Current Biology 17, 43–48, January 9, 2007 ©2007 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2006.10.058

Report

# From Social to Genetic Structures in Central Asia

Raphaëlle Chaix,[1,2,*] Lluís Quintana-Murci,[3] Tatyana Hegay,[4] Michael F. Hammer,[5] Zahra Mobasher,[5] Frédéric Austerlitz,[6] and Evelyne Heyer[1]
[1] Unité d'Eco-Anthropologie
Centre National de la Recherche UMR 5145
Université Paris 7
Musée de l'Homme
Paris 75116
France
[2] Department of Statistics
University of Oxford
OX1 3TG Oxford
United Kingdom
[3] Centre National de la Recherche FRE 2849
Institut Pasteur
Paris 75015
France
[4] Institute of Immunology
Academy of Sciences
Tashkent 700060
Uzbekistan
[5] Department of Ecology and Evolutionary Biology
Division of Biotechnology
University of Arizona
Tucson, Arizona 85721
[6] Laboratoire Ecologie Systématique et Evolution
Centre National de la Recherche Scientifique UMR 8079
Université Paris-Sud
Orsay 91405
France

## Summary

**Pastoral and farmer populations, who have coexisted in Central Asia since the fourth millennium B.C. [1], present not only different lifestyles and means of subsistence but also various types of social organization. Pastoral populations are organized into so-called descent groups (tribes, clans, and lineages) and practice exogamous marriages (a man chooses a bride in a different lineage or clan). In Central Asia, these descent groups are patrilineal: The children are systematically affiliated with the descent groups of the father. By contrast, farmer populations are organized into families (extended or nuclear) and often establish endogamous marriages with cousins [2–4]. This study aims at better understanding the impact of these differences in lifestyle and social organization on the shaping of genetic diversity. We show that pastoral populations exhibit a substantial loss of Y chromosome diversity in comparison to farmers but that no such a difference is observed at the mitochondrial-DNA level. Our analyses indicate that the dynamics of patrilineal descent groups, which implies different male and female sociodemographic histories, is responsible for these sexually-asymmetric genetic patterns. This molecular signature of the pastoral social organization disappears over a few centuries only after conversion to an agricultural way of life.**

## Results and Discussion

We compared the genetic diversity of the HVS-1 sequence from the maternally inherited mitochondrial DNA (mtDNA) in 12 pastoral and nine farmer populations from Central Asia, as well as the diversity of six short tandem repeats (STRs) from the nonrecombining region of the paternally inherited Y chromosome (NRY) in 11 pastoral and seven farmer populations. Uzbek populations, who used to be pastoral nomads, were considered in this study as farmers. Indeed, they settled in the different parts of Uzbekistan from the 16$^{th}$ century, and since then, their social organization underwent a transition characterized by a considerable loss of the social importance of descent groups' organization. In addition, some groups of Uzbeks have adopted endogamous marriages and other aspects of the social organization that characterizes farmers' groups (S. Jacquesson, personal communication).

For both genetic systems, we compared different estimators of genetic diversity and demographic growth between pastoral and farmer populations (Tables 1 and 2). In the case of the mtDNA sequence, the heterozygosity $H$ and the mean number of pairwise differences $\pi$, which both estimate within-population diversity, were high in pastoral populations (median $H$ = 0.99, median $\pi$ = 5.29) and in farmer populations (median $H$ = 0.99, median $\pi$ = 5.32), with no significant difference between these two groups of populations for both $H$ and $\pi$ (bilateral Wilcoxon test, $p > 0.1$ for both statistics). We found a low level of genetic differentiation among pastoral populations as well as among farmer populations ($\phi_{st}$ = 0.01 in both cases, $p > 0.1$). The relationship between geographic and genetic distances in both groups of populations is shown in Figure S1 in the Supplemental Data available with this article online. Moreover, both groups of populations exhibited significantly negative Tajima's $D$ (−1.90 and −1.76, respectively, in pastoral and farmer populations, $p > 0.1$), which is a signature of demographic growth in the case of neutrally evolving genetic systems. In contrast with mtDNA data, the heterozygosity ($H$) inferred from Y chromosome STRs in pastoral populations was significantly lower than in farmer populations (0.86 and 0.99, respectively, $p < 0.01$). Similarly, $\pi$ was lower in pastoral than in farmer populations (2.86 and 3.59, respectively, $p < 0.01$). In addition, pastoral populations present a much higher level of population differentiation compared to that observed among farmer populations ($R_{st}$ = 0.19 and 0.06, respectively, $p < 0.01$). This higher differentiation is not the result of higher geographic distances among pastoral populations, as

