

# Y-Chromosomal Variation in the Czech Republic

F. Luca,<sup>1</sup> F. Di Giacomo,<sup>2</sup> T. Benincasa,<sup>1</sup> L.O. Popa,<sup>3</sup> J. Banyko,<sup>4</sup> A. Kracmarova,<sup>5</sup>  
P. Malaspinga,<sup>2</sup> A. Novelletto,<sup>1,2\*</sup> and R. Brdicka<sup>5</sup>

<sup>1</sup>Department of Cell Biology, University of Calabria, Rende, Italy

<sup>2</sup>Department of Biology, University "Tor Vergata", Rome, Italy

<sup>3</sup>National Museum of Natural History, Bucharest, Romania

<sup>4</sup>University of P. J. Safarik, Kosice, Slovak Republic

<sup>5</sup>Institute for Haematology and Blood Transfusion, Prague, Czech Republic

**KEY WORDS** Y chromosome; peopling of Europe; genetic dating; microsatellite variation

**ABSTRACT** To analyze the contribution of the Czech population to the Y-chromosome diversity landscape of Europe and to reconstruct past demographic events, we typed 257 males from five locations for 21 UEPs. Moreover, 141 carriers of the three most common haplogroups were typed for 10 microsatellites and coalescent analyses applied. Sixteen Hg's characterized by derived alleles were identified, the most common being R1a-SRY<sub>10831</sub> and P-DYS257\*(xR1a). The pool of haplogroups within I-M170 represented the third most common clade. Overall, the degree of population structure was low. The ages for Hg I-M170, P-DYS257\*(xR1a), and R1a-SRY<sub>10831</sub> appeared to be comparable and compatible with their presence during or soon after the LGM. A signal of popu-

lation growth beginning in the first millennium B.C. was detected. Its similarity among the three most common Hg's indicated that growth was characteristic for a gene pool that already contained all of them. The Czech population appears to be influenced, to a very moderate extent, by genetic inputs from outside Europe in the post-Neolithic and historical times. Population growth post-dated the archaeologically documented introduction of Neolithic technology and the estimated central value coincides with a period of repeated changes driven by the development of metal technologies and the associated social and trade organization. *Am J Phys Anthropol* 132:132–139, 2007. © 2006 Wiley-Liss, Inc.

The male-specific portion of the human Y chromosome (MSY) represents, together with mtDNA, an uniparentally inherited polymorphic system. This property is also associated with the ability of the MSY to detect high levels of structure within and between populations. In fact, not only the male-specific portion of the human Y chromosome (MSY) is represented in numbers 1/4 and 1/3 than the autosomes and the X chromosome, respectively, but it is also influenced by the action of additional factors that further reduce its effective population size. These include variance in family size and heritability of reproductive success (Austerlitz and Heyer, 2000; Heyer et al., 2005), both of which could be influenced by social rank and other cultural features in addition to biological determinants (Zerjal et al., 2003). The overall result is a very fast divergence attributable to stochastic factors which, in populations at the sub-continental scale and in the short term, can largely override limited gene flow and the possible action of natural selection.

Major advances in reconstructing the particular realization of the build-up of MSY diversity in the whole world, as well as in more local populations include: a) the ever-increasing detail of MSY phylogeny based on indel and other binary markers with low recurrence of mutation (Y Chromosome Consortium, 2002; Jobling and Tyler-Smith, 2003); b) the estimation of the antiquity of lineages based on either the infinite site mutation model for binary markers with negligible recurrence of mutation (UEP) or the stepwise mutation model for microsatellite variability into the UEP lineages, with a variety of methods (Wilson and Balding, 1998; De Knijff, 2000; Stumpf and Goldstein, 2001; Di Giacomo et al., 2004; Zhivotovsky et al., 2004; Luca et al., 2005); c) the use of

coalescent methods to quantify processes of population growth and infer events of subdivision from the observed distribution of molecular types (Weale et al., 2002; Kasparaviciute et al., 2004).

Evidence for a strong structure of the extant populations of Europe for the MSY has been accumulating by analyses of independent sets of binary indel and Single Nucleotide Polymorphisms (SNP) markers (Rosser et al., 2000; Semino et al., 2000). More recent studies, considering both spatial frequency patterns of haplogroups (Hg) with varying degrees of phyletic affinity and their age estimates, have identified clear patterns in the geographic distributions of haplotypes or Hg's as well as sharp genetic boundaries (Cruciani et al., 2004; Rootsi et al., 2004; Semino et al., 2004). They have proposed models for the underlying population movements, with a mainly prehistorical temporal assignment. In addition, analyses focused on specific Hg's or limited geographic regions displayed the outcome of post-Neolithic population processes (Di Giacomo et al., 2004; Pericic et al., 2005). Finally, the analysis of a large database of micro-

Grant sponsor: PRIN-MIUR 2005.

\*Correspondence to: Andrea Novelletto, Department of Biology, University of Rome "Tor Vergata", via della Ricerca Scientifica, snc, 00133 Rome, Italy. E-mail: novelletto@bio.uniroma2.it

Received 21 February 2006; accepted 8 August 2006.

DOI 10.1002/ajpa.20500

Published online 31 October 2006 in Wiley InterScience (www.interscience.wiley.com).



**Fig. 1.** Map of the Czech Republic showing the sampling locations.

satellite data has confirmed the main picture of the MSY geography in Europe, but has also highlighted the signature of more recent population events (Kayser et al., 2005; Roewer et al., 2005).

Here we report on the results of the characterization of MSY diversity in a population of central Europe as obtained with a sampling scheme that detected population substructure in other European countries (Malaspina et al., 2000; Stefan et al., 2001; Brion et al., 2004; Roewer et al., 2005). Indeed, the area here investigated is crucial to understanding the origin of the present-day genetic landscape of the Continent, as it lies on routes that had inevitably to be involved in a wide range of processes, including repopulation events from glacial refugia, population inputs from the Asian-European border, as well as local expansions triggered by major technological advances.

In this paper we specifically tested the hypothesis that MSY diversity in the Czech Republic retains a signature of a sudden population expansion as documented in the local archaeological record.

## MATERIALS AND METHODS

### Subjects

We studied an overall number of 257 males collected in the five sampling locations shown in Figure 1. These are arranged along a West-to-East transect in the southern part of the Czech Republic. The linear distances between consecutive locations are, from West to East, 65, 65, 65, and 50 km, respectively. Full informed consent was obtained from all participants in this study.

### DNA typings

We used 21 UEP markers, i.e. SRY<sub>10831</sub>, M78, M201, DYS221<sub>136</sub>, M170, M253, P37, M26, M223, M267, M172, M67, M9, LLY22g, Tat, DYS257(p27), M56, M157, M87, the YAP element insertion/deletion polymorphism, and the rearrangement detected by probe p12f2. These allow the recognition of the haplogroups (Hg's) listed in Table 1. The Y Chromosome Consortium (2002) nomenclature

for the I clade was revised according to new phylogeny reported by Underhill et al. (2005). Chromosomes that cannot be assigned to any of the above Hg's are classified as Y\*(xA,DE,G,I,J,K).

