

## LANGUAGE DIVERSITY OF THE JAPANESE ARCHIPELAGO AND ITS RELATIONSHIP WITH HUMAN DNA DIVERSITY

Naruya Saitou and Timothy A. Jinam

---

### Introduction

The Japanese Archipelago stretches over 4000 km from north to south, and is the homeland of three human populations; Ainus, Mainlanders, and Okinawans. The origins of these people have been studied for a long time. The standard theory based on craniofacial data is the "dual structure model" propagated by Japanese physical anthropologists, notably Hanihara (1991). According to this model, first migrants to the Japanese Archipelago came from somewhere in Southeast Asia in Upper Paleolithic age more than 30,000 years ago, and were probably ancestors of the Jomon people. The second wave of migration is believed to have taken place later in the Yayoi period (3000-1700 BP), where the people arrived from Northeast Asia. Indigenous Jomon people and new migrants in and after Yayoi period gradually mixed with each other. This model provides a reasonable explanation for the morphological similarity between the Ainu people of Hokkaido at the northernmost main island of Japanese Archipelago, and the Okinawan people in Southwest Archipelago, despite of large geographical distance. Similarity of these peoples was already noticed more than one hundred years ago by von Baelz (1911) who proposed the Ainu-Okinawan common origin theory.

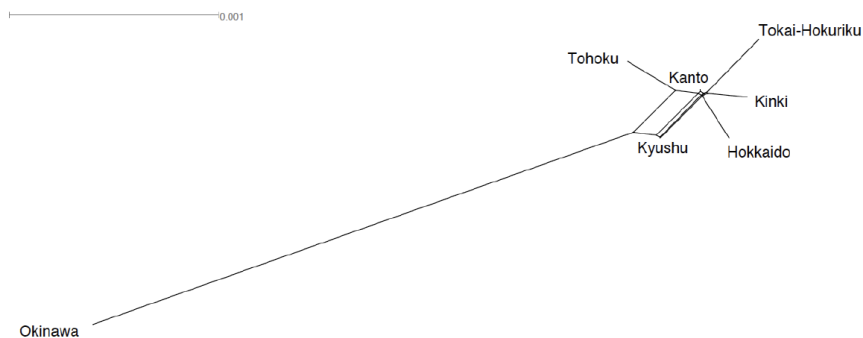
We recently determined genome-wide single-nucleotide polymorphisms (SNPs) for Ainus and Ryukyuan (see Figure 1), and analysed these data with existing data sets (Japanese Archipelago Human Population Genetics Consortium 2012; Jinam *et al.* 2015). Major findings were: (1) Recent admixture with Mainlanders was observed for more than one third of Ainu individuals; (2) Ainus and Okinawans are tightly clustered with a high statistical support followed by Mainlander and Korean in the phylogenetic tree of East Asian populations; (3) Ainu population probably experienced admixture with other populations distributed north of Hokkaido. We further analysed these data, and found that Ainu people are genetically different from Mainlanders living in Tohoku, northern part of the Honshu Island, and the estimated age of starting admixture between Ainu and Mainlander ancestral populations was ca. 6th century A.D. (Jinam *et al.* 2015). We also determined partial nuclear genome DNA sequences of Jomon period people from DNA extracts used for mitochondrial DNA haplotype determination (Kanzawa-Kiriyama *et al.* 2013), and found that the Ainu people are genetically closest to the Jomon people, and



Figure 1

the mainlanders are sandwiched between Jomon and East Eurasian continental populations (Kanzawa-Kiriyama *et al.*, 2016).

Another aspect is the genome diversity of Japanese Archipelago Mainlander populations. Yamaguchi-Kabata *et al.* (2008) reported genome-wide SNP data analysis of 7,000 Japanese, and they computed  $F_{st}$  distances among people from seven geographical areas (see Figure 1). Figure 2 is a phylogenetic network based on these genetic distances using Neighbour-net (Bryant and Morton 2004). The Okinawan population is far apart from the remaining six populations, as expected. Kyushu area, which is geographically closest to Okinawa, is also genetically closest to Okinawa. Interestingly, Tohoku area, which is



**Figure 2**

geographically far apart from Okinawa, shares a short split with the Okinawa population. If modern Tohoku people inherited DNAs of the Emishi who lived in Tohoku area more than 1,000 years ago, this pattern suggests a genetic affinity between the Emishi people and the ancestral Okinawans.

Nakaoka *et al.* (2013) analysed HLA data of ten regions in Japan, and found various interesting diversity among these regions. First of all, Okinawa was quite different from the remaining nine regions. Shikoku and Hokuriku were outliers and were different with each other. Shikoku was somewhat closer to Okinawa, while Hokuriku was most distant from Okinawa.

We recently examined the genome-wide SNP data of the Izumo population who are distributed relatively close to East Korea (see Figure 1). It should be noted that Izumo area is in the Chugoku-Shikoku District which was not examined by Yamaguchi-Kabata *et al.* (2008). Figure 3 is a Principal Component Analysis based on preliminary 21 Izumo individual data with four other East Asian populations: Tokyo (Nishida *et al.* 2008), Korea (Bae *et al.* 2011), North China (HapMap Consortium, 2005), and South China (Teo *et al.* 2009). The PCA coordinates of three continental populations (Koreans, Northern Chinese, and Southern Chinese) and Japanese Mainlanders in Tokyo shown in Figure 3 are in good agreement with their geographical relationship. In contrast, Izumo people are genetically more apart from Koreans than Japanese Mainlanders in Tokyo, although Izumo is geographically between Tokyo and Korea. The PCA plot of Figure 3 suggests the existence of some heterogeneous migrants to the Japanese Archipelago Mainland.