Perspectives

Structural and functional genomics in Old World camels—where do we stand and where to go

Pamela A. Burger† and Elena Ciani‡

†Department of Interdisciplinary Life Sciences, Research Institute of Wildlife Ecology, University of Veterinary Medicine, 1160 Vienna, Austria
‡Department of Biosciences, Biotechnologies and Biopharmaceutics, University of Bari “Aldo Moro”, Bari, Italy

Key words: genome, livestock, phenome, polymorphism

Introduction

Understanding the structure and function of genomes is important for investigating interactions among genes, between genes and the environment, and deciphering complex traits. With climate change and increasing human population size, it is more important than ever to conserve biodiversity and to improve efficiency in animal health and production using fewer natural resources. This can be achieved by increasing the diversification of locally adapted (livestock) species and by improving the ability to use genotypes to accurately predict relevant adaptive and production phenotypes (Clark et al., 2020). A high-quality reference genome assembly is a prerequisite to initiate functional genome annotation. Significant progress has been made in this direction by sequencing whole animal genomes, detecting sequence variants, associating them to phenotypic traits, and using genomic variation to select for predicted genetic differences in routinely measured traits. Here, we provide a brief overview of the current situation and future challenges in structural and functional genomics in Old World camels.

Implications

Genome-to-phenome research will boost
• our knowledge on molecular mechanisms underlying adaptive and production phenotypes,
• whole-genome-enabled animal selection
• genetic diversity conservation.

Where Do We Stand in Structural and Functional Genomics in Camels

In Old World camels, de novo whole-genome sequencing including shot-gun, long reads, and structural Hi-C mapping was used to build high-quality reference genomes on chromosome levels (Elbers et al., 2019; Ming et al., 2020 and references therein). Multiomic, i.e., transcriptomic and proteomic analyses of dromedary (Alvira-Iraizoz et al., 2021) and Bactrian camel (Wu et al., 2014) kidney tissues complemented and advanced these important first steps towards understanding the structure and function of camelid genomes. Building on these newly available high-quality reference genomes, polymorphism and structural and functional analyses of targeted regions were performed, and initial genomewide association studies were launched. In the following, we briefly summarize the main structural and functional research areas developed so far in Old World camels, and example studies and corresponding references are presented in Table 1.

Structure and polymorphisms in immune response genes

Major parts of the immunome have been resolved including the major histocompatibility complex (MHC), Natural Killer cell and T-cell receptor genes, and the organization of two immunoglobulin light-chain loci. Comparative analyses of innate and adaptive immune response (IR) genes showed a general low diversity among Old World camels.

Environmental adaptation

Comparative genomics revealed complex pathways related to desert adaptations, including fat and water metabolism, stress responses to heat, aridity, intense ultraviolet radiation, and dust. Genes under potential selection in African and Asian dromedaries were involved in inflammatory responses of bacterial and fungal infections as well as the immune system and the circadian rhythm.

