

ORIGINAL ARTICLE

Revisiting the peopling of Japan: an admixture perspective

Rita Rasteiro¹ and Lounès Chikhi^{1,2}

The first inhabitants of Japan, the Jomon hunter-gatherers, had their culture significantly modified by that of the Yayoi farmers, who arrived at a later stage from mainland Asia. How this change took place is still debated, but it has been suggested that modern Japanese are the product of an admixture between these two populations. Here, we applied for the first time an admixture approach to study the Jomon–Yayoi transition, using Y-chromosomal data published earlier. Our results suggest that the Neolithic transition, in this part of the world, probably took place by a process of demic diffusion. We also show that for two populations that could not have contributed to this process, our approach is able to detect inconsistencies when they are used as parental populations. However, despite these promising results, we could not locate precisely the geographical origin of the Yayoi in mainland Asia, as different potential sources gave similarly good results. This suggests that more loci would be required for a better understanding of the peopling of Japan.

Journal of Human Genetics (2009) 54, 349–354; doi:10.1038/jhg.2009.39; published online 8 May 2009

Keywords: admixture; Japan; Jomon; Neolithic; Yayoi; Y-chromosome

INTRODUCTION

The development and spread of farming, referred to as the 'Neolithic transition,' was one of the major demographic events of human prehistory.¹ This process took place independently in different geographical areas, each one most likely associated with different demographic changes and with different domesticated animals and plants. In principle, each of these changes can be described as a process by which at least two human groups (Paleolithic hunter-gatherers (HG) and Neolithic farmers) admixed to different extents. These processes can be seen as admixture models and although they have been used to study the Neolithic transition in Europe,^{2–4} this has not been the case for Asia. Here, we focus on Eastern Asia, where the transition to agriculture has long been controversial, specifically regarding the prehistory of Japan.^{5–8}

Archeological data suggest that there were probably two migratory waves of incoming people, both from the Asian continent to Japan. The first migration took place c. 38 000–37 000 BP, before the Pleistocene land bridges were submerged,⁹ and later gave rise to the Jomon culture ($\geq 12 000$ BP).¹⁰ Although they were an HG society, the Jomon were the holders of one of the oldest pottery cultures known in the world and probably also led a sedentary or semi-sedentary life, well before showing any clear evidence of having developed agriculture.^{1,11} A long time after this period, c. 2300 BP, a second wave of people, together with a 'wet rice culture,' weaving and metalwork, entered the southern Kyushu island (Figure 1), through the Korean Peninsula,¹² and then spread northeastward, starting the Yayoi period.

The transformation and the replacement models represent the two opposite extremes of the demographic models that have been proposed to explain the peopling of Japan and the contribution of both Jomon and Yayoi populations to modern Japanese. While the latter model claims that modern Japanese should be descendents of the incoming Yayoi who replaced completely the Jomon people,⁵ the former entails a movement of the Yayoi culture and ideas rather than people, with consequently no genetic contribution of the Yayoi to modern Japanese.⁶ However, reality must have been less extreme and currently it is widely accepted that the modern Japanese are the result of an admixture between the two populations that produced both the Jomon and Yayoi cultures. This was suggested by Hanihara⁷ and Matsumura⁸ on the basis of dental and cranial characteristics, and more recently by a number of authors who used genetic data,^{13–17} including ancient DNA.^{18,19}

As one of the few points on which all studies agree is that at least two human groups admixed at some point in the past, a simplified way to explain the data is the use of an admixture approach. However, one of the limitations, of most of admixture models, is that they usually ignore genetic drift since the admixture event. This is why we used an approach²⁰ that has already been applied to address the Neolithic transition in Europe^{2,21,22} and where drift is explicitly accounted for. We expect that the admixture process varied geographically, as the incomers (early farmers) were meeting and admixing with the local populations and their descendants were themselves mixing with other populations. Although the admixture process must

¹Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, Oeiras, Portugal and ²UMR CNRS 5174 Evolution and Diversité Biologique, Université Paul Sabatier, 118 route de Narbonne, Bât. 4R3 b2, Toulouse cedex 4, France

Correspondence: R Rasteiro, Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, 2780-156 Oeiras, Portugal.

E-mail: rrasteiro@igc.gulbenkian.pt

Received 8 January 2009; revised 6 April 2009; accepted 8 April 2009; published online 8 May 2009



Figure 1 Map of the Japanese Islands. Approximate geographical locations of the Japanese populations analyzed in this study. The other samples used as parental are not represented on the map.

have been complex, we can predict that a correlation should exist between the admixture level at a particular location, measured by the contribution of one parental population, and the geographical distance from that parental population, as has been shown in Europe.² We also expect that this relationship would not hold, if the same analysis was performed using parental populations that could not have contributed. We note here that to carry out the admixture analysis, two modern populations are chosen to approximate the haplotype frequencies of the original parental populations (Jomon and Yayoi). The choice of these parental populations is based on archeological evidence and is described in the Materials and methods section.

