Estimating the Population Mutation Rate from a de novo Assembled Bactrian Camel Genome and Cross-Species Comparison with Dromedary ESTs

PAMELA A. BURGER AND NICOLA PALMIERI

From the Institut für Populationsgenetik, Vetmeduni Vienna, Veterinärplatz 1, 1210 Wien, Austria/Europe

Address correspondence to P. A. Burger at the address above, or e-mail: Pamela.Burger@vetmeduni.ac.at.

Abstract

The Bactrian camel (Camelus bactrianus) and the dromedary (Camelus dromedarius) are among the last species that have been domesticated around 3000–6000 years ago. During domestication, strong artificial (anthropogenic) selection has shaped the livestock, creating a huge amount of phenotypes and breeds. Hence, domestic animals represent a unique resource to understand the genetic basis of phenotypic variation and adaptation. Similar to its late domestication history, the Bactrian camel is also among the last livestock animals to have its genome sequenced and deciphered. As no genomic data have been available until recently, we generated a de novo assembly by shotgun sequencing of a single male Bactrian camel. We obtained 1.6 Gb genomic sequences, which correspond to more than half of the Bactrian camel’s genome. The aim of this study was to identify heterozygous single-nucleotide polymorphisms (SNPs) and to estimate population parameters and nucleotide diversity based on an individual camel. With an average 6.6-fold coverage, we detected over 116,000 heterozygous SNPs and recorded a genome-wide nucleotide diversity similar to that of other domesticated ungulates. More than 20,000 (85%) dromedary expressed sequence tags successfully aligned to our genomic draft. Our results provide a template for future association studies targeting economically relevant traits and to identify changes underlying the process of camel domestication and environmental adaptation.

Key words: single-nucleotide polymorphism, whole-genome sequencing, camelidae, nucleotide diversity, comparative genomics

During the process of domestication and breed formation, farm animals have encountered both natural and artificial selections. This created a huge diversity of phenotypes among domesticated species. Moreover, this has triggered marked genetic adaptation to different environmental and anthropogenic, very specific, conditions. Livestock and other domestic animals therefore represent a unique resource to understand the genetic bases of phenotypic variation and genetic adaptation (Andersson and Georges 2004). The development of “next-generation” sequencing technologies and high-throughput genotyping platforms tremendously advanced whole-genome analysis in domestic animals and nonmodel species (Nawy 2012). Just as new advances in sequencing methodologies have made it possible to identify the basic changes during the transformation from a wild species to a domesticated animal. This has been investigated in chicken (Rubin et al. 2010), dog (vonHoldt et al. 2010), pig (Larson et al. 2010), and sheep (Kijas et al. 2009).

One of the key technologies in the genomic analyses of domestic animals are high-density single-nucleotide polymorphism (SNP) arrays. These usually biallelic markers are distributed over the entire genome and have been found to occur every 100–300 bp in humans (Brown 2002). With a proper coverage to discriminate genuine polymorphisms from sequencing errors (Nielsen et al. 2011), these SNPs have been highly useful to reveal some of the genetic factors underlying human disease (Hirschhorn and Daly 2005). For domestic animals, the characterization of SNPs can be used to determine the genotype associated with specific favored phenotypes, for example, in genome-wide association studies as done in pigs (Amaral et al. 2009), cattle (Eck et al. 2009; Kawahara-Miki et al. 2011) and chicken (Marklund and
Material and Methods

We performed whole-genome shotgun sequencing of a single male Bactrian camel for the identification of heterozygous SNPs and the estimation of populations parameters and sequencing error rate. Genomic DNA (5–10 µg) was extracted from two ethylene diamine tetra-acetic acid blood samples of a single male Bactrian camel originating from the Austrian Zoo Herberstein using the DNaseel blood & tissue kit (Qiagen; Vienna, Austria). Sequencing library preparation (Genomic DNA Sample Prep Kit; Illumina) was performed according to the manufacturer’s protocol. The shotgun sequencing of paired-end (PE) reads (2 × 101 bp) was carried out with an Illumina Genome Analyzer IIx. In the shotgun sequencing method, the DNA is sheared into fragments, which are randomly distributed over the entire genome. The assembly of large genomes with many repetitive sequences such as in mammals (Ridley 1996; Alkan et al. 2011) is challenging and therefore PE sequencing with predefined insert sizes of 400–600 bp were used for the assembly process (Jaffe et al. 2003).