We used the following sequential typing scheme to determine Hg frequencies. YAP (Hammer and Horai, 1995) and DYS257 (Hammer et al., 1998) were typed in all subjects. p12f2 (Rosser et al., 2000) was typed on all YAP(-)/DYS257(G). All subjects producing a positive p12f2 amplification were typed with multiplex 1 (see below). LLY22g, Tat, and M9 (Underhill et al., 1997; Zerjal et al., 1997) were tested sequentially on all subjects carrying ancestral alleles at all markers in multiplex 1. Finally, all unclassified subjects were tested for SRY<sub>10831</sub> to detect Hg A. Within each Hg the remaining markers were assayed when appropriate: SRY<sub>10831</sub> (Kwok et al., 1996) within P-DYS257; M267, M172, and M67 (Malaspina et al., 2001; Underhill et al., 2001) within J-p12f2; M253, M223 (multiplex 2), and P37 (YCC, 2002) within I-M170; DYS221<sub>136</sub> (Hammer et al., 2001) within G-M201; M78 (Underhill et al., 2000) within DE-YAP; M56, M87, and M157 (Underhill et al., 2000) within R1a-SRY<sub>10831</sub>. These latter were used to search for additional variation within R1a-SRY<sub>10831</sub>, as originally described in Asia, but they turned out to be monomorphic and thus uninformative in this population. On the whole, in this scheme Hg P-DYS257\*(xR1a) is relatively ill-defined. However it has been shown that the bulk of these chromosomes in Europe are also derived at M173, P25, and M269 (Jobling and Tyler-Smith, 2003).

The YAP insertion and p12f2 rearrangement were typed by PCR, followed by direct visualization on agarose gels. M9, LLY22g, Tat, M78, and DYS257 were typed by restriction with HinfI, HindIII, Hsp92II, AciI, and BanI, respectively. M67, M172, DYS221, and M267 were typed by ASO probe hybridization (Di Giacomo et al., 2003, 2004). P37, M56, M87, and M157 were also typed by ASO hybridization, with probes and conditions listed in Table 2.

Multiplex reactions 1 and 2 were developed as fast assays for three and two loci, respectively (Table 3). Each set of loci was multiplexed in a 20  $\mu$ l reaction with

TABLE 1. Haplogroup absolute and relative frequencies in the five Czech sampling locations

| Haplogroup               | Location |       |          |         |       | Total |
|--------------------------|----------|-------|----------|---------|-------|-------|
|                          | Klatovy  | Pisek | J.Hradec | Trebitz | Brno  |       |
| DE-YAP*(xE3b1)           |          |       |          |         |       |       |
| N                        | 0        | 0     | 1        | 1       | 0     | 2     |
| %                        | 0        | 0     | 2.0      | 2.0     | 0     | 0.8   |
| E3b1-M78                 |          |       |          |         |       |       |
| N                        | 2        | 1     | 4        | 3       | 3     | 13    |
| %                        | 4.2      | 1.5   | 8.2      | 6.1     | 6.5   | 5.1   |
| G-M201*(xG2)             |          |       |          |         |       |       |
| N                        | 1        | 0     | 1        | 0       | 0     | 2     |
| %                        | 2.1      | 0     | 2.0      | 0       | 0     | 0.8   |
| G2-P15                   |          |       |          |         |       |       |
| N                        | 3        | 4     | 2        | 2       | 0     | 11    |
| %                        | 6.3      | 6.2   | 4.1      | 4.1     | 0     | 4.3   |
| I-M170*(xI1,I2a,I2b1)    |          |       |          |         |       |       |
| N                        | 0        | 3     | 2        | 0       | 0     | 5     |
| %                        | 0        | 4.6   | 4.1      | 0       | 0     | 1.9   |
| I1-M253                  |          |       |          |         |       |       |
| N                        | 2        | 5     | 1        | 1       | 4     | 13    |
| %                        | 4.2      | 7.7   | 2.0      | 2.0     | 8.7   | 5.1   |
| I2a-P37                  |          |       |          |         |       |       |
| N                        | 7        | 6     | 2        | 2       | 2     | 19    |
| %                        | 14.6     | 9.2   | 4.1      | 4.1     | 4.3   | 7.4   |
| I2a1-M26                 |          |       |          |         |       |       |
| N                        | 0        | 0     | 0        | 2       | 1     | 3     |
| %                        | 0        | 0     | 0        | 4.1     | 2.2   | 1.2   |
| I2b1-M223                |          |       |          |         |       |       |
| N                        | 3        | 2     | 2        | 0       | 0     | 7     |
| %                        | 6.3      | 3.1   | 4.1      | 0       | 0     | 2.7   |
| J-p12f2*(xJ1,J2)         |          |       |          |         |       |       |
| N                        | 0        | 1     | 0        | 0       | 2     | 3     |
| %                        | 0        | 1.5   | 0        | 0       | 4.3   | 1.2   |
| J1-M267                  |          |       |          |         |       |       |
| N                        | —        | —     | —        | —       | —     | —     |
| %                        | —        | —     | —        | —       | —     | —     |
| J2-M172*(xJ2f)           |          |       |          |         |       |       |
| N                        | 1        | 1     | 1        | 3       | 0     | 6     |
| %                        | 2.1      | 1.5   | 2.0      | 6.1     | 0     | 2.3   |
| J2f-M67                  |          |       |          |         |       |       |
| N                        | 0        | 1     | 0        | 1       | 1     | 3     |
| %                        | 0        | 1.5   | 0        | 2.0     | 2.2   | 1.2   |
| K-M9*(xP,N)              |          |       |          |         |       |       |
| N                        | 0        | 1     | 0        | 0       | 0     | 1     |
| %                        | 0        | 1.5   | 0        | 0       | 0     | 0.4   |
| N-LLY22g*(xN3)           |          |       |          |         |       |       |
| N                        | —        | —     | —        | —       | —     | —     |
| %                        | —        | —     | —        | —       | —     | —     |
| N3-Tat                   |          |       |          |         |       |       |
| N                        | 0        | 2     | 1        | 1       | 0     | 4     |
| %                        | 0        | 3.1   | 2.0      | 2.0     | 0     | 1.6   |
| P-DYS257*(xR1a)          |          |       |          |         |       |       |
| N                        | 11       | 19    | 13       | 16      | 13    | 72    |
| %                        | 22.9     | 29.2  | 26.5     | 32.7    | 28.3  | 28.0  |
| R1a-SRY <sub>10831</sub> |          |       |          |         |       |       |
| N                        | 17       | 19    | 16       | 17      | 19    | 88    |
| %                        | 35.4     | 29.2  | 32.7     | 34.7    | 41.3  | 34.2  |
| Y*(xA,DE,G,I,J,K)        |          |       |          |         |       |       |
| N                        | 1        | 0     | 3        | 0       | 1     | 5     |
| %                        | 2.1      | 0     | 6.1      | 0       | 2.2   | 1.9   |
| Total                    |          |       |          |         |       |       |
| N                        | 48       | 65    | 49       | 49      | 46    | 257   |
| %                        | 100.0    | 100.0 | 100.0    | 100.0   | 100.0 | 100.0 |

