

Intra-population SNP analysis identifies signatures for grazing stress tolerance in Egyptian local goats

J.M. Mwacharo¹, A.R. Elbeltagy², E.-S. Kim³, A.M. Aboul-Naga², B. Rischkowsky¹ & M.F. Rothschild³

¹Small Ruminant Genetics and Genomics Group, International Center for Agricultural Research in the Dry Areas (ICARDA), P. O. Box 5689 Addis Ababa, Ethiopia.

j.mwacharo@cgiar.org (Corresponding author)

²Animal Production Research Institute (APRI), Agriculture Research Centre (ARC), Ministry of Agriculture. Nadi Elsaid Street, Dokki, Cairo, Egypt.

³Department of Animal Science, Iowa State University 2255 Kildee Hall, Ames, IA 50011-3150. USA.

Summary

The need to support agricultural industries and human livelihoods in dryland environments using adapted livestock is becoming an unavoidable priority as prospects of global warming become even more evident. Here, using the integrated haplotype score (*iHS*) approach, we analysed genotype data generated using the caprine 50K Illumina SNP BeadChip from 366 individuals of four local Egyptian desert goat populations to investigate the genetic basis of tolerance to grazing stress. The analysis identified 13 selection sweep regions that spanned 68 genes across eight chromosomes. The strongest selection sweeps were observed on chromosomes 6 and 12 which, spanned 21 and 17 genes, respectively. Though a majority of the identified genes were associated with multiple functions, functional analysis of the 38 genes found in the two strongest candidate regions, revealed the three most significant gene ontology functional term clusters were associated with biological processes of kidney function and development, cellular response to oxidative stress and intra-cellular calcium homeostasis. This suggests that water and mineral homeostasis and the proper maintenance of cellular responses to external stimuli and processes that modulate calcium homeostasis and other signal transduction pathways are key functions in tolerance to grazing stress. Our results could prove useful in genomic selection and identification of genes involved in desert grazing stress tolerance.

Key words: adaptation, caprine, genotyping, SNP, selection sweeps,

Introduction

Drylands (dry sub-humid, semi-arid, arid and hyper-arid environments) cover more than 41.3% of the earth's land surface, are found in all continents and support more than 51% of the global livestock population. There is evidence showing that the actual land mass that can be considered arid or semi-arid is increasing (UNCCD, 2010). It is generally agreed that the adoption of any strategy(ies) to increase agricultural production in drylands provides a much needed route to adapt to the effects of rapid climate change, which otherwise will become a major obstacle to attaining sustainable livelihoods (<http://www.fao.org/docrep/012/i0372e/i0372e01.pdf>).

However, the development of livestock populations that are resilient enough to adapt to the effects of global warming while supporting livestock production in dryland communities, have many biological challenges. By their very nature, adaptation traits are complex and difficult to improve using conventional animal breeding. However, genomic tools offer opportunities to overcome these challenges. Grazing stress, which characterizes dryland environments, incorporates a complex of interacting biophysical stressors including thermal stress, direct solar radiation, physical exhaustion and acute and chronic water and feed (quality and quantity) shortages. It can be hypothesized that the genomes of long-term dryland dwelling local livestock populations have been shaped by grazing stress. Therefore, the genomes of desert dwelling livestock populations that have evolved under long-term grazing stress are very likely to carry unique chromosomal regions that confer grazing stress tolerance. Identification of such regions may prove valuable in developing livestock breeds and populations that can be on the frontline of aligning breeding goals and objectives to climate change. Here, we investigated the genetic basis for tolerance to grazing stress using Egyptian local desert goats through a genome-wide analysis approach. Our results reveal the complexity of genome architecture with respect to adaptation.

