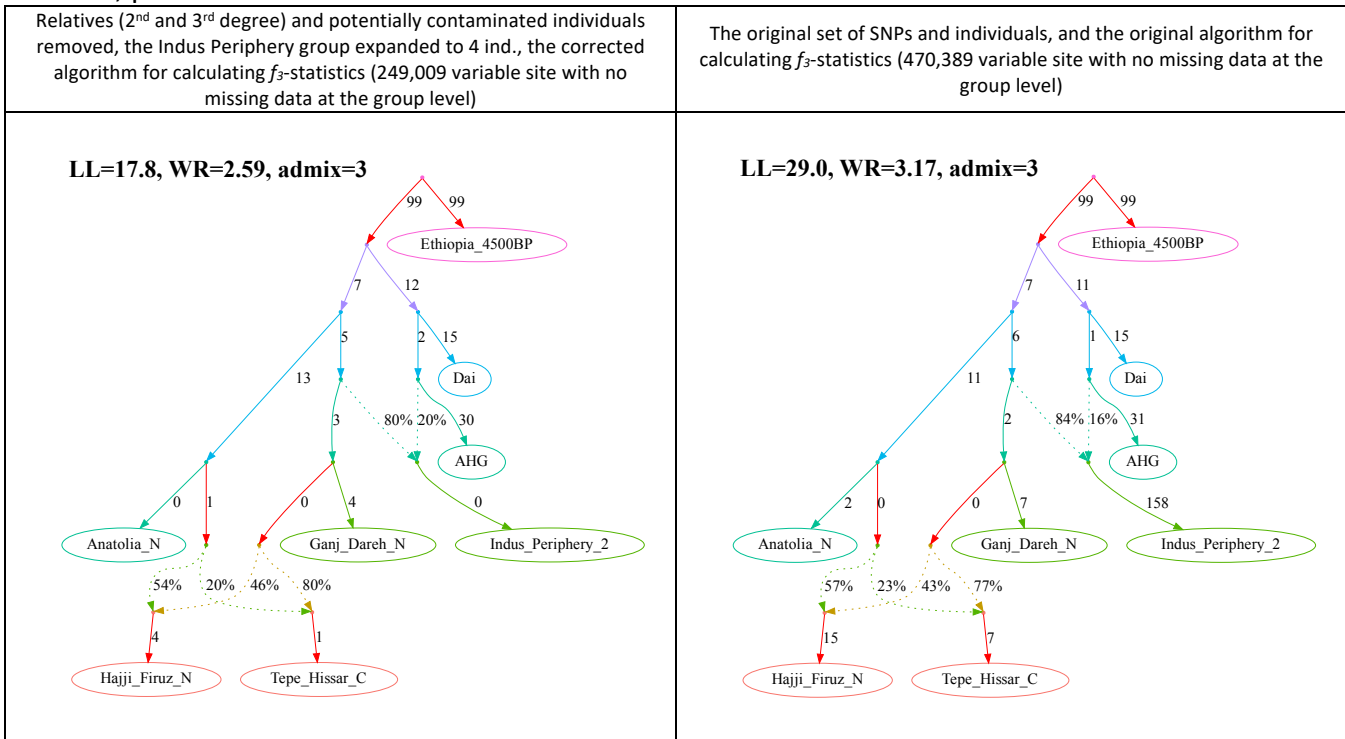


Figure 3—source data 5. Published admixture graph from Shinde *et al.* (2019) and alternative graphs found with *findGraphs* (8 populations, 3 admixture events) for the modified group composition and using the updated algorithm for calculating *f*-statistics. The graphs were also re-fitted on the original set of SNPs/individuals and using the original algorithm for calculating *f*-statistics. Model parameters (graph edges) that were inferred to be unidentifiable are plotted in red.

