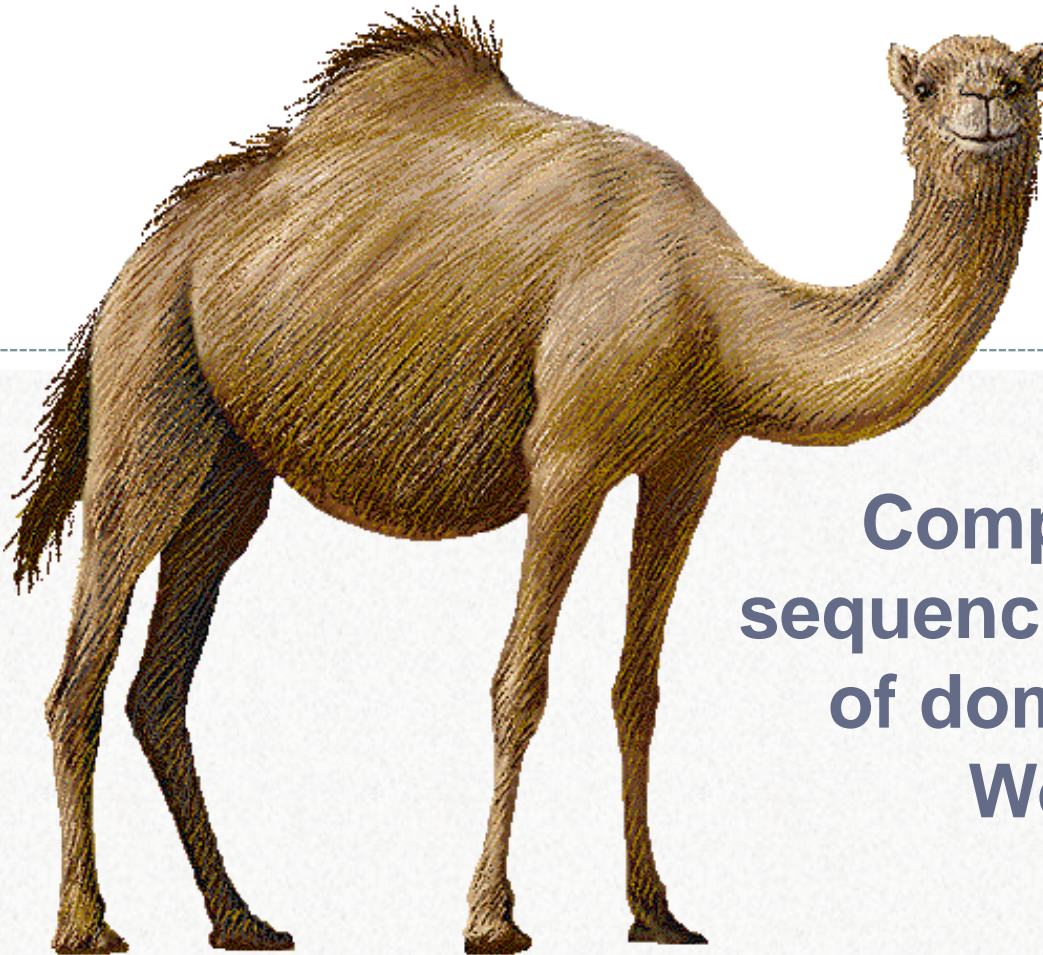


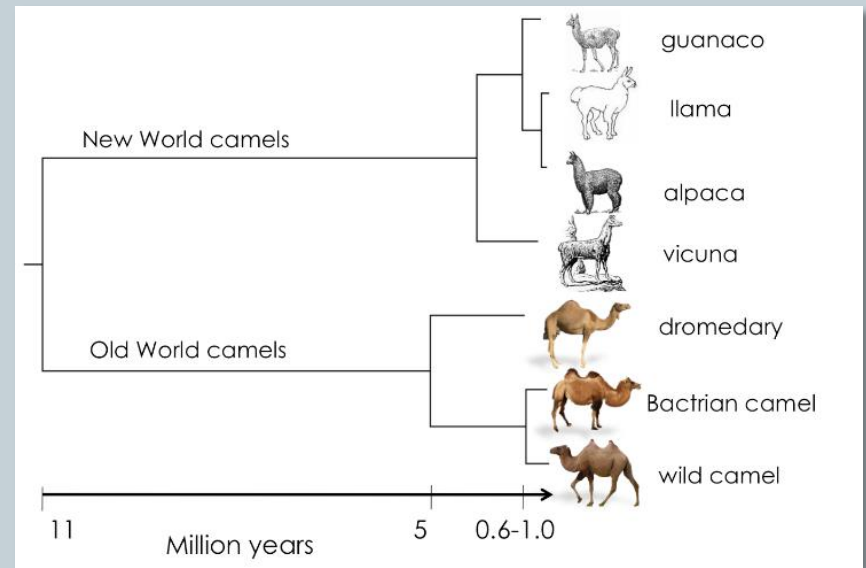
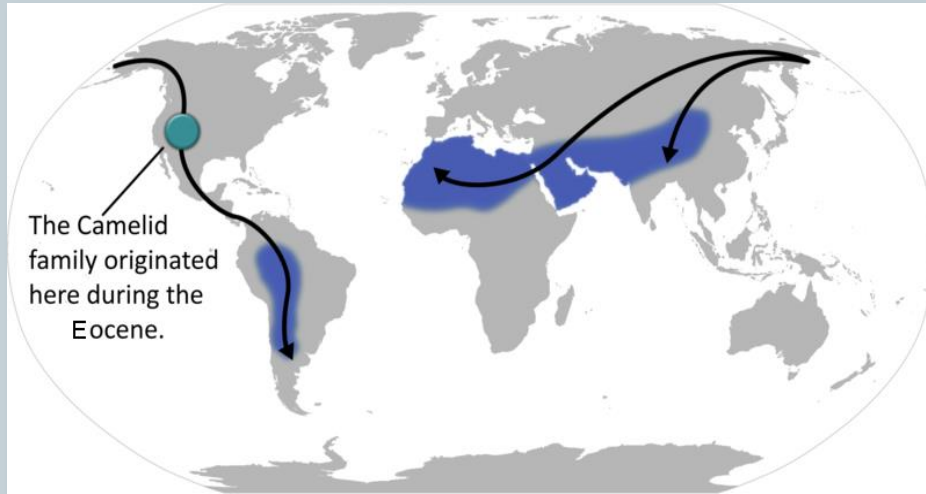
Bob Fitak

PI: Dr. Pamela Burger

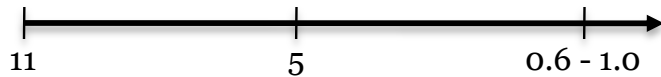
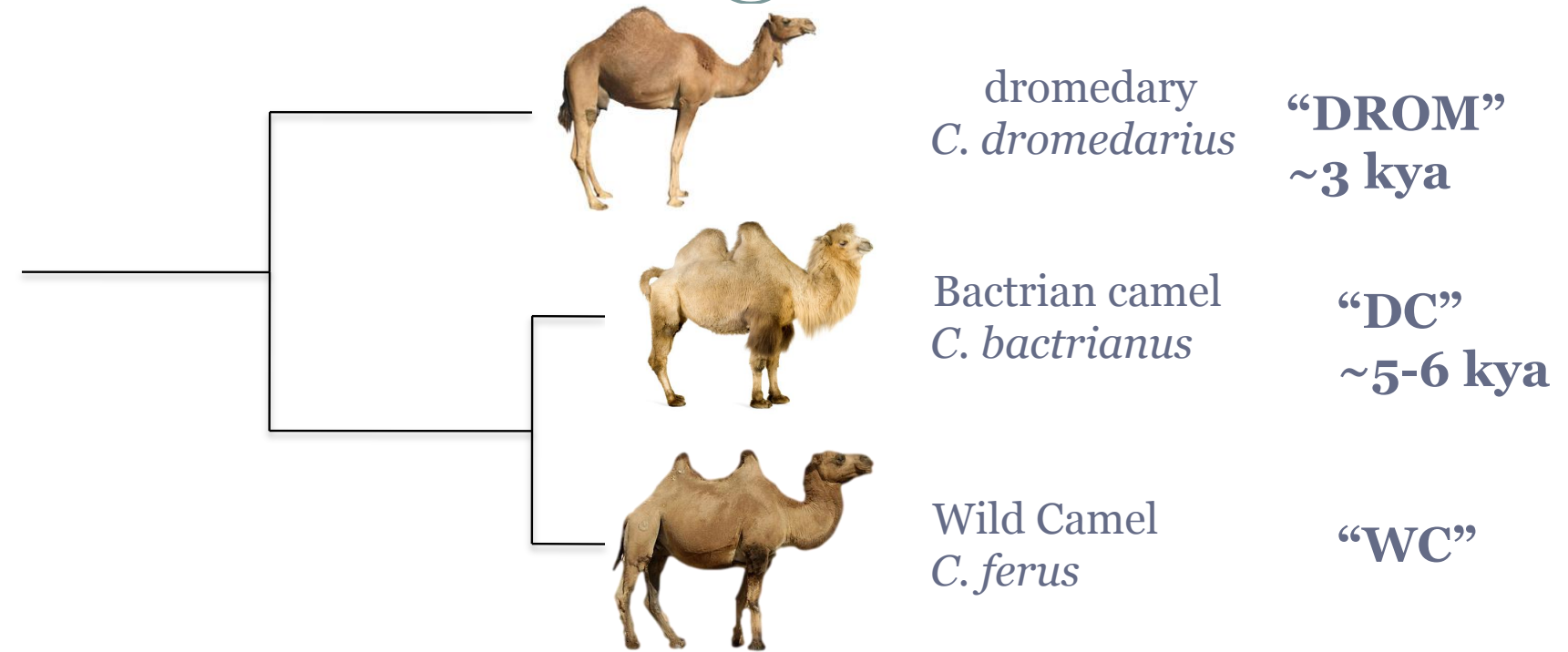


Complete genome re-sequencing reveals patterns of domestication in Old World camelids

Camels of the World



Old World Camels (*Camelini*)



Million Years

Why Camels?