Thus, the aim of this study was to determine whether an admixture approach could be fruitful to studying the Neolithic transition in Japan. To do this we analyzed Y-chromosomal data from the literature, using different ‘parental’ populations, to test different hypotheses. In a first set of analyses, the parental populations were chosen among a set of Asian populations (see below for details). The data were also analyzed by using, as a negative test, populations that were unlikely to have contributed to the gene pool of modern Japanese, namely a European (Sardinia) and a geographically closer (Oceania) population, and for which comparable Y-chromosomal data were available. Altogether, we show that admixture models can indeed provide interesting insights into the peopling of Japan. In particular, our results strongly suggest that the Yayoi immigrants spread by a process similar to the demic diffusion, first proposed for Europe by Ammerman and Cavalli-Sforza.²³

MATERIALS AND METHODS

Populations used

The analyses presented in this study were based on published non-recombining Y-chromosome data of Japanese and other Asian populations. A total of 275 individuals, representing each of the Japanese islands (Figure 1), were analyzed: Ainu (20), Aomori (26), Shizuoka (61), Tokushima (70), Kyushu (53) and Okinawa (45). All the Japanese data were published by Hammer *et al.*,¹⁵ except the Ainu data that were pooled with data from Tajima *et al.*²⁴ Mainland Asian data¹⁵ were obtained for populations from Northeast (441), Southeast (683)

and Central (419) Asia and also a sample from Korea²⁵ (43). We also used two additional populations, Sardinia²⁶ (77) and Oceania¹⁵ (209), as parental populations in the admixture model used (see below). Y-chromosome binary haplogroups were defined by the analysis of the binary polymorphisms described in Hammer *et al.*¹⁵ The Y-chromosome lineages from Japan, mainland Asia, Korea, Oceania and Sardinia followed the haplogroups nomenclature of the Y Chromosome Consortium.²⁷

The admixture model

The admixture method used assumes that an ‘admixed’ or ‘hybrid’ population (H), of size N_H , is the result of the admixture of two independent parental populations, P_1 and P_2 , of size N_1 and N_2 , T generations ago, with respective contributions p_1 and p_2 ($p_2=1-p_1$). After the admixture event, the three populations are isolated and assumed to evolve independently under pure genetic drift. The advantage of this model, and of the associated inference methods, is that (i) the three populations have different N_i (where i can be 1, 2 or h) and (ii) drift and admixture are separated. It is important to note that, by explicitly accounting for drift after the admixture event, the method allows for present-day samples from the parental populations to have drifted significantly from the original unknown parental populations. In addition, the method does not fix the original parental allele frequencies. Instead, they were allowed to vary and this uncertainty is explicitly taken into account. A Bayesian full-likelihood method based on this model was developed by Chikhi *et al.*,²⁰ implemented in the LEA (Likelihood-based Estimation of Admixture) software freely available at www.igc.gulbenkian.pt/static/docs/LEA.zip.²⁸ LEA implements a Monte Carlo Markov Chain algorithm to jointly infer all the parameters of the admixture model, including the ancestral allelic configurations that are compatible with the present, observed allelic frequencies. For each analysis, LEA was run for 300 000 steps, as it has been shown that it is enough to reach an equilibrium for the Y-chromosomal data.^{2,20}

Choice of parental populations

For simplicity and consistency, the P_1 population was always used to represent the HG or Jomon, whereas the population P_2 was used to represent the farmers of the Yayoi period. Hence, the parameter p_1 represents the ‘Jomon’ contribution, at the moment of admixture, whereas p_2 would represent the ‘Yayoi’ contribution. However, like all admixture methods, it requires that these parental populations be defined. Although it is unlikely that today’s populations are direct descendants from any of the original groups, we can use current archeological and anthropological data to identify populations that are likely to be less admixed, and use them as descendants from the original parental populations. It is noteworthy that if there has been a lot of admixture in these parental populations, the general effect should be to blur the original signal, and make it less clear. Therefore, any signal observed today should be an indication that some information is still present in the data. Although the Jomon culture has almost been replaced across Japan, there are some indigenous minority ethnic groups who live in the peripheral areas of Japan, which are considered descendants of this ancient culture.^{7,16,17,24} Those are the Ainu people, in the northern part of the Hokkaido Island, and the Ryukyuan, in the southern Ryukyu Islands. Moreover, the Ainu lived in relative isolation until the end of the nineteenth century²⁹ and show unique physical characteristics; such as hairiness, wavy hair and deep-set eyes, which are different from those of most Japanese. In contrast, the Ryukyuan kingdom had past relations with mainland Japan since medieval times, with possibly frequent gene flow,^{30,31} but it is thought to have nevertheless maintained genetic differentiation from mainland Japan.³² For these reasons, the admixture analyses were performed using either the Ainu or the Ryukyuan, the latter represented by the Okinawa sample, as descendants of the P_1 population, in the different analyses. For the descendants of the Yayoi (considered the P_2 population), different parental populations from mainland Asia were also used, namely NEA (northeast Asia), SEA (southeast Asia), CAS (central Asia) and Korea.