*Correspondence: chaix@stats.ox.ac.uk

Table 1. Sample Descriptions and Estimators of Genetic Diversity from the mtDNA Sequence

| Population | n | Location | Long | Lat | *H* | $\pi$ | *D* | *pD* | $P_S$ | *C* | Origin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Karakalpaks | 20 | Uzbekistan | 58 | 43 | 0.99 | 5.29 | −1.95 | 0.01 | 0.90 | 1.05 | [21] |
| Karakalpaks (On Tört Uruw) | 53 | Uzbekistan/Turkmenistan border | 60 | 42 | 0.99 | 5.98 | −1.92 | 0.01 | 0.70 | 1.20 | This study |
| Karakalpaks (Qongirat) | 55 | Karakalpakia | 59 | 43 | 0.99 | 5.37 | −2.01 | 0.01 | 0.82 | 1.15 | This study |
| Kazakhs | 50 | Karakalpakia | 63 | 44 | 0.99 | 5.23 | −1.97 | 0.01 | 0.88 | 1.11 | This study |
| Kazakhs | 55 | Kazakhstan | 80 | 45 | 0.99 | 5.66 | −1.87 | 0.01 | 0.69 | 1.25 | [22] |
| Kazakhs | 20 |  | 68 | 42 | 1.00 | 5.17 | −1.52 | 0.05 | 1.00 | 1.00 | [21] |
| Kyrgyzs | 20 | Kyrgyzstan | 74 | 41 | 0.97 | 5.29 | −1.38 | 0.06 | 0.55 | 1.33 | [21] |
| Kyrgyzs (Sary-Tash) | 47 | South Kyrgyzstan, Pamir | 73 | 40 | 0.97 | 5.24 | −1.95 | 0.01 | 0.49 | 1.52 | [22] |
| Kyrgyzs (Talas) | 48 | North Kyrgyzstan | 72 | 42 | 0.99 | 5.77 | −1.65 | 0.02 | 0.77 | 1.14 | [22] |
| Turkmen | 51 | Uzbekistan/Turkmenistan border | 59 | 42 | 0.98 | 5.48 | −1.59 | 0.04 | 0.53 | 1.42 | This study |
| Turkmen | 41 | Turkmenistan | 60 | 39 | 0.99 | 5.20 | −2.07 | 0.00 | 0.73 | 1.21 | [23] |
| Turkmen | 20 |  | 59 | 40 | 0.98 | 5.28 | −1.71 | 0.02 | 0.75 | 1.18 | [21] |
| Dungans | 16 | Kyrgyzstan | 78 | 41 | 0.94 | 5.27 | −1.23 | 0.12 | 0.31 | 1.60 | [21] |
| Kurds | 32 | Turkmenistan | 59 | 39 | 0.97 | 5.61 | −1.35 | 0.05 | 0.41 | 1.52 | [23] |
| Uyghurs | 55 | Kazakhstan | 82 | 47 | 0.99 | 5.11 | −1.91 | 0.01 | 0.62 | 1.28 | [22] |
| Uyghurs | 16 | Kyrgyzstan | 79 | 42 | 0.98 | 4.67 | −1.06 | 0.15 | 0.63 | 1.23 | [21] |
| Uzbeks (Nord) | 40 | Karakalpakia | 60 | 43 | 0.99 | 5.49 | −2.03 | 0.00 | 0.68 | 1.21 | This study |
| Uzbeks (South) | 42 | Uzbekistan: Surkhandarya | 67 | 38 | 0.99 | 5.07 | −1.96 | 0.01 | 0.81 | 1.14 | [23] |
| Uzbeks (South) | 20 | Uzbekistan | 66 | 40 | 0.99 | 5.33 | −1.82 | 0.02 | 0.90 | 1.05 | [21] |
| Uzbeks (Khorezm) | 20 | Uzbekistan: Khorezm | 61 | 42 | 0.98 | 5.32 | −1.62 | 0.04 | 0.70 | 1.18 | [21] |
| Tajiks (Yagnobi) | 20 |  | 71 | 39 | 0.99 | 5.98 | −1.76 | 0.02 | 0.90 | 1.05 | [21] |

The pastoral populations are in the gray area; the farmer populations are in the white area. n, sample size; long, longitude; lat, latitude; *H*, heterozygosity; $\pi$, mean number of pairwise differences; *D*, Tajima's D; *pD*, probability that *D* is significantly different from zero; Ps, proportion of singletons; *C*, mean number of individuals carrying the same mtDNA sequence.

presented in Figure S2. The rate of demographic growth, which was estimated with the coalescent-based software Batwing [5, 6], was found to be lower in pastoral than in farmer populations, the difference being marginally significant ($r$ = 1.004 and 1.008, respectively, p = 0.056).

