23 November, 2010

High Connectivity among Argali Sheep from Afghanistan and Adjacent Countries: A Noninvasive Assessment Using Neutral and Candidate Gene Microsatellites

G. Luikart[1,2,3], S. Amish[2], J. Winnie[4], A. Beja-Pereira[3], R. Godinho[3], F.W. Allendorf[1], R. Harris[4]

[1]Flathead Lake Biological Station, University of Montana, Polson, MT 59860, USA
[2]Division of Biological Sciences, University of Montana, Polson, MT 59860, USA
[3]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal
[4]Department of Ecosystem and Conservation Science, University of Montana, Missoula, MT 59812, USA

Abstract
We quantified population connectivity and genetic variation in the Marco Polo sub-species of argali mountain sheep (*O. a. polii*) by genotyping neutral and candidate adaptive microsatellite loci in 172 individuals sampled noninvasively from five study areas in Afghanistan, China, and Tajikistan. Heterozygosity and allelic richness were generally high (mean H = 0.67, mean A = 6.1), but were significantly lower in the China study area (H = 0.61, P < 0.001; A = 4.9, P < 0.01). We tested for signatures of natural selection by genotyping microsatellite loci in candidate genes (eight of the 17 loci). One immune system gene (TCRG4) showed an excess of rare alleles compared to neutral expectations. Another immune system gene (GLYCAM-1) showed excessive differentiation (high F_{ST}) between areas. Estimates of genetic differentiation were similar (F_{ST} = 0.035 versus 0.033) with and without the loci deviating from neutrality, suggesting that selection is not a primary driver of overall molecular variation, and that candidate gene loci can be used for connectivity monitoring, as long as selection tests are conducted to avoid biased gene flow estimates. Adequate protection of argali and maintenance of connectivity will require monitoring and international cooperation because argali exhibit high gene flow across international borders.

Keywords: Bottlenecks, fragmentation, *Ovis ammon,* wild sheep, Pamir Mountains, natural selection, adaptation, disease, noninvasive monitoring, mountain ungulate

Introduction
Genetic assessments and monitoring are increasingly crucial for delineating population boundaries and movement corridors, and possibly for understanding adaptation to environmental change in extreme environments (Shackleton 1997; Schwartz et al. 2007). Availability of candidate adaptive gene markers (e.g. Kohn et al. 2006) along with neutral loci could make feasible the assessment of both adaptive and neutral connectivity, i.e. gene flow and adaptation to changing environments (Black et al. 2001).

Argali are an ecologically and economically important species, but are increasingly threatened throughout their range. The Marco Polo subspecies of argali (*Ovis ammon polii*) is among the largest wild sheep and is perhaps the most charismatic wild animal in the Pamir

Mountains of Tajikistan, China, Kyrgyzstan, and Afghanistan (Fedosenko & Blank 2005). The Marco Polo subspecies is categorized as Near Threatened on the IUCN Red List (IUCN 2009). Argali in the Pamir Mountains are important because of their role as a flagship species for the entire ecosystem (Schaller and Kang 2008). However, populations are susceptible to human pressures including poaching, displacement, competition, and disease transmission from livestock, as well as possible habitat fragmentation (Shackleton 1997). Due to their innate habitat preferences, argali are generally distributed patchily, with areas of inappropriate habitat separating populations. Unlike the related mountain sheep (e.g., *O. canadensis*) of North America, argali are generally believed to be willing to traverse long distances, possibly across seemingly inhospitable terrain. Thus the degree to which populations are truly fragmented, either demographically or genetically, is often an open question.

Unfortunately, argali are among the most difficult of wild ungulates to study due to their wary nature, choice of remote and precipitous habitats, and low population density. Argali, unlike North American wild sheep, are cursorial and will move long-distances to escape predators or disturbance. Little is known about argali movements or migrations because they are difficult to capture, tag or collar, and track. In part because they are intolerant of human disturbance, argali typically live only where human access is difficult or infrequent.

Marco Polo argali in the Wakhan Corridor of Afghanistan are found only in a small section of the Big Pamir Mountains, in the eastern Little Pamirs, and in the Wakhjir Valley; thus, possible isolation among populations is a legitimate concern (Figure 1). In contrast, it is likely that populations in Tajikistan are more contiguous in nature (Weinberg et al. 1997). The exact status of Marco Polo argali populations in China is uncertain, although they are known to exist in relatively high numbers in most of the Taxkorgan Nature Reserve in Kashi Prefecture, Xinjiang (Schaller et al. 1987; Gong et al., 2007; Schaller & Kang 2008).

Noninvasive and remote genetic sampling (Taberlet et al. 1999; Beja-Pereira et al. 2009) facilitates research on elusive species such as argali. Molecular genetic markers with metapopulation models allow assessment of migration rates over the recent past (dozens to hundreds of generations, assuming migration-drift equilibrium), and also current migration rates by identifying actual immigrants, e.g. using individual-based assignment tests (Cornuet et al. 1999; Paetkau et al. 2004) or by quantifying immigrant ancestry (Wilson & Rannala 2003). Molecular markers also allow detection of recent population size reductions or bottlenecks (e.g., Cornuet & Luikart 1996).

Detecting loci under selection is important because selection can bias estimates of population genetic parameters, e.g. F_{ST}. Detecting selection signatures also can help infer if a population has experienced a recent selection or stress event such as environmental change or disease die-off, which could help infer the cause of population declines. For example, Simões et al. (2008) studied the genetic response to selection and detected both a reduced effective population size (increased drift at multiple microsatellite loci) and directional selection (F_{ST}-outlier effects at a single microsatellite locus) during the adaptation to a new environment in populations of *Drosophila*. The authors suggest that selection at a single locus was associated with adaptive challenges that increased mortality, contributing to genome-wide drift and reduced effective population size.

Selection can be detected as extremely high (or low) genetic differentiation (F_{ST}) between populations for a single locus compared to neutral loci. Researchers have developed "F_{ST} outlier" tests (Beaumont & Nichols 1996; Antão et al. 2008) and shown that they have reasonable power (Beaumont & Balding 2004) to detect directional selection between

populations. Selection can also be detected using neutrality tests within populations (Watterson 1978). For example, Paterson (1998) detected even allele frequencies at *MHC* in a population of Soay sheep (*Ovis aries*). Microsatellites in genes affected by selection will undergo genetic hitch-hiking and show the selection signature of the gene under selection. This approach has been implemented in large-scale genome scans of thousands of loci to identify genes or genome regions under selection (Payseur et al. 2002; Vasemägi et al. 2005).

We genotyped neutral and candidate adaptive gene microsatellites from fecal DNA sampled from five study areas within three countries with resident Marco Polo argali populations (Afghanistan, China, and Tajikistan) to assess the genetic and demographic status of argali across the region. Our specific objectives were to 1) quantify connectivity of argali among countries and among the three remaining areas with argali in Afghanistan, 2) test for reduced variation and bottleneck signatures within study areas, and 3) test for signatures of selection at immune system genes that might result from adaptive differentiation or stress events such as disease die-offs.

Methods

Study area

Although geographers have not agreed on precise boundaries, the Pamir Mountains are generally viewed as constituting the eastern portion of Tajikistan, the northern half of the Wakhan Corridor in Afghanistan, and the southwest corner of Xinjiang, China. This arid (annual precipitation typically ~ 130mm/yr) and high elevation (highest peaks > 7,000m) range is typified by broad valleys and relatively gentle slopes ("pamir" refers to broad, grassy plateau-like topography).

We sampled from argali in five different locations within the Pamir Mountains; three within Wakhan District, Badakhshan Province, Islamic Republic of Afghanistan, one within Gorno-Bakakhshan, Tajikistan, and one within Taxkorgan County, Kashi Prefecture, Xinjiang Autonomous Region, People's Republic of China. We also collected samples from the Karichinai Valley in Khunjerab National Park, Pakistan, but because none yielded DNA these are not discussed further. We termed our five study areas (Fig. 1) the Big Pamir Mountains, the Little Pamir Mountains, and the Wakhjir Valley (all within Afghanistan), the Murghab study area (in southern Gorno-Badakhshan Province, Tajikistan), and the Taxkorgan study area (in Xinjiang, China). All five study areas were typified by rolling hills and rugged mountains at elevation of 3,900 to 5,300 m, vegetated by arid, steppe vegetation communities of grasses and low-lying shrubs.

