



Toward a Camel HD BeadChip: the Illumina[®] Agricultural Greater Good Initiative 2019

E. Ciani, S. Brooks, F. Almathen, A. Eggen, P. Burger

From *genetics* to *genomics*:

a paradigm common to most livestock species

Phase 0

Establishment of families with recorded pedigree and phenotypes

Phase I

Good reference genome sequence assembly, anchored to chromosomes

Catalogue of genomic variation (SNPs, SVs)

Phase II

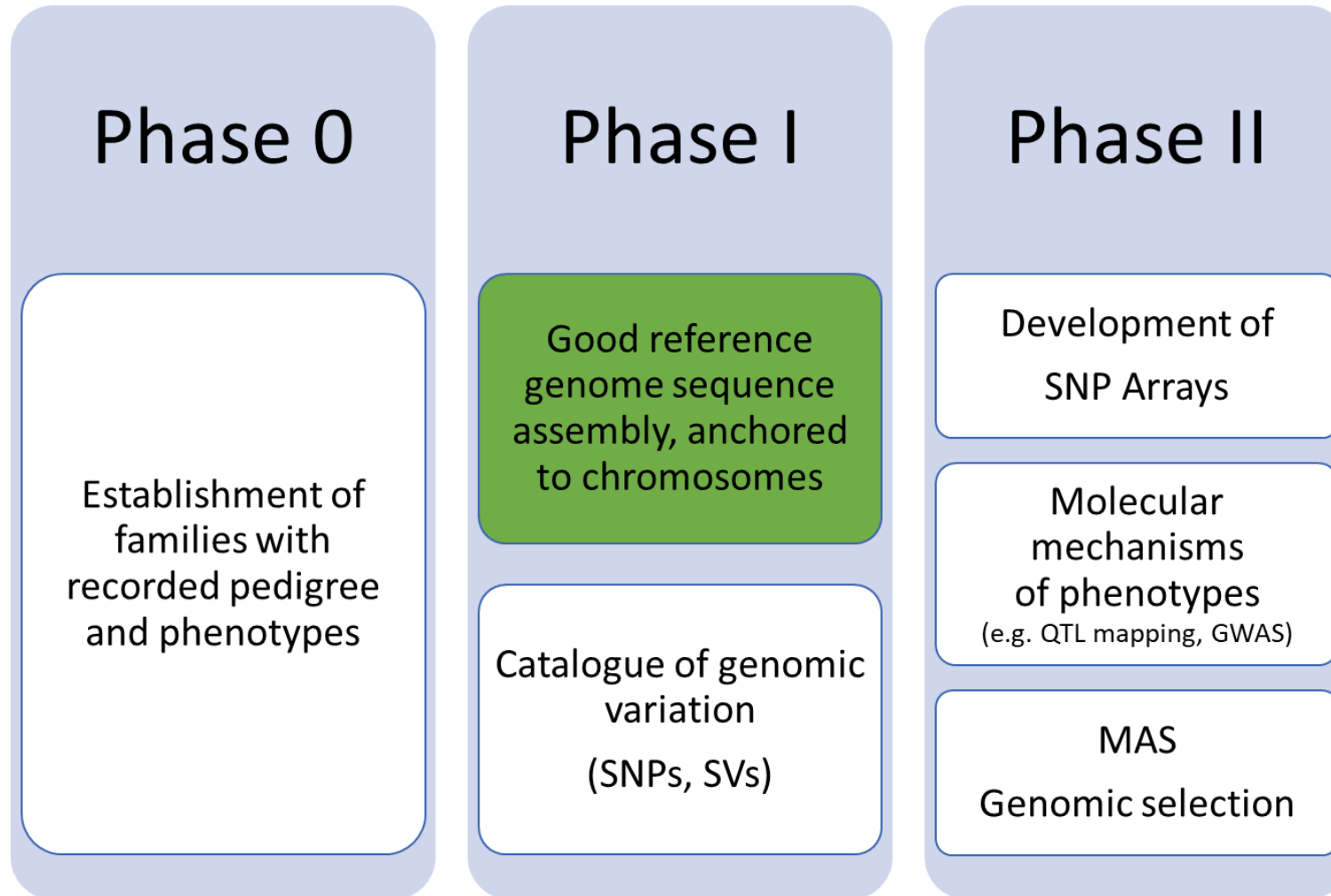
Development of SNP Arrays

Molecular mechanisms of phenotypes
(e.g. QTL mapping, GWAS)