2.5 mM MgCl<sub>2</sub>, 0.125 mM dNTPs, 0.375 ng/μl BSA, and 0.5 U Taq Polymerase for 33 cycles (94°C, 20 s; 56°C, 1 min; 72°C, 1 min). One micro liter of each reaction was then reamplified with one of the original primers and two internal fluorescently labeled allele-specific primers (Table 3) in a 10 μl reaction with 2.0 mM MgCl<sub>2</sub>, 0.20

mM dNTPs, 0.5 ng/μl BSA, and 0.5 U Taq Polymerase for 18 cycles (94°C, 30 s; 42°C, 1 min; 72°C, 30 s). Two micro liters of the product were mixed with 12 μl of formamide, loaded on an ABI 310 automated sequencer and analyzed with the GeneScan software with Tamra as internal standard. Allele states were identified by black (ancestral) or blue (derived) peaks of locus-specific size.

We also typed 141 subjects carrying Hg's I\*(xI1,I2a,I2b1), I1-M253, I2a-P37, I2a1-M26, I2b1-M223, P-DYS257\*(xR1a), and R1a-SRY<sub>10831</sub> with the following microsatellite markers (Butler et al., 2002): YCA2A, YCA2B, YCAIIa, YCAIIb, DYS385A and B, DYS388, DYS391, DYS426, DYS439, DYS460, and H4 (individual haplotype data available at [www2.bio.uniroma2.it/biologia/laboratori/lab-geneticaumana/geneuma-pubb.htm](http://www2.bio.uniroma2.it/biologia/laboratori/lab-geneticaumana/geneuma-pubb.htm)).

### Data analysis

Diversity indexes and AMOVA computations were obtained with the Arlequin 2.000 package (Schneider et al., 2000). To assay the gradient of Hg differentiation, we carried out spatial autocorrelation analysis using the program SAAP (Sokal and Oden, 1978). We performed several runs using different numbers of distance classes to faithfully represent the geographical distances among samples, yet retaining a meaningful number of comparisons in each class.

Dating estimates of Hg antiquity based on microsatellite diversity were obtained with the program BATWING (Wilson and Balding, 1998) under two different conditions: in the first one, all settings for prior distributions were as described (Arredi et al., 2004; Di Giacomo et al., 2004). In the second one, the priors for α (the rate of increase of population size) and β (the time of start of population growth) were relaxed to UNIFORM (0.0, 0.04) and UNIFORM (0.0, 1.0), respectively, in order to explore the signature of population growth which is present in the data. Mutation rates at the 10 STR loci were given GAMMA (2,1000) as prior. We already validated these settings (Luca et al., 2005) by observing that they are able to predict figures for mutation rates obtained in father-son transmissions (Gusmao et al., 2005) and that using lower priors for mutation rates produces convergent results.

For Hg I, information on the phylogenetic relationships between internal haplogroups was used, with the "infsites" option, which permits only STR trees consistent with the Single Nucleotide Polymorphisms (SNP) data. Two I2a1-M26 haplotypes were not considered, as the multirepeat deletion at YCAIIb found on these chromosomes does not conform to the stepwise mutation model assumed in the method. Each BATWING run consisted of 5,000 Markovian steps, after a warmup of 1,000.

A multidimensional representation of population affinities was obtained by Correspondence Analysis (as implemented in SPSS) on the listing of data from Poland, Germany (Kayser et al., 2005), Croatia, Serbia and Bosnia (Marjanovic et al., 2005) and Italy (our unpublished data). This required some approximation in Hg assignment across studies, e.g. lumping of F\*(xJ2,K), G\*, J1 and the uncharacterized Hg's into a single category.

### RESULTS

Haplogroup frequencies are reported in Table 1. The diversity between population samples was very low, with an undetectable overall F<sub>st</sub>. R1a-SRY<sub>10831</sub> is the single

TABLE 2. Probe sequences and hybridization and washing temperatures for four loci assayed by dot-blot hybridization

| Locus             | PCR primers             | Allele       | AS0-PROBE  | T (°C)   |
|-------------------|-------------------------|--------------|--|----------|
| P37               | YCC, 2002               | ANC.<br>DER. | TTG GTT CAT AGT GTA AA<br>TTG GTT CAC AGT GTA AA       | 44<br>44 |
| M56 <sup>1</sup>  | Paracchini et al., 2002 | ANC.<br>DER. | GAT TAC GAA GAA AGG AG<br>GAT TAC GAT GAA AGG AG       | 45<br>45 |
| M87 <sup>1</sup>  | Paracchini et al., 2002 | ANC.<br>DER. | GAA TCT TAT ATT TTT GT<br>GAA TCT TAC ATT TTT GT       | 40<br>40 |
| M157 <sup>1</sup> | Paracchini et al., 2002 | ANC.<br>DER. | AAC AAA AAC AAC CAC AAA T<br>AAC AAA AAC CAC CAC AAA T | 45<br>46 |

<sup>1</sup> Multiplexed in a single reaction.

TABLE 3. Primer sequences and conditions for the allele-specific assays of multiplex PCR reactions 1 and 2

| Locus       | Reaction 1                             |                                      | Reaction 2                              |         | Product size (bp) |
|-------------|--|--------------------------------------|---|---------|-------------------|
|             | Primer concentration <sup>1</sup> (μM) | Primers                              | Primer concentration (μM)               | Product |                   |
| Multiplex 1 | M26                                    | 0.10                                 | HEX-ATTCAGTGTCTCTGTC                    | 0.15    | 83                |
|             |  |                                      | FAM-ATTCAGTGTCTCTGT                     | 0.11    |                   |
|             |  |                                      | CCAGTGTGTAAGTTTTATTACAATTT <sup>2</sup> | 0.15    |                   |
|             | M170                                   | 0.08                                 | HEX-CTTAAAAATCATTGTTCA                  | 0.20    | 96                |
|             |  |                                      | FAM-CTTAAAAATCATTGTTCC                  | 0.06    |                   |
|             |  |                                      | CCAATTACTTTCACATTTAAGACC <sup>3</sup>   | 0.15    |                   |
| M201        | 0.08                                   | HEX-GTACCTATTACGAAAAAC               | 0.15                                    | 152     |                   |
|             |  | FAM-GTACCTATTACGAAAAA                | 0.15                                    |         |                   |
|             |  | TATGCATTTGTTGAGTATATGTC <sup>2</sup> | 0.15                                    |         |                   |
| Multiplex 2 | M223                                   | 0.17                                 | HEX-GATAAAATTTACTTACAGTC                | 0.15    | 160               |
|             |  |                                      | FAM-GATAAAATTTACTTACAGTT                | 0.08    |                   |
|             |  |                                      | CCTTTTTGGATCATGGTTCTT <sup>4</sup>      | 0.15    |                   |
|             | M253                                   | 0.10                                 | HEX-ATAGATAGCAAGTTGAC                   | 0.19    | 135               |
|             |  |                                      | FAM-ATAGATAGCAAGTTGAT                   | 0.15    |                   |
|             |  |                                      | CAGCTCCACCTCTATGCAGTTT <sup>5</sup>     | 0.15    |                   |

<sup>1</sup> For both primers; primer sequences as in Underhill et al. (2000).