Adaptation to heat stress

Camels have developed adaptive mechanisms to heat stress including biochemical and physiological processes. Molecular
| Structural or functional genomic area | Main addressed topic | Reference |
|--------------------------------------|----------------------|-----------|
| **Genome assembly, annotation, and structure** | High-quality chromosome-level Bactrian camel reference genome, immunome | Ming et al. (2020), *Mol Ecol Res* doi:10.1111/1755-0998.13141 |
| | Improved annotation of dromedary reference genome, immunome | Lado et al. (2020), *BMC Genom* doi:10.1186/s12864-020-06990-4 |
| | High-quality chromosome-level dromedary reference genome | Elbers et al. (2019), *Mol Eol Res* doi:10.1111/1755-0998.13020 |
| | Bactrian camel genome assembly at scaffold level | Wu et al. (2014), *Nat Comm* doi:10.1038/ncomms6188 |
| | Dromedary genome assembly at scaffold level | Fitak et al. (2016), *Mol Ecol Res* doi:10.1111/1755-0998.12443 |
| | Bactrian camel genome assembly at scaffold level | Jirimutu et al. (2012), *Nat Comm* doi:10.1038/ncomms3089 |
| | Radiation hybrid (RH) clones to prepare RH dromedary genome map | Perelman et al. (2018), *Sci Rep* doi:10.1038/s41598-018-20223-5 |
| **Transcriptome, general expression studies** | Multiomic analysis of dromedary kidney | Alvira-Iraizoz et al. (2021), *Com Biol* doi:10.1038/s42003-020-02327-3 |
| | Differential expression of Bactrian camel renal medulla | Wu et al. (2014), *Nat Comm* doi:10.1038/ncomms6188 |
| | Expressed sequence tags of dromedary | Al-Swailem et al. (2010), *PLoS One* doi:10.1371/journal.pone.0010720 |
| **Polymorphism and structure in immune response genes** | Cytotoxic Effector Proteins in camels | Futas et al. (2021), *Genes* doi:10.3390/genes12020304 |
| | Immune response genes related to Middle East Respiratory Syndrome (MERS) in dromedaries | Lado et al. (2021), *Cells* doi:10.3390/cells10061291 |
| | Immune response genes related to Crimean-Congo hemorrhagic fever virus infection in dromedaries | Lado et al. (2021), *Cells* doi:10.3390/cells11010008 |
| | Immune response genes related to Middle East Respiratory Syndrome (MERS) in dromedaries | Lado et al. (2020), *BMC Genom* doi:10.1186/s12864-020-06990-4 |
| | Genome analysis of dromedary T-cell receptor gamma (TRG) locus | Antonacci et al. (2020), *Dev Com Imm* doi:10.1016/j.dev.imm.2020.103614 |
| | Camel adaptive immune response receptor reservoir | Ciccarese et al. (2019), *Front Gen* doi:10.3389/fgene.2019.00997 |
| | Natural Killer Cell receptor genes in camels | Futas et al. (2019), *Front Gen* doi:10.3389/fgene.2019.00620 |
| | Major histocompatibility complex in camels | Plasil et al. (2019), *HLA* doi:10.1111/tan.13510 |
| | | Plasil et al. (2019), *Cells* doi:10.3390/cells8101200 |
| | | Plasil et al. (2016), *BMC Genom* doi:10.1186/s12864-016-2500-1 |
| **Environmental adaptation** | Differentially selected genes between African and Asian dromedaries | Lado et al. (2020), *Commun Biol* doi:10.1038/s42003-020-1098-7 |
| | Adaptation to desert environment in Bactrian camels | Wu et al. (2014), *Nat Comm* doi:10.1038/ncomms6188 |
| **Adaptation to heat stress, heat shock proteins** | Dromedary as livestock model for heat resistant | Tibary et al. (2020), *Theriogen* doi:10.1016/j.theriogenology.2020.05.046 |
| | Molecular adaptation to heat stress in dromedaries | Hotter et al. (2019), *Front Gen* doi:10.3389/fgene.2019.00588 |
| | Thermotolerance of camel somatic cells | Saadelin et al. (2019), *J Adv Res* doi:10.1016/j.jare.2019.11.009 |
| | Sequence analysis of heat shock protein beta-1 in dromedary | Manee et al. (2017), *PLoS One* doi:10.1371/journal.pone.0189905 |
| | Sequence and expression of heat shock protein 90-alpha | Saeed et al. (2015), *Int J Biol Macromol* doi:10.3390/ijms13074214 |
| **Genomics pathways selected in domestication** | Selection signals during domestication in camels | Fitak et al. (2020), *Comm Biol* doi:10.1038/s42003-020-1039-5 |
aspects involved different heat shock protein expression patterns.

**Genomic pathways involved in domestication**

Selection signals in camels revealed the neural crest cell and thyroid hormone pathways to be involved in the process of domestication.

**Morphological phenotypes, growth, and coat color**

Low-density single-nucleotide polymorphism analysis revealed genes related to growth in Iranian dromedaries. Polymorphisms in *MC1R* and *ASIP* genes were associated with coat color variation in dromedaries.

### Production phenotypes, milk, and performance

Proteomic approaches resolved the complexity of the milk protein fraction of dromedary, Bactrian camels, and hybrids. Cryptic and alternative splice sites in mRNAs in camel milk were investigated as well as the structure and variation in alpha-, beta-, and kappa-casein genes. Genes related to chondrogenesis, energy balance, and urinary system development were under selection between packing and racing dromedaries.

