

The East Asian linguistic phylum: A reconstruction based on language and genes,  
with implications for the ethnolinguistic prehistory of Japan and Trans-Eurasian

George van Driem and Gyaneshwer Chaubey  
University of Bern and Benares Hindu University

The Trans-Himalayan, Austro-Tai, Austroasiatic and Hmong-Mien language families together make up the East Asian linguistic phylum. The geographical distribution and the centre of diversity of each language family together with the relative chronology of branching of each of the four family trees permits inferences about the dispersal and likely geographical origin of each of these language families. Twenty years ago, it was observed that the global geography of language families often corresponded with the distribution of Y chromosomal markers. The ubiquitous albeit not universal correlation of paternal lineages with the geographical distribution of language families suggested that a subset of the paternal ancestors of many language communities also happened to represent the linguistic ancestors of these populations. The Trans-Himalayan, Austroasiatic, Austro-Tai and Hmong-Mien language families are each found to correlate with one of four subclades of the Y chromosomal haplogroup O occurring in high frequency in the modern populations speaking these languages. Data from both historical linguistics and population genetics enable us to present a reconstruction of the founding dispersals of these language families. We conclude with an excursus on the implications of these findings for our understanding of the ethnolinguistic prehistory of Japan and the Altaic or “Trans-Eurasian” linguistic phylum.