<sup>2</sup> Same as forward primer of Underhill et al. (2000).

<sup>3</sup> Same as reverse primer of Underhill et al. (2000).

<sup>4</sup> Same as reverse primer of Underhill et al. (2001).

<sup>5</sup> Same as reverse primer of Cinnioglu et al. (2004).

Hg showing the highest frequency in the easternmost sample of Brno. This is in line with the frequency and continental distribution of this Hg reported by Rosser et al. (2000), who showed a sharp decline over the transect here examined. In our data this trend is detected by spatial autocorrelation (Moran I decreasing from 0.026 to -0.967), though it is not statistically significant. Given the linear arrangement of the samples, we also simply regressed Hg frequency on geographic distance, with insignificant results ( $r = 0.296$ ;  $P = n.s.$ ). Furthermore, no clear trend in the parallel decrease in the frequency of P-DYS257\*(xR1a) was detected ( $r = 0.791$ ;  $P = n.s.$ ). These results point towards a remarkable homogeneity among the five population samples, that can be treated as a single super-sample.

Haplogroup frequencies in the overall sample are in agreement with those reported in the literature for the population of the Czech Republic or neighbouring nations. E3b1-M78 frequency is in the range (2.9%–8%) reported in countries of central Europe (Cruciani et al., 2004; Semino et al., 2004). The frequency of G-M201 (inclusive of G2) agrees with a pattern of leveled frequencies always lower than 4.4% in central Europe (Semino et al., 2000). We show here that  $\gg 80\%$  of the entire haplogroup G-M201 is accounted for by the derived Hg G2-P15. The G2-P15 frequency among Czechs is slightly lower than that found in continental Italy and Greece (6.3 and 6.6%,

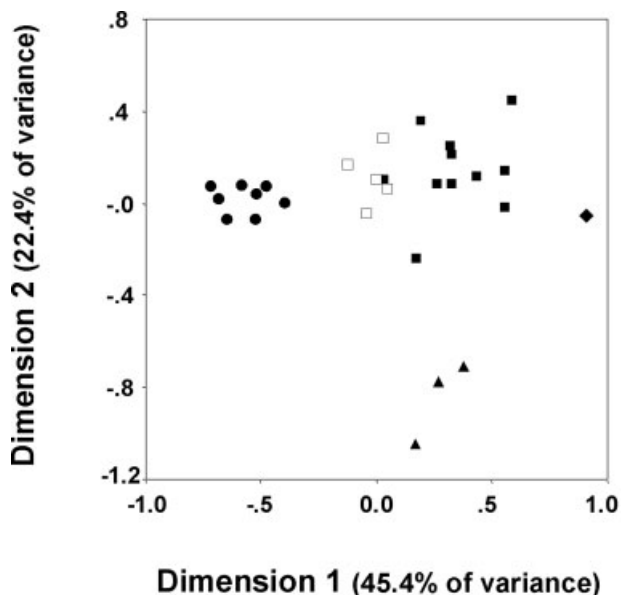


Fig. 2. Correspondence Analysis. Two-dimensional plot of the distribution of populations according to their Hg frequencies. ●: Polish samples; ■: German samples; ◆: Italian sample; ▲: Balkan samples; □: Czech samples.

TABLE 4. STR haplotype diversity for three major Hg's and averages (2.5th to 97.5th percentiles) of posterior distributions from coalescent analyses for Hg age, natural rate of increase of the population (Alpha), and time since beginning of population growth (Beta), under different prior conditions

|                        | I-M170              | I1-M253 <sup>1</sup> | I2a-P37 <sup>1</sup> | I2b1-M223 <sup>1</sup> | P-DYS257*(xR1a)     | R1a-SRY <sub>10831</sub> |
|------------------------|---------------------|----------------------|----------------------|------------------------|---------------------|--------------------------|
| Typed                  | 34                  | 9                    | 14                   | 7                      | 43                  | 62                       |
| H <sup>2</sup>         | 26                  | 7                    | 10                   | 6                      | 36                  | 41                       |
| Mean                   | 4.58 ± 2.30         | 2.03 ± 1.25          | 1.94 ± 1.17          | 2.38 ± 1.47            | 3.52 ± 1.83         | 2.98 ± 1.58              |
| pairwise<br>difference |                     |                      |                      |                        |                     |                          |
| Diversity              | 0.978 ± 0.014       | 0.917 ± 0.092        | 0.923 ± 0.060        | 0.952 ± 0.096          | 0.989 ± 0.008       | 0.966 ± 0.013            |
| Stringent priors       |                     |                      |                      |                        |                     |                          |
| Age <sup>3</sup>       | 534 (261–1,041)     | 340 (60–1,041)       | 218 (65–548)         | 227 (53–623)           | 581 (22–1,392)      | 469 (224–1,036)          |
| Alpha                  | 0.040 (0.031–0.050) |                      |                      |                        | 0.042 (0.031–0.050) | 0.043 (0.031–0.050)      |
| Beta <sup>3</sup>      | 58 (34–90)          |                      |                      |                        | 93 (61–141)         | 84 (53–132)              |
| Relaxed priors         |                     |                      |                      |                        |                     |                          |
| Age <sup>3</sup>       | 497 (234–998)       | 325 (68–998)         | 218 (72–548)         | 229 (54–635)           | 398 (183–889)       | 348 (176–697)            |
| Alpha                  | 0.023 (0.003–0.039) |                      |                      |                        | 0.031 (0.016–0.040) | 0.032 (0.015–0.040)      |
| Beta <sup>3</sup>      | 97 (28–217)         |                      |                      |                        | 150 (97–222)        | 125 (82–191)             |

<sup>1</sup> Age as midpoint between minimum and maximum.

<sup>2</sup> Number of different haplotypes.

<sup>3</sup> In generations.

respectively) (Di Giacomo et al., 2003). The frequencies of Hg I-M170 and its internal Hg's here found are very similar to those previously reported. However, we also found carriers of I2a1-M26, also in this population associated with the diagnostic YCAIIb 11 allele (Ciminelli et al., 1995; Quintana-Murci et al., 1999). This brings the SNP diversity index for the entire I clade to 0.738, i.e. the third highest among the large series of European populations examined (Rootsi et al., 2004). Both Hg's J-p12f2 and N3-Tat display frequencies lower than 5% in the overall sample.