Agricultural

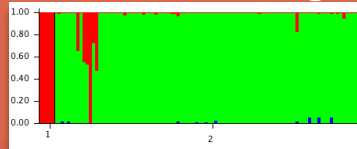


Not a camel product!

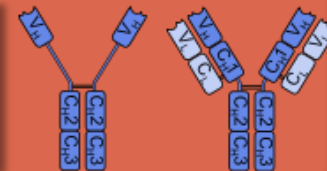


Conservation

< 1000 remaining



Evolutionary



Why Camels?

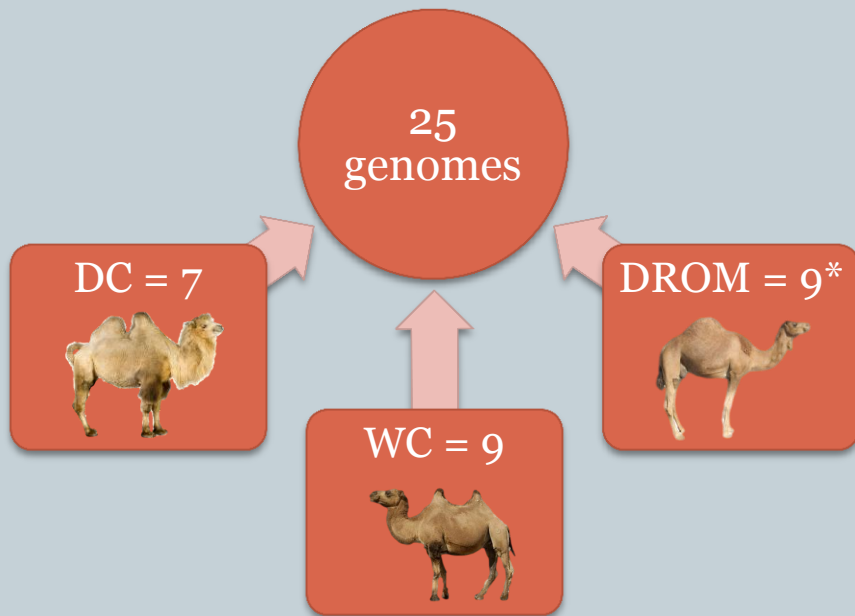


Project Aims:



- De novo assembly and annotation of the dromedary genome
- Identify SNPs
 - High-throughput chip
- Demographic history
- Natural and artificial selection (domestication)
 - Economically relevant traits

Camel Genome Resequencing

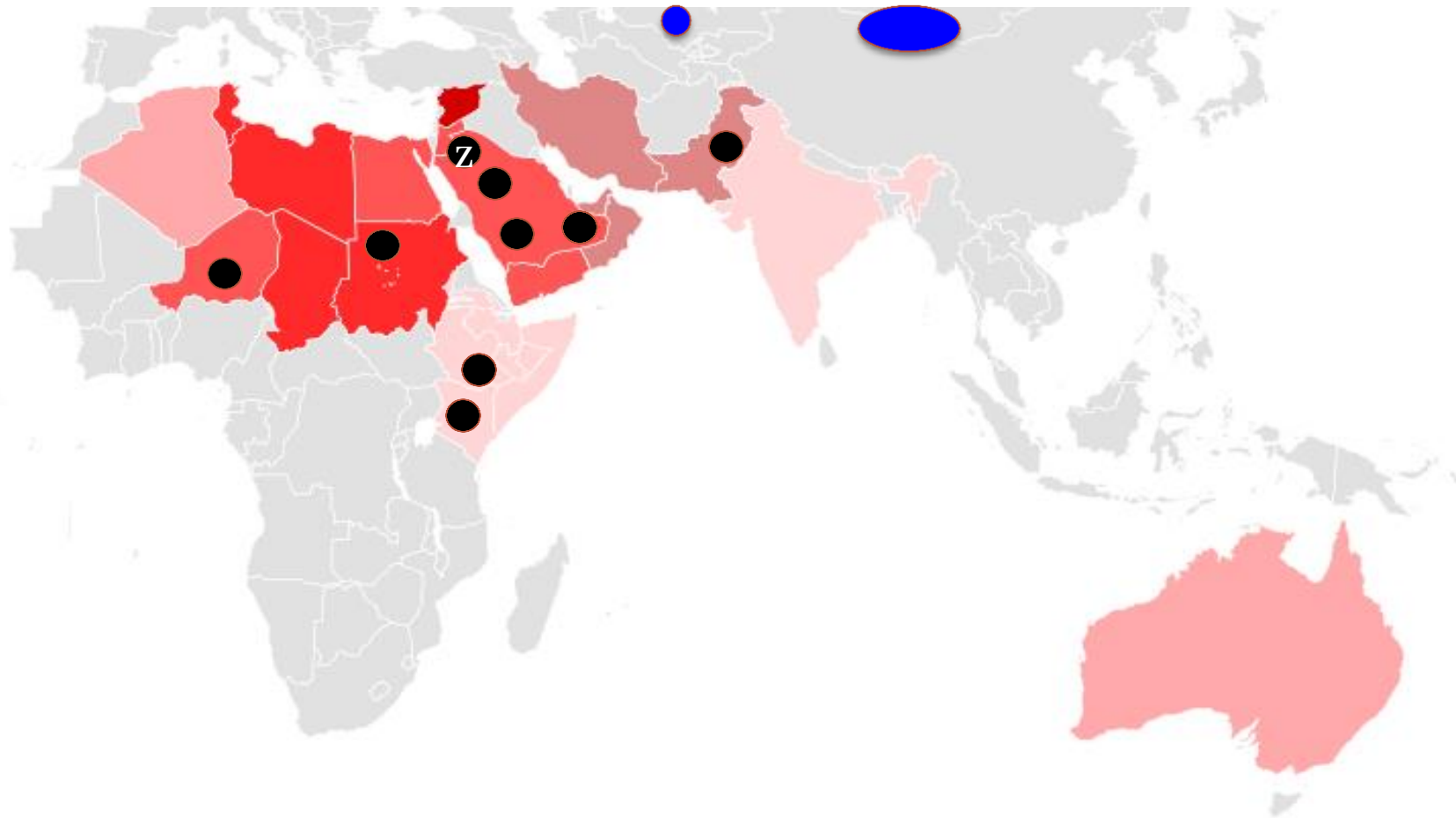


- Hi seq 2000



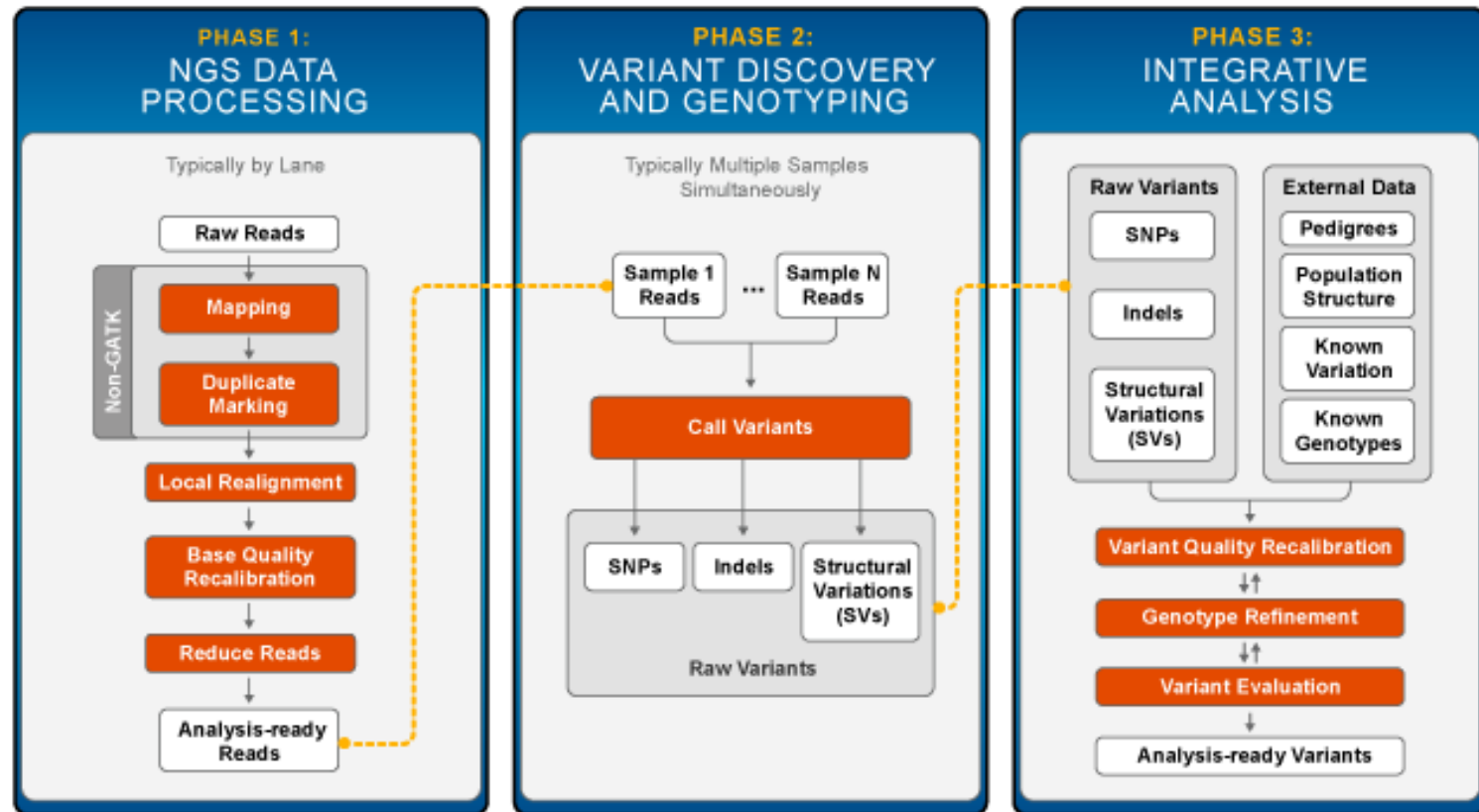
- 2 x 100 PE reads, ~500 bp insert size
 - 1 Lane per individual
 - ✦ ~15x coverage
 - DROM
 - ✦ 3 lanes 2 x 100
 - ✦ 5kb insert MP library

