

Validation of identity-by-descent tools applied to WES and WGS data from a Norwegian *SOD1*-family

Maria Fahlström¹, Geir J. Braathen¹, Ole-Bjørn Tysnes^{2,3}, Daphne van Oosten⁴, Øystein Holla¹, Magnus Dehli Vigeland⁵, Trygve Holmøy^{6,7}, Helle Høyen^{1,6}

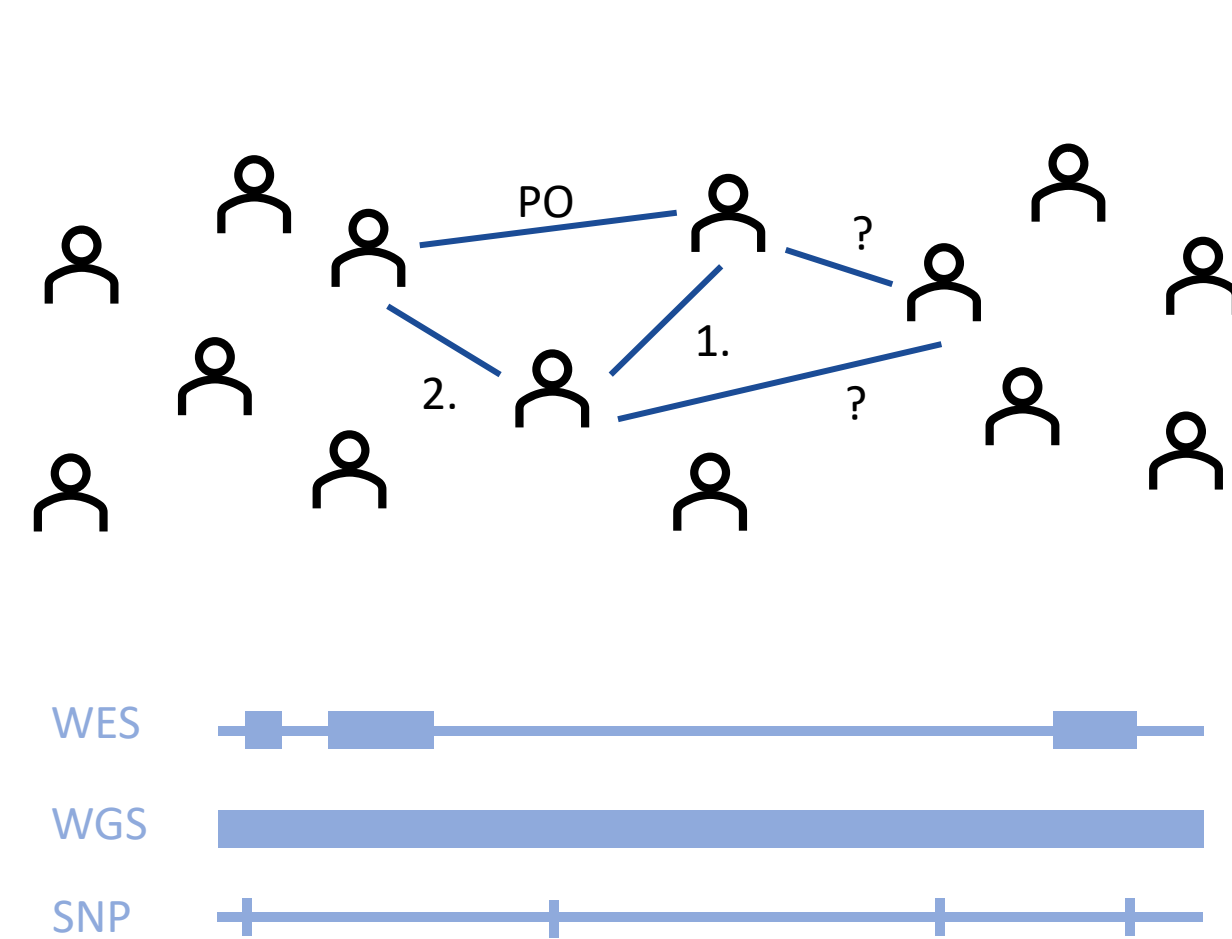
¹Department of Medical Genetics, Telemark Hospital Trust, Skien, Norway, ²Department of Neurology, Haukeland University Hospital, Bergen, Norway, ³Clinical Department 1, University of Bergen, Bergen, Norway, ⁴Department of Neurology, UMC Utrecht Brain Center, Utrecht, The Netherlands, ⁵Department of Forensic Sciences, Oslo University Hospital, Oslo, Norway, ⁶Faculty of Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway, ⁷Department of Neurology, Akershus University Hospital, Lørenskog, Norway.

