

# Y Chromosome Evidence for Anglo-Saxon Mass Migration

Michael E. Weale,<sup>\*1</sup> Deborah A. Weiss,<sup>†1</sup> Rolf F. Jager,<sup>\*‡</sup> Neil Bradman,<sup>\*</sup> and Mark G. Thomas<sup>\*</sup>

<sup>\*</sup>The Centre for Genetic Anthropology, Departments of Biology and Anthropology, University College London, University of London; <sup>†</sup>Department of Anthropology, University of California, Davis; and <sup>‡</sup>Faculteit Biologie, Vrije Universiteit, Amsterdam, The Netherlands

British history contains several periods of major cultural change. It remains controversial as to how much these periods coincided with substantial immigration from continental Europe, even for those that occurred most recently. In this study, we examine genetic data for evidence of male immigration at particular times into Central England and North Wales. To do this, we used 12 biallelic polymorphisms and six microsatellite markers to define high-resolution Y chromosome haplotypes in a sample of 313 males from seven towns located along an east-west transect from East Anglia to North Wales. The Central English towns were genetically very similar, whereas the two North Welsh towns differed significantly both from each other and from the Central English towns. When we compared our data with an additional 177 samples collected in Friesland and Norway, we found that the Central English and Frisian samples were statistically indistinguishable. Using novel population genetic models that incorporate both mass migration and continuous gene flow, we conclude that these striking patterns are best explained by a substantial migration of Anglo-Saxon Y chromosomes into Central England (contributing 50%–100% to the gene pool at that time) but not into North Wales.

## Introduction

Following depopulation during the last glacial maximum and subsequent resettlement by hunter-gatherers ca. 7000 B.C., the history of Britain has been marked by a series of cultural transitions. These include the appearance of sedentary agricultural communities (the Neolithic transition) (ca. 4000 B.C.), the arrival and spread of Late Bronze-Iron Age and Celtic material culture (ca. 1000–100 B.C.), Roman occupation and influence (A.D. 43–410), the rise of Anglo-Saxon language and culture (ca. A.D. 400–800), Viking invasions and influence (ca. . 800–1000), and the Norman Conquest (A.D. 1066) (Kearney 1989; Hunter and Ralson 1998; Davies 1999).

The use of migration as an explanation for cultural transitions has varied greatly over the past 100 years and remains controversial (Clark 1966; Chapman 1997; Burmeister 2000; Shennan 2000). Before the 1960s, archaeological evidence for cultural change (such as changes in pottery type) was often interpreted as *prima facie* evidence for substantial immigration. The processual school or New Archaeology that emerged in the 1960s and 1970s rejected this view, arguing firstly that the adoption of new cultures could occur through trade or by the influx of a small ruling elite with minimal or no impact on the gene pool (the “elite dominance” model of Renfrew [1987]) and secondly that if no positive evidence for migration could be found then explanations based on non-migrational internal forces were

more parsimonious and therefore preferable (Adams 1968; Adams, Van Gerven, and Levy 1978). More recently, this antimigrationist stance has been questioned (Anthony 1990; Härke 1998) and migrational models reconsidered (e.g., Chapman and Hamerow 1997; Burmeister 2000).

These changes in archaeological opinion have had a particular impact on interpretations of cultural transitions in Britain (Clark 1966). Because of its geographical location on the northwestern edge of Europe, migrations and invasions from the continental mainland were once considered to be the obvious explanation for cultural transition (e.g., Hawkes and Hawkes 1942). Today, cultural changes before the Roman invasion, which are for the most part lacking in historical records, are no longer interpreted as automatically implying migration, and for these changes the role of migration remains unresolved. For cultural transitions after the Roman invasion, historical records argue against large migrations coinciding with either Roman occupation or the Norman Conquest, and the prevailing view is therefore that these later events represent examples of elite dominance (Kearney 1989; Davies 1999). Historical and archaeological research argue for some degree of Viking settlement in both East Anglia and the Midlands in the 9th century A.D. but against a substantial displacement of the existing people during this period (Richards 2000).

Today, the most hotly debated of all the British cultural transitions is the role of migration in the relatively sudden and drastic change from Romano-Britain to Anglo-Saxon Britain (Hamerow 1997; Burmeister 2000). This transition was once widely accepted as providing clear evidence for a mass migration from continental Europe and the near-complete replacement of the indigenous population in England (Leeds 1954; Myres 1986). Stories of migration are included in the writings of Gildas (ca. A.D. 540) and Bede (A.D. 731) and hinted at in Anglo-Saxon sagas, such as *Beowulf* (Davies 1999). Archaeological evidence confirmed a rapid rise

<sup>1</sup> Both these authors contributed equally to this work.

Key words: Y chromosome, unique event polymorphisms, microsatellite haplotypes, British population history, genetic anthropology, population genetic models.

Address for correspondence and reprints: Michael Weale, The Centre for Genetic Anthropology, Departments of Biology and Anthropology, University College London, University of London, Darwin Building, Gower Street, London WC1E 6BT, United Kingdom. E-mail: m.weale@ucl.ac.uk.

*Mol. Biol. Evol.* 19(7):1008–1021. 2002

© 2002 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

of continental culture in England (Esmonde-Cleary 1993) and suggested a contemporaneous desertion of continental Germanic settlements. More recently, however, authors have questioned the evidence for large-scale immigration (Crawford 1997; Hamerow 1997) and continental emigration (Näsman 1988; Petersen 1991) and emphasized the continuity of the Romano-British population in England. The sudden change to an Anglo-Saxon culture has been attributed instead to rapid acculturation and indigenous developments, with only a small number of Germanic immigrants (perhaps a male military elite) settling in Britain (Arnold 1984; Hodges 1989; Higham 1992). The contribution of Anglo-Saxon immigration to the modern English gene pool thus remains uncertain.

Genetic data comprise an obvious source of information to help resolve these issues. Previous studies examining biological variation in Britain have identified various patterns of genetic variation. These include studies on blood groups (Bodmer 1993; Mascie-Taylor and Lasker 1996), serum proteins and isoenzymes (Cartwright, Hargreaves, and Sunderland 1977; Mastana et al. 1993), HLA genes (Papiha, Duggan Keen, and Rodger 1985; Bodmer 1993), and multiple classical genetic markers (Falsetti and Sokol 1993; Cavalli-Sforza, Menozzi, and Piazza 1994; Mastana and Sokol 1998), as well as on patterns of disease incidence, such as phenylketonuria (Tyfield, Osborn, and Holton 1997), multiple sclerosis (Poser 1994), skin cancer (Long, Darke, and Marks 1998), and haemochromatosis (Merryweather-Clarke et al. 1997). These data have been interpreted as reflecting historical migrations and settlement patterns, but formal testing of alternative migratory models has not been attempted.

The non-recombining portion of the Y chromosome and the mitochondrial genome are useful sources of data because they provide exceptionally detailed high-resolution haplotypes, allowing fine definition of the underlying gene genealogies. The Y chromosome, which is much larger, is particularly useful because it has many slowly mutating biallelic markers to help resolve genealogical clades as well as rapidly mutating microsatellite markers to aid in the dating of very recent events (Thomas et al. 1998; Kayser et al. 2001). The extra information provided by these high-resolution haplotypes facilitates the fitting of population genetic models. Although the resulting demographic inferences are based on only a single locus, increasing the effects of evolutionary variance derived from chance differences in the genealogy, such systems are still useful because they are the only ones that allow sex-specific demographic inferences to be made.

Previous studies of mtDNA and Y chromosome variation across Europe have reported evidence of Paleolithic and Neolithic expansions reflected in large-scale clines (Torroni et al. 1998; Casalotti et al. 1999; Hill, Jobling, and Bradley 2000; Malaspina et al. 2000; Richards et al. 2000; Rosser et al. 2000; Semino et al. 2000; Simoni et al. 2000), but these studies did not consider the effects of historical migrations on more local patterns of genetic variation. Helgason et al. (2000) ex-

amined Y chromosome and mtDNA variation in the modern Icelandic population to assess the relative proportions of Scandinavian and Celtic ancestry stemming from historical migrations, whereas Wilson et al. (2001) compared Y chromosome, X chromosome, and mtDNA variation in eight population samples (including the Llangefni, Norway, and Friesland samples reported here) to investigate genetic changes associated with cultural transitions in North Wales and Orkney, two areas at the fringes of the British Isles. Through a comparison of signature haplotypes, Wilson et al. (2001) found evidence for Celtic male ancestry in the North Welsh and/or both Celtic and Scandinavian (Viking) male ancestry in the modern Orcadian population. Further comparisons of these British samples with Basque data suggested that the male Celtic genetic component was Paleolithic in origin, and therefore, that subsequent cultural transitions in North Wales were not associated with substantial incoming male gene flow. However, the study of Wilson et al. did not directly address the effects of cultural transitions in other areas of Britain.