We trimmed the raw reads with a modified Mott algorithm implemented in the PoPoolation software package (Kofler et al. 2011), where low-quality bases with a QPHRED value below 20 at the ends of the reads are removed (Table 1). Using FastQC (Babraham bioinformatics; Cambridge, UK; www.bioinformatics.babraham.ac.uk), we performed an additional quality check of the trimmed reads. As there is no specific Bactrian camel reference genome available, we generated a de novo assembly with CLC Assembly Cell 4.0.1beta (CLC bio; Aarhus, Denmark; www.iclebio.com/). In a first step, we created contig sequences by assembling all the trimmed reads using the paired-end information to resolve longer repeats. Second, we aligned the PE reads with BWA 0.5.7 (Li and Durbin 2009) using the de novo assembled Bactrian camel contig sequences as reference. We applied the following filtering steps on the resulting alignment: we removed 1) reads with a mapping quality QPHRED score lower than 20 and 2) reads that did not map in proper pairs. For downstream analysis, we consequently kept only those reads that mapped uniquely and in proper pairs (Table 1). Identification of heterozygous SNPs was carried out with BCFTools (http://samtools.sourceforge.net/samtools.shtml) implemented in SAMtools 0.1.14 (Li et al. 2009). To be conservative and to avoid false SNPs in repetitive regions,  

Table 1 Summary of read and mapping statistics

| Description                                      | Number     | % of total |
|--------------------------------------------------|------------|------------|
| Number of untrimmed reads in pairs              | 115,328,191| 81.3%      |
| Number of trimmed reads in pairs                 | 93,705,190 | 61.2%      |
| Number of trimmed single reads                   | 15,663,023 | 10.1%      |
| Average length of trimmed reads (bp)             | 88.94      |            |
| Total number of PE reads used for mapping         | 187,410,380|            |
| Total number of reads mapped                      | 115,870,996| 61.8%      |
| Total number of reads mapped in proper pairs      | 114,746,608| 61.2%      |

a One read of the PE reads (given in boldface) was removed due to low quality; note that the trimmed reads in pairs and single do not sum up to 100% because those PE reads that have been discarded as a pair during the trimming process are missing in the count.

b After applying a filtering step for a minimum mapping quality of QPHRED 20.
we employed Repeatmasker 3.2.8 (http://repeatmasker.org) to detect repetitive regions in the \textit{de novo}-assembled Bactrian camel genome and removed SNPs overlapping with these regions using BEDTools 2.15.0 (Quinlan and Hall 2010). Finally we calculated exact two-sided 95% confidence intervals on the remaining set of SNPs using the Clopper-Pearson interval method (Clopper and Pearson 1934) based on the assumption that each SNP is drawn from a binomial distribution. We removed SNPs whose allele frequency was outside the limits of the confidence intervals. To determine the within-individual genetic diversity \( \pi \) (Nei and Li 1979; Nei and Kumar 2000), we jointly estimated the scaled mutation rate \( \theta = 4N_e \mu \) (Johnson and Slatkin 2006; milkRho (Haubold et al. 2010). This program computes the maximum likelihood (ML) estimators of \( \theta, \varepsilon, \) and \( q \) (recombination rate) based on the heterozygous polymorphisms obtained through shotgun sequencing of single diploid individuals. Using this approach, joint estimates for \( \theta, \varepsilon, \) and \( q \) can be calculated without an external (possibly inaccurate) assumption of a sequencing error rate (Lynch 2008). A critical point in the estimation of \( \pi \) is the coverage of sites; it is essential that both nucleotides at a heterozygous locus are sequenced at least twice, so there is sufficient information to distinguish between genuine polymorphisms and sequencing errors. Therefore, we considered only positions with a minimum 4-fold coverage for the ML estimations. In the mlRho program, confidence intervals are determined by calculating the values of an estimator where the likelihood is 2 log units below the maximum (Haubold et al. 2010). For a comparative genomic analysis between the two Old World camelid species, we used BLAT (Kent 2002) to align the available dromedary ESTs (Al-Swailem et al. 2010; Otu, H. personal communication; http://camel.kacst.edu.sa) with the Bactrian camel contig sequences. Finally, we plotted and edited our results with the R software package version 2.11.1 (Hornik 2011).

