BWA-mem is not the best aligner for ancient DNA short reads

Adrien Oliva[1], Raymond Tobler[1,2], Bastien Llamas[1,2,3], Yassine Souilmi[1,2,3,*]

[1] Australian Centre for Ancient DNA, School of Biological Sciences, Faculty of Sciences, The University of Adelaide, Adelaide SA 5005, Australia

[2] The Environment Institute, Faculty of Sciences, The University of Adelaide, Adelaide SA 5005, Australia

[3] National Centre for Indigenous Genomics, Australian National University, Canberra, ACT 0200, Australia

*Corresponding Author: Y.S. yassine.souilmi@adelaide.edu.au

Abstract

Xu and colleagues (Xu et al., 2021) recently suggested a new parameterisation of *BWA-mem* (Li, 2013) as an alternative to the current standard *BWA-aln* (Li and Durbin, 2009) to process ancient DNA sequencing data. The authors tested several combinations of the *-k* and -r parameters to optimise *BWA-mem*'s performance with degraded and contaminated ancient DNA samples. They report that using *BWA-mem* with *-k* 19 *-r* 2.5 parameters results in a mapping efficiency comparable to *BWA-aln* with *-l* 1024 *-n* 0.03 (i.e. a derivation of the standard parameters used in ancient DNA studies; (Schubert et al., 2012)), while achieving significantly faster run times.

We recently performed a systematic benchmark of four mapping software (i.e. *BWA-aln*, *BWA-mem*, *NovoAlign* (http://www.novocraft.com/products/novoalign), and *Bowtie2* (Langmead and Salzberg, 2012) for ancient DNA sequencing data and quantified their precision, accuracy, specificity, and impact on reference bias (Oliva et al., 2021). Notably, while multiple parameterisations were tested for *BWA-aln*, *NovoAlign*, and *Bowtie2*, we only tested *BWA-mem* with default parameters.

Here, we use the alignment performance metrics from Oliva et al. to directly compare the recommended *BWA-mem* parameterisation reported in Xu et al. with the best performing alignment methods determined in the Oliva et al. benchmarks, and we make recommendations based on the results.

Methods

We investigated the alignment performance of the parameterisation recommended by Xu et al., i.e. *-k* 19 and *-r* 2.5 (hereafter called BWA9) against several of the best performing strategies identified in Oliva et al. (namely, BWA1, BWA2, BWA8, Novo1IUPAC, Novo2IUPAC, and Novo2, see Table 1 for parameter settings).

Following the analytical framework of Oliva et al., our benchmark is based on simulated reads (including fragmentation, damage, and sequencing errors typical for ancient DNA samples; see (Oliva et al., 2021)) that were generated for each of the following three samples from the 1000 Genome Project (1000 Genomes Project Consortium et al., 2015) dataset, each coming from a distinct population, and were aligned to reference genome GRCh37:

- Individual *HG00119* from the British in England and Scotland population; GBR; labelled Europe in this study.
- Individual *NA19471* from the Luhya population in Webuye, Kenya; LWK; labelled Africa in this study.

- *Individual HG00513* from the Han Chinese population in Beijing, China; CHB; labelled East Asia in this study.

In addition to quantifying read alignment precision (i.e. the proportion of correctly aligned reads relative to all aligned reads) and proportion of aligned reads (i.e. the fraction of aligned reads relative to the total number of simulated reads) for each strategy, we tested the specificity (i.e., the fraction of unmapped reads) of these strategies for two sets of potential contaminants—i.e. bacterial and dog reads—that were also used in Oliva et al., 2021.

Results

BWA9 had a slight improvement in the proportion of total reads aligned relative to *BWA-mem* using default settings (BWA8), but this came at the cost of consistently lower precision (Figure 1). These precision differences are particularly accentuated for reads between 30 to 60bp, the range of read lengths that is typical of ancient DNA. As demonstrated here and in more detail in our recent alignment software benchmark (Oliva et al., 2021), *BWA-aln* (BWA1 and BWA2) is the most precise alignment method amongst the tested strategies, having moderately higher precision relative to *BWA-mem* for shorter reads while mapping a much higher percentage of reads overall (Oliva et al., 2021; van der Valk et al., 2021).