This study is the first to analyze data from an east-west transect across Central England and North Wales to evaluate evidence of male population migration under a wide range of flexible population genetic models. Samples were collected in seven towns along this transect, and a combination of slowly evolving biallelic markers (so-called Unique Event Polymorphisms or UEPs) and rapidly evolving microsatellites on the Y chromosome were typed to look for evidence of local or small-scale genetic transitions. We compared the data with samples from Friesland and Norway to look for evidence of male immigration from the continent. In addition to comparing signature haplotypes among population samples, we applied novel model-based methods to make inferences about both the possible timing and extent of male continental migration into Central England.

## Materials and Methods

### Population Samples and Genotyping

Buccal swabs were collected from 313 males in the British towns of North Walsham, Fakenham, Bourne, Southwell, Ashbourne, Abergele, and Llangefni (fig. 1). These towns were selected because they lie approximately 50 km apart along an east-west transect of Britain and are long established market towns (mentioned in the Domesday Book of A.D. 1086 with current populations of 5,000–10,000) that are less likely to be influenced by recent migration than large cities (Pooley and Turnbull 1998). We apply the labels “East Anglia” to the North Walsham and Fakenham samples, “Midlands” to the Bourne, Southwell, and Ashbourne samples, “North Wales” to the Abergele and Llangefni samples, and “Central England” to the combined Midlands and East Anglia samples, although these labels are for convenience and only designate the general geographical area of these samples. Samples were acquired if both the donor and the donor’s paternal grandfather were born within 30 km of one of these market towns. For

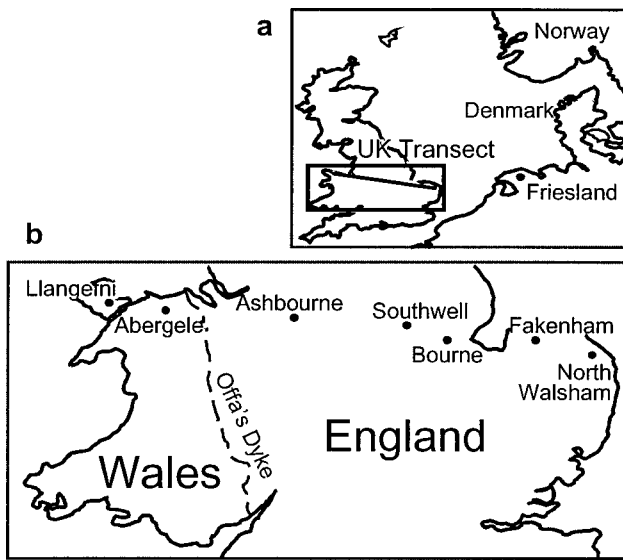


FIG. 1.—(a) Map of Britain and neighboring area. (b) Enlargement showing towns within Britain along east-west transect.

comparison, we also collected DNA samples from 94 males in Friesland (northern Netherlands) and 83 males in Norway, two nearby locations with different roles in Britain's immigration history. Friesland is thought to be one of the source locations for Anglo-Saxon immigration both because of its geographical location and because Frisian is considered to be the closest extant language to Old English (Nielsen 1985). Norway represents one source of the Viking invaders. Samples were collected anonymously, and informed consent was obtained from all individuals before samples were taken.

Standard phenol-chloroform DNA extractions were performed. Six microsatellites (DYS19, DYS388, DYS390, DYS391, DYS392, DYS393) and 11 of the biallelic UEP markers (92R7, M9, M13, M17, M20, SRY+465, SRY4064, SRY10831, sY81, Tat, YAP) were typed and analyzed, as described by Thomas, Bradman, and Flinn (1999). One additional UEP marker, 12f2, was typed, as described by Rosser et al. (2000). Microsatellite repeat sizes were assigned according to the nomenclature of Kayser et al. (1997). Haplogroups were defined by the 12 UEP markers according to a nomenclature modified from Rosser et al. (2000) and Weale et al. (2001) and are presented in figure 2. The correspondence between this nomenclature and that proposed by the Y Chromosome Consortium (2002) is as follows: hg1 = P\*(xR1a), hg2 = BR\*(xDE,JR), hg3 = R1a1, hg4 = DE\*(xE), hg7 = A3b2, hg8 = E3a, hg9 = J, hg16 = N3, hg20 = O2b, hg21 = E\*(xE3a), hg26 = K\*(xL,N3,O2b,P), hg28 = L, hg29 = R1a\*, hg37 = Y\*(xBR,A3b2).

#### Statistical and Population Genetic Analysis

We tested for significant population differentiation using the Exact Test for Population Differentiation of Raymond and Rousset (1995). Unbiased genetic diversity,  $h$ , and its standard error were calculated using the formulae given by Nei (1987, pp. 178–180). We quan-

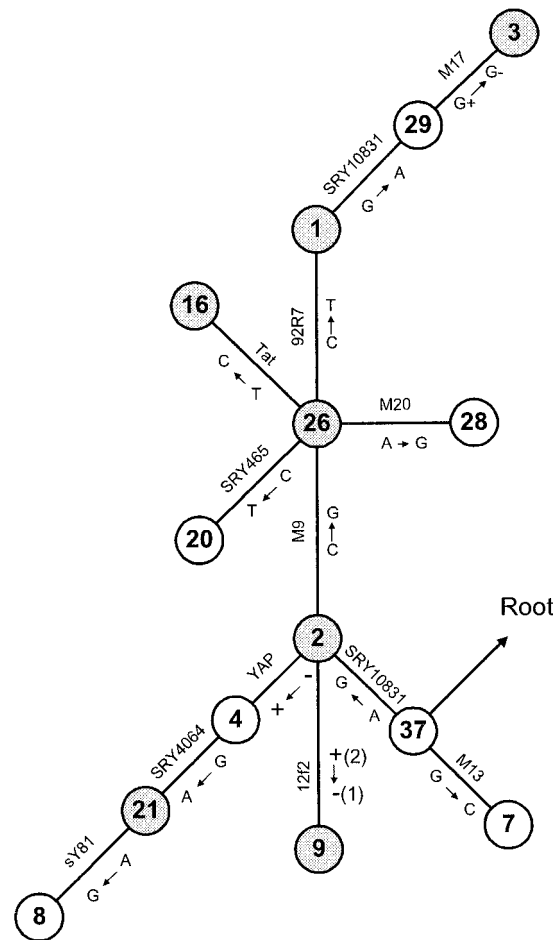


FIG. 2.—Y chromosome haplogroup network defined by the 11 UEP markers used in this study following a nomenclature modified from Rosser et al. (2000) and Weale et al. (2001). The root has been deduced by comparison with other great ape species (Underhill et al. [2000]; P. Underhill, personal communication for position of SRY10831). Shading indicates haplogroups observed in this study.

tified patterns of genetic differentiation using the genetic distance measures  $F_{ST}$  and  $R_{ST}$  estimated from analysis of molecular variance (AMOVA)  $\Phi_{ST}$  values (Reynolds, Weir, and Cockerham 1983; Michalakis and Excoffier 1996). Confidence intervals for these statistics were constructed using bootstrap estimates of standard errors based on resampling haplotypes according to observed population frequencies. We visualized patterns of genetic differentiation using principal coordinates analysis performed on a similarity matrix calculated as  $1 - R_{ST}$ . Values along the main diagonal of the similarity matrix, representing the similarity of each population sample to itself, were calculated from the estimated genetic distance between two copies of the same population sample (for  $\Phi_{ST}$ -based  $F_{ST}$  and  $R_{ST}$  values, the resulting self-similarity values simplify to  $n/(n - 1)$ , where  $n$  is the sample size).

We explored population genetic models that could explain our data using two methods of inference. The first method involved full-likelihood Bayesian inference of genetic and demographic parameters under population splitting and growth using the BATWING program

(URL: <http://www.maths.abdn.ac.uk/~ijw>), extended from the Markov chain Monte Carlo (MCMC) algorithm presented by Wilson and Balding (1998). Microsatellite mutation likelihoods were calculated using an unbounded symmetric stepwise mutation model. The priors chosen for major BATWING parameters are summarized in table 1. Locus-specific priors for the mutation rate per generation were based on observed mutations for these loci, as reported in Heyer et al. (1997), Bianchi et al. (1998), and Kayser et al. (2000), combined with a standard exponential pre-prior. As a precautionary measure, DYS388 was excluded from BATWING analysis (and also from the Monte Carlo likelihood method described later) because no published data on observed meioses are available for this locus, and therefore, no direct verification exists of stepwise mutation behavior. Unique-event mutations inferred from binary marker data were used to condition the possible trees but otherwise did not contribute to the likelihood. Population splitting was modeled under strict fission with no subsequent background migration. Population growth was modeled as an exponential from an initially constant effective population size. Weakly informative priors were given to other parameters to aid in the convergence of the MCMC process. The prior for initial effective population size covers the values commonly assumed for the global Y chromosome effective population size as well as lower values to compensate for this being a regional sample and for representing the effective size before growth. The prior for population growth rate per generation is very flat and gives support to a very wide range of possible values (extending beyond the 2.5% and 97.5% quantiles) to reflect the uncertainty in relating growth in effective population size to growth in real (census) population size. All other parameters, such as the start-of-growth date and the population split dates, were given flat, uninformative priors.