The position of Czech samples in the genetic landscape of central Europe emerges in Correspondence Analysis (Fig. 2). In the first dimension the Czech samples are clustered and positioned in between the Polish and German samples. In fact Hg's R1a and N3-Tat (more frequent in Poland) have the largest negative loadings on this dimension, whereas Hg's K\*-M9 and J2\*-M172 (more frequent in Germany and Italy) have the largest positive ones. The second dimension is mainly contributed by Hg's I-M170, E3b1-M78, and K\*-M9 (negative loadings) vs. Hg P-DYS257\*(xR1a). Thus, the position of the three Balkan populations in Figure 2 is in line with the focal distribution of Hg's I and E3b1 in this area (Cruciani et al., 2004; Pericic et al., 2005).

To obtain information on past processes involving the Czech population, we analyzed the diversity at 10 microsatellite loci in the three most common Hg's (Table 4). The 10 loci here used produced a very high STR haplotype diversity, only slightly diminished in I1-M253, I2a-P37, and I2b1-M223. I1-M253 and I2b1-M223 indeed shared YCAII 21-19 as the most common pattern (Rootsi et al., 2004), but this affinity was not extended to the rest of the haplotype, in line with the reviewed phylogeny. In fact, I2b1-M223 shared DYS391, DYS426, and H4 modal alleles with I1-M253, but it also shared DYS388, DYS385b, DYS426, and H4 modal alleles with I2a-P37. Mean pairwise difference among STR haplotypes within each Hg was highest for I-M170, but it was strongly reduced in each of its subclades.

We then applied coalescent dating methods under a model of population growth in the post-glacial period, summarized in the prior distributions for  $\alpha$  and  $\beta$ . Under these conditions, the ages of Hg's I-M170, P-DYS257\*(xR1a), and R1a-SRY<sub>10831</sub> turned out to be almost com-

parable, while the ages of I subclades were remarkably younger than the entire Hg. In our analysis the ages of I2a-P37 and I2b1-M223 were similar and approximately one half that of the entire Hg I-M170 clade, whereas previous results (Rootsi et al., 2004) identified I2b1-M223 (formerly I1c) as the oldest I subclade.

To exclude that strict ranges for  $\alpha$  and  $\beta$  priors played an overwhelming role on the reconstruction of the coalescent process, we relaxed the assumptions on these parameters by using flat priors. In these conditions Hg I-M170 as a whole emerges as the oldest, in line with its presence during or soon after the LGM (Rootsi et al., 2004). The results confirmed a growth rate in a narrow range among the three Hg's, between 0.023 and 0.032, i.e. a still marked growth rate though lower than before, paralleled by a more ancient start of growth. The beginning of growth ranges between 97 and 150 generations ago, i.e. between the second and first millennium B.C. by considering 30 years/generation (Arredi et al., 2004). The ages of Hg's P-DYS257\*(xR1a) and R1a-SRY<sub>10831</sub> appear to be only slightly younger but nevertheless place their origins most likely in the Upper Paleolithic. In summary, figures for the rate and start of population growth were similar for the three Hg's, and indicated that growth was characteristic for the whole population in a relatively more recent period.

## DISCUSSION

We explored diversity of the male-specific portion of the human Y chromosome (MSY) in five closely spaced Czech population samples. The Hg's P-DYS257\*(xR1a) and R1a-SRY<sub>10831</sub> establish a major divide across central Europe, with a line roughly extending from the Adriatic to the Baltic seas, potentially crossing the transect examined here. This line separates high frequencies of R1a-SRY<sub>10831</sub> to the East from low frequencies to the West, with an opposite trend for P-DYS257\*(xR1a) (Malaspina et al., 2000; Rosser et al., 2000; Semino et al., 2000). The same line was also detected in an analysis of population differentiation at seven microsatellite loci commonly used in forensics (Roewer et al., 2005). Kayser et al. (2005) found this sharp genetic boundary to coincide with the German-Polish border, and interpreted it as the result of massive population movements associated with

World War II, superimposed on pre-existing continent-wide clines. Here we show a very similar composition among five Czech sub-samples. Equating P-DYS257\* (xR1a) to R1\*(xR1a1), and R1a-SRY<sub>10831</sub> to R1a1, our sub-samples display frequencies intermediate between the Polish and German ones. The overall ratio between the two Hgs is 1.22 (34.2/28.0) as compared to <1 in Germany and >3 in Poland. Thus, the Czech Republic appears to be affected by a much smoother frequency shift, if any, supporting the interpretation by Kayser et al. (2005) for a very recent origin of the German-Polish discrepancy.

The frequency patterns of other Hg's contribute to the description of the Czech population. First, Hg I-M170 frequency is in the range reported for other central European populations (Semino et al., 2000; Rootsi et al., 2004; Kayser et al., 2005), with the exception of Balkans (Marjanovic et al., 2005). All its three common subclades are represented, and also I2a1-M26 is found, enlarging to the East the area where this latter Hg is found within Europe, though at low frequencies (Capelli et al., 2003; Rootsi et al., 2004). These findings further increase the SNP diversity within Hg I-M170 and raise questions on whether the presence of one or more of its internal haplogroups (before or after the LGM) has been obscured in some areas of Europe by the overlay of other Hg's.

Also, the frequency of YAP+ chromosomes and the large proportion accounted by E3b1-M78 within this clade, is in line with other central European results, signed by Neolithic or post-Neolithic range expansions within the Continent (Cruciani et al., 2004).

The low frequency of J-p12f2 and its subclades confirms the low gene flow with the Mediterranean, from where this Hg spread into southern Europe (Marjanovic et al., 2005; Capelli et al., 2006 and references therein). King and Underhill (2002) associated this Hg with Neolithic inputs into Europe by observing that its presence predicts the distribution of painted pottery. Also, the missing Hg J1-M267 denotes very poor contacts with the Middle East. N3-Tat is a marker of populations who originally spoke Uralic languages, that is found at high frequency in the Baltic region (Zerjal et al., 1997). Its low frequency in the Czech population is in line with a reduced genetic contribution from northern Europe (Jobling and Tyler-Smith, 2003).

Overall, these results identify the Czech population as one influenced to a very moderate extent by genetic inputs from outside Europe in the post-Neolithic and historical times. It thus may represent an ideal population to draw inferences on geographically confined processes that might have occurred also in other parts of central Europe.