When comparing specificity against potential contaminants, BWA9 has a near identical specificity to the default *BWA-mem* parameterisation (BWA8) for dog reads, and slightly poorer specificity when testing with bacterial reads, but both parameterisations perform considerably worse than the tested *NovoAlign* (Novo1IUPAC, Novo2IUPAC, and Novo2) and *BWA-aln* (BWA1 and BWA2) strategies for dog reads (Figure 2).

Finally, comparing running times of the two *BWA-mem* parameterisations for each of the three simulated human datasets showed that BWA9 is slightly quicker than BWA8 (Figure 3), confirming the results of ref. (Xu et al., 2021).

Conclusion

Xu et al. report that *BWA-mem* produces alignment results that are comparable to a derivation of the widely used BWA-aln in the ancient DNA field. Consequently, they recommend the use of a specific non-default *BWA-mem* parameterisation for ancient DNA studies because of its superior runtime and specificity relative to *BWA-aln*. However, we find that this parameterisation actually decreases alignment precision relative to *BWA-mem* using default settings for sequencing reads shorter than 70 bases, which are particularly abundant in ancient DNA samples. Moreover, *BWA-mem* is consistently outperformed by *BWA-aln* under the tested parameterisations for both precision and the proportion of reads mapped, and also had greatly improved specificity when the DNA contamination came from a phylogenetically related organism (i.e. a dog in the present study). Crucially, Oliva et al. have demonstrated that improvements in these alignment metrics are also complemented by a reduction in reference genome bias—an alignment-related bias that can inflate false positives and is particularly problematic for ancient DNA studies.

Accordingly, despite having improved run times, we do not recommend that *BWA-mem* is prioritised over *BWA-aln* for research using short reads—such as ancient DNA, cell-free DNA, and forensic research fields. If run time is an issue for researchers, we recommend the use of *NovoAlign* using the free default parameterisation, so long as an appropriate IUPAC reference can be generated. Readers interested in more detailed discussion of these issues are directed to refs. (Oliva et al., 2021; Poullet and Orlando, 2020; Schubert et al., 2012; van der Valk et al., 2021) for recent benchmarks of different alignment strategies using short reads.

Acknowledgments

We thank the Novocraft Technologies team for providing access to their proprietary *NovoAlign*.

Funding: This work was supported by the Australian Research Council [ARC DP190103705], AO was supported by an ARC PhD Scholarship [ARC IN180100017], RT was supported by an ARC DECRA Fellowship [ARC DE190101069], BL was supported by an ARC Future Fellowship [ARC FT170100448].

Conflict of Interest: The authors declare no conflict of interest.

Data Accessibility

The scripts used to create the used datasets in this study are available in the github repository at: https://github.com/AdrienOliva/Benchmark-aDNA-Mapping.

References

1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., Abecasis, G.R., 2015. A global reference for human genetic variation. Nature 526, 68–74.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv [q-bio.GN].

