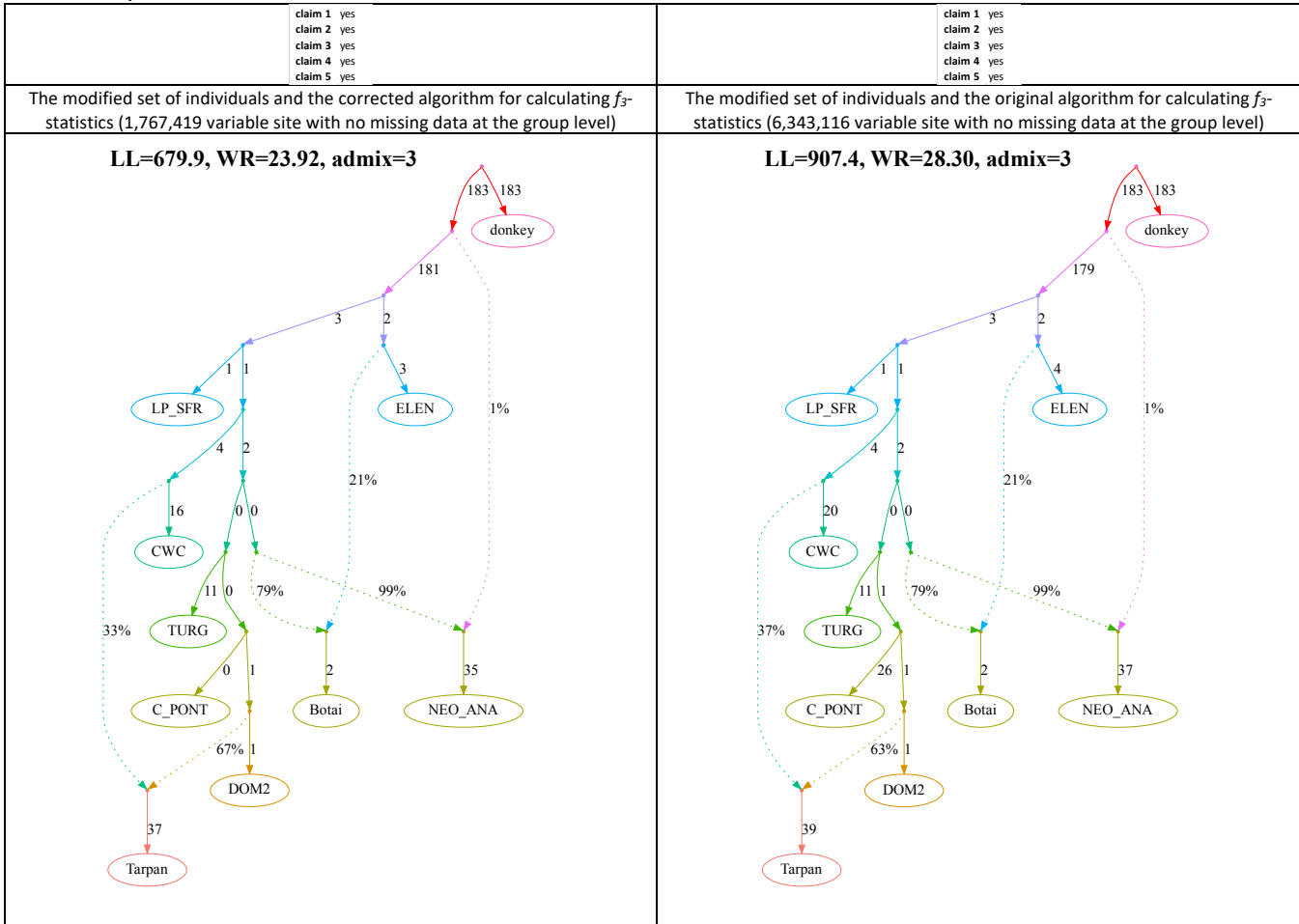


**Figure 3—source data 7.** Published admixture graphs from Librado *et al.* (2019) and alternative graphs found with *findGraphs* (10 populations, 3 to 5 admixture events) for the modified group composition and using the updated algorithm for calculating  $f$ -statistics. The graphs were also re-fitted on the original set of SNPs/individuals and using the original algorithm for calculating  $f$ -statistics. Selected alternative graphs found with *findGraphs* when more admixture events were allowed (from 6 to 9) are also shown. Model parameters (graph edges) that were inferred to be unidentifiable are plotted in red.

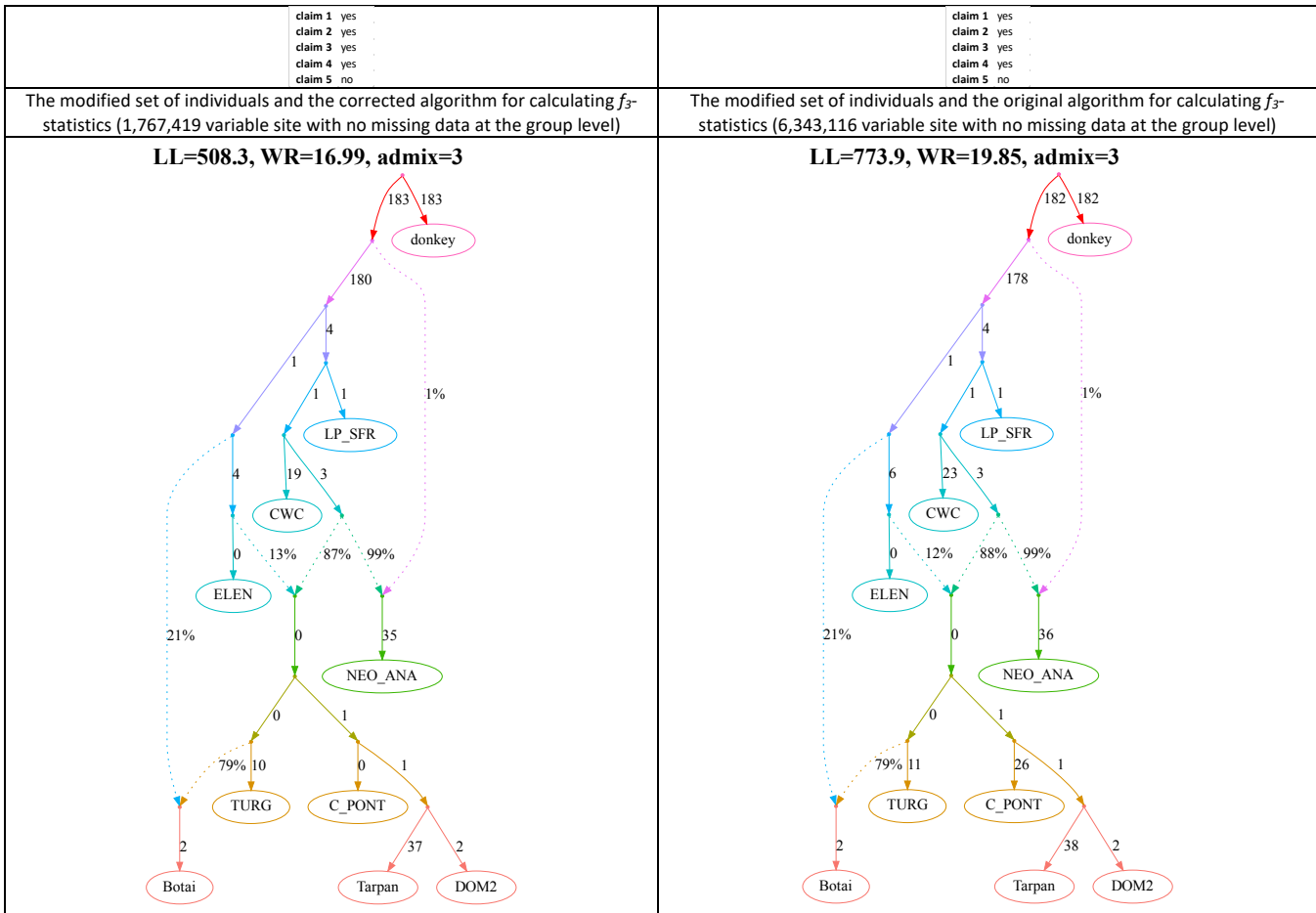
**a, published model, 3 admixture events**



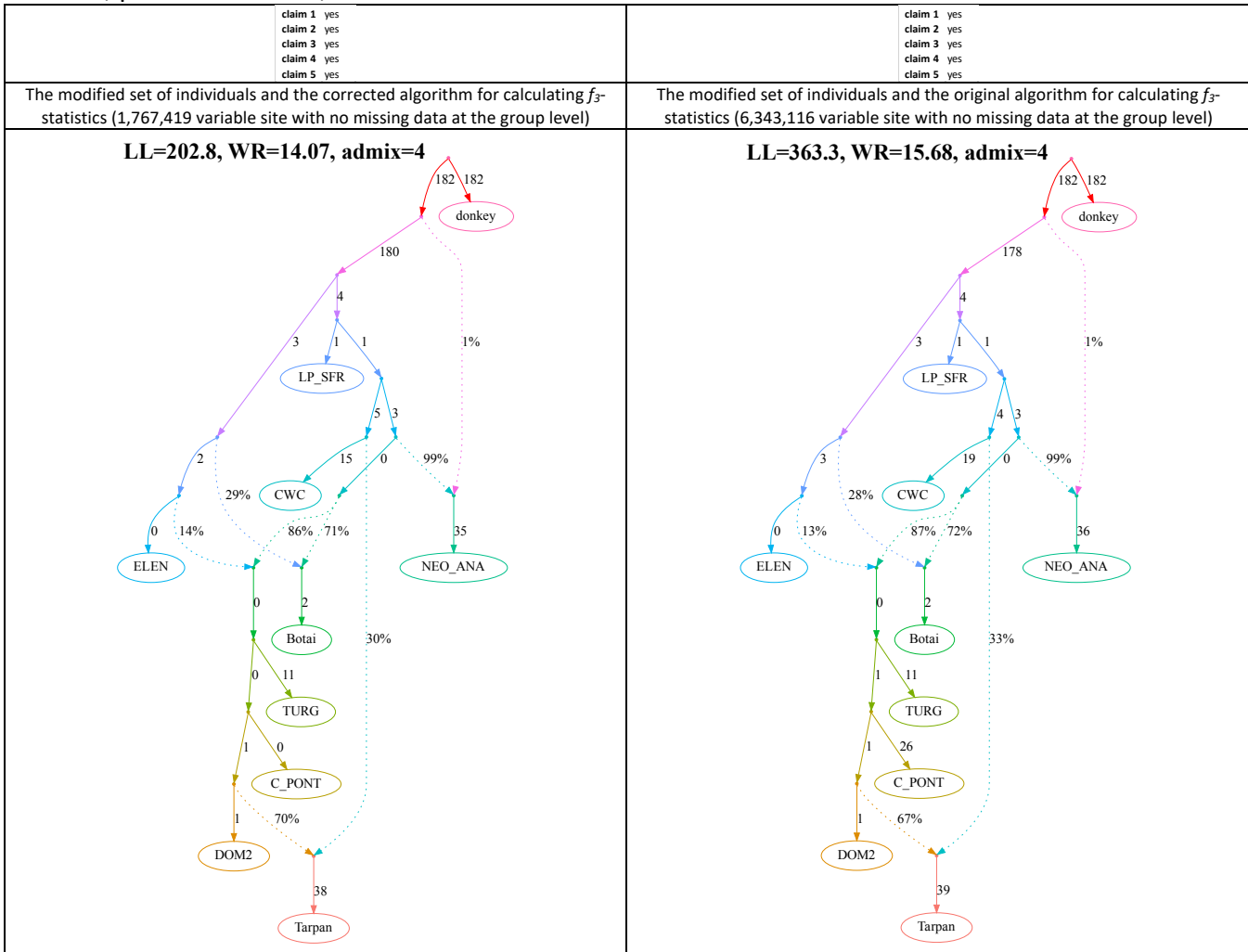
Claims by Librado *et al.* 2021 relying on the admixture graph:

- 1) NEO-ANA-related admixture is absent in DOM2;
- 2) DOM2 and C-PONT are sister groups;
- 3) there is no gene flow connecting the CWC and the cluster associated with Yamnaya horses and horses of the later Sintashta culture whose ancestry is maximized in the Western Steppe (DOM2, C-PONT, TURG);
- 4) there was a gene flow from a deep-branching ghost group to NEO-ANA;
- 5) Tarpan is a mixture of a CWC-related and a DOM2-related lineage.