The myostatin gene, a negative regulator of skeletal muscle mass in animals that has a role in determining muscular hypertrophy, was sequenced in camels, and its expression levels in different muscle tissue were established. The gut microbiome was characterized by high-throughput sequencing of 16S rRNA in Bactrian camels at different age and gastrointestinal segments.
Finally, genes and quantitative trait loci (QTLs) related to other physiological functions were identified involving insulin resistance and hematological traits in Bactrian camels, loss of keratin genes in three evolutionary lineages of mammals (including dromedary), and the oxytocin-neurophysin I (OXT) gene with its regulatory regions.

Where Do We Need to Go in Structural and Functional Genomics

To advance structural and functional genomics, that is, genome-to-phenome research in Old World camels, we suggest to follow the path of the Functional Annotation of Animal Genomes (FAANG) project (Clark et al., 2020 and references therein). Focusing on 1) pangenomes and comparative genomics, 2) large-scale genotype-to-expression characterization, 3) large phenotype collections, and 4) germplasm banking will accelerate genome-to-phenome research to improve the use and conservation of Old World camels’ genomic diversity in relation with adaptation, production, and sustainability traits. Initial pangenomic studies have been done, with notable examples represented by whole-genome re-sequencing of 105 Asian Camelus bactrianus genomes (NCBI-SRA accession number SRP107089) and hundreds of African and Asian Camelus dromedarius genomes (Illumina Eleventh Agricultural Greater Good Initiative). While these projects covered major geographic regions/countries and identified millions of single nucleotide polymorphisms (SNPs), amenable to implementation within array-based genotyping platforms, future studies shall focus on in-depth analyses of within-population/within-country genetic diversity, to capture fine-scale genomic variation and population structures. This will assist prioritization of in vitro germplasm conservation of endangered stocks. The availability of medium-density SNP arrays and the perspective for high-density formats will foster more detailed genomic sequence variation catalogs. Further reducing genotyping platform costs will allow affordable technology access to less-developed countries and low-added value industries such as milk and meat commodities in marginal areas.

The wealth of data that are (will be) available for the two domestic Old World camel species deserves further functional and comparative genome mining, also exploiting existing whole-genome sequence collections for the wild counterpart (Camelus ferus). This will improve gene annotation, understanding of protein domain architectures and their evolution across taxa, as well as the identification of conserved, or peculiar, regulatory elements and signaling pathways.

Over the last decade, genome wide association study (GWAS) has allowed associating genetic loci to various livestock quantitative traits; similar expectations exist for Old World camels, provided that reliable large-scale phenotyping can successfully be implemented in these species. Notably, this is not trivial under the currently practiced extensive/semi-extensive, highly-mobile farming systems. Understanding the underlying functional mechanisms in camels has been limited so far, hampering the translation of the associations into action. Studies of expressed QTL (eQTL) enrichment among trait-associated variants have highlighted the importance of gene expression regulation in phenotypic variability. Hence, the next step in transcriptomic studies is to integrate large data from different biological layers, that is, by ex-post integrating eQTL information in GWAS results based on co-localization principles or by ex-ante integrating information whether an SNP regulates the expression of a gene into GWAS. This will increase the power to identify trait-associated loci and provide mechanistic insights. Further large-scale efforts in extending the current Old World camels transcriptome dataset in terms of the number and type of examines tissues and animals under different eco-climatic conditions number SRP107089).

Where Do We Need to Go in Structural and Functional Genomics

To advance structural and functional genomics, that is, genome-to-phenome research in Old World camels, we suggest to follow the path of the Functional Annotation of Animal Genomes (FAANG) project (Clark et al., 2020 and references therein). Focusing on 1) pangenomes and comparative genomics, 2) large-scale genotype-to-expression characterization, 3) large phenotype collections, and 4) germplasm banking will accelerate genome-to-phenome research to improve the use and conservation of Old World camels’ genomic diversity in relation with adaptation, production, and sustainability traits. Initial pangenomic studies have been done, with notable examples represented by whole-genome re-sequencing of 105 Asian Camelus bactrianus genomes (NCBI-SRA accession number SRP107089) and hundreds of African and Asian Camelus dromedarius genomes (Illumina Eleventh Agricultural Greater Good Initiative). While these projects covered major geographic regions/countries and identified millions of single nucleotide polymorphisms (SNPs), amenable to implementation within array-based genotyping platforms, future studies shall focus on in-depth analyses of within-population/within-country genetic diversity, to capture fine-scale genomic variation and population structures. This will assist prioritization of in vitro germplasm conservation of endangered stocks. The availability of medium-density SNP arrays and the perspective for high-density formats will foster more detailed genomic sequence variation catalogs. Further reducing genotyping platform costs will allow affordable technology access to less-developed countries and low-added value industries such as milk and meat commodities in marginal areas.