Inferences based on STR variation in the three most common Hg's obtained with coalescent methods deserve careful evaluation. First, even though our sampling was carried out in a limited geographic area, it returned age estimates for I-M170, P-DYS257\*(xR1a), and R1a-SRY<sub>10831</sub> similar to those obtained in reports with a wider geographical coverage (Kivisild et al., 2003; Cruciani et al., 2004; Rootsi et al., 2004; Semino et al., 2004; Pericic et al., 2005). These estimates paralleled the trend in mean pairwise difference distribution which, however, weights equally mutation rates across STR loci. Our coalescent estimates are associated with very large confidence intervals (up to three fold the average) and, as such, can only be used with caution to locate in time the age of Hg's (molecules). Conservatively, one can simply conclude that the Czech population harbors a large part of the STR variation generated in each Hg. With two sets of

prior assumptions, the ages of the three most common Hg's turned out to be largely overlapping, and compatible with their presence during or soon after the LGM.

However, a local signal emerged from the distribution of this diversity, i.e. that of a fast and recent population growth, which persists even after relaxing the prior assumptions and is similar for the three Hg's. This is summarized by the parameters  $\alpha$  and  $\beta$  and their relatively narrow confidence intervals (up to 1.5 fold the average). Estimation of the  $\beta$  parameter most likely locates the beginning of this process in the first millennium B.C., with confidence intervals that are barely compatible with the archaeologically documented introduction of Neolithic technology in this area (Haak et al., 2005). At least for the female lineage, these authors found a little genetic contribution to the present European gene pool from the first farmers settled in the area. Independently from the relevance of these data for reconstructing the genetics of Europe in the early Neolithic (Barbujani and Chikci, 2006), our central value for population growth coincides with a later period of repeated changes in the material cultures in this geographic region, driven by the development of metal technologies and the associated social and trade organization (Childe, 1957; Piggott, 1965; Kristiansen, 1998; Kruta, 2000).

In conclusion, the combined use of SNP and STR markers allowed us to explore different time horizons for the age of molecules and for the process of population growth (Barbujani and Chikci, 2006; Torroni et al., 2006). In fact, our data for the Czech population favor a model in which the age of the most common MSY molecules and their presence in the same gene pool largely predate a consistent population growth. Similar results have been obtained for Lithuania (Kasperaviciute et al., 2004). Both regions lie at the north-western and northern edge, respectively, of the putative homeland (central and south-eastern Europe) of an aboriginal quota of the molecular MSY diversity. This offers an unprecedented opportunity to test alternative models for a continental pattern of diversity which is arranged along the southeast-to-northwest axis (Cavalli-Sforza et al., 1994; Currat and Excoffier, 2005). The question whether this could be the result not only of a single demic diffusion, but also of the demographic increases affecting pre-existing local gene pools is still open (Renfrew, 1979). Examples of recent growth of pre-existing gene pools adding complexity to simple demic diffusion models are provided by mtDNA Hg's HV and H1 (Torroni et al., 1998), as well as Y chromosomal Hg R-SRY<sub>2627</sub> (Hurles et al., 1999). In order to arrive at a more complete picture, future research efforts should be directed towards the study of additional populations from central and south-eastern Europe with genetic markers and statistical methods which enable inferences on population dynamics. Populations in which a major and recent admixture can be excluded represent an ideal setting to perform this search.

## ACKNOWLEDGMENTS

J.B. and L.O.P. were fellows under the NATO "Outreach" programme.

## LITERATURE CITED

Arredi B, Poloni ES, Paracchini S, Zerjal T, Fathallah DM, Makrelouf M, Pascali VL, Novelletto A, Tyler-Smith C. 2004.