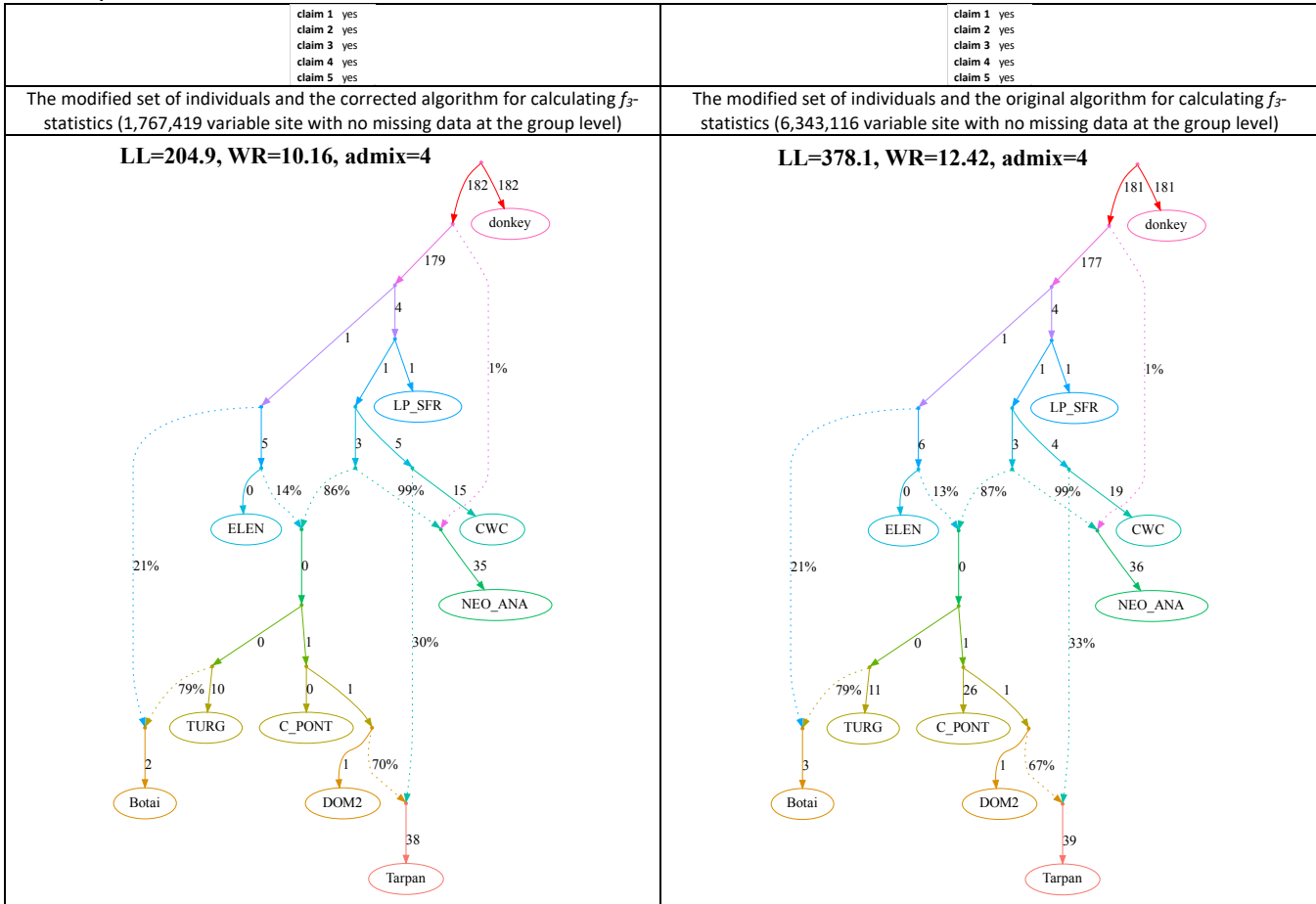
**b**, an alternative model with 3 admixture events fitting significantly better than the published one



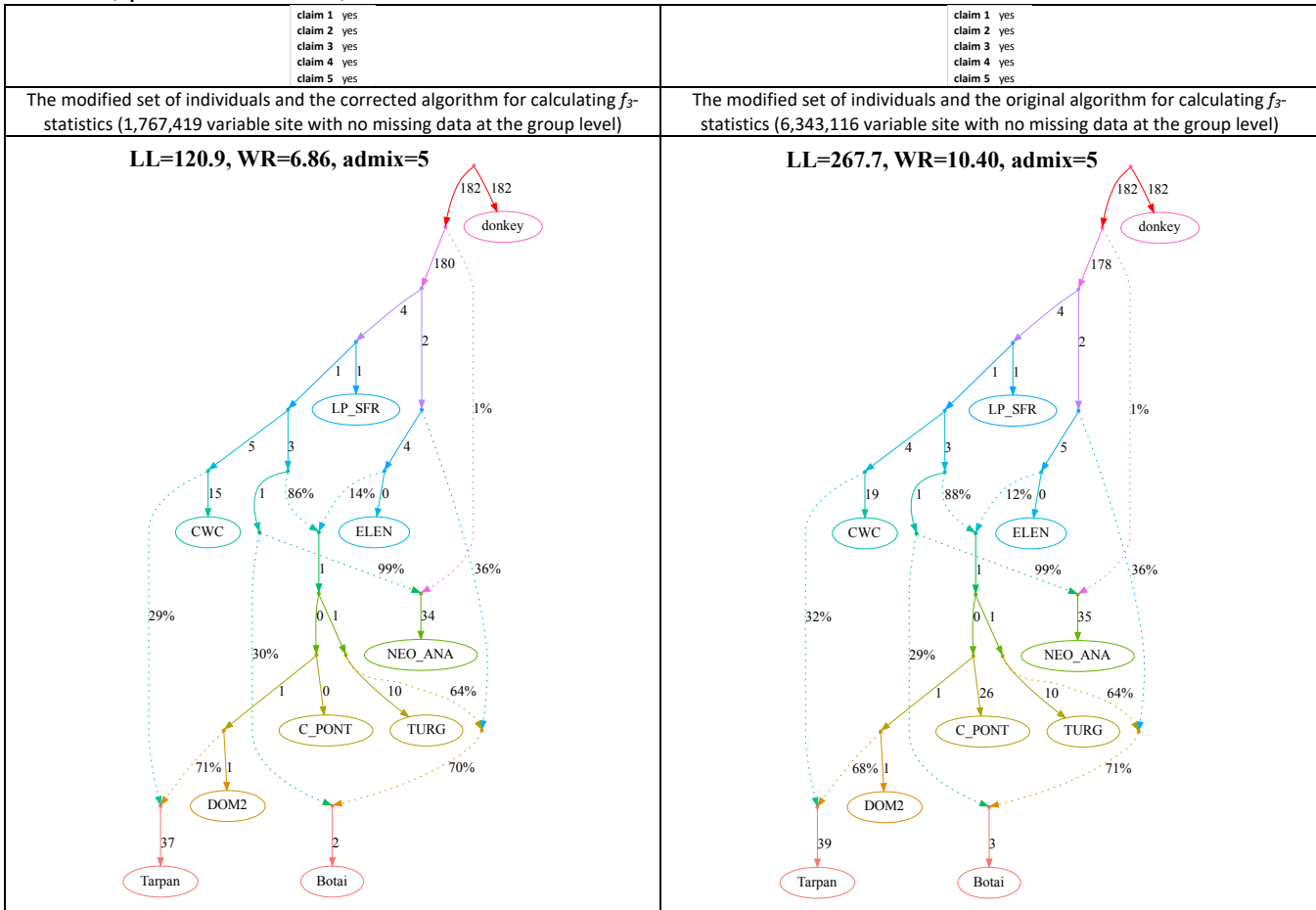
c, published model, 4 admixture events



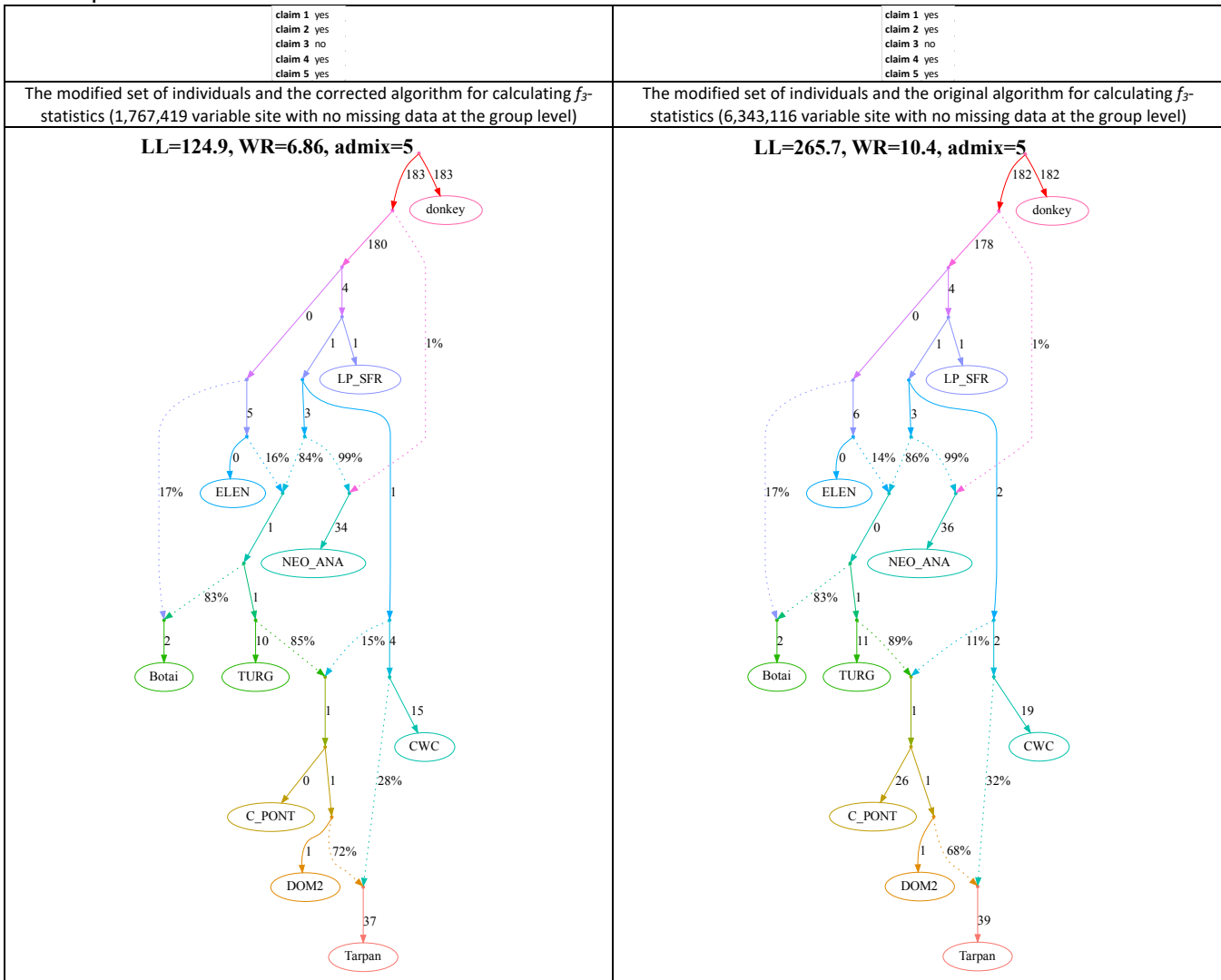
d, an alternative model with 4 admixture events fitting not significantly worse than the published one