Altogether, our mtDNA results indicated that both farmer and pastoral populations show high levels of within-population genetic diversity and low levels of among-population differentiation, and both populations are experiencing a rapid demographic growth. By contrast, Y chromosome data revealed substantial differences between the two groups of populations, with pastoral populations exhibiting significantly lower levels of within-population diversity, significantly higher levels of among-population differentiation, and a tendency toward a lower rate of demographic growth compared to farmer populations. To understand the discrepancy observed in the two uniparentally inherited genetic systems and especially the loss of Y chromosome diversity in pastoral populations, we investigated the distribution of genetic diversity within populations by performing multidimensional scaling analyses (MDS) based on matrices of distance between the Y STR haplotypes or between the mitochondrial sequences of a sample, including ethnological information when available. In the case of the Y chromosome, we used as input distance the sum of squared size difference between STR haplotypes.

Table 2. Sample Descriptions and Estimators of Genetic Diversity from the Y Chromosome STRs

| Population | *n* | Location | Long | Lat | *H* | $\pi$ | r | $P_S$ | *C* | Origin |
|---|---|---|---|---|---|---|---|---|---|---|
| Karakalpaks (On Tört Uruw) | 54 | Uzbekistan/Turkmenistan border | 60 | 42 | 0.86 | 3.40 | 1.002 | 0.24 | 2.84 | [10] |
| Karakalpaks (Qongïrat) | 54 | Karakalpakia | 59 | 43 | 0.91 | 3.17 | 1.003 | 0.28 | 2.35 | [10] |
| Kazakhs | 50 | Karakalpakia | 63 | 44 | 0.85 | 2.36 | 1.004 | 0.16 | 2.78 | [10] |
| Kazakhs | 38 | Kazakhstan: Almata, KatonKaragay Karatutuk, Rachmanovsky Kluchi | 68 | 42 | 0.78 | 2.86 | 1.004 | 0.26 | 2.71 | [24] |
| Kazakhs | 49 | South East Kazakhstan | 77 | 40 | 0.69 | 1.56 | 1.012 | 0.22 | 3.06 | [25] |
| Kyrgyzs | 41 | Kyrgyzstan: central (mixed) | 74 | 41 | 0.88 | 2.47 | 1.004 | 0.41 | 1.86 | [24] |
| Kyrgyzs (Sary-Tash) | 43 | South Kyrgyzstan, Pamir | 73 | 40 | 0.45 | 1.30 | 1.003 | 0.12 | 4.78 | [25] |
| Kyrgyzs (Talas) | 41 | North Kyrgyzstan | 72 | 42 | 0.94 | 3.21 | 1.002 | 0.39 | 1.78 | [25] |
| Mongolians | 65 | Mongolia: Ulaanbaatar | 90 | 49 | 0.96 | 3.37 | 1.009 | 0.38 | 1.81 | [24] |
| Turkmen | 51 | Uzbekistan/Turkmenistan border | 59 | 42 | 0.67 | 1.84 | 1.006 | 0.27 | 3.00 | [10] |
| Turkmen | 21 | Turkmenistan: Ashgabat | 59 | 40 | 0.89 | 3.34 | 1.006 | 0.48 | 1.62 | [24] |
| Dungans | 22 | Kyrgyzstan: Alexandrovka, Osh | 78 | 41 | 0.99 | 4.13 | 1.005 | 0.82 | 1.10 | [24] |
| Kurds | 20 | Turkmenistan: Bagyr | 59 | 39 | 0.99 | 3.59 | 1.009 | 0.80 | 1.11 | [24] |
| Uyghurs | 33 | Kazakhstan: Almaty, Lavar | 79 | 42 | 0.99 | 3.72 | 1.007 | 0.67 | 1.22 | [24] |
| Uyghurs | 39 | South East Kazakhstan | 79 | 43 | 0.99 | 3.79 | 1.008 | 0.77 | 1.15 | [25] |
| Uzbeks (North) | 40 | Karakalpakia | 60 | 43 | 0.96 | 3.42 | 1.005 | 0.48 | 1.54 | [10] |
| Uzbeks (South) | 28 | Uzbekistan: Kashkadarya | 66 | 40 | 1.00 | 3.53 | 1.008 | 0.93 | 1.04 | [24] |
| Tajiks (Yagnobi) | 22 | Tajikistan: Penjikent | 71 | 39 | 0.87 | 2.69 | 1.012 | 0.45 | 1.69 | [24] |

The pastoral populations are in the gray area; the farmer populations are in the white area. n, sample size; long, longitude; lat, latitude; *H*, heterozygosity; $\pi$, mean number of pairwise differences; r, growth-rate estimate with Batwing; $P_S$, proportion of singletons; *C*, mean number of individuals carrying the same Y STR haplotype.