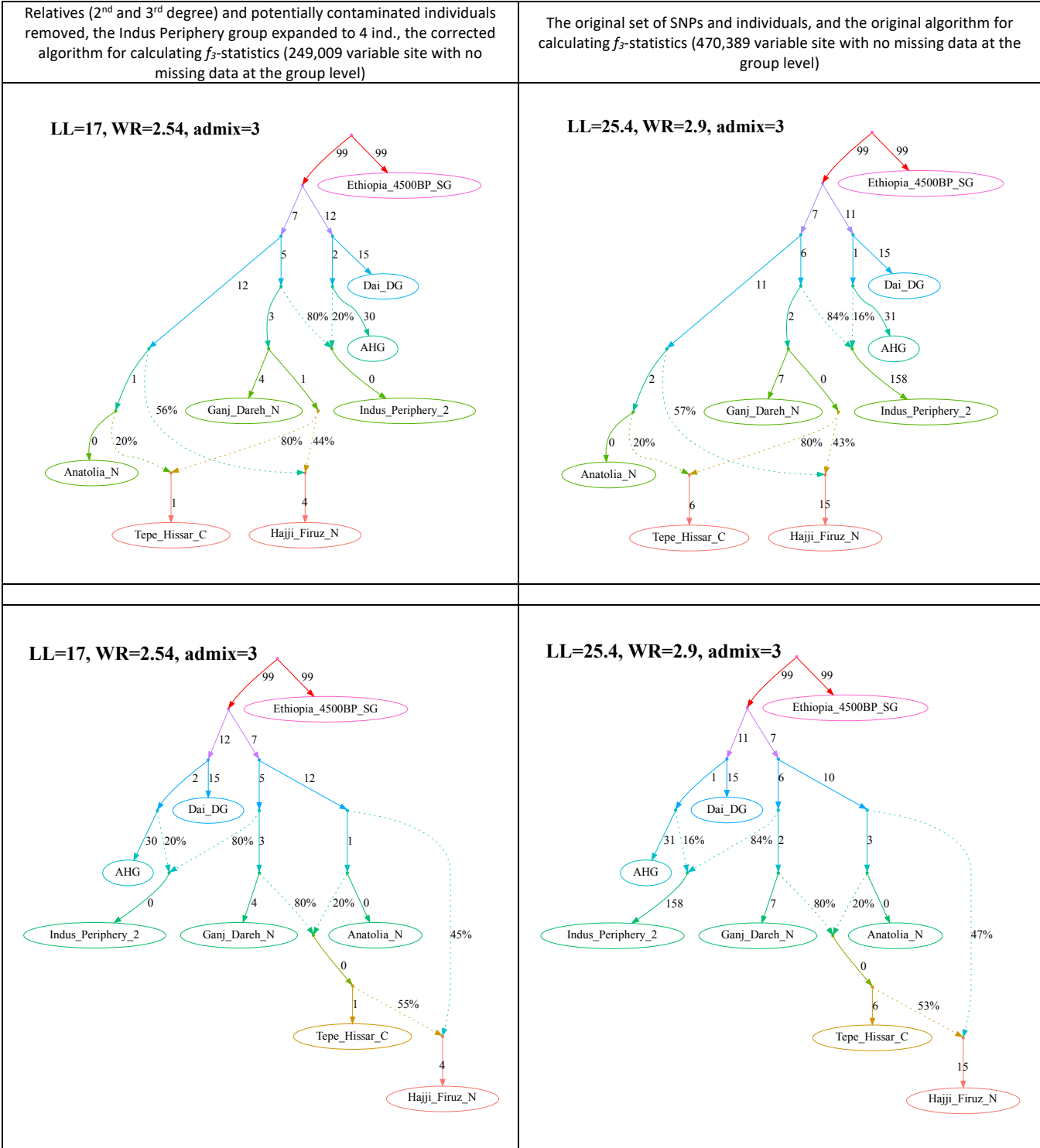
a, published model with 3 admixture events



A claim by Shinde et al. 2019 relying on the admixture graph:

Primary ancestry in the Indus Periphery group forms the deepest branch in the Iranian Neolithic clade composed of the Indus Periphery, Ganj Dareh Neolithic, Hajji Firuz Neolithic, and Tepe Hissar Chalcolithic groups.

b, alternative models fitting nominally better than the published one and confirming all of its important topological details