MAS
Genomic selection

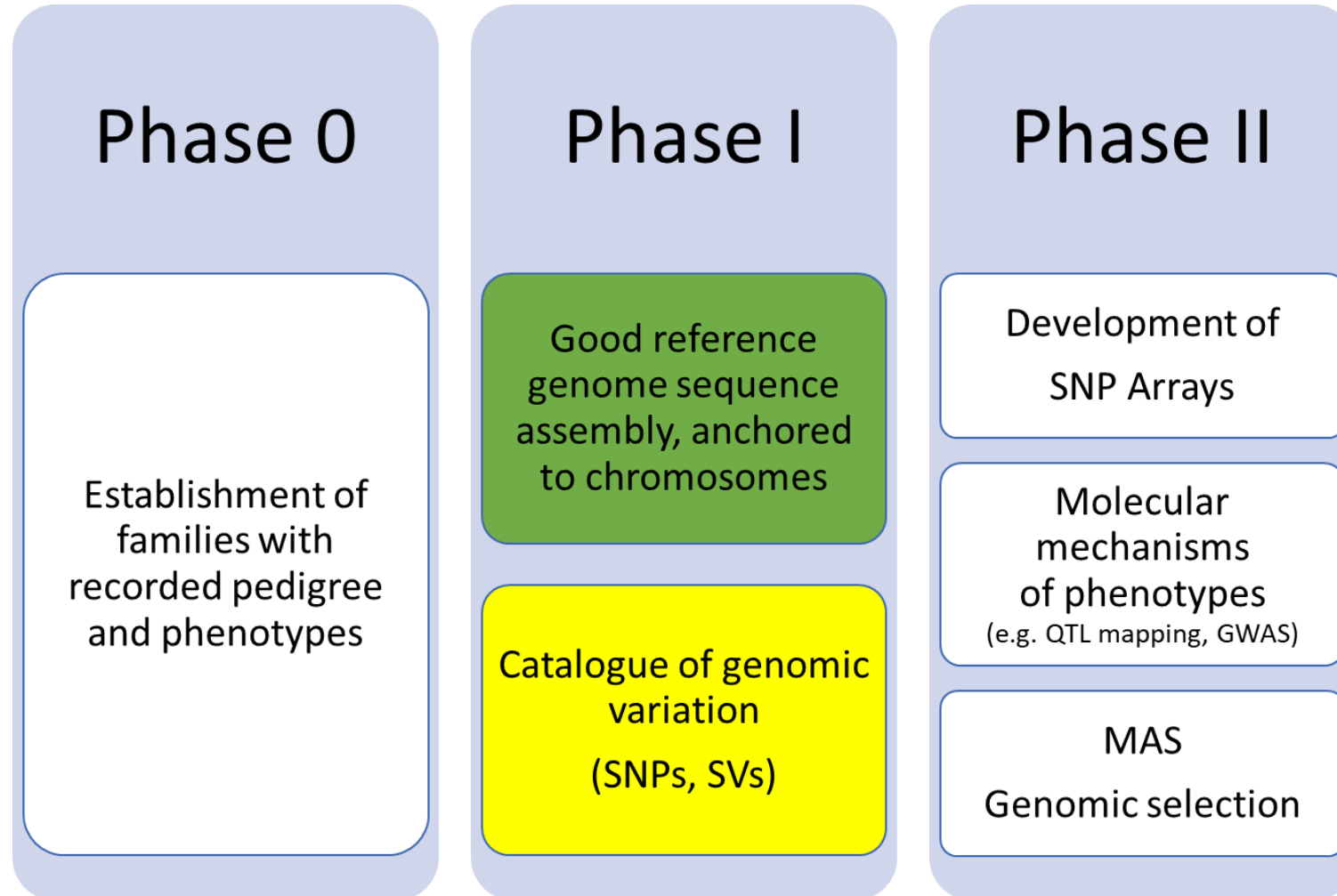
From *genetics* to *genomics*:

a paradigm common to most livestock species



From *genetics* to *genomics*:

a paradigm common to most livestock species



...resequencing dromedary genomes...