BATWING suffers from the drawback that inference is based on a model of population differentiation that involves population splitting only, with no subsequent migration. In contrast, the MIGRATE method of Beerli and Felsenstein (2001) and the GENETREE method of Bahlo and Griffiths (2000), both also based on coalescent modeling, allow constant background migration but do not allow population splitting or mass migration at a single point in time. Nielsen and Wakeley (2001) have presented a coalescent-based MCMC inference method that allows both population splitting and migration, but it is not suitable for microsatellite data and does not allow for population growth or for additional mass migration events that may occur at specific points in time. We therefore developed an alternative inference method that allowed us to explore more flexible models under a range of historical scenarios involving both background and mass migration in the presence of population splitting and growth. This method proceeds through Monte Carlo likelihood estimation based on summary statistics describing genetic differentiation. In what follows, we use the term background migration to refer to continuous gene flow occurring in each generation and mass migration to refer to a single

**Table 1**  
**Prior Distributions of Major BATWING Parameters and Posterior Values Based on a Splitting Model Between Central England and Friesland.**

PARAMETER	PRIOR	PRIOR QUANTILES			POSTERIOR QUANTILES		
		2.50%	50%	97.50%	2.50%	50%	97.50%
DYS19 mutation rate	Gamma(3,1459)	0.00042	0.0018	0.0050	0.0013	0.0025	0.0043
DYS390 mutation rate	Gamma(5,929)	0.0017	0.0050	0.011	0.0017	0.0033	0.0056
DYS391 mutation rate	Gamma(3,878)	0.00070	0.0030	0.0082	0.00066	0.0014	0.0026
DYS392 mutation rate	Gamma(2,878)	0.00028	0.0019	0.0063	0.00058	0.0012	0.0022
DYS393 mutation rate	Gamma(1,878)	0.000029	0.00079	0.0042	0.00067	0.0013	0.0025
Initial effective population size	Gamma(1,1,0.0001)	371	7,903	38,960	125	356	987
Growth rate	Gamma(1,01,1)	0.026	0.703	3.71	0.029	0.08	2.02
Time growth starts (in generations)	Uniform(0,TMRCA <sup>a</sup> )	—	—	—	47	91	192

<sup>a</sup> TMRCA: time to most recent common ancestor.

**Table 2**  
**Haplogroup Frequencies and Genetic Diversities in Seven British Towns, Friesland and Norway**

HAPLOGROUP	NORTH WALES		MIDLANDS			EAST ANGLIA		FRIESLAND ( <i>n</i> = 94)	NORWAY ( <i>n</i> = 83)
	Llangefni ( <i>n</i> = 80)	Abergele ( <i>n</i> = 18)	Ashbourne ( <i>n</i> = 54)	Southwell ( <i>n</i> = 70)	Bourne ( <i>n</i> = 12)	Fakenham ( <i>n</i> = 53)	N. Walsham ( <i>n</i> = 26)		
hg1.....	0.888	0.556	0.648	0.643	0.667	0.566	0.577	0.553	0.265
hg2.....	0.038	0.056	0.222	0.186	0.333	0.415	0.308	0.340	0.446
hg3.....	0.013	—	0.037	0.057	—	—	0.038	0.074	0.217
hg9.....	0.013	—	0.037	0.057	—	—	0.038	0.011	0.024
hg16.....	—	—	—	—	—	—	—	—	0.036
hg21.....	0.038	0.389	0.056	0.057	—	0.019	0.038	0.021	0.012
hg26.....	0.013	—	—	—	—	—	—	—	—
Genetic diversity, <i>h</i> ..	0.212 (±0.0607)	0.569 (±0.0707)	0.535 (±0.0660)	0.550 (±0.0606)	0.485 (±0.1059)	0.517 (±0.0301)	0.591 (±0.0767)	0.578 (±0.0339)	0.690 (±0.0290)

NOTE.—Dash indicates haplogroup was not observed. Genetic diversities (based on haplogroup frequencies) are given ± standard error.

large gene flow event. A program was written (available from M.E.W.) to simulate the coalescent under growth from an initially constant effective population size, allowing a single split  $T_S$  generations ago of the parent population into two descendent populations, A and B, and subsequent background migration at a rate  $m$ , where a proportion  $m$  migrates from population A to B and simultaneously from B to A in each generation. A further extension to the program allowed for a single unidirectional migration event from B to A at a time  $T_F$  generations ago, such that lineages in A immediately after this event had a probability  $F$  of having just migrated from B. Final sample sizes in each population were set to match those obtained in our study, and microsatellite repeat sizes in these samples were simulated using the unbounded symmetric stepwise mutation model. Parameters for population growth and mutation rates were set as fixed constants based on BATWING posterior modal values and were validated firstly by checking that the simulated within-population microsatellite variances were similar to observed values and secondly through comparison with BATWING results under scenarios involving no background migration. These parameters were fixed as follows: initial effective population size = 300; start of growth = 80 generations before present (BP); growth rate = 0.06; ratio of effective sizes of population A to population B = 1:1; mutation rates—DYS19 = 0.0023, DYS390 = 0.0030, DYS391 = 0.0012, DYS392 = 0.0011, DYS393 = 0.0012. The sensitivity of inferences to the initial effective population size was assessed by setting its value an order of magnitude higher (to 3000). Inferences considered here were little affected, and unaffected when  $(\delta\mu)^2$  was used as a summary statistic (see below).

Confidence intervals (95%) for the remaining parameters  $T_S$ ,  $m$ ,  $T_F$ , and  $F$  describing population differentiation were estimated by setting constant values for all but one of the parameters and finding the upper and lower limits for the remaining parameter of interest. Preset values for the split date  $T_S$  were chosen to reflect one of the three scenarios: (1) Island model ( $T_S = \infty$ ), (2) Neolithic ( $T_S = 240$  generations BP or 6,000 years BP assuming 25 years per generation), and (3) Anglo-Saxon ( $T_S = 60$  generations BP or 1,500 years BP assuming 25 years per generation). Preset values for the

background migration rate  $m$  were set at one of two extremes: (1) an implausibly low value of  $m = 0$ , and (2) an implausibly high value of  $m = 0.1\%$ . The latter value is estimated from migration statistics to and from the European Economic Area as a whole over the past 25 years (source: <http://www.homeoffice.gov.uk/rds/index.htm>). We regard this as a figure well in excess of realistic values for continuous background migration both because it is based on figures for the whole of Europe (i.e., not just Friesland) and because we expect background migration in recent times to be very high as a result of modern trends in communication and travel. In models that included the parameters  $F$  and  $T_F$ ,  $F$  was always the parameter of interest and  $T_F$  was set under an Anglo-Saxon scenario ( $T_F = 60$ ).

Confidence limits (95%) for the parameter of interest were found using  $R_{ST}$  as a summary statistic of population differentiation, such that 2.5% of the simulated  $R_{ST}$  values were equal to or more extreme than the observed  $R_{ST}$  value when the parameter was set at one of the two limits, based on 10,000 Monte Carlo iterations of the coalescent simulation. Confidence limits were determined to an accuracy of the least significant digit stated. We also obtained confidence limits using  $(\delta\mu)^2$  rather than  $R_{ST}$  as the summary statistic (Goldstein et al. 1995) and found these to be similar although generally wider than those obtained using  $R_{ST}$ .

## Results

### Patterns of Genetic Differentiation

The UEP markers defined seven haplogroups (hereafter abbreviated to hg) (fig. 2 and table 2), whereas the UEP plus microsatellite markers defined 150 haplotypes (table 3). Hg1 and hg2 predominate in the samples, with hg1 occurring at the highest frequencies in North Wales and hg2 occurring at the highest frequencies in East Anglia, Friesland, and Norway. Hg2 is almost completely absent from North Wales. Except for the high frequency of hg21 in Abergele, the observed haplogroup frequencies match their geographic distribution in Europe (Rosser et al. 2000). The high frequency of hg21 in Abergele may indicate high genetic isolation and drift in some parts of North Wales, a hypothesis that is consistent both with the tight clustering of Abergele haplotypes within

**Table 3**  
**UEP + Microsatellite Haplotype Frequencies and Genetic Diversities in Seven British Towns, Friesland, and Norway**

