12-1-2011

Afghan Genetic Mysteries

Bernard Dupaigne

National Museum of Natural History, CNRS UMR 7206, Paris, France, bdupaigne@mnhn.fr

Recommended Citation

Open access pre-print, subsequently published as Dupaigne, Bernard (2011) "Afghan Genetic Mysteries," Human Biology: Vol. 83: Iss. 6, Article 10.
Available at: http://digitalcommons.wayne.edu/humbiol_preprints/19

This Open Access Preprint is brought to you for free and open access by the WSU Press at DigitalCommons@WayneState. It has been accepted for inclusion in Human Biology Open Access Pre-Prints by an authorized administrator of DigitalCommons@WayneState.
Dear Editor,

I am a cultural anthropologist and I have focused a large part of my research on Afghanistan populations. To my surprise and despite the international attention given to this country in the last three decades, I have noted the absence of significant genetic studies intended to depict its complex population history. Available articles mainly concern those ethnic groups located across the borders, meaning that any conclusion about Afghanistan is an indirect inference, as only few Afghans living in Afghanistan, or recently emigrated from Afghanistan, have been sampled (Agarwal et al. 1976; Rahimi et al., 1977; Goedde et al., 1977a, 1977b; Hirth et al., 1979,1982; Benkmann et al. 1980; Berti et al. 2005Hohoff et al. 2006; Lacau et al. 2011).

Afghanistan is at the crossroad of many civilisations, where Central Asian and Indo-European populations (coming from the Iranian plateau and the Indian sub-continent) have met and sometimes admixed (Bruk 1955; Orywal 1983; Reichert 1986).

As Afghanistan has been practically isolated from all the neighbouring regions from the 18th century to the beginning of the 20th century, more than elsewhere today spoken languages well identify existing ethnic groups that are fully conscious of their identity. Therefore, linguistic diversity can be used as an excellent proxy to ethnic identity. In total, there are at least thirty-two different languages and dialects in Afghanistan, belonging to four different linguistic families. Twelve languages are specific to this country: Parachi; Munji, Sangeship-Eshkashimi (Badakhshan province); Kati, Askuni, Waigali, Prasuni (Nurestan province); Pashayi; Ningalami, Watapuri, Sawi, Tirahi (Farhadi 1969, 1970). For more
details, the generally accurate guide “The Ethnologue” can be consulted and, to this end, I will report the capitalized three-letter accession code they use in each section (Lewis 2009).

Generally, the definition of Afghan ethnic groups corresponds to usual ethnological criteria: a specific political or social organization, a marked preference for endogamous marriages, the feeling of sharing a common culture, values, and memories and a marked geographic continuity of the territory each group inhabits (Dupaigne 2002).

Before passing to a systematic description of the ethnic groups living in the country, I would like to highlight that, traditionally, it was possible to marry a woman of a group considered as inferior, while the reverse was impossible. Furthermore, while marriages between Turkmens and Tajiks or between Uzbeks, Tajiks and Arabs could be envisaged, marriages between Pashtuns tribes were unlikely between tribes far apart from each other and almost impossible with any other ethnic group.

Almost no Afghan ethnic group can be considered fully genetically homogeneous, in view of the wars, the deportation of populations, the alliances and its possible past integration in more powerful groups after a defeat. Further, the tradition of stronger groups to accept among them impoverished families or clans that worked unpaid in exchange of protection until a final integration may distort a trustworthy genetic inference about ancestry (Centlivres 1988).

I understand that researchers potentially interested in Afghans have been discouraged by the political instability of the region and by its cultural complexity. While I am unable to fix the first problem, I will describe the different groups, trying to provide a list of hypothesis that can be investigated with molecular markers and highlight the cases where such investigation is likely to be unsuccessful. Anyway, and most importantly, the chronology of the arrival of the different ethnic groups, largely unknown, would be enlightened by population genetics. Several useful studies could be achieved before the official departure of
allied forces scheduled for 2014. The foreign influence cannot be limited to that of conquering armies, and the scientific field open to research is immense.

Concerning the genetic sampling, the size and the ethnical diversity of the population of Kabul and other major cities would often suffice and travelling to remote regions would not be required. Moreover and without mentioning many well-equipped hospitals, the universities of Kabul, Jalalabad, Mazar-i Sharif, Herat, or Kapisa would be eager to cooperate.

