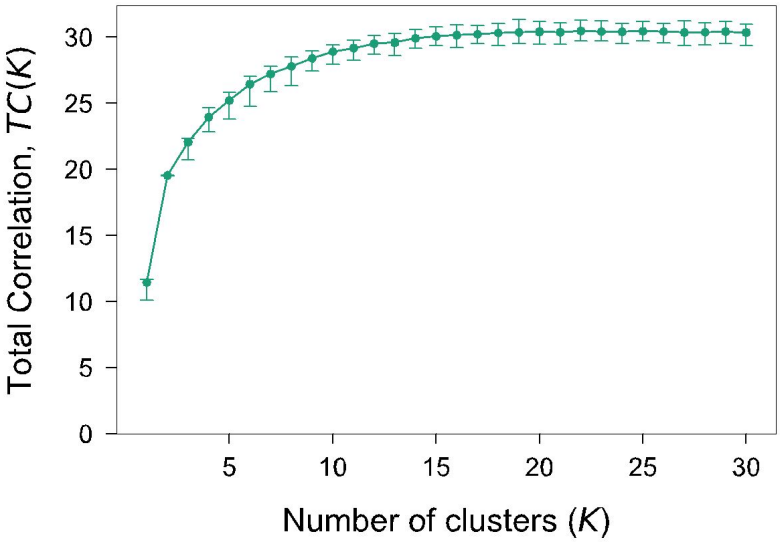
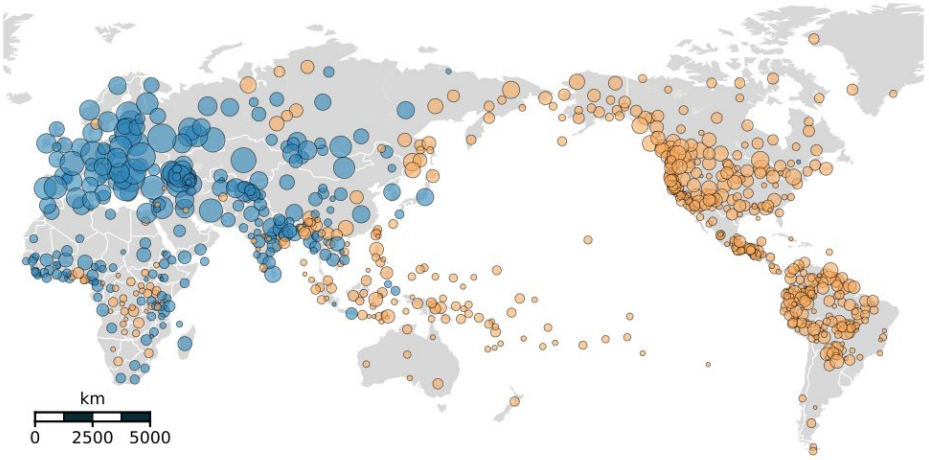
Results for the new data in “maps of myths”

- 958 traditions with their geographical locations (longitude and latitude)
- 2583 motifs
    - “I120a” is duplicated. It has two ids, 103927 / 103940
    - 103940 is taken as it has more traditions than 103927
- Out of 958 traditions, 901 traditions that have at least ten motifs are selected for CorEx analysis. Thus, the size of the new binary matrix is 901 by 2583.
- Previously, we use the data of 683 traditions and 1638 motifs.