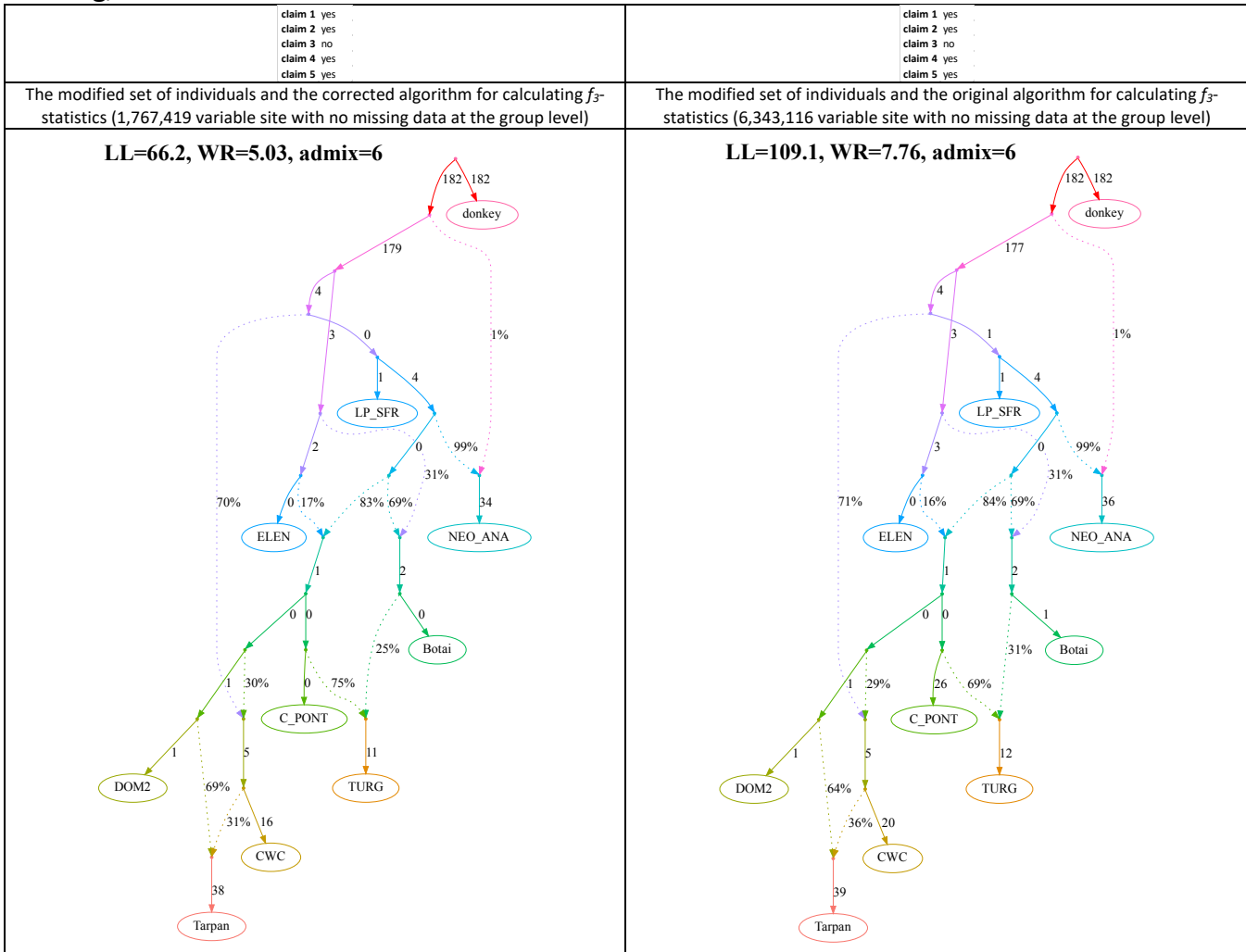
e, published model, 5 admixture events



**f**, an alternative model with 5 admixture events fitting not significantly worse than the published one



**g, selected models with 6 admixture events**



claim 1 no  
 claim 2 yes  
 claim 3 no  
 claim 4 no  
 claim 5 yes

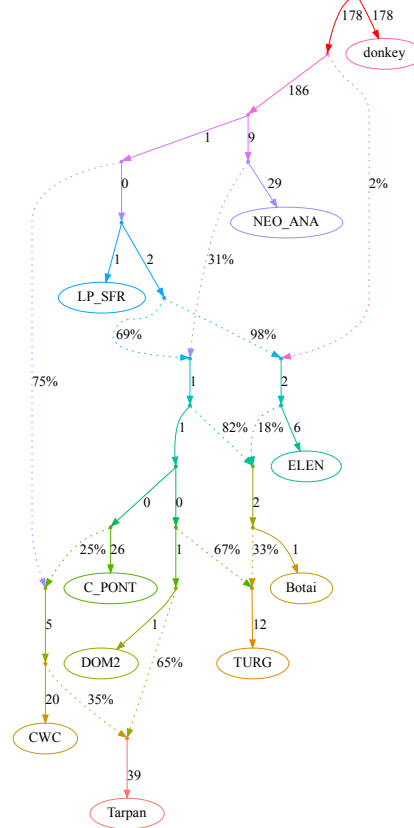
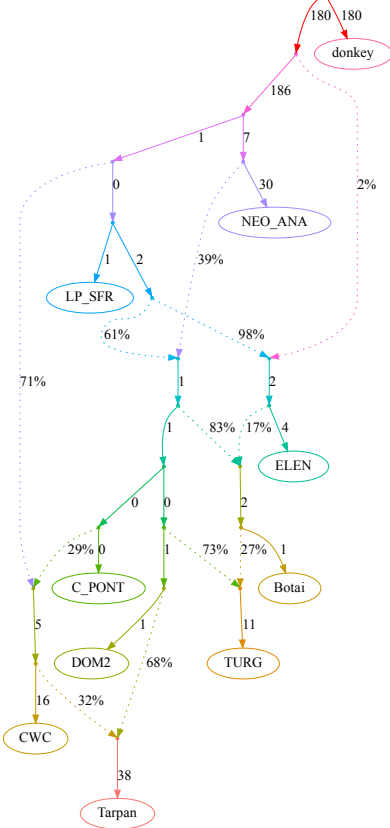
claim 1 no  
 claim 2 yes  
 claim 3 no  
 claim 4 no  
 claim 5 yes

The modified set of individuals and the corrected algorithm for calculating  $f_3$ -statistics (1,767,419 variable site with no missing data at the group level)

The modified set of individuals and the original algorithm for calculating  $f_3$ -statistics (6,343,116 variable site with no missing data at the group level)

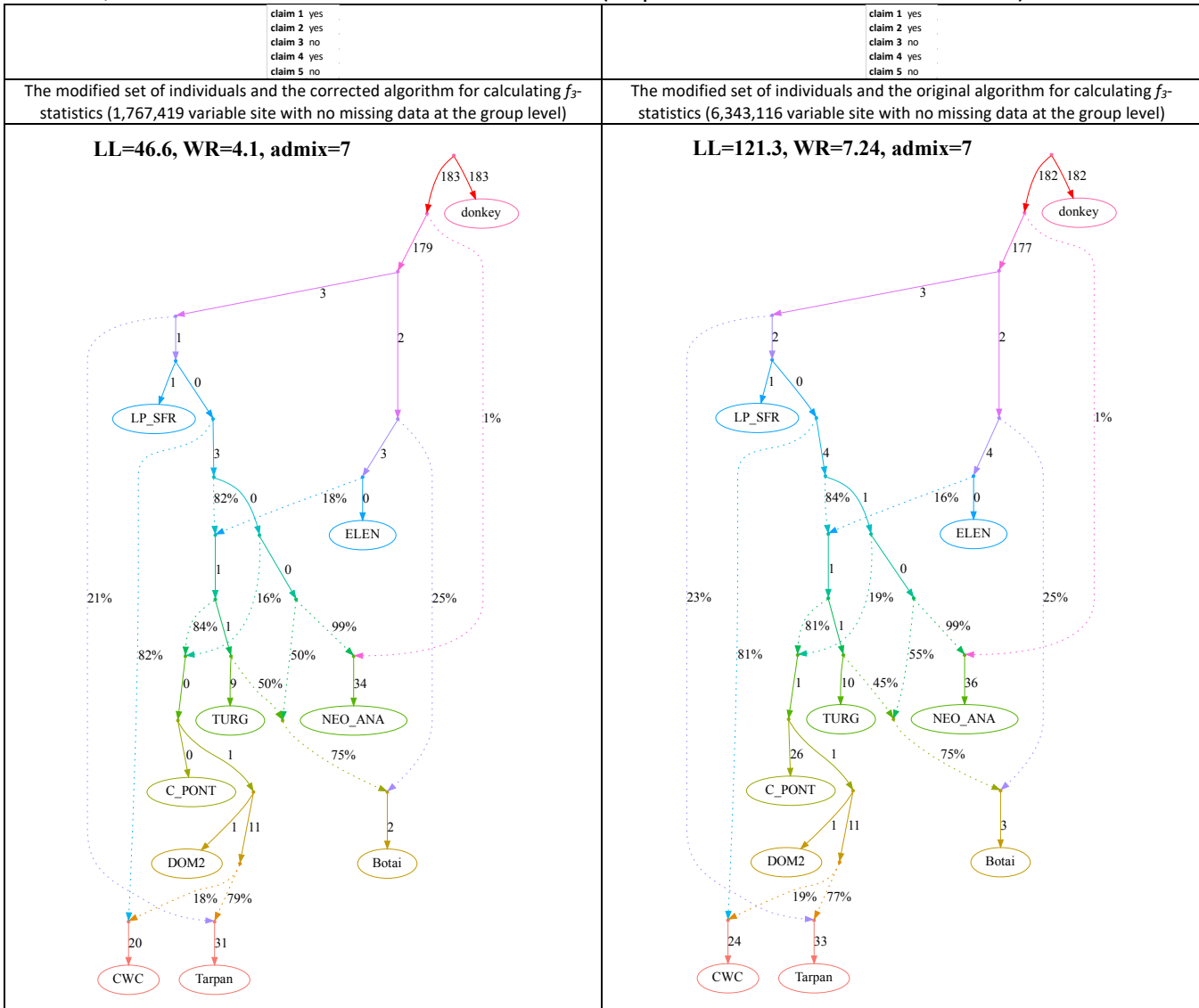
**LL=68, WR=5.28, admix=6**

**LL=131.6, WR=8.1, admix=6**





**h, selected models with 7 admixture events (all plausible models with WR < 5 SE)**



claim 1 yes  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 yes

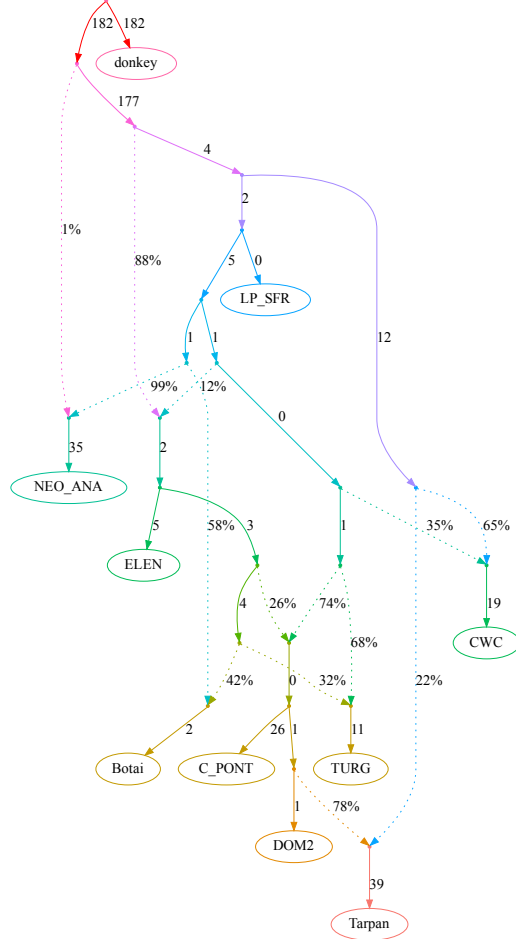
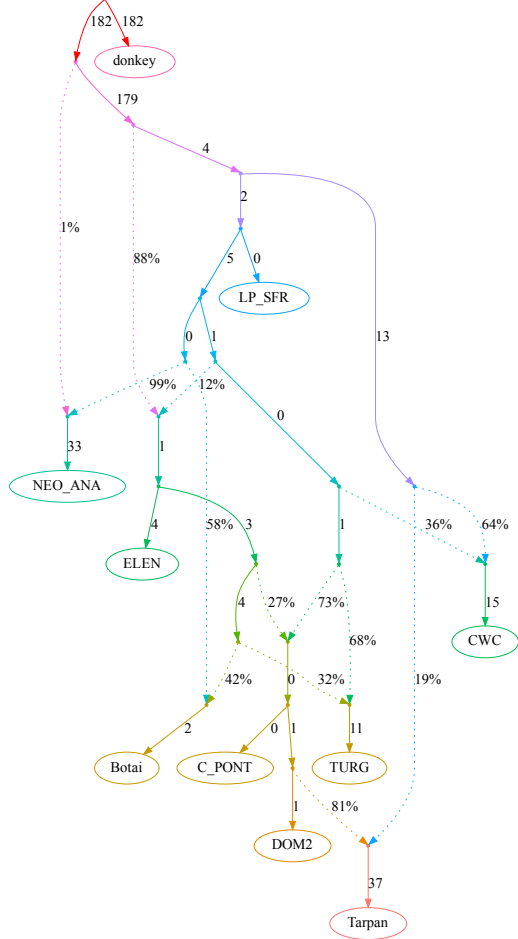
claim 1 yes  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 yes

The modified set of individuals and the corrected algorithm for calculating  $f_3$ -statistics (1,767,419 variable site with no missing data at the group level)