While the reader may conclude that Afghanistan is just a virtual name corresponding to a territory inhabited by tribes having contrasting and independent political aspirations, my experience shows that Afghans of any ethnical belonging, although still attached to their ancient loyalties, consider themselves as citizens of a unique Afghan nation.

Before passing to the systematic description of the different ethnic groups, I would like to thank Pierre Darlu for providing useful references about genetic investigations and Nancy Wise for translating my text into English. If this letter persuades any anthropologist to start a genetic survey in Afghanistan, I will be glad to answer any inquiry sent through the editorial office of Human Biology. Meanwhile, figure 1 can be used to geographically locate some groups I mention.

**Major ethnic groups:**

**The Pashtuns** (PBU, PBT)

The Pashtuns, constituting around the 40% of the population, represent the dominant ethnic group in Afghanistan (“Afghan” generally is a synonym of “Pashtun”). They initiated the modern Afghan State in 1747 when, during a big tribal meeting (*loya djirga*) in Kandahar, Ahmad Shah, chief of the Abdalis tribe, was elected *emir* (king) of the southern tribes (Bellew...
1880, Caroe 1958). They live in the south and southeast of the country, but also in the “tribal belt” of West Pakistan. The Pashtuns (called Pathans in Pakistan) speak Pashto, a language belonging to the oriental group of Iranian languages, which was recognised in 1936 as one of the two official languages of Afghanistan (the other being the Farsi).

The existence of Pashtuns is first mentioned in 982 CE in the geographic treatise *Hodûd al-ʿAlam* (1937). This manuscript actually refers to those inhabiting West Pakistan, in the area of the Suleiman Mountains, but it does not mean that other Pashtun groups were not already living in Afghanistan. Originally followers of the Iranian fire religions (Zoroastrism, Mazdéism, Manichaeism), all Pashtuns converted to Islam between the 7th and the 11th century as a consequence of the Arab conquests.

Nowadays, the Pashtuns are grouped in two political confederations, called Durranis in the south (areas of Kandahar and Quetta) and Ghilzays in the east (areas of Nangarhar and Paktya). They are respectively composed by about ten main tribes (Alekzays, Alikozays, Andars, Atchazays, Barakzays, Ishakzays, Mohammedzays, Nurzays, Popalzays, Sadozays) and by about eight main tribes (Ahmadzays, Hotakis, Ishakzays, Kharotis, Musakhêls, Suleymankhêls, Tarakis, Zadrans). Some eastern tribes (Jajis, Kakars, Khogianis, Mangals, Safis, Turis) as well as the Afridis, Shinwaris, Mohmands and Waziris do not belong to such confederations (Lal, 1846).

I here provide a detailed picture of the social organization, because each Pashtun tribe (or clan) claims to descend from a single common male ancestor. This traditional belief might be genetically tested, even though the traditional habit of adoption, and the various conquests and tribal wars, are likely to complicate the picture.

We note that in the 19th century some British authors argued that the Pashtuns were the descendants of one of the ten lost tribes of Israel (Burnes 1834), but genetic research still have to find any evidence of such claim (Qamar 2002; Parfitt, 2003).
The political and religious group of the Taleban (literally, students of religion) is mainly, but not exclusively, constituted by Pashtuns.

THE PERSANOPHONES:

The Tajiks (PRS)
The Tajiks represent about 25-30% of the whole Afghan population. They are believed to be the most ancient inhabitants of Afghanistan (Masson 1964-1965; Dupree 1973). They are a sedentary population of Persian tradition, spread almost all over the country. Differently from Pashtuns, Tajiks are not organised in a tribal system. They are farmers, shopkeepers, civil servants and artisans, naming themselves by their region or valley of origin: Andarabis, Badakhchis, Heratis, Logaris, Panjcheris, Wakhis, etc. In Afghanistan, the Tajiks are generally known as Farsiwan (literally meaning: “who speak the Persian language”) because they speak Farsi (or Dari, the official name of the literary language). The Farsi is also the language of the central administration, medias, traders, and shares with Pashto the status of official language of the country.