Ht #	Microsatellite Haplotype <sup>a</sup>	Llan. (n = 80)	Aber. (n = 18)	Ash. (n = 54)	South. (n = 70)	Bour. (n = 12)	Fak. (n = 53)	N. Wal. (n = 26)	Fries. (n = 94)	Nor. (n = 83)	Total Counts
<b>Haplogroup I</b>											
1	14-12-24-11-13-13	0.275	0.056	0.130	0.186	—	0.113	0.231	0.128	0.060	72
2	14-12-23-11-13-13	0.063	—	0.111	0.086	0.083	0.094	0.077	0.170	0.024	43
3	14-12-24-10-13-13	0.025	—	0.130	0.129	—	0.075	—	0.021	0.036	27
4	14-12-25-11-13-13	0.113	0.111	0.037	0.014	0.083	0.038	—	—	0.012	18
5	14-12-23-10-13-13	0.025	—	0.037	0.014	0.167	0.019	0.077	0.021	—	12
6	14-12-24-11-13-14	0.088	0.056	—	—	—	—	—	0.011	0.012	10
7	15-12-24-11-13-13	0.063	0.056	0.037	—	—	—	—	—	0.012	9
8	15-12-24-10-13-13	—	—	0.019	—	—	—	0.038	0.032	0.012	6
9	14-12-24-12-13-13	0.025	—	—	—	0.167	—	—	—	0.012	5
10	15-12-23-11-13-13	—	0.056	—	0.029	—	0.019	—	—	—	4
11	14-12-25-11-13-12	—	—	0.056	0.014	—	—	—	—	—	4
12	14-12-25-10-13-13	0.013	—	—	—	—	0.057	—	—	—	4
13	14-12-24-11-12-13	—	0.056	—	0.014	—	0.038	—	—	—	4
14	14-12-24-11-11-13	0.025	—	—	—	—	—	—	—	—	4
15	15-12-25-11-13-13	0.038	0.056	—	—	—	—	—	—	—	4
16	14-13-23-11-13-13	0.025	—	0.019	—	—	—	—	—	—	3
17	14-12-25-11-14-13	0.013	—	—	—	—	0.038	—	—	—	3
18	14-12-24-11-13-12	0.013	—	—	0.014	—	—	—	0.011	—	3
19	14-12-23-11-13-14	—	—	—	0.014	—	0.019	—	0.021	0.012	3
20	14-12-24-10-13-12	—	—	—	—	0.083	—	0.077	—	—	3
21	14-11-23-11-13-13	—	—	—	0.014	—	—	—	—	—	2
22	15-12-23-10-13-13	—	—	—	—	—	—	—	—	—	2
23	14-13-24-11-13-13	—	—	—	—	—	—	—	0.021	—	2
24	14-12-26-11-14-14	0.025	—	—	—	—	—	—	—	—	2
25	14-12-24-10-13-14	—	—	0.019	—	—	—	0.038	—	—	2
26	14-12-26-11-14-13	—	0.056	—	—	—	—	—	0.011	—	2
27	14-12-25-12-13-13	0.025	—	—	—	0.083	—	—	—	—	2
28	12-12-23-11-13-13	—	—	—	0.014	—	—	—	—	—	2
29	14-12-22-11-13-13	—	—	—	0.014	—	0.019	—	—	—	2
30	16-12-25-11-13-13	—	—	—	—	—	—	—	—	—	2
31	11-12-24-11-13-13	0.013	—	—	0.014	—	—	—	—	0.012	1
32	13-12-24-10-13-14	—	—	—	—	—	—	—	—	—	1
33	14-12-24-11-14-12	—	—	—	—	—	—	—	0.011	—	1
34	14-12-24-11-14-13	0.013	—	—	—	—	—	—	—	—	1
35	14-12-24-11-14-14	—	0.056	—	—	—	—	—	—	—	1
36	14-12-21-11-13-13	—	—	—	—	—	—	—	—	—	1
37	14-12-24-12-13-14	—	—	—	—	—	—	—	0.011	—	1
38	14-12-22-12-13-13	—	—	—	—	—	—	—	0.011	—	1
39	14-12-23-11-14-13	—	—	—	—	—	—	0.038	—	0.012	1
40	12-12-24-11-13-13	—	—	—	0.014	—	—	—	—	—	1
41	14-12-25-11-13-14	0.013	—	—	0.014	—	—	—	—	—	1
42	13-12-23-10-13-13	—	—	—	—	—	0.019	—	—	—	1
43	14-12-23-10-14-13	—	—	—	—	—	—	—	—	—	1
44	14-12-26-10-13-13	—	—	—	—	—	—	—	0.011	—	1
45	14-12-24-10-14-13	—	—	0.019	—	—	—	—	—	—	1
46	14-12-24-10-13-15	—	—	—	—	—	—	—	0.011	—	1
47	14-12-23-11-13-12	—	—	—	—	—	—	—	—	0.012	1

**Table 3**  
Continued.

Ht #	Microsatellite Haplotype <sup>a</sup>	Llan. (n = 80)	Aber. (n = 18)	Ash. (n = 54)	South. (n = 70)	Bour. (n = 12)	Fak. (n = 53)	N.Wal. (n = 26)	Fries. (n = 94)	Nor. (n = 83)	Total Counts
48	14-13-24-10-12-13	—	—	0.019	—	—	—	—	—	—	1
49	12-12-24-10-13-13	—	—	—	0.014	—	—	—	—	—	1
50	14-14-24-10-13-12	—	—	—	—	—	—	—	0.011	—	1
51	14-14-25-10-13-13	—	—	—	—	—	—	—	0.011	0.012	1
52	14-12-23-12-13-13	—	—	—	—	—	—	—	0.011	—	1
53	14-12-22-11-13-12	—	—	—	—	—	—	—	—	0.012	1
54	15-12-23-12-13-13	—	—	—	—	—	0.019	—	—	—	1
55	14-12-24-10-12-13	—	—	—	—	—	—	—	0.011	—	1
56	13-12-24-10-14-13	—	—	—	—	—	—	—	0.011	—	1
57	15-12-24-11-13-14	—	—	0.019	—	—	—	—	—	—	1
58	15-12-24-11-15-13	—	—	—	0.014	—	—	—	—	—	1
59	14-12-23-11-11-13	—	—	—	—	—	—	—	0.011	—	1
Haplogroup 2											
60	14-14-22-10-11-13	—	—	0.093	0.029	0.083	0.094	0.077	0.053	0.084	27
61	14-14-23-10-11-13	—	—	0.019	0.014	0.083	0.113	—	0.032	0.133	23
62	15-14-23-10-11-13	—	—	—	—	—	0.019	—	0.021	0.072	9
63	15-14-22-10-11-13	—	—	—	—	0.083	—	—	0.011	0.048	6
64	15-13-23-10-12-14	—	—	—	0.014	—	—	0.077	0.011	0.024	6
65	15-13-23-10-12-15	—	—	—	0.029	—	0.019	0.077	0.021	—	5
66	15-15-23-10-11-13	—	—	0.019	—	—	0.019	—	0.011	—	4
67	16-14-22-10-11-13	—	—	—	0.014	—	—	—	—	0.012	3
68	14-14-22-10-11-14	—	—	—	—	—	0.019	—	0.011	0.012	3
69	17-13-25-11-11-13	0.025	—	—	—	—	—	—	—	0.012	2
70	16-13-24-10-11-13	—	—	—	—	—	0.038	—	—	—	2
71	16-13-23-10-12-15	—	—	—	—	—	—	—	0.021	—	2
72	14-14-23-11-11-13	—	—	—	—	—	—	—	0.011	0.012	2
73	16-13-23-10-12-13	—	—	—	0.014	—	—	—	0.011	—	2
74	14-14-22-11-11-12	—	—	0.019	—	—	0.019	—	—	—	2
75	15-14-22-11-11-13	—	—	0.019	—	—	0.019	—	—	—	2
76	15-13-24-10-12-15	—	—	—	—	0.083	—	—	—	—	2
77	14-14-22-11-11-13	0.013	—	—	—	—	—	—	—	0.012	1
78	15-12-22-10-10-13	—	—	—	—	—	0.019	—	—	—	1
79	15-12-22-11-11-14	—	—	—	0.014	—	—	—	—	—	1
80	15-13-22-10-11-13	—	—	—	—	—	—	—	0.011	—	1
81	15-13-22-10-11-14	—	—	—	—	—	—	0.038	—	—	1
82	15-13-23-09-12-14	—	—	—	—	—	—	0.038	—	—	1
83	15-13-23-10-12-13	—	—	—	—	—	—	—	0.011	—	1
84	14-14-22-09-11-13	—	—	—	—	—	—	—	0.011	—	1
85	13-14-22-10-11-13	—	0.056	—	—	—	—	—	—	—	1
86	15-13-23-11-11-13	—	—	0.019	—	—	—	—	0.011	—	1
87	15-13-23-11-12-14	—	—	—	—	—	—	—	—	—	1
88	15-12-21-10-11-14	—	—	—	0.014	—	—	—	—	—	1
89	15-14-22-09-11-13	—	—	—	—	—	0.019	—	—	—	1
90	14-13-22-10-11-15	—	—	—	—	—	—	—	0.011	—	1
91	14-16-22-10-11-14	—	—	—	—	—	—	—	0.011	—	1
92	14-12-22-10-11-13	—	—	0.019	—	—	—	—	—	—	1

Table 3  
Continued.