Samples



GATK Analysis Pipeline

Calling Variants with the GATK



Dromedary Genome





Dromedary Genome



Dromedary Genome: k=64, Error-corrected Reads

Number of scaffolds (>500 bp)	37,752
Total length	2.06 Gb
N ₅₀	1.48 Mb
Gaps	150,386
Total gap length	53.4 Mb
Error-free bases	91.8 %
CEGMA	98.4% CEGS (100% partial)

Abyss 1.3.6 : Simpson et al. 2009

Reapr 1.0.17 : Hunt et al. 2013

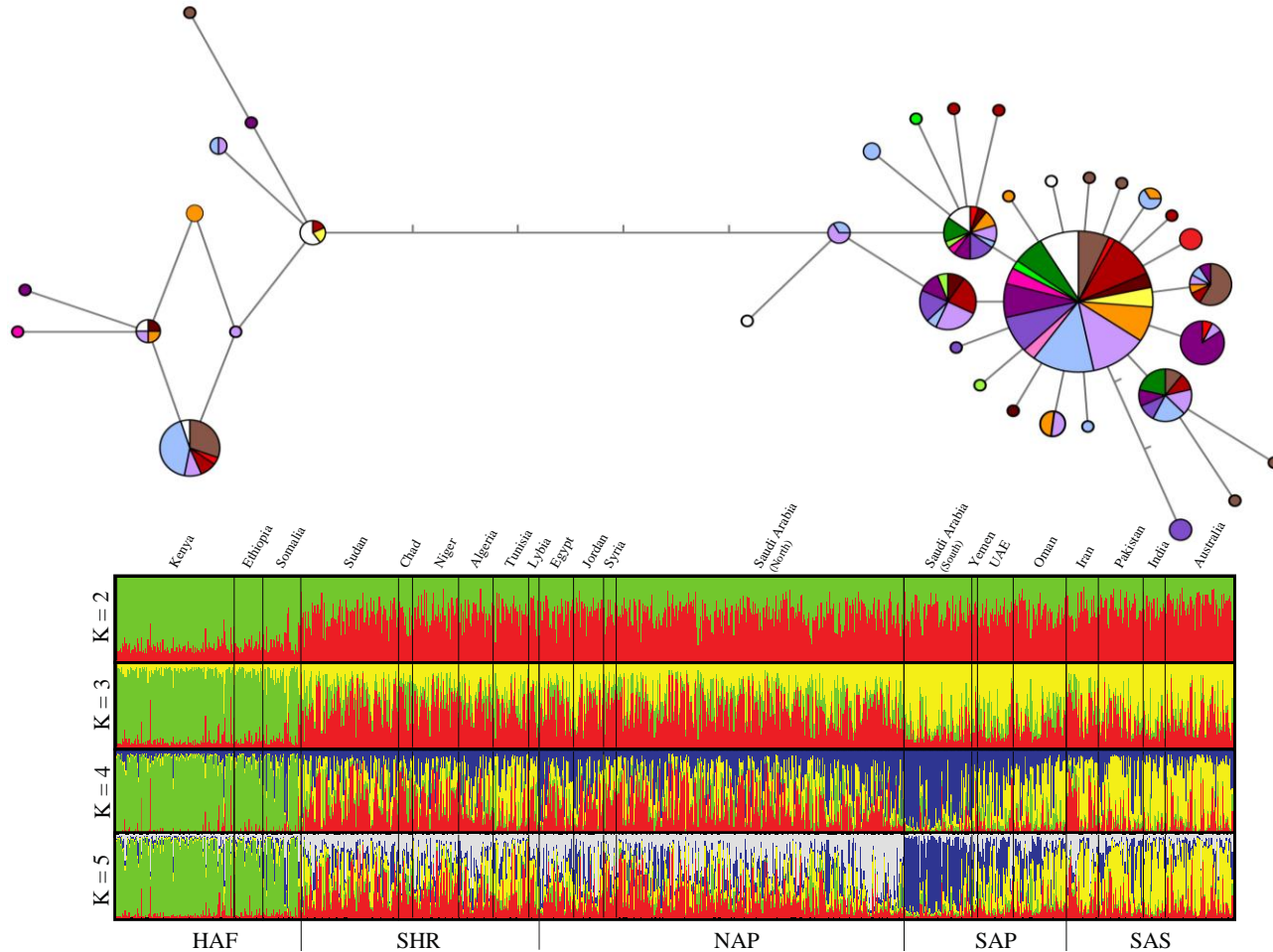
Dromedary Genetics





Dromedary Genetics

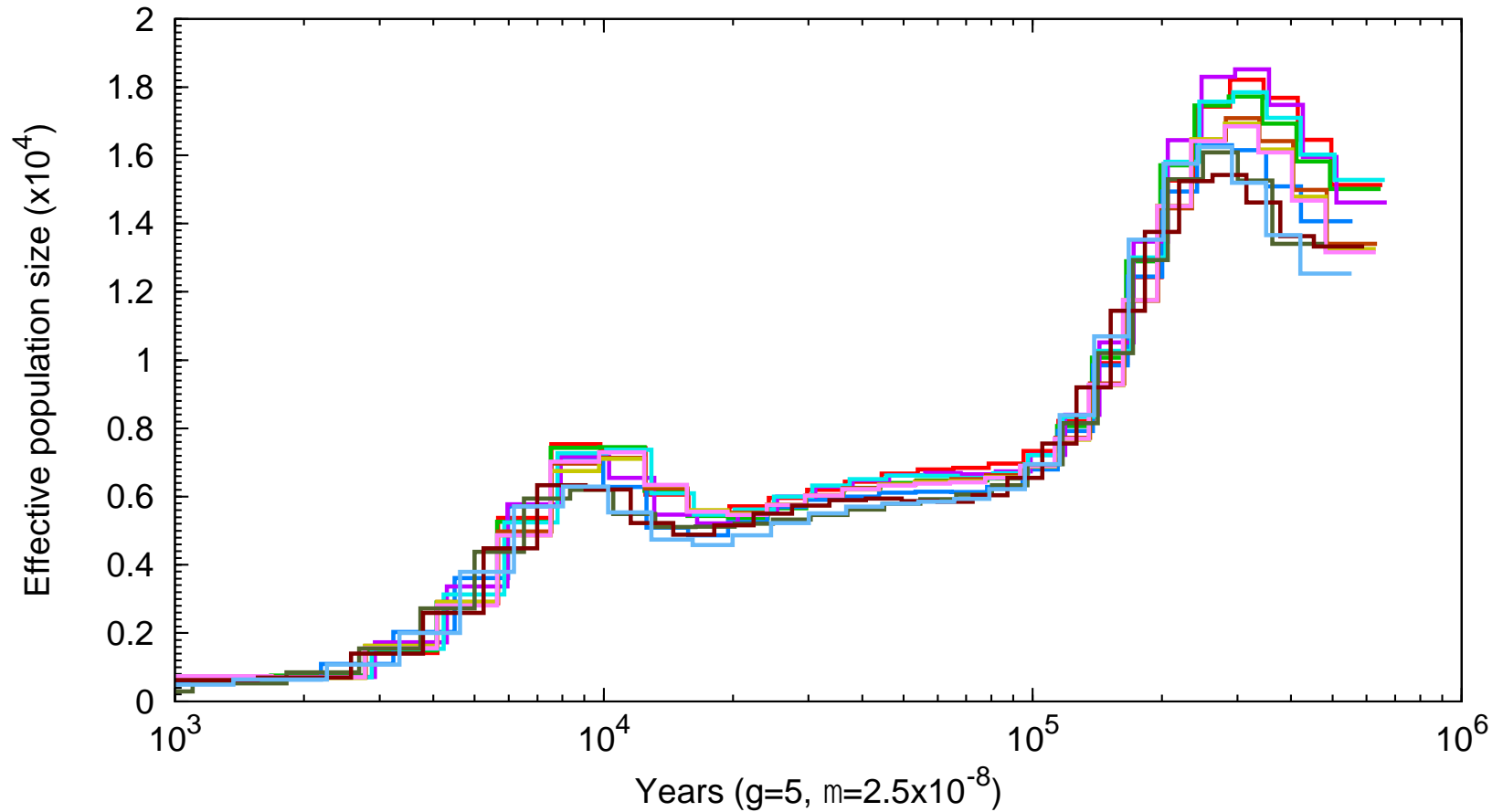
1160 dromedaries, 16 countries; mtDNA 862 bp, 17 microsats