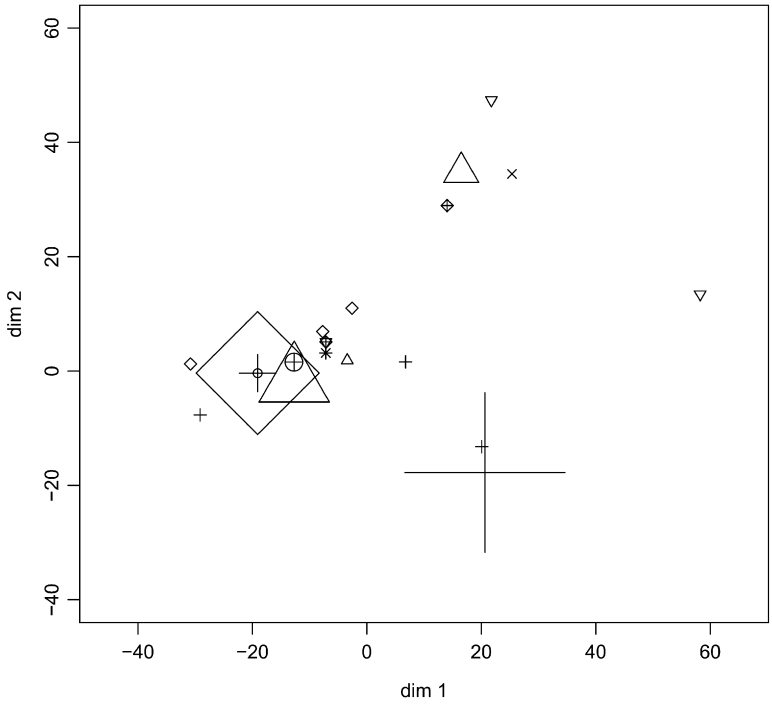


Figure 1. Multidimensional Scaling Analysis based on the Matrix of Distance between Y STR Haplotypes in a Specific Pastoral Population: The Karakalpaks On Tört Uruw

Each point represents a given Y STR haplotype, and the represented distance between two Y STR haplotypes is the bidimensional projection of the sum of squared size difference between these two haplotypes. The different symbols represent the different clans, e.g., a cross for the Ömir clan of the Keneges tribe. The size of the points is proportional to the number of individuals from the same clan carrying the same haplotype. Thirteen individuals from the Ömir clan of the Keneges tribe are carrying the same haplotype, ten individuals from the Qarasyraq clan of the Mangyt tribe are carrying the same haplotype (diamond), and ten individuals from the Nökis clan of the Keneges tribe are carrying the same haplotype (triangle). Stress = 10.1%.

The analyses of pastoral populations, as presented here in the case of the Karakalpaks On Tört Uruw (Figure 1), revealed the existence of clusters of individuals belonging to the same clan and having exactly the same Y STR haplotype. We termed these clusters of genetically identical individuals belonging to the same descent group (lineage or clan) "identity cores." In the present example, one can observe for the Y chromosome, an identity core in the Ömir clan of the Keneges tribe (grouping 13 individuals) and another identity core in the Qarasyraq clan of the Mangyt tribe (grouping ten individuals). The network illustrating the phylogenetic relationships between the Y STR haplotypes of this sample is provided in Figure S3. One can observe that a clan is made of an identity core and a few haplotypes at minor frequencies, and these haplotypes arise either by mutation (those located next to the identity core in the MDS) or by migration and adoption (those located far away from the identity core).

The observed identity cores were specific to Y chromosome data and mainly restricted to pastoral populations. Indeed, very few of them were observed for the Y haplotypes of farmers or the mtDNA data of either pastoral or farmer populations (see in Figure 2 the MDS from mtDNA data for the Karakalpaks On Tört Uruw). Moreover, the average number of individuals carrying the same haplotype (*C*) was much higher for Y chromosome data in pastoral populations than in farmer populations (2.71 and 1.15, respectively, $p < 0.01$, see Table 2 for the values per population). It was also higher than the average *C* values for mtDNA data in both group of populations (1.19 and 1.21, respectively, in both cases, see Table 1 for the values per population). Moreover, when considering Y chromosome data in pastoral populations, we observed a lower proportion of singletons ($P_S$) in comparison to farmer populations (27% and 77% of singletons, respectively, $p < 0.01$) as well as in comparison to the proportion of singletons observed from mtDNA data in both groups of populations (74% and 68%, respectively).