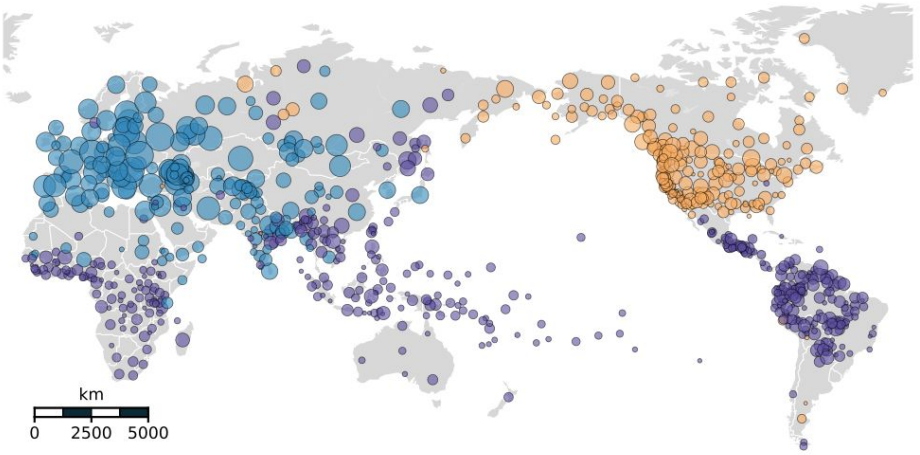
1. Total correlation still saturates after ten clusters



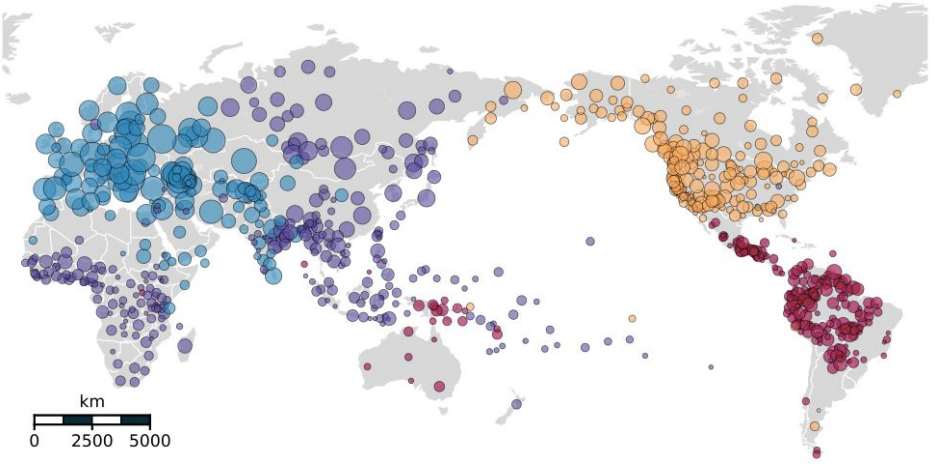
2. CorEx clusterings (K=2, 3, 4, …, 12)
    - K=2