e-mail: maria.fahlstrom@sthf.no

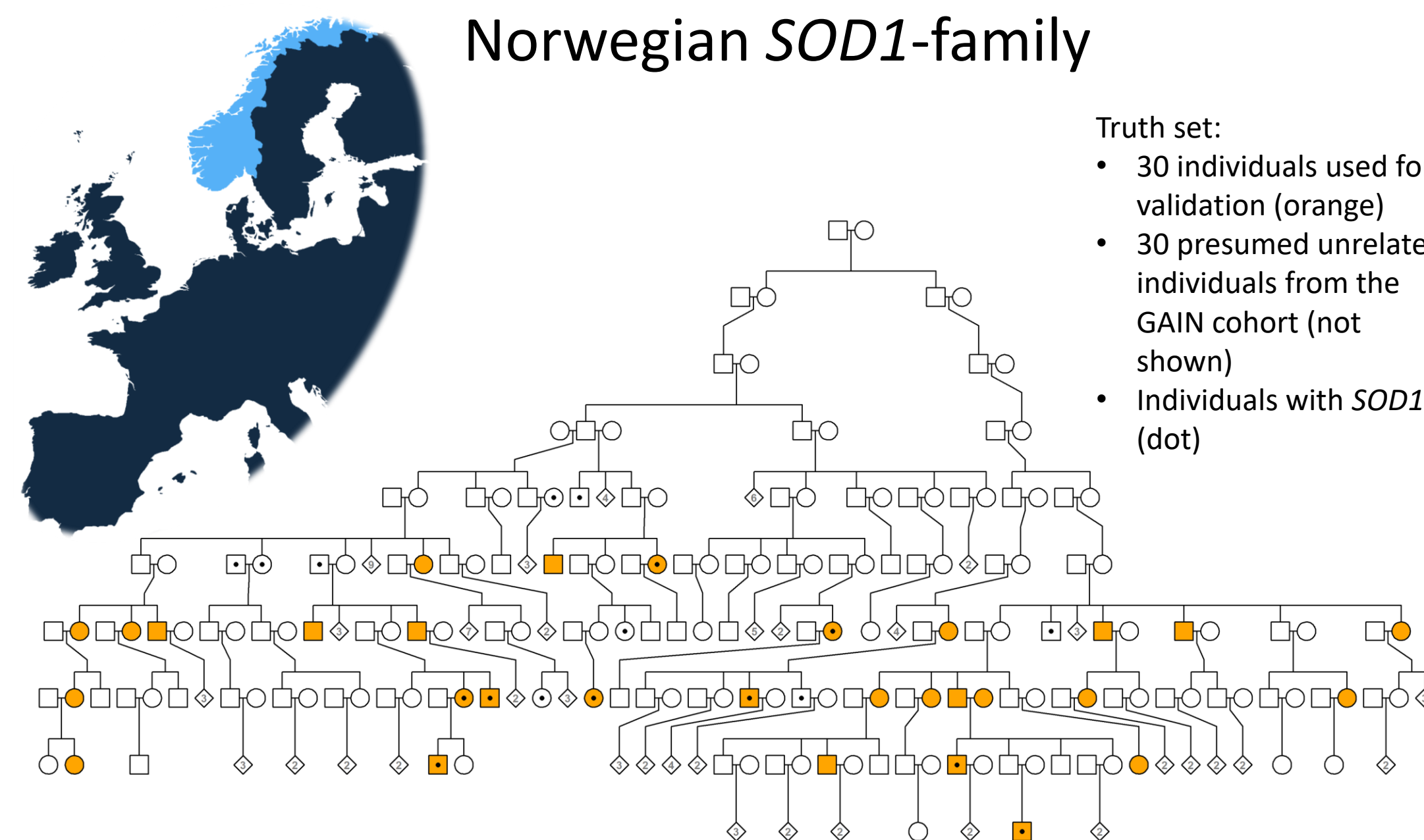
INTRODUCTION

The GAIN study (Genetics of ALS in Norway) began in 2019 and has so far enrolled over 700 ALS patients. On inclusion, participants report both close and distant relatives with ALS. As many are unaware of more distant relations, alternative measures to identify familial aggregation are needed for ALS genetics. Our first 600 samples underwent whole-exome sequencing (WES). However, most distant relatedness detection tools are optimized for SNP arrays or whole-genome sequencing (WGS), making their performance on WES uncertain.