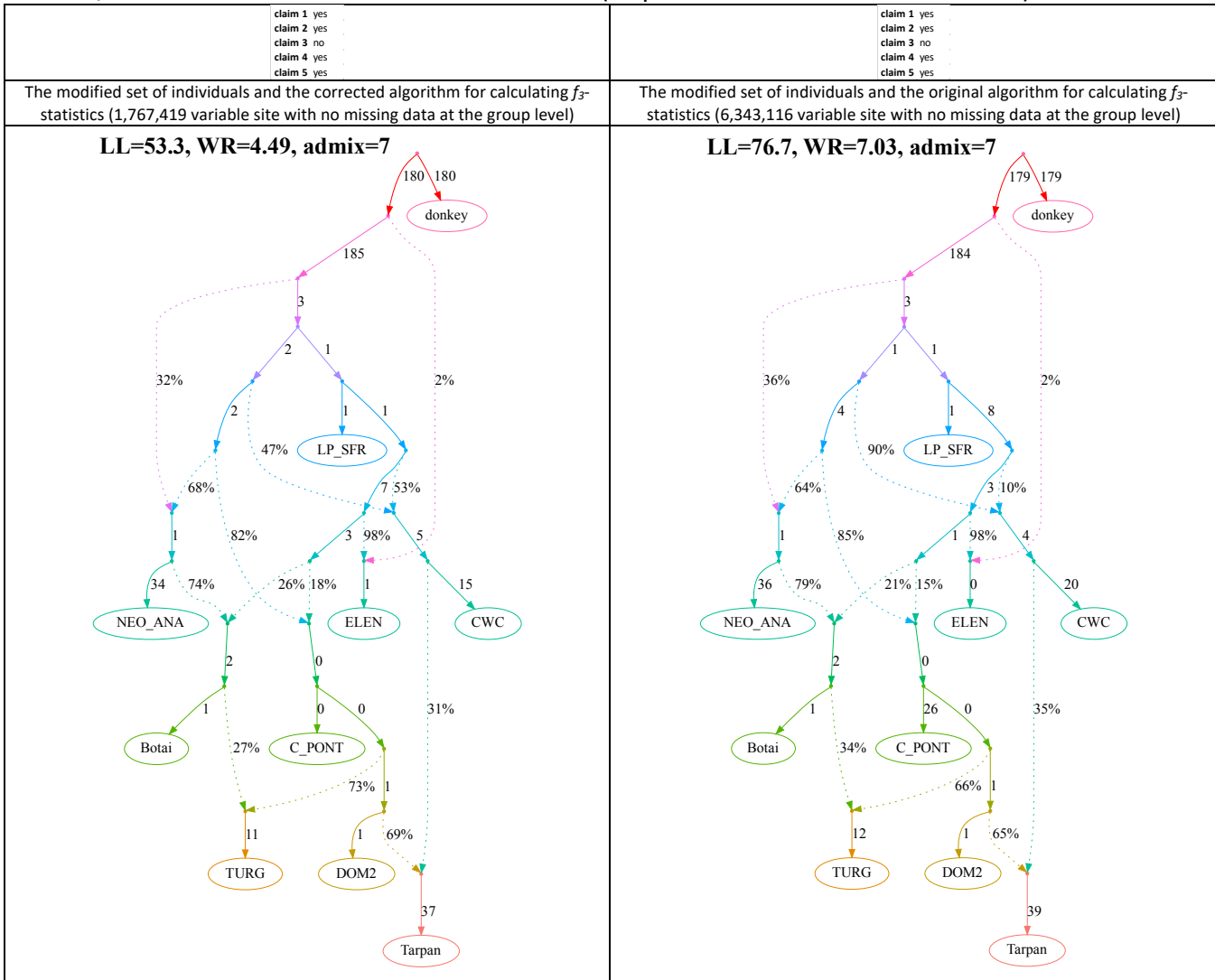
The modified set of individuals and the original algorithm for calculating  $f_3$ -statistics (6,343,116 variable site with no missing data at the group level)

**LL=36, WR=4.12, admix=7**

**LL=45.9, WR=4.48, admix=7**



i, selected models with 7 admixture events (all plausible models with  $WR < 5 SE$ )



claim 1 no  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 no

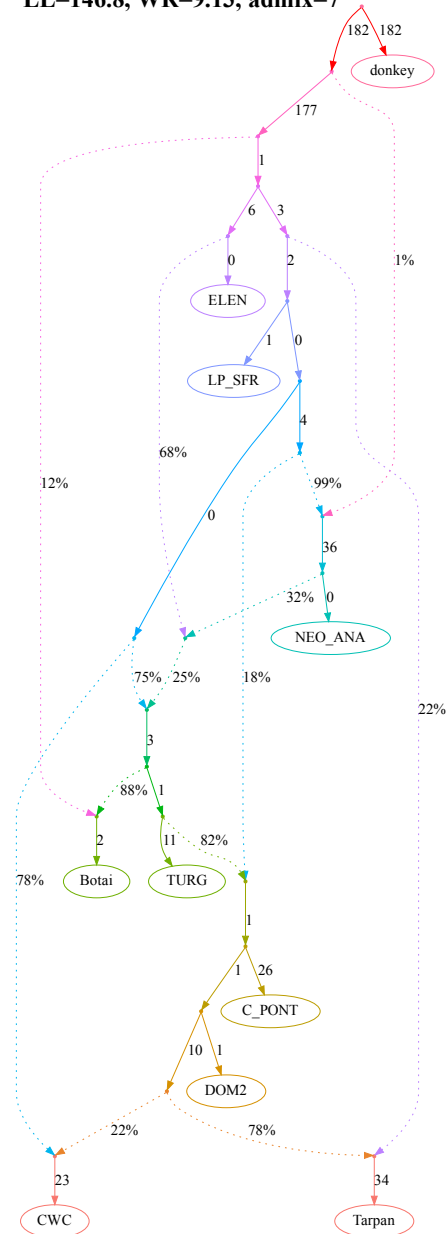
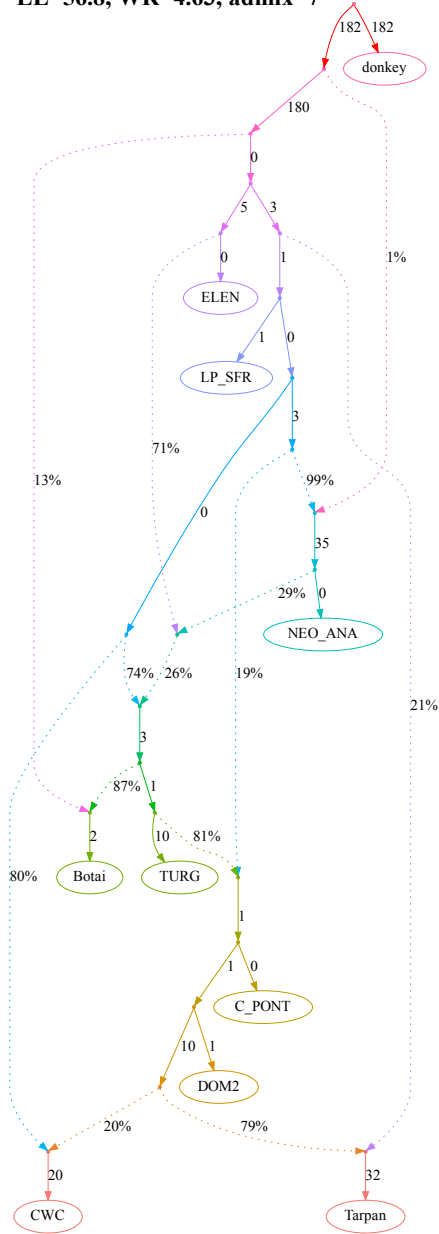
claim 1 no  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 no

The modified set of individuals and the corrected algorithm for calculating  $f_3$ -statistics (1,767,419 variable site with no missing data at the group level)

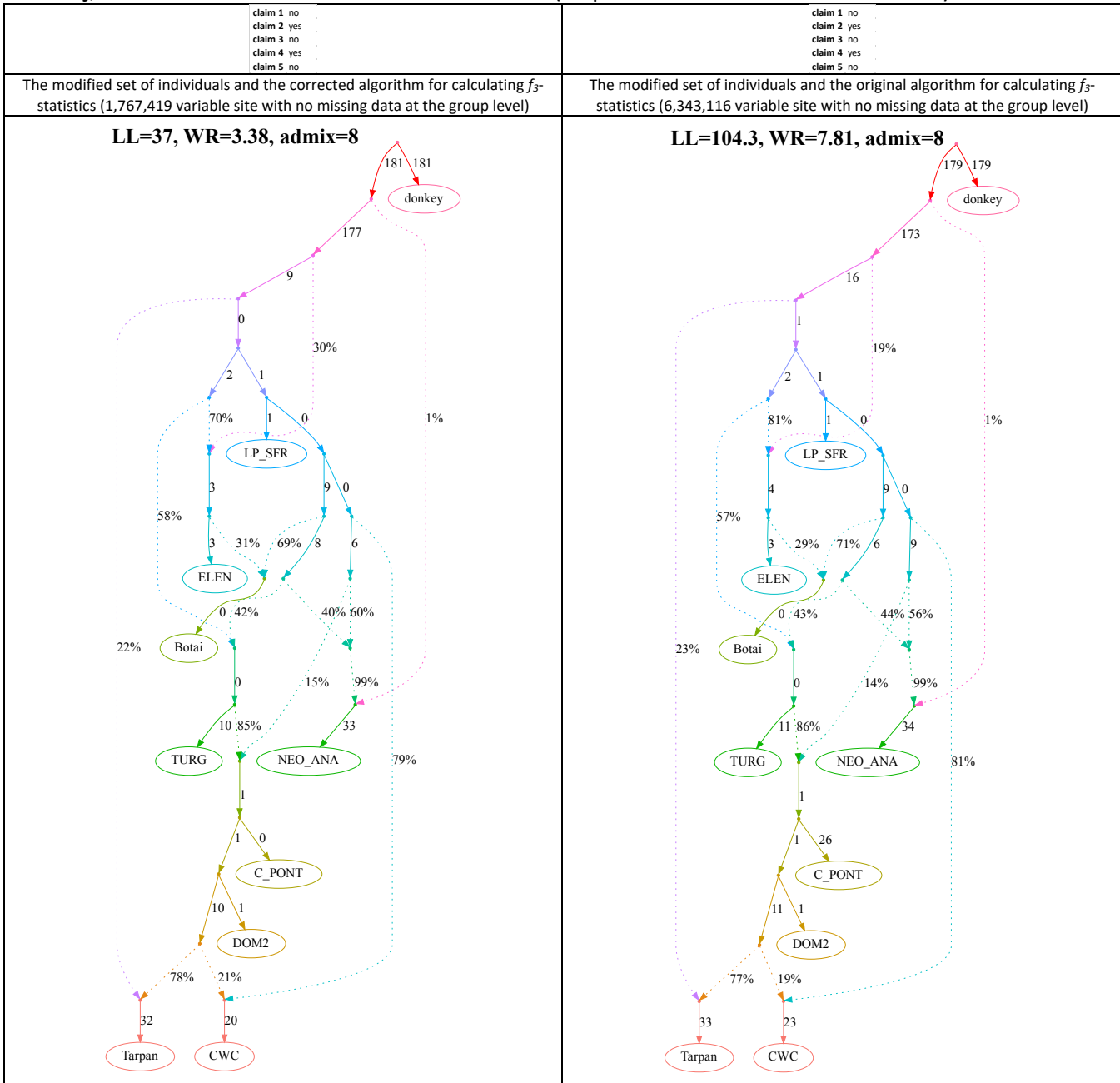
The modified set of individuals and the original algorithm for calculating  $f_3$ -statistics (6,343,116 variable site with no missing data at the group level)

**LL=56.8, WR=4.63, admix=7**

**LL=146.8, WR=9.13, admix=7**



**j**, selected models with 8 admixture events (all plausible models with  $WR < 4 SE$ )



claim 1 no  
 claim 2 yes  
 claim 3 yes  
 claim 4 yes  
 claim 5 no

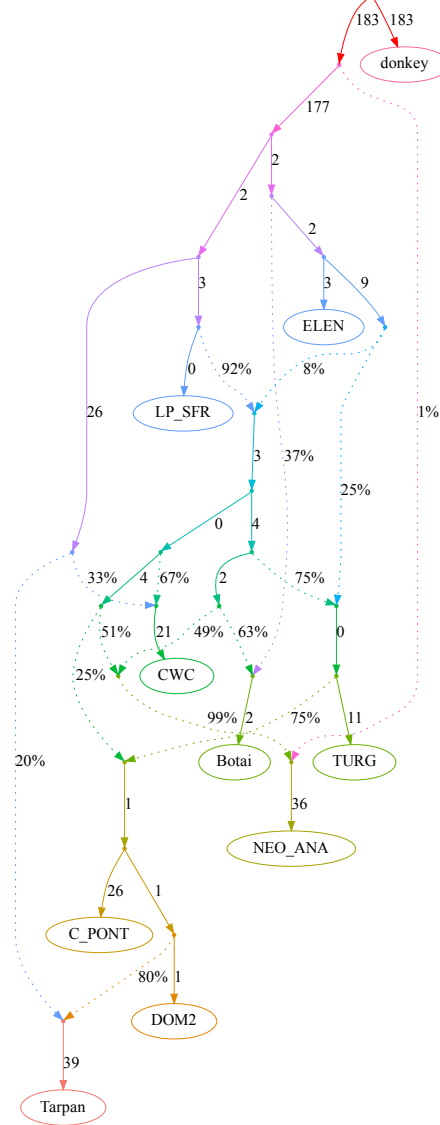
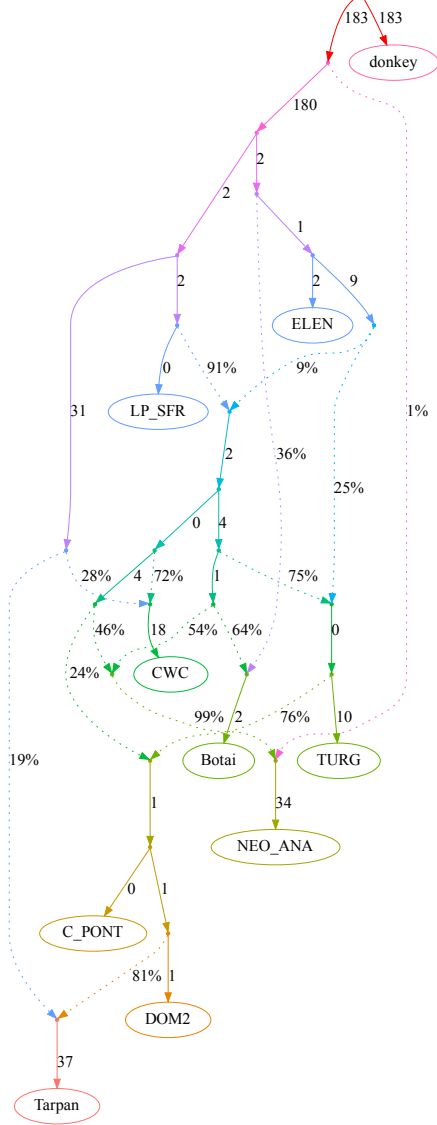
claim 1 no  
 claim 2 yes  
 claim 3 yes  
 claim 4 yes  
 claim 5 no