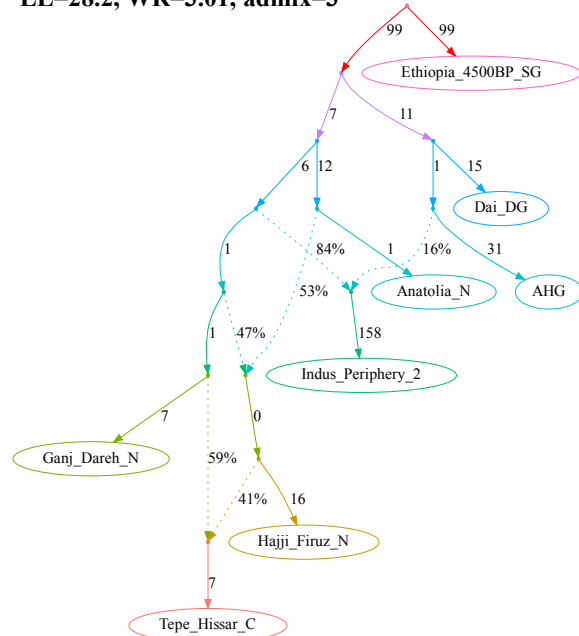
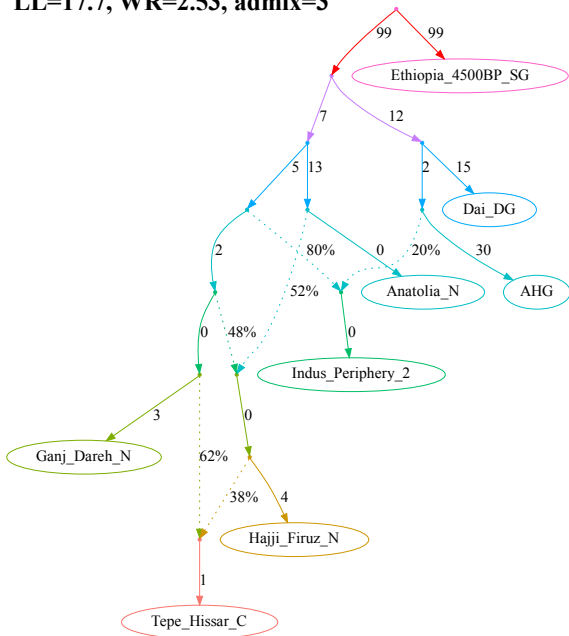


Relatives (2nd and 3rd degree) and potentially contaminated individuals removed, the Indus Periphery group expanded to 4 ind., the corrected algorithm for calculating f_3 -statistics (249,009 variable site with no missing data at the group level)

The original set of SNPs and individuals, and the original algorithm for calculating f_3 -statistics (470,389 variable site with no missing data at the group level)

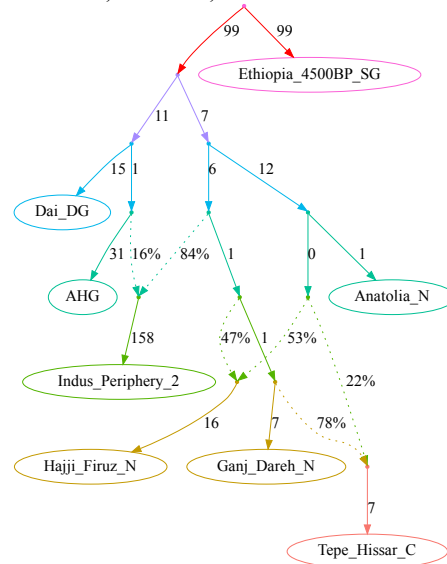
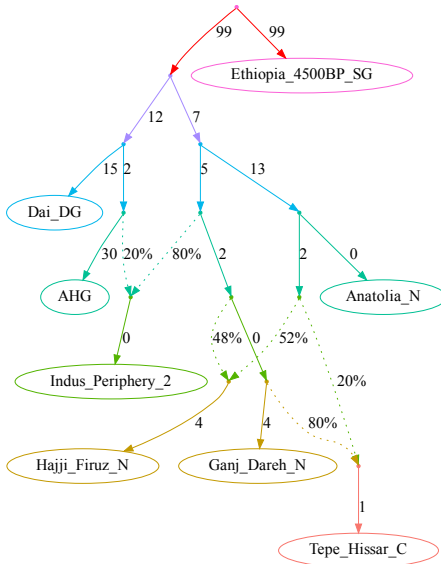
LL=17.7, WR=2.53, admix=3

