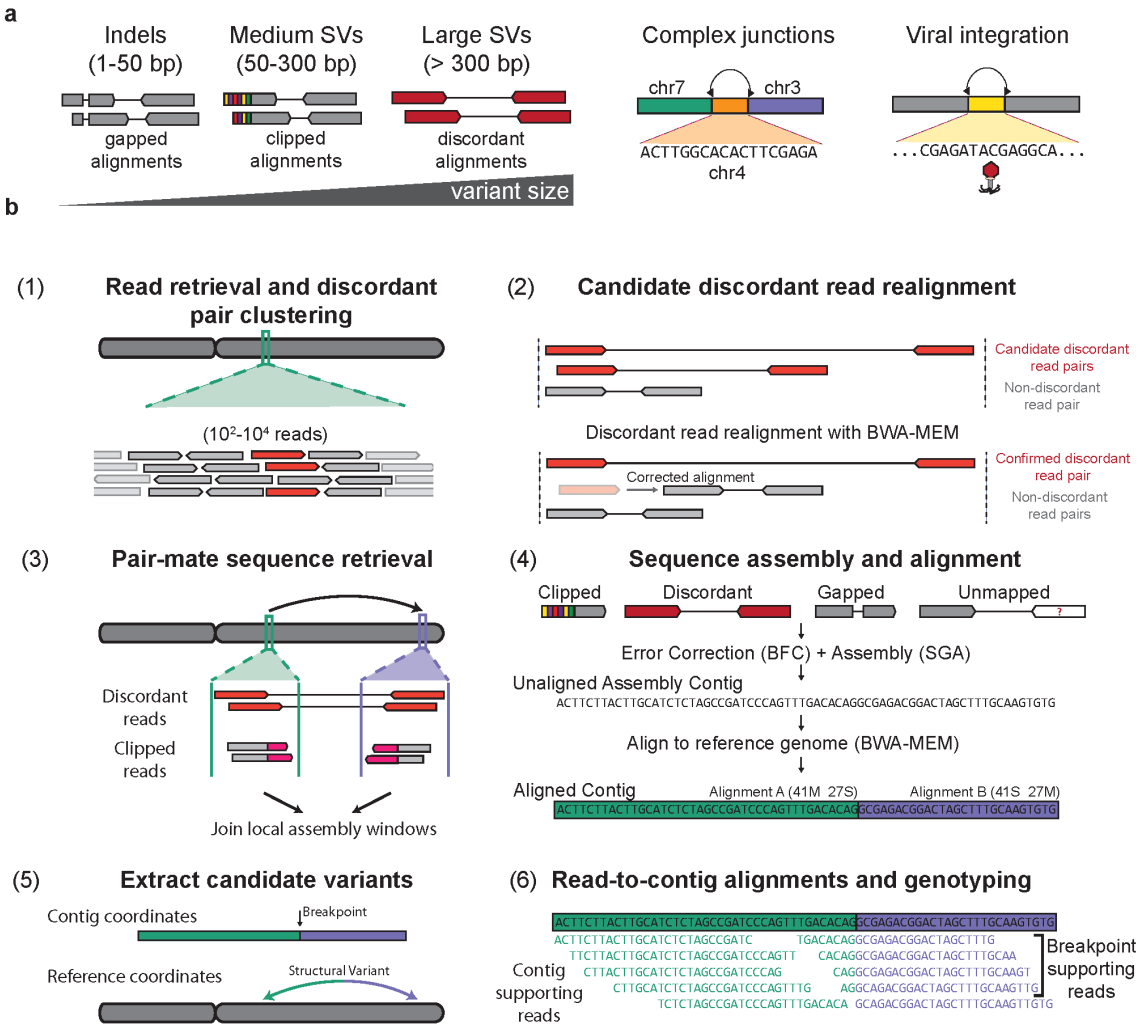

Figure 1: Overview of the SvABA structural variation detection tool. **a)** (left) SvABA assembles with SGA aberrantly aligned sequence reads that may reflect an indel or SV. Such reads include gapped alignments (for indels), clipped alignments (for medium and large SVs) and discordant read pairs (for large SVs). In additional to detecting indels and SVs, SvABA can identify complex rearrangement junctions (middle) and sites of viral integration (right). **b)** The workflow for the SvABA pipeline. (1) Reads within a small window are extracted from one or multiple BAM files and discordant reads are clustered. (2) Discordant reads are re-aligned to the reference to remove pairs that have a candidate non-discordant alignment. (3) The discordant read clusters are used to identify additional regions where reads should be extracted. (4) The sequences are error corrected with BFC and assembled with SGA into contigs. Contigs are immediately aligned to the reference with BWA-MEM. (5) Contigs with multi-part alignments or gapped alignments are parsed to extract candidate variants. (6) Sequence reads are aligned to the contig and to the