The presence of identity cores for Y chromosome data in pastoral populations is unlikely to be the result of sampling bias because special attention was paid to sample only unrelated men in all populations (i.e., all individuals were unrelated for at least two generations back in time). These identity cores could instead be the direct consequence of the internal dynamics of patrilineal descent groups in pastoral populations (the population is divided in tribes, each tribe in clans, and each clan in lineages). Indeed, three aspects of this internal dynamics may explain our observations: First, descent groups are regularly submitted to lineal fissions (also called lineage segmentations): When a descent group (lineage or clan) reaches a given genealogical depth, it splits into two new descent groups. An important point is that these fissions do not occur at random. Indeed, closely related men tend to remain in the same descent group, and each newly created group will thus have a common paternal ancestor that is more recent than the paternal ancestor of its parent descent group. This phenomenon certainly reduces the Y chromosome genetic diversity of the newly created descent groups, whereas if the fission had occurred randomly, it would not be the case (see Smouse [7] for a theoretical study of the impact of lineal fissions in the Yanomama). Secondly, this process is reinforced by the strong genetic drift that is likely to occur in these small descent groups, and the virtual absence of migration of men among descent groups further exacerbates the strength of genetic drift. Thus, although some Y chromosome haplotypes might go to extinction, others might reach rapidly high frequencies within a descent group and thus give rise to the so-called identity cores. Third, the demographic stochasticity inherent to small groups of individuals is likely to lead some descent groups to extinction and thus

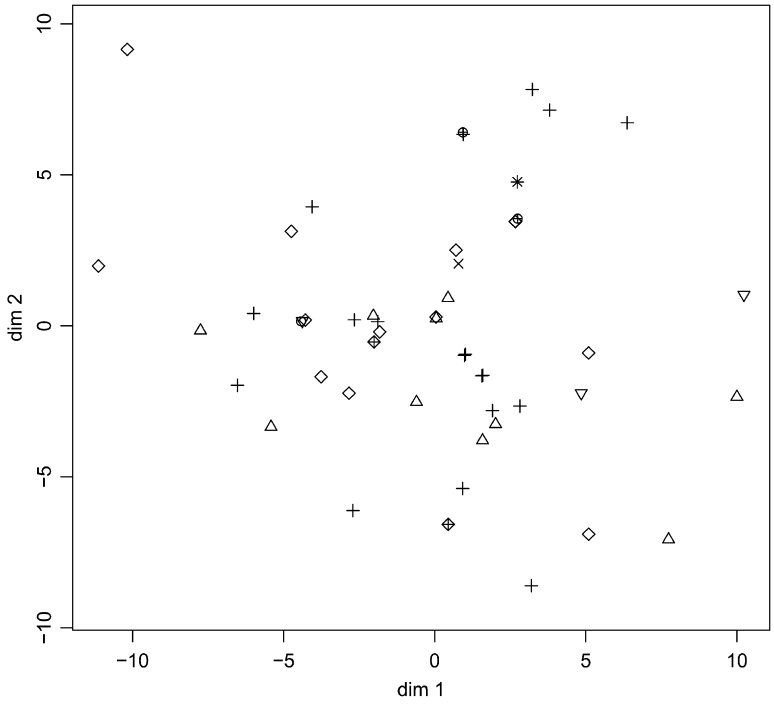


Figure 2. MDS Based on the Number of Differences between the Mitochondrial Sequence in the Same Pastoral Population: The Karakalpaks On Tört Uruw

Stress = 19.8%. The symbols are the same as in Figure 1.

reduce the Y chromosome genetic diversity of the population when considered as a whole. In this view, these demographic and genetic processes may thus explain not only the existence of the identity cores at the Y chromosome level in pastoral populations but also their overall lower Y chromosome diversity compared to farmer populations. Accordingly, these processes might be the basis of the higher genetic differentiation for the Y chromosome, as observed among pastoral populations. Other phenomena might have also contributed to the observed patterns of diversity in pastoral populations. Indeed, pastoral men could have experienced a higher variance of reproductive success than farmers. Polygyny was traditionally practiced in pastoral populations, with the number of wives often depending on the wealth of the husband. In addition, it could have been practiced among Central Asian Muslim populations of farmers since the 10$^{th}$ century. However, polygyny was forbidden in former USSR [4] and was not observed when we collected demographic data in four pastoral populations (Qongïrat, On Tört Uruw, Kazakhs, and Turkmen) nor in the Uzbeks (farmers). Moreover, we did not observe any substantial differences in the variance of the number of effective children between the Uzbeks and the four pastoral populations (results presented in Table S3). Finally, in pastoral populations, local adaptation processes giving advantage to some men carrying particular Y chromosome haplotype might also have contributed to the observed diversity patterns in these populations. However, this hypothesis was not supported by our demographic data because we did not observe any significant differences in reproductive success between men belonging to identity cores and men exhibiting rare haplotypes (data not shown).

The similarity of the levels of mitochondrial diversity observed in pastoral and farmer populations might be a consequence of the complex rules of exogamy in practice in pastoral populations. Traditionally, a man must choose a bride in such a way that he will not be sharing with her a common ancestor on the paternal lineage for a given number of past generations. This number is usually close to the genealogical depth of a lineage (five to ten generations depending on the population, for example, seven generations for the Kazakhs) so that in practice, the bride usually belongs to a different lineage (the so-called lineage exogamy) [4, 8]. This rule has at times been replaced by a clanic exogamy, such as in the Karakalpaks, with the obligation of choosing a bride in a different clan (but in the same tribe) [2]. These rules of exogamy imply that, at each generation, a significant number of women migrate from one descent group to another. This intense migration process reduces the strength of genetic drift within each descent group for mtDNA, whose polymorphism is thus not affected by population structure.