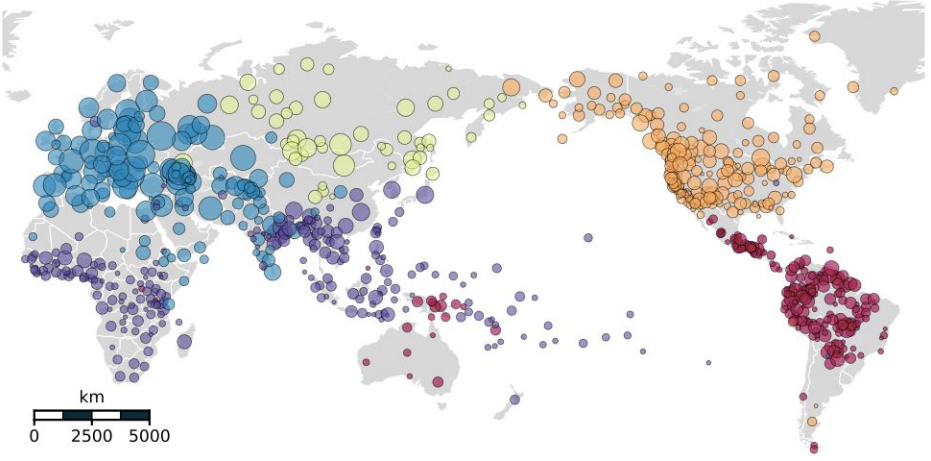
- K=3



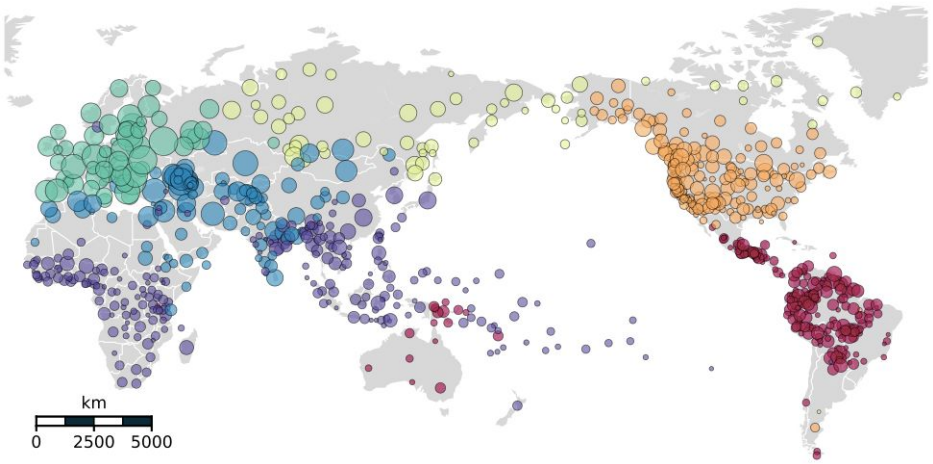
- K=4



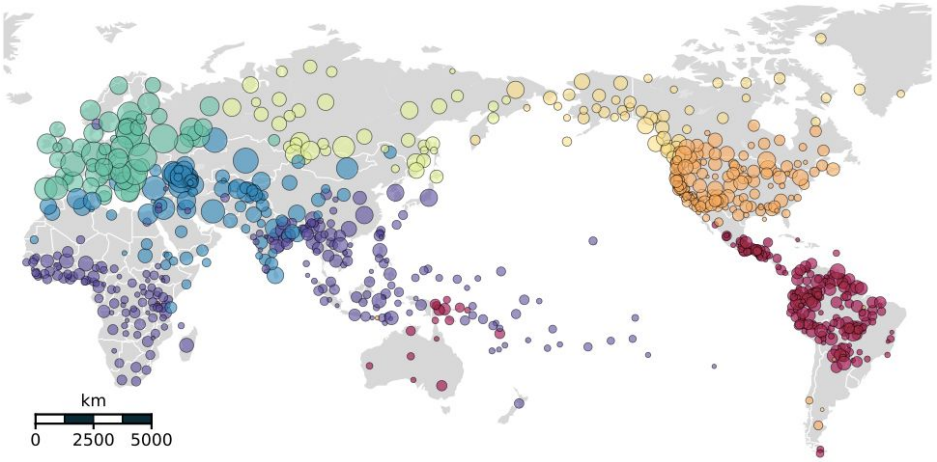
- K=5



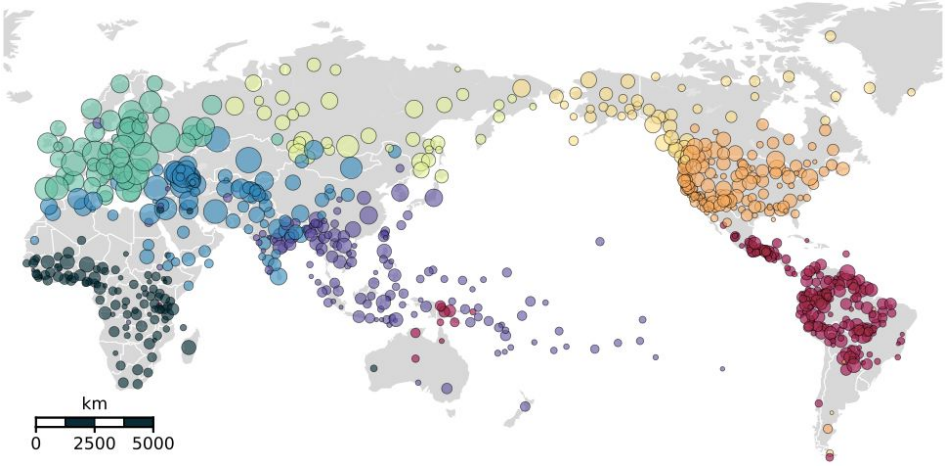
- K=6



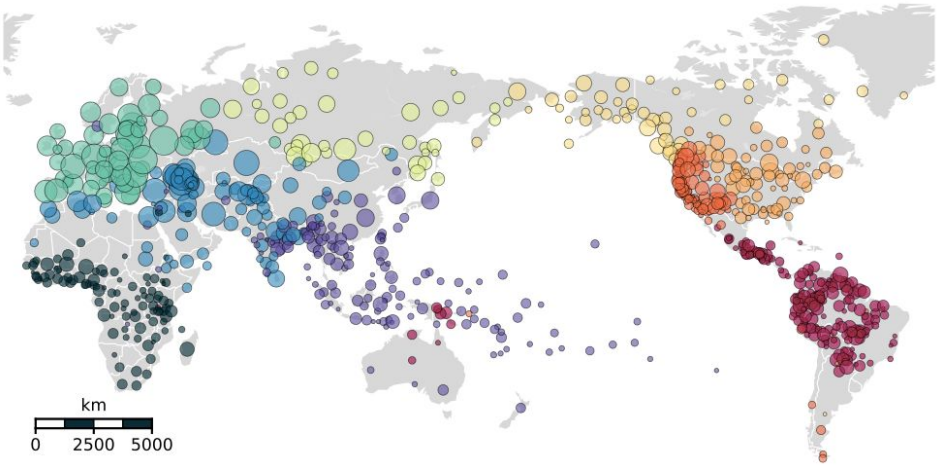