Small ancient Tajiks groups (Pamiri Tajiks in Figure 1), different from each other and from the majority of other Tadjiks can be found in the mountain valleys of Pamir in the Northeast provinces of Badakhshan and Wakhan (Kussmaul, 1965). They speak archaic Iranian dialects, called Eshkashemi (SGL), Munji (MNJ), Rochani, Shughni (SGH), Yaghnobi, Wakhi (WBL). These isolates would permit a more refined genetic analysis of the original Tajik settlers (main Tajiks are often mixed), even though a loss of genetic diversity might be expected.

The Aymaqs (AIQ)
Living in the isolated mountains of west central Hindu-Kush and accounting for around 600,000 people, they are also called Farsiwan (speaking farsi), in spite of the fact that the word used to identify them, Aymaq, is of Mongol origin. They do not call themselves in such a way, preferring to use toponyms related to the original location of their tribe: Taymûris (around the cities of Herat and Qala-e no), Jamchidis (north of Herat) Taymanis, Ferozkohis (Mandersloot 1971). The kind of farsi they speak contains many Turkish terms borrowed from neighbouring populations. Differently from the Tajiks (another Farsiwan people), they are endogamous and remain attached to their tribal affiliations. They have a semi-nomadic lifestyle, and could be the descendants of the Ghorids who had established an empire from Afghanistan to India in the 12th and 13th centuries.

The Baluchs (BGN)

There are about 250,000 Baluchs in Afghanistan. They originally came from the north of the Caspian Sea. Arab geographers reported their presence during the 10th century in the province of Seistan (south-west of Afghanistan, between Iran and Pakistan) where they still live in their majority. Some smaller groups live in the northwest of Afghanistan and in Turkmenistan (Pehrson 1966). The Baluchs speak an ancient Iranian language and share a semi-nomadic lifestyle (Planhol 1993). Their tribal organisation is very strong and differentiated, most of them being from Rakhshani tribe. They are all Sunni Moslems, even those living in Iran.

The Ormuris (ORU) and Paratchi (PRC)

The Ormuris only consist of few families, either living around the town of Baraki-Barak in the province of Logar or near Kaniguram in the province of Waziristan (Pakistan). They speak Ormuri, an ancient language, nearly extinct, of the south-eastern Iranian group (Kieffer 1977).
Until recent times, the Ormuris did not exchange wives with their neighbours and their maternal lineages should be of interest.

The Paratchi people, who speak a language of the same name, can still be found only in some isolated villages of the Patchaghan valley (district of Nijrab, Kapisa province) and in the Shutul valley (Parwan province).

The Hazaras (HAZ)

The Hazaras live mainly in the fairly isolated mountainous regions of central Afghanistan and account for around 10% of the Afghans. They speak a Persian dialect, the Hazaragi. Because of an “Asiatic” facial morphology, it has often been suggested that the Hazaras might be the descendants of soldiers of the army of Genghis Khan. Genghis Khan had conquered the region between 1221 and 1227 CE after having unified various Central Asian tribes, like the Mongols, the Turks and the Tatars. They came as soldiers, without families, and some of them could eventually have settled in the Afghan valleys and married local women. In this way, they might be part of the ethnogenesis of the Hazaras. If such unsupported belief had to be considered as reliable, it should be pointed out that Genghis Khan’s cavalry-men were chosen among allied and conquered tribes, and were likely to be genetically heterogeneous, meaning that the genetic roots of the Hazaras may be multiple (Qamar 2002). More likely, the Hazaras are the descendants of nomad Turkish groups moving from the Altai to the West (Farhadi 2010). They could have been either the Huns Hephtalites (who annihilated the Kouchan empire in the 5th century), or the Tû-Kiue (a group that came to Afghanistan from the North in the 6th century), or even an admixed population having its roots in the various oriental Turkish groups forced to migrate elsewhere because “pushed” by the Mongol conquest.
Modern Hazaras use the districts names corresponding to the area they inhabit (Jaghorri, Behsud, Timuri, Day Tchopan, Day-Kundi, Day-Mirdad, Day-Zangi, Sheikh-Ali, Yakaolang). We still do not know if they used to form well-constituted endogamic groups or if they admixed with groups of ancient Iranian lineage established in the region since more distant times (Poladi 1989; Mousavi 1998; Monsutti 2005).