Al-Mathen / Charruau et al. submitted



Demographic History

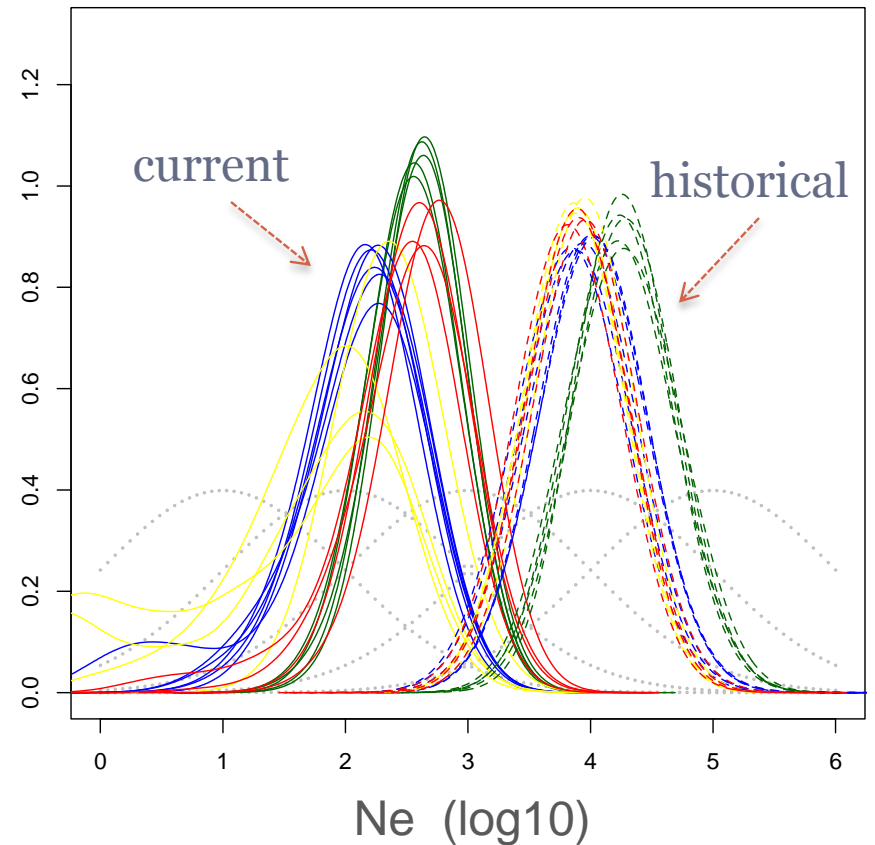
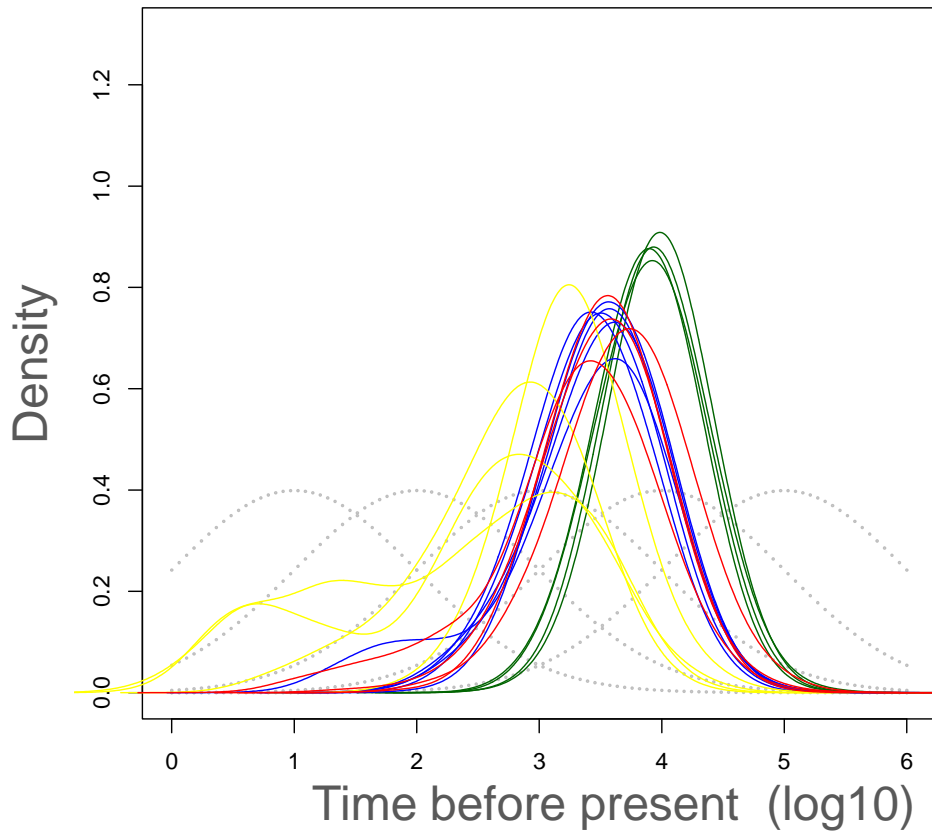




Demographic History



17 microsatellites, ~1000 dromedaries

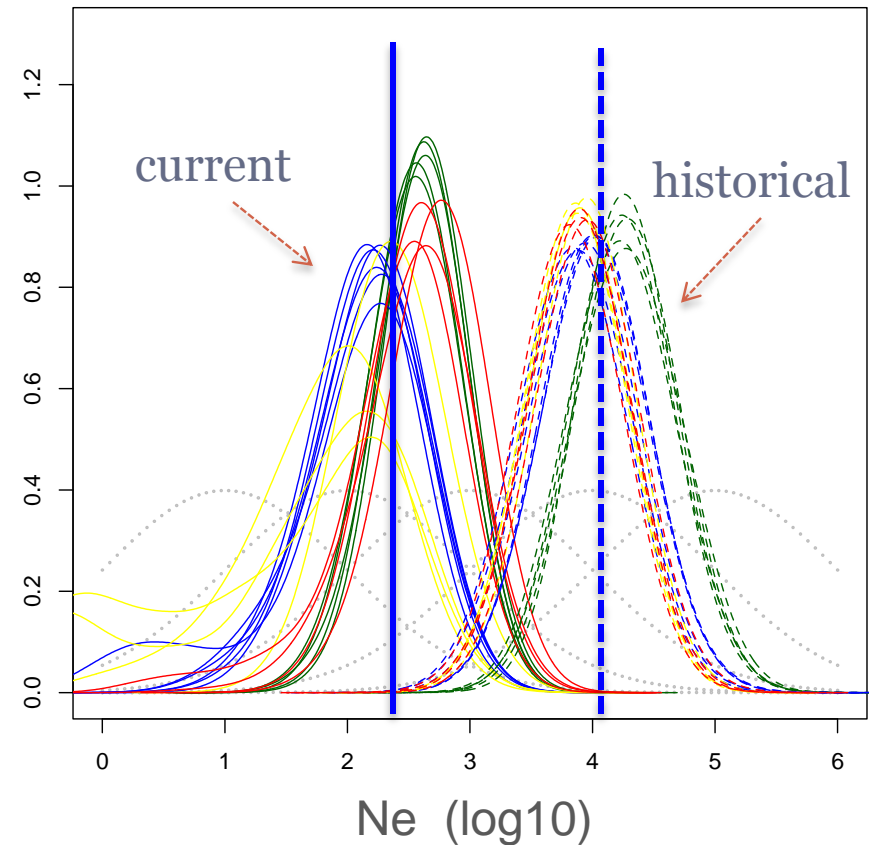
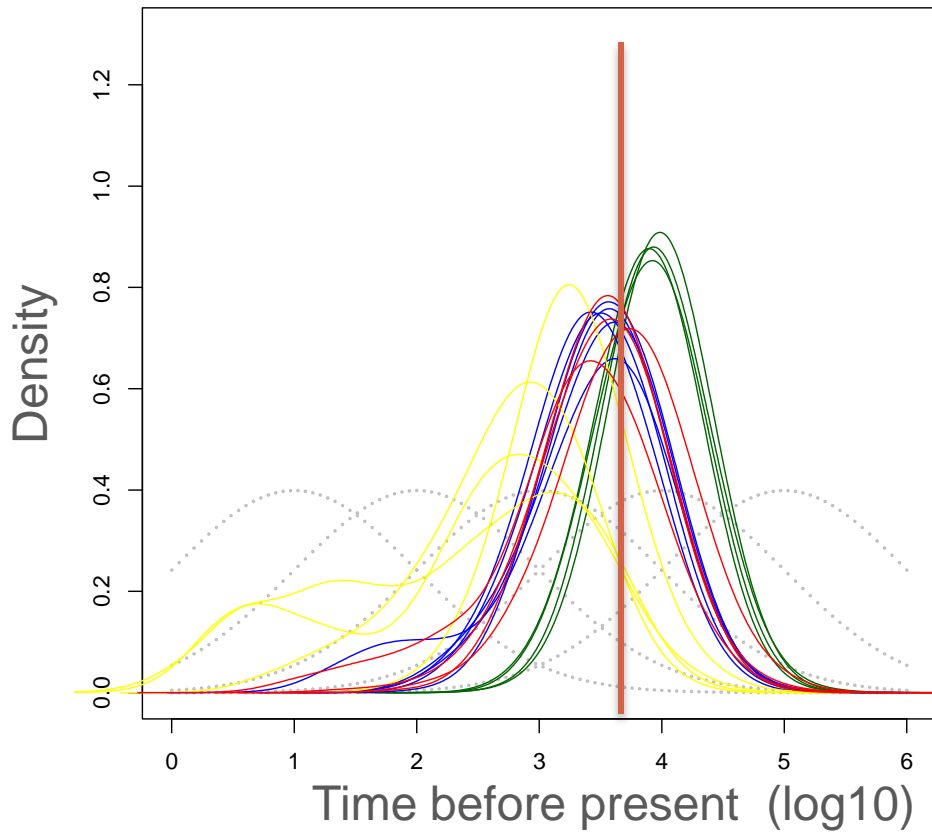