To determine whether our approach was robust to incorrect specification of the parental populations, we also used as P_2 two populations that are unlikely to have contributed to the gene pool of the Japanese: one from Europe (Sardinia) and the other from a closer geographical area, Oceania. We expected

that there should be no correlation (or at least much less) between admixture and geographical distances in these cases.

Altogether, each of the four Japanese ‘admixed’ populations (Aomori, Shizuoka, Tokushima, Kyushu) were analyzed using two populations for P_1 (Ainu and Okinawa), six populations for P_2 (including Sardinia and Oceania), making a total of 12 different sets of admixture analyses. In addition, for each admixture analysis, the parental populations were also considered as ‘admixed’ populations. For example, we used the Ainu as P_1 , as H , against the six different P_2 populations. This kind of analysis allowed us to quantify the uncertainty around the estimation of p_1 , as the hybrid and one parental population (here P_1) are exactly identical. Thus, the p_1 posteriors should always have a mode equal or very close to one, with a variance related to both the sample size and drift since the admixture event. Of course, when the Ainu and Okinawa are used as ‘pseudo-hybrids’ the corresponding posteriors were not used in the regression analysis described below.

Calculating drift

The LEA software also allowed us to estimate genetic drift since the admixture event in the three populations, through the parameters $t_i = T/N_i$, where i corresponds to 1 (Jomon parental population), 2 (Yayoi parental population) or h (Japanese hybrid population). Populations that had developed agriculture earlier would have increased in size earlier and would thus exhibit lower amounts of drift since the admixture event. Consequently, if the admixture model is consistent, the t_1 values should in general be higher than the t_2 values, whereas t_h values should be more variable across populations.

Spatial variation of admixture: regression analysis

To detect, quantify and assess the significance of any geographical trend in the admixture proportions across Japan, we used a linear regression approach similar to that used by Chikhi *et al.*² The idea is to determine whether there is a correlation between the ‘Yayoi contribution,’ measured by p_2 , and the geographical distance from the population used for P_2 . For each location sampled in Japan, we computed a geographical distance from the sample used as P_2 and then estimated a linear regression between this distance and p_2 . To account for the uncertainty around p_2 , we followed the resampling approach used by Chikhi *et al.*² For each of the Japanese samples, one p_2 value was randomly sampled from the corresponding posterior distribution. This process was repeated 1000 times to obtain the empirical distribution of regression lines. This was done independently, for each set of admixture analyses performed, using a particular pair of parental populations. A similar approach was used for t_h , to determine whether drift in the admixed populations was also correlated with the geographical distance. The geographical distance was calculated as a straight line from the central point of the area corresponding to the population used as P_2 (for example, Korea to Seoul for Korea), taking into account an entering in Japan from Korea, through Kyushu. It is worth noting that the spatial points used are necessarily non-independent, because of local gene flow, as was, for instance, noted by Sokal *et al.*³³ As a consequence, allele or haplotype frequencies are spatially autocorrelated, and hence violate the assumption of independence necessary to calculate the significance of a linear regression, using classical approaches. This is why we did not perform such tests, and used the values to represent the relationship between the parameter of interest and geographical distance.

F_{ST} analysis

The genetic structure of the populations was also described using F_{ST} values. These values were computed with the equation $F_{ST} = (\bar{H}_T - \bar{H}_S) / \bar{H}_T$,³⁴ using the Okinawans and Ainu against all other Asian populations.

RESULTS

Admixture proportions

Figure 2 shows, for the different pairs of parentals tested, the mode of the p_1 posterior distributions (represented in Figure 3a), where p_1 represents the HG contribution to the Japanese populations. The modes represent the most probable values, but as the p_1 distributions are wide they should be interpreted with caution. In fact, earlier

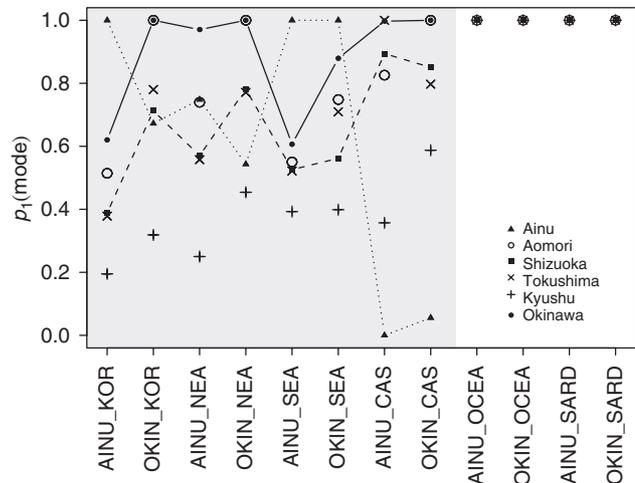


Figure 2 Jomon contribution, across Japan. Mode of the p_1 posterior distributions, for all the Japanese populations analyzed, with p_1 representing the hunter-gatherers Jomon contribution to modern Japanese. In the x axis are represented the parental populations used (P_1 followed by P_2). The letter codes are as follows: AINU—Ainu, OKIN—Okinawa, KOR—Korea, NEA—northeast Asia, SEA—southeast Asia, CAS—central Asia, OCEA—Oceania and SARD—Sardinia.