An analogy can be made between the pastoral social organization and the SLOSS (Single Large or Several Small) model, which compares the evolution of the genetic diversity in a panmictic population (Single Large, thereafter denoted SL) and in a population of same size divided in several isolated demes (Several Small, thereafter denoted SS) [9]. This model predicts that each deme of a SS population will rapidly fix a different allele. Thus, on a short-term basis, this model predicts a higher diversity in a SS population, but on a long-term basis, because of the extinction of local demes, it predicts a reduction of diversity in such populations compared to an SL population. The pastoral population resembles, as far as the men are concerned, to an SS population (each descent group being a deme, with no migration between demes), which in the long-term loses Y chromosome diversity because of the joint action of genetic drift and the lineage-extinction process (heterozygosity $H$ = 0.86, mean number of individuals carrying

the same haplotype $C = 2.71$, percentage of singletons $P_S = 27\%$). On the other hand, because of high interdeme female migration rates, the female pastoral population is similar to a SL population and thus preserves in the long term a high mitochondrial diversity ($H = 0.99$, $C = 1.19$, $P_S = 74\%$).

Finally, a closer investigation of the diversity in Uzbeks yielded insights on the demographic processes resulting from a transition in lifestyle. Indeed, since the 16th century, the different Uzbeks groups, who used to be pastoral nomads, have progressively adopted a sedentary way of life based on agriculture. They show diversity values ($H$, $\pi$, $P_S$, $C$) for the Y chromosome similar to that of the other farmer populations (see Table 2); this means that after a few centuries of settlement, the Y chromosome diversity of these two populations has not kept the genetic signature of the pastoral social organization. This is especially striking in the case of the Uzbeks from the South of Uzbekistan whose settlements date as far back as the 16th century compared to the Uzbeks from the North of the country (settlements dating 17–18th centuries). This pattern is also consistent with the clear reduction of genetic kinship (measured from Y chromosome STRs) for the Uzbeks within descent groups in comparison to Kazakhs, Turkmens, and Karakalpaks [10]. Two demographic processes may have acted jointly to yield this rapid transition: (1) a social transition in the Uzbeks, that is a progressive dissolution of their descent groups after the early loss of their nomadic way of life from the 16th century on [2] and, in the South of Uzbekistan, a reorganization in families establishing endogamous marriages, as in traditional farmers (S. Jacquesson, personal communication); (2) an intensification of gene flow from traditional farmer populations into the Uzbek populations. Indeed, the genetic differentiation between Uzbeks and farmer populations appears less than the differentiation between pastoral and farmer populations. For the Y STRs, the $R_{ST}$ between the North Uzbeks and the farmer populations amounted to 0.05 (and that between the South Uzbeks and the farmer populations amounted to 0.03), whereas the $R_{ST}$ between each of the eight pastoral populations and the farmer populations ranged from 0.06 to 0.39 (median equalled 0.11). For mitochondrial DNA, the genetic distances ($\phi_{st}$) between traditional farmers and the four Uzbeks populations tend also to be smaller (0–0.014, median = 0.005) than between traditional farmers and the 12 pastoral populations under study (0.001–0.047, median = 0.012). These observations are consistent with the theory of an actual demic agricultural diffusion [11], even at such microgeographic scale, involving real migrations of farmers followed by admixture rather than a simple diffusion of technologies, as also observed in Indian populations who recently converted to an agriculture-based way of life [12].

In conclusion, this study demonstrates how the cultural subdivision in patrilineal descent groups has left its footprints on Y chromosome diversity of pastoral populations without disturbing mitochondrial diversity. Indeed, the male population is experiencing a demographic history of lineal fissions of descent groups without subsequent migrations between descent groups, and this leads to the so-called identity cores and to a reduction of Y chromosome diversity. Conversely, the female population undergoes at each generation massive migration flows between descent groups (lineages or clans) as a result of the social rules of exogamy in practice in these populations and therefore prevents the social structure from imprinting mitochondrial structure. More specifically, the case study of Uzbeks' diversity clearly shows that such a molecular signature in Y chromosome diversity is short-lived and can disappear within a few centuries after the disintegration of descent groups.

