Commentary

From molecular genetics to archaeogenetics

Colin Renfrew

McDonald Institute for Archaeological Research, Downing Street, Cambridge CB2 3ER, United Kingdom

Applying molecular genetics to questions of early human population history, and hence to major issues in prehistoric archaeology, is becoming so fruitful an enterprise that a new discipline—*archaeogenetics*—has recently come into being. That many of its applications have so far related to prehistoric Europe (1) is due in part to the detailed archaeological attention devoted to Europe by a series of nineteenth and twentieth century scholars (2). It is also due in part to the early application of a specific demographic model, the "wave of advance" (3), to explain the chronological patterning that emerged as farming spread across Europe at the onset of the neolithic period (4) and to elucidate the structuring resulting from an early principal components analysis of the classical genetic markers for Europe (5, 6). The application of DNA sequencing, permitting female lineages to be investigated through mtDNA (7) and male lineages through the Y chromosome (8), has already brought a series of new questions into perspective, generating lively debate (9, 10). The time is ripe, therefore, for more closely focused regional studies, devoted to specific historical problems. The paper by Wilson *et al.* (11) in this issue of PNAS breaks new ground in investigating one such early demographic episode, the Viking conquest of the Orkney Islands (Fig. 1) in the ninth century A.D. It also raises a number of general problems that emerge when reconstructing demographic history.

The archaeological record in the Orkney Islands (12) bears out the picture conveyed in the Norse sagas (13) that Viking princelings from Scandinavia took control of Orkney, establishing the dynasty of the Norse earls. Because that record indicates considerable continuity from the preceding Pictish period as well as Norse innovations, it has always been a matter for surprise that the surviving place names of Orkney so comprehensively reflect the Norse language of the Viking incomers, with hardly any surviving Pictish toponyms (14). These toponyms do, however, survive in the Highlands of the Scottish mainland immediately to the south. It is thus highly interesting that Wilson *et al.* (11) find their Orkney Y chromosome sample to be intermediate between their Irish and Welsh samples on the one hand (which they assume to be representative of pre-Norse Orkney also) and, on the other, the sample from Norway (the Viking homeland). This is a very suggestive finding. Note, however, that the term "Celtic" with which the authors designate the Irish and Welsh samples is a linguistic one that could also be used for the (pre-Norse) Pictish population of Orkney whose little-understood language is currently assigned to the Celtic language family (15).

A further important development in archaeogenetics reflected in this paper is the remarkable long-term continuity in the use of surnames as secure indicators of paternal lineage, as has previously been observed in Ireland by Hill *et al.* (16), where significant genetic differences were noted between Gaelic and non-Gaelic surname samples. In one province (Connaught) the Gaelic surname samples showed a frequency of 98% for haplogroup 1, relating to the Atlantic Modal Haplotype discussed in the Orkney study reviewed here (11). A consideration of Orcadian surnames, excluding those associated with Scottish settlers subsequent to the fourteenth century A.D., allows 38% of the (male) chromosomes to be identified as Scandinavian in origin.

The archaeological record bears out the picture conveyed in the Norse sagas that Viking princelings from Scandinavia took control of Orkney.

It is unfortunate, however, that the general underlying similarity in the mtDNA haplogroup distributions in European populations (17) was reflected in an apparent lack of structure in the samples analyzed, so that no evidence is available to indicate whether an equivalent female population from Norway accompanied the male migrants inferred from the Y chromosome haplogroup frequencies. There may, however, be more work to be done here because the principal components analysis undertaken on the mtDNA data (figure 2 of ref. 11) shows Orkney more than twice as distant from the Basque sample than is the Norwegian sample when mtDNA is considered. It is perhaps disappointing that Orcadian samples were not included in the analysis undertaken for microsatellites on the X chromosome, where Basque and Norway are well differentiated in the principal component analysis.