In this study, we analyzed WES data to estimate distant relationships for four identity-by-descent (IBD) tools: DRUID¹, GRAPE², TRIBES³ and XIBD⁴. In addition, we compared their performance on WES and WGS data.

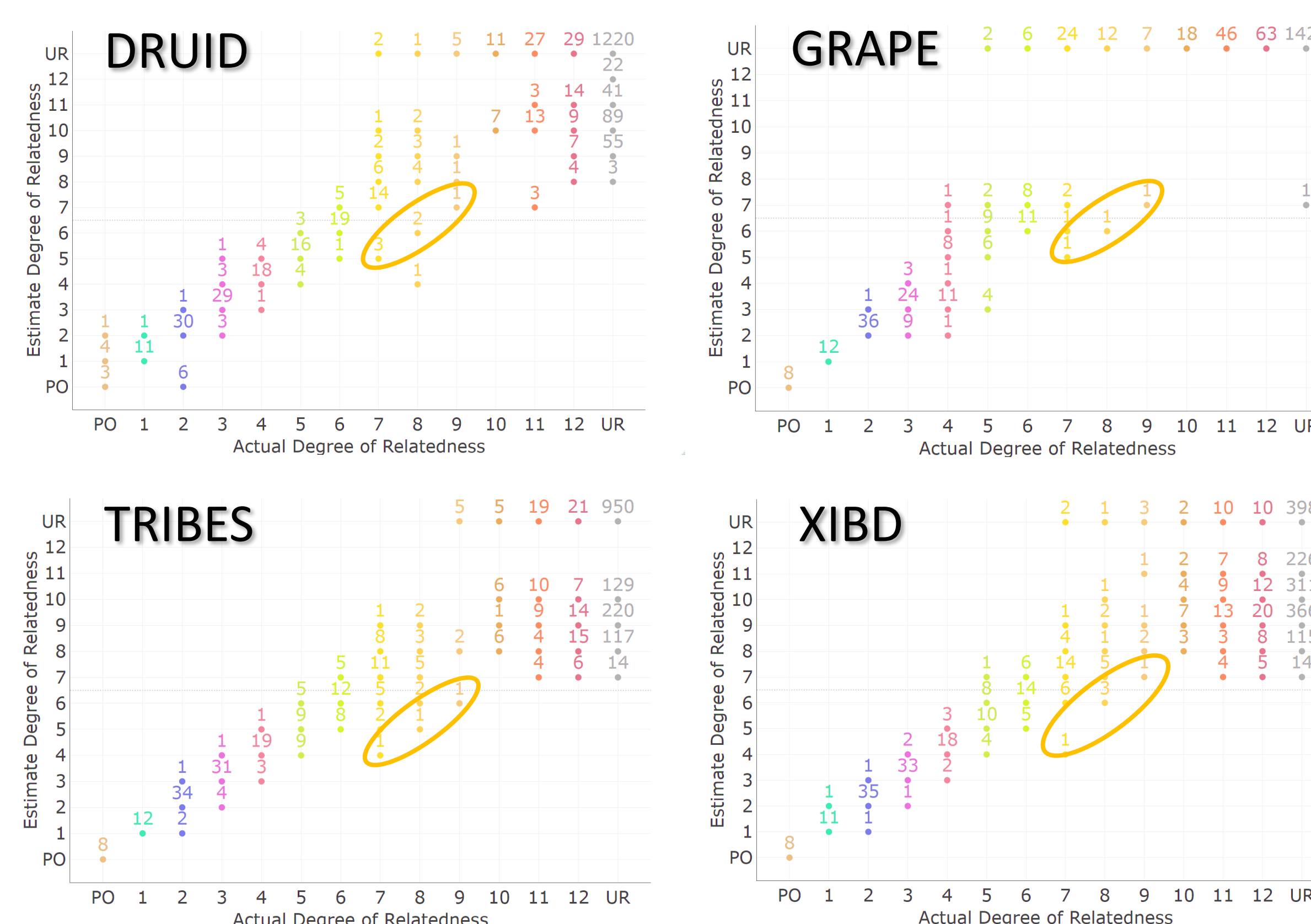
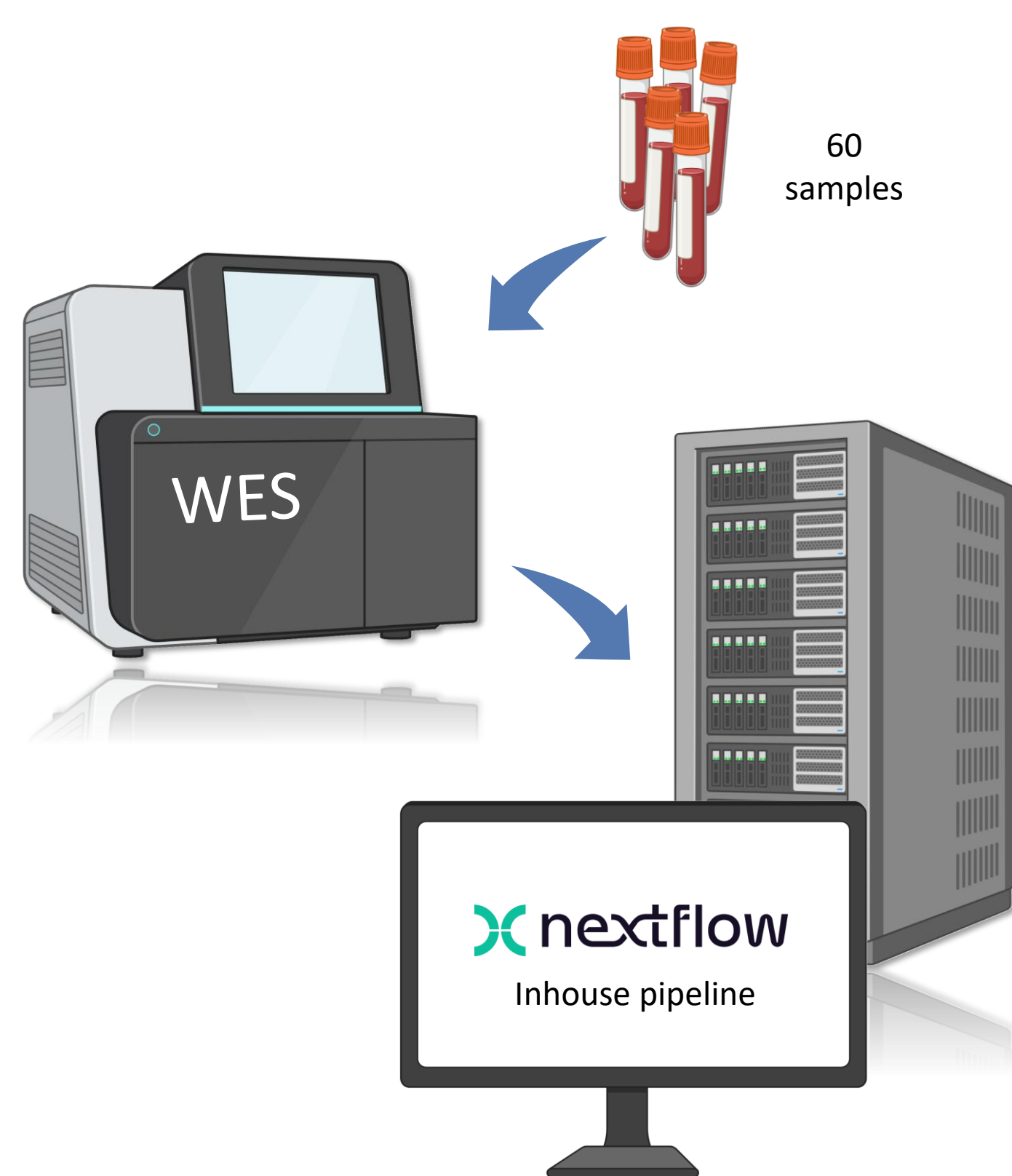