simulation results suggest that any point estimate (median, mean or mode) should be interpreted with care²⁰ and we should rather focus on spatial trends, if any, across the populations.² Although it was difficult to infer precise contributions, this figure shows a clear difference in the results, whether Asian or non-Asian populations are used as P_2 : (i) with Sardinia or Oceania, the p_1 modes are equal to one, indicating no contribution of these populations to the Japanese populations and (ii) with Asian populations, the modal p_1 values are widespread and are most of the time higher than 0.2, suggesting a variable but significant contribution of the HG. There is one exception, when we use the CAS population as P_2 , the p_1 modes for the Ainu are very low, which is counterintuitive. A closer look at these posteriors shows that they are very flat and hence that there is little information to infer p_1 .

Even though there is an uncertainty on p_1 values for specific populations (see Figure 3a), a geographical trend in the ‘Yayoi’ contribution (p_2) is found (Figure 3b). The randomization approach applied to test this trend is summarized in Table 1 (and Supplementary Figure SI 1a), where the correlation coefficients obtained from the linear regressions are represented. Although the values are small, this figure again shows a clear signal. First, when the Sardinian or Oceanian samples are used as P_2 , which are unlikely to have contributed to the Japanese gene pool, there is a positive or no correlation, suggesting that the contribution of these two populations increases (or stays constant) with geographical distance from their current location. Second, in all the other analyses (that is, when the P_2 populations are Korea, NEA, SEA and CAS), the correlation coefficients values are negative; that is, p_2 logically decreases with geographical distance from the population used as P_2 . In addition, with the exception of SEA as P_2 , when comparing the analyses of Ainu vs Okinawa as P_1 , the most negative correlation values seem to be associated with the Ainu. This trend is, however, only close to significance ($P=0.058$).

Drift

In Figures 4a and b, the amount of drift between the present-day samples of the populations used as P_1 or P_2 , and the ancestral populations, is represented through t_1 and t_2 respectively. Comparing

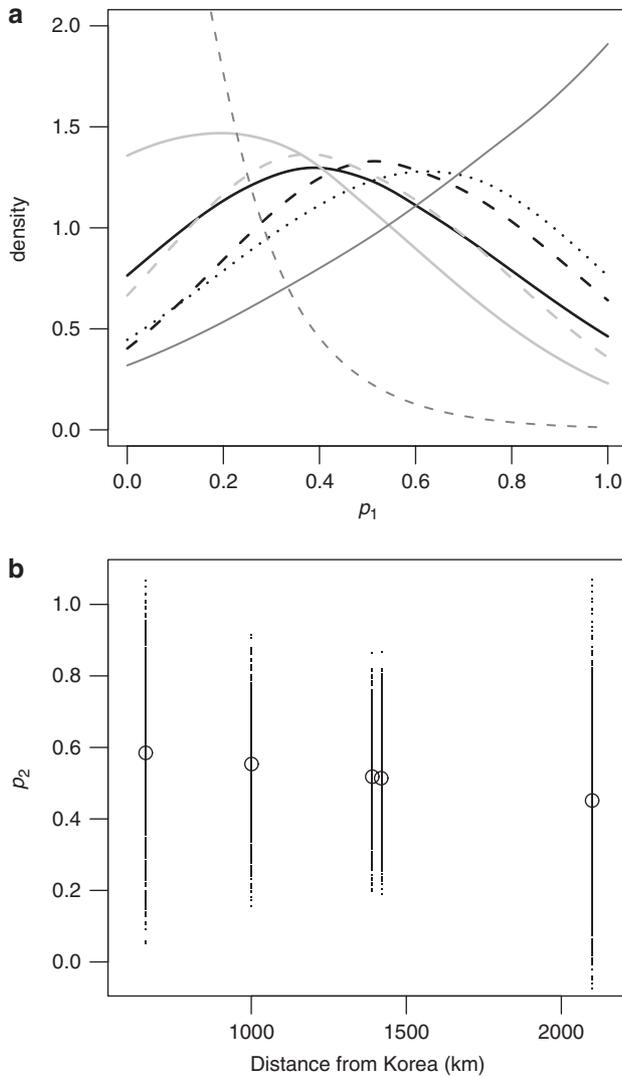


Figure 3 Jomon and Yayoi contributions, across Japan. **(a)** Posterior distributions of p_1 for all Japanese populations (Kyushu—gray, Tokushima—dashed gray, Shizuoka—black, Aomori—dashed black and Okinawa—dotted black) and parental populations (thinner lines: Ainu—dark gray and Korea—dashed dark gray) used. Each curve corresponds to the analysis of a specific hybrid population. **(b)** Linear regression of p_2 against geographical distance from P_2 (Korea). The circles represent the mean value for each population. These analyses were carried out using the Ainu and Korean populations as P_1 and P_2 , respectively.

the t_1 and t_2 distributions clearly indicates that the two estimates are extremely different. In fact, in all the analyses done, the t_1 modal values, of all the Japanese populations studied, were always greater than t_2 and, at the same time, were higher if we used the Ainu sample as P_1 , instead of Okinawa (data not shown). Moreover, the t_1 posteriors were wider and had more variable modal values than t_2 , especially when using the Ainu as P_1 , but nevertheless were similar to each other.