Ht #	Microsatellite Haplotype <sup>a</sup>	Llan. (n = 80)	Aber. (n = 18)	Ash. (n = 54)	South. (n = 70)	Bour. (n = 12)	Fak. (n = 53)	N. Wal. (n = 26)	Fries. (n = 94)	Nor. (n = 83)	Total Counts
93	15-14-24-10-11-13	—	—	—	—	—	0.019	—	—	—	1
94	15-14-24-10-12-15	—	—	—	—	—	—	—	0.011	—	1
95	15-15-22-10-11-13	—	—	—	—	—	—	—	0.011	0.012	1
96	15-15-22-10-12-13	—	—	—	—	—	—	—	0.011	—	1
97	14-16-22-09-11-14	—	—	—	0.014	—	—	—	0.011	—	1
98	14-15-22-10-11-14	—	—	—	—	—	—	—	0.011	—	1
99	14-15-22-10-11-13	—	—	—	—	—	—	—	0.011	—	1
100	16-13-23-10-13-13	—	—	—	0.014	—	—	—	—	—	1
101	14-12-21-10-11-14	—	—	0.019	—	—	—	—	—	—	1
102	16-13-24-10-12-14	—	—	—	0.014	—	—	—	—	—	1
103	14-14-22-10-12-13	—	—	—	—	—	—	—	0.011	—	1
104	17-13-22-10-11-14	—	—	—	—	—	—	—	—	0.012	1
105	17-13-23-11-11-14	—	—	—	—	—	—	—	0.011	—	1
106	17-13-24-10-12-15	—	—	—	—	—	—	—	—	0.012	1
Haplogroup 3											
107	16-12-25-11-11-13	—	—	0.019	0.014	—	—	—	—	0.036	5
108	15-12-25-10-11-13	—	—	—	—	—	—	—	—	0.048	4
109	15-12-25-11-11-13	—	—	—	—	—	—	—	0.011	0.024	3
110	16-12-25-10-11-13	—	—	0.019	—	—	—	—	—	0.012	2
111	16-12-24-10-11-13	—	—	—	0.014	—	—	—	—	0.012	2
112	15-12-26-11-11-13	—	—	—	—	—	—	—	—	0.024	2
113	17-12-25-10-11-13	—	—	—	0.014	—	—	—	—	—	1
114	15-12-25-11-11-14	—	—	—	—	—	—	—	0.011	—	1
115	15-10-26-10-11-13	—	—	—	—	—	—	—	—	0.012	1
116	15-14-25-11-11-13	—	—	—	—	—	—	—	0.011	—	1
117	16-09-25-10-11-13	—	—	—	—	—	—	—	0.011	—	1
118	16-10-25-10-11-13	—	—	—	—	—	—	—	—	—	1
119	16-10-26-10-11-13	0.013	—	—	—	—	—	—	0.011	—	1
120	15-10-25-10-11-13	—	—	—	—	—	—	—	—	—	1
121	15-12-24-11-11-13	—	—	—	—	—	—	—	—	0.012	1
122	16-12-25-10-11-14	—	—	—	0.014	—	—	0.038	—	—	1
123	14-12-25-10-11-13	—	—	—	—	—	—	—	—	—	1
124	16-12-26-11-11-13	—	—	—	—	—	—	—	0.011	—	1
125	17-10-24-10-11-13	—	—	—	—	—	—	—	—	0.012	1
126	17-12-23-10-11-13	—	—	—	—	—	—	—	0.011	—	1
127	17-12-23-11-11-14	—	—	—	—	—	—	—	—	0.012	1
Haplogroup 9											
128	14-16-23-10-11-12	—	—	—	0.029	—	—	—	—	0.012	3
129	16-15-24-10-11-12	0.013	—	—	—	—	—	—	—	—	1
130	14-15-26-10-11-12	—	—	—	—	—	—	0.038	—	—	1
131	14-14-23-10-11-12	—	—	—	0.014	—	—	—	—	—	1
132	14-16-24-10-11-12	—	—	—	0.014	—	—	—	—	—	1
133	15-15-23-10-11-12	—	—	0.019	—	—	—	—	—	—	1
134	15-15-24-10-11-12	—	—	0.019	—	—	—	—	—	—	1
135	15-17-23-09-11-13	—	—	—	—	—	—	—	—	0.012	1
136	16-15-23-11-11-13	—	—	—	—	—	—	—	0.011	—	1



**Table 3**  
Continued.

Ht #	Microsatellite Haplotype <sup>a</sup>	Llan. (n = 80)	Aber. (n = 18)	Ash. (n = 54)	South. (n = 70)	Bour. (n = 12)	Fak. (n = 53)	N.Wal. (n = 26)	Fries. (n = 94)	Nor. (n = 83)	Total Counts
Haplogroup 16											
137	15-12-23-11-14-14	—	—	—	—	—	—	—	—	0.012	1
138	14-12-23-11-14-14	—	—	—	—	—	—	—	—	0.012	1
139	14-12-24-11-14-14	—	—	—	—	—	—	—	—	0.012	1
Haplogroup 21											
140	13-12-24-11-11-13	0.025	0.222	—	—	—	—	—	—	0.012	7
141	13-12-24-10-11-13	—	0.056	0.019	0.029	—	—	—	0.021	—	6
142	13-12-25-10-11-13	—	—	—	0.014	—	0.019	—	—	—	2
Haplogroup 26											
143	14-12-24-10-11-13	—	—	—	0.014	—	—	—	—	—	1
144	13-12-23-10-11-13	—	—	0.019	—	—	—	—	—	—	1
145	13-12-24-10-11-14	—	0.056	—	—	—	—	—	—	—	1
146	13-12-22-10-11-13	—	—	0.019	—	—	—	—	—	—	1
147	13-12-24-12-11-13	0.013	—	—	—	—	—	—	—	—	1
148	13-12-23-10-12-13	—	—	—	—	—	—	0.038	—	—	1
149	14-12-23-10-11-13	—	0.056	—	—	—	—	—	—	—	1
Haplogroup 26											
150	13-12-23-10-12-13	0.013	—	—	—	—	—	—	—	—	1
Genetic diversity, <i>h</i>		0.899 ± 0.0239	0.954 ± 0.0394	0.949 ± 0.0145	0.945 ± 0.0169	0.970 ± 0.0443	0.954 ± 0.0123	0.935 ± 0.0327	0.952 ± 0.0134	0.964 ± 0.0096	

NOTE.—Microsatellite haplotypes within haplogroups are arranged in decreasing order of total frequency. Dash indicates haplotype was not observed. Genetic diversities (based on haplotype frequencies) are given ± standard error.

<sup>a</sup> Microsatellite haplotypes are defined by a string of six numbers giving the repeat size at loci DYS19, DYS388, DYS390, DYS391, DYS392, and DYS393, respectively.

**Table 4**  
**Genetic Distances and *P*-values**

	Llangefni	Abergele	Ashbourne	Southwell	Bourne	Fakenham	N. Walsham	Friesland	Norway
Llangefni . . . .		0.008*	<0.001*	<0.001*	<0.001*	<0.001*	<0.001*	<0.001*	<0.001*
Abergele . . . . .	0.123		0.002*	<0.001*	0.048*	<0.001*	<0.001*	<0.001*	0.002*
Ashbourne . . . .	0.168	0.081		0.707	0.093	0.188	0.080	0.181	<0.001*
Southwell . . . .	0.126	0.063	-0.013		0.082	0.233	0.141	0.115	<0.001*
Bourne . . . . .	0.202	0.082	-0.034	-0.031		0.117	0.061	0.172	0.237
Fakenham . . . .	0.234	0.112	-0.004	0.006	-0.030		0.032*	0.153	0.037*
N. Walsham . . .	0.174	0.082	-0.024	-0.024	-0.052	-0.017		0.551	0.008*
Friesland . . . .	0.184	0.097	-0.001	0.004	-0.031	-0.007	-0.017		<0.001*
Norway . . . . .	0.310	0.175	0.081	0.093	0.067	0.038	0.065	0.044	

NOTE.—Upper right triangle reports *P*-values of pairwise exact tests for population differentiation (Raymond and Rousset 1995) based on UEP + microsatellite haplotype frequencies. Asterisk indicates a significant value ( $P < 0.05$ ). Lower left triangle reports AMOVA-based  $R_{ST}$  values (Michalakis and Excoffier 1996) based on microsatellite haplotypes defined from six loci: DYS19, DYS388, DYS390, DYS391, DYS392, and DYS393.

this haplogroup and with the lower genetic diversity found in Llangefni.

Several patterns emerge from analyses of genetic distance and population differentiation (table 4 and fig. 3). Firstly, little genetic differentiation exists among the Central English towns. The only significant difference in haplotype frequencies occurs between the neighboring towns of Fakenham and North Walsham ( $P = 0.032$  for individual pairwise comparison or  $P = 0.035$  for combined test of all English towns). Although more hg1 and fewer hg2 chromosomes were observed in the Midlands than in East Anglia (table 2), Mantel tests on the genetic versus geographic distance correlation and linear regression on hg1 and hg2 frequencies reveal no significant within-England clinal patterns.