Can identity-by-descent (IBD) relatedness tools optimized for SNP or WGS data be used with WES data for estimation of distant relationships?



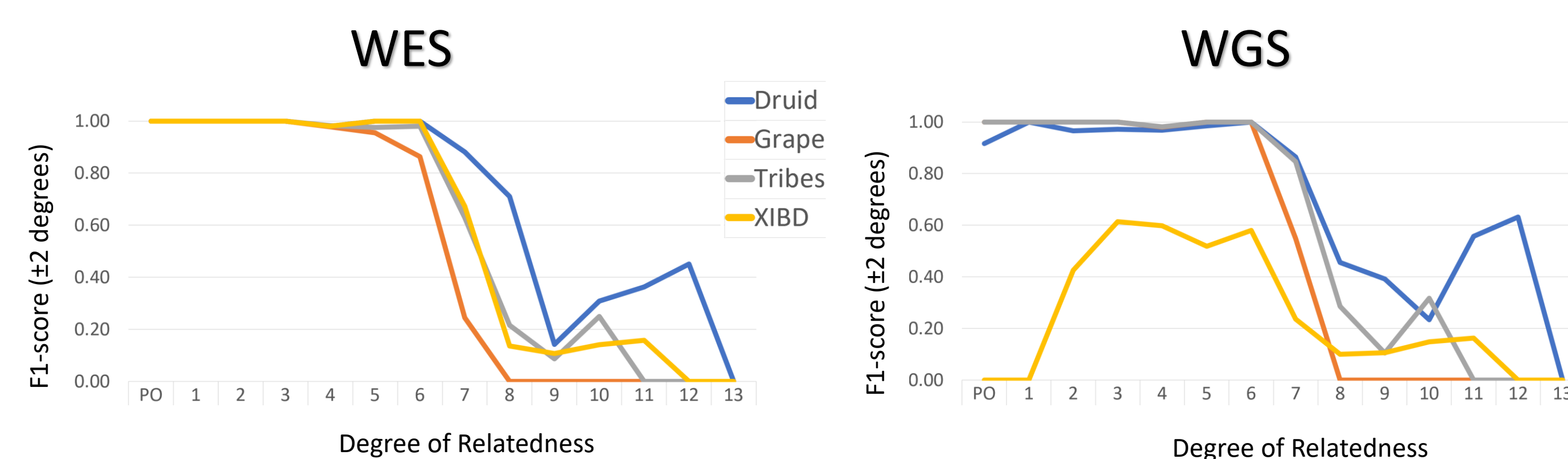
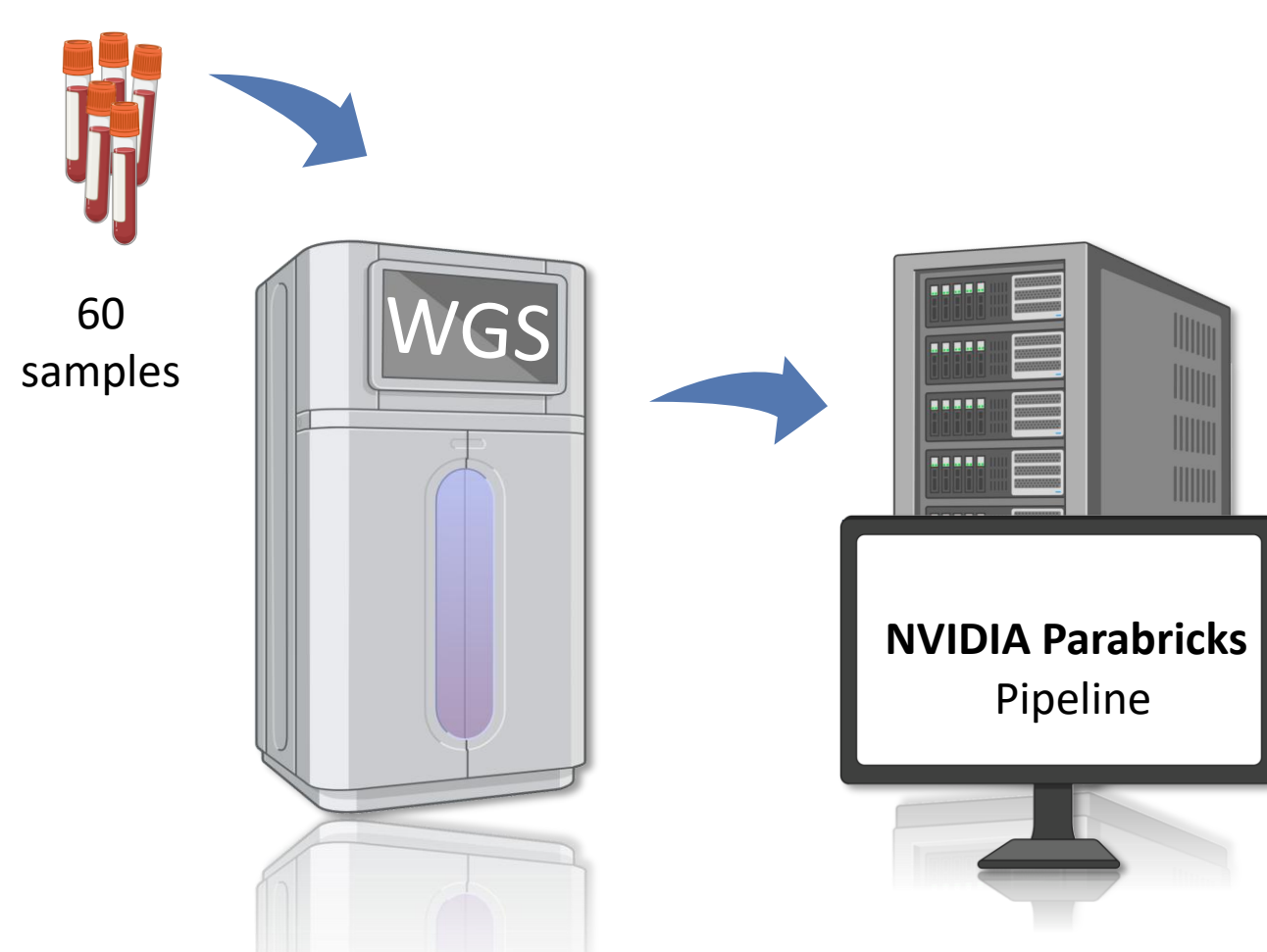
Tool	Method	Pipeline	Original input
DRUID	IBD segment estimation combined with graphs	1. QC filtering* 2. Linkage disequilibrium pruning* 3. Phasing – Beagle 5.5* 4. IBD estimation – GERMLINE 1.5.3* 5. DRUID	SNP array
GRAPE	IBD segment estimation	1. QC filtering 2. Linkage disequilibrium pruning 3. Phasing – Beagle 4.1 4. IBD estimation – IBS GERMLINE RaPID 5. Relatedness estimation – ERS ERS+KING	SNP array
TRIBES	IBD segment estimation	1. QC filtering 2. Linkage disequilibrium pruning 3. Phasing – Beagle 4.1 4. IBD estimation – GERMLINE 1.5.3 5. Masking of artefactual IBD 6. Relatedness estimation	WGS
XIBD	IBD segment estimation	1. QC filtering* 2. XIBD 3. Relatedness estimation*	SNP array