Pamela Burger

Institute of Population Genetics, University of Veterinary Medicine (Vetmeduni)
Vienna, **Austria**.



Samantha A. Brooks

Department of Animal Sciences, University of Florida
Gainesville, FL, **USA**.



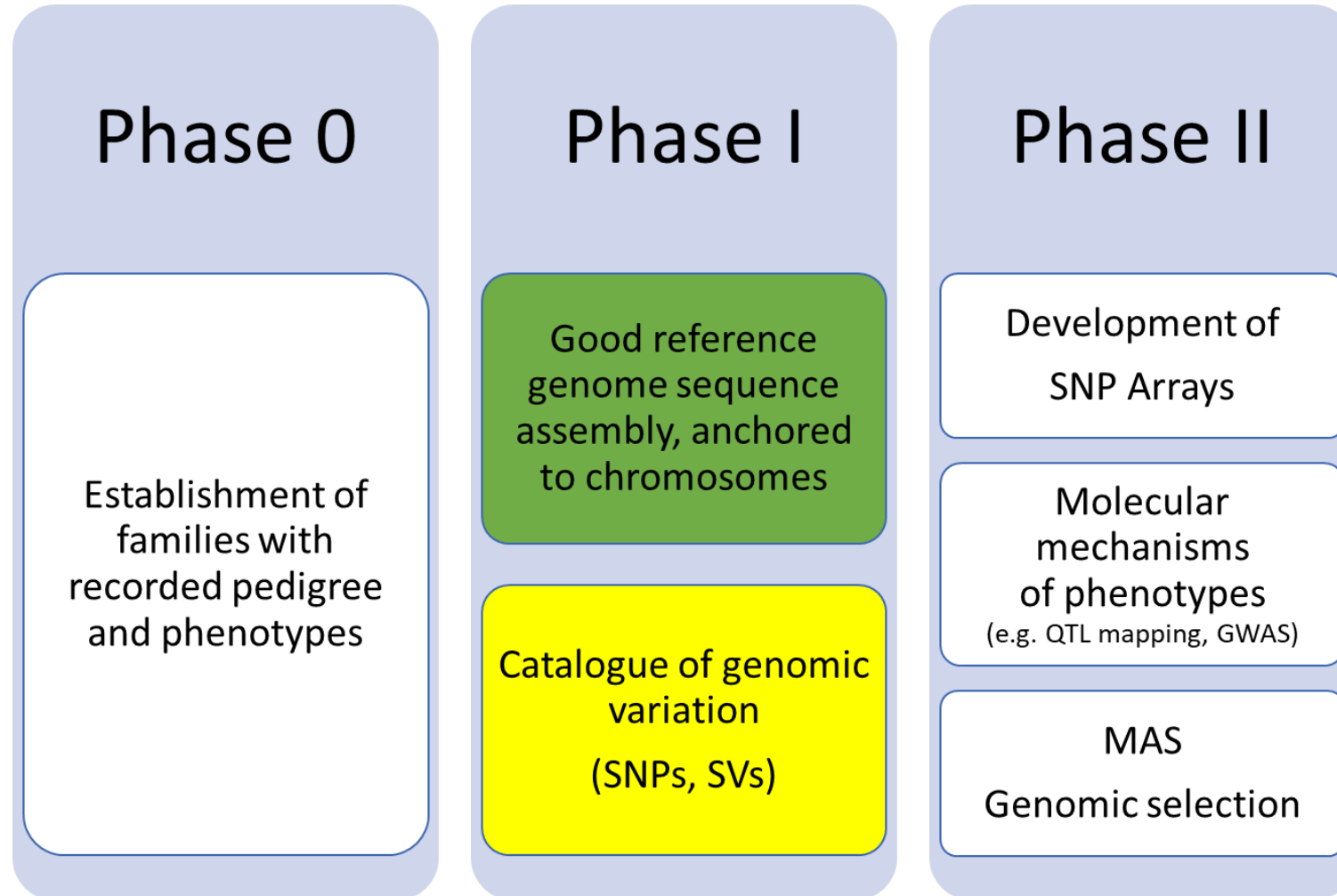
Faisal Almathen

King Faisal University
Hofuf, Al Ahsa, **Kingdom of Saudi Arabia**

Chair of the International Camel Consortium for Genetic Improvement and Conservation.

From *genetics* to *genomics*:

a paradigm common to most livestock species



...moving toward a collaborative effort...



Horse Genome Project



Infinium Microarrays for Agriculture Consortia

Illumina BeadChips are the foundation of more than 25 different consortia products, with new groups forming every year.

Why funding large collaborative dromedary research is so hard ?

- Not perceived as a livestock species by many western countries (and western funding agencies)
- Not perceived as a strategical asset by many national governments in «dromedary countries»
- Absence, or very poor, lobby action by breeder associations (when existing)
- At national level, highly fragmented production chain
- Highly fragmented scientific community
- Poor interaction between the scientific community and policy-makers

Collaborating for the Greater Good

Applications for the 2020 Agricultural Greater Good Initiative grants are currently being accepted. The deadline for all 2020 applications is December 6, 2019.

