REVIEW

Japan considered from the hypothesis of farmer/language spread

Elisabeth de Boer¹, Melinda A. Yang², Aileen Kawagoe³ and Gina L. Barnes⁴*

¹Faculty of East Asian Studies, Ruhr-Universität Bochum, Bochum, Germany, ²Department of Biology, University of Richmond, Richmond, Virginia, USA, ³Department of Social Studies, New International School of Japan, Tokyo, Japan and ⁴SOAS, University of London, London, UK

*Corresponding author. Email: gb11@soas.ac.uk

Abstract

Formally, the Farming/Language Dispersal hypothesis as applied to Japan relates to the introduction of agriculture and spread of the Japanese language (between ca. 500 BC–AD 800). We review current data from genetics, archaeology, and linguistics in relation to this hypothesis. However, evidence bases for these disciplines are drawn from different periods. Genetic data have primarily been sampled from present-day Japanese and prehistoric Jōmon peoples (14,000–300 BC), preceding the introduction of rice agriculture. The best archaeological evidence for agriculture comes from western Japan during the Yayoi period (ca. 900 BC–AD 250), but little is known about northeastern Japan, which is a focal point here. And despite considerable hypothesizing about prehistoric language, the spread of historic languages/dialects through the islands is more accessible but difficult to relate to prehistory. Though the lack of Yayoi skeletal material available for DNA analysis greatly inhibits direct study of how the pre-agricultural Jōmon peoples interacted with rice agriculturalists, our review of Jōmon genetics sets the stage for further research into their relationships. Modern linguistic research plays an unexpected role in bringing Izumo (Shimane Prefecture) and the Japan Sea coast into consideration in the populating of northeastern Honshu by agriculturalists beyond the Kantō region.

Keywords: Japanese language; Yayoi agriculture; Jōmon genetics; language/farming hypothesis; Mumun–Yayoi migrations

Media summary: How Jōmon genetics, Mumun migrations, and Japanese dialect differences inform farming/language spread hypothesis in Japan

Introduction

The Jōmon period (see Figure 1) is one of the longest post-Palaeolithic archaeological periods in the world, sustained by a hunting–gathering–fishing–horticultural subsistence pattern (Habu, 2004). The introduction of rice agriculture into the Genkai Bay area of North Kyushu (Figure 1) from around 1000 BC marks the beginning of the Yayoi period. Wet-rice technology was introduced by Mumun (‘plain pottery’) culture migrants from the southern Korean Peninsula relatively early, ca. 1000 BC, in the Peninsula’s Mumun period (1500–300 BC). These migrants were not Yayoi, they were Mumun.

The term Yayoi has four uses, which can create much confusion. First, it is the designation of the period beginning with the introduction of rice agriculture around 1000 BC until the advent of the Mounded Tomb Culture in the third century AD. Yayoi is a period designation exclusive to Japan; it includes both farmers and hunter–gatherers and entails the agricultural transition in a time-transgressive and regionally disparate process. Second, ‘Yayoi people’ may refer to anyone living in the Japanese Islands in the Yayoi period, or third, Yayoi may refer specifically to admixed people

© The Author(s), 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.
(Mumun + Jōmon in varying in proportions and across great distances). Fourth, Yayoi may indicate acculturation: the adoption of (rice) agriculture (and other continental material culture) by Jōmon-lineage people in the Yayoi period. All of these conflicting aspects of Yayoi must be kept in mind and clearly defined in any discussion.
Upon the Mumun migrants’ arrival, it is assumed from artefactual, skeletal morphological, and genetic evidence that many took partners from the resident Jōmon population, leading to mixed descendants – and to cultural transformation. These generational ‘admixed’ farmers then spread out from North Kyushu into the rest of the Japanese Islands, meeting more local Jōmon-lineage peoples with further admixing taking place. It is common in the literature to label the admixed farmers as ‘Yayoi’ because they carried the new culture based on rice, while increasingly recognizing that Jōmon-lineage peoples (distinguished as Epi-Jōmon in the north, but included within ‘Yayoi’ in the south) survived well into historical times without substantial admixture. The mixing of these two populations is recognized as the Dual Structure Hypothesis of the modern mainland Japanese population (Hanihara, 1991, 2000).

The first section below reports on Jōmon genetics as a basis for further research on Yayoi admixtures. Despite the 13,000-year Jōmon history entailing multiple cultural formations and temporal trends, archaeologists tend to treat the Jōmon as a homogeneous people unadmixed with mainland East Asians and Austronesians since their separation from each other in the Palaeolithic. Genetic research is disabusing us of that approach, revealing potential Jōmon contacts with coastal mainland East Asians before the arrival of Mumun migrants; these data mirror archaeological evidence of contact. This is enlightening for its own sake, particularly in tracing connections between the Jōmon and continental peoples. However, in terms of investigating Yayoi admixtures, it makes the issue much more complicated. As more Yayoi-period genomic data becomes available, we may then be able to track the regional spread of continental genes through the Japanese archipelago. The addition of new Yayoi-period genomic data is vital, as new waves of immigration from the Korean Peninsula in the succeeding Kofun and Nara periods have enhanced the continental contribution in the modern Japanese, making it difficult to use present-day genomes to effectively study population movement and admixture in the Yayoi period.

The second section focusses on the spread of agriculture from west to east through the Japanese Islands. The major topic of study in this spread has been the adoption of irrigated wet-rice agriculture. It has been assumed that this technology was spread by admixed farmer-migrants. In contrast, new studies are documenting instances of non-irrigated (swamp, dryland) rice and millet cultivation that may be earlier (e.g. Nasu & Momohara, 2016; Shitara & Takashe [sic], 2014). These authors assume that local Final Jōmon and Epi-Jōmon peoples were the cultivators, although some may have become farmers of irrigated wet rice later. The relationships between cultivation technologies and ethnic populations are often unclear, making the next step in the analysis, on language spread, difficult.