When we analyzed the t_h estimates, they appeared to be highly variable (data not shown) and to display a geographical trend. As for p_2 , we applied a regression of these estimates against the geographical distance, from the population used as P_2 , to all the possible combinations of parentals. The correlation coefficients obtained from these linear regressions are represented in Table 1 (and Supplementary Figure SI 1b). All the correlation coefficients are positive (except with Ainu vs Sardinia), showing that the t_h values increase as

Table 1 Spatial variation of admixture and drift

P_2	P_1			
	Ainu		Okinawa	
	r_{p_2}	r_{t_h}	r_{p_2}	r_{t_h}
NEA	-0.08730	0.03413	-0.24240	0.64510
CAS	-0.04634	0.06180	-0.18400	0.48580
SEA	-0.17550	0.06690	-0.11460	0.65360
Korea	-0.13770	0.06445	-0.19810	0.66640
Oceania	0.16190	0.03664	0.24639	0.62980
Sardinia	0.15670	-0.02782	-0.00095	0.56270

Abbreviations: NEA, northeast Asia; CAS, central Asia; SEA, southeast Asia. Correlation values (r), for the p_2 and t_h regression analyses described in the Materials and methods section, for the parental population pairs used (represented by P_1 and P_2).

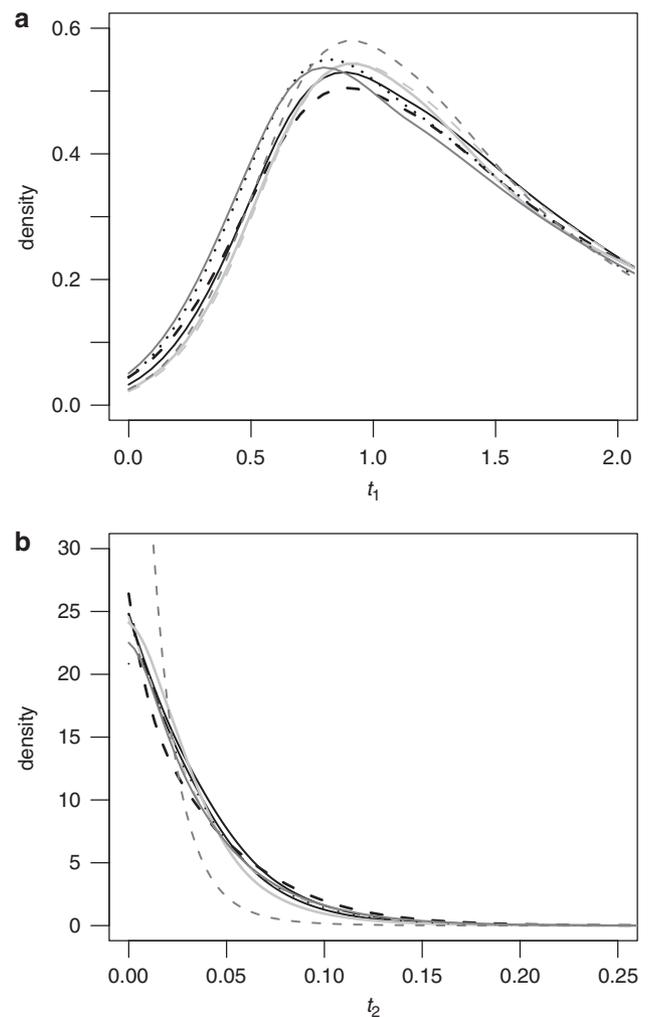


Figure 4 Distributions of the t_i 's for all Japanese populations. **(a)** Posterior distributions of t_1 . The different curves represent the amount of genetic drift, since the admixture event, between the present sample of Ainu and the ancestral populations of HG (Jomon) that interbred with the incoming farmers (Yayoi). **(b)** Posterior distributions of t_2 . As in **a**, but for the drift between the Korean and Yayoi populations instead. The color codes are as in Figure 3.

geographical distance increases, but at the same time are much smaller when we use the Ainu as P_1 , on the order of 0.034–0.067 vs the 0.486–0.666 for the Okinawans.

Table 2 Population differentiation with Ainu and Okinawans populations

	F_{ST}	
	Ainu	Okinawa
Ainu	0	0.096
Aomori	0.159	0.018
Shizuoka	0.195	0.029
Tokushima	0.210	0.042
Kyushu	0.219	0.047
Okinawa	0.096	0
Korea	0.329	0.131
NEA	0.213	0.107
CAS	0.308	0.174
SEA	0.311	0.125

Abbreviations: NEA, northeast Asia; CAS, central Asia; SEA, southeast Asia.