The Big Pamir Mountains extend from approximately 37º N to 37º 20' N latitude and 72º 45'E to73º 30' E longitude, and are bordered on the south by the Wakhan River and on the north by the Panj (Amu Darya) River, which also forms the border between Afghanistan and Tajikistan. Within the Big Pamirs, we sampled only from approximately 1000 km^2 centered on 37º N latitude and 73º E longitude. The Little Pamir Mountains, located approximately 150 km east of the Big Pamir Mountains, are centered on 37º 23' N latitude and 74º 20' E longitude. The Wakhjir Valley, located approximately 37º 03' N latitude and 74º 30' E longitude, forms a small spur of Afghanistan that separates Xinjiang, China from the Northern Areas of Pakistan. Our Murghab study area in south-eastern Tajikistan extended along an east-west line of approximately 45 km located about east-west located about 10 to 30 km north of the Afghan border. Our Taxkorgan study area was located within the Taxkorgan Nature Reserve in Taxkorgan County, Kashi Prefecture in Xinjiang, and was centered on approximately 37º 23' N latitude and 75º 20' E longitude.

Geographic distance between study areas ranged from a minimum of ~15 km between Little Pamirs and Murghab to a maximum of 212 km between Taxkorgan and the Big Pamirs (Fig. 1). Distance within Afghanistan ranged from 38 km between the Wakhjir Valley and the Little Pamirs to 164 km between the Big Pamirs and the Little Pamirs.

Sampling

All field work was conducted either on foot, horseback, or "yak-back". Because argali move frequently through difficult terrain and our own movements were circumscribed by the valley systems separated by steep ridges, we made no attempt to impose a standardized geographic sampling regime. Instead, we attempted to survey for argali by walking to high vantage points to search for animals during early morning and late afternoon time periods.

Whenever we encountered fecal pellets we were certain had been freshly deposited by argali, we collected 3 fecal pellets from each pellet group (i.e. pellet pile). We only collected pellets adjacent to each other within the group, reducing to inconsequential the probability of >1 individual argali being represented within each individual sample. We avoided collecting from a pellet group that were scattered over more than approximately a $0.1m^2$ area, or that appeared to have been deposited while the animal was moving. We took GPS locations for each sample (unless samples were within a few paces of an existing GPS fix, in which case we recorded the same location), and noted the date, time, and name of the collector.

Fecal pellets were stored in sterile 30cm centrifuge tubes with securely fitting screw-tops to which internal "sporks" were attached (which allowed individual handling of each sample without risk of contamination; Evergreen Scientific, Los Angeles, CA, USA). We placed 3 fecal pellets into approximately 6 parts of 95% ethyl alcohol (ETOH) for each part fecal material, and stored them at room temperature for 1-4 months before extraction. We collected and extracted DNA from one pellet from each of ~250 pellet groups.

DNA extraction, genotyping, and sexing

Genetic lab work was conducted in two laboratories. Initial work was undertaken at CTM/CIBIO (Centro de Testagem Molecular), Portugal, where fecal samples were extracted and eight STR markers were co-amplified in three multiplex PCR reactions as described in Harris et al. (2010). All samples were initially screened twice to quantify the quality of the nuclear DNA for producing genotype data. Samples that showed reliable amplifications (peak height > 50 units and identical genotypes at 8 loci) in this first step were selected to continue the genotyping process. Selected samples were independently re-genotyped 3 to 4 times for all loci.

The remaining lab work and genetic analysis were conducted at the University of Montana Conservation Genetics Laboratory, Missoula, Montana, USA. Three loci were genotyped in both labs as a data quality check. In total, we genotyped 18 microsatellite loci including 10 putatively neutral loci, and eight loci located in candidate (function) gene introns (except for one candidate locus (ADCYAP-1) located a few hundred base pairs upstream from the candidate gene). The candidate gene loci (ADCYAP-1, KRT2, MHC2 (i.e., OLADRBps), TCRG4, IFNG, MMP9, GLYCAM-1, LIF) are described in Luikart et al. (2008a, b). All loci have some immune system function, except for KRT2 which codes for keratin, a molecule in horn and hair.

Ten of the 18 microsatellite loci were screened only in the Montana laboratory, along with the amelogenin sex-identification locus (Pidancier et al. 2006). A multiplex and a single PCR were optimized and 10ul reactions were performed on MJR PTC200 thermocyclers using

touch-down profiles. Each reaction contained: 2.5ul of template DNA, 4.5ul of QIA multiplex mix (Qiagen), and either 1ul of 10x primer mix, or 1ul of 2pM forward and reverse primers. Two different touch-down profiles with 35 cycles were used, one with an initial annealing temperature of 63°C stepping down to 58°C, and another starting at 58°C and stepping down to 53°C. Fluorescently labeled DNA fragments were visualized on an ABI 3130xl automated capillary sequencer (Applied Biosystems) in the Murdock DNA Sequencing Facility at the University of Montana. Allele sizes were determined using the ABI GS600LIZ ladder (Applied Biosystems). Chromatograms were viewed and analyzed using GeneMapper software v3.7 (Applied Biosystems).

Sex was determined by PCR amplification of the amelogenin gene as in Pidancier et al. (2006). Two PCR products (~315 and ~359bp) were obtained for males but only the longer product for females. Consensus genotypes were based on multiple sample runs and the following rules: for a sample to be heterozygous at a locus both alleles had to be observed twice; for a sample to be homozygous the allele had to be observed three times. Ten percent of samples were randomly chosen, re-extracted, and repeat genotyped to monitor for errors. No genotype differences or errors were detected. Due to the large size of the fragments at the amelogenin locus consensus genotypes were determined as above with the following changes: heterozygotes (males) were provisionally accepted if a male band was only observed once; homozygotes where only the female band was observed less than 3 times (e.g. of 3 independent PCRs) were classified as of unknown gender, and genotypes where only the male band was observed 3 or more times were accepted as males.

Data analysis

Individuals were identified from consensus genotypes from at least four (and up to eight) repeat PCRs for each locus. Consensus genotypes were identified as in previous work (Luikart et al. 2008a; Harris et al. 2010). The probability that two unrelated individuals (or two random siblings) would have identical genotypes (P_{ID}) was computed using Dropout (McKelvey & Schwartz 2005) and Api-Calc (Ayres & Overall 2004). Principal correspondent analysis (PCA) and multilocus genotype matching were conducted in GENALEX (Peakall & Smouse 2006) to identify outliers due to potential genotyping errors or non-argali samples, and to identify identical genotypes. Loci contributing significantly more unique individuals than expected were found with DROPOUT (McKelvey & Schwartz 2005). We estimated expected heterozygosity, tested for gametic (linkage) disequilibrium, and assessed departures from Hardy–Weinberg proportions using exact tests and a Markov chain as implemented in GENEPOP 3.4 (Raymond & Rousset 1995). Allelic richness estimates were corrected for sample sizes using rarefaction (Kalinowski 2005). We quantified genetic differentiation among study areas using exact tests for allele frequency differences and using GENEPOP 3.4.

We tested for reduced allelic richness and reduced heterozygosity (e.g. in study areas with low variation) using Wilcoxon's signed-ranks test. This is a nonparametric test for paired comparisons that is appropriate and powerful when homologous loci are examined in related populations. We tested for genetic signatures of recent population bottlenecks using heterozygosity excess (i.e., deficit of rare alleles) across multiple neutral loci (Cornuet & Luikart 1996, Luikart & Cornuet et al. 1998). We used two mutation models (the stepwise model SMM; and two-phase model TPM with 80% SMM and 20% multi-step mutations with variance of 12) to cover the range of likely mutations models for microsatellite loci (Piry et al. 1999).