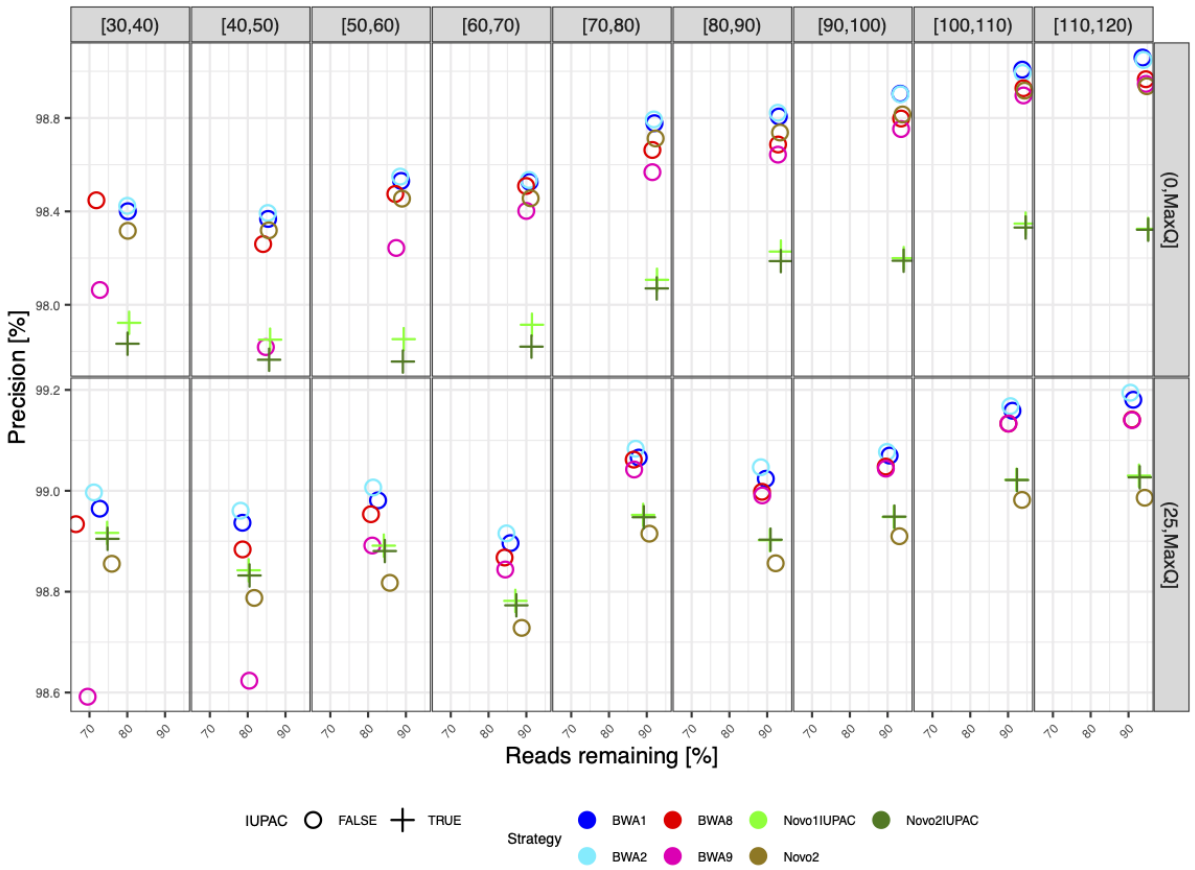
Oliva, A., Tobler, R., Cooper, A., Llamas, B., Souilmi, Y., 2021. Systematic benchmark of ancient DNA read mapping. Brief. Bioinform.

Poullet, M., Orlando, L., 2020. Assessing DNA Sequence Alignment Methods for Characterizing Ancient Genomes and Methylomes.

Schubert, M., Ginolhac, A., Lindgreen, S., Thompson, J.F., Al-Rasheid, K.A.S., Willerslev, E., Krogh, A., Orlando, L., 2012. Improving ancient DNA read mapping against modern reference genomes. BMC Genomics 13, 178.

van der Valk, T., Pečnerová, P., Díez-del-Molino, D., Bergström, A., Oppenheimer, J., Hartmann, S., Xenikoudakis, G., Thomas, J.A., Dehasque, M., Sağlıcan, E., Fidan, F.R., Barnes, I., Liu, S., Somel, M., Heintzman, P.D., Nikolskiy, P., Shapiro, B., Skoglund, P., Hofreiter, M., Lister, A.M., Götherström, A., Dalén, L., 2021. Million-year-old DNA sheds light on the genomic history of mammoths. Nature.