- K=7



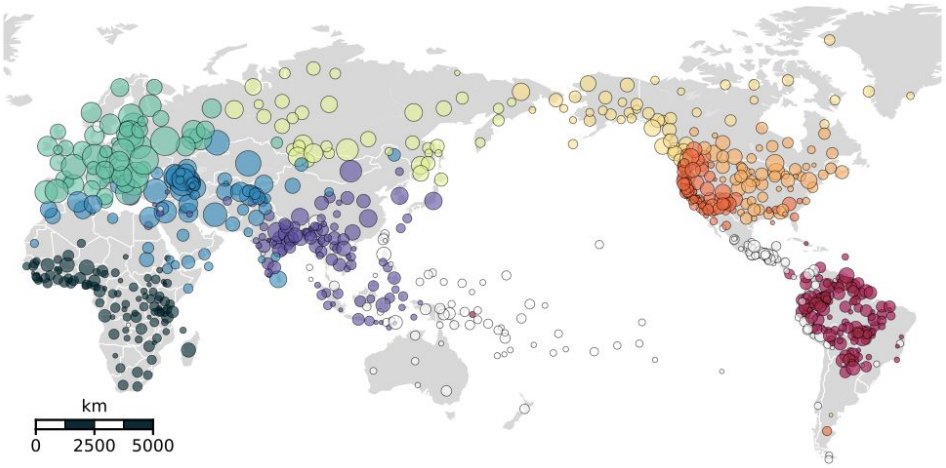
- K=8



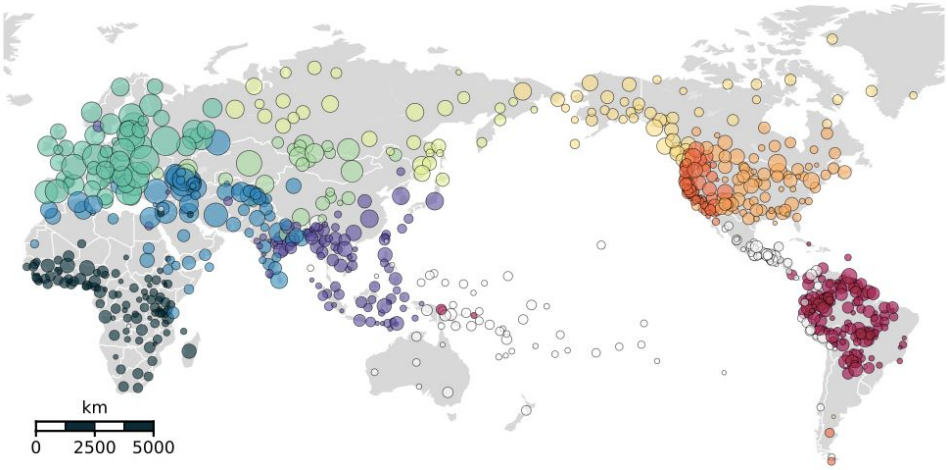
- K=9



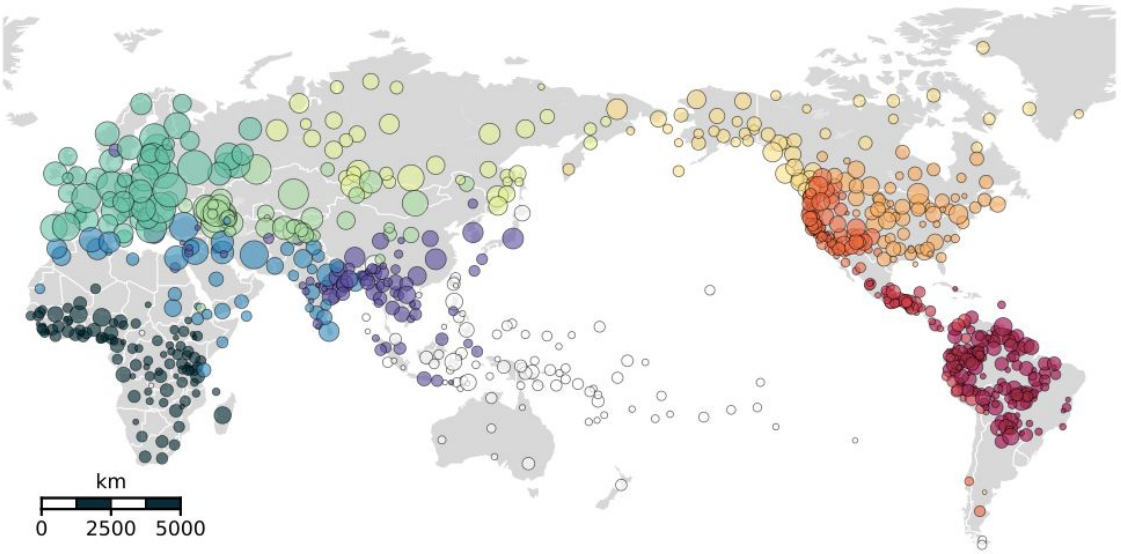
- K=10



- K=11

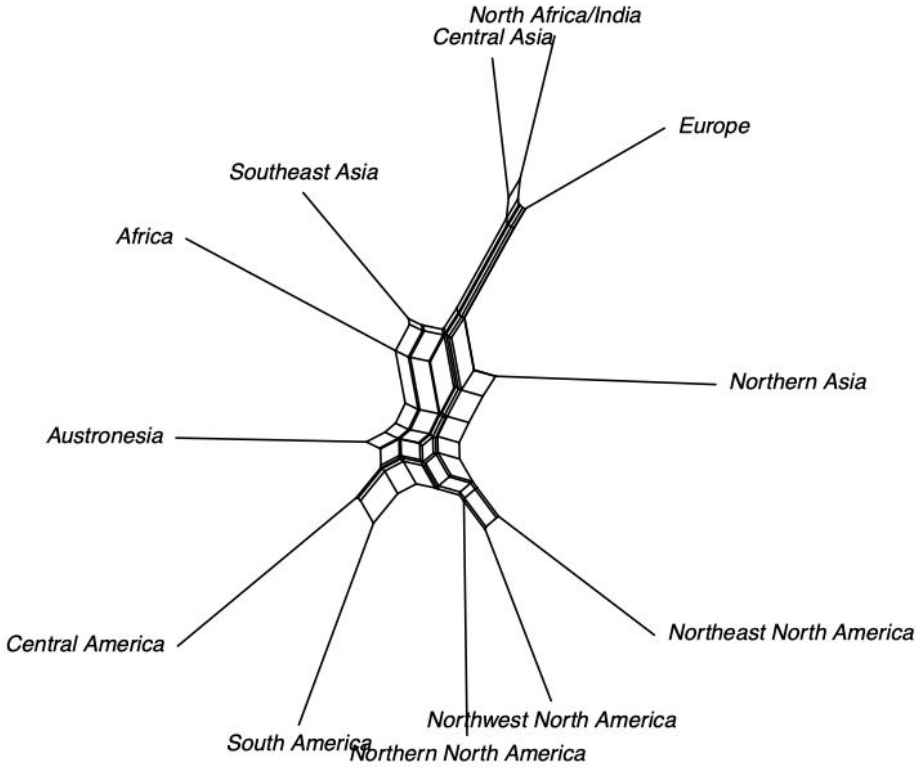


- K=12