Connectivity among populations was assessed using indices of genetic differentiation (F_{ST}) as well as the number of migrants (Nm) estimated from both equilibrium models and assignment test approaches that do not assume equilibrium. Equilibrium migration rate models included the private alleles method (in GENEPOP 3.4), and the F_{ST} method assuming an island model of migration. Two nonequilibrium methods included a Bayesian assignment test approach (BayesAss; Wilson & Rannala 2003) and an individual-based the assignment test of Rannala & Mountain (1997) coded in GeneClass 2.0 (Piry et al. 2004).

We tested for locus-specific signatures of selection in two ways. First we tested for excess evenness of allele frequencies at an individual locus within populations using BOTTLENECK, which gives a probability for each locus being at mutation-drift equilibrium (Cornuet & Luikart 1996). We also plotted the probability values of each locus to help assess genome-wide patterns caused by demographic events (e.g. bottlenecks) that affect all loci similarly. Selection on individual loci can cause an excess of deficit of heterozygosity (i.e. rare alleles) compared to mutation-drift equilibrium expectations.

Second, we tested each locus for an excessively high or low F_{ST} compared to the mean observed F_{ST} by using F_{ST}-outlier tests (Beaumont & Balding 2004, implemented in Antão et al. 2008). An excessively high F_{ST} at a locus compared to thousands of simulated neutral loci indicates possible divergent selection; an excessively low F_{ST} suggests possible balancing selection.

Results

One locus (*MAF*226) out of 18 loci deviated from Hardy-Weinberg proportions ($P < 0.01$) with a deficit of heterozygotes (Table 1). *MAF*226 deviated in two study areas and had a null allele, and therefore was removed from the study. Eleven pairs of loci deviated significantly ($P < 0.01$) from gametic disequilibrium; However among our 1,360 tests, approximately 13.6 deviations were expected by chance alone. No pair of loci deviated strongly from gametic disequilibrium in more than one population. The P_{ID} for identifying sibling pairs was estimated to be less than one in one hundred thousand (10^{-5}) in all study areas for the 17 loci, making power high for resolving between random individuals and sibs (Table 2).

Genetic variation

Heterozygosity ranged from a low of 0.61 in Taxkorgan (China) to a high of 0.69 in Murghab (Table 2). Allelic richness ranged from 4.8 in Taxkorgan to 6.7 in Murghab. Taxkorgan had significantly lower heterozygosity and allelic richness than each other study areas ($P<0.05$; Wilcoxon signed-ranks test). Little Pamir and Wakhjir had the second lowest heterozygosity and allelic richness, respectively.

Genetic structure and connectivity

Mean F_{ST} for the 17 loci was 0.035 among the five study areas. Mean F_{ST} for the nine putatively neutral loci (0.033) was similar to F_{ST} for the candidate loci (0.04), so most results below are reported for all 17 loci, unless stated otherwise. Pairwise mean F_{ST}'s ranged from 0.008 (between Murghab and Big Pamir) up to 0.055 (between Taxkorgan and Wakhjir). Taxkorgan had the highest pairwise F_{ST}'s ranging from a low of 0.033 (with Murghab) to 0.055 (with Wakhjir). Murghab had the lowest F_{ST}'s ranging from only 0.008 with Big Pamir, to the 0.033 with Taxkorgan

We obtained estimates of 5.1 and 6.6 migrants per generation using the private alleles method and F_{ST}-based method (assuming an island model), respectively. The mean frequency of private alleles p(1) was 0.026. The Bayesian approach (BayesAss) for estimating the current number of migrants did not yield informative results because there was not enough information in the data given the relatively low F_{ST} (Faubet et al. 2007), despite our fairly large number of loci with high heterozygosity.

Six highly probable immigrant individuals (P >0.99) were identified in four of the five populations using the individual-based assignment test of Rannala & Mountain (1997). The probable migrants included the following: two in Murghab (from Big Pamir), one in Little Pamir (from Wakhjir), one in Big Pamir (from Murghab), and two in Wakhjir (1 from Big Pamir and 1 from Little Pamir). The estimated probability of the each of the six putative immigrants actually being an immigrant ranged from 99.90% for the immigrant in Little Pamir to 99.95% for the immigrant into Big Pamir. When we lowered the stringency criterion for identification of a migrant (from P >0.99 to P >0.90), five additional migrants were identified, including two into Murghab, two in Wakhjir, and one in Big Pamir.

No immigrants were identified in the China study area of Taxkorgan. In fact, only one individual of 38 from Taxkorgan could potentially be an immigrant, but the probability of that individual being a resident from Taxkorgan was still 11% (Fig. 2). If we lower the criterion of certainty for the identification of an immigrant (from P >0.99 to P >0.90) Taxkorgan still showed no evidence for immigrants unlike all other study areas (e.g., Fig. 2).

Selection tests and F_{ST}

The *TCRG4* gene microsatellite had a significant excess ($P < 0.01$) of rare alleles (i.e., uneven allele frequency distribution compared to neutral expectations) in both the Murghab and the Little Pamir study areas. None of the nine neutral loci or the other seven candidate gene loci deviated from expected allele frequencies under mutation-drift equilibrium (Fig. 3).

Two candidate gene microsatellite loci had an F_{ST} value significantly different from neutral expectations. *GLYCAM*-1 had a significantly higher F_{ST} ($F_{ST} = 0.068$; $P = 0.02$) and *ADCYAP-1* had a significantly lower F_{ST} ($F_{ST} = 0.002$, $P = 0.03$) than expected under neutrality. Neither F_{ST} deviation was significant at the 0.01 level. No neutral loci gave evidence of selection or deviated from mutation-drift equilibrium.

Estimates of genetic differentiation were similar with and without the three outlier loci (*GLYCAM-1*, *ADCYAP-1*, and *TCRG4*): the mean F_{ST} decreased slightly from $F_{ST} = 0.035$ for all 17 loci to $F_{ST} = 0.033$ for the 14 loci with no selection signature. Pairwise F_{ST}, computed after removing outliers, changed most for the Taxkorgan area in China. For example, F_{ST} declined from 0.051 to 0.033 when removing the three outlier microsatellites in candidate genes.

Discussion

Our study of neutral and candidate adaptive genes in argali populations suggests relatively high variation within, and low differentiation among populations compared to other mountain sheep (e.g., Gutierrez-Espeleta 2000; Worley et al. 2004; Epps et al. 2005; Hogg et al. 2006; Luikart et al. 2008a). This is similar to results obtained in Mongolia using mtDNA of argali populations which had been assumed to represent different subspecies (Tserenbataa et al. 2004).

Genetic variation and bottlenecks
Heterozygosity and allelic richness were high compared to many of the same loci genotyped in other wild sheep, in which mean heterozygosity is approximately $H = 0.60$ or lower (Ozut 2001; Epps et al. 2005; Hogg et al. 2006; Luikart et al. 2008a, 2008b), and even lower in other wild ungulate species (e.g., Gebremedhin et al. 2009 and papers cited therein). This suggests that Marco Polo argali populations have relatively high effective population sizes and that our study areas are not yet isolated or inbred, as has been feared (Shackleton 1997; Harris et al. 2010).

Nonetheless, the significantly lower heterozygosity and allelic richness in Taxkorgan (China) compared to our other study areas suggests this population is smaller (Schaller et al. 1987, 2008), and perhaps relatively more isolated than the other populations. The Taxkorgan population is near the southeastern edge of the range of argali and there is a long fence (350 km) near the Tajik-Chinese border (Schaller et al. 2008) that could potentially reduce connectivity of the Chinese argali with other populations in Murghab and the Afghan Pamirs; however it is uncertain if the fence is a barrier because, for example, it is not continuous (e.g. open on some high slopes) and argali might jump over it in some locations.