Materials and methods

Whole blood was collected from 366 individuals from four local Egyptian desert goat populations (Barki = 150, Saidi = 60, Farafra = 72 and Nubian = 84). The animals were sampled at random from farmers' flocks where anthropic selection is modest. Genotyping was performed on the Illumina CaprineSNP52K BeadChip assay at GeneSeek Inc (<http://www.neogen.com/Genomics/>). Signatures of selection were studied using the integrated haplotype score (*iHS*; Voight et al 2006) approach using the *rehh* package in R. The *iHS* detects long conserved stretches of chromosomal regions relative to neutral expectations, that are likely associated with selection constraints, which in our case was grazing stress tolerance. The *iHS* score was computed for each SNP with $MAF \geq 0.03$. The $-\log_{10}(P\text{-value}) > 4.0$ (equivalent to $P\text{-value} < 0.0001$), was used to define significant *iHS* scores. Candidate regions were identified if at least two SNPs passed the significance threshold. If multiple SNPs passed the threshold, a distance of 0.25 Kb up- and down-stream of the most significant SNP was used to define candidate region intervals. This distance was chosen based on the rate of change in the mean pairwise linkage disequilibrium statistic (r^2) binned over distance intervals across autosomes of the study populations. The candidate regions were investigated for the presence of functional genes against the *Capra hircus* V1 (ARS1 (GCF_001704415.1)) genome assembly. Functional annotation and Gene Ontology (GO) term enrichment analysis was performed with ENRICHR (Kuleshov et al. 2016).

Results and discussion

We observed one candidate selection sweep region located on chromosome XX that spanned no annotated genes. Such regions have also been reported in cattle (Xu et al. 2015; Iso-Touru et al. 2016) and can be attributed to the incomplete annotation of the genomes. Genome-wide, we identified 68 genes mapping to 13 candidate selection sweep regions across eight chromosomes (CHI5, CHI6, CHI7, CHI11, CHI12, CHI16, CHI20 and CHI25; Figure 1). The large number of identified candidate regions and genes suggest that tolerance to grazing stress involves a

complex network of genes acting interactively and which influence multiple traits (Lv et al. 2014; Kemper et al. 2014). Lv et al. (2014) observed that adaptation is the product of the interaction of several complex traits that most often are controlled by several genes, while Kemper et al. (2014) noted that selection for complex traits leaves little or no classic selection sweeps owing to weak selection acting on several sites across the genome.

The strongest selection sweeps were on CHI6 and CHI12 and were defined by 19 and 18 significant SNPs, respectively (Figure 1). The region on CHI6 spanned 21 genes (Table 1) while 17 genes occurred in the region on CHI12 (Table 1). We hypothesize that these two strong candidate regions and their associated genes could be the primary drivers behind the adaptation of the study populations to grazing stress.

Functional annotation analysis of the 38 genes found on CHI6 and CHI12 revealed several functional term clusters; the top four most significant terms were associated with the biological processes of kidney function and development (GO:0035502 - metanephric part of ureteric bud development; GO:0001657 - ureteric bud development), response to oxidative stress (GO:0034614 - cellular response to reactive oxygen species) and intra-cellular calcium homeostasis (GO:0099093 - mitochondrial calcium release) (Table 2). Interestingly, these three biological processes are closely associated with the major components of grazing stress.

Renal vasodilation, ion transmembrane transport, water-salt metabolism, bicarbonate absorption, water retention and reabsorption are key functions for the kidney and renal cortex (Gatalica et al. 2008) and selection sweeps spanning genes regulating water retention and reabsorption in renal cells and blood vessels in the kidney have been detected in desert adapted sheep (Yang et al. 2017) and camelidae (Wu et al. 2014). Similarly, long term exposure to thermal stress and direct solar radiation results in the production of reactive oxygen species (e.g. O_2^- , OH and H_2O_2) which can damage various cellular components (proteins, lipids, nucleic acids and key organelles such as mitochondria) therefore impairing essential molecular functions such as nucleic acid synthesis and repair, cell subdivision and energy biosynthesis (See reviews by Slimen et al. 2016; Al-Dawood 2017). Oxidative stress can also result in apoptosis and necrosis. Intracellular Ca^{2+} signaling is versatile and fundamental for the regulation of multiple cellular processes, including cell metabolism, differentiation, proliferation, secretion, gene activation and cell death. Mitochondria modulates the amplitude and the spatio-temporal patterns of Ca^{2+} signaling which controls cellular energy metabolism (ATP production), survival and death (Patergnani et al. 2011). Our findings therefore suggest that the proper functioning of kidneys to maintain water and electrolyte balance, and of multiple cellular processes to maintain genome integrity in the face of oxidative stress and the action of mitochondria through the modulation of Ca^{2+} signals are central to grazing stress adaptation. Our findings could prove useful in genomic selection and in the identification of genes to improve grazing stress tolerance and other economic traits, to support livestock production industry in regions that are likely to be affected by worsening climates.