The third section deals with language spread. Prehistoric language as such is inaccessible, but the fact that nowadays all languages on the islands except Ainu are related to Japanese makes it absolutely clear that the population that brought wet-rice agriculture to Japan must have been speakers of proto-Japanese. (Proto-Japanese is the ancestral language from which all modern varieties of Japanese descended.) Proto-Japanese split into several dialects, and then Ryūkyūan split off from the Kyūshū dialect. Evidence for the latter scenario has been growing in recent years. (For an overview see De Boer 2017.) Although the Ryūkyūan varieties form independent languages now, their split from the Kyūshū dialect makes them in origin dialects of Japanese.

Proto-Japanese is postulated to have developed on the continent (Whitman, 2011; Miyamoto, 2016; Robbeets, 2005, 2017), and it is hypothesized to have had spread throughout the archipelago with the admixed Yayoi people practising irrigated rice agriculture. This accords well with the farming/language dispersal hypothesis (Renfrew, 1987; Bellwood & Renfrew, 2002). Problems, however, include exactly when and how this spread took place, whether it was indeed only associated with rice agriculture or also with other grain crops, and what Jōmon or Epi-Jōmon language(s) might have been replaced or pushed out by which dialect of Japanese. With the general lack of Yayoi DNA and written linguistic data, it is difficult to match food production techniques, genomic identity and language type.
Jōmon genetics

The transition from the Jōmon to the Yayoi periods marked a transition in the material culture of Japan, with the appearance and increase of tools, artefacts and archaeobotanical remains associated with agriculture. Studies have long associated this transition with the migration of Mumun people through the Korean peninsula into the closest island in the Japanese archipelago, Kyushu (Hammer & Horai, 1995; Hanihara, 1991). Early assessments using cranial and dental morphology and genetics of humans and their commensals led to the Dual Structure model (Hanihara, 1991), for which a key component is the argument that present-day Japanese are a mixture of populations indigenous to Japan since the Palaeolithic and later migrants from mainland Asia. Many studies, mostly using present-day genomes of indigenous populations in the archipelago (i.e. Ainu, Ryūkyū) as a proxy for Jōmon-related ancestry, found genetic patterns that suggest that the present-day Japanese populations are admixed, supporting that aspect of the Dual Structure model (Jeong, Nakagome, & Di Rienzo, 2016a; Jinam et al., 2015; Nakagome et al., 2015).

Advances in DNA sequencing methods have allowed access to genetic material for studying population demographic history in the Japanese archipelago. Uniparental markers were some of the earliest used to infer population history; these include mitochondrial DNA (mtDNA), which exists outside the nucleus, and the non-recombining portion of the Y-chromosome. The primary units considered when studying these markers are haplogroups, which are variants of a DNA segment that derives from the same common ancestor, and they can provide information on a population’s genetic history.

Since uniparental transfer results in no recombination, it is straightforward to reconstruct phylogenetic trees and assign a haplogroup to the mtDNA or Y-chromosome region (Haber, Mezzavilla, Xue, & Tyler-Smith, 2016). Sets of haplogroups are now well annotated and easily identifiable in newly sequenced individuals using many tools and databases (e.g. ISOGG, 2019; Weissensteiner et al., 2016). If they are identified to be at high frequency in a localized area but in low frequency elsewhere, the origin of the haplogroup, or the common ancestor, can be inferred to have derived from that region. While the lack of recombination allows simple haplogroup assignment, it also means that it is a single data point representing population history, and stochastic events or sex-biased migration patterns can lead to inaccurate inference of the true population history (Nordborg, 1998; Rosenberg & Nordborg, 2002).

In contrast, the nuclear genome undergoes recombination, which means that small segments of each chromosome have effectively independent histories. Study of a single diploid genome (i.e. the 23 pairs of chromosomes in the cell nucleus) can be interpreted as study of small portions of thousands of different individuals, resulting in high statistical power for inferring the history of the population(s) represented by those individuals. Thus, sampling genome-wide is imperative for assessing the accuracy of the demographic relationships inferred from uniparental markers.

The ability to extract and sequence ancient DNA (aDNA) has revolutionized the study of human prehistory, allowing direct sampling of past populations and revealing genetic variation no longer observed today. Post-mortem degradation of DNA makes it very difficult to sequence aDNA, so the greater numbers of mtDNA makes it much easier to extract and sequence. The earliest ancient DNA studies in humans were for mtDNA (e.g. Krings et al., 1997), including from a Jōmon individual as early as 1989 (Horai et al., 1989) and Yayoi individuals as early as 1995 (Oota, Saitou, Matsushita, & Ueda, 1995). In fact, the majority of ancient DNA studies on Jōmon individuals have been on mtDNA (Schmidt & Seguchi, 2014).

Only recently have ancient nuclear genomes been retrieved from individuals from Japan. Ancient genomes have been published from human remains at different sites (Sanganji, Ikawazu, and Funadomari associated with Jōmon material culture; Shimo-Motoyama associated with Yayoi material culture; Figure 1). Two partial genomes sequenced from the Sanganji shellmound in Tōhoku were of 3000-year-old individuals (Kanzawa-Kiriyama et al., 2017). Genome-wide data have also been retrieved for a 2700-year-old individual from Ikawazu (McColl et al., 2018) and two 3900-year-old individuals from the Funadomori site in Hokkaido (Kanzawa-Kiriyama et al., 2019).
nuclear genomes is still sparse, but a rough understanding of the ancestry associated with the Jōmon is taking shape with each study.

Characterizing Jōmon-period ancestry is fundamental to understanding the genetic composition of humans that lived during the Yayoi period, who ostensibly represent various degrees of mixture between Jōmon and continental peoples. This in turn will improve our understanding of population migration on the archipelago and its role in the spread of agriculture across the archipelago. Here, we highlight some key features and emerging questions observed to date on Jōmon-related ancestry, which allow us to better understand the challenges (and rewards!) in interpreting the genetics of the Yayoi in future studies.

Analyses of both uniparental (Y and mtDNA) and nuclear DNA of present-day populations have highlighted that ancestry persists in Japan that is deeply diverged from ancestry found in mainland East Asian populations. The distribution of haplogroup frequencies across multiple populations is often used to determine where haplogroups originated. For example, those that occur in high frequency in present-day populations from Japan but low frequency elsewhere (e.g. N9b, M7a, G1b) were proposed as candidate haplogroups of Jōmon origin (Adachi, Shinoda, Umetsu, & Matsumura, 2009; Adachi et al., 2011; Kanzawa-Kiriyama, Saso, Suwa, & Saitou, 2013; Kivisild et al., 2002).