The Illumina Agricultural Greater Good Initiative grants, launched in 2011, are awarded annually. This program spurs critically needed research that will increase the sustainability, productivity, and nutritional density of agriculturally important crop and livestock species. Grant recipients receive donations of Illumina products to support their projects.

[Application details and criteria](#)

[Terms and Conditions](#)



Suheil Gaouar,
University of Tlemcen,
Algeria



Deadline for 2020 application is
December 6, 2019!

Project outline



20 tera-bases of Illumina NovaSeq sequencing data:

- ~ 400 samples, representative of different countries, at a 15X depth, by using Illumina proprietary library preparation kits
- 25 selected samples at a 35X depth, by using an Illumina partner service (10X Genomics[®] Linked-Read sequencing) → access conventional NGS-unmappable regions, provide resolution of haplotypes, better detection and characterization of structural variants (Garcia et al., 2016).

Capitalization of already existing sequence data:

VETMEDUNI, Vienna, Austria (Pamela Burger)
University of Florida, USA (Samantha Brooks)
King Faisal University (Faisal Almathen)
University of Bari (Elena Ciani)

9 WGS
14 WGS
6 WGS
122 DD-RAD genomes

NCBI BioProjects

PRJEB32117
PRJNA543338
PRJNA512907
PRJNA416700
PRJNA416695
PRJEB15365
PRJNA317727
PRJNA310822
PRJNA269961
PRJNA269274
PRJEB6946
PRJNA234474
PRJNA82161
PRJNA62507
PRJNA20873

Additional contributions are welcomed!



Project outline



Sample collection

Material Transfer agreements

DNA isolation & shipment

Centralized lab (VETMEDUNI)