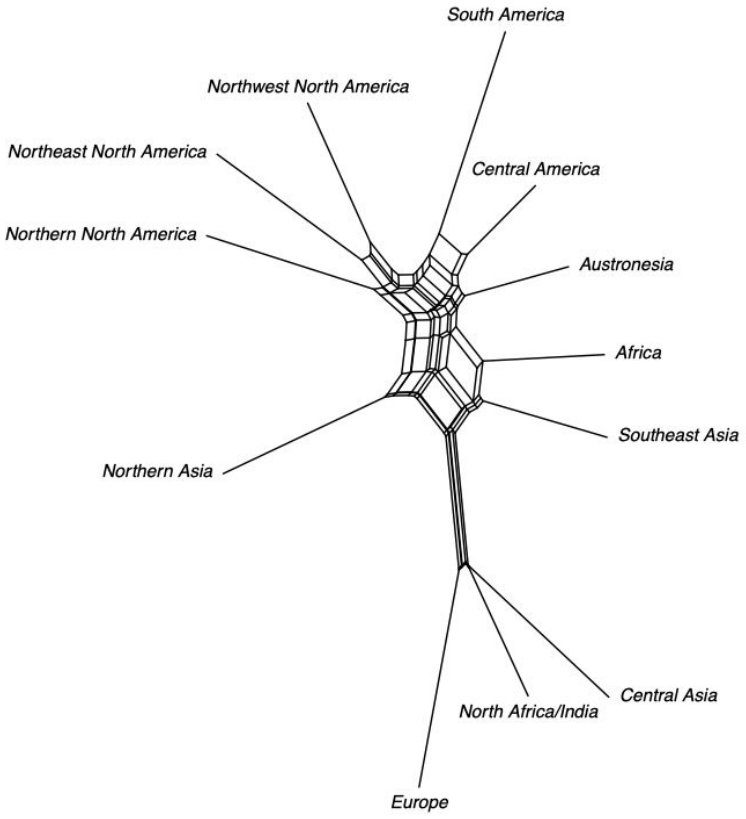


- Clusters
    - A new cluster that covers North Africa and India emerges
    - Two clusters in Western North America become a cluster in the new data
- Things to check
    - Phylogenetic network
    - AMOVA, NMI with bioregions (Needs to collect language family affiliation for new traditions)
    - Similarity decaying
    - Spatial patterns (Motif classification was changed by Berezkin, http://mapsofmyths.com/motif-types)
    - Multiscale Backbone