- A predominantly Neolithic origin for Y-chromosomal DNA variation in North Africa. *Am J Hum Genet* 75:338–345.
- Austerlitz F, Heyer E. 2000. Allelic association is increased by correlation of effective family size. *Eur J Hum Genet* 8:980–985.
- Barbujani G, Chikhi L. 2006. Population genetics: DNAs from the European Neolithic. *Heredity* 97:84, 85.
- Brion M, Quintans B, Zarrabeitia M, Gonzalez-Neira A, Salas A, Lareu V, Tyler-Smith C, Carracedo A. 2004. Micro-geographical differentiation in Northern Iberia revealed by Y-chromosomal DNA analysis. *Gene* 329:17–25.
- Butler JM, Schoske R, Vallone PM, Kline MC, Redd AJ, Hammer MF. 2002. A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. *Forensic Sci Int* 129:10–24.
- Capelli C, Redhead N, Abernethy JK, Gratrix F, Wilson JF, Moen T, Hervig T, Richards M, Stumpf MP, Underhill PA, Bradshaw P, Shaha A, Thomas MG, Bradman N, Goldstein DB. 2003. A Y chromosome census of the British Isles. *Curr Biol* 13:979–984.
- Capelli C, Redhead N, Romano V, Cali F, Lefranc G, Delague V, Megarbane A, Felice AE, Pascali VL, Neophytou PI, Poulli Z, Novelletto A, Malaspina P, Terrenato L, Fellous M, Thomas MG, Goldstein DB. 2006. Population structure in the Mediterranean basin: A Y chromosome perspective. *Ann Hum Genet* 70:207–225.
- Cavalli Sforza LL, Menozzi P, Piazza A. 1994. The history and geography of human genes. Princeton: Princeton University Press.
- Childe VG. 1957. The dawn of European civilization, 6th ed. London: Routledge Kegan Paul.
- Ciminelli BM, Pompei F, Malaspina P, Hammer M, Persichetti F, Pignatti PF, Palena A, Anagnou N, Guanti G, Jodice C, Terrenato L, Novelletto A. 1995. Recurrent simple tandem repeat mutations during human Y-chromosome radiation in Caucasian subpopulations. *J Mol Evol* 41:966–973.
- Cinnioglu C, King R, Kivisild T, Kalfoglu E, Atasoy S, Cavalleri GL, Lillie AS, Roseman CC, Lin AA, Prince K, Oefner PJ, Shen P, Semino O, Cavalli-Sforza LL, Underhill PA. 2004. Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet* 114:127–148.
- Cruciani F, La Fratta R, Santolamazza P, Sellitto D, Pascone R, Moral P, Watson E, Guida V, Colomb EB, Zaharova B, Lavinha J, Vona G, Aman R, Cali F, Akar N, Richards M, Torroni A, Novelletto A, Scozzari R. 2004. Phylogeographic analysis of haplogroup E3b (E-M215) Y chromosomes reveals multiple migratory events within and out of Africa. *Am J Hum Genet* 74:1014–1022.
- Curat M, Excoffier L. 2005. The effect of the Neolithic expansion on European molecular diversity. *Proc Biol Sci* 272:679–688.
- De Knijff P. 2000. Messages through bottlenecks: On the combined use of slow and fast evolving polymorphic markers on the human Y chromosomes. *Am J Hum Genet* 67:1055–1061.
- Di Giacomo F, Luca F, Anagnou N, Ciavarella G, Corbo RM, Cresta M, Cucci F, Di Stasi L, Agostiano V, Giparaki M, Loutradis A, Mammi' C, Michalodimitrakis EN, Papola F, Pedicini G, Plata E, Terrenato L, Tofaneli S, Malaspina P, Novelletto A. 2003. Clinal patterns of human Y chromosomal diversity in continental Italy and Greece are dominated by drift and founder effects. *Mol Phylogenet Evol* 28:387–395.
- Di Giacomo F, Luca F, Popa LO, Akar N, Anagnou N, Banyko J, Brdicka R, Barbujani G, Papola F, Ciavarella G, Cucci F, Di Stasi L, Gavrila L, Kerimova MG, Kovatchev D, Kozlov AI, Loutradis A, Mandarino V, Mammi' C, Michalodimitrakis EN, Paoli G, Pappa KI, Pedicini G, Terrenato L, Tofaneli S, Malaspina P, Novelletto A. 2004. Y chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe. *Hum Genet* 115:357–371.
- Gusmao L, Sanchez-Diz P, Calafell F, Martin P, Alonso CA, Alvarez-Fernandez F, Alves C, Borjas-Fajardo L, Bozzo WR, Bravo ML, Builes JJ, Capilla J, Carvalho M, Castillo C, Catanesi CI, Corach D, Di Lonardo AM, Espinheira R, Fagundes de Carvalho E, Farfan MJ, Figueiredo HP, Gomes I, Lojo MM, Marino M, Pinheiro MF, Pontes ML, Prieto V, Ramos-Luis E, Riancho JA, Souza Goes AC, Santapa OA, Sumita DR, Vallejo G, Vidal Rioja L, Vide MC, Vieira da Silva CI, Whittle MR, Zabala W, Zarrabeitia MT, Alonso A, Carracedo A, Amorim A. 2005. Mutation rates at Y chromosome specific microsatellites. *Hum Mutat* 26:520–528.
- Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tanzer M, Vilems R, Renfrew C, Gronenborn D, Alt KW, Burger J. 2005. Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science* 310:1016–1018.
- Hammer MF, Horai S. 1995. Y chromosomal DNA variation and the peopling of Japan. *Am J Hum Genet* 56:951–962.
- Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, Templeton AR, Zegura SL. 1998. Out of Africa and back again: Nested cladistic analysis of human Y chromosome variation. *Mol Biol Evol* 15:427–441.
- Hammer MF, Karafet TM, Redd AJ, Jarjanazi H, Santachiara-Benerecetti S, Soodyall H, Zegura SL. 2001. Hierarchical patterns of global human Y-chromosome diversity. *Mol Biol Evol* 18:1189–1203.
- Heyer E, Sibert A, Austerlitz F. 2005. Cultural transmission of fitness: Genes take the fast lane. *Trends Genet* 21:234–239.
- Hurles ME, Veitia R, Arroyo E, Armenteros M, Bertranpetit J, Perez-Lezaun A, Bosch E, Shlumukova M, Cambon-Thomsen A, McElreavey K, Lopez De Munain A, Rohl A, Wilson IJ, Singh L, Pandya A, Santos FR, Tyler-Smith C, Jobling MA. 1999. Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of a Y-chromosomal DNA polymorphism. *Am J Hum Genet* 65:1437–1448.
- Jobling MA, Tyler-Smith C. 2003. The human Y chromosome: An evolutionary marker comes of age. *Nat Rev Genet* 4:598–612.
- Kasperaviciute D, Kucinskas V, Stoneking M. 2004. Y chromosome and mitochondrial DNA variation in Lithuanians. *Ann Hum Genet* 68:438–452.
- Kayser M, Lao O, Anslinger K, Augustin C, Bargel G, Edelmann J, Elias S, Heinrich M, Henke J, Henke L, Hohoff C, Illing A, Jonkisz A, Kuzniar P, Lebioda A, Lessig R, Lewicki S, Maciejewska A, Monies DM, Pawlowski R, Poetsch M, Schmid D, Schmidt U, Schneider PM, Stradmamn-Bellinghausen B, Szibor R, Wegener R, Wozniak M, Zoledziwska M, Roewer L, Dobosz T, Ploski R. 2005. Significant genetic differentiation between Poland and Germany follows present-day political borders as revealed by Y-chromosome analysis. *Hum Genet* 117:428–443.
- King R, Underhill PA. 2002. Congruent distribution of Neolithic painted pottery and ceramic figurines with Y-chromosome lineages. *Antiquity* 76:707–714.
- Kivisild T, Rootsi S, Metspalu M, Mastana S, Kaldma K, Parik J, Metspalu E, Adojaan M, Tolk HV, Stepanov V, Golge M, Usanga E, Papiha SS, Cinnioglu C, King R, Cavalli-Sforza L, Underhill PA, Vilems R. 2003. The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *Am J Hum Genet* 72:313–332.
- Kristiansen K. 1998. Europe before history. Cambridge, UK: Cambridge University Press.
- Kruta V. 2000. Les Celtes Histoire et dictionnaire. Paris: Robert Laffont S A.
- Kwok C, Tyler-Smith C, Mendonca BB, Hughes I, Berkovitz GD, Goodfellow PN, Hawkins JR. 1996. Mutation analysis of the 2 kb 5' to SRY in XY females and XY intersex subjects. *J Med Genet* 33:465–468.
- Luca F, Basile M, Di Giacomo F, Novelletto A. 2005. Independent methods for evolutionary genetic dating provide insights into Y-chromosomal STR mutation rates confirming data from direct father-son transmissions. *Hum Genet* 118:153–165.
- Malaspina P, Cruciani F, Santolamazza P, Torroni A, Pangrazio A, Akar N, Bakalli V, Brdicka R, Jaruzelska J, Kozlov A, Malyarchuk B, Mehdi SQ, Michalodimitrakis E, Varesi L, Memmi MM, Vona G, Vilems R, Parik J, Romano V, Stefan M, Stenico M, Terrenato L, Novelletto A, Scozzari R. 2000. Patterns of male-specific inter-population divergence in Europe West Asia and North Africa. *Ann Hum Genet* 64:395–412.
- Malaspina P, Tsopanomalou M, Duman T, Stefan M, Silvestri A, Rinaldi B, Garcia O, Giparaki M, Plata E, Kozlov AI, Barbujani G, Vernesi C, Papola F, Ciavarella G, Kovatchev D, Kerimova MG, Anagnou N, Gavrila L, Veneziano L, Akar N, Loutradis A,