Beyond the specific context of pastoral populations and Central Asia, this study more generally completes previous investigations of the impact of traditional social organizations on the shaping of genetic diversity. For example, a high genetic microdifferentiation between tribes, at the protein level, has been observed in Amerindians [13], and asymmetrical patterns of male and female dispersions have been observed among patrilocal and matrilocal populations from Northern Thailand [14]. Altogether, these studies clearly indicate that social organization and lifestyle are important factors influencing the evolution of genetic diversity in human populations. In this view, genome-wide sequencing and the genotyping of data in larger population panels are needed for further development of the "ethnogenetics" which uses genetic data to enhance our knowledge of the social organization of human populations, as a complement to ethnographic studies. Moreover, our study also underscores the necessity of understanding better the potential impact of social processes on the long-term evolution of our species: To which extent could social organization have modified the action of natural selection or the intensity of drift during recent human evolution? Further molecular data as well as new theoretical models are now needed for tackling such questions.

## Experimental Procedures

### DNA Samples and Genetic Markers

We studied the HVS-1 sequence from the maternally inherited mitochondrial DNA (mtDNA) in 12 pastoral and nine farmer populations from Central Asia, including five new samples from Karakalpakia (Karakalpaks On Tört Uruw, Qongirat, Kazakhs, Turkmen, and Uzbeks). Informed consent was obtained from all participants. A detailed description of all samples including geographical coordinates is provided in Table 1. The poly-C region, also named the "hyperhypervariable" region, located between sites 16179 and 16195, was excluded. We also studied six STRs (DYS388, DYS389I, DYS390, DYS391, DYS392, and DYS393) from the nonrecombining region of the paternally inherited Y chromosome (NRY) in 11 pastoral and seven farmer populations (detailed description in Table 2). The HVS-1 sequences and Y chromosome haplotypes corresponding to the five new samples from Central Asia (Karakalpaks On Tört Uruw, Qongirat, Kazakhs, Turkmen, and Uzbeks) are available in Tables S1 and S2. For the populations sampled in Karakalpakia, we also collected some ethnological information about each participant in the study including the name of his lineage, clan, and tribe as well as demographic data (number of effective sons and daughters of each participant in the study).

### Statistical Analyses

We used Arlequin software version 2001 [15] to estimate the heterozygosity ($H$) and the mean number of pairwise differences ($\pi$) for both genetic systems as well as the genetic distances between populations ($\phi_{st}$, based on the number of pairwise differences between mtDNA sequences and $R_{st}$, based on the sum of squared size differences between haplotypes of Y STRs [16]). In the case of the mtDNA sequence, we also computed Tajima's $D$ [17].

The absolute growth rate $r$ was estimated from Y STRs with Batwing [6, 7]. We assumed a model of continuous exponential growth (in this model, $N = N_0 r^t$, where $N$ is the present population size, $N_0$ is the initial size, and $t$ is the time measured in generations). Mutations were assumed to occur under a stepwise mutation model (SMM). Uniform prior distributions were used for $\theta$ and $\Omega$ ($\theta = 2N_e\mu$ and $\Omega = 2N_e\alpha$, where $N_e$ is the present effective size of populations, $\mu$ is the per-generation and per-microsatellite mutation rate and $\alpha$ is the exponential growth rate so that $\alpha = \ln(r)$). We assumed a mutation rate of 0.21% per microsatellite and per generation [18]. The average number of individuals carrying the same haplotype (*C*) and the proportion of singletons (*Ps*) were computed with a Perl script. To compare the two groups of populations, we estimated the median of each of these statistics within pastoral populations and within farmer populations. We also performed bilateral Wilcoxon tests [19] to assess whether the observed differences in diversity or differentiation between farmer and pastoral populations were significant. To investigate the distribution of diversity within population, we performed MDS (by using R version 2.3.1) based on matrix of distance (sum of squared size difference between Y STR haplotypes and number of pairwise differences between mtDNA sequences) computed by using a Perl script. We drew the phylogenetical relationships between YSTR haplotypes by using the reduced median method implemented in the software Network [20].

### Supplemental Data

Supplemental Data include three figures and three tables and can be found with this article online at http://www.current-biology.com/cgi/content/full/17/1/43/DC1/.

### Acknowledgments

We wish to thank all the participants in the study. We also acknowledge Svetlana Jacquesson for helpful discussions on ethnological data, Bruno Toupance and Julia Huggins for technical assistance, and the four anonymous reviewers for helpful comments. This work was supported by the Centre National de la Recherche grant project OHLL and the European Science Foundation project OMLL.

