1 Purpose

Generate a .bbl file from a list of citations, so that references can be included within a document. Simply copy the reference list from the .bbl file that BibTeX generates and paste it into the main .tex file (and delete the associated \bibliography and \bibliographystyle commands).

2 Example

Once bibtex is run, the .bbl file will contain the following:

```
\begin{thebibliography}{1}
\bibitem{snpsift}
P.~Cingolani, V.~M. Patel, M.~Coon, T.~Nguyen, S.~J. Land
    , D.~M. Ruden, and
  X.~Lu.
\newblock {{U}sing {D}rosophila melanogaster as a {M}odel
    for {G}enotoxic
  {C}hemical {M}utational {S}tudies with a {N}ew {P}
      rogram, {S}np{S}ift}.
\newblock {\em Front Genet}, 3:35, 2012.
\end{thebibliography}
```

Simply copy and paste this code into the main document.

3 Citing bioinformatic tools

1. SnpSift[4]
2. BWA-MEM[6]
3. FastQC[2]
4. SAMtools and BCFtools[7]
5. Trimmomatic[3]
6. bam.iobio[9]

We sequenced 10 samples on 10 lanes on an Illumina HiSeq 2000, aligned the resulting reads to the hg19 reference genome with BWA-MEM[6], applied GATK[8] base quality score recalibration, indel realignment, duplicate removal, and performed SNP and INDEL discovery and genotyping across all 10 samples simulatenously using standard hard filtering parameters or variant quality score recalibration according to GATK Best Practices recommendations [5, 10].

4 Other citations

1. GWAS on Australian aboriginal population [1]

References

[1] D. Anderson, H. J. Cordell, M. Fakiola, R. W. Francis, G. Syn, E. S. Scaman, E. Davis, S. J. Miles, T. McLeay, S. E. Jamieson, and J. M. Blackwell. First genome-wide association study in an Australian aboriginal population provides insights into genetic risk factors for body mass index and type 2 diabetes. *PLoS ONE*, 10(3):e0119333, 2015.

[2] S. Andrews. FastQC A Quality Control tool for High Throughput Sequence Data.

[3] A. M. Bolger, M. Lohse, and B. Usadel. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120, Aug 2014.

[4] P. Cingolani, V. M. Patel, M. Coon, T. Nguyen, S. J. Land, D. M. Ruden, and X. Lu. Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Front Genet*, 3:35, 2012.

[5] M. A. DePristo, E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.*, 43(5):491–498, May 2011.

[6] H. Li. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv e-prints*, March 2013.

[7] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16):2078–2079, Aug 2009.

[8] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.*, 20(9):1297–1303, Sep 2010.

[9] C. A. Miller, Y. Qiao, T. DiSera, B. D'Astous, and G. T. Marth. bam.iobio: a web-based, real-time, sequence alignment file inspector. *Nat. Methods*, 11(12):1189, Dec 2014.

[10] G. A. Van der Auwera, M. O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S. Gabriel, and M. A. DePristo. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics*, 11(1110):1–11, Oct 2013.