The most prominent of these is N9b, which is the most frequent mtDNA haplogroup assigned to Jōmon individuals. An estimate of the time to the most recent common ancestor (tMRCA) is ca. 22,000 years ago, which is older than the tMRCA for several other haplogroups found in East Asians (Adachi et al., 2011). The non-recombining region of the Y-chromosome has also been extensively studied in Japan. Y-haplogroup D1b is exceptionally interesting as it is found in high frequency in the Japanese archipelago and the Tibetan Plateau but not elsewhere (Hammer & Horai, 1995; Tajima et al., 2004; Watanabe et al., 2019). The estimated tMRCA for D1b is ca. 19,400 years ago (Hammer et al., 2006), similar to that estimated for N9b mtDNA. These dates, if corresponding to the expansion of Jōmon ancestry on the archipelago, suggest that Jōmon ancestry separated fairly early from mainland East Asian ancestry.

One mtDNA haplogroup found in many Jōmon-associated individuals (e.g. Kanzawa-Kiriyama et al., 2013; Adachi et al., 2011; Kim et al., 2010) as well as Yayoi-associated individuals (Igawa et al., 2009) is D4, which is widespread in northern East Asia both in past and present-day populations (e.g. Tanaka et al., 2004; Zhao et al., 2010; Hong et al., 2014; Jeong et al., 2018; Lan et al., 2019). Its widespread presence, however, suggests that it was found in the common ancestral population of all East Asians, making it difficult to evaluate its role in more recent migrations and interactions.

In genomic studies of ancient Asians, both a 40,000-year-old individual found in Tianyuan Cave, Beijing, China (Fu et al., 2013; Yang et al., 2017) and 8000-year-old individuals from Southeast Asia associated with Hōabinian foragers (McColl et al., 2018) show an early split time from other East Asians, including from ancient Jōmon genomes (Kanzawa-Kiriyama et al., 2017, 2019). Thus, we denote them as ‘basal Asians’. In contrast, humans living in Southeast Asia, the eastern steppes of Siberia and the southwest edges of the Tibetan Plateau from the Neolithic to historic periods (e.g. Sikora et al., 2019; Jeong et al., 2016b) share a closer relationship to present-day East Asians than the Jōmon individuals, confirming that Jōmon are distantly related to mainland East Asian ancestry (Yang, MA and Fu, Q unpublished observations).

In fact, the Jōmon sampled thus far probably separated earlier than the split time of East Asians and the early Asian population that contributed to Native Americans (Gakuhari et al., 2019; Kanzawa-Kiriyama et al., 2017, 2019). Estimates of when ancestral Native Americans separated from the East Asian lineage is ca. 36,000–25,000 years ago (Moreno-Mayar et al., 2018), and a direct estimate of the split time between the Jōmon and mainland East Asians ranges from 38,000–18,000 years ago (Kanzawa-Kiriyama et al., 2019). These estimates are broad, indicating there is still high uncertainty as to the exact date. However, studies of both uniparental and nuclear markers all highlight Jōmon ancestry separated from mainland East Asian ancestry in the Palaeolithic, with dates ranging from 38,000–18,000 years ago.
Box: Three-way Analysis of Genetic Relationships

Figure 2. F4-tests surveying the relationship of populations with East Asian ancestry (on the x-axis, ‘X’) to the Ikawazu Jōmon, relative to ancient Southeast Asian foragers (Hōabinhian, HB), an ancient Tibetan (Chokhopani, CH), and an inland ancient northern East Asian (Shamanka_EN, SH) (graphs by Melinda Yang). Numerical values for the data visualized here can be found in Table S3. F4-statistics are tests of symmetry of the form f4(Out, P1; P2, P3), where it is assumed that P2 and P3 form a clade with respect to P1. Our statistics take the form (A) f4(Mbuti, HB; Ikawazu Jōmon, ‘X’ East Asians) or (B, C) f4(Mbuti, Ikawazu Jōmon; CH or SH, ‘X’ East Asians), with the underlying tree structure verified by other genetic analyses. Central African Mbuti were used as the outgroup (Out).

Key: tree models showing relationships of f4 values are placed to the right of the plot: f4 > 0 corresponds to the highest tree; f4 = 0 corresponds to the middle tree (clarified by dashed grey arrow); and f4 < 0 corresponds to the lowest tree for each figure. The presence of gene flow in the model is indicated by the red arrow. A significantly positive value highlights gene flow between the outgroup (P1) and ‘X’ East Asians (P3). A significantly negative value highlights gene flow between the outgroup (P1) and the other population in the tree (P2). Two standard errors are indicated by the thin grey bar, while one standard error is indicated by the thick grey bar. Points are coloured (colour online) by region of origin, with darker colours for ancient populations and lighter colours for present-day populations. East Asians for whom all or part of their population live near coastal East Asia are shaded grey. *Ancient islanders were sampled from Southeast Asian and Southwest Pacific islands, but were previously shown to share a close relationship to present-day Austronesians (McColl et al., 2018; Skoglund et al., 2016).

Interpretation: in Figure 2A, 8000-year-old Hōabinhian (HB) foragers from Laos and Malaysia are an outgroup to the ikawazu Jōmon individual and present-day and ancient populations of East Asian ancestry. A connection between the ikawazu Jōmon individual and the Hōabinhians would result in a
A key aspect of the observed genome-wide patterns above is that these Jōmon individuals do not share any relationship with sampled ‘basal Asians’. In admixture tests (Patterson et al., 2012), two populations that form a clade (e.g. the Jōmon and mainland present-day East Asians) may sometimes share excess ancestry with an outgroup population that potentially indicates partial shared demographic history, probably via a population closely related to the outgroup. However, Jōmon individuals show no excess connection to ‘basal Asians’, including the Southeast Asian Hòabìnhians relative to present-day East Asians, as indicated by the lack of significantly negative values in the Box (Figure 2A).