\section*{Results and Discussion}

Whole-genome shotgun sequencing was used to detect heterozygous SNPs and to estimate population parameters in a single male Bactrian camel. For this purpose, we generated 20 Gb of raw reads by PE read sequencing (\( 2 \times 101 \) bp) on an Illumina Genome Analyzer IIx. After the trimming process, most of the reads displayed a length between 97 and 101 bp (Figure 1a), and the average read length was 88.9 bp (Table 1). The quality \( Q_{\text{PHRED}} \) scores across all bases of the trimmed reads were generally high ranging from 30 to 38, with a decrease to 28 toward the end of the reads (Figure 1b). A summary statistics of the trimming and mapping results is given in Table 1. In absence of a species-specific reference genome, we created a \textit{de novo} assembly and obtained 1.57 Gb of genomic sequence, which corresponds to more than half of the Bactrian camel genome assuming an average mammalian genome size of 3 Gb. The Bactrian camel \textit{de novo} genome assembly consisted of 781 462 contigs with a minimum and maximum contig size of 500 and 36 400 bp, respectively, and an average GC content of 39.6% as shown in Table 2. The N50, which describes that 50\% of the assembly consists of contigs equal or larger than this value, was 2814 bp. The distribution of the contig lengths is displayed in Figure 2. The complete assembly containing all contigs in fasta format has been submitted to the European Nucleotide Archive project (accession number PRJEB407).

Paired-end read mapping against the \textit{de novo}-generated reference assembly resulted in 114.7 million reads (61.1\%) that were uniquely mapped in proper pairs. This seemingly moderate mapping result might be due to the stringent filtering criteria, that is, only reads with a minimum mapping quality of 20 were used (Figure 3). In addition, we have to consider the high frequency of repetitive sequences in the mammalian genome (Charlesworth et al. 1994; Ridley 1996) that renders the mapping of the reads more difficult. Here, we show the number of reads mapping to a contig proportional to the contig lengths (Figure 4). On average, we reached a 6.6-fold coverage of the Bactrian camel genome including only positions up to a coverage of 15-fold (adjusted to approximately twice the average read depth). We used these stringent criteria to account for gene duplications and copy number variations. Applying the same rigorous conditions for SNP calling and using only sites with maximum coverage of 15-fold, we detected 116 313 heterozygous SNPs, whose major allele frequencies lay within a two-sided 95\% confidence interval (see Material and Methods). Recently, 2 129 442 heterozygous SNPs have been described in the domestic Bactrian camel genome (Jirimutu et al. 2012). However, as these sequences have not been made publicly accessible, we could not cross-check our data. Therefore, we compared our results with other domesticated ungulates for which SNP data are available. Whole-genome resequencing of a single \textit{Bos taurus} animal for SNP detection (Eck et al. 2009) resulted in 749 091 heterozygous and 1 694 546 homozygous polymorphisms in comparison with the cattle reference genome bosTaur 4.0 (Elsik et al. 2009). The comparatively lower numbers of heterozygous SNPs in the Bactrian camel analyzed in this study and in the single Fleckvieh cattle (Eck et al. 2009) might be due to the fact that little more than a half of the Bactrian genome has been sequenced and to the low genome sequencing depths of 6.6- and 7.4-fold, respectively. It has been estimated that a 20- to 30-fold coverage is necessary to discriminate 99\% of the heterozygous variants (Li et al. 2008). For sites with a low coverage, the probability of sampling just one of the parental chromosomes is relatively high, which consequently would lead to an underestimation of the heterozygosities and nucleotide diversity (Lynch 2008). Also, both sequenced individuals were male and therefore missing the X-chromosomal heterozygous polymorphisms.