The modified set of individuals and the corrected algorithm for calculating  $f_3$ -statistics (1,767,419 variable site with no missing data at the group level)

The modified set of individuals and the original algorithm for calculating  $f_3$ -statistics (6,343,116 variable site with no missing data at the group level)

**LL=36.5, WR=3.68, admix=8**

**LL=52.2, WR=5.84, admix=8**



k, selected models with 8 admixture events (all plausible models with WR < 4 SE)



claim 1 yes  
 claim 2 yes  
 claim 3 yes  
 claim 4 yes  
 claim 5 yes

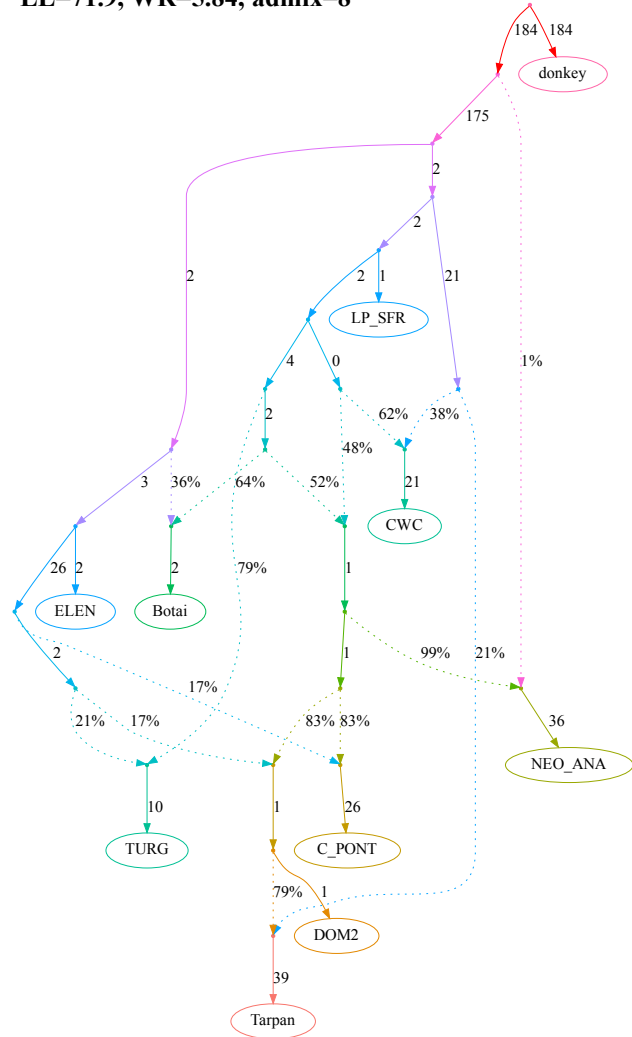
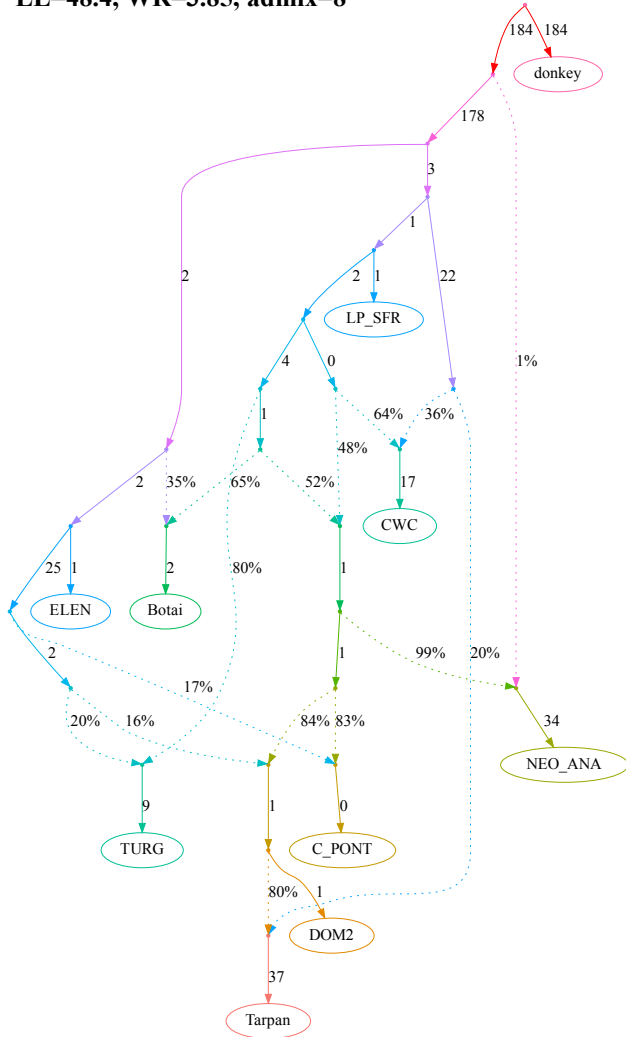
claim 1 yes  
 claim 2 yes  
 claim 3 yes  
 claim 4 yes  
 claim 5 yes

The modified set of individuals and the corrected algorithm for calculating  $f_3$ -statistics (1,767,419 variable site with no missing data at the group level)

The modified set of individuals and the original algorithm for calculating  $f_3$ -statistics (6,343,116 variable site with no missing data at the group level)

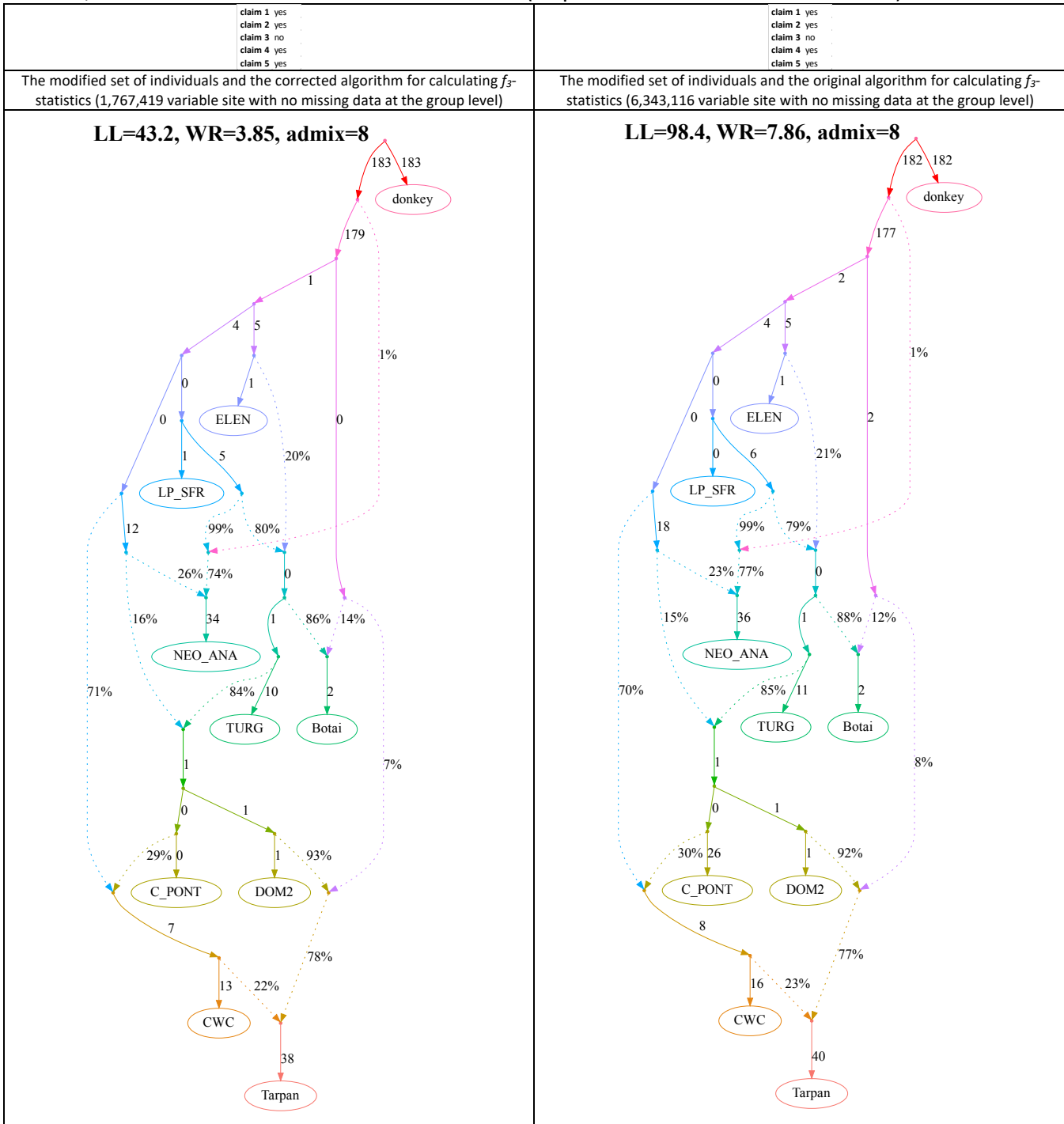
**LL=48.4, WR=3.85, admix=8**

**LL=71.9, WR=5.84, admix=8**

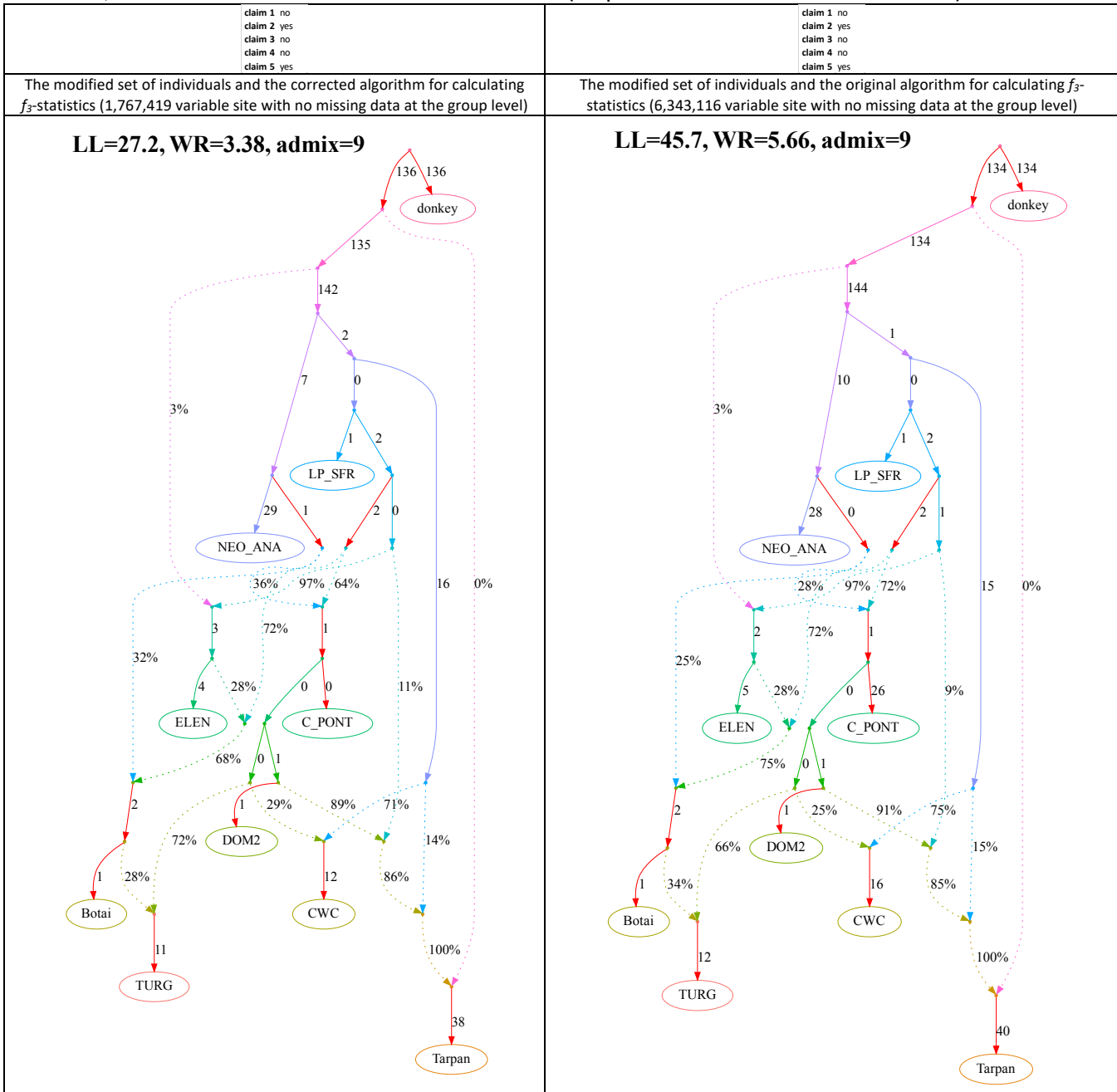




I, selected models with 8 admixture events (all plausible models with WR < 4 SE)