LL=28.2, WR=3.01, admix=3

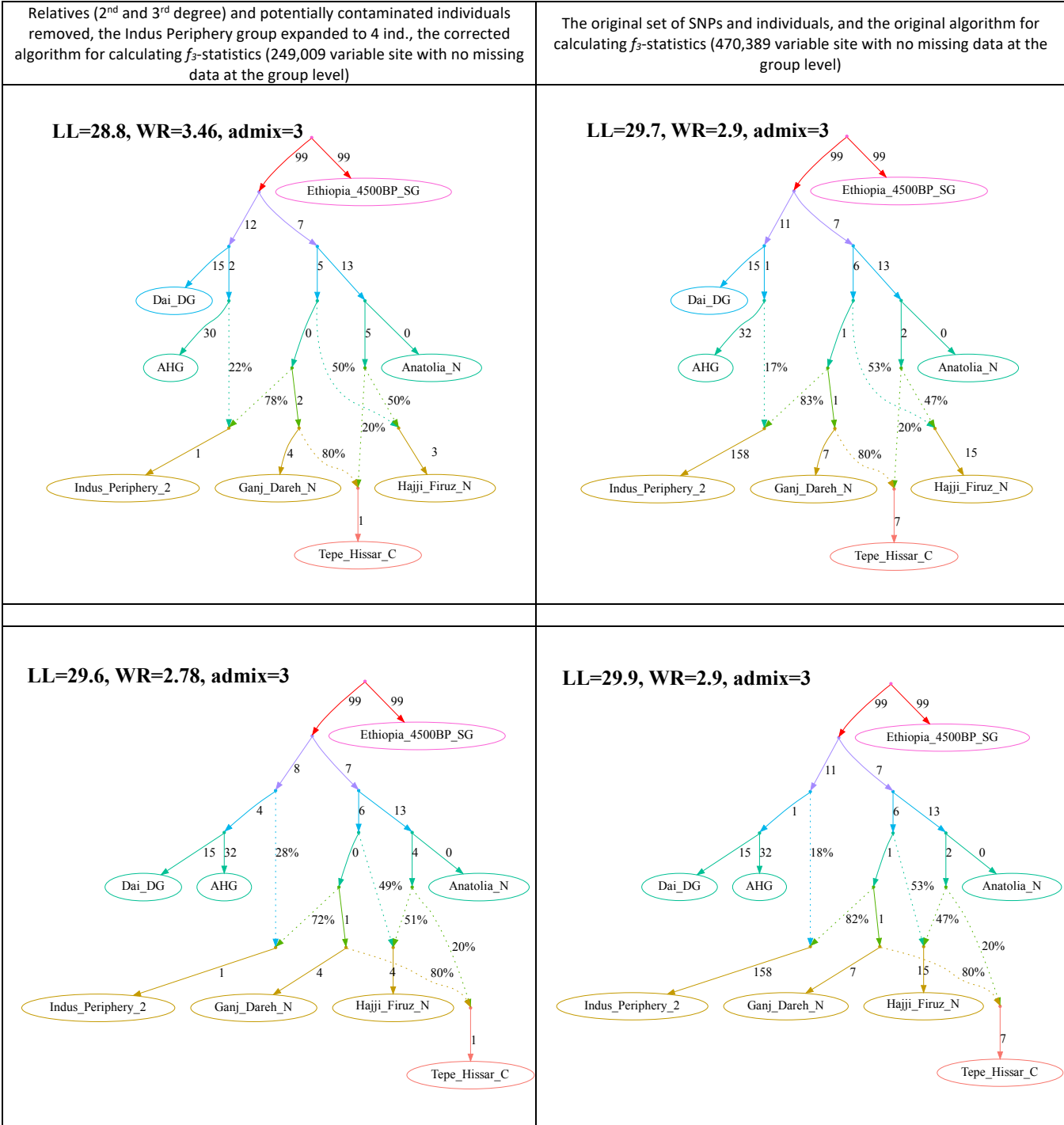


LL=17.7, WR=2.53, admix=3

LL=28.1, WR=3.01, admix=3

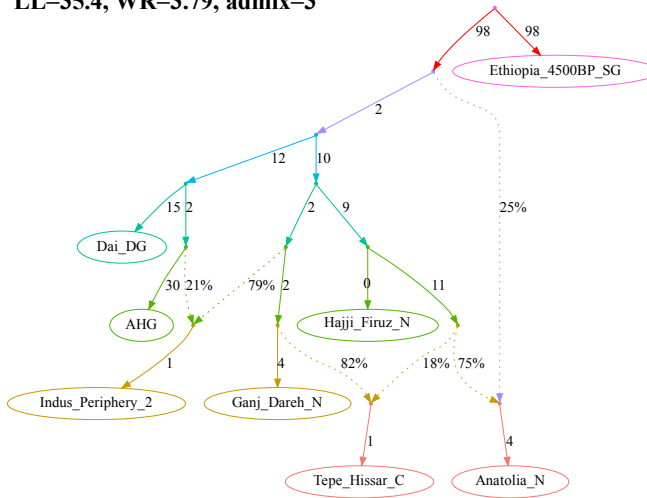


c, alternative models fitting not significantly worse than the published one and differing from it in important ways



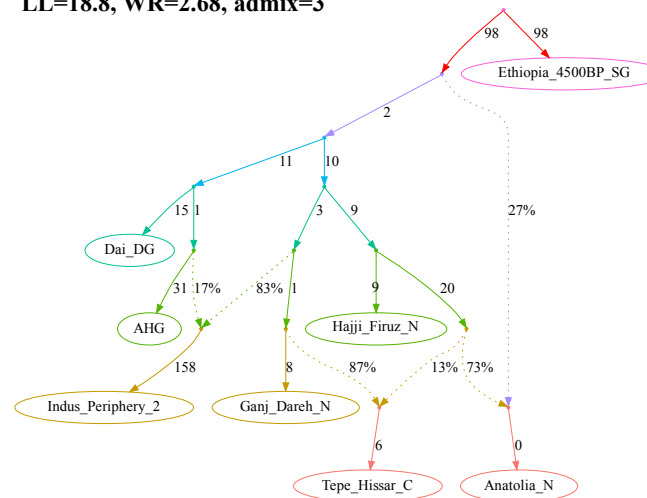
Relatives (2nd and 3rd degree) and potentially contaminated individuals removed, the Indus Periphery group expanded to 4 