F_{ST}

Table 2 (see also Supplementary Figure SI 2) shows higher F_{ST} values in the pairwise comparisons involving the Ainu, compared with those involving the Okinawans. It also shows, that for both sets of pairwise comparisons, the F_{ST} values increase when the geographic distance increases in a southwestern direction from the northern tip of Honshu (Aomori) toward Kyushu. This is particularly interesting as the Ainu and Okinawan are located on opposite sides of the Japanese archipelago and hence of this axis. We also note that this trend of F_{ST} values shows a clear and sudden increase when the samples are taken from the Asian continent, starting with Korea. This is particularly clear despite the fact that, when we consider only the Japanese populations, the F_{ST} values involving Okinawans and Ainu are on different scales, the first set of values all being below 0.1 and the others all being above 0.1. In fact, the F_{ST} between the Ainu and Okinawans (0.096) is the smallest F_{ST} value among the pairs involving the Ainu (which vary from 0.096 to 0.219 in Kyushu), but it is the largest among the pairs involving the Okinawans (which vary from 0.018 in Aomori to 0.096 against the Ainu). Another consequence of this cline is that the Okinawan population seems genetically close to the northernmost populations of Honshu, but strangely not to the Ainu, the only sampled population north of Honshu. Indeed, the F_{ST} with the latter is higher than the F_{ST} values between Okinawans and all other Japanese populations. Thus, the Okinawans appear to be the Japanese who are genetically closest to the Ainu from the Ainu viewpoint, whereas it is exactly the opposite from the Okinawans viewpoint. It is as if the Okinawans were ‘virtually’ located in northern Honshu, and the Ainu were genetically close to them but had been submitted to significant drift.

DISCUSSION

Dual origins of Japanese

Although our results may only reflect the paternal history of the Japanese, they confirm the idea that a significant admixture took place and thus do not support either the replacement or the transformation models between the incoming Yayoi and the local Jomon. Indeed, in the replacement model the estimate of p_1 should be equal to zero (or at least very close, because of statistical uncertainty), whereas in the transformation model p_1 would be close or equal to 1. This is clearly not what we observe. Moreover, our results show a decreasing geographical trend in the Yayoi contribution across Japan, when populations are sampled in a southwest–northeast direction (Figures 3a and b). These results agree with a model in which the first farmers

entered in Japan from Korea, through the closest island (Kyushu), and then spread across most of Japan moving to the northeast (until the geographical limits of the Honshu Island). During the expansion of farmers, it is expected that the rise in population density, because of food production, should lead to a more limited drift (as populations were larger). This can be seen in the gradient observed with the t_h estimates, which suggests that drift is higher in the northernmost populations (with a maximum observed in the Ainu), where the archeological record suggests a later arrival of agriculture.¹¹ In addition, the differences encountered between t_1 and t_2 estimates ($t_1 \gg t_2$) are consistent with a model of an expanding population that dispersed to a less populated area; that is, the P_2 populations (Yayoi) increased in size earlier in time, having suffered a lower amount of drift. Interestingly, the t_1 values are higher than those found in Europe,² which could be because of an earlier introduction of agriculture in Europe.

During the admixture process, it is important to note that the indigenous populations, the Jomon, who admixed with the Yayoi, were probably genetically differentiated from each other across the Japanese islands. How differentiated they were, some 2000 years ago, is difficult to say, but this pre-admixture differentiation should have some implication in the analyses, and their interpretation. The fact that all t_1 posteriors were very similar to each other suggests that even if there was differentiation between HG populations, before the arrival of the Yayoi, they were not dramatically different, compared with the amount of drift that occurred since the admixture event. This indirectly shows that our model, despite its simplicity, captures important aspects of the ‘Neolithic transition’ in the Japanese archipelago. However, we note that, depending on whether we use the Ainu or Okinawans as P_1 , the t_1 estimates are rather different, suggesting a much higher drift when the Ainu are used. This, together with the fact that the Ainu have much higher F_{ST} than the Okinawans against all other populations, suggests that the Ainu have probably had a much lower effective size than the Okinawans. It could be because of a greater isolation, a later and more limited influence of agriculture or a combination of both.

The continental origin of the Yayoi farmers

Several hypotheses have been suggested regarding the geographic origin of the Asian populations, which gave rise to the Yayoi, even though it is usually accepted that they probably entered Japan through South Korea. Nevertheless, skull and teeth morphology inference⁷ and classical markers^{13,16} support an NEA origin.