In particular, comparison of the Southeast Asian Hòabìnhians from Laos and Malaysia with mainland East Asians from East Asia and a Jōmon individual (Ikawazu) shows that the East Asians and Jōmon ancestry are equidistant from Hòabìnhian ancestry (Figure 2A), indicating that the Jōmon do not share a special relationship with Hòabìnhians as previously suggested (McColl et al., 2018). Tests of genetic similarity do not show Hòabìnhians or the Jōmon sharing exceptionally high genetic similarity with each other (Figure S1). Shared ancestry between present-day Japanese and recent Southeast Asians are best explained by a common process – gene flow from mainland East Asia, a phenomenon well-characterized by tests of the Dual Structure model in Japan and by observations of gene flow into Southeast Asia (Kanzawa-Kiriyama et al., 2017, 2019; Lipson et al., 2018). In the future, individuals from eastern Eurasia will almost certainly help to clarify how Jōmon ancestry first entered the Japanese archipelago, but the current sampling has not yet shed light on who they might have been.

Connections do exist between the Jōmon and East Asians who live along the coast of East Asia. The first Jōmon mtDNA sequenced in 1989, from the Early Jōmon Urawa-I site in the Kantō region (Figure 1), shared a close relationship to present-day Southeast Asians that suggests common ancestry with Southeast Asians (Horai et al., 1989). This haplogroup, called E1a1a, is prevalent in Austronesian-speaking populations of Taiwan and Southeast Asia who share a closer relationship to East Asians than other Southeast Asians. A ca. 8000-year-old individual sampled from Liang Island off the southeast coast of China in Fujian province also shares this haplogroup (Ko et al., 2014). Study of the Liang Island individual’s nuclear genome indicates a close genetic relationship to Austronesians (Yang, MA and Fu, Q unpublished observations). Present-day Austronesians, such as the Ami and Atayal from Taiwan (Mallick et al., 2016), share connections with the Ikawazu Jōmon relative to more inland ancient East Asians (Figure 2B and C), potentially indicating admixture between the Jōmon and southern East Asian populations. Comparisons with nuclear genomes from Fujian in southern China show similar connections (Yang, MA and Fu, Q unpublished observations). Thus, rather than Southeast Asian, the primary signal seems to be associated with Austronesians, who derive from an East Asian population associated with southern China.

Siberian connections, particularly in Hokkaido, have also been a recurring theme of Jōmon-related genetic studies. Several of the mtDNA haplogroups prevalent in the Japanese and Jōmon but rare in mainland East Asians, especially G1b, can be found in southeastern Siberians, highlighting potential contacts along the coast in northern East Asia as well (Sato et al., 2009; Adachi et al., 2011).
Pre-modern Ainu show similar mtDNA haplogroups to populations in the Lower Amur region of Siberia and populations belonging to the early historic Okhotsk culture of Hokkaido (600–1200 AD) (Adachi, Kakuda, Takahashi, Kanzawa-Kiriyama, & Shinoda, 2018). Studies of nuclear DNA from Early Neolithic populations in the Primorye region of Russia highlight that the Jōmon are an outgroup to all mainland East Asians, including these more northern populations in Siberia (Sikora et al., 2019; de Barros Damgaard et al., 2018). In the same comparisons set up for coastal southern East Asians above, it can be shown that ancient and present-day coastal northern mainland East Asians dating up to 7700 years ago share connections with the Ikawazu Jōmon relative to more inland East Asians (Figure 2B, C). Ancient northern Siberian ancestry prevalent during the Palaeolithic notable for both its closer relationship with European-related rather than Asian-related ancestry and its impact on Native American ancestry (Raghavan et al., 2014; Sikora et al., 2019) is not found in mainland East Asians or the Jōmon, which emphasizes that the connections are specific to coastal mainland East Asians and the Jōmon.

One explanation for a connection between the Jōmon and coastal East Asians could be that the Jōmon were not completely isolated from mainland East Asians. By 3900 years ago, the date of the oldest Jōmon nuclear genome sampled (from Funadomari, Figure 1), Austronesians were rapidly expanding into islands in the Pacific (Tsang, 1992; Spriggs, 2011). The main patterns observed both in past mtDNA studies and in recent genome-wide studies of the Jōmon all seem to highlight coastal connections, which may suggest that the Jōmon experienced gene flow with populations deriving from mainland East Asia prior to any contact associated with migration of Mumun immigrants from the Korean peninsula. This is supported by archaeological studies that track artefact commonalities resulting from trade and contact along the coast (Bausch, 2017), some of which date back to the Palaeolithic (Morisaki, 2015). This is logical given that the Jōmon were inveterate deep-sea fishers given their harpoon technology, and scores of dugout canoes have been excavated from archaeological sites (Habu, 2010) — although it is not clear how seaworthy they were.

To address the relationship between population movement and the spread of farming in the Japanese archipelago, Jōmon and Mumun ancestry must be well characterized. Here in this section, we have briefly highlighted some key features of Jōmon ancestry using ancient mtDNA and nuclear genomes. First, Jōmon ancestry diverged fairly early from that of mainland East Asians. Second, they do not show notable connections to currently sampled basal Asians, such as the Tianyuan individual or Hōabinhians. Third, they share coastal connections with coastal populations in northern and southern East Asia. This suggests that one of the major assumptions about the Jōmon may not be true — that they were genetically isolated since the first Palaeolithic migrations to the Japanese archipelago until the Mumun migration at the end of the Jōmon period. If the Jōmon themselves are already partially admixed, then characterizing increased gene flow from the mainland in populations from the Yayoi or more recent historic periods will be substantially more complex.

Advances that utilize rare alleles or long haplotypes (Lawson & Falush, 2012; Schiffels et al., 2016) are increasing the power of demographic analyses focused on more closely related populations, including in Japan (Takeuchi et al., 2017), but these are still difficult to apply in the realm of ancient DNA where obtaining data with high enough coverage is still rare. Present-day genomes are informative on recent history (Takeuchi et al., 2017) but make it difficult to resolve questions about periods as early as the Yayoi. Ancient genomes directly from the Yayoi are needed to clarify early population movement associated with the spread of farming. Such sampling is just starting, where two nuclear genomes from the Shimo-Motoyama rock shelter site in northwest Kyushu dating to the Terminal Yayoi show evidence of admixture between Jōmon- and mainland East Asian-related ancestry, in a region where populations were thought to be of unadmixed Jōmon-lineage (Shinoda et al., 2019). Characterizing the role admixture played during the Yayoi requires understanding of the features of Jōmon ancestry described above to help us understand how to contextualize these and future data.