Demographic History

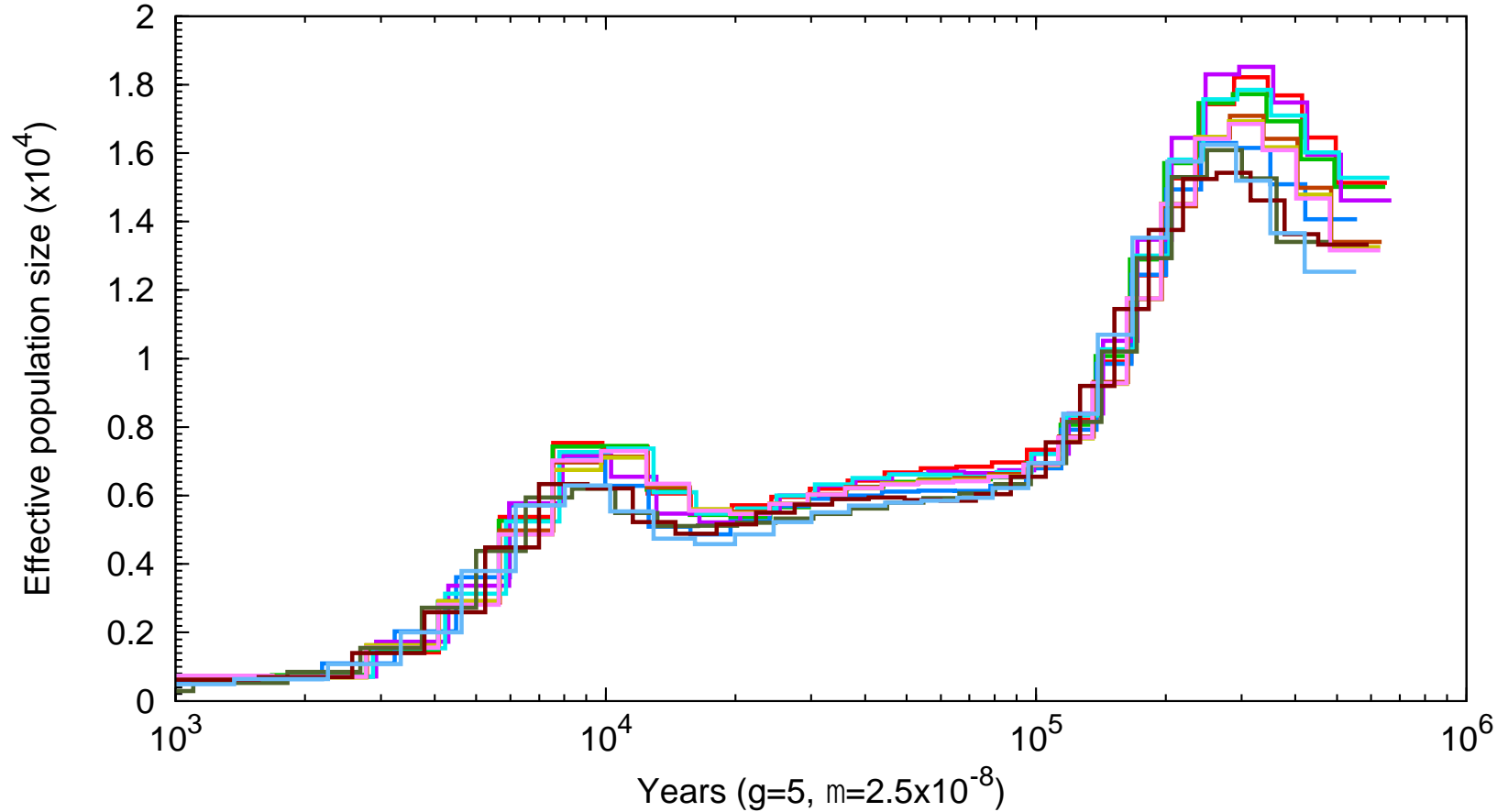


17 microsatellites, ~1000 dromedaries



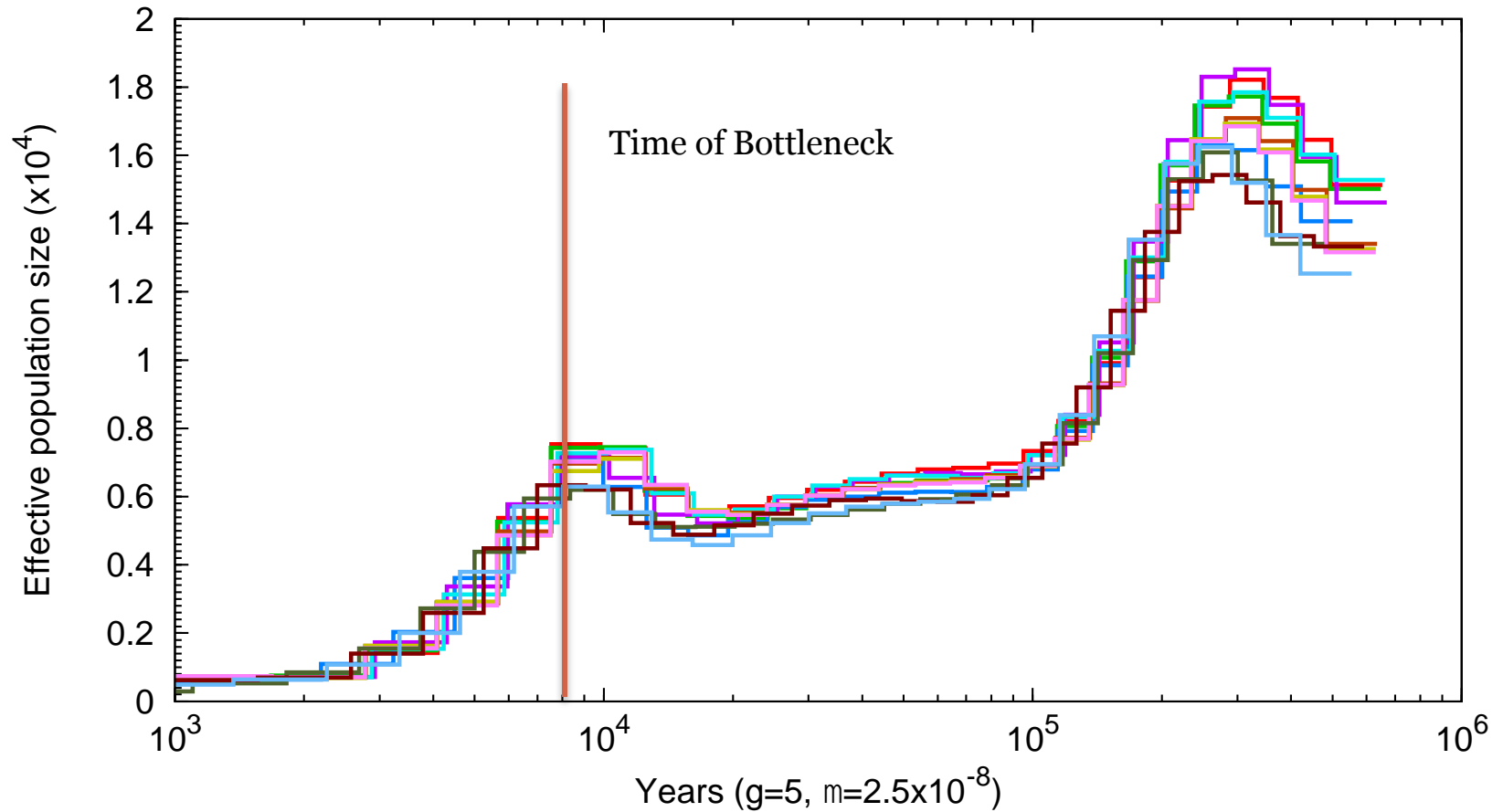


Demographic History