3. Phylogenetic network
    - Cosine distance

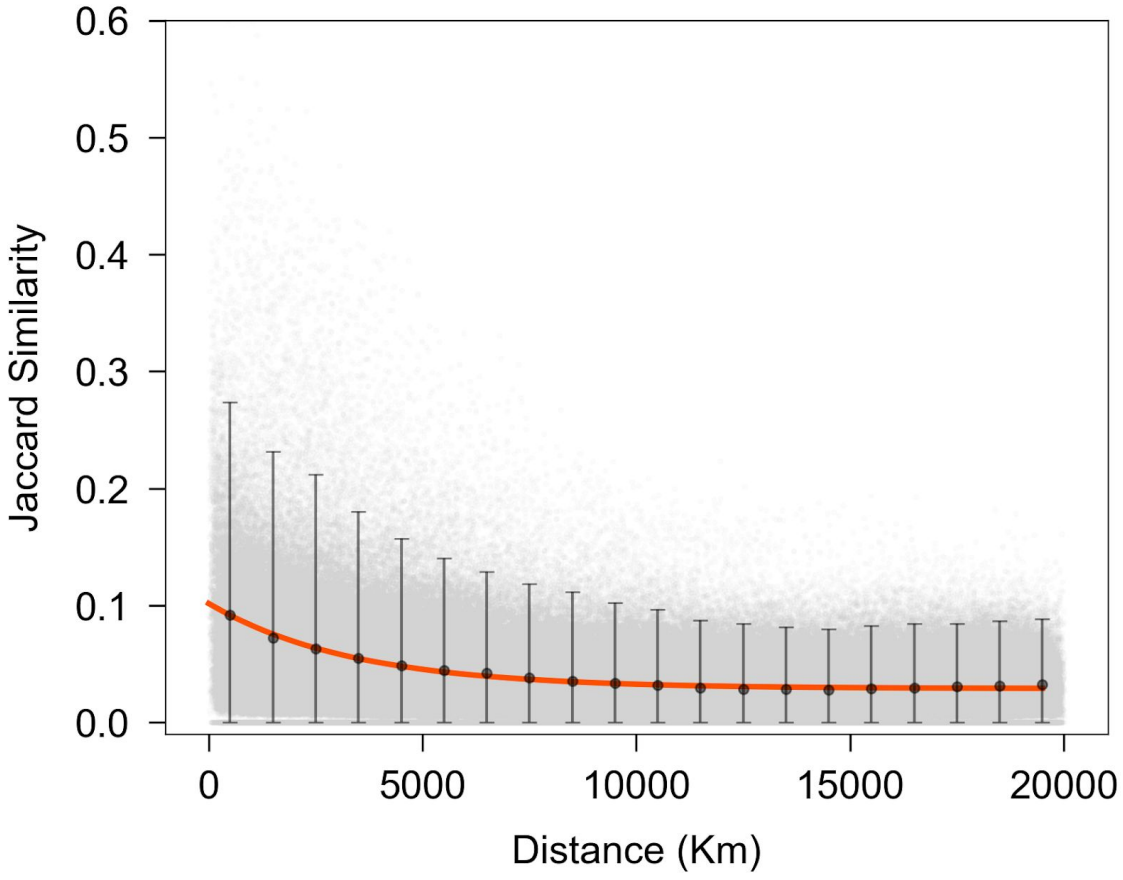


    - Hamming distance



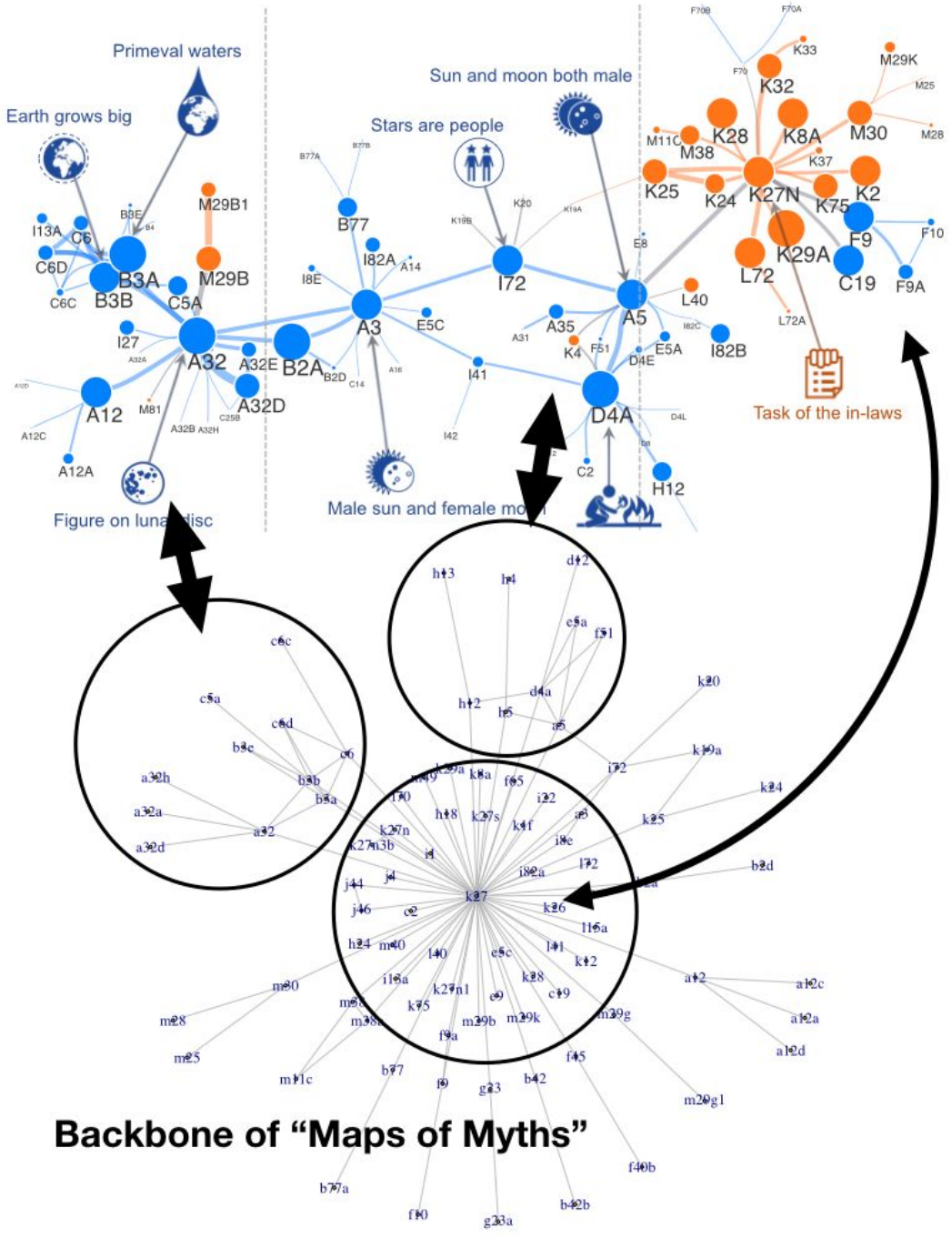
- Network structures of two distances are similar
    - Africa – Austronesia
    - Africa – Southeast Asia – Central Asia, …

4. Similarity decaying



Characteristic distance is 3364 km, which is larger than the previous characteristic distance 2572 km.

5. Motif backbone



- The backbone of “maps of myths” is dominated by the motif “k27”, which is newly added in the data. “k27” is about “Competitions and difficult tasks” that is the parent of “k27n” which is the center of the previous backbone.

- The previous separations are captured even in the backbone of “maps of myths.”

6. Summary

- CorEx clusterings, phylogenetic networks, and similarity decaying patterns are similar in both data, the previous data we used and the new data from “maps of myths”
- However, due to the increase of data coverage and the changes in motif classification, additional works would be needed to obtain figures for our manuscript, if we use the new data.