More recently, Hammer *et al.*¹⁵ placed the Yayoi farmers as having originated in SEA. However, with the same data as Hammer *et al.*¹⁵ (with the exception of the Korean data), the admixture model we used could not establish with so much accuracy the continental origins of these populations, but our results suggest that they entered Japan through the Korean Peninsula. It may be important to note that our approach is model-based and has been tested on simulated data,²⁰ whereas the conclusion reached by Hammer *et al.*¹⁵ was based on visual patterns of allele or haplogroup frequencies and were neither justified by any statistical test, nor by analyses of simulated data. Thus, our approach is not ‘just’ confirming established results, but rather adding more solid results to conclusions whose statistical validity was not determined. In addition, when we found that the exact location of the Yayoi cannot be ascertained with certainty, whereas Hammer *et al.*¹⁵ assert that they arrived from SEA, one should question the strength of the latter statement. This should not be taken as a criticism of the Hammer *et al.*¹⁵ study, which provided both new results and hypotheses to test. Rather, what our results show is that it might be

possible, using an admixture approach, to test different hypotheses, something that has not been done so far. Indeed, our method was able to identify populations that clearly could not have contributed to the modern Japanese gene pool at that time (namely Sardinia and Oceania). If some of the Asian parental populations that were analyzed had generated results similar to those of Sardinia and Oceania, they could have been identified as unlikely parentals. This type of results was not observed, which suggests that these data do not contain enough information to clearly identify the most likely descendent of one of the parental populations of the modern Japanese, namely the Yayoi. This is not necessarily surprising as the Y-chromosome represents only one set of linked markers. We believe thus that until more *loci* are obtained this question may not be easily answered, and should remain open. If we have contributed to making this statement, we feel that a significant step will have been taken.

In summary, our results support at least one admixture event in the peopling of Japan, namely the spread of Yayoi farmers by a process of demic diffusion, similar to the one in Europe during the Neolithic age^{2,23,35}. We suggest that when the Yayoi men entered Japan, and brought with them agriculture and new technologies, they also raised the carrying capacity of the area first colonized, leading to an increase in the size of the newly admixed populations. When this area could no longer support the increased population, their descendents expanded into new territories, repeating the admixture process. By the time the geographic limits of Japan were reached (northeastward until the Hokkaido Island and southwestward in the Ryukyu Islands), there was a gradual dilution of the Yayoi's gene pool. However, in spite of having detected the presence of Jomon and Yayoi contributions in the Japanese populations, the method we used was not capable of locating precisely the area of origin of the ancestral populations, and different populations seemed to produce similarly consistent results. Nevertheless, the general approach appears to provide interesting and promising results, which should open new avenues for research.

ACKNOWLEDGEMENTS

RR was supported by a Fundação para a Ciência e Tecnologia (FCT) grant (ref. SFRH/BD/30821/2006). LC was partly funded by the CNRS, the FCT grant PTDC/BIA-BDE/71299/2006, and the Institut Français de la Biodiversité, Programme Biodiversité de l'Océan Indien (CD-AOOI-07-003). LC travels, between Toulouse and Lisbon, were partly funded by the Programme d'Actions Universitaires Intégrées Luso-françaises 2007/2008, and were made possible thanks to Professor A Coutinho and B Crouau-Roy. LEA calculations were performed, with the help of P Fernandes, using the High Performance Computing Centre (HERMES, FCT grant H200741/re-equip/2005). We are grateful to V Sousa and to anonymous reviewers for helpful comments.

- 1 Bellwood, P. *First Farmers: The Origins of Agricultural Societies* (Blackwell Publishing, Oxford, 2005).
- 2 Chikhi, L., Nichols, R. A., Barbujani, G. & Beaumont, M. A. Y genetic data support the Neolithic demic diffusion model. *Proc. Natl Acad. Sci. USA* **99**, 11008–11013 (2002).
- 3 Currat, M. & Excoffier, L. The effect of the Neolithic expansion on European molecular diversity. *Proc. Biol. Sci.* **272**, 679–688 (2005).
- 4 Dupanloup, I., Bertorelle, G., Chikhi, L. & Barbujani, G. Estimating the impact of prehistoric admixture on the genome of Europeans. *Mol. Biol. Evol.* **21**, 1361–1372 (2004).
- 5 Cavalli-Sforza, L. L., Menzies, P. & Piazza, A. *The History and Geography of Human Genes* (Princeton University Press, Princeton, 1994).