The absence of strong bottleneck signatures, even in Taxkorgan, and the reasonably high allelic richness suggest no evidence of recent or severe reductions in population size. Power for detecting severe reductions is reasonably high when using seventeen highly variable microsatellite loci and 38 individuals (Cornuet & Luikart 1996; Luikart & Cornuet 1998), as we have from Taxkorgan. Thus if the Taxkorgan population has become genetically bottlenecked or increasingly isolated, which seems likely, the signal might not detectable if the isolation was recent (e.g. <2-4 argali generations ago); Bottleneck signals can take several generations to become detectable if the bottleneck effective size (N_e) remains fairly large (e.g. >50; Figure 3 in Cornuet & Luikart 1996). Genetic bottleneck signatures also might be obscured by immigration if recent immigration has occurred.

Differentiation and connectivity
The genetic differentiation (F_{ST}) in argali is similar to or lower than other mountain sheep sampled at similar spatial scales in North America. For example, in desert bighorn sheep (*O. canadensis nelsoni*) from Arizona, F_{ST}'s ranged from 0.04 to 0.20 (Gutierrez-Espeleta 2000). Furthermore, over a geographic distance of only 5 km, F_{ST} ranged from 0.046 to 0.113 in desert bighorn sheep populations without and with a barrier (e.g. road), respectively (Epps et al. 2005). Worly et al. (2004) found that genetic differentiation in thinhorn sheep populations (*Ovis dalli*) from western Canada was similar to that reported in desert bighorn sheep (Gutierrez-Espeleta et al. 2000). Rocky Mountain bighorn sheep (*O. canadensis canadensis*) also show F_{ST} of ~0.11 at this spatial scale (e.g. 40 km across Glacier National Park, Luikart et al. 2008a).

The lower differentiation among argali is consistent with their more cursorial nature compared to North American sheep. Argali tend to migrate and run for long distances following a threat rather than running up into steep escape terrain as do bighorn sheep. Argali are built more for running, and having longer legs, than North American sheep. They also will move across large valleys which is less common in North American sheep.

The highest differentiation (F_{ST}) for Taxkorgan among our study areas is consistent with the significantly reduced genetic variation there (compared with other study areas), and suggests increased isolation or a smaller population size in Taxkorgan. We recommend additional studies of Taxkorgan and other argali on the Chinese side of international borders, along with monitoring of genetic variation to ensure early detection of population declines, isolation, or

recruitment problems, which could potentially be developing. The lowest genetic differentiation in Murghab (F_{ST}'s 0.008 to 0.033) is consistent with Murghab having the highest genetic variation and being centrally located in the heart of the distribution range of Marco Polo argali.

To estimate gene flow, mean F_{ST} values can give only very rough estimates of average number of migrants per generation, and only if populations are near mutation-drift equilibrium (Whitlock & McCauley 1999). Because most natural populations are seldom near equilibrium, and violate other assumptions, our estimates of ~5 or 6 migrants per generation must be interpreted with great caution; the actual number of migrants could be far higher, for example.

Current (contemporary) gene flow can be detected from the identification of actual migrants by using individual-based assignment tests (Paetkau et al. 2004). For example, Taxkorgan had no detectable immigrants (out of 38 individuals sampled) suggesting relatively low connectivity. The identification of putative immigrants in all other populations suggests they currently are not isolated. The threshold of 99% certainty for identification of a migrant could be viewed as overly stringent. With our sample size of 172 individuals (and $\alpha < 0.01$), we expect 1.7 migrants to be identified by chance alone (as false positives); whereas we identified 6 migrants ($P > 0.99$). Because we detected 6 probable migrants, it seems likely that several true migrants exist and that most populations, except perhaps Taxkorgan in China, have current migration rates greater than zero. The use of 95% certainty ($\alpha < 0.05$) for each individual assignment resulted in identification of 11 probable migrants when ~8 were expected by chance alone.

Selection and adaptation

We detected evidence for selection only at candidate gene loci, not at neutral loci, suggesting candidate gene approaches can potentially identify loci under selection when using noninvasive sampling in wild sheep. The *GLYCAM-1* microsatellite showed higher F_{ST} than neutral expectations (Fig. 2). This could potentially result from selection at this locus or at other genes nearby such as IFNg (<20 centimorgans away from *GLYCAM-1* in domestic sheep), which has been associated with parasite load in sheep (Coltman et al. 2001) and other ungulates (Ezenwa et al. 2010). *GLYCAM-1* function involves mediating the trafficking of blood-born lymphocytes into secondary lymph nodes, and also is expressed in the mammary gland of lactating mammals (Hou et al. 2000; Rasmussen et al. 2002). Further studies are needed to assess if *GLYCAM-1* genotypes are associated with disease resistance in sheep.

ADCYAP-1 showed lower F_{ST} than neutral expectations, which could potentially reveal balancing selection for certain alleles across populations. The *ADCYAP-1* gene (adenylate cyclase-activating polypeptide) is involved in regulation of production of interleukin 6 that activates the production of T-helper cell 2 (Th2) cytokines involved in defense against helminths and other extracellular parasites (Mosmann & Sad 1996). *ADCYAP-1* was recently found to be associated with nematode parasite infection in domestic sheep (Crawford et al. 2006), and heterozygotes had lower parasite loads in wild bighorn sheep (Luikart et al. 2008b). If argali suffer from disease (e.g. transmitted from livestock) it is plausible, though speculative, that parasites and disease in the Pamirs have lead to selection at *ADCYAP-1*. Future research is needed to assess potential effects of disease in argali and on *ADCYAP-1*.

Removal of the two F_{ST} outlier loci caused little change in mean F_{ST} among the five study areas (from 0.035 to 0.30, without *ADCYAP-1* and *GLYCAM-1*). Similarly, removal of the locus (TCRG4) with a heterozygosity-excess had little effect on our multilocus F_{ST} estimates among study areas.

Interestingly, removal of *GLYCAM-1* decreased F_{ST} between Taxkorgan (China) and other study areas. For example, F_{ST} from ~0.05 to ~0.04 when we removed *GLYCAM-1* when comparing Taxkorgan with Big Pamir or Taxkorgan to Wakhjir (Figure 2B, 2C). Removal of *GLYCAM-1* did not substantially reduce F_{ST} between other study areas, suggesting this *GLYCAM-1* gene contributes substantially to the relatively high multi-locus F_{ST} observed between Taxkorgan and other study areas. The *MHC* locus also had relatively high F_{ST} between Taxkorgan and other study areas (e.g. Figure 2C). These observations raise the speculative hypothesis that some disease-related selection differential exists between Taxkorgan and other study areas.

How could selection tests and genotyping of both neutral and candidate adaptive loci help advance conservation genetics studies? Many candidate adaptive loci will behave as neutral loci, and thus can be used to assess genetic variation, differentiation (F_{ST}) and demography (Nm and change in N_e). For example, most allozyme loci and candidate gene loci show little evidence of selection. However, if a locus reliably shows a response to selection, it could be used to monitor to detect adaptive challenges (e.g. disease die-offs or environmental change) or to identify adaptively-differentiated populations that have exceptionally high F_{ST} only at candidate genes associated with selection gradients (e.g. disease or temperature). Future developments in genomics will allow noninvasive analyses of hundreds of neutral and candidate adaptive genes, which will help detect population declines and perhaps infer their cause (Simões et al. 2008; Allendorf et al. in review).

Conclusions

Our study illustrates the potential usefulness of genotyping both neutral and candidate adaptive loci, which can facilitate inferences about both demographic status (migration and bottlenecks) and selection events such a disease epizootics and environmental change. Our study suggests that candidate gene loci can be used for connectivity monitoring as long as "outlier tests" are conducted to avoid using non-neutral loci when estimating parameters (e.g. F_{ST}) that can be biased by natural selection. Future noninvasive studies will include 100s of loci (e.g., SNPs) in candidate genes thanks to advances in genotyping technologies that work well with partially-degraded DNA, such as qPCR assays , which we are developing for use in Fluidigm SNP chip arrays (Allendorf et al. 2010).