Received: August 16, 2006
Revised: October 4, 2006
Accepted: October 20, 2006
Published: January 8, 2007

### References

1. Cavalli-Sforza, L.L., Menozzi, P., and Piazza, A. (1994). The History and Geography of Human Genes (Princeton, NJ: Princeton University Press).
2. Jacquesson, S. (2002). Parcours ethnographiques dans l'histoire des deltas. In Karakalpaks et Autres Gens de l'Aral: Entre Rivages et Déserts, *Volume 10*, S. Jacquesson and V. Fourniau, eds. (Tachkent - Aix-en-Provence: Edisud), pp. 51–92.
3. Krader, L. (1963). Peoples of Central Asia, *Volume 26* (Bloomington, IN: Indiana University Publications).
4. Krader, L. (1963). Social Organization of the Mongol-Turkic Pastoral Nomads, *Volume 20* (The Hague: Indiana University Publications).
5. Wilson, I.J., and Balding, D.J. (1998). Genealogical inference from microsatellite data. Genetics *150*, 499–510.
6. Wilson, I.J., Weale, M.E., and Balding, D.J. (2003). Inferences from DNA data: Population histories, evolutionary processes, and forensic match probabilities. Journal of the Royal Statistical Society Series A *166*, 1–33.
7. Smouse, P.E., Vitzthum, V.J., and Neel, J.V. (1981). The impact of random and lineal fission of the genetic divergence of small human groups: A case study among the Yanomama. Genetics *98*, 179–197.
8. Ghasarian, C. (1996). Introduction à l'étude de la Parenté (Paris: Editions du Seuil).
9. Frankham, R., Ballou, J.D., and Briscoe, D.A. (2003). Introduction to Conservation Genetics (Cambridge: Cambridge University Press).
10. Chaix, R., Austerlitz, F., Khegay, T., Jacquesson, S., Hammer, M.F., Heyer, E., and Quintana-Murci, L. (2004). The genetic or mythical ancestry of descent groups: Lessons from the Y chromosome. Am. J. Hum. Genet. *75*, 1113–1116.
11. Ammerman, A., and Cavalli-Sforza, L.L. (1984). The Neolithic Transition and the Genetics of Population in Europe (Princeton, NJ: Princeton University Press).
12. Cordaux, R., Deepa, E., Vishwanathan, H., and Stoneking, M. (2004). Genetic evidence for the demic diffusion of agriculture to India. Science *304*, 1125.
13. Neel, J.V., and Ward, R.H. (1970). Village and tribal genetic distances among American Indians, and the possible implications for human evolution. Proc. Natl. Acad. Sci. USA *65*, 323–330.
14. Hamilton, G., Stoneking, M., and Excoffier, L. (2005). Molecular analysis reveals tighter social regulation of immigration in patrilocal populations than in matrilocal populations. Proc. Natl. Acad. Sci. USA *102*, 7476–7480.
15. Schneider, S., Roessli, D., and Excoffier, L. (2000). Arlequin ver. 2000: A software for Population Genetics Data Analysis (Switzerland: Genetics and Biometry Laboratory, University of Geneva).
16. Excoffier, L., Smouse, P.E., and Quattro, J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics *131*, 479–491.
17. Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics *123*, 585–595.
18. Heyer, E., Puymirat, J., Dieltjes, P., Bakker, E., and de Knijff, P. (1997). Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. Hum. Mol. Genet. *6*, 799–803.
19. Sokal, R.R., and Rohlf, F.J. (2001). Biometry (New York: W.H. Freeman and Company).
20. Bandelt, H.J., Forster, P., Sykes, B.C., and Richards, M.B. (1995). Mitochondrial portraits of human populations. Genetics *141*, 743–753.
21. Comas, D., Plaza, S., Wells, R.S., Yuldaseva, N., Lao, O., Calafell, F., and Bertranpetit, J. (2004). Admixture, migrations, and dispersals in Central Asia: Evidence from maternal DNA lineages. Eur. J. Hum. Genet. *12*, 495–504.
22. Comas, D., Calafell, F., Mateu, E., Perez-Lezaun, A., Bosch, E., Martinez-Arias, R., Clarimon, J., Facchini, F., Fiori, G., Luiselli, D., et al. (1998). Trading genes along the silk road: mtDNA sequences and the origin of central Asian populations. Am. J. Hum. Genet. *63*, 1824–1838.
23. Quintana-Murci, L., Chaix, R., Wells, R.S., Behar, D.M., Sayar, H., Scozzari, R., Rengo, C., Al-Zahery, N., Semino, O., Santachiara-Benerecetti, A.S., et al. (2004). Where west meets east: The complex mtDNA landscape of the southwest and Central Asian corridor. Am. J. Hum. Genet. *74*, 827–845.
24. Zerjal, T., Wells, R.S., Yuldasheva, N., Ruzibakiev, R., and Tyler-Smith, C. (2002). A genetic landscape reshaped by recent events: Y-chromosomal insights into central Asia. Am. J. Hum. Genet. *71*, 466–482.
25. Perez-Lezaun, A., Calafell, F., Comas, D., Mateu, E., Bosch, E., Martinez-Arias, R., Clarimon, J., Fiori, G., Luiselli, D., Facchini, F., et al. (1999). Sex-specific migration patterns in Central Asian populations, revealed by analysis of Y chromosome short tandem repeats and mtDNA. Am. J. Hum. Genet. *65*, 208–219.