**m**, selected models with 9 admixture events (all plausible models with  $WR < 4 SE$ )



claim 1 no  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 no

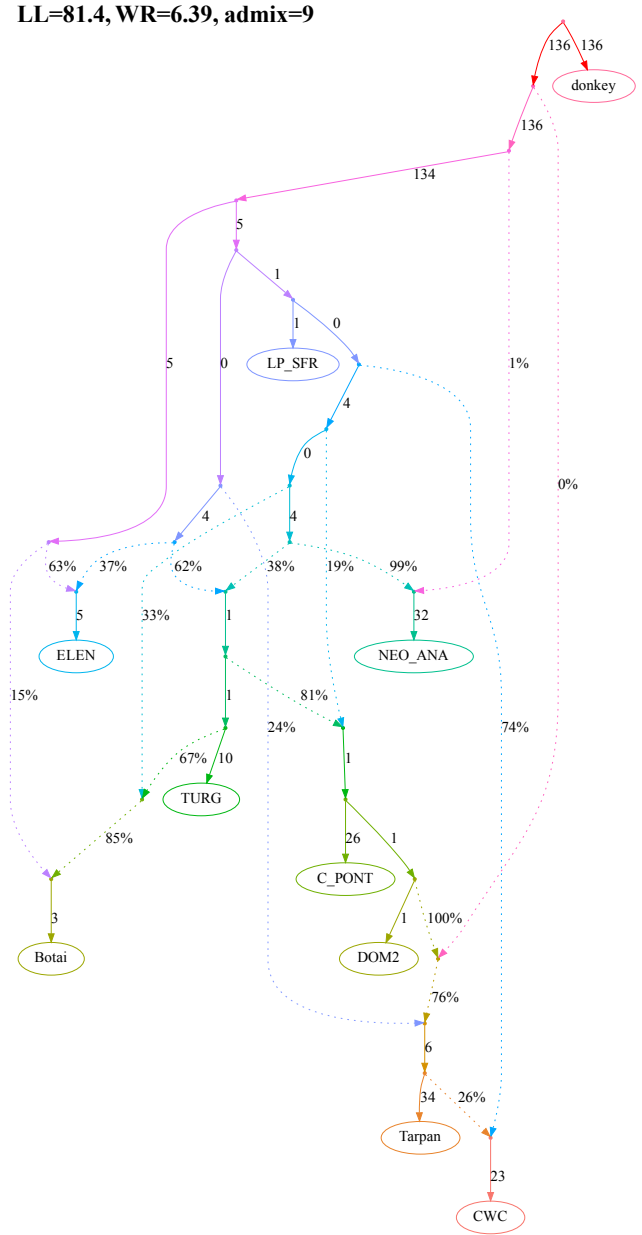
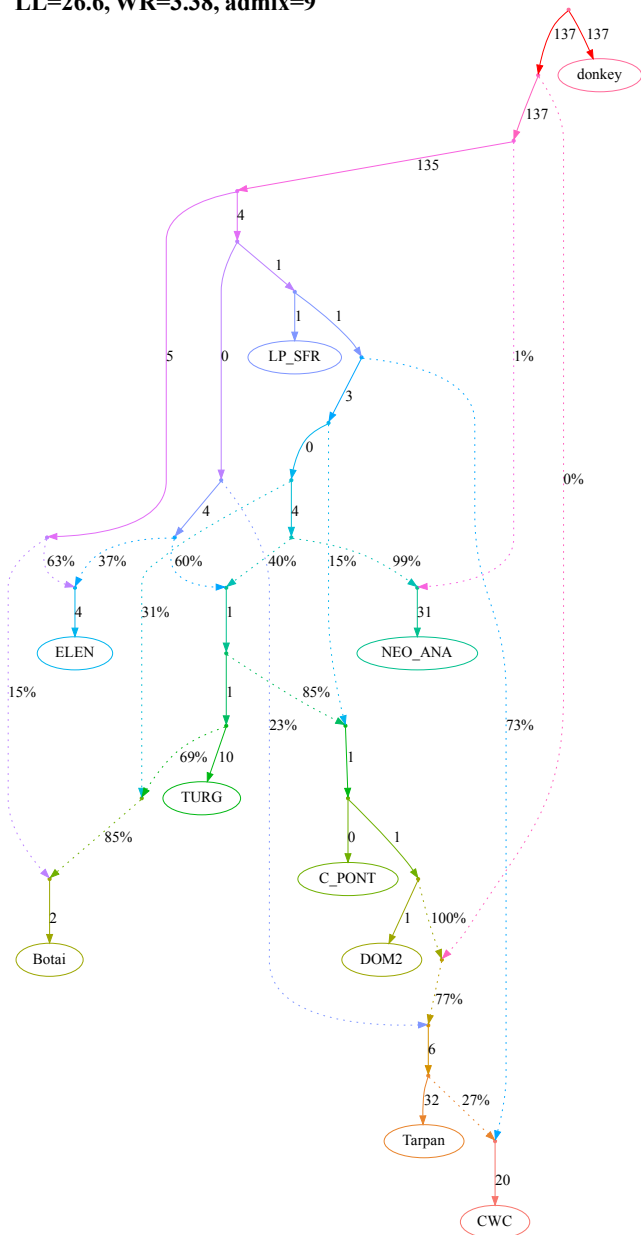
claim 1 no  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 no

The modified set of individuals and the corrected algorithm for calculating  $f_3$ -statistics (1,767,419 variable site with no missing data at the group level)

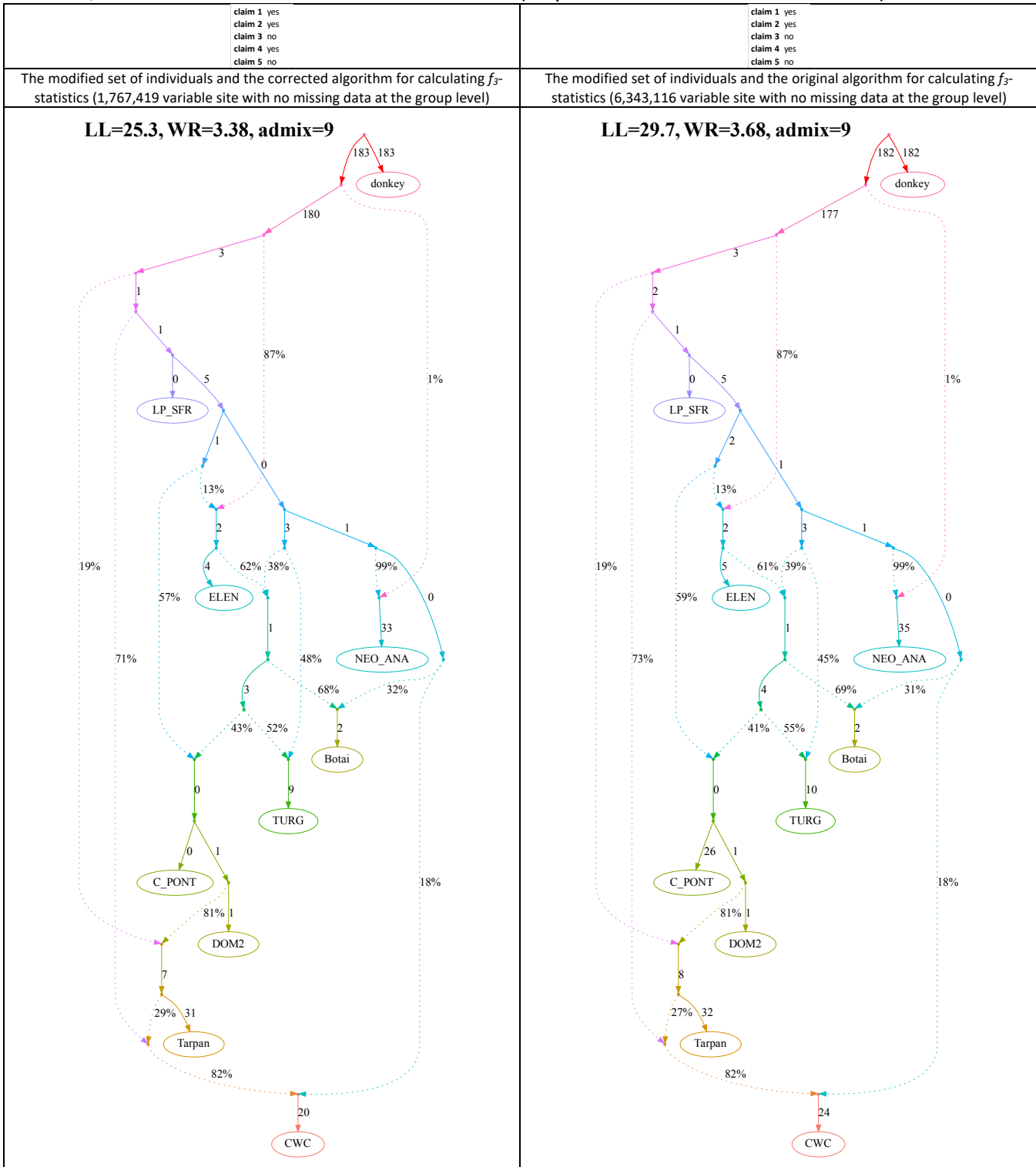
The modified set of individuals and the original algorithm for calculating  $f_3$ -statistics (6,343,116 variable site with no missing data at the group level)

**LL=26.6, WR=3.38, admix=9**

**LL=81.4, WR=6.39, admix=9**



n, selected models with 9 admixture events (all plausible models with WR < 4 SE)



claim 1 no  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 no

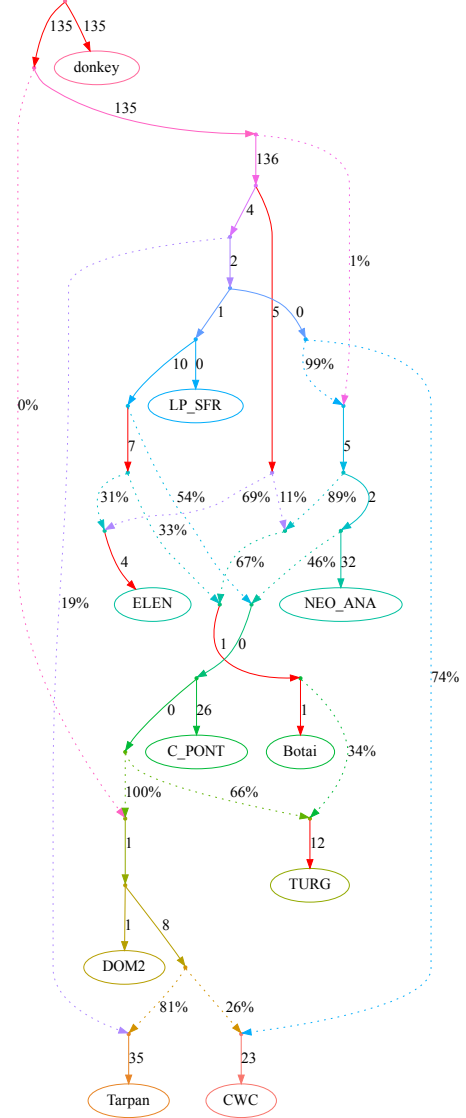
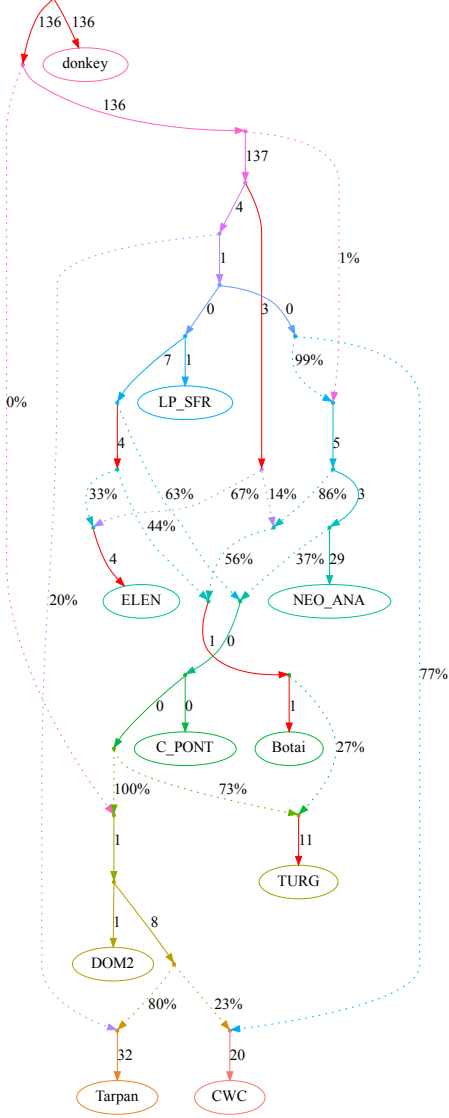
claim 1 no  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 no