Argali populations appear to have high genetic variation and connectivity in the Pamirs within Wakhan District of Afghanistan, and Murghab (Tajikistan), but potentially are becoming isolated in Taxkorgan, China. We recommend additional studies, including genetic and demographic monitoring of connectivity, along with disease status, to help maintain connectivity and ensure persistence of argali populations. The establishment of an international nature reserve and coordination program involving Afghanistan, China, and Tajikistan (e.g. Schaller 2007) would help maintain connectivity and facilitate conservation of argali, their habitat, and other species in the region.

Acknowledgements

Our study was part of the Afghanistan Biodiversity Conservation Program of the Wildlife Conservation Society (WCS), supported by the United States Agency for International Development (USAID). For field assistance we thank B. Habib, Z. Moheb, Sabir, and A. Khairzad. We also thank S. Kondratov, K.J. Zhang, S. Ostrowski, D. Bedunah, S. Nikzad, Z. Ejlasi, I. Farahmand, Q. Sahar, K. Sediqi, K. Sidiqi, G. Sediq, A. Ahamad, R. King, and L.

Yook. We thank A. Dehgan, P. Zahler, P. Smallwood, P. Bowles, and A. Simms for advice and administrative assistance. AB-P was supported by SFRH/BPD/17822/2004 RG by SFRH/BPD/36021/2007, and this work was supported by POCI/CVT/567558/2004 all from Fundacao para a Ciencia e Tecnologia (FCT), Portugal. G.L. and F.W.A were supported in part by a grant from U.S. National Science Foundation (Grant DEB 074218). GL was partially supported by grants from the Walton Family foundation, CIBIO, the University of Porto, Portugal, and a research grant PTDC/BIA-BDE/65625/2006 from the Portuguese Science Foundation.

Literature Cited

Allendorf FW, Hohenlohe P, Luikart G (2010) Genomics and the future of conservation. Nature Reviews Genetics. 11:697-709.

Antão TA, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular adaptation based on an Fst-outlier method. *BMC Bioinformatics*, **9**, 323.

Ayres KL, Overall ADJ (2004) API-CALC 1.0: a computer program for calculating the average probability of identity allowing for substructure, inbreeding and the presence of close relatives. *Molecular Ecology Notes* **4**, 315–318.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. Lond. B Biol. Sci*. **263**, 1619–1626.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology,* **13**, 969–980

Beja-Pereira A, Oliviera R, Alves PC, Schwartz M, Luikart G (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular Ecology Resources,* **9**, 1279-1301.

Black WC, Baer CF, Antolin MF, DuTeau NM (2001) Population genomics: genome-wide sampling of insect populations. *Annu. Rev. Entomol*. **46**, 441–469.

Crawford AM et al. (2006) Discovery of quantitative trait loci for resistance to parasitic nematode infection in sheep: I. Analysis of outcross pedigrees. *BMC Genom*, **7**, 178.

Kalinowski S (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes,* **5**, 187–189

Cornuet J-M, Luikart G (1996) Description and evaluation of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144,** 2001-2014.

Cornuet J-M, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes for selecting or excluding populations as origins of individuals. *Genetics*, **153**, 1989-2000.

Ezenwa et al. (2010) Hidden consequences of living in a wormy world: nematode-induced immune-suppression facilitates tuberculosis invasion in African buffalo. *American Naturalist,* In press.

Epps et al. (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters*, **8,** 1029–1038.

Faubet P, Waples RW, Gaggiotti OE (2007) Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Molecular Ecology*, **16**, 1149–1166.

Fedosenko AK, Blank DA (2005) *Ovis ammon*. *Mammalian Species,* **773**, 1-15.

Gebremedhin B, et al. (2009) Conservation status of the critically endangered Walia ibex (*Capra walie*): evidence from genetic data and environmental parameters. *Animal Conservation*, **12**, 89-100.

Gong MH, ZG Dai, ZH Zeng, Q Zhang, YL Song (2007) A preliminary survey of population size and habitats of Marco Polo sheep (*Ovis ammon polii*)in Taxkorgan Nature Reserve, Xinjiang, China. *Acta Theriologica Sinica* **27**, 317 -324. (In Chinese with English abstract).

Greenwood, PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, **28**, 1140–1162.

Gutierrez-Espeleta GA, Kalinowski ST, Boyce WM, and Hedrick PW (2000) Genetic variation and population structure in desert bighorn sheep: implications for conservation. *Conservation Genetics*, **1**, 3–15.

Harris RB, Winnie J Jr, Amish SJ, Beja-Pereira A, Godinho R, Costa V, Luikart G (2010) Population estimation of argali (*Ovis ammon*) in the Afghan Pamir using capture-recapture modeling from fecal DNA. *Journal of Wildlife Management*.74: 668-677

Hogg JT, SH Forbes, BM Steele, G Luikart (2006) Genetic rescue of an insular population of large mammals. *Proceedings of the Royal Society,* **273,** 1491-1499.

Hou Z, Bailey JP, Vomachka AJ, Matsuda M, Lockefeer JA, Horseman ND: (2000) Glycosylation-dependent cell adhesion molecule 1 (*GLYCAM* 1) is induced by prolactin and suppressed by progesterone in mammary epithelium. **141**, 4278-4283.

IUCN (2009) IUCN Red List of Threatened Species; section by Harris RB & Reading R. IUCN Species Survival Commission, Gland Switzerland. Version 2009. <www.iucnredlist.org>.

Kalinowski S (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, **5**, 187–189.

Kohn MH, WJ Murphy, EA Ostrander, RK Wayne (2006). Genomics and Conservation Genetics. *Trends Ecology and Evolution,* **21**, 629-637.

Luikart G, J-M Cornuet (1998) Empirical evaluation of a test for detecting recent historical population bottlenecks. *Conservation Biology*, **12**, 228-237.

Luikart G, PR England, D Tallmon, S Jordan, P Taberle (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981-994.

Luikart G, Zundel S, Rioux D, Miquel C, Keating KA, Hogg JT, Steele B, Foresman K, Taberlet P (2008a) Low genotyping error rates for microsatellite multiplexes and noninvasive fecal DNA samples from bighorn sheep. *Journal of Wildlife Management*, **72**, 299-304.

Luikart G, Pilgrim K, Visty J, Ezenwa VO, Schwartz MK (2008b) Candidate gene microsatellite variation is associated with parasitism in wild bighorn sheep. *Biology Letters*, **4**, 228-231.

McKelvey KS, Schwartz MK (2005) DROPOUT: a program to identify problem loci and samples for noninvasive genetic sampling in a capture-mark-recapture framework. *Molecular Ecology Notes,* **5**, 716-718.

Magomedo M-R, Akhmedov, EG, Wall WA, Subbotin, AE (2003) Current status and population structure of argalis (*Ovis ammon* L., 1758) in Central Asia. *Beitrage zur Jagd und Wildforschung* **28,** 151-163.

Mosmann RR, Sad S (1996) The expanding universe of T-cell subsets: Th1, Th2 and more. *Immunol. Today* **17**, 138–146.