The remaining parts of this paper focus on populations in the Yayoi to Nara periods, examining (1) how archaeological shifts indicate a complex role for migration and admixture that was probably region-specific and (2) how similarities in dialect types hint at prehistoric migration of admixed Yayoi populations along the Japan Sea coast, which only stresses the need for more data on ancient Yayoi DNA.
The coastal genetic connections proposed are not directly tied to the farmer/language spread hypothesis because they are found in the Jōmon and thus would be prior to the spread of farming. Instead, they are a cautionary point – to understand farmer/language spread, we need to understand who were the pre-existing populations (the Jōmon). If their ancestry is not as genetically isolated as has been argued, then future studies need to account for that.

**Agricultural transition in Japan**

In dealing with this problem, first we must define ‘agriculture’. This is not as easy as it appears, since the Jōmon were experienced at plant management, leading to several local domesticated species. Thus, they can be called hunter–gatherer–fisher–horticulturalists. However, the plants that they domesticated or at least managed – soybeans, adzuki, Perilla, sweet chestnuts and Japanese millet (*Echinochloa utilis*) among others – were not suitable for intensive cultivation as carbohydrate sources to become staples (Crawford, 2011; Barnes, 2015: 111–112, and references therein). Crawford would rather discard the distinction between horticulture and agriculture, preferring to deal with proportionalities in ‘resource production’, which includes not only food but other strategic resources such as lacquer and hemp for fibres. Here, we use ‘agriculture’ to refer to the farming of major starch-grain crops: rice (*Oryza sativa*) and millets (*Setaria italica* and *Panicum miliaceum*) in Early Yayoi, with barley and wheat added in later Yayoi.

The Jōmon subsistence system was interrupted by the importation of rice and millets from the Korean Peninsula in the Mumun Period, heralding the start of the Yayoi period in the Japanese Islands. The beginnings of rice agriculture in each region of Japan have been meticulously tracked, first from its source on the continent into North Kyushu (Miyamoto, 2017, 2018, 2019) and then eastwards through the archipelago (Fujio, 2004, 2009, 2013, 2014, 2017). Proto-Japanese speaking Mumun migrants (Whitman, 2011) coming into North Kyushu interacted with Jōmon peoples there, creating an initial admixed Yayoi population which then spread throughout the Japanese islands as they occupied new lands for farming.

For each region, Fujio has identified a ‘hazy’ period of possible cultivation capped by the clear introduction of irrigated rice agriculture (Figure 3); he distinguishes this hazy period, however, from the clear adoption of irrigated rice agriculture and the appearance of the continental cultivation toolkit, moated settlements and moated burials. These elements have been treated as the ‘Yayoi package’ (Mizoguchi, 2013) or ‘Yayoi complex’ (Miyamoto, 2016, 2019). Within most of western Honshu Island, Fujio expects that grain cultivation (rice and particularly millets) preceded the formal package. If we assume that the package was instituted by migrating farmers, the hazy periods and non-irrigated rice-farming periods could represent the gradual conversion of Jōmon-lineage peoples to an agricultural way of life. By 380 BC, however, the ‘Yayoi package’ entailing bronze and iron usage was fully established west of the ‘Waist of Honshu’ (between Wakasa Bay and Ise Bay, Figure 1).

Beyond the Waist of Honshu, the uptake of farming was far more diverse. Current research identifies phases of Jōmon pottery usage during the advent of agriculture in the Kinai, Chūbu central mountains, and Kantō regions where cultivation of dry-field crops also preceded irrigated rice agriculture (Shitara & Takashe, 2014; Shitara & Fujio, 2014, pp. 12–13; Takase, 2014; Barnes, 2019). Thus, some farmers in the Yayoi period might have been *neither* genetically admixed with Mumun (being Jōmon-lineage) nor Japanese speakers. Rice agriculture was established in Hokuriku and Tōhoku from 380 BC. However, settlements at the northern tip of Honshu failed by 100 BC, with occupants reverting to a hunting and gathering lifestyle. Thus, North Tōhoku is included with Hokkaido in the Epi-Jōmon period, which lasted to 700 AD, but we will refer to Jōmon-lineage people in other parts of Tōhoku also as Epi-Jōmon.

According to Figure 3, irrigated rice agriculture was instituted by 380 BC in Hokuriku and 300 BC in South Kantō, but not until 220 BC in Chūbu and 50 BC in eastern Tōkai. Interestingly, it reached North and South Kantō via different routes. The specific ceramic style accompanying the introduction of rice into Gunma Prefecture in northwest Kantō had its origin near the Japan Sea coast in Niigata Prefecture (Baba, 2008). It most likely arrived there through diffusion along the Japan Sea coast from western Honshu, a completely different route from that arriving in Kanagawa Prefecture along the Pacific seaboard.
North of the Toné River, the essential elements of the Yayoi 'package' are mainly missing until ca. 150 AD (Fujio, 2013). Thus, it is assumed that cultivation techniques for rice and millet in South and Central Tōhoku were previously borrowed by Epi-Jōmon peoples living in this area and the Japanese language was spread by later movements of people in the Kofun period.

In the second and third centuries AD, state formation processes began with the widespread adoption of mounded tomb building for those Yayoi elites who had risen to high status within the developing agricultural polities. The succeeding Kofun ('old tomb') period (Barnes, 2007) witnessed the establishment of a centralized Yamato state in the fifth-century Kinai region. This was also a century of a second wave of immigration from the Korean Peninsula, stimulated by warfare among Peninsular states. In the fifth and sixth centuries, migrations of elites from western Honshu augmented the Kantō population and introduced equestrian culture.