Secondly, in contrast to the Central English towns, the two North Welsh towns show highly significant differences, both from each other and from the five Central English towns. Llangefni has a very high frequency of hg1, tightly clustered around the modal haplotype (haplotype #1 in table 3; haplotype 1.15 in Wilson et al. 2001) that has been found at high frequency in other Atlantic populations, including Ireland (Hill, Jobling, and Bradley 2000) and the Orkneys (Wilson et al. 2001).

Both Llangefni and Abergele have very low frequencies of hg2 and dispersed haplotypes within this haplogroup, a pattern which is consistent with these haplotypes entering the North Welsh populations through separate, infrequent admixture events.

Thirdly, no significant differences in haplotype frequencies exist between Friesland and any of the Central English towns. Comparisons between Norway and the Central English towns, on the other hand, are all significant, apart from Bourne ( $P = 0.237$ ), which may be explained by the small number of samples collected from this town ( $n = 12$ ). Furthermore, bootstrap tests on  $R_{ST}$  values revealed that the Central English (all five towns combined) are significantly more closely related to the Frisians than they are to the North Welsh (Llangefni:  $P < 0.001$ ; Abergele:  $P = 0.046$ ) or to the Norwegians ( $P = 0.005$ ). Both Friesland and Norway are significantly different from the North Welsh towns. Similar results were obtained using  $F_{ST}$  values based on haplogroup frequencies, but tests on  $F_{ST}$  values based on haplotype frequencies were not significant because of the large number of singletons at this level. Taken together, these results suggest considerable male-line commonality between Central England and Friesland.

Wilson et al. (2001) identified two haplotypes—(1) 2.47 (haplotype #60 in our table 3), and (2) 3.65 (haplotype #107 in our table 3)—that proved useful in inferring a Viking contribution to the Orcadian gene pool, although they noted that it might not be possible to distinguish 2.47 from an Anglo-Saxon contribution in other parts of Britain. We compared the frequencies found in the Central English, Frisian, and Norwegian samples of (1) the 2.47 and 3.65 haplotypes on their own, (2) these two haplotypes plus their one-step mutational neighbors, and (3) these two haplotypes plus their one-step networks (defined as all haplotypes within a sample connected to the named haplotype by a series of one-step mutations via observed intermediate haplotypes). As suggested by the results in the previous paragraph, in each case the frequency distribution in Central England more closely matched that in Friesland than that in Norway. Thus, neither of these two haplotypes provided any positive evidence of a (Norwegian) Viking contribution to the Central English gene pool that could not be ex-

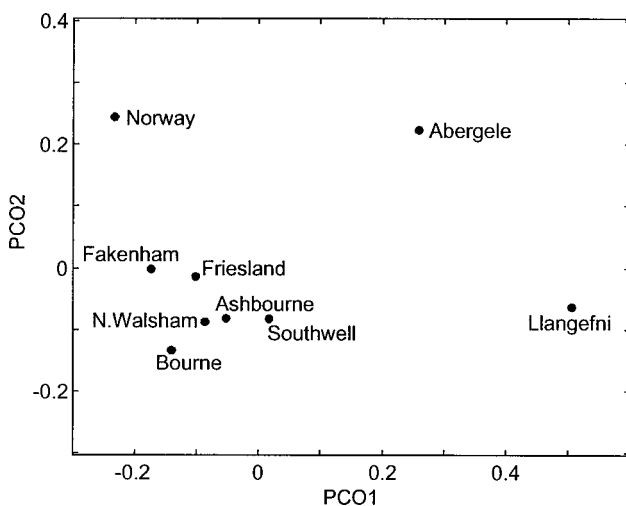


FIG. 3.—Principal coordinates plot of  $R_{ST}$  values reported in table 4. The first and second projections (PCO1 and PCO2) explain 59% and 20% of the variation, respectively.

plained by a substantial contribution originating in Friesland only.

### Population Genetic Models

We explored various population genetic models (see *Materials and Methods*) to evaluate whether or not a large Anglo-Saxon migration event is needed to explain the extremely high Central English-Frisian affinity. We started with a simple model of population fission with no background migration. We found a 95% credible interval for the split date using BATWING of 0–88 generations (0–2,200 years BP, assuming 25 years per generation), which corresponded well with the 95% confidence interval from the Monte Carlo likelihood method with no background migration (0–91 generations or 0–2,275 years BP assuming 25 years per generation).

Next, we looked at the levels of background migration, operating continuously from generation to generation, needed to maintain the Central English-Frisian genetic similarity under two other scenarios not involving Anglo-Saxon mass migration. Under an Island Model scenario (constant background migration between two populations that split at  $T_S = \infty$ ) the 95% confidence interval for  $m$ , estimated from the Monte Carlo likelihood method, is 0.3%–50% (where 50% indicates complete panmixia and is a maximum value for  $m$ ). The same result (to the significant digit given) is found under a Neolithic mass migration scenario (population split 240 generations BP). We note that a figure of  $m = 0.3\%$  is three times higher than the figure we estimated as representing an implausibly high value for  $m$ , well in excess of realistic values, based on migration statistics to and from the European Economic Area as a whole over the past 25 years. If we set  $m$  at the implausibly high value of 0.1%, the 95% confidence interval for a Central English-Frisian split date is 0–97 generations (0–2,425 years BP, assuming 25 years per generation). The figure of 97 generations BP represents an extreme upper limit for the migration event in this case both because it is based on such an implausibly high value for background migration and because it requires the most severe mass migration event imaginable, namely a 100% replacement of the English Y chromosome pool.

Next, we assumed that an Anglo-Saxon migration event did take place 60 generations ago (i.e., 1,500 years BP assuming 25 years per generation) and asked how big an event would be needed to explain the Central English-Frisian genetic similarity. If the Central English and Frisian populations were very different at the time of the event, a larger mass migration would be needed. We therefore started by assuming complete genetic identity of the two populations at the time of the Neolithic (i.e., a Central English-Frisian population split 240 generations BP). Assuming no background migration, the 95% confidence interval of the proportion  $F$  of the Central English population derived from an Anglo-Saxon mass migration event is 65%–100%. If a background migration rate since the Neolithic of  $m = 0.1\%$  is allowed, the 95% confidence interval for  $F$  widens to 50%–100%. This result is unchanged if a 30-year gen-

eration time is assumed (i.e., an Anglo-Saxon migration event 50 rather than 60 generations ago).

### Discussion

Our results indicate the presence of a strong genetic barrier between Central England and North Wales and the virtual absence of a barrier between Central England and Friesland. Any attempt to explain these results in terms of demographic history and migration needs to encompass both these findings satisfactorily. The Central English-North Welsh barrier cannot be explained purely as a simple isolation-by-distance phenomenon because it contrasts strongly with the lack of evidence for a cline among the five widely separated English towns. Our findings are particularly striking, given the high resolution and rapid mutation rate of the Y chromosome haplotypes on which they are based. These allow genetic barriers, if they exist, to be clearly defined.

The best explanation for our findings is that the Anglo-Saxon cultural transition in Central England coincided with a mass immigration from the continent. Such an event would simultaneously explain both the high Central English-Frisian affinity and the low Central English-North Welsh affinity. If we use a rate of 0.1%, as observed over the past 25 years, to represent an extremely high value for continuous background migration between Central England and continental Europe, then we estimate that an Anglo-Saxon immigration event affecting 50%–100% of the Central English male gene pool at that time is required. We note, however, that our data do not allow us to distinguish an event that simply added to the indigenous Central English male gene pool from one where indigenous males were displaced elsewhere or one where indigenous males were reduced in number. Furthermore, although our models assume a single instantaneous migration event, we would also expect a more gradual process lasting several generations but still resulting in the same degree of admixture (a picture which may fit the historical data better [Härke 2002]) to produce very similar genetic patterns.

We accept that our data do not prove conclusively that an Anglo-Saxon mass migration event took place. If a background migration rate of 0.3% is allowed between Central England and Friesland, then the need for a mass migration event disappears. However, we note that this is an extremely high rate even by modern standards and would have to have been maintained continuously over thousands of years. A background migration rate of 0.3% would imply that one in six of today's Central English males descend from Frisians (or a population identical to Frisians) that emigrated to England after the Anglo-Saxon period and that an equal proportion of today's Frisians descend from English in a like manner. We also note that under a unidirectional gene flow model involving immigration into Central England only, the rate of background migration would then have to double to be at least 0.6% on a continuous basis.

It is also true that a mass migration event could have occurred outside the Anglo-Saxon migration period because the 95% confidence interval for a Central En-

glish-Frisian split extends as far back as 425 B.C. (if one allows a background migration rate of 0.1% and a generation time of 25 years). Archaeology and the testimony of Caesar combine to suggest an immigration of the Belgae, a Celtic tribe from northern Gaul, into central southern England (Hampshire and West Sussex) between 100 and 80 B.C. (Hawkes 1968; Cunliffe 1988, pp. 147–149; Cunliffe 1991, pp. 108–110). Furthermore, although Friesland lay outside the maximum extent of the Roman Empire, small numbers of Frisian mercenaries were recruited by the Romans and stationed as far north as Hadrian's Wall (Breeze and Dobson 1978, pp. 139–140; Collingwood, Wright, and Tomlin 1995, p. 501). However, most historians would see these movements, if they would acknowledge them at all, as preludes to post-Roman Anglo-Saxon migration, and it would be odd indeed to deny the latter while at the same time assigning an extremely large mass migration status to the former.