The most respected among the Hazaras, are the Seyyeds. Men wear a traditional black turban and can be easily identified. The Seyyeds claim to descend from the Prophet Mohammed, and are strictly endogamous. A genetic study would certainly shed some light on their paternal lineage in order to verify any affiliation with those found in the Arabian Peninsula.

**The Moghols (MHJ)**

Considered by cultural anthropologists as true descendants of Mongol groups (Schurmann 1962), the small group called locally “moghols” (Zerjal 2003) speak a form of Mongol, mixed with Persian language (Iwamura 1961). They are established in eight villages of western Afghanistan, located in the southeast of Herat city (villages of Kundur, Karez-e Molla, Burya-baf, Bêdawi and Zimi). They are also to be found in the province of Maymana (Do-rodi and Morcha-gol villages), and near Obeh, in the province of Ghor (Zaman-abad).

**THE “TURKISH PEOPLES”:**

Many different populations speaking languages of the Turco-oriental group live in Afghanistan, mainly in the Northern provinces (Jarring 1939). Until the military conquest in 1882 by Abd-or Rahman, the founder of the modern Afghan state, they lived in small kingdoms, independent from the emir of Bukhara, and from the Pashtun chiefs of southern Afghanistan (Krueger 1963; Toltova 1963; Heywood 1998).
Nowadays, smaller groups are located in well-defined districts, and were semi-nomads until quite recently: Qarakalpak (KAA), Kazakhs (KAZ), Qarluqs, Tatars, Türks near Herat, Mazar-i Sharif and Kondoz cities; also Uygûris (UIG) in Badakhshan province. They still remain attached to their tribal affiliations (Centlivres 1975).

**The Uzbeks (UZS)**

The Uzbeks (~10% of the whole population of Afghanistan) are probably among the descendants of the ancient Qipchak populations established in Central Asia, and came from the North before the 16th century (Burnes 1834; Clavijo 1928). Nomads or sedentary, they occupy the rich loess land of the northern plains, from Maymana to Khanabad. While not all of them have a memory of their ancient tribal affiliations, some still identify themselves as members of tribes: Burka, Durmen, Kipchak, Lakay, Mangit, Ming, Muytan, Qattaghan, Qungrat. At the end of the 19th century, the Uzbek kingdoms (Maymana, Sheberghan, Andkhoy, Balkh, Khulm, Kondoz and Fayzabad) were annexed to the newly constituted Afghan nation and Pashtuns nomads were installed in these former kingdoms. The two populations would not normally intermarry (Irwin 2010; Revazov 1983).

**The Afshars (AZB)**

The Afshars arrived in Afghanistan following the conquest of Nadir Shah in 1737 and, since then, have always lived in the Chandawol neighbourhood of Kabul. Descendants of a Turk speaking group, they belong to Shia Muslim branch (Ligeti 1957). Generally well educated, they soon became civil servants, soldiers or bodyguards at Court and kept their identity. A genetic comparative study with Iranian populations would be possible in order to identify more precisely their roots.
The Turkmens (TUK)

Turkmens (approximately 2% of the whole population) are probably the descendants of the ancient Oghûz tribes of Mongolia. The Oghûz were already in the 7th century organized as a federation of twenty-four clans established around the Aral Sea and the Caspian Sea. From the 10th century onwards, they were converted to Islam (König 1962).

Turkmens speak a language belonging to the south-western group of the Turkish languages. The main stream of Turkmens arrived in Afghanistan between 1928 and 1933. They settled near Herat, and along the left bank of the Amu-Daria River, between the towns of Andkhoy and Kondoz. They still keep a tribal organization: the Tekkes (the dominant group in nearby Turkmenistan) near Herat, the Yomuds (numerous in east Iran), the Karas at Andkhoy, the Qarqins, the Ersaris who are the more numerous in Afghanistan, around Aqcha, and, finally, the Alielis, Egdyrs, Chaudors, Göklens, Salors and Saryks (Irons 1975).