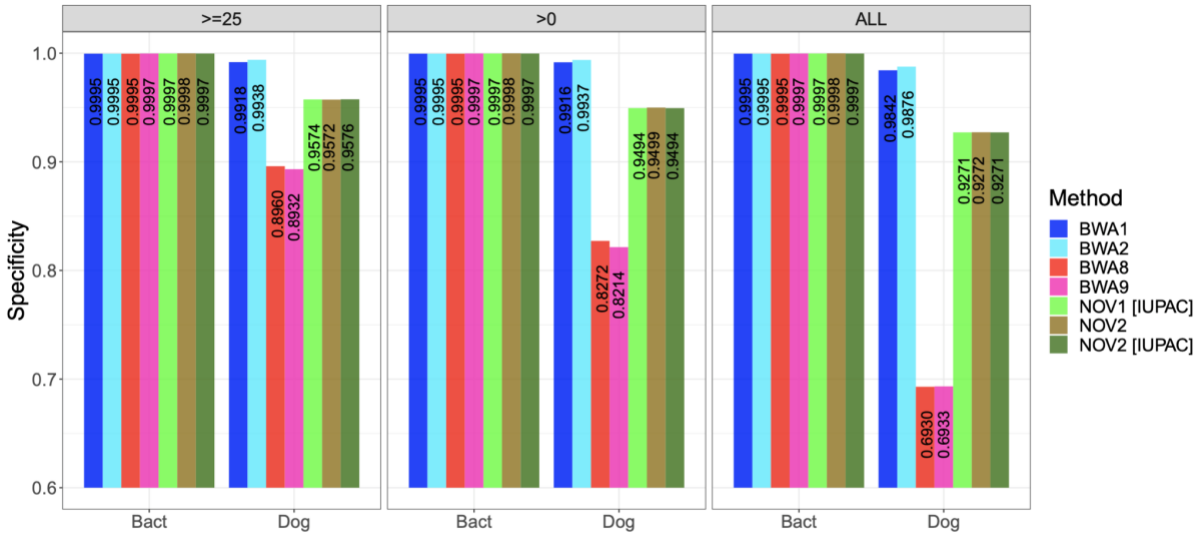
Xu, W., Lin, Y., Zhao, K., Li, H., Tian, Y., Ngatia, J.N., Ma, Y., Sahu, S.K., Guo, H., Guo, X., Xu, Y.C., Liu, H., Kristiansen, K., Lan, T., Zhou, X., 2021. An efficient pipeline for ancient DNA mapping and recovery of endogenous ancient DNA from whole-genome sequencing data. Ecol. Evol. 11, 390–401.



119

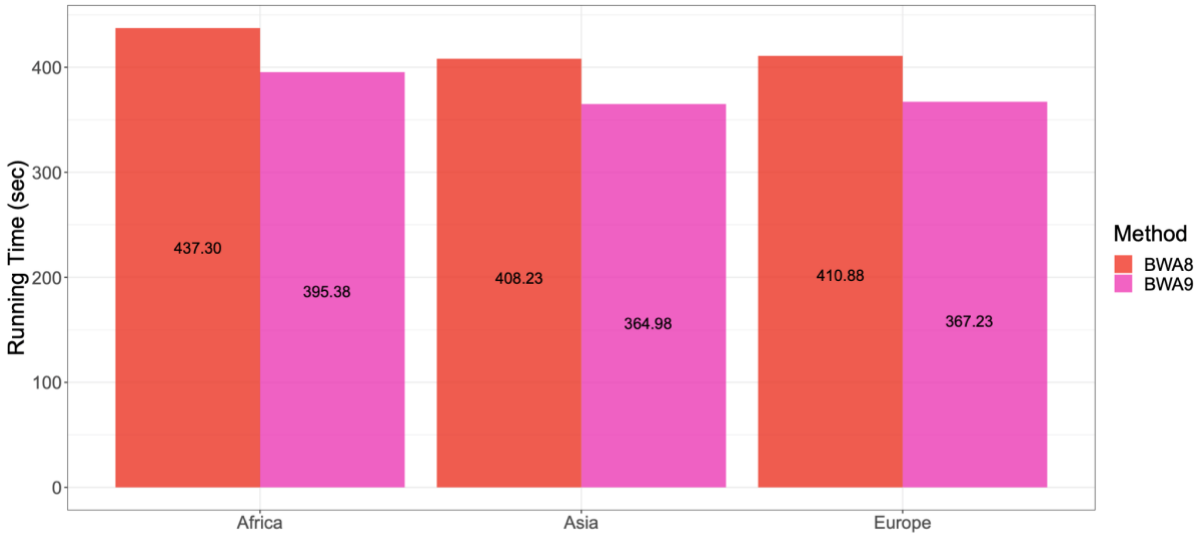
120 **Figure 1. Alignment precision relative to read length and mapping quality for the simulated East**
121 **Asian sample.** Results are shown for seven parameterisations of four different alignment software,
122 including an IUPAC reference-based alignment for a subset of the *NovoAlign* parameterisations (see
123 key). BWA9 is the *BWA-mem* strategy recommended by Xu et al., 2021, with parameter details for the
124 other strategies provided in Table 1. The panels in each row show results after applying the specific
125 mapping quality filter, which results in the removal of all reads below the required mapping quality.
126 Results were similar for the simulated European and African samples and are shown in Appendix
127 Figures 1 and 2, respectively.