Finally, we accept that our inferences are based on population genetic analyses that assume a particular model of microsatellite evolution under selective neutrality and growth and that departures from these assumptions may influence our results. However, we note that the accuracy of the mutation model is diminished in importance by the small number of generations that would allow new mutations to accumulate since Anglo-Saxon times and also that any selective sweeps would also have to have been very recent in order to have influenced our conclusions greatly, especially because the effects of such sweeps would partly be accommodated by our model of exponential population growth. In addition, the estimates provided by BATWING for effective population sizes at the time of the Anglo-Saxon migration event are very small (table 1). Thus, a large amount of error caused by drift is already allowed for by our BATWING and Monte Carlo likelihood analyses. We do not presume an exact correspondence between real and effective population size dynamics or between the real population history of England, which has seen many different changes in size, and our simple model of exponential growth. However, we note that the posterior mode of the effective population growth rate per generation provided by BATWING (6.0%) almost exactly matches the real estimated population growth rate averaged over the past 1,500 years (Hatcher 1977, pp. 1348–1530; Wrigley and Schofield 1989, pp. 1541–1871; Härke 2002), whereas our 95% credible interval covers both the lower average growth rates of A.D. 500–1750 (approximately 3% per 25 years, albeit with large fluctuations) and the higher average growth rates of A.D. 1750–2000 (approximately 23% per 25 years).

Anglo-Saxon settlements and culture appeared throughout England but, importantly, did not extend into North Wales, where many of the original Celtic Britons living in England are thought to have fled (Kearney 1989; Davies 1993, 1999). Conflict between the Welsh and Anglo-Saxon kingdoms continued over a long period. Offa's Dyke (an earthwork barrier 240 km long) was constructed ca. A.D. 790 and provided a well-defined boundary between England and Wales. The lin-

guistic, cultural, and political separation of the two regions lasted at least until A.D. 1282 when Edward I of England defeated the Welsh King Llywelyn II (Davies 1993). Our results suggest that this separation has also restricted male-mediated gene flow between the two regions over the past approximately 1500 years.

Comparisons of Central English and Norwegian haplotypes reveal no evidence of distinctive common signature haplotypes indicative of Viking origin, in contrast to Orcadian-Norwegian comparisons (Wilson et al. 2001). However, the Vikings who may have settled in East Anglia and the Midlands are thought to have been predominantly from Denmark, rather than Norway (Richards 2000). Previously published data suggest that the Danish have greater Y chromosome genetic affinity with the English than with the Norwegians (Malaspina et al. 2000; Rosser et al. 2000). However, the Danish-German border is believed to be another source location of the Anglo-Saxons (Kearney 1989; Davies 1999), so any Danish Viking influence on the English gene pool may prove difficult to distinguish from Anglo-Saxon influence. Further studies within Scandinavia and elsewhere are needed to resolve this issue.

This study shows that the Welsh border was more of a genetic barrier to Anglo-Saxon Y chromosome gene flow than the North Sea. Remarkably, we find that the resultant genetic differentiation is still discernible in the present day. These results indicate that a political boundary can be more important than a geophysical one in population genetic structuring and that informative patterns of genetic differentiation can be produced by migration events occurring within historical times.

## Acknowledgments

We thank all of the Y chromosome donors who contributed to this study and M. A. Strøksnes, Drs. P. Clark, M. Gee, P. R. Gage, S. J. Keen, M. R. Greenwood, J. D. Peterson, Kirtley, A. Hawkes, L. J. Moore, and M. Sears for helping to collect the samples. We thank Heinrich Härke, Lounès Chikhi, and two anonymous referees for valuable comments on an earlier draft of this paper. M.G.T. is supported by a Nuffield Foundation grant (NUF-NAL).

## LITERATURE CITED

- ADAMS, W. Y. 1968. Invasion, diffusion, evolution? *Antiquity* **42**:194–215.
- ADAMS, W. Y., D. P. VAN GERVEN, and R. S. LEVY. 1978. The retreat from migrationism. *Annu. Rev. Anthropol.* **7**:483–532.
- ANTHONY, D. W. 1990. Migration in archaeology: the baby and the bathwater. *Am. Anthropol.* **92**:895–914.
- ARNOLD, C. J. 1984. *From Roman Britain to Saxon England*. Croom Helm, London.
- BAHLO, M., and R. C. GRIFFITHS. 2000. Inference from gene trees in a subdivided population. *Theor. Popul. Biol.* **57**:79–95.
- BEERLI, P., and J. FELSENSTEIN. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. USA* **98**:4563–4568.