Sequencing

Illumina Inc. Lab

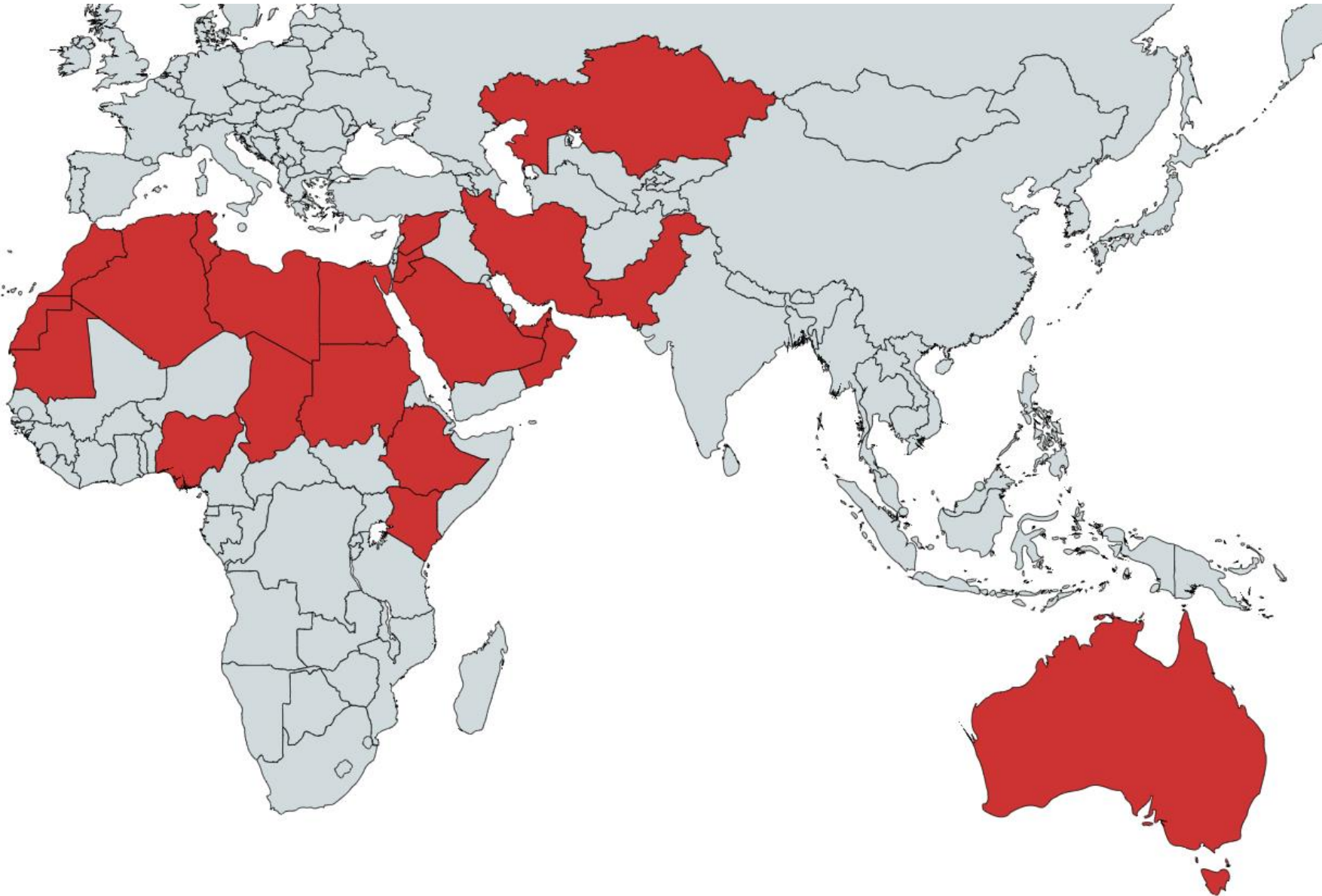
Data storage & accession

Elixir infrastructure ?

Data analysis

Multi-country bio-informatic remote working group

Sample collection - ongoing



Still few missing countries:
Somalia
Niger
Mali

Open for collaboration!!!