- Michalodimitrakis EN, Terrenato L, Novelletto A. 2001. A multi-step process for the dispersal of a Y chromosomal lineage in the Mediterranean area. *Ann Hum Genet* 65:339–349.
- Marjanovic D, Fornarino S, Montagna S, Primorac D, Hadziselimovic R, Vidovic S, Pojskic N, Battaglia V, Achilli A, Drobnic K, Andjeljinovic S, Torroni A, Santachiara-Benerecetti AS, Semino O. 2005. The peopling of modern Bosnia-Herzegovina: Y-chromosome haplogroups in the three main ethnic groups. *Ann Hum Genet* 69:757–763.
- Paracchini S, Arredi B, Chalk R, Tyler-Smith C. 2002. Hierarchical high throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. *Nucleic Acids Res* 30:e27.
- Pericic M, Lauc LB, Klaric IM, Rootsi S, Janicijevic B, Rudan I, Terzic R, Colak I, Kvesic A, Popovic D, Sijacki A, Behluli I, Dordevic D, Efremovska L, Bajec DD, Stefanovic BD, Villems R, Rudan P. 2005. High-resolution phylogenetic analysis of southeastern Europe traces major episodes of paternal gene flow among Slavic populations. *Mol Biol Evol* 22:1964–1975.
- Piggott S. 1965. *Ancient Europe from the beginnings of agriculture to classical antiquity*. Edinburgh: Edinburgh University Press.
- Quintana-Murci L, Semino O, Poloni ES, Liu A, Van Gijn M, Passarino G, Brega A, Nasidze IS, Maccioni L, Cossu G, al-Zahery N, Kidd JR, Kidd KK, Santachiara-Benerecetti AS. 1999. Y-chromosome specific YCAII, DYS19 and YAP polymorphisms in human populations: A comparative study. *Ann Hum Genet* 63:153–166.
- Renfrew C. 1979. *Before civilization the radiocarbon revolution and prehistoric Europe*. Cambridge, UK: Cambridge University Press.
- Roewer L, Croucher PJ, Willuweit S, Lu TT, Kayser M, Lessig R, de Knijff P, Jobling MA, Tyler-Smith C, Krawczak M. 2005. Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. *Hum Genet* 116:279–291.
- Rootsi S, Magri C, Kivisild T, Benuzzi G, Help H, Bermisheva M, Kutuev I, Barac L, Pericic M, Balanovsky O, Pshenichnov A, Dion D, Grobei M, Zhivotovsky LA, Battaglia V, Achilli A, Al-Zahery N, Parik J, King R, Cinnioglu C, Khusnutdinova E, Rudan P, Balanovska E, Scheffrahn W, Simonescu M, Brehm A, Goncalves R, Rosa A, Moisan JP, Chaventre A, Ferak V, Furedi S, Oefner PJ, Shen P, Beckman L, Mikerezi I, Terzic R, Primorac D, Cambon-Thomsen A, Krumina A, Torroni A, Underhill PA, Santachiara-Benerecetti AS, Villems R, Semino O. 2004. Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. *Am J Hum Genet* 75:128–137.
- Rosser ZH, Zerjal T, Hurler ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Corte-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Golge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskaskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Norby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previdere C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA. 2000. Y-chromosomal diversity within Europe is clinal and influenced primarily by geography rather than language. *Am J Hum Genet* 66:1526–1543.
- Schneider S, Roessli D, Excoffier L. 2000. *ARLEQUIN v 2000: A software for population genetics data analysis*. Switzerland: Genetics and Biometry Laboratory, University of Geneva.
- Semino O, Magri C, Benuzzi G, Lin AA, Al-Zahery N, Battaglia V, Maccioni L, Triantaphyllidis C, Shen P, Oefner PJ, Zhivotovsky LA, King R, Torroni A, Cavalli-Sforza LL, Underhill PA, Santachiara-Benerecetti AS. 2004. Origin diffusion and differentiation of Y-chromosome haplogroups E and J: Inferences on the Neolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet* 74:1023–1034.
- Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiae M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA. 2000. The genetic legacy of Paleolithic *Homo sapiens* in extant Europeans: A Y chromosome perspective. *Science* 290:1155–1159.
- Sokal RR, Oden NL. 1978. Spatial autocorrelation analysis in biology. I Methodology. *Biol J Linn Soc* 10:199–228.
- Stefan M, Stefanescu G, Gavrilu L, Terrenato L, Jobling MA, Malaspina P, Novelletto A. 2001. Y chromosome analysis reveals a sharp genetic boundary in the Carpathian region. *Eur J Hum Genet* 9:27–33.
- Stumpf MPH, Goldstein DB. 2001. Genealogical and evolutionary inference with the human Y chromosome. *Science* 291:1738–1742.
- Torroni A, Achilli A, Macaulay V, Richards M, Bandelt HJ. 2006. Harvesting the fruit of the human mtDNA tree. *Trends Genet* 22:339–345.
- Torroni A, Bandelt HJ, D'Urbano L, Lahermo P, Moral P, Sellitto D, Rengo C, Forster P, Savontaus ML, Bonne-Tamir B, Scozzari R. 1998. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am J Hum Genet* 62:1137–1152.
- Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ. 1997. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* 7:996–1005.
- Underhill PA, King RJ, Chow C-ET, Kivisild T, Rootsi S. 2005. New phylogenetic relationships for Y chromosome haplogroup I: Reappraising its pattern of regional subdivision. In: Mellars P, Stringer C, Bar-Yosef O, Boyle K, editors. *Rethinking the human revolution: New behavioural and biological perspectives on the origin and dispersal of modern humans*. Cambridge, UK: McDonald Institute for Archaeological Research. p 73–75.
- Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* 65:43–62.
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonne-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Davis RW, Feldman MW, Cavalli-Sforza LL, Oefner PJ. 2000. Y chromosome sequence variation and the history of human populations. *Nat Genet* 26:358–361.
- Weale ME, Weiss DA, Jager RF, Bradman N, Thomas MG. 2002. Y chromosome evidence for Anglo-Saxon mass migration. *Mol Biol Evol* 19:1008–1021.
- Wilson IJ, Balding DG. 1998. Genealogical inference from microsatellite data. *Genetics* 150:499–510.
- Y Chromosome Consortium. 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* 12:339–348.
- Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefelhovel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjidmaa D, Sajantila A, Salo P, Crawford MH, Ginter EK, Evgrafov OV, Tyler-Smith C. 1997. Genetic relationships of Asians and Northern Europeans revealed by Y-chromosomal DNA analysis. *Am J Hum Genet* 60:1174–1183.
- Zerjal T, Xue Y, Bertorelle G, Wells RS, Bao W, Zhu S, Qamar R, Ayub Q, Mohyuddin A, Fu S, Li P, Yuldasheva N, Ruzibakiev R, Xu J, Shu Q, Du R, Yang H, Hurler ME, Robinson E, Gerelsaikhan T, Dashnyam B, Mehdi SQ, Tyler-Smith C. 2003. The genetic legacy of the Mongols. *Am J Hum Genet* 72:717–721.
- Zhivotovsky LA, Underhill PA, Cinnioglu C, Kayser M, Morar B, Kivisild T, Scozzari R, Cruciani F, Destro-Bisol G, Spedini G, Chambers GK, Herrera RJ, Yong KK, Gresham D, Tournev I, Feldman MW, Kalaydjieva L. 2004. The effective mutation rate at Y chromosome short tandem repeats with application to human population-divergence time. *Am J Hum Genet* 74:50–61.