSC547 Project Pitch
Janet Li & Armita Safa
Genome Sequencing

- Figuring out the order of DNA nucleotides

ACGCTGCAT...
Genomic Variants

- Defined as alterations in the genome
- Impact evolutionary processes [Weischenfeldt et al, 2013]
- Traits and appearance – hair color, blood type, etc
- Susceptibility to diseases
  - Genomic variants have been associated with cognitive disabilities, predispositions to obesity, cancer and other maladies [Sudmant et al., 2015]
  - Large recurrent microdeletions associated with schizophrenia [Stefansson et al., 2008]
  - The role of rare structural variants in the genetics of autism spectrum disorders [Kusenda et al., 2009]
Genomic Variants

- Single Nucleotide Variant
- Structural Variants
  - Deletion
  - Insertion
  - Duplication
  - Inversion
- Translocation
SV Detection

---

Genome Biol 20, 246 (2019)
https://doi.org/10.1186/s13059-019-1828-7
Example: Circos plots

Image: https://training.galaxyproject.org/training-material/topics/visualisation/tutorials/circos/tutorial.html
Example: Graph-based view

Graph View: SequenceTubeMap

<table>
<thead>
<tr>
<th>Color</th>
<th>Trackname</th>
<th>Show Track</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>chr12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ins_chr12:80849878&amp;chr12:80849878_NONE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ins_chr12:80853221&amp;chr12:80853221_NONE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ins_chr12:80853303&amp;chr12:80853303_NONE</td>
<td></td>
</tr>
<tr>
<td></td>
<td>inv_chr12:80842171&amp;chr12:80861781</td>
<td></td>
</tr>
</tbody>
</table>

Image: http://demo.momig.tokyo/#graph
Visualizing Multiple SVs

● It’s often interesting to compare genomic differences between individuals
● Trio analysis (parents + child)
  ○ Was an SV inherited, or did it occur as a de novo alteration?
● Tumour/normal comparisons
  ○ Which SVs are present in tumour tissue and may be driving more mutations?
Project

- View and compare SVs in multiple genomes
- Show positions of genes related to a certain disease
- Potential dataset sources:
  - Genome in a Bottle (variant calls)
  - 1000 Genomes Project (variant calls)
  - DisGeNET (disease-gene associations)