Ozut D (2001) Conservation genetics of Anatolian mouflon (*Ovis gmelinii anatolica*, Valenciennes, 1858). M.Sc. Middle East Technical University, Ankara, Turkey.
Peakall R., Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288–295.
Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, **13**, 55–65.
Paterson S (1998) Evidence for balancing selection at the major histocopatibility complex in a free-living ungulate. *Journal of Heredity*, **89**, 289- 294.
Payseur BA, Cutter AD, Nachman M W (2002) Searching for evidence of positive selection in the human genome using patterns of microsatellite variability. *Mol. Biol. Evol.* **7**, 1143–1153.
Pidancier N, Jordan S, Luikart G, Taberlet P (2006) Evolutionary history of the genus Capra (Mammalia, Artiodactyla): Discordance between mitochondrial DNA and Y-chromosome phylogenies. *Molecular Phylogenetics and Evolution,* **40**, 739-349.
Piry S, Luikart G, Cornuet J-M (1999) Bottleneck: A computer program for detecting recent reductions in effective population size from allele frequency data. *J. Heredity*, **90**, 502-503.
Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. *J Hered*. **95**, 536-9.
Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the USA*, **94**, 9197–9201.
Rasmussen LK, Johnsen LB, Petersen TE, Sorensen ES (2002) *GLYCAM*-1 mRNA is expressed in the mammary gland as splicing variants and encodes various aberrant truncated proteins. *Immunol Lett,* **83**, 73-75.
Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution,* **49**, 1280-83.
Taberlet P, LP Waits, G Luikart (1999) Non-invasive genetic sampling: look before you leap. *Trends in Ecology and Evolution*, **14**, 323-327.
Schaller B (2007) A proposal for a Pamir International Peace Park. USDA Forest Service *Proceedings RMRS*, **49**, 227-231.
Schaller GB, Li H, Lu H, Ren JR, Qiu MJ, Wang HB (1987) Status of large mammals in the Taxkorgan Reserve, Xinjiang, China. *Biological Conservation*, **42**, 53-71.
Schaller GB, AL Kang (2008) Status of Marco Polo sheep *Ovis ammon polii* in China and adjacent countries: conservation of a vulnerable subspecies. *Oryx*, **42**, 100–106.
Schwartz MK, G Luikart, RS Waples (2007) Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution*, **22,** 25-33.
Shackleton DM (editor & compiler) (1997) Wild Sheep and Goats and their relatives: status survey and conservation action plan for Caprinae. IUCN, Gland, Switzerland, and Cambridge UK. 390 pp.
Simões P, Pascual M, Santos J, Rose MR, Matos M (2008) Evolutionary dynamics of molecular markers during local adaptation: a case study in *Drosophila subobscura*. *BMC Evol Biol*. **8**, 66-77.

Tserenbataa T, Ramey II RR, Ryder OA, Quinn TW, Reading RP (2004) A population genetic comparison of argali sheep (*Ovis ammon*) in Mongolia using the ND5 gene of mitochondrial DNA; implications for conservation. *Molecular Ecology* **13,** 1333-1339.

Vasemägi A, Nilsson J, Primmer CR (2005) Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar L.*). *Molecular Biology and Evolution*, **22**, 1067–1076.

Weinberg PI, AK Fedosenko, AB Arabuli, A Myslenkov, AV Romashin, I Voloshina, N Zheleznov 91997) The Commonwealth of Independent States (former USSR). Pages 172-193 in: D. S. Shackleton (ed.) Wild sheep and goats and their relatives: Status survey and conservation action plan for Caprinae. IUCN/SSC Caprinae Specialist Group, IUCN, Gland, Switzerland and Cambridge, UK. 390 pp.

Whitlock MC, and DE McCauley 1999. Indirect measures of gene flow and migration: FST doesn't equal 1/(4Nm+1). Heredity **82**, 117-125.

Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177-1191.

Worley K, J Carey, A Veitch, DW Coltman (2006) Detecting the signature of selection on immune genes in highly structured populations of wild sheep (*Ovis dalli*). *Molecular Ecology,* **15,** 623-637.

Table 1. Characteristics of the 18 loci tested. The top locus (MAF226) was excluded due to a null allele and high mean F_{IS}. The next nine loci are presumed to be selectively neutral and are not near coding genes. N is the mean number of individuals genotyped at the locus among the five study areas. A is the mean allelic richness among study areas.

Locus Name	N	A	SE	H_e	SE	F_{IS}
MAF226	33.0	5.40	0.75	0.72	0.01	0.43
MAF36	34.2	6.00	0.71	0.74	0.01	-0.03
MAF48	33.2	7.60	0.24	0.81	0.01	0.02
MAF209	33.6	4.60	0.51	0.67	0.02	0.09
FCB304	33.8	6.00	0.84	0.69	0.05	-0.01
FCB266	34.4	6.20	0.86	0.76	0.03	0.05
HH62	33.8	7.80	0.73	0.72	0.02	0.10
MAF33	34.4	5.60	0.51	0.53	0.04	0.16
MAF65	33.6	8.60	0.81	0.81	0.02	0.07
ILST30	33.4	2.80	0.37	0.48	0.02	0.15
ADCYAP-1	34.4	6.80	0.58	0.79	0.01	-0.03
KRT2	34.2	9.80	0.20	0.84	0.02	-0.01
MHC2	29.2	6.60	0.68	0.76	0.04	0.06
TCRG4	32.2	5.40	0.93	0.53	0.09	0.07
IFNG	34.2	1.20	0.20	0.01	0.01	-0.02
MMP9	32.8	6.40	0.51	0.74	0.02	0.07
GLYCAM-1	30.0	9.20	1.16	0.79	0.04	-0.04
LIF	29.8	3.80	0.20	0.65	0.01	0.00

Table 2. Characteristics of genetic variation in each of the five study areas.

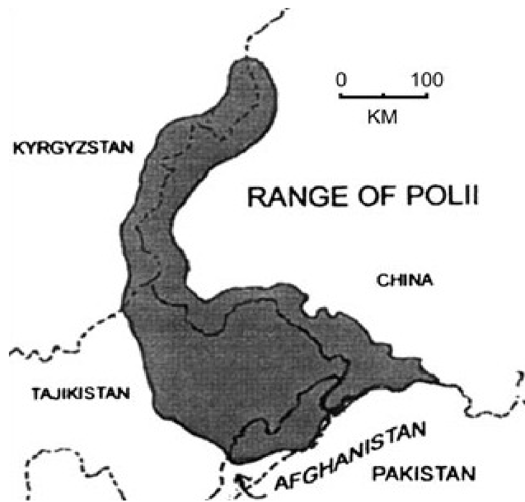
Study Area	***N***	***H*$_e$ (SE)**	***A***	***F*$_{IS}$**	**P$_{IDsibs}$**
Big Pamir	63	0.68 (0.05)	6.40	0.010	8.6E-07
Taxkorgan (China)	38	0.61 (0.05)	4.88	-0.027	6.1E-06
Little Pamir	29	0.65 (0.06)	6.65	0.050*	2.1E-06
Murghab (Tadjik.)	24	0.69 (0.05)	7.31	0.028	8.2E-07
Wakhjir	18	0.69 (0.05)	5.94	0.017	1.0E-06
Mean	**34.4**	**0.67**	**6.12**	**0.040**	**2.2E-6**

*p<0.05; *N* is the number of individuals sampled. H_e is mean expected heterozygosity. *A* is allelic richness (mean number of alleles per locus corrected for sample size *N*; Kalinowski 2005). P_{IDsibs} is the probability of identity among sibs (Waits et al. 2001).

Table 3. F_{ST} between all pairs of sampling areas. All 17 loci included below the diagonal. Only the eight candidate gene loci above. Largest values are in bold.

	Big Pamir	**Taxkorgan**	**Little Pamir**	**Murghab**	**Wakhjir**
Big Pamir	-	**0.075**	0.038	0.007	0.020
Taxkorgan	0.051	-	**0.060**	**0.058**	**0.055**
Little Pamir	0.039	**0.053**	-	0.029	0.016
Murghab	0.008	0.040	0.034	-	0.009
Wakhjir	0.018	**0.053**	0.032	0.010	-

A)



B)

