Additional notes concerning the samples.

Our Tharu samples D254 and D260 have been previously published by Reich et al. as Tharu\_107\_R1 and Tharu\_108\_R1, respectively.

Our Velama sample VELZ260 has been previously published by Reich et al as Velama\_184\_R2

Our sample labelled D238 and reported as Tharu is in fact a Brahmin sample from Uttar Pradesh (see update of the Supplementary table 1).

There are four first degree relatives groups in our sample: [Dusadh A393, A394, A395]; [Dusadh A387, A388]; [Dharkar HA030, HA039] and [Kanjar evo\_32, Kol 321e]. Following the publication we have identified that sample evo\_32 was erroneously labelled as Kanjar before any genetic analyses. We hereby re-label the sample as belonging to Kol population (see update of the Supplementary table 1).

There are three second degree relatives groups in our sample: [Dusadh A392 is equally related to the three Dusadh samples A393, A394 and A395]; [Dharkars HA039, HA041]; [Kanjar evo\_37 and Dharkar HA023]. Again the last pair needs further explanation. The Dharkar and Kanjar practice a nomadic lifestyle and were living side by side at the time of sampling. As the ethnic border between the two is permeable we cannot rule out neither our error during sample collection and/or subsequent labelling nor shifted self-identity.

The identified issues of mislabelling do not have any notable impact on the results as discussed in the published article.

Similarly the few relatives have only marginal effect at the level of detail not considered in the paper. These effects include the emergence of Dusadh relatives group specific ancestry component starting k=14 in the Admixture analysis.

References

Reich, D., Thangaraj, K., Patterson, N., Price, A.L., and Singh, L. (2009). Reconstructing Indian population history. Nature 461, 489–494.