The Nara period began when a national capital, Heijō (Figure 1), was built in today’s Nara Prefecture. At this point, however, the Yamato state was limited to the Kinai region, but alliances were made with outlying chieftains in Kyushu and Kantō. By the mid-seventh century these figures were integrated into a court hierarchy that represented an expanded Yamato state. However, this state did not extend further north than Kantō; peoples who occupied the Tōhoku region were beginning to be called emishi in court documents. The Nara state actively pursued military activities to bring this region under court control, although they ultimately failed and the region remained under the rule of prominent emishi.
‘Emishi’ in the Nara period is understood to have been a word that designated people who lived in the north beyond the reach of the state. It was neither an ethnonym nor referred to a way of life but was an administrative term – and so encompassed everyone living beyond state jurisdiction regardless of subsistence, language, or genetic background. We assume that those *emishi* who practised rice agriculture spoke Japanese, but Ainu place names surviving from Niigata and Fukushima northwards clearly indicate that the Ainu language was being spoken by some *emishi* (see Hudson, 1994, fig. 1). Historical documents record the need for interpreters at least once during the state expansion into the north (the Emishi Wars 774–811), which ended abruptly with the court abandoning its military campaigns (Friday, 1997); the Epi-Jōmon included among the *emishi* were very likely speakers of an Ainu-related language, and most linguists view Ainu as a relic Jōmon language. It is clear that evidence of the Ainu language predates the emergence of the Ainu people in historical documents in the seventeenth century (Hudson, 1999).

Knowledge of the subsistence system of Tōhoku in the Kofun and Nara periods is not well developed. The Kantō became a prominent horse-breeding area, supplying the Court with horses and cavalry, leading to the development of the medieval samurai (Farris, 1992; Friday, 1997). Further north, Epi-Jōmon, admixed Yayoi descendents and Kofun-period pioneers formed the *emishi* who opposed the Nara Court’s efforts at subjugation. By the ninth century, powerful local warlords from south-western clans such as the Abe and Kiyohara ruled large territories in the north on behalf of the Heian Court, and finally in the twelfth century, the Northern Fujiwara established their rule over Tōhoku. It is significant that, despite the court origins of this family, the coffins of the first three Northern Fujiwara rulers contained only a little rice and no continental millets, but large quantities of barnyard millet (*E. utilis*) (Hudson, 1999, p. 76). This suggests that continental-type agriculture was not a prominent feature of Tōhoku farming even into the Medieval period, with barnyard millet harking back to Jōmon horticultural practices.

The languages of Japan

As in many other areas of the world, migrations in connection with the spread of agriculture in Japan are thought to have led to large-scale replacement of the languages of indigenous hunter–gatherers (Robbeets & Savelyev, 2017). The only remnant today of the many languages of hunter–gatherers that must have been present in Japan in the Jōmon period is formed by Ainu, an (almost) extinct language on the northeastern island of Hokkaidō. The presence of Ainu place names in Honshū, from Niigata and Fukushima Prefectures northwards, indicates that this region too was once an area where an Ainu-related language was spoken. All other languages in the archipelago now belong to the Japanese language family.

It is likely that at least initially diversification in the language was low, as is often the case when a group of migrants spreads out over a new territory. Dialect borders that later developed in Japan often follow natural barriers such as mountain ranges that impede communication between neighbouring areas. The best-known dialect border for instance, which divides the dialects up into an eastern Japanese and a western Japanese dialect group based on differences in vocabulary and grammar, follows the Hida and Kiso mountain ranges in central Honshū. Most of these differences, for instance those in grammar, do not go back further than the fourteenth century, and many are linked to linguistic influence radiating out from the old capitals of Nara and Kyōto. Some other dialect borders, on the other hand, may have links to the prehistoric migrations that spread agriculture through the islands.

The unexpected dialect type of the northeast

An example of such an unexpected dialect distribution is the fact that the phonology and tone system of the Tōhoku dialect resemble those of Izumo, on the Japan Sea coast in the west of Honshū, more than those of the adjacent Kantō region.

The tone system of West Kantō, for instance, belongs to the so-called Chūrin or ‘middle circle’ type, in which the division of the words of the language in tone classes is different from that of the Tōhoku
region, which has a so-called Gairin or 'outer circle' type tone system. These names are derived from the geographical positions of the types relative to each other (Figure 4). An earlier work (De Boer, 2010) analysed the Gairin-type merger pattern of the proto-Japanese tone classes as the result of an innovation that other dialects lacked. As such, a Gairin-type system can develop multiple times independently, as the tonal assimilations that gave rise to the Gairin system are natural (there are, for instance, also two regions with Gairin A tone systems in Kyūshū and Shizuoka), but the agreements between the dialects of Izumo and Tōhoku go deeper.

Within the innovative Gairin type, a further innovation occurred, resulting in a sub-type where high tone shifts away from syllables that contain /i/ or /u/. It can be seen that this type (Gairin B in Figure 4) developed in Izumo in the centre of the region, whereas the older Gairin type (Gairin A) was preserved in the periphery. The Gairin B innovation is also found in the Tōhoku region, except in two areas far removed from the Japan Sea coast. This distribution suggests that the Gairin B innovation was introduced on the Japan Sea coast side, and spread eastward from there.

The vowel systems of both Izumo and Tōhoku have centralized /i/ and /u/ and raised /e/ and /o/ so that these vowels are all close together and no longer maximally opposed to each other as they are in other Japanese dialects. In central Izumo and in the Tōhoku region (except for a small area far removed from the Japan Sea coast), this has resulted in mergers between vowels (Figure 5). This

---

Figure 4. Map of the Japanese tone systems. Adapted by EdB from Wurm and Hattori (1981): no page number.
geographical distribution, too, suggests that an innovation that originated in Izumo was introduced to the northeast by way of the Japan Sea coast.

The East Kantō dialect (Figure S2) takes a somewhat intermediate position between the West Kantō dialect and the Tōhoku dialect. Some centralization of /i/ and /u/ is present, and the fact that the tonal distinctions have disappeared is most likely the result of a mixture of West Kantō (Chűrin) and Tōhoku (Gairin) influences. From other areas in Japan it is known that confusion between adjacent tone systems with different merger patterns of the tone classes can lead to collapse of the system. The northeastern toneless zone includes not only the East Kantō dialect, but also the Pacific side of southern Tōhoku (Figure 4). On the Japan Sea side of southern Tōhoku the reflexes are mixed Chűrin and Gairin B, but the system has not collapsed.