List of References

- Al-Dawood A., 2017. Towards heat stress management in small ruminants – A review. *Ann. Anim. Sci.* 17: 59-88.
- Gatalica, Z., S.L., Lilleberg, M.S., Koul, T., Vanecek, O., Hes, B., Wang, & M. Michal, 2008. COX2 gene polymorphisms and protein expression in renomedullary interstitial cell tumors. *Hum. Pathol.* 39: 1495-1504.
- Iso-Touru, T. M., Tapio, J., Vilkki, T., Kiseleva, I., Ammosov, Z., Ivanova, R., Popov, M., Ozerov, J.,

- Kantanen, 2016. Genetic diversity and genomic signatures of selection among cattle breeds from Siberia, eastern and northern Europe. *Anim. Genet.* 47: 647-657.
- Kemper, K.E., S.J. Saxton, Bolormaa S, Hayes BJ, Goddard ME (2014). Selection for complex traits leaves little or no classic signatures of selection. *BMC Genomics* 15: 246.
- Kuleshov MV, M.R. Jones, A.D. Rouillard, N.F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S.L. Jenkins, K.M. Jagodnik, A. Lachmann, M.G. McDermott, C.D. Monteiro, G.W. Gundersen, A. Ma'ayan, 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* 44: W90–W97.
- Lv F.-H., S. Agha, J. Kantanen, L. Colli, S. Stucki, J.W. Kijas, S. Joost, M.H. Li, P. Ajmone-Marsan, 2014. Adaptations to climate-mediated selective pressures in sheep. *Mol. Biol. Evol.* 31: 3324–3343.
- Petergnani S., J.M. Suski, C. Agnoletto, A. Bononi, M. Bonora, E.D. Marchi, C. Giorgi, S. Marchi, S. Missiroli, F. Poletti, A. Rimessi, J. Duszynski, M.R. Wieckowski, P. Pinton, 2011. Calcium signaling around mitochondria associated membranes (MAMs). *Cell Commun. Signal.* 9: 19.
- Slimen I.B., T. Najar, A. Ghram, M. Abdrrabba, 2016. Heat stress effects on livestock: molecular, cellular and metabolic aspects, a review. *J. Anim. Physiol. Anim. Nutr.* 100: 401-412.
- UNCCD, 2010. United Nations Decade for Deserts and the Fight Against Desertification. United Nations Convention to Combat Desertification, Bonn, Germany.
- Voight B.F., S. Kudravalli, X. Wen, J.K. Pritchard, 2006. A map of recent positive selection in the human genome. *PLoS Biol.* 4: 446-458.
- Wu H., X. Guang, M.B. Al-Fageeh, J. Cao, S. Pan, H. Zhou, L. Zhang, M.H. Abutarboush, Y. Xing, Z. Xie, A.S. Alsharqeti, Y. Zhang, Q. Yao, B.M. Al-Shomrani, D. Zhang, J. Li, M.M. Manee, Z. Yang, L. Yang, Y. Liu, J. Zhang, M.A. Altammami, S. Wang, L. Yu, W. Zhang, S. Liu, L. Ba, C. Liu, X. Yang, F. Meng, S. Wang, L. Li, E. Li, X. Li, K. Wu, S. Zhang, J. Wang, Y. Yin, H. Yang, A.M. Al-Swailem, J. Wang, 2014. Camelid genomes reveal evolution and adaptation to desert environments. *Nat. Commun.* 5: 5188.
- Xu L., D.M Bickhart, J.B. Cole, S.G. Schroeder, J. Song, C.P. Tassell, T.S. Sonstegard, G.E. Liu, 2015. Genomic signatures reveal new evidences for selection of important traits in domestic cattle. *Mol. Biol. Evol.* 32: 711 –725.
- Yang J., W.R. Li, F.H. Lv, S.G. He, S.L. Tian, W.F. Peng, Y.W. Sun, Y.X. Zhao, X.L. Tu, M. Zhang, X.L. Xie, Y.T. Wang, J.Q. Li, Y.G. Liu, Z.Q. Shen, F. Wang, G.J. Liu, H.F. Lu, J. Kantanen, J.L. Han, M.H. Li, M.J Liu, 2016. Whole-genome sequencing of native sheep provides insights into rapid adaptations to extreme environments. *Mol. Biol. Evol.* 33: 2576-2592.