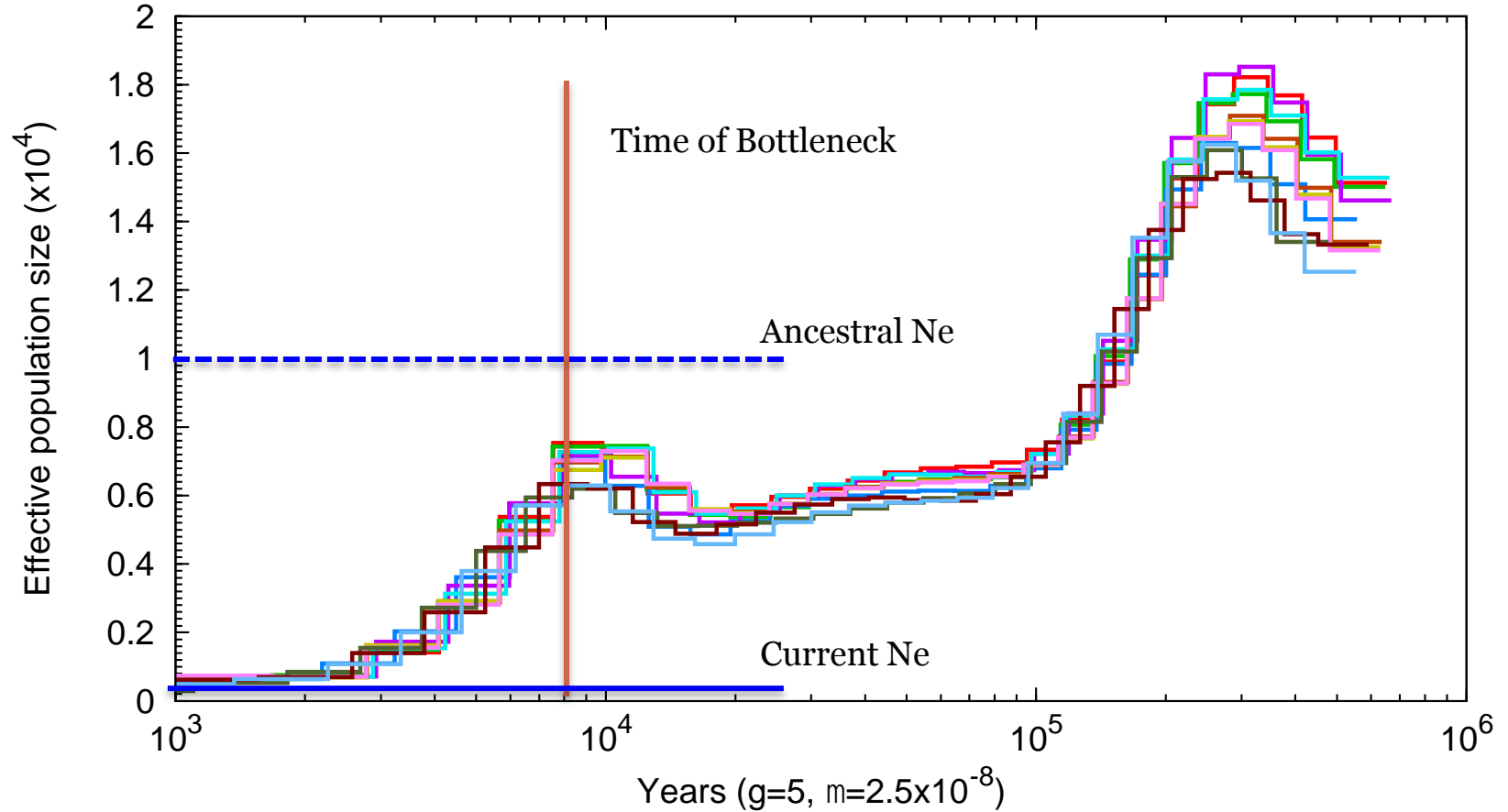


Demographic History





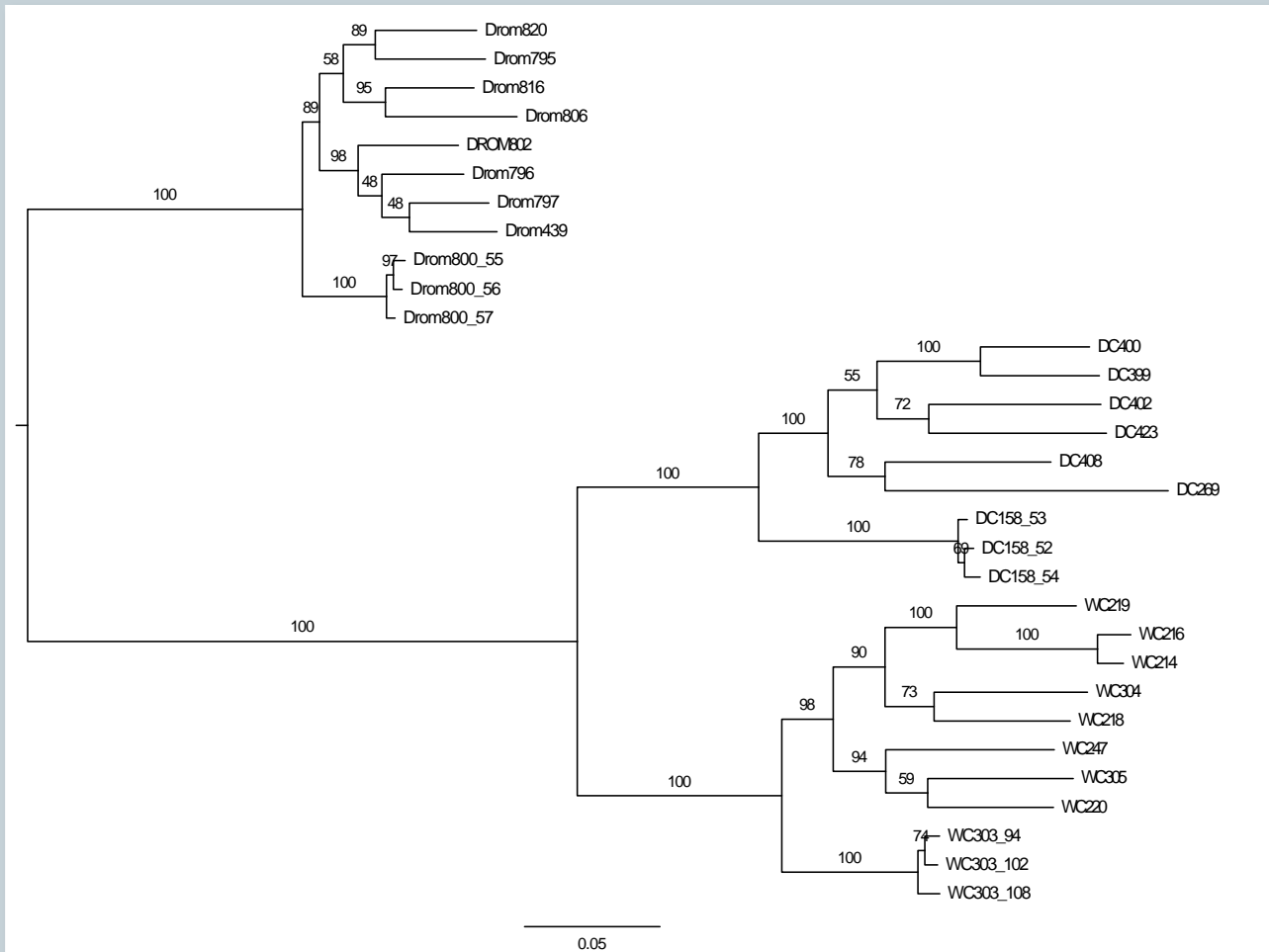
Demographic History



Wild and Domestic Bactrian Camels

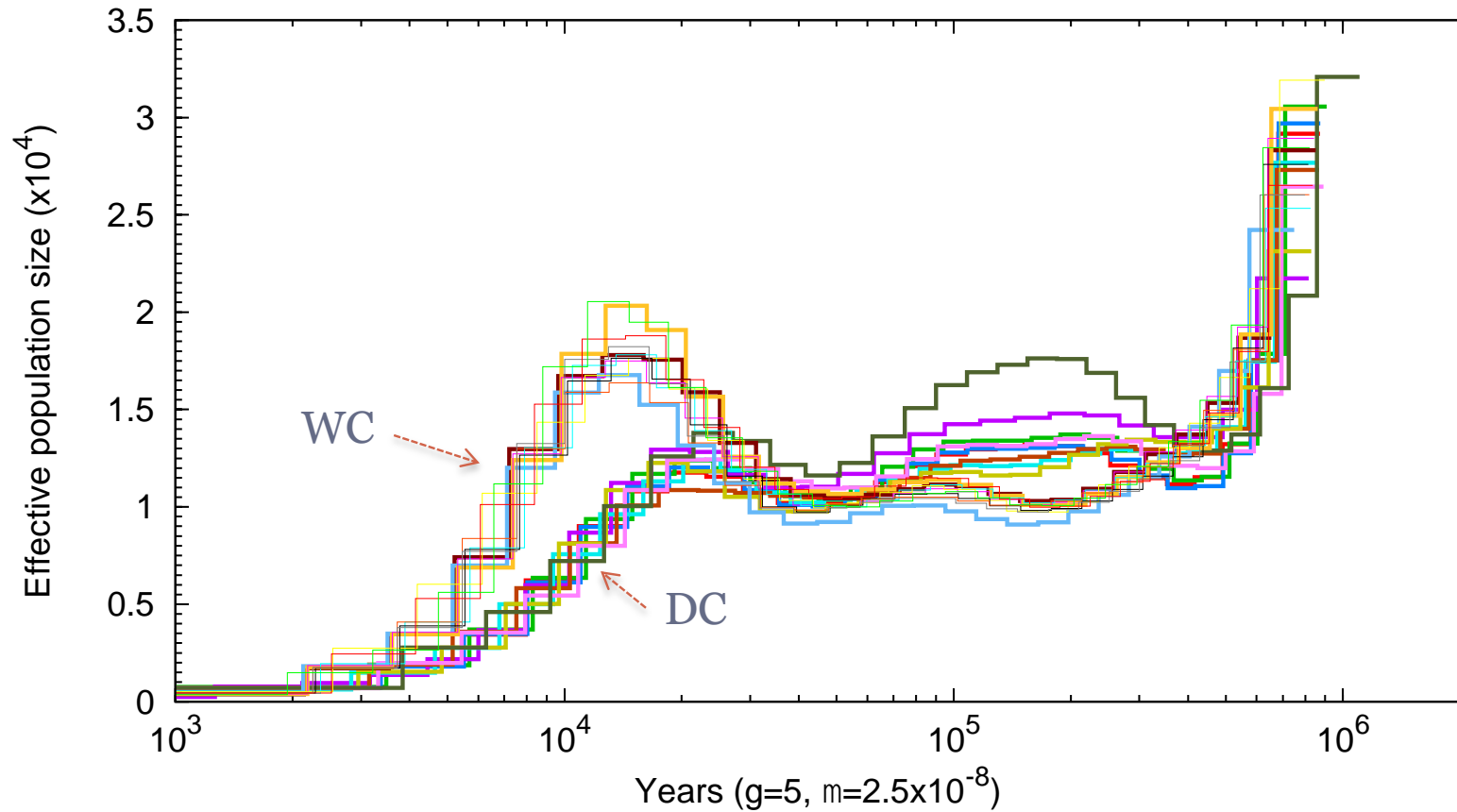