Until 1884 and the takeover of Merv by Russian troops, Turkmen living around the Aral Sea used to raid east Iranian villages, capturing peasants and selling them as slaves on the markets of Khiva and Bukhara. Genetic admixture has thus happened. Further more, being generally wealthy, due to their successful carpet trade, they could sometimes afford having a second wife from neighbouring villages, not necessarily belonging to their own group.

The Kirghiz (KIR)

Mainly living in the sovereign state of Kirghizstan and in the Pamirs Mountains across the Chinese border, there are only a few Kirghiz in Afghanistan. Their arrival from China is rather recent, before 1895 and around 1920 and 1949. Living all the year at an altitude of 4000 m. in valleys accessible only in summer, they breed yaks and sheep (Dor 1975, Dor & Naumann 1978). Such geographic isolation might be of interest for demographers and geneticists to study recent drift and bottleneck phenomena with reference to Kirghiz living in
Kirghizstan (Facchini 1997; Gurgey 2000). In 1979, a majority left to Pakistan as refugees, establishing in the area of Chitral. In 1982, they were given asylum in Kara Gündüz village of eastern Turkey, before commuting to larger cities. Those Kirghiz (estimate 2,000) that remained in Afghanistan can be easily found in the Pamir area.

**THE “INDIANS”:**

**The Brahuis (BRH)**

The Brahuis live in South-west Afghanistan, in Seistan and Registan. They belong to a very ancient population, having pre-Aryan roots and maintaining a strong cultural identity. They speak a Dravidian language of the family of southern India’s ancient languages (region of Dekkan) and of the north of Ceylon. Linguists believe that their arrival chronologically precedes the one of the Baluchs. Traditionally desert nomads, they breed racing dromedaries.

**The Pachays (AEE; GLH; PSI; PSH)**

This appellation groups the populations of various valleys in the provinces of Laghman, Konar and Kapisa in the east of Afghanistan (Wutt 1980). They are the descendants of a Dardic indo-Aryan group and have been isolated until recent times: in each valley a different dialect is spoken. Mutual intelligibility is very low.

Dardic dialects belonging to the same group are spoken in some other villages: Gawarbati (GWT) and Sawi (SDG) in the Konar valley. Tirahi (TRA) is still spoken by a small group of the same name in a few villages of Nangarhar province

**The Nurestanis**

This interesting population inhabits four isolated mountain valleys in the northeast of the country. In each valley, Ashkun (ASK), Bashgal, Prasun (PRN) and Waygal (WBK), a different dialect is spoken. Such dialects, mutually understandable to some extent, are very ancient (Fussmann 1972). Both have Iranian and Indian features and are related to the ancient pre-Vedic language of India before it split into Indian and Iranian branches. Until the end of the 19th century when they became part of the Afghan kingdom, the Nurestanis lived nearly in total isolation, as they were not yet Moslems (Robertson, 1896). To the eyes of anthropologists, they have a special interest as almost perfect cultural and geographical isolates. Nowadays, it is impossible to reach the valleys, as they have become strongholds of insurgents, local and foreigners.

**The Gojurs (GJU)**

They are constituted by small groups of endogamous nomads of Indian origin. Generally sheep farmers, they speak a specific language, Gojur, and live in the highlands of the eastern provinces (Kapisa, Nangarhar) and in the northeast (Baghlan).

**The Jats**

*Jat* means caste in India. This term identifies small and marginalized groups of itinerant artisans of Indian origin. The *Gorbats*, a similar group originally coming from Iran, speak Persian (Rao 1982).

**The Sikhs and the Hindus**

In the 19th century, Sikhs and Afghans fought together for the possession of Peshawar, now capital of the eastern Pashtuns of Pakistan. From 1835 to 1843, the city was administered and
beautified by Paolo Avitabile, the Italian lieutenant in the 1st infantry regiment of Napoleon who became general in the army of Ranjit Singh, the Sikh king of Penjab.

Easily recognisable by their turban, the Sikhs used to live in their majority in the south and east of Afghanistan (Kandahar, Kabul, Jalalabad, Khost), where they were mostly shopkeepers, specialised in Indian tissues and jewels.

In the same way, 20,000 to 30,000 Hindus used to live in Afghanistan, as tissue traders in southern cities. In 1832, Alexander Burnes (1834) pretended that all commerce activities in Central Asia, from Astrakhan and Machhad to Calcutta, were in their hands.