Table 1. The strongest candidate selection sweep regions and associated genes identified by *iHS* analysis in four Egyptian goat populations.

Chromosome	Location of genomic region (bp)		Number of significant SNPs	Genes present
	Start	Stop		
6	30,246,187	47,044,407	19	<i>UNC5C, BMPR1B, CCSER1, MMRN1, SNCA, HERC3, HERC6, PPM1K, PKD2, SPP1, FAM184B, LAP3, MED28, DCAF16, NCAPG, SLIT2, PACRGL, KCNIP4, PPARGC1A, ZCCHC4, ANAPC4</i>
12	34,295,846	36,783,669	18	<i>ENSBTAG00000019545, MIPEP, TNFRSF19, NUP58, MTMR6, SACS, SGCG, ENSBTAG00000048237, MICU2, LATS2, XPO4, CRYL1, CENPJ, RNF17, MPHOSPH8, PARP4, ATP12A</i>

Table 2. Significantly enriched functional term clusters and their significance as determined by ENRICH analysis of genes in the two strongest candidate regions identified by *iHS* analysis.

Functional (GO) term cluster	<i>P</i> - value	Associated genes
Metanephric part of ureteric bud development	0.00047	<i>SLIT2, PKD2</i>
Ureteric bud development	0.00041	<i>SLIT2, PKD2</i>
Cellular response to reactive oxygen species	0.00112	<i>PKD2, PPARGC1A, SNCA</i>
Mitochondrial calcium release	0.00288	<i>MICU2, PKD2</i>

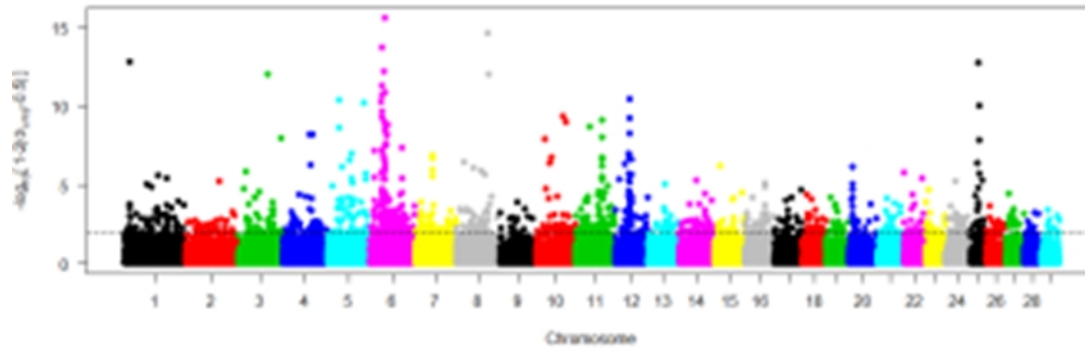


Figure 1. Manhattan plots for the genome-wide distribution of SNPs following selection sweep analysis based on *iHS* analysis.