To catalog the genomic diversity within a single individual, we used the program mlRho (Haubold et al. 2010) and obtained a likelihood estimation of the population mutation rate \( \theta = 4N_e \mu \) of \( 1.29 \times 10^{-3} \) (CI: \( 1.28 \times 10^{-3} \); \( 1.30 \times 10^{-3} \)). As the calculations of the estimators are based on assembled shotgun reads, we used only positions with a minimum 4-fold coverage (Haubold et al. 2010). Second-generation
Figure 1. Length and quality distribution of trimmed reads. (A) After trimming, most of the reads showed a length between 97 and 101 bp. (B) The per base quality of the trimmed reads ranged above a QPHRED score of 20. For each position a Box-and-Whisker plot was calculated with FastQC. The upper line displays the median value, while the central line represents the mean quality per base.
sequencing techniques are prone to error rates, possibly in the order of the targeted genetic diversity. Therefore, this software additionally computes a likelihood estimation of the sequencing error rate $\varepsilon$. In the case of the shotgun-sequenced Bactrian camel genome $\varepsilon$ equaled $6.64 \times 10^{-4}$, which was 10 times less than the genuine polymorphisms. The nucleotide diversity $\pi$, also known as the average pairwise nucleotide differences in a sample of DNA sequences (Nei and Li 1979; Nei and Kumar 2000), is a measure of the average genome-wide heterozygosity and can be used as an estimator of the scaled mutation rate $\theta = 4N_e \mu$, assuming random mating and the infinite sites mutation model (Hamilton 2009). In mlRho, $\pi$ is interpreted as $\theta$ without any numerical changes (Haubold et al. 2010). At low coverages, the ML estimates can be downwardly biased (Lynch 2008); therefore, we calculated the nucleotide diversity also with a minimum coverage of 6-, 8-, and 10-fold. However, we saw only a slight increase of $\theta$ at higher coverages, while the estimated error rates remained stable (Figure 5), rendering our results of the within-individual nucleotide diversity as robust. Moreover, our results are comparable to the heterozygosity $(1.0 \times 10^{-3})$ estimated across the whole genome of a domestic Bactrian camel (Jirimutu et al. 2012). Individual-based estimates of heterozygosity can provide information on relative levels of inbreeding in nonrandom mating populations (Lynch 2008). We observed that compared with other domesticated ungulates, the nucleotide diversity in a single Bactrian camel is higher than that in cattle $(9.4 \times 10^{-4}$; Eck et al. 2009) but similar to pig $(1.1–2.1 \times 10^{-3}$; Amaral et al. 2009).

To gain comparative genome-wide information within the Old World camelids, we used the available dromedary ESTs (AL-Swailem et al. 2010; Otu, H. personal communication; http://camel.kacst.edu.sa), which were already assembled in contigs corresponding to genes. We aligned these EST contigs against the Bactrian camel contig sequences and found an overlap of 20,014 (84.8%; see online supplementary Table 1) out of 23,602 genes (AL-Swailem et al. 2010), suggesting that most of the homologous genes are recovered in our genome assembly. Further analyses will be necessary to identify one-to-one ortholog pairs.

**Table 2** Summary of the Bactrian camel assembly statistics

| Total contig bases | 1,568,773,587 bp (1.57 Gb) |
|--------------------|----------------------------|
| Number of contigs  | 781,462                    |
| Number of contigs $> 1$ kb | 510,192 (65.3%) |
| Shortest contig    | 500 bp                      |
| Longest contig     | 36,400 bp                   |
| Average contig length | 2008 bp                    |
| N50 contig length  | 2814 bp                     |
| N90 contig length  | 897 bp                      |
| Average read depth | 6.56-fold                   |

N50/90 refers to a length-weighted median such that 50/90% of the genome is contained in contigs of the indicated size or greater.

$^a$Here we present the total number of bases in the assembly with a base quality $Q_{PHRED}$ score greater than 20 ($Q_{PHRED} > 20$).

$^b$The average read depth was calculated using reads up to a maximal coverage of 15.

![Figure 2. Distribution of the contig length.](chart.png) Distribution of the contig length. While 34.7% of the contigs’ lengths are below 1 kb, the length-weighted median N50 (Table 1) demonstrates that 50% of the assembly consists of contigs with a length equal to or greater than 2814 bp.
Figure 3. Distribution of the mapping quality. Only PE reads with a minimum mapping quality \( q \) of 20 and mapped in proper pairs were used for the analyses. 91.5% of the PE reads show \( q = 60 \). The slightly increased number of reads with \( q \) between 20 and 30 are due to mapping hits in repetitive regions (data not shown).