ind., the corrected algorithm for calculating f_3 -statistics (249,009 variable site with no missing data at the group level)

LL=35.4, WR=3.79, admix=3

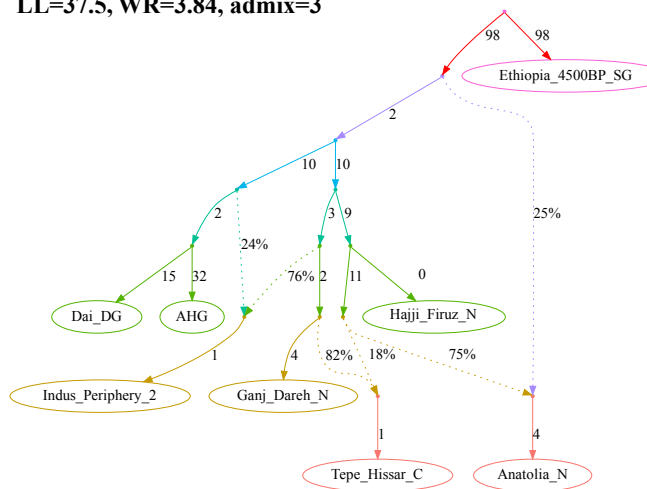


The original set of SNPs and individuals, and the original algorithm for calculating f_3 -statistics (470,389 variable site with no missing data at the group level)

LL=18.8, WR=2.68, admix=3



LL=37.5, WR=3.84, admix=3



LL=19.1, WR=2.68, admix=3