The paper by Wilson *et al.* (11) raises other challenging issues that have yet to be resolved. In the first place, it poses explicitly the question of the extent to which major cultural transitions, as documented in the archaeological record, involved the movement of people or simply of ideas. And of course the authors have successfully shown that there was indeed significant gene flow accompanying the Norse conquest of the Orkneys. But the effectiveness of such analysis inevitably depends on the existence of diagnostic criteria that would distinguish the populations of the receptor and donor areas, here the Orkneys and Norway, at the time in question. They have indeed documented that for the relevant male markers in those two areas, but the matter remains open on the female side in view of the current lack of distinctive parameters (as between Orkney and Norway) where the mtDNA data are concerned. When they make the observation that "patterns of Y chromosome variation indicate that Neolithic and Iron Age transitions in the British Isles occurred without large-scale male movements," one is entitled to apply the same strict criteria. If, for example, the populations situated on both sides of the English Channel had broadly similar Y chromosome haplotype frequencies immediately before the neolithic transition, it is perfectly possible in theory that a very substantial population movement could have taken place across the Channel without significantly changing the haplotype frequencies on either the French or the English side. These may be difficult matters to investigate, but it

See companion article on page 5078.

*E-mail: dap38@cam.ac.uk.

www.pnas.org/cgi/doi/10.1073/pnas.091084198

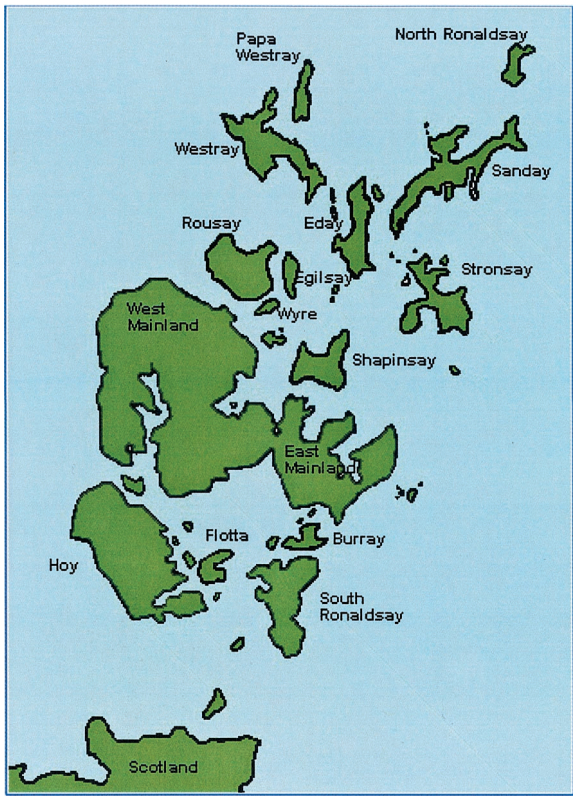

Fig. 1 The Orkney Islands located north of Scotland. Archaeogenetic data suggest that Viking settlements left substantial genetic as well as cultural influence on this Scottish archipelago. [Reproduced with permission from www.orknet.co.uk (Copyright 1997, Orknet).]

should be observed that the analysis here has proceeded without the use of samples from mainland Britain: the samples considered are from peripheral islands (Orkney, Ireland, and Anglesey). Although the question that they pose is an entirely valid one, there would clearly be need of a more ambitious sampling strategy to begin to formulate a definitive answer. The matter is underlined, so far as the neolithic is concerned, by the circumstance that the Orkney Islands may well have lacked any permanent population until the arrival of neolithic settlers (18). This, like the inception of the neolithic period in Crete (19), is one instance where the movement of females as well as males seems an indispensable assumption!