3 Species ML Tree





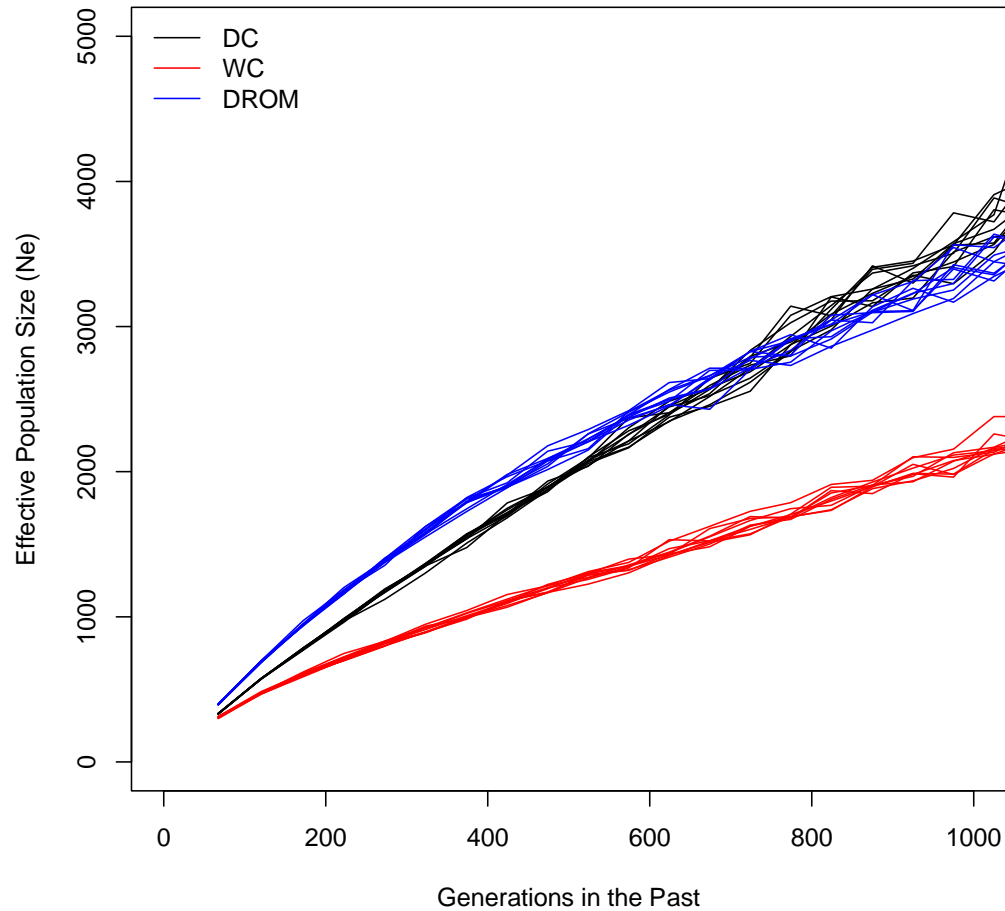
Demographic History



SNeP



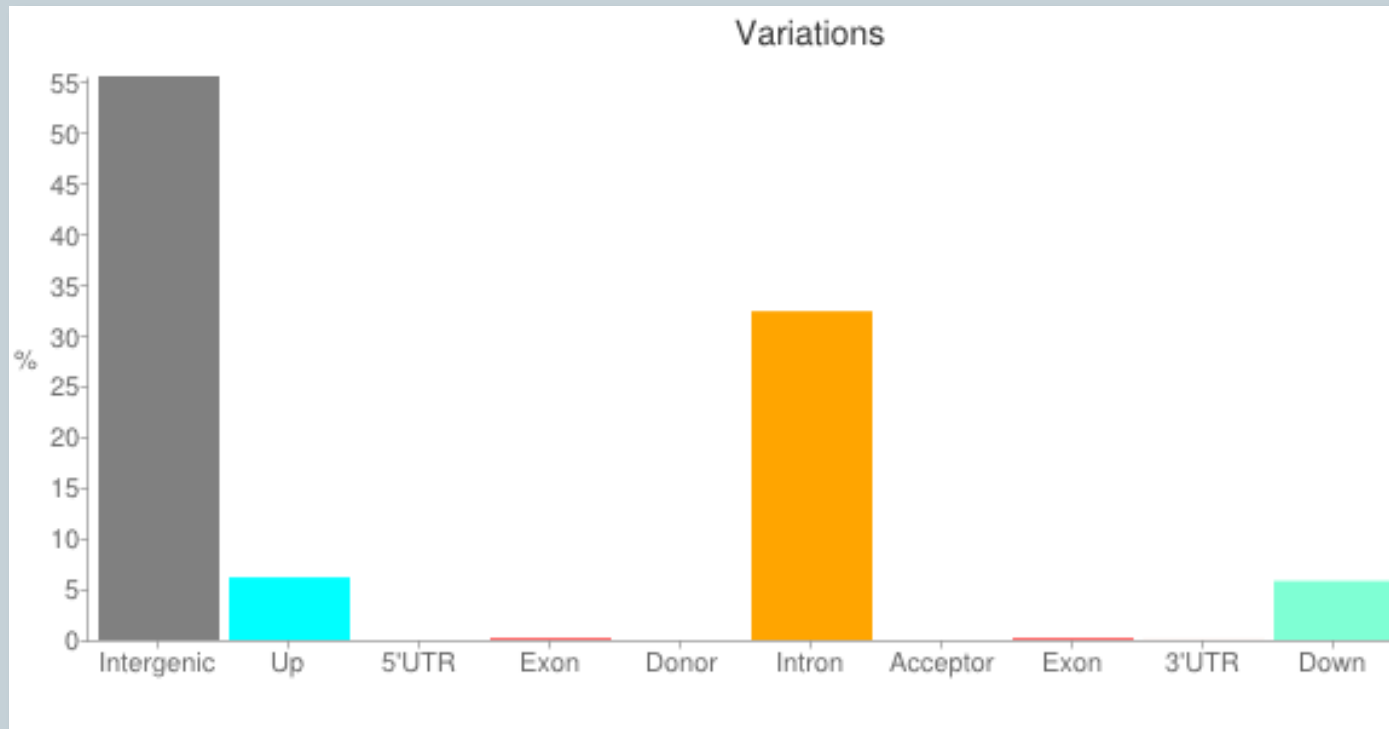
10 permutations of 20K SNPs
(40 Longest Scaffolds)