128



129

130 **Figure 2. Specificity of all tested alignment methods.** Bacterial and dog reads were aligned to the
131 GRCh37 reference using the seven tested parameterisations of four different alignment software,
132 including an IUPAC reference-based alignment for a subset of the *NovoAlign* parameterisations (see
133 key). The specificity corresponds to the number of unmapped reads, with higher values being better.



134

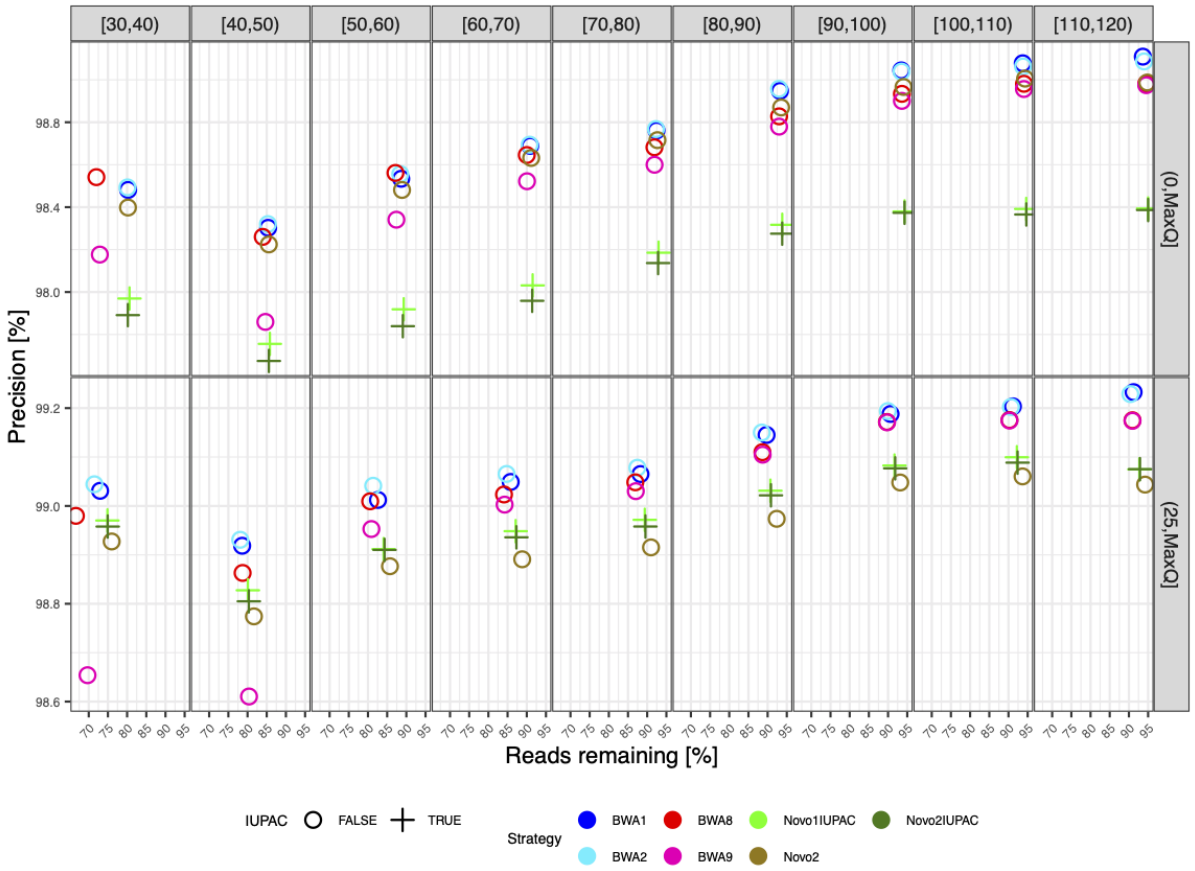
135 **Figure 3: Execution time for each of the *BWA-mem* strategies.** The execution time (walltime) in
136 seconds of BWA8 (default parameters) and BWA9 (Xu et al. parameterisation; -k 19 -r 2.5) based on
137 1.5 million simulated reads.