DRUID and XIBD were set up as Nextflow pipelines where * were added for ease of use and automation. For code, please check: [STHF_MEDGEN / GAIN_study / DRUID - GitHub](#) and [STHF_MEDGEN / GAIN_study / xibd - GitHub](#)



Within two degrees, all four tools exhibited strong precision and sensitivity up to 5th degree. For 6th degree GRAPE's sensitivity dropped, while TRIBES and XIBD maintained both precision and sensitivity. DRUID differed from the other tools in few exact matches for the first three degrees and higher precision for 7th degree and above.

3rd cousins being married caused overestimation of degrees of relatedness for their children and grandchildren for all four tools.



$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}}$$

CONCLUSION

WES data can be effectively used to detect relatedness up to 6th degree with IBD tools. Choice of tool should depend on purpose: GRAPE is preferred when false positives must be minimized, whereas TRIBES is optimal for maximizing the number of detected matches. XIBD runs considerably slower than the other tools and is recommended only when X-chromosome IBD is of specific interest. For detecting relationships beyond the 6th degree, DRUID in combination with TRIBES is advisable. Switching from WES to WGS data benefits GRAPE. TRIBES and DRUID show similar performance, while XIBD performs notably worse.

¹Ramstetter MD, Shenoy SA, Dyer TD et al. Inferring Identical-by-Descendant Sharing of Sample Ancestors Promotes High-Resolution Relative Detection. Am J Hum Genet. 2018 Jul 5;103(1):30-44. ²Medvedev A, Lebedev M, Ponomarev A, et al. GRAPE: genomic relatedness detection pipeline. F1000Res. 2023 Apr 13;11:589. ³Twine NA, Szul P, Henden L et al. TRIBES: A user-friendly pipeline for relatedness detection and disease gene discovery. bioRxiv 2019;686253. ⁴Henden L, Wakeham D, Bahlo M. XIBD: software for inferring pairwise identity by descent on the X chromosome. Bioinformatics. 2016 Aug 1;32(15):2389-91.