- 6 Mizoguchi, Y. in *Prehistoric Hunter-Gatherers in Japan* (eds. Akazawa, T. and Aikens, C.M.) 107–136 (University of Tokyo Press, Tokyo, 1986).
- 7 Hanihara, K. Dual structure model for the population history of the Japanese. *Jpn. Rev.* **2**, 1–33 (1991).
- 8 Matsumura, H. Differentials of Yayoi immigration to Japan as derived from dental metrics. *Homo* **52**, 135–156 (2001).
- 9 Pope, K. O. & Terrell, J. E. Environmental setting of human migrations in the circum-Pacific region. *J. Biogeogr.* **35**, 1–21 (2008).
- 10 Ono, A., Sato, H., Tsutsumi, T. & Kudo, Y. Radiocarbon Dates and Archaeology of the Late Pleistocene in the Japanese Islands. *Radiocarbon* **44**, 447–494 (2002).
- 11 Highman, C. in *The Human Past: World Prehistory and the Development of Human Societies* (ed. Scarre, C.) 234–263 (Thames and Hudson, London, 2005).
- 12 Jin, H., Kwak, K., Hammer, M. F., Nakahori, Y., Shinka, T., Lee, J. *et al.* Y-chromosomal DNA haplogroups and their implications for the dual origins of the Koreans. *Hum. Genet.* **114**, 27–35 (2003).
- 13 Sokal, R. R. & Thomson, B. A. Spatial genetic structure of human populations in Japan. *Hum. Biol.* **70**, 1–22 (1998).
- 14 Hammer, M. F. & Horai, S. Y chromosomal DNA variation and the peopling of Japan. *Am. J. Hum. Genet.* **56**, 951–962 (1995).
- 15 Hammer, M. F., Karafet, T. M., Park, H., Omoto, K., Harihara, S., Stoneking, M. *et al.* Dual origins of the Japanese: common ground for hunter-gatherer and farmer Y chromosomes. *J. Hum. Genet.* **51**, 47–58 (2006).
- 16 Omoto, K. & Saitou, N. Genetic origins of the Japanese: a partial support for the dual structure hypothesis. *Am. J. Phys. Anthropol.* **102**, 437–446 (1997).
- 17 Horai, S., Murayama, K., Hayasaka, K., Matsubayashi, S., Hattori, Y., Fucharoen, G. *et al.* mtDNA polymorphism in East Asian Populations, with special reference to the peopling of Japan. *Am. J. Hum. Genet.* **59**, 579–590 (1996).
- 18 Horai, S., Kondo, R., Murayama, K., Hayashi, S., Koike, H. & Nakai, N. Phylogenetic affiliation of ancient and contemporary humans inferred from mitochondrial DNA. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **333**, 409–417 (1991).
- 19 Oota, H., Saitou, N., Matsushita, T. & Ueda, S. A genetic study of 2,000-year-old human remains from Japan using mitochondrial DNA sequences. *Am. J. Phys. Anthropol.* **98**, 133–145 (1995).
- 20 Chikhi, L., Bruford, M. W. & Beaumont, M. A. Estimation of admixture proportions: a likelihood-based approach using Markov chain Monte Carlo. *Genetics* **158**, 1347–1362 (2001).
- 21 Belle, E. M. S., Landry, P. & Barbujani, G. Origins and evolution of the Europeans' genome: evidence from multiple microsatellite loci. *Proc. Biol. Sci.* **273**, 1595–1602 (2006).
- 22 Chikhi, L. in *Examining the Farming/Language Dispersal Hypothesis* (eds. Bellwood, P. & Renfrew, C.) 435–447 (McDonald Institute for Archaeological Research, Cambridge, 2003).
- 23 Ammerman, A. J. & Cavalli-Sforza, L. L. *The Neolithic Transition and the Genetics of Populations in Europe* (Princeton Univ. Press, Princeton, 1984).
- 24 Tajima, A., Hayami, M., Tokunaga, K., Juji, T., Matsuo, M., Marzuki, S. *et al.* Genetic origins of the Ainu inferred from combined DNA analyses of maternal and paternal lineages. *J. Hum. Genet.* **49**, 187–193 (2004).
- 25 Xue, Y., Zerjal, T., Bao, W., Zhu, S., Shu, Q., Xu, J. *et al.* Male demography in East Asia: a north-south contrast in human population expansion times. *Genetics* **172**, 2431–2439 (2006).
- 26 Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S., Beckman, L. E. *et al.* The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* **290**, 1155–1159 (2000).
- 27 YCC. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res.* **12**, 339–348 (2002).
- 28 Langella, O., Chikhi, L. & Beaumont, M. LEA (likelihood-based estimation of admixture): a program to estimate simultaneously admixture and time since the admixture event. *Mol. Eco. Notes* **1**, 357–358 (2001).
- 29 Hudson, M. The Linguistic Prehistory of Japan: some archaeological speculations. *Anthropol. Sci.* **102**, 231–255 (1994).
- 30 Toma, T., Tsunehiko, H., Hajime, S., Kuniaki, H. & Hajime, I. Metric dental diversity of Ryukyu Islanders: a comparative study among Ryukyu and other Asian populations. *Anthropol. Sci.* **115**, 119–131 (2007).
- 31 Haneji, K., Hanihara, T., Sunakawa, H., Toma, T. & Ishida, H. Non-metric dental variation of Sakishima Islanders, Okinawa, Japan: a comparative study among Sakishima and neighboring populations. *Anthropol. Sci.* **115**, 35–45 (2007).
- 32 Yamaguchi-Kabata, Y., Nakazono, K., Takahashi, A., Saito, S., Hosono, N., Kubo, M. *et al.* Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am. J. Hum. Genet.* **83**, 445–456 (2008).
- 33 Sokal, R. R., Harding, R. M. & Oden, N. L. Spatial patterns of human gene frequencies in Europe. *Am. J. Phys. Anthropol.* **80**, 267–294 (1989).
- 34 Nei, M. F-statistics and analysis of gene diversity in subdivided populations. *Ann. Hum. Genet.* **41**, 225–233 (1977).
- 35 Barbujani, G., Sokal, R. R. & Oden, N. L. Indo-European origins: a computer-simulation test of five hypotheses. *Am. J. Phys. Anthropol.* **96**, 109–132 (1995).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)