The modified set of individuals and the corrected algorithm for calculating  $f_3$ -statistics (1,767,419 variable site with no missing data at the group level)

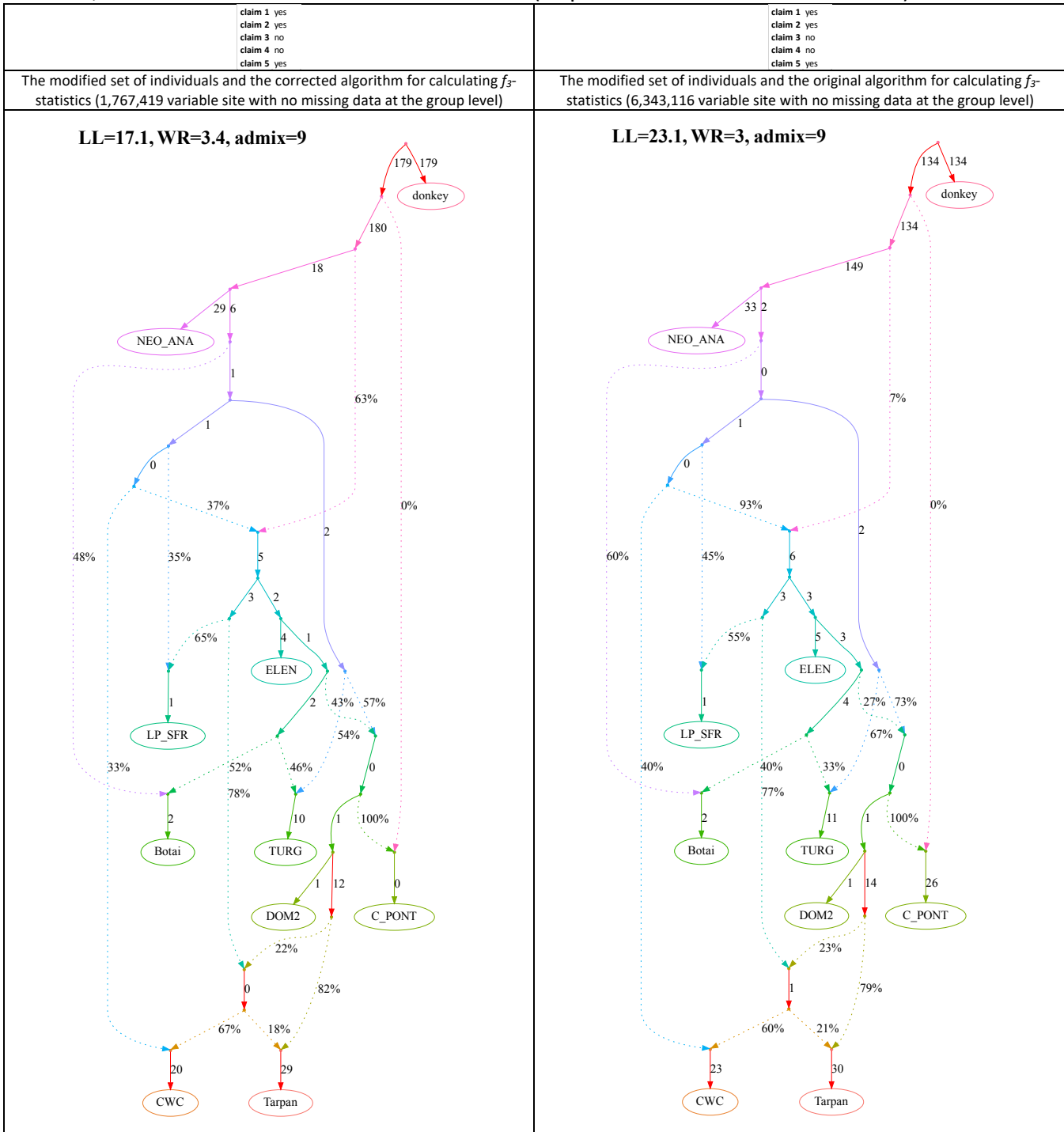
The modified set of individuals and the original algorithm for calculating  $f_3$ -statistics (6,343,116 variable site with no missing data at the group level)

**LL=26.5, WR=3.38, admix=9**

**LL=56.8, WR=5.64, admix=9**



o, selected models with 9 admixture events (all plausible models with WR < 4 SE)



claim 1 yes  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 no

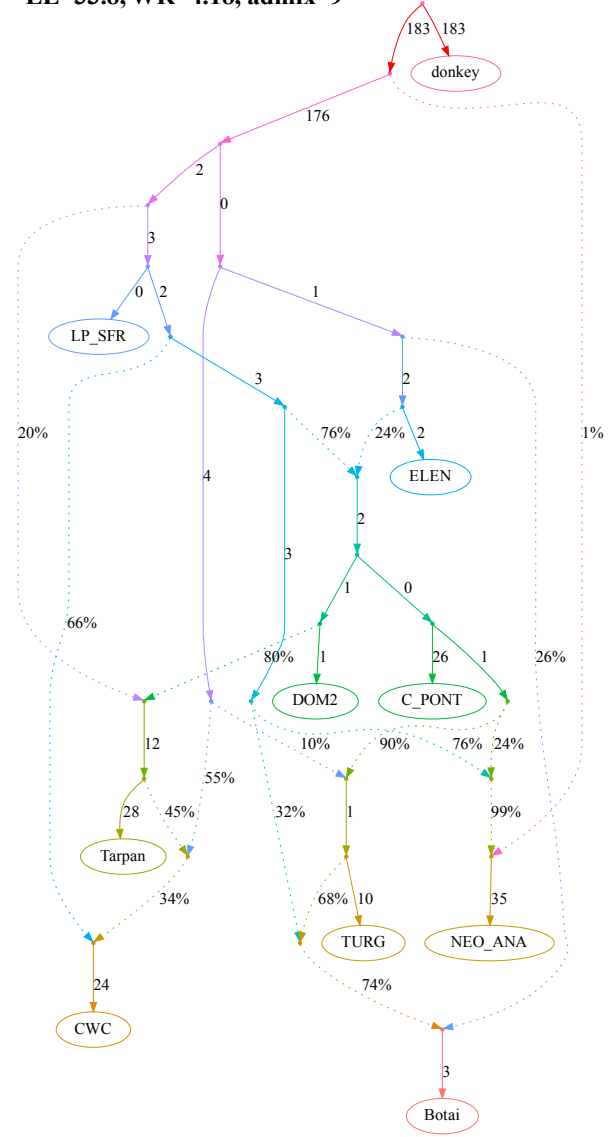
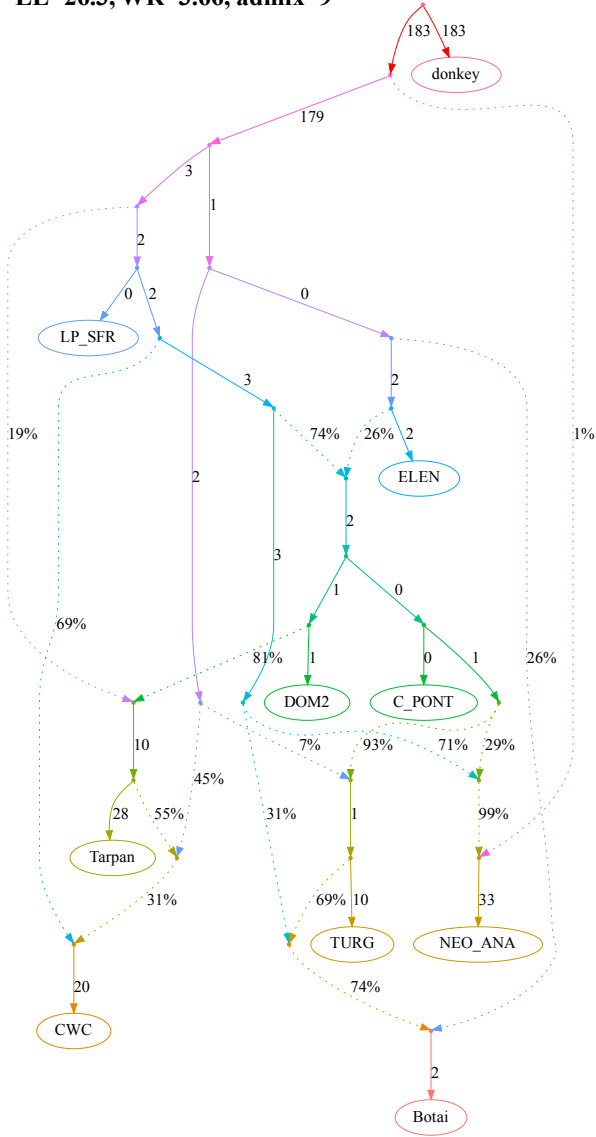
claim 1 yes  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 no

The modified set of individuals and the corrected algorithm for calculating  $f_3$ -statistics (1,767,419 variable site with no missing data at the group level)

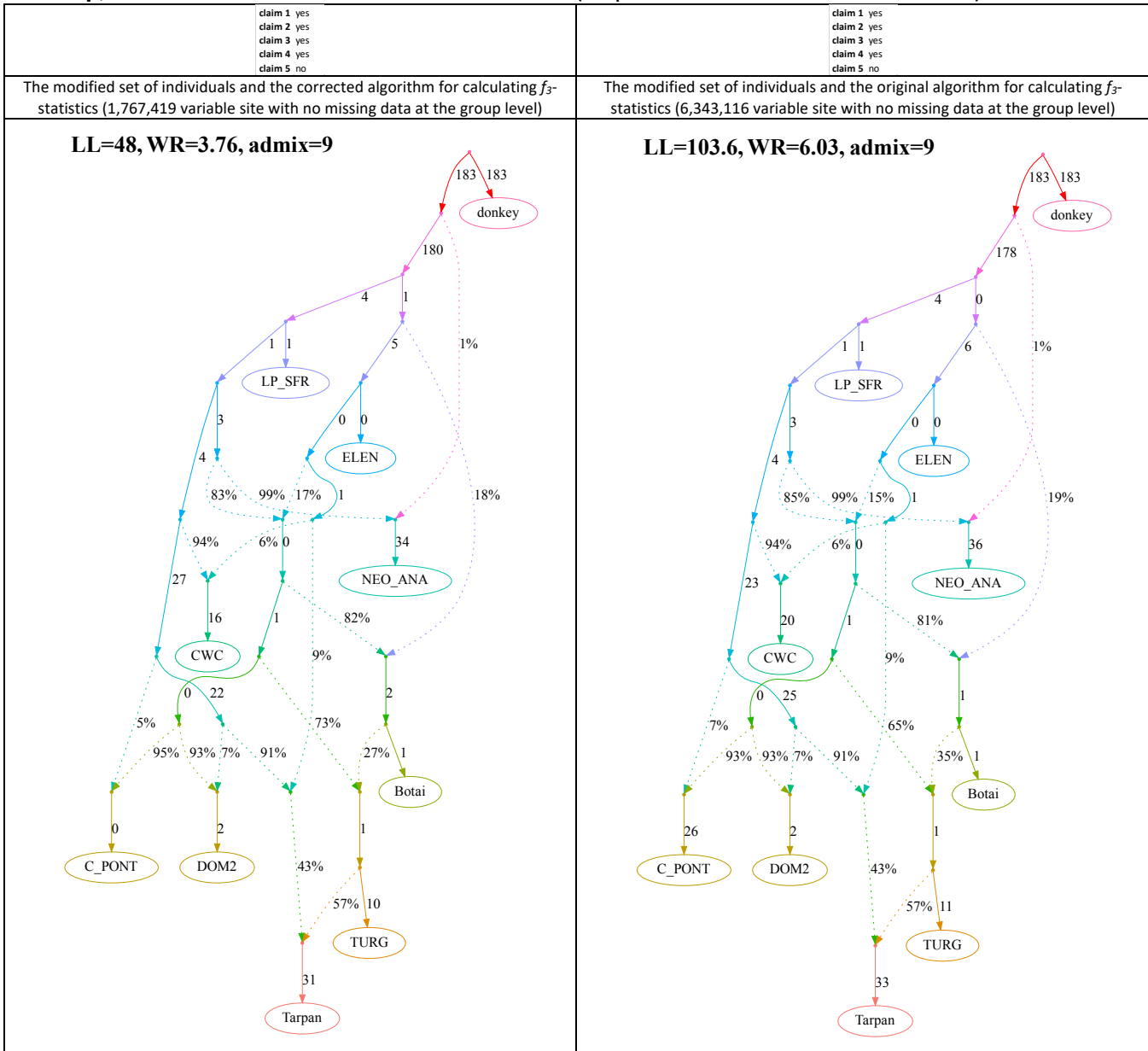
The modified set of individuals and the original algorithm for calculating  $f_3$ -statistics (6,343,116 variable site with no missing data at the group level)