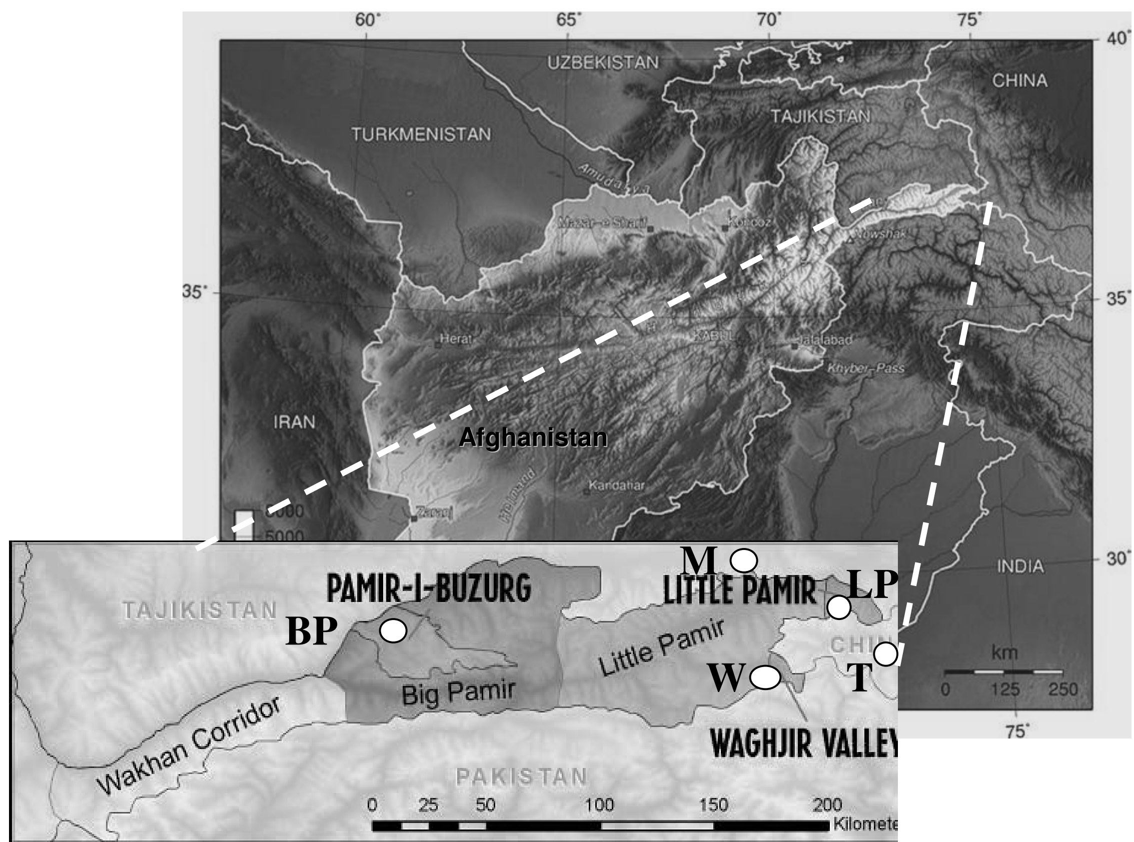


Figure 1. Map of (A) distribution of the subspecies (*O. a. polli*) (from Schaller and Kang 2008), and (B) the study areas in the Pamir Mountains of Afghanistan, Tajikistan, and China, showing national boundaries and major rivers. White dots show approximate locations of our Murghab (M) study area in southern Tajikistan, the Taxkorgan (T) study area in Xinjiang, China, three study areas in Afghanistan: the Big Pamir (BP), Little Pamir (LP), Wakhjir (Waghjir) Valley (W).

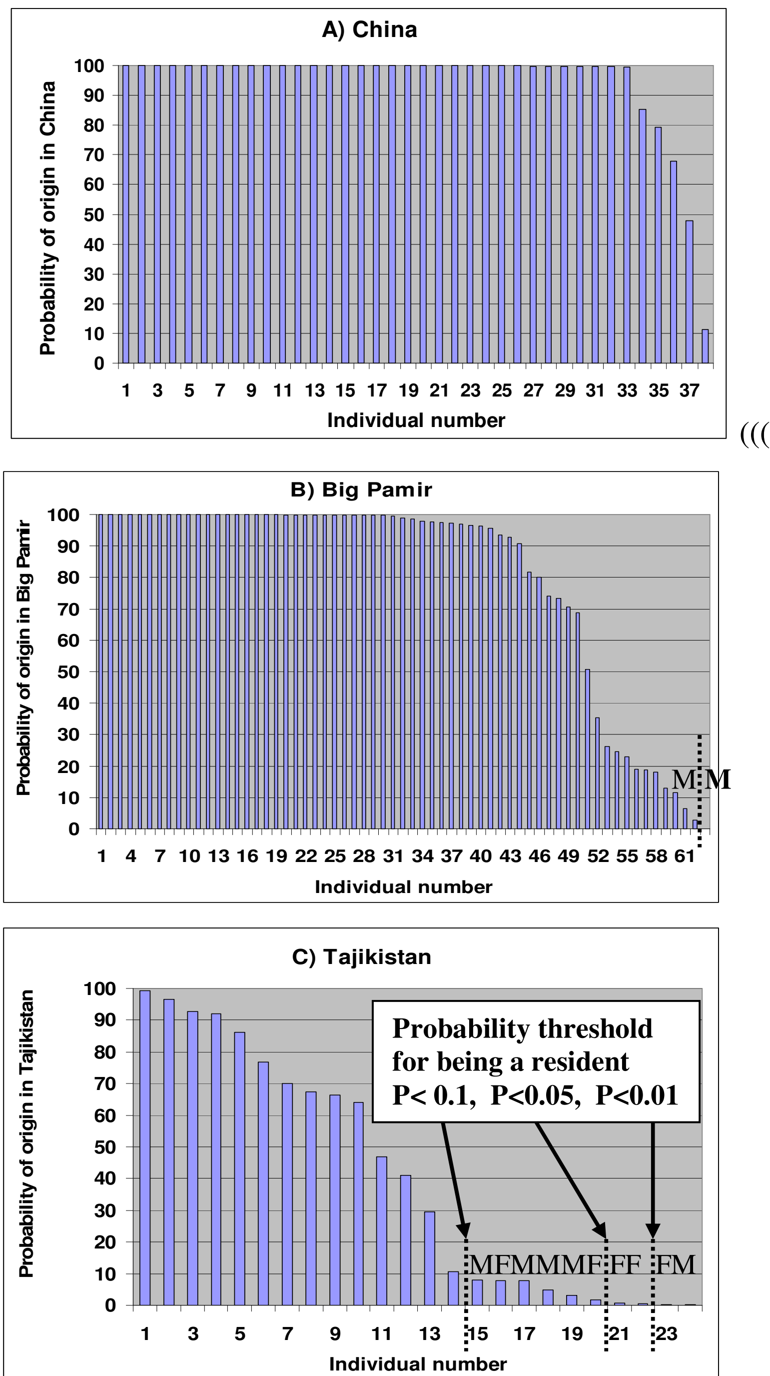
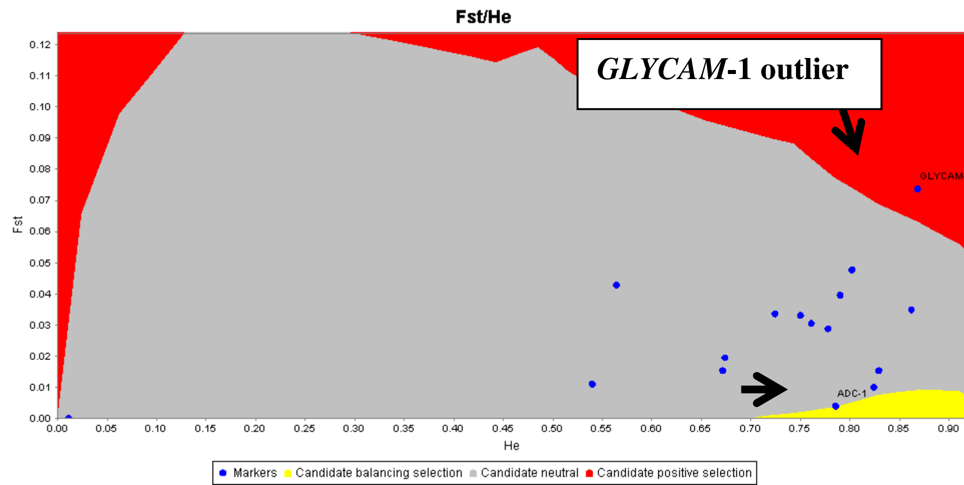