138 **Table 1. Different alignment parameterisations tested.**

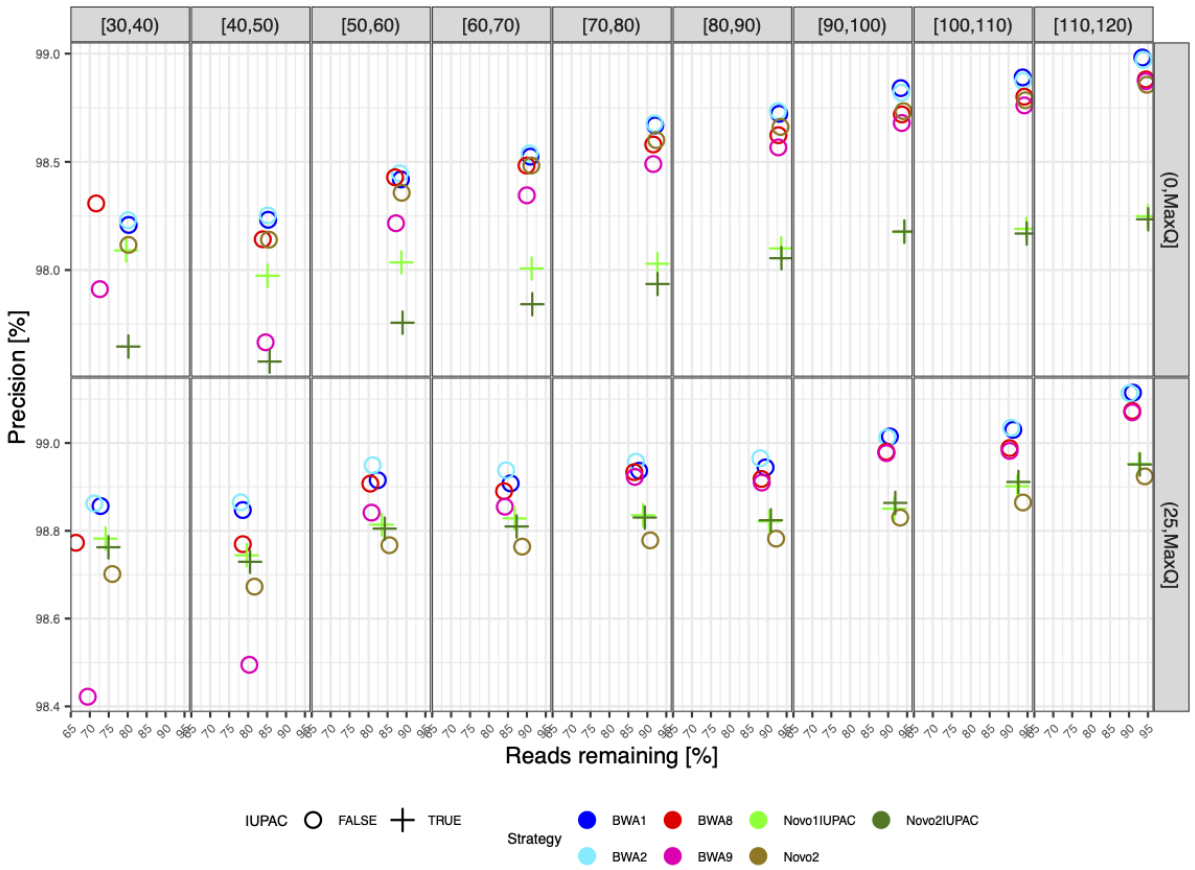
Method	Software	Parameterisation
BWA1	*BWA-aln*	-l 1024 -n 0.01 -o 2
BWA2	*BWA-aln*	-l 1024
BWA8	*BWA-mem*	default
BWA9	*BWA-mem*	-k 19 -r 2.5
Novo1IUPAC	*NovoAlign*	-k
Novo2(IUPAC)*	*NovoAlign*	default

139 *Used with and without the IUPAC reference (Novo2 and NovoIUPAC).

Appendix



Appendix Figure 1: Alignment precision relative to read length and mapping quality for the simulated European sample. See Figure 1.



144

145 **Appendix Figure 2: Alignment precision relative to read length and mapping quality for the**

146 **simulated African sample.** See Figure 1.s