What does the opposition between the West Kantō tone and vowel system and the northern Tōhoku tone and vowel system and the intermediate zone between them mean for the way in which the northeast was settled by Japanese-speaking farmers? There are a number of things to consider. Vocabulary can be easily borrowed long distance from other dialects, as long as there is contact (for instance through trade along the coast). Phonology, or the sound structure of a language, on the other hand, usually spreads to adjacent areas, with which there is sustained contact. If it spreads to distant areas, migration is the most likely cause, although it is possible for similarities in phonology between far-flung regions to be the result of parallel independent development, if the innovations are easily repeatable and cross-linguistically common.

The similarities between Izumo and Tōhoku (and to a lesser extent the Noto Peninsula) come in shared sets (see Tables S1 and S2; and De Boer, forthcoming), which makes parallel independent development unlikely. It suggests movement of people along the coast in at least two different periods: first from Izumo to the area of the Noto peninsula (present-day Ishikawa and Toyama Prefectures), spreading the set of features listed in Table S1; and later from Izumo to the Tōhoku coast, spreading the set of features listed in Table S2.

**Izumo in the Late Yayoi period**

Although Izumo (Shimane Prefecture) is nowadays relatively poor and sparsely populated, this was not always the case. Izumo was one of the most powerful confederacies of the Mid and Late Yayoi periods.
It was the great rival of Yamato and the focal point of a wide maritime trade network that included the Japan Sea coast, the Ryūkyūs, Kyushu, Korea and China (Torrance, 2016). Izumo formed alliances with other adjacent regions along the Japan Sea coast; Watanabe (1995) speaks of an ‘Izumo cultural zone’ which he places in the Late Yayoi period.

The corner-projected mound burials typical for Izumo in this period are also found on the Noto Peninsula and in Toyama Prefecture, where they stem from 100–250 AD (Maeda, 2007, p. 6). This makes it likely that the presence of the Izumo vowel system on the Noto Peninsula and in Toyama dates back to the Mid to Late Yayoi period, suggested not only by the congruence of burial types (Figure S3) but also the presence of the Izumo-style vowel mergers in Toyama (Figure 5). The Izumo-style tone systems (Gairin A and Gairin B) had apparently not yet developed, as the tone system of the Noto Peninsula does not share these innovations (Figure 4). Some other innovations have occurred in the local tone systems since then, but not the same as those shared by Izumo and Tōhoku.

The fact that the Izumo tonal innovations are today present in the northern Tōhoku region, and in a mixed form along the southwestern Tōhoku coast, means that migrations from Izumo to these areas must have taken place after the Gairin B innovations developed. The presence of both Gairin A and Gairin B in the northeast may indicate migrations in different periods. It may also be related to different points of departure from Izumo, as political and economic prominence in Izumo fluctuated historically between the eastern and western parts (Piggott, 1989; Torrance, 2016). The Gairin B tonal innovation is most advanced in western Izumo (Hirako, 2017), meaning that it most likely started there and only gradually spread to eastern Izumo. In the Tōhoku region too, the Gairin B system was most likely present in a smaller area in earlier times. The tone shifts typical of this system continue to spread to adjacent areas: recent fieldwork (Boiko, 2018) has shown that, by now, the Gairin A area on the Shimokita Peninsula has disappeared, and instead, the Gairin B tone shifts have been adopted.

**Looking for the time of the migrations**

Exploring the similarity of the Izumo and Tōhoku dialects, it is necessary to examine opportunities for migration of Izumo dialect speakers into the northeast. Although the very first forays of rice farmers from western Japan to the northern tip of Honshū already took place in the Middle Yayoi period, these settlements were later abandoned. The full ‘Yayoi package’ reached the coastal areas of Ishikawa, Toyama and western Niigata as early as 300 BC. It is possible that rice farmers settling along the coast departed from Izumo, but there is no direct evidence for this. Clear evidence of influence from Izumo is more recent: the Izumo-style corner-projected burial mounds in Ishikawa and Toyama date from 100–250 AD. While the full Yayoi ‘package’ appeared briefly in the central mountains (Figure 3: #1, Chūbu district), it also intruded into western Gunma Prefecture at this time. The ceramics accompanying the irrigated-rice culture package derived from the Jōetsu Basin in western Niigata and spread through the river valleys of Nagano Prefecture into western Gunma (Baba, 2008).

It is unlikely that the rice farmers who moved inland from western Niigata spoke a dialect with the features typical of Izumo. Linguistic influence from Izumo in the form of migration was not only later, it also may have been initially limited to Ishikawa (i.e. the Noto Peninsula) and Toyama, as Izumo-style mound burials have not been found in Niigata. If there was already some Izumo influence on the dialect of Niigata at the time when rice farmers moved inland to Gunma, it would have been on the vowel system only. The tonal innovations in the Izumo dialect had not yet occurred, judging from the fact that these are not present in the dialects of Ishikawa and Toyama. What applies to the dialect of Niigata also applies to Nagano (Chūbu) and Gunma (West Kantō), which was settled from Niigata, and to Tochigi and Ibaraki (East Kantō), which were all settled overland. If the Tōhoku region was entirely settled overland from these areas, the similarities with the Izumo dialect remain unexplained.

Can migration of Izumo dialect speakers into Tōhoku in later periods explain the similarities between the dialects? Piggott (1989) examined the history of Izumo throughout the Late Yayoi and Kofun periods, distinguishing between eastern and western Izumo. Izumo’s unique mound-burial culture of Late Yayoi gave way to cultural and political inroads from Kibi to the south and Yamato itself,
first into western Izumo. The chieftains of eastern Izumo maintained trade relations with northern Koshi (western Niigata) into the sixth century, but by the 540s the Izumo chieftains had all allied with Yamato (Piggott, 1989, pp. 59–60). Torrance (2016, p. 4), on the other hand, argues that Izumo remained an important and independent presence along the Japan Sea coast, at least until the late sixth or early seventh century.