After the Taleban take over of 1994-1996, both the Hindus and the Sikhs had to leave Afghanistan, but some returned after 2001 and rebuilt their religious sites (Dupree 2001).

Both Sikh and Indus speak Hindustani (HIN), Punjabi (PAN), Lahnda (LAH) or Sindhi (SND).

THE SEMITES:

The Arabs

Arabs groups had arrived in Afghanistan at the very beginning of the Arab conquest (7th – 11th centuries) and then, again, in 1366 CE from Syria and Iraq in the army of Tamerlan.

Many had settled in the north of the Amu-Daria river (in Uzbekistan and Tajikistan) (Vinnikov 1940; Doerfer 1969), or in the ancient province of Bactriane (north of Afghanistan) where, nowadays, they can be estimated at a number of more than 100,000 individuals. They are semi-nomads and breed cattle, are Sunnites Moslems and speak Persian (Barfield 1981).

Four isolated villages of Afghan Arabs (ABH) have kept their Arab language, though mixed with Persian terms, and are strictly endogamous: Hassan-abad, north-east of Shiberghan,
Sultan-aregh, near Aqcha, Yakh-dan, and Khoshhal-abad, near Dawlat-abad, Balkh province (Farhadi 1970). The communities living in these villages call themselves Qoresh Arabs, thus claiming a supposed belonging to the Qoresh tribe of the Prophet Mohammed (Sana 1975; Kieffer 1980; Dupaigne 1982). A genetic test may verify any possible relatedness as, according to their tradition, they came from Iraq after the conquest of Baghdad by Tamerlan in 1358 (Debec 1967).

The Jews

According to a very vague verse in the Bible (Zechariah 6:8), they could have lived in Afghanistan in the region of Khorasan (“Land in the North”) (Fischel 1945; Goldstein 2010). Concerning ancient times nothing else is known besides an interesting oral tradition handed down in Kabul by Alexander Burnes (1834), claiming that, after the destruction of Jerusalem, a Jewish tribe was deported by Nabuchodonosor II, king of Babylon, to the province of Ghor in 587 BCE. Such tradition is somewhat corroborated by the finding, in a cemetery near Jam (Ghor province), of graves with Jewish inscriptions, dated from 1012 to 1249 CE, that is at the time of the Ghor empire that collapsed in the 13th century (Bruno 1963; Gnoli 1964). After that, it is likely that they moved elsewhere, but it is impossible to know if they left any descendants. Other Jews arrived in Herat from Mashhad (city east of Iran) in 1839 escaping massacres. After the soviet revolution, starting in 1924, many other Jews came to Afghanistan from different areas of Central Asia, namely the cities of Khiva, Bukhara and Samarkand. At the beginning of the 20th century, Herat still hosted an important Jewish Persian-speaking community, while smaller groups could be found in Bala Morghab, Qala-e No, Maymana, Andkhoy, Kandahar and Kabul.
Starting with 1951, many Jews left Herat to Israel or the United States, and a complete emigration took place in 1978 after the communist takeover. Today, there is only one Jew left in Afghanistan, the guardian of the synagogue in Kabul.

To sample DNA from Jews of Afghan ancestry, a researcher should contact the numerous (~15) Afghan synagogues located in Israel, mainly in Tel-Aviv, or with the Afghan synagogue existing in New York City. Also, a “bukhariot” quarter, created in 1892 and called Rehovoth, exists in Jerusalem and Afghan Jews live close to it (Bar’am-Ben Yossef, 1998).

Bernard Dupagne
National Museum of Natural History, Paris
CNRS UMR 7206
Paris, France
Literature cited


Iwamura, S. 1961. The Zirni manuscript: a Persian-Mongolian glossary and grammar. Kyoto University, Results of the Kyoto University Scientific Expedition to the Karakoram and Hindukush 1955.


Vinnikof. 1940. Les Arabes en URSS. *Sovetskaja Etnografija*. IV.


Figure 1
Map of some ethnical groups of Afghanistan (redrawn from Gille 1984, courtesy of Alain Marigo). A more detailed map can be found in Dupree 1980 at the page 58.