RESEARCH ARTICLE

Open Access

Developing a 670k genotyping array to tag ~2M SNPs across 24 horse breeds



Robert J. Schaefer¹, Mikkel Schubert², Ernest Bailey³, Danika L. Bannasch⁴, Eric Barrey⁵, Gila Kahila Bar-Gal⁶, Gottfried Brem⁷, Samantha A. Brooks⁸, Ottmar Distl⁹, Ruedi Fries¹⁰, Carrie J. Finno⁴, Vinzenz Gerber¹¹, Bianca Haase¹², Vidhya Jagannathan¹³, Ted Kalbfleisch¹⁴, Tosso Leeb¹³, Gabriella Lindgren¹⁵, Maria Susana Lopes¹⁶, Núria Mach⁵, Artur da Câmara Machado¹⁶, James N. MacLeod³, Annette McCoy¹⁷, Julia Metzger⁹, Cecilia Penedo¹⁸, Sagi Polani⁶, Stefan Rieder¹⁹, Imke Tammen¹², Jens Tetens^{20,21}, Georg Thaller²⁰, Andrea Verini-Supplizi²², Claire M. Wade¹², Barbara Wallner⁷, Ludovic Orlando^{2,23}, James R. Mickelson²⁴ and Molly E. McCue^{1*}

Still some issues on DNA quality and quantity...

Genotyping success in blood versus hair root DNA

DNA samples genotyped on the MNEc2M array were primarily derived from blood, but also came from hair roots. Both DNA sources had failed samples, however, a substantially higher fraction of hair root samples produced poor genotyping rates. Hair root DNA also tended to have lower DNA concentrations when samples were re-hydrated. To maximize information, we chose to genotype all submitted samples, even though samples from hair roots did not meet minimal DNA quantity guidelines specified by Affymetrix, though samples were dropped if they did not meet genotyping quality control thresholds. To determine if sample origin (blood versus hair roots), DNA quantity, or both, were associated with failure to pass genotyping quality control metrics, a logistic regression for sample success on DNA concentration (Pico Green) and blood/hair status was performed (Additional file [14](#): Figure S6). It is clear that higher DNA concentrations increased the probability of genotyping success. However, while blood samples were more likely to produce passing samples regardless of DNA concentration, at adequate concentrations, hair root samples were also highly likely to produce passing genotypes.

Best library preparation protocol to adopt

- ***TruSeq DNA PCR-Free:***

- Remove the need for PCR (save times & remove bias)

- Ability to sequence challenging regions (GC-rich regions, promoters, and repetitive content)

- Excellent coverage quality for deep insight into the genome (reduce gaps)

- 1 μ g DNA**



OR

- ***Nextera DNA Flex:***

- Fastest Illumina library prep workflow (~3.5 hours)

- Flexibility toward variations in sample type, DNA input amount, & application

- Optimized library prep performance, generating reliable results

- 100-500 ng DNA**



Expected impacts



- Improvement of the current **reference genome** (contiguity, completion, accuracy)
- Comprehensive **catalogue of variants** (multiple pipelines including Illumina® Dragen) → toward a dromedary “pangenome”
- Deeper understanding of **evolutionary processes** (domestication, inter-specific hybridization, dispersal and selection) that shaped the camel genomes
- **HD SNP Array** development (Illumina® BeadChip) → valuable resource for promoting genomic research and downstream applications
- **Phenotype-genotype association** studies → Limited by availability of phenotype records!!!
- Selection of reduced panels for routine applications (DNA-based **identification and parentage** tests..)
- Boost the establishment of (cross-border?) reference population(s) for **genomic selection**
- improve **productivity** (peri-urban systems), promote better **genetic management** and selection of more resilient animals (extensive desert systems)

Acknowledgements

Project partners

S. Brooks - University of Florida, USA

F. Almathen - King Faisal University, Saudi Arabia

P. Burger - Vetmeduni, Austria

A. Eggen - Illumina Agrigenomics, France

Collaborators

Bioinformatic group at DBBB, University of Bari, Italy

Fabio Marroni - ApplideGenomics & University of Udine, Italy

Population genomics group at University of Palermo, Italy

Elixir (Italian node) Prof. Graziano Pesole

Many sample providers all around the world!

The logo for Illumina, featuring the word "illumina" in a lowercase, sans-serif font. The letter "i" is orange, and the rest of the letters are grey. A registered trademark symbol (®) is located to the upper right of the "a".



Thank you for your attention !