There is another important finding in the paper by Wilson *et al.* that is both interesting in itself and leads them to an argument whose inferential foundations may be questioned. They rightly emphasize the strong similarity in the Y chromosome haplogroup frequencies between the Basque country, and the Welsh and Irish samples. The three in consequence cluster closely on the principal components diagram for the Y chromosome data. This observation leads the authors to the following conclusion: "in the British Isles the Neolithic transition did not entail a major demographic shift. Accordingly, farming may have spread in Britain more through cultural transmission than through some form of gene flow." Later in the paper they observe: "This is in sharp contrast with the mtDNA pattern [in the principal components analysis] in which the [Irish and Welsh] populations are closer to the centre of the plot, indicating that they have undergone more female-mediated gene flow from other European populations than the Basques have. Thus at least one of the cultural transitions in the British Isles since the Upper Palaeolithic must have involved a demic component on the female side." I suspect, however, that these arguments rest on two uncertain premises that illustrate the general difficulties in interpretation of all archaeogenetic data.

The first problem is the inference that, if the Irish, Welsh, and Basque Y chromosome haplotype frequencies are closely similar today and may have been so in Upper Palaeolithic times, then no significant gene flow into Ireland and Wales in the male line occurred at the onset of the neolithic. As noted above, very significant gene flow could have occurred at that time without notable impact on haplotype frequencies if the donor and receptor populations were themselves not distinguishable in that respect. Such may well have been the case. The second problem lies with the mitochondrial data and with the conclusion that the female-mediated gene flow inferred must have occurred "*since the Upper Palaeolithic*." This conclusion rests on the implicit assumption that much of the variability now seen in mtDNA haplogroup distributions entered Europe since the Upper Paleolithic, an assumption developed in the original "wave of advance" model (4, 5) but one contested in subsequent mtDNA studies (20).

These remarks are not intended as a criticism of the paper by Hill *et al.* (16), nor of the original "wave of advance" model for demic diffusion, but rather to suggest the need for a "second generation" wave of advance model that will take into account not only the cultural interactions between the incoming farmers (initially from Anatolia to Greece) and the indigenous population, but also the genetic and demographic consequences of the intermarriages between the two groups. Already the Y chromosome data produced in the important paper by Semino *et al.* (21) clearly show a decline from south-east to north-west Europe in the frequency of the supposed "neolithic" haplogroup. If we imagine that, through assimilation and intermarriage, an actively farming community in a region to which farming had recently spread contained, T centuries after the inception there of farming, a genetic input of X% (say 10%) from the indigenous mesolithic population and retained (100-X)% (i.e., 90%) genetic input deriving from the adjacent source area from which the most recent stage of the spread occurred, we have the basis for a model, the Staged Population-Interaction Wave of Advance (SPIWA). Such assumptions could yield an exponential decline across Europe (along the direction of spread) in the frequency of the "incomer-farmer" genes as against the indigenous mesolithic genes (which are assumed as a first approximation to be homogeneous). The SPIWA model addresses the same general problem as the "wave of advance" demic diffusion model of Ammerman and Cavalli-Sforza (3) but brings into direct consideration the interactions between the indigenous mesolithic population and the incoming farmers whose demographic progress was considered in the original model. The "neolithization" process is here viewed as a series of successive steps or stages, in each of which the incoming farming population interacts (culturally and genetically) with the local mesolithic population. Population growth takes place with the inception of farming as in the original model, but the fall-off in gene flow (and the clinal reduction) is exponential rather than linear, a pattern more in keeping with recent Y chromosome work (21).

Moreover, it should be observed that, if the indigenous gene frequencies happened at the outset to be the same as those of the incoming neolithic farmers, the entire "wave of advance" could take place, precisely as in the original model, without any impact on the haplotype frequencies at all. Very significant gene flow could indeed occur, as in the original model, but the similarities or equivalence between donor and receptor haplotype frequencies would make the process invisible to gene frequency analysis.

It is not, of course, suggested here that such crude models could approximate to the complex reality of prehistoric Europe, but simply that we need to construct further models appropriate to that task. The paper by Wilson *et al.*, with its well-defined regional focus, certainly draws attention to many of the right questions. With the increasing availability of data for both male (Y chromosome) and female (mtDNA) variability in Europe (21, 20), the way will increasingly be open for useful regional studies of this kind.