Figure 4. Number of mapped reads distributed over the contig lengths. The number of reads mapped to a contig increases with the length of the contigs.
Acknowledgments

P. B.; Austrian Science Fund FWF (P21084-B17) to P. B. fellowship of the Austrian Academy of Sciences (11506) to A. Nolte for library preparation and sequence generation. Kofler, T. Burger and Palmieri • Bactrian Camel Whole-Genome Shotgun Sequencing

Supplementary material can be found at http://www.jhered.org/.

Funding

Austrian Programme for Advanced Research and Technology-fellowship of the Austrian Academy of Sciences (11506) to PB; Austrian Science Fund FWF (P21084-B17) to PB.

Acknowledgments

We thank V. Nolte for library preparation and sequence generation and R. Koller, T. Visinovska, and M. Kapun for bioinformatic support. Specifically, we thank R. Pichler (Zoo Herberstein) for sample collection and H. Otu for additional information on the dromedary ESTs.

References

Abdussamad AM, Kalla DJU, Maigandi SA. 2012. Bringing camels into focus: a photo-essay on dromedaries in the niger-niger corridor. In: Knoll EM, Burger PA, editors. Camels in Asia and North Africa. Interdisciplinary perspectives on their significance in past and present. Vienna: Austrian Academy of Science Press. p. 225–230.

Agrawal RP, Jain S, Shah S, Chopra A, Agarwal V. 2011. Effect of camel milk on glycemic control and insulin requirement in patients with type 1 diabetes: 2-years randomized controlled trial. Eur J Clin Nutr. 65:1048–1052.

Alkan C, Sahjadian S, Eichler EE. 2011. Limitations of next-generation genome sequence assembly. Nat Methods. 8:61–65.

Al-Swaleh AM, Shehata MM, Abu-Duhair FM, Al-Yamani EJ, Al-Busalahd KA, Al-Arawi MS, Al-Khider AY, Al-Muhaimed AN, Al-Qattani FH, Manee MM, et al. 2010. Sequencing, analysis, and annotation of expressed sequence tags for Camelus dromedarius. PLoS ONE. 5:e10720.

Amaral AJ, Megens HJ, Kerstens HH, Heuven HC, Dibbits B, Crooijmans RP, den Dunnen JT, Groenen MA. 2009. Application of massive parallel sequencing to whole genome SNP discovery in the porcine genome. BMC Genomics. 10:374.

Amaral AJ, Ferretti I, Megens HJ, Crooijmans RP, Nie H, Ramos-Onsins SE, Perez-Encisco M, Schook LB, Groenen MA. 2011. Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA. PLoS ONE. 6:e14782.

Andersson L, Georges M. 2004. Domestic-animal genomics: deciphering the genetics of complex traits. Nat Rev Genet. 5:202–212.

Brown TA. 2002. Genomes. Oxford: Bios Scientific Publishers.

Bullert R. 1990. The camel and the wheel. New York: Columbia University Press.

Charlesworth B, Sniegowski P, Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature. 371:215–220.

Clopper C, Pearson ES. 1934. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika. 26:404–413.

Eck SH, Benet-Pagés A, Fliškowskii K, Meitinger T, Fries R, Strom TM. 2009. Whole genome sequencing of a single Bos taurus animal for single nucleotide polymorphism discovery. Genome Biol. 10:R82.

Elshl CG, Tellam RL, Worley KC, Gibbs RA, Muzny DM, Weinstock GM, Adelson DJ, Eichler EE, Elnitski L, Guigo R, et al. 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science. 324:522–528.

Fan B, Du ZQ, Gorbach DM, Rothschild MF. 2010. Development and application of high-density SNP arrays in genomic studies of domestic animals. Asia-Austr J Anim Sci. 23:833–847.

Grigson C. 2012. Camels, copper and donkeys in the early iron age of the southern levant: timna revisited. Levant. 44:82–100.

Hamilton MB. 2009. Population genetics. West Sussex: Wiley-Blackwell.

Haubold B, Pfaffelhuber P, Lynch M. 2010. mlRho - a program for estimating the population mutation and recombination rates from shotgun-sequenced diploid genomes. Mol Ecol. 19:277–284.