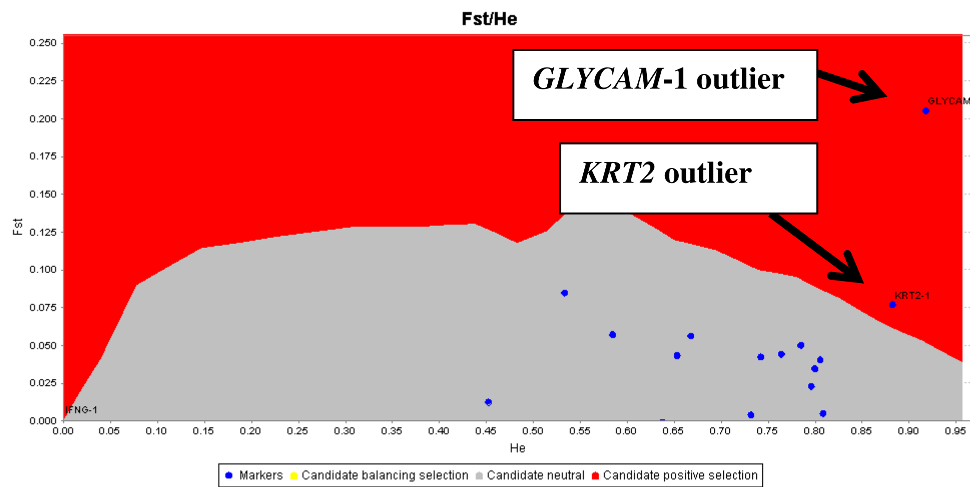
Figure 2. Assignment test estimates of the probability of local origin (Y-axis) of each individual in the study area from which it was sampled. Each vertical bar represents one individual (X-axis). A) In China (Taxkorgan), the individual genotype least likely to originate from China (#38) still had an estimated 11% probability of originating there. In the Big Pamir, B), one individual (#63) had a probability of only 0.04% of originating locally, and was therefore considered an immigrant. In C) the Tajikistan (Murghab) study area, two individuals (#23 and

#24) had a very low probability of local origins (< 0.3%). The letters M and F designate male and female individuals with a reasonably low probability of being local residents. It is surprising that 3 of 4 of the high probability ($P<0.05$) immigrants in the Murghab area are females, because males are thought to disperse more in wild sheep. In Murghab, 10 of the 24 individuals were males. In China, 24 of the 38 individuals were males; in Big Pamir, 28 of 63 were males.

A. All five study areas



B. Taxkorgan versus Big Pamir



C. Taxkorgan versus Wakhjir

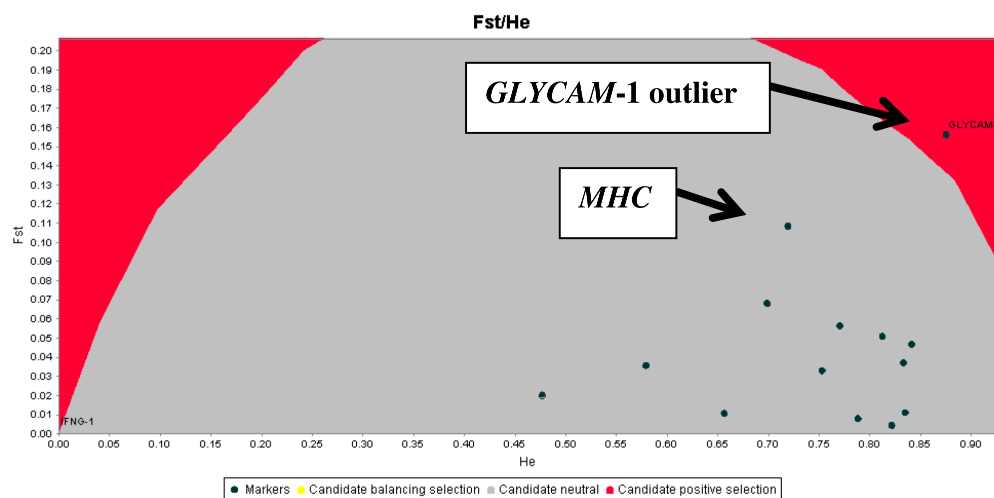
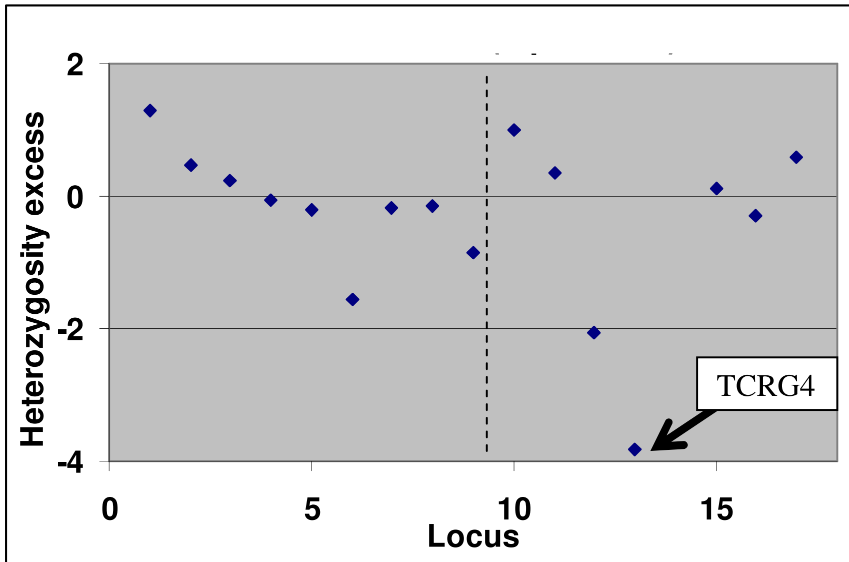


Figure 3. F_{ST}-outlier test results showing: A) the *GLYCAM*-1 locus with excessively high F_{ST} among A) all five study areas, B) between Taxkorgan versus Big Pamir, and C) Taxkorgan versus Wakhjir. Dots represent loci. Light grey area (with most dots) represents the expected area for neutral loci (99%). *MHC* has the second highest F_{ST} in B) and C) (see arrow).

A.



B.

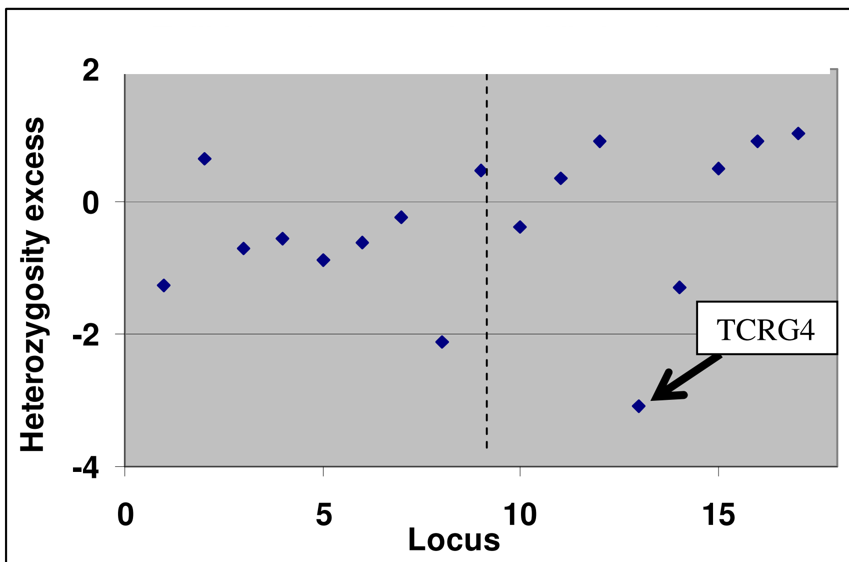


Figure 4. One locus (TCRG4, see arrow) had a significant deviation from mutation-drift equilibrium, i.e., a deficit of heterozygosity (also called an excess of rare alleles), in two populations: *A*. Little Pamir, and *B*. Murghab (Tajikistan). Loci (dots) at mutation-drift equilibrium will have zero heterozygosity excess (y-axis). Loci are in the same order as listed in Table 1. Vertical dashed line separates the 9 neutral (1-9) and 8 candidate adaptive gene loci (10-17).