Two-humped Camels: SNPs



- ~3.1 million SNPs
- Mean $F_{st} = 0.176$



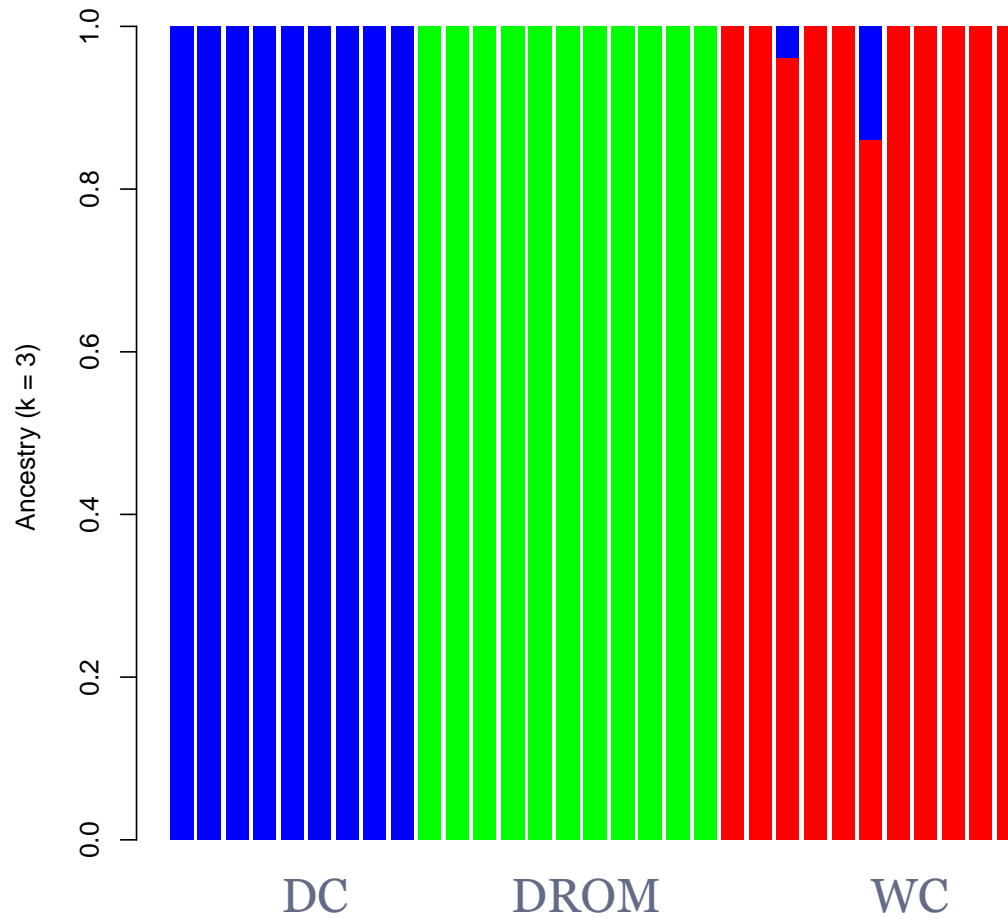
Two-humped Camels: SNPs



- SNPs with Z-score > 3
 - ~34,000 SNPs (0.13% coding)
 - dn/ds = 1.7

GENE ID	Description
UMODL1	olfaction
TBC1D8B	Vesicle biogenesis, calcium ion binding
ANO9	Trans-membrane calcium ion transport
TSHZ3	Chromatin/DNA binding zinc finger
LOC102523665	ATP-dependent RNA helicase

Camel Admixture



In progress:



- Dromedary
 - Complete genome annotation, ~~SNP Discovery~~
- Wild and Domestic Bactrian Camels
 - Tests for selection (HKA/MK-based tests; Liu et al. 2014)
- 3 species, high-density genotyping chip
 - Economically relevant traits



Acknowledgements

Pamela Burger

Elmira Mohandesan

Marlies Dolezal

PopGen @Vetmeduni

Jukka Corandaer

Pablo Orozco-terWengel

Michael Bruford Lab

Mongolian Academy of
Sciences



FWF

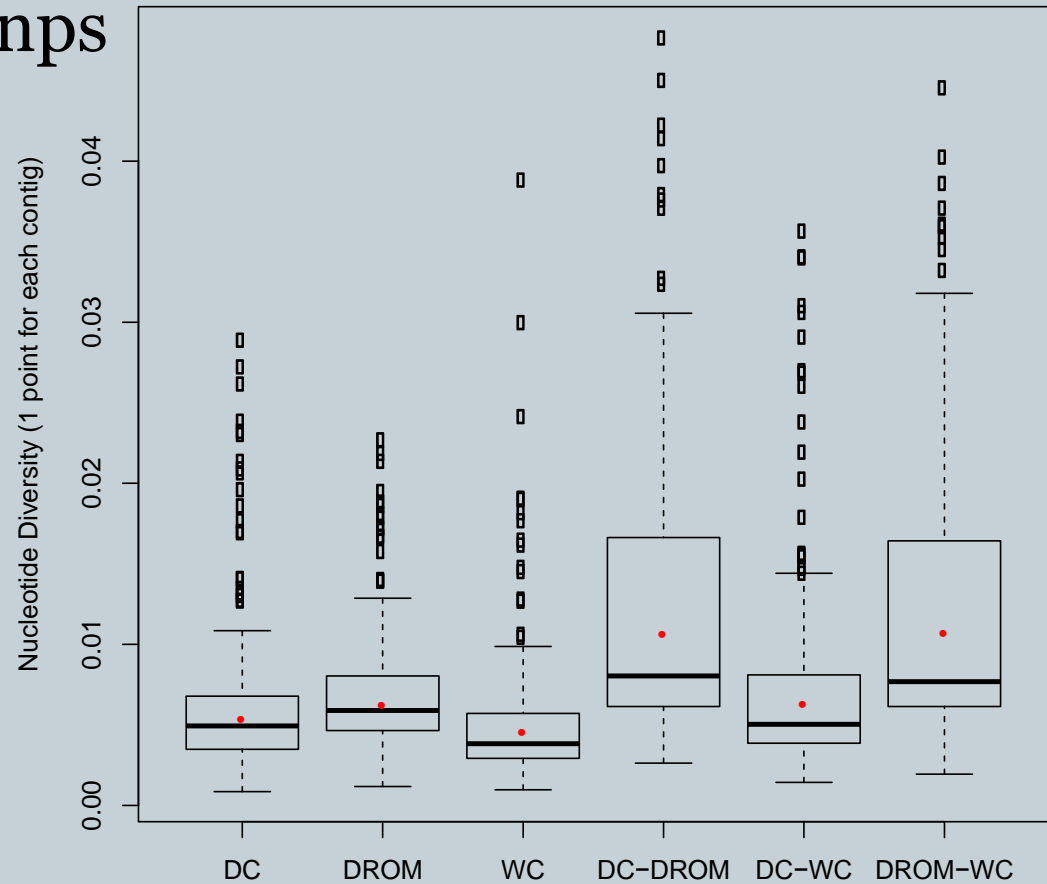


EUROPEAN
SCIENCE
FOUNDATION
GENOMIC-RESOURCES

Two-humped Camels: SNPs

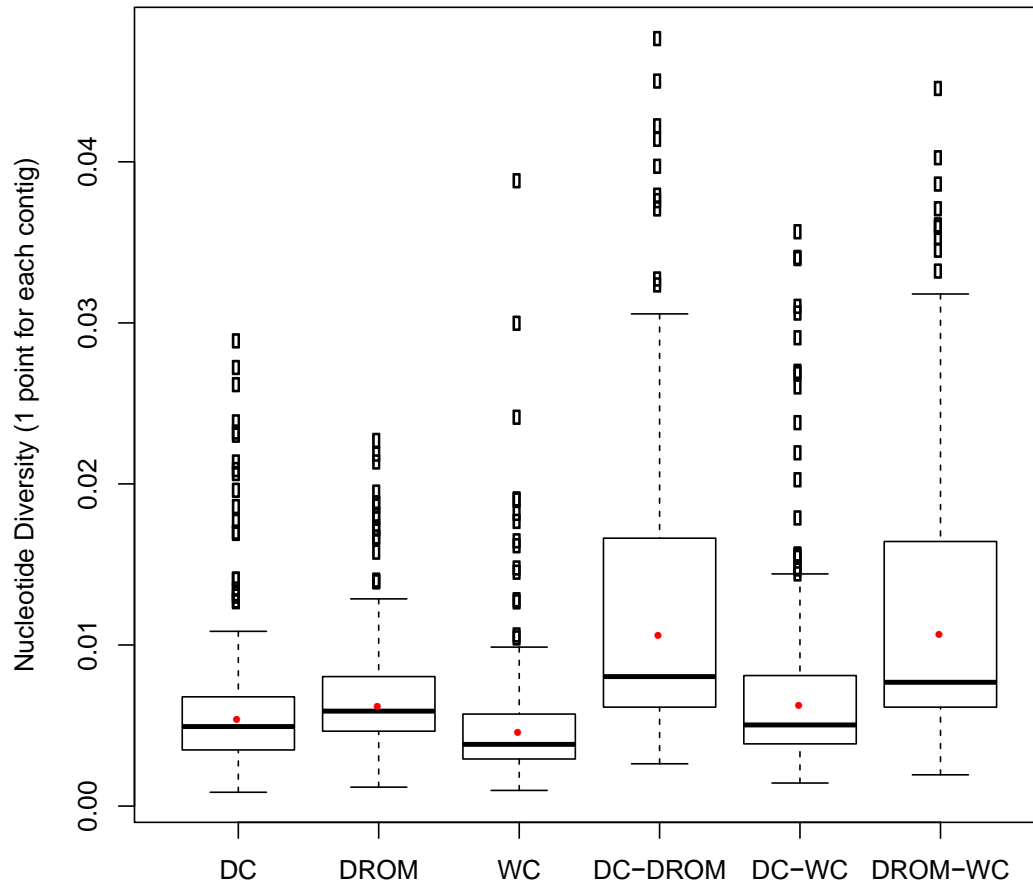


- 3098017 DC-WC snps