1. Renfrew, C. (2000) in *Archaeogenetics: DNA and the Population Prehistory of Europe*, eds. Renfrew, C. & Boyle, K. (McDonald Institute, Cambridge, U.K.), pp. 3–12.
2. Trigger, B. G. (1989) *A History of Archaeological Thought* (Cambridge Univ. Press, Cambridge, U.K.).
3. Ammerman, A. J. & Cavalli-Sforza, L. L. (1973) in *The Explanation of Culture Change: Models in Prehistory*, ed. Renfrew, C. (Duckworth, London), pp. 343–58.
4. Ammerman, A. J. & Cavalli-Sforza, L. L. (1984) *The Neolithic Transition and the Genetics of Population in Europe* (Princeton Univ. Press, Princeton).
5. Menozzi, P., Piazza A. & Cavalli-Sforza, L. L. (1978) *Science* **201,** 786–792.
6. Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. (1994) *The History and Geography of Human Genes* (Princeton Univ. Press, Princeton).
7. Richards, M., Corte-Real, H., Forster, P., Macaulay, V., Wilkinson-Herbots, H., Demaine, A., Papiha, S., Hedges, R., Bandelt, H. J. & Sykes, B. (1996) *Am. J. Hum. Genet.* **59,** 185–203.
8. Underhill, P. A., Shen, P., Lin, A. A., Jin, L., Passarino, G., Yang, W. H., Kauffman, E., Bonne-Tamir, B., Bertranpetit, J., Francalacci, P., *et al.* (2000) *Nat. Genet.* **26,** 358–361.
9. Barbujani, G., Bertorelle, G. & Chikhi, L. (1998) *Am. J. Hum. Genet.* **62,** 488–491.
10. Richards, M. & Sykes B. (1988) *Am. J. Hum. Genet.* **62,** 491–492.
11. Wilson, J. F., Weiss, D. A., Richards, M., Thomas, M. G., Bradman, N. & Goldstein, D.B. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 5078–5083. (First Published April 3, 2001; 10.1073/pnas.071036898)
12. Morris, C. (1985) in *The Prehistory of Orkney*, ed. C. Renfrew (Edinburgh Univ. Press, Edinburgh), pp. 210–243.
13. Palsson, H. & Edwards P. (1981) *Orkneyinga Saga*: *The History of the Earls of Orkney* (Penguin Books, Harmondsworth, U.K.).
14. Marwick H. (1929) *The Orkney Norn* (Oxford Univ. Press, Oxford).
15. Forsyth, K. S. (1997) *Language in Pictland*: *The Case against "Non-Indo-European Pictish"* (De Keltiche Draak, Utrecht, The Netherlands).
16. Hill, E. W., Jobling, M. A. & Bradley, D.G (2000), in *Archaeogenetics: DNA and the Population Prehistory of Europe*, eds. Renfrew, C. & Boyle, K. (McDonald Institute, Cambridge, U.K.), pp. 203–208.
17. Simoni, L., Calafell, F., Bertranpetit, J. & Barbujani, G. (2000), in *Archaeogenetics: DNA and the Population Prehistory of Europe*. eds. Renfrew, C. & Boyle, K. (McDonald Institute, Cambridge, U.K.), pp. 131–138.
18. Saville, A. (2000) in *Neolithic Orkney in its European Context*, ed. Ritchie, A. (McDonald Institute, Cambridge, U.K.), pp. 91–100.
19. Renfrew, C. (1996). *Cretan Stud.* **5,** 1–27.
20. Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellitto, D., Cruciani, F., Kivisild, T., *et al.* (2000) *Am. J. Hum. Genet.* **67,** 1251–1276.
21. Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S., Beckman, L. E., De Benedictis, G., Francalacci, P., Kouvatsi, A., Limborska, S., *et al.* (2000) *Science* **290,** 1155–1159.