The wealth of data that are (will be) available for the two domestic Old World camel species deserves further functional and comparative genome mining, also exploiting existing whole-genome sequence collections for the wild counterpart (Camelus ferus). This will improve gene annotation, understanding of protein domain architectures and their evolution across taxa, as well as the identification of conserved, or peculiar, regulatory elements and signaling pathways.

Over the last decade, genome wide association study (GWAS) has allowed associating genetic loci to various livestock quantitative traits; similar expectations exist for Old World camels, provided that reliable large-scale phenotyping can successfully be implemented in these species. Notably, this is not trivial under the currently practiced extensive/semi-extensive, highly-mobile farming systems. Understanding the underlying functional mechanisms in camels has been limited so far, hampering the translation of the associations into action. Studies of expressed QTL (eQTL) enrichment among trait-associated variants have highlighted the importance of gene expression regulation in phenotypic variability. Hence, the next step in transcriptomic studies is to integrate large data from different biological layers, that is, by ex-post integrating eQTL information in GWAS results based on co-localization principles or by ex-ante integrating information whether an SNP regulates the expression of a gene into GWAS. This will increase the power to identify trait-associated loci and provide mechanistic insights. Further large-scale efforts in extending the current Old World camels transcriptome dataset in terms of the number and type of examines tissues and animals under different eco-climatic conditions number SRP107089).

Finally, genes and quantitative trait loci (QTLs) related to other physiological functions were identified involving insulin resistance and hematological traits in Bactrian camels, loss of keratin genes in three evolutionary lineages of mammals (including dromedary), and the oxytocin-neurophysin I (OXT) gene with its regulatory regions.

Where Do We Need to Go in Structural and Functional Genomics

To advance structural and functional genomics, that is, genome-to-phenome research in Old World camels, we suggest to follow the path of the Functional Annotation of Animal Genomes (FAANG) project (Clark et al., 2020 and references therein). Focusing on 1) pangenomes and comparative genomics, 2) large-scale genotype-to-expression characterization, 3) large phenotype collections, and 4) germplasm banking will accelerate genome-to-phenome research to improve the use and conservation of Old World camels’ genomic diversity in relation with adaptation, production, and sustainability traits. Initial pangenomic studies have been done, with notable examples represented by whole-genome re-sequencing of 105 Asian Camelus bactrianus genomes (NCBI-SRA accession number SRP107089) and hundreds of African and Asian Camelus dromedarius genomes (Illumina Eleventh Agricultural Greater Good Initiative). While these projects covered major geographic regions/countries and identified millions of single nucleotide polymorphisms (SNPs), amenable to implementation within array-based genotyping platforms, future studies shall focus on in-depth analyses of within-population/within-country genetic diversity, to capture fine-scale genomic variation and population structures. This will assist prioritization of in vitro germplasm conservation of endangered stocks. The availability of medium-density SNP arrays and the perspective for high-density formats will foster more detailed genomic sequence variation catalogs. Further reducing genotyping platform costs will allow affordable technology access to less-developed countries and low-added value industries such as milk and meat commodities in marginal areas.

The wealth of data that are (will be) available for the two domestic Old World camel species deserves further functional and comparative genome mining, also exploiting existing whole-genome sequence collections for the wild counterpart (Camelus ferus). This will improve gene annotation, understanding of protein domain architectures and their evolution across taxa, as well as the identification of conserved, or peculiar, regulatory elements and signaling pathways.