**LL=26.3, WR=3.66, admix=9**

**LL=33.8, WR=4.18, admix=9**



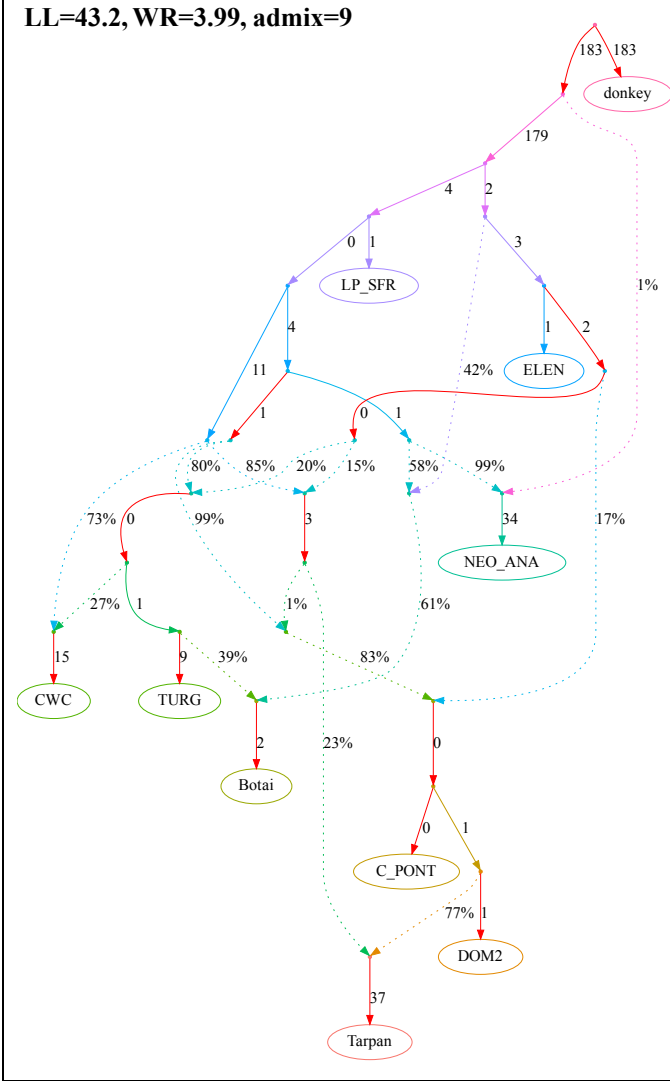
**p**, selected models with 9 admixture events (all plausible models with WR < 4 SE)





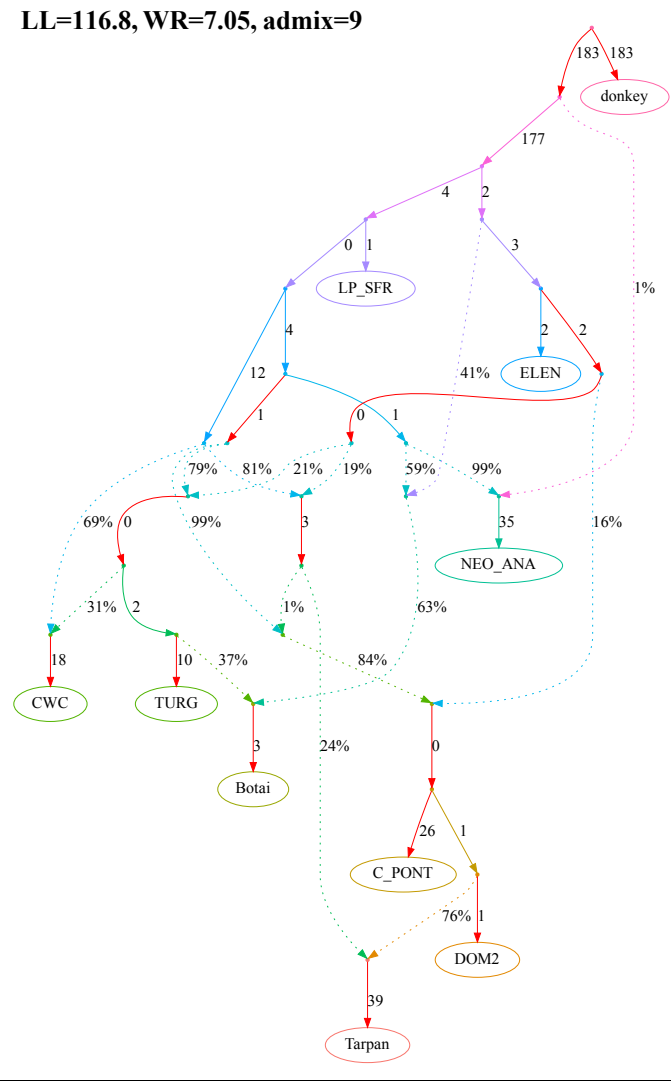
claim 1 yes  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 yes

The modified set of individuals and the corrected algorithm for calculating  $f_3$ -statistics (1,767,419 variable site with no missing data at the group level)

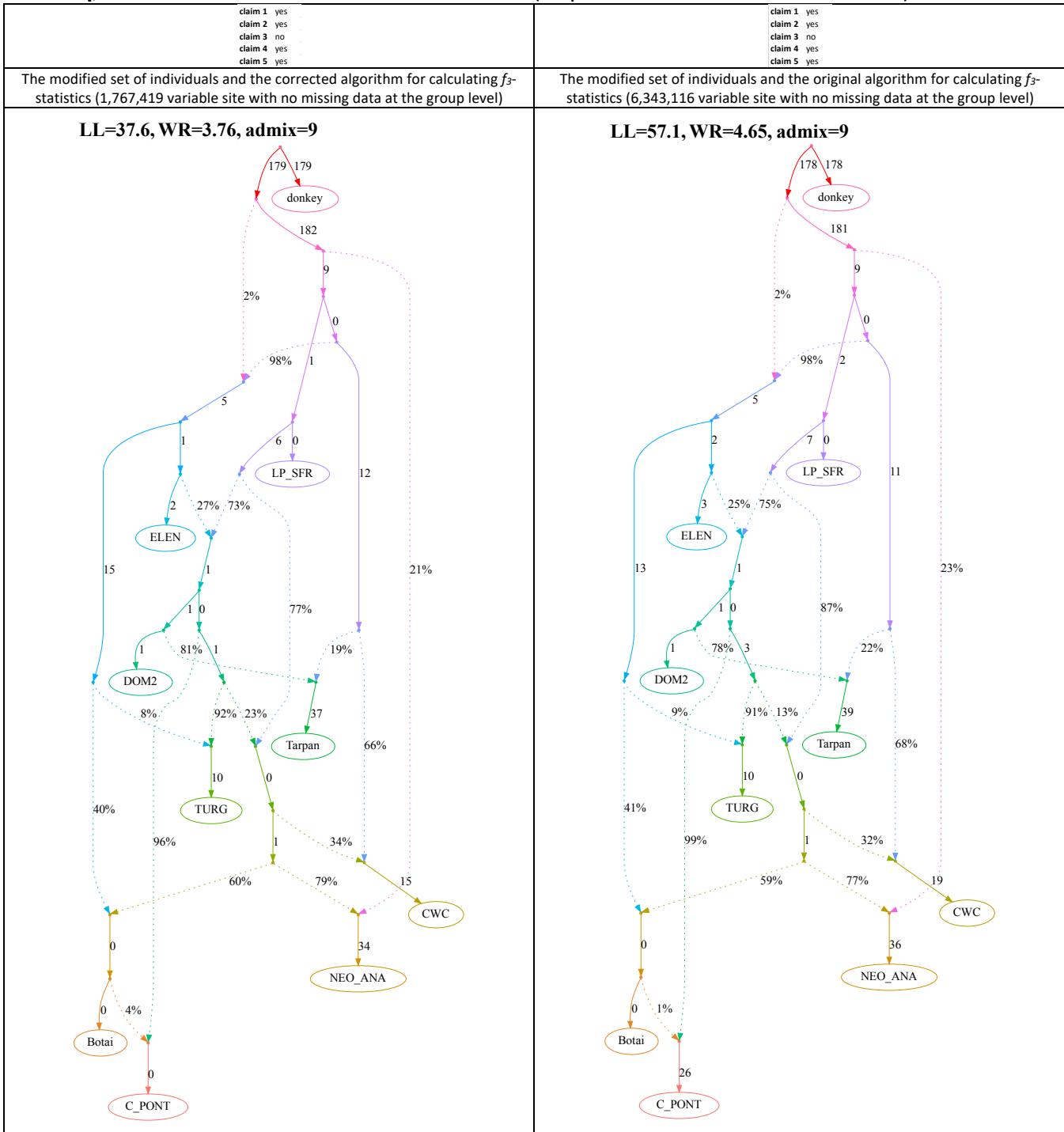


claim 1 yes  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 yes

The modified set of individuals and the original algorithm for calculating  $f_3$ -statistics (6,343,116 variable site with no missing data at the group level)



**q**, selected models with 9 admixture events (all plausible models with WR < 4 SE)



claim 1 yes  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 no

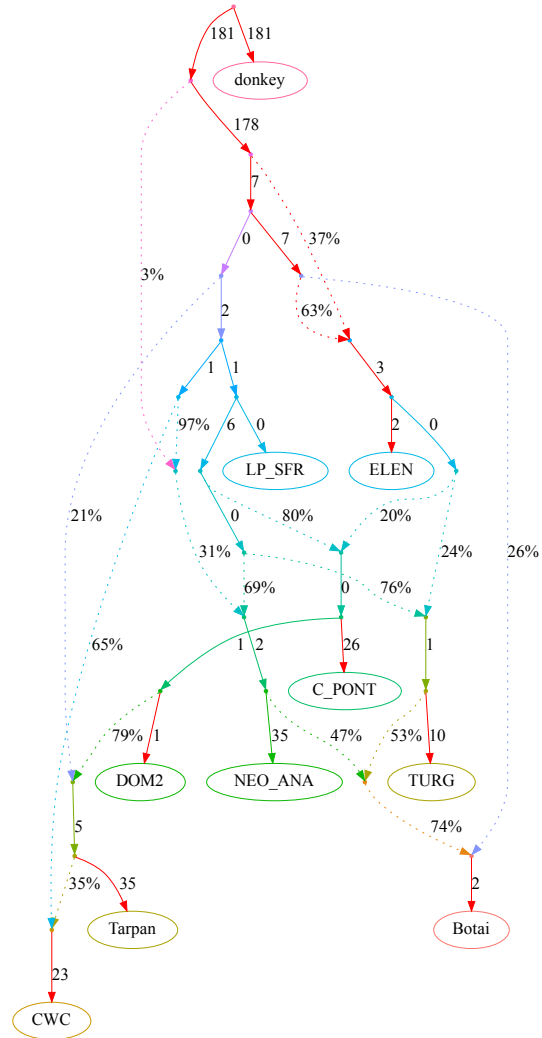
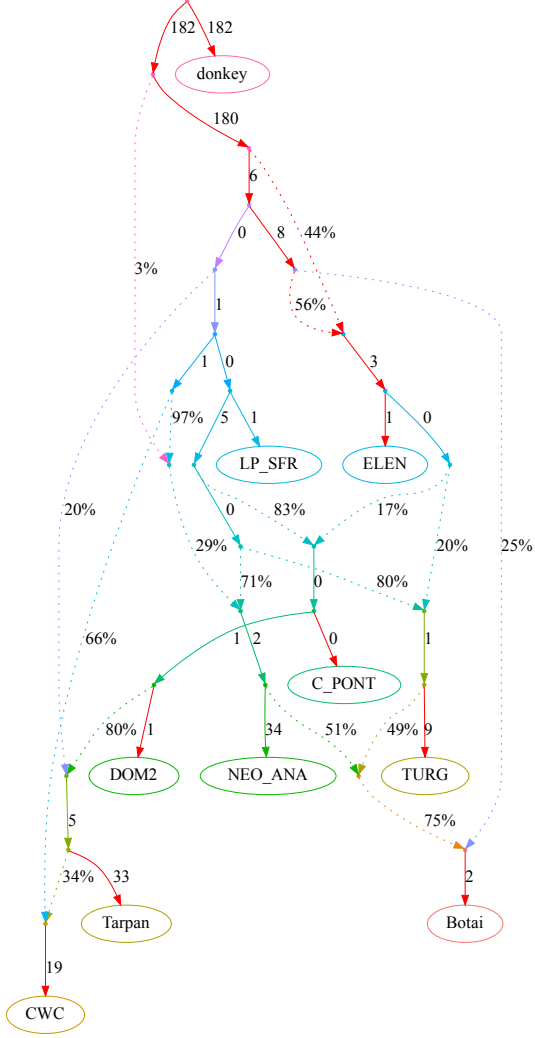
claim 1 yes  
 claim 2 yes  
 claim 3 no  
 claim 4 yes  
 claim 5 no

The modified set of individuals and the corrected algorithm for calculating  $f_3$ -statistics (1,767,419 variable site with no missing data at the group level)

The modified set of individuals and the original algorithm for calculating  $f_3$ -statistics (6,343,116 variable site with no missing data at the group level)

**LL=31.3, WR=3.78, admix=9**

**LL=66, WR=5.61, admix=9**



r, selected models with 9 admixture events (all plausible models with WR < 4 SE)