- BIANCHI, N. O., C. I. CATANESI, G. BAILLIET, V. L. MARTINEZ-MARIGNAC, C. M. BRAVI, L. B. VIDAL-RIOJA, R. HERRERA, and J. S. LOPEZ-CAMELO. 1998. Characterization of ancestral and derived Y-chromosomal haplotypes of New World populations. *Am. J. Hum. Genet.* **63**:1862–1871.
- BODMER, W. F. 1993. The genetics of Celtic populations. *Proc. Br. Acad.* **82**:37–57.
- BREEZE, D. J., and B. DOBSON. 1978. *Hadrian's Wall, Revised edition.* Pelican, Harmondsworth.
- BURMEISTER, S. 2000. Archaeology and migration: approaches to an archaeological proof of migration. *Curr. Anthropol.* **41**:539–567.
- CARTWRIGHT, R. A., H. J. HARGREAVES, and E. SUNDERLAND. 1977. Serum protein and isoenzyme polymorphisms from Nottingham, England. *Hum. Biol.* **49**:629–640.
- CASALOTTI, R., L. SIMONI, M. BELLEDI, and G. BARBUJANI. 1999. Y-chromosome polymorphisms and the origins of the European gene pool. *Proc. R. Soc. Lond. B* **266**:1959–1965.
- CAVALLI-SFORZA, L. L., P. MENOZZI, and A. PIAZZA. 1994. *The history and geography of human genes.* Princeton University Press, Princeton, NJ.
- CHAPMAN, J. 1997. The impact of modern invasions and migrations on archaeological explanation. Pp. 11–20 in J. CHAPMAN and H. HAMEROW, eds. *Migrations and invasions in archaeological explanation.* BAR International Series 664, Oxford, UK.
- CHAPMAN, J., and H. HAMEROW, eds. 1997. *Migrations and invasions in archaeological explanation.* BAR International Series 664, Oxford, UK.
- CLARK, G. 1966. The invasion hypothesis in British archaeology. *Antiquity* **40**:172–189.
- COLLINGWOOD, R. G., R. P. WRIGHT, and R. S. O. TOMLIN. 1995. *The Roman inscriptions of Britain, Vol. 1, 2nd edition.* Sutton, Stroud.
- CRAWFORD, S. 1997. Britons, Anglo-Saxons and the Germanic burial ritual. Pp. 45–72 in J. CHAPMAN and H. HAMEROW, eds. *Migrations and invasions in archaeological explanation.* BAR International Series 664, Oxford, UK.
- CUNLIFFE, B. 1988. *Greeks, Romans and Barbarians: spheres of interaction.* Batsford, London.
- . 1991. *Iron Age communities in Britain.* 3rd edition (Archaeology of Britain series). Routledge, London.
- DAVIES, J. 1993. *A History of Wales.* Penguin Press, London.
- DAVIES, N. 1999. *The Isles: a history.* Macmillan, London.
- ESMONDE-CLEARY, A. S. 1993. Approaches to the differences between Late Romano-British and Early Anglo-Saxon archaeology. *Anglo-Saxon Stud. Archaeol. Hist.* **6**:57–63.
- FALSETTI, A. B., and R. R. SOKOL. 1993. Genetic structure of human populations in the British Isles. *Ann. Hum. Biol.* **20**: 215–229.
- GOLDSTEIN, D. B., A. RUIZ LINARES, L. L. CAVALLI-SFORZA, and M. W. FELDMAN. 1995. Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl. Acad. Sci. USA* **92**:6723–6727.
- HAMEROW, H. 1997. Migration theory and the Anglo-Saxon “identity crisis.” Pp. 33–44 in J. CHAPMAN and H. HAMEROW, eds. *Migrations and invasions in archaeological explanation.* BAR International Series 664, Oxford, UK.
- HÄRKE, H. 1998. Anthropologists and migrations: a problem of attitude? *Curr. Anthropol.* **39**:19–45.
- . 2002. Kings and warriors: population and landscape in early medieval Britain. Pp. 145–175 in P. SLACK and R. WARD, eds. *The peopling of Britain: the shaping of a human landscape.* Oxford University Press, Oxford.
- HATCHER, J. 1977. *Plague, population and the English economy.* Palgrave Macmillan, London.
- HAWKES, C. F. C. 1968. New thoughts on the Belgae. *Antiquity* **42**:6–16.
- HAWKES, J., and C. HAWKES. 1942. *Prehistoric Britain.* Penguin, Harmondsworth, United Kingdom.
- HELGASON, A., S. SIGURDARDÓTTIR, J. NICHOLSON, B. SYKES, E. W. HILL, D. G. BRADLEY, V. BOSNES, J. R. GULCHER, R. WARD, and K. STEFÁNSSON. 2000. Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am. J. Hum. Genet.* **67**:697–717.
- HEYER, E., J. PUYMIRAT, P. DIELTJES, E. BAKKER, and P. DE KNIJFF. 1997. Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Hum. Mol. Genet.* **6**:799–803.
- HIGHAM, N. 1992. *Rome, Britain and the Anglo-Saxons.* Seaby, London.
- HILL, E. W., M. A. JOBLING, and D. G. BRADLEY. 2000. Y-chromosome variation and Irish origins. *Nature* **404**:351–352.
- HODGES, R. 1989. *The Anglo-Saxon achievement: archaeology and the beginnings of English society.* Duckworth, London.
- HUNTER, J., and I. RALSON, eds. 1998. *The archaeology of Britain.* Routledge, London.
- KAYSER, M., S. BRAUER, G. WEISS, W. SCHIEFENHÖVEL, P. A. UNDERHILL, and M. STONEKING. 2001. Independent histories of human Y chromosomes from Melanesia and Australia. *Am. J. Hum. Genet.* **68**:173–190.
- KAYSER, M., A. CAGLIÀ, D. CORACH et al. (30 co-authors). 1997. Evaluation of Y-chromosomal STRs: a multicenter study. *Int. J. Legal Med.* **110**:125–133.
- KAYSER, M., L. ROEWER, M. HEDMAN et al. (14 co-authors). 2000. Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *Am. J. Hum. Genet.* **66**:1580–1588.
- KEARNEY, H. 1989. *The British Isles: a history of four nations.* Cambridge University Press, Cambridge.
- LEEDS, E. T. 1954. The growth of Wessex. *Oxoniensia* **19**:45–60.
- LONG, C. C., C. DARKE, and R. MARKS. 1998. Celtic ancestry, HLA phenotype and increased risk of skin cancer. *Br. J. Dermatol.* **138**:627–630.
- MALASPINA, P., F. CRUCIANI, P. SANTOLAMAZZA et al. (24 co-authors). 2000. Patterns of male-specific inter-population divergence in Europe, West Asia and North Africa. *Ann. Hum. Genet.* **64**:395–412.
- MASCIE-TAYLOR, C. G. N., and G. W. LASKER. 1996. Further notes on possible changes in the geographic distribution of ABO and Rh blood groups in Great Britain. *Hum. Biol.* **68**: 473–478.
- MASTANA, S. S., R. JAYASEKARA, P. FISHER, R. J. SOKOL, and S. S. PAPIHA. 1993. Genetic polymorphism of orosomucoid (ORM) in populations of the United Kingdom, Indian subcontinent, and Cambodia. *Jpn. J. Hum. Genet.* **38**:289–296.
- MASTANA, S. S., and R. J. SOKOL. 1998. Genetic variation in the East Midlands. *Ann. Hum. Biol.* **25**:43–68.
- MERRYWEATHER-CLARKE, A. T., J. J. POINTON, J. D. SHEARMAN, and K. J. H. ROBSON. 1997. Global prevalence of putative haemochromatosis mutations. *J. Med. Genet.* **34**: 275–278.
- MICHALAKIS, Y., and L. EXCOFFIER. 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142**: 1061–1064.
- MYRES, J. N. L. 1986. *The English settlements.* Oxford University Press, Oxford.
- NÄSMAN, U. 1988. Den folkvandringstiden krisen i Sydskan-dinavien, inklusive Öland och Gotland. Pp. 227–255 in U.

- NÄSMAN and J. LUND, eds. *Folkevandringstiden i Norden: En krisetid mellem ældre og yngre jernalder*. Aarhus Universitetsforlag, Aarhus.
- NEI, M. 1987. *Molecular evolutionary genetics*. Columbia University Press, New York.
- NIELSEN, H. F. 1985. Old English and the continental Germanic languages: a survey of morphological and phonological interrelations. 2nd edition. Institut für Sprachwissenschaft der Universität Innsbruck, Innsbruck.
- NIELSEN, R., and J. WAKELEY. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics* **158**:885–896.
- PAPIHA, S. S., M. F. DUGGAN KEEN, and R. S. E. RODGER. 1985. DR antigens and Bf allotypes in northeast England, UK. *Hum. Hered.* **35**:246–250.
- PETERSEN, P. V. 1991. Nye fund af metalsager fra yngre germansk jernalder: Detektorfund og danefæ fra perioden 1966–88. Pp. 49–65 in P. MORTENSEN and B. M. RASMUSSEN, eds. *Fra Stamme til Stat i Danmark 2: Høvdingesamfund og Kongemagt*. Aarhus Universitetsforlag, Aarhus.
- POOLEY, C., and J. TURNBULL. 1998. *Migration and mobility in Britain since the eighteenth century*. UCL Press, London.
- POSER, C. M. 1994. The dissemination of multiple-sclerosis—a viking saga—a historical essay. *Ann. Neurol.* **36**:S231–S243.
- RAYMOND, M., and F. ROUSSET 1995. An exact test for population differentiation. *Evolution* **49**:1280–1283.
- RENFREW, C. 1987. *Archaeology and language: the puzzle of Indo-European origins*. Jonathan Cape, London.
- REYNOLDS, J., B. S. WEIR, and C. C. COCKERHAM. 1983. Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* **105**:767–779.
- RICHARDS, J. D. 2000. *Viking age England*. Stroud, Tempus.
- RICHARDS, M., V. MACAULAY, E. HICKEY et al. (37 co-authors). 2000. Tracing European founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum. Genet.* **67**:1251–1276.
- ROSSER, Z. H., T. ZERJAL, M. E. HURLES et al. (63 co-authors). 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am. J. Hum. Genet.* **67**:1526–1543.
- SEMINO, O., G. PASSARINO, P. J. OEFNER et al. (17 co-authors). 2000. The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* **290**:1155–1159.
- SIMONI, L., F. CALAFELL, D. PETTENER, J. BERTRANPETIT, and G. BARBUJANI. 2000. Geographic patterns of mtDNA diversity in Europe. *Am. J. Hum. Genet.* **66**:262–278.
- SHENNAN, S. 2000. Population, culture, history, and the dynamics of culture change. *Curr. Anthropol.* **41**:811–835.
- THOMAS, M. G., N. BRADMAN, and H. M. FLINN. 1999. High throughput analysis of 10 microsatellite and 11 binary polymorphisms on the human Y-chromosome. *Hum. Genet.* **105**:577–581.
- THOMAS, M. G., K. SKORECKI, H. BEN-AMI, T. PARFITT, N. BRADMAN, and D. B. GOLDSTEIN. 1998. Origins of Old Testament priests. *Nature* **394**:138–140.
- TORRONI, A., H. BANDELT, H. D'URBANO et al. (11 co-authors). 1998. mtDNA analysis reveals a major Late Paleolithic population expansion from southwestern to north-eastern Europe. *Am. J. Hum. Genet.* **62**:1137–1152.
- TYFIELD, L. A., M. J. OSBORN, and J. B. HOLTON. 1997. Sequence variation at the phenylalanine hydroxylase gene in the British Isles. *Am. J. Hum. Genet.* **60**:388–396.
- UNDERHILL, P. A., P. SHEN, A. A. LIN et al. (21 co-authors). 2000. Y chromosome sequence variation and the history of human populations. *Nat. Genet.* **26**:358–361.
- WEALE, M. E., L. YEPISKOPSYAN, R. F. JAGER, N. HOVHANNISYAN, A. KHUDOYAN, O. BURBAGE-HALL, N. BRADMAN, and M. G. THOMAS. 2001. Armenian Y chromosome haplotypes reveal strong regional structure within a single ethno-national group. *Hum. Genet.* **109**:659–674.
- WILSON, I. J., and D. J. BALDING. 1998. Genealogical inference from microsatellite data. *Genetics* **150**:499–510.
- WILSON, J. F., D. A. WEISS, M. RICHARDS, M. G. THOMAS, N. BRADMAN, and D. B. GOLDSTEIN. 2001. Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc. Natl. Acad. Sci. USA* **98**:5078–5083.
- WRIGLEY, E. A., and R. S. SCHOFIELD. 1989. *The population history of England*. Cambridge University Press, Cambridge.
- THE Y CHROMOSOME CONSORTIUM. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res.* **12**:339–348.

EDWARD HOLMES, reviewing editor

Accepted January 25, 2002