Hirschhorn JN, Daly MJ. 2005. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet. 6:95–108.

Hornik K. 2011. The R FAQ. ISBN 3-900051-08-9. Available from: URL http://CRAN.R-project.org/doc/FAQ/R-FAQ.html.

Jaffe DB, Butler J, Gnerre S, Anarde MA, Nix DA, Lindblad-Toh K, Mikkelsen DT, Meyers MA, Zhang XH, Bernstein H, et al. 2013. Genome sequence assembly and annotation of the domestic horse genome. Nat Commun. 4:2264.

Johnson PL, Slavkin M. 2006. Inference of population genetic parameters in metagenomics: a clean look at messy data. Genome Res. 16:1320–1327.
Kawahara-Miki R, Tsuda K, Shiwa Y, Arai-Kichise Y, Matsumoto T, Kanesaki Y, Oda S, Ebihara S, Yajima S, Yoshikawa H, et al. 2011. Whole-genome resequencing shows numerous genes with nonsynonymous SNPs in the Japanese native cattle Kuchinoshima-Ushi. BMC Genomics. 12:103.

Kent WJ. 2002. BLAT—the BLAST-like alignment tool. Genome Res. 12:656–664.

Kijas JW, Townley D, Dalrymple BP, Heaton MP, Maddox JF, McGrath A, Wilson P, Ingersoll RG, McCulloch R, McWilliam S, et al. 2009. A genome wide survey of SNP variation reveals the genetic structure of sheep breeds. PLoS ONE. 4:e4668.

Kofler R, Orozco-terWengel P, De Maio N, Pandey RV, Noite V, Futschik A, Kosiol C, Schlötterer C. 2011. PoPoolation: a toolbox for population genetic analysis of next generation sequencing data from pooled individuals. PLoS ONE. 6:e15925.

Larson G, Liu R, Zhao X, Yuan J, Fuller D, Barton L, Dobney K, Fan Q, Gu Z, Liu X, et al. 2010. Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. Proc Natl Acad Sci USA. 107:7686–7691.

Li H, Ruan J, Durbin R. 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. 18:1851–1858.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 25:2078–2079.

Lynch M. 2008. Estimation of nucleotide diversity, disequilibrium coefficients, and mutation rates from high-coverage genome-sequencing projects. Mol Biol Evol. 25:2409–2419.

Marklund S, Carlborg O. 2010. SNP detection and prediction of variability between chicken lines using genome resequencing of DNA pools. BMC Genomics. 11:665.

Mayklemans S, Baral TN, Retamozzo VC, De Baetselier P, De Genst E, Kinne J, Leonhardt H, Magez S, Nguyen VK, Revers H, et al. 2009. Camnelid immunoglobulins and nanobody technology. Vet Immunol Immunopathol. 128:178–183.

Nawy T. 2012. Rare variants and the power of association. Nat Methods. 9:324.

Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA. 76:5269–5273.

Nei M, Kumar S. 2000. Molecular Evolution and Phylogenetics. New York: Oxford University Press.

Nielsen R, Paul JS, Albrechtsen A, Song YS. 2011. Genotype and SNP calling from next-generation sequencing data. Nat Rev Genet. 12:443–451.

Peters J, von den Driesch A. 1997. The two-humped camel (Camelus bactrianus): new light on its distribution, management and medical treatment in the past. J Zool. 242:651–679.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26:841–842.

Ridley M. 1996. Evolution. Cambridge (MA): Blackwell Science Inc.

Rubin CJ, Zody MC, Eriksson J, Meadows JR, Sherwood E, Webster MT, Jiang L, Ingrman M, Sharpe T, Ka S, et al. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. Nature. 464:587–591.

Uerpmann HP, Uerpmann M. 2002. The appearance of the domestic camel in south-east Arabia. J Oman Stud. 12:235–260.

Vonholdt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, et al. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature. 464:898–902.

Wenbo L, Dongfeng L, Jianfeng L, Siurui C, Lujiang Q, Jiangxia Z, Guiyun X, Ning Y. 2011. A genome-wide snp scan reveals novel loci for egg production and quality traits in white leghorn and brown-egg dwarf layers. PLoS One. 6:e28600.

Received February 22, 2012; First decision April 17, 2013; Accepted January 25, 2013