Over the last decade, genome wide association study (GWAS) has allowed associating genetic loci to various livestock quantitative traits; similar expectations exist for Old World camels, provided that reliable large-scale phenotyping can successfully be implemented in these species. Notably, this is not trivial under the currently practiced extensive/semi-extensive, highly-mobile farming systems. Understanding the underlying functional mechanisms in camels has been limited so far, hampering the translation of the associations into action. Studies of expressed QTL (eQTL) enrichment among trait-associated variants have highlighted the importance of gene expression regulation in phenotypic variability. Hence, the next step in transcriptomic studies is to integrate large data from different biological layers, that is, by ex-post integrating eQTL information in GWAS results based on co-localization principles or by ex-ante integrating information whether an SNP regulates the expression of a gene into GWAS. This will increase the power to identify trait-associated loci and provide mechanistic insights. Further large-scale efforts in extending the current Old World camels transcriptome dataset in terms of the number and type of examines tissues and animals under different eco-climatic conditions number SRP107089).

Finally, genes and quantitative trait loci (QTLs) related to other physiological functions were identified involving insulin resistance and hematological traits in Bactrian camels, loss of keratin genes in three evolutionary lineages of mammals (including dromedary), and the oxytocin-neurophysin I (OXT) gene with its regulatory regions.
conditions will boost our understanding of the functional consequences of genetic variation.

Particular challenges in Old World camels to achieve research-to-practice goals include 1) the need for data sharing policies/practices, 2) human resources empowerment, 3) a better integrated scientific community and stronger interactions among different stakeholder categories, and finally 4) coordinated national/international investments for successful long-term management of the required infrastructures and facilities for camel research and development.

**Conclusion**

Old World camels can count now on high-quality reference genome assemblies and hence are ready—from a scientific perspective—to move to a research-to-practice level, while benefiting from advanced “omic” technologies and approaches for better functional annotations of regulatory elements and pathways. This will pave the way to omic-empowered genomic selection, genome-enabled management, and ultimately, to understanding, conserving, and utilizing Old World camels’ genomic diversity. Now, it is primarily the national governments’ turn to promote the establishment of strong public–private partnerships for large-scale, routine phenome characterization and to integrate research results into sustainable breeding management and conservation practices.

**Acknowledgments**

PAB acknowledges funding from the Austrian Science Fund project P29623-B25. The authors declare no conflict of interest.

**Literature Cited**

Alvira-Iraizoz, F., Gillard, B.T., Lin, P., Paterson, A., Pauža, A.G., Ali, M.A., Alabsi, A.H., Burger, P.A., Hamadi, N., Adem, A., et al. 2021. Multiomic analysis of the Arabian camel (*Camelus dromedarius*) kidney reveals a role for cholesterol in water conservation. Commun. Biol. 4(1):779. doi:10.1038/s42003-021-02327-3

Clark, E.L., A.L. Archibald, H.D. Daetwyler, M.A.M. Groenen, P.W. Harrison, R.D. Houston, C. Kühn, S. Lien, D.J. Macqueen, J.M. Reecy, et al. 2020. From FAANG to fork: application of highly annotated genomes to improve farmed animal production. Genome Biol. 21(1):285. doi:10.1186/s13059-020-02197-8

Elbers, J.P., M.F. Rogers, P.L. Perelman, A.A. Proskuryakova, N.A. Serdyukova, W.E. Johnson, P. Horin, J. Corander, D. Murphy, and P.A. Burger. 2019. Improving Illumina assemblies with Hi-C and long reads: an example with the North African dromedary. Mol. Ecol. Resour. 19(4):1015–1026. doi:10.1111/1755-0998.13020

Ming, L., Z. Wang, L. Yi, M. Batmunkh, T. Liu, D. Siren, J. He, N. Juramt, T. Jambl, Y. Li, and Jirimutu. 2020. Chromosome-level assembly of wild Bactrian camel genome reveals organization of immune gene loci. Mol. Ecol. Resour. 20(3). doi:10.1111/1755-0998.13141

Wu H., X. Guang, M.B. Al-Fageeh, J. Cao, S. Pan, H. Zhou, L. Zhang, M.H. Abutarboush, Y. Xing, Z. Xie, et al. 2014. Camelid genomes reveal evolution and adaptation to desert environments. Nat. Commun. 5:5188. doi:10.1038/ncomms6188. Erratum in: Nat Commun. 2015;6:6107.