Under the circumstances it is possible that groups from Izumo, avoiding Yamato pressure, would migrate northward along the Japan Sea coast. Diffusion into Tōhoku through mountain basins and over passes could have taken place, but were the numbers enough for linguistic change throughout Tōhoku? The mixed nature (and collapse) of the tone systems in south and central Tōhoku suggest that there was some influence but not enough to determine the dialect. The linguistic influence in northern Tōhoku, however, was much stronger.

Hudson (2017) mentions that until the fifth century the pottery types of the northern Tōhoku region and Hokkaidō were identical. From the fifth to the seventh centuries, however, this pottery type starts to disappear from northern Tōhoku, and is from then on restricted to Hokkaidō only. This development has taken to mean that the Epi-Jōmon population among the emishi who, from the evidence of place-names, must have spoken an Ainu-related language, moved away from northern Honshū into Hokkaidō in that period. According to Matsumoto (2018, p. 158), after the late fifth century there are no traces of habitation in the northern Tōhoku for about a century, after which a new population arrives in the late sixth century. The new arrivals were archaeologically indistinguishable from Kofun cultures elsewhere in Japan (Matsumoto, 2018, p. 159).

If the people arriving in the late sixth century were from Izumo, then northern Tōhoku would have been sparsely populated at that time, as the Epi-Jōmon population had for a large part moved away. That means that the new population could rapidly spread out over the entire area, which would explain why the Izumo-style tone system was preserved well there, in contrast to the situation in south and central Tōhoku. In south and central Tōhoku there was interference not only from other dialects but most likely also from languages of the local Epi-Jōmon population, who were no longer hunter-gatherers but settled farmers by then. The relatively low internal diversity of the Tōhoku dialects may be attributed to the overall later spread of Japanese to the northeast compared with other areas of mainland Japan (Inoue, 1992).

Migration of people from Izumo to the northeast may even have left a genetic trace. Saitō (2017, p. 127) uses a principal component analysis (PCA), in which nuclear DNA from modern Tōhoku individuals is compared with nuclear DNA from modern Ainu, Ryūkyūan and mainland Japanese individuals (Figure S4). A PCA simplifies complex datasets, creating summaries that emphasize the components explaining the maximum amount of variance. In population genetics, this method is often a quick and useful tool to examine population relationships, as the maximum amount of variance is often attributed to shared population history. However, it is not a formal test of shared ancestry.

Saitō concludes that, in contrast to Ryūkyūans, the Tōhoku population has very low traces of Ainu-related ancestry, which was used as a proxy for Jōmon-related ancestry. He also includes a PCA showing the position of Izumo individuals relative to Okinawan and Kantō individuals (Saitō, 2017, p. 155), and he remarks upon the fact that these individuals occupy the same relative position as the Tōhoku individuals (Figure S4), namely they are to the lower right of the average Mainland Japanese individual.

It is better to be cautious in drawing conclusions until more detailed comparisons from different areas of Japan are available, but the initial results are definitely interesting in light of the dialectal similarities between Izumo and the Tōhoku region.

**Concluding remarks**

Confronted with the Yayoi farmers and their new mode of subsistence, the different ethnic groups that were already present in Japan could react in different ways: assimilation and adoption of the new
techniques, specialization in products with which to trade with the new farming populations, resistance or withdrawal. The available options differed per region and per period, influencing the way in which the Japanese language spread.

The present-day dialect distribution in northern Honshū suggests that a considerable part of the early historic emishi population were speakers of an Izumo-type dialect. There certainly will have been Epi-Jomon people speaking an Ainu-related language among the emishi ranks, but what seems clear is that the Izumo-type dialect of Tōhoku must have already had a strong foothold in the northeast before the period of the Emishi Wars of the late eighth century. That dialect is proposed to have spread either with the introduction of rice agriculture by admixed Yayoi peoples in Mid to Late Yayoi, or in the expansion of the mounded tomb culture in the Kofun period.

This study has highlighted a strong need to bring Japanese scholarship on the history and archaeology of Tōhoku into consideration with the farming/language dispersal hypothesis. The complexity of population movements in this area is illustrated by Hakomori (2013). Also needed is ancient DNA sequencing at fine temporal and geographic scales. With increasingly sophisticated techniques for quantifying genetic admixtures (cf. Chikhi, Nichols, Barbujani, & Beaumont, 2002; Dupanloup, Bertorelle, Chikhi, & Barbujani, 2004; Wollstein & Lao 2015), maybe Yayoi- to Nara-period skeletal material from the Kantō and Tōhoku regions will enlighten us on the interactions between indigenous peoples and migrants in the spread of agriculture to the northeast. The deep ancestry associated with the Jōmon will aid in highlighting changing proportions of different types of Asian-related ancestry in these different regions, although it behooves us to be alert to Jōmon continental interaction prior to the Yayoi period.

**Acknowledgements.** We would like to thank Qiaomei Fu for allowing us to cite unpublished work on ancient human genetics of Neolithic East Asia.

**Supplementary Material.** The supplementary material for this article can be found at [https://doi.org/10.1017/ehs.2020.7](https://doi.org/10.1017/ehs.2020.7).

**Author contributions.** Gina Barnes designed the project and wrote the Introduction and Agriculture section; Elisabeth de Boer contributed the Language section; Melinda Yang contributed the section on Jōmon genetics; all three contributed comments and revisions to the main text and Conclusions; Aileen Kawagoe acted as researcher and discussant, with many of her comments incorporated herein.

**Financial support.** The linguistic part of this project has received funding from the European Research Council under the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 677317.

**Publishing ethics.** As a review article, elements of our individual research and previous publications have been included in the wider scholarly context.

**Conflicts of interest.** None of the authors have conflicts of interest to declare.

**Research transparency and reproducibility.** The genetic data shown are all previously published, as described in Table S5. Data in Table S5 are presented in a format that can be read by the software Admixture (version 412; Patterson et al., 2012) where genotyped data are available for the ‘1240K’ SNP panel and can be found in a publicly available database: v42.4, David Reich Lab Harvard University (Reich, 2020). Data used to visualize Figures 2 and S1 are reproduced in Tables S3 and S4. Scripts used to conduct the genetic analyses and visualize the data can be made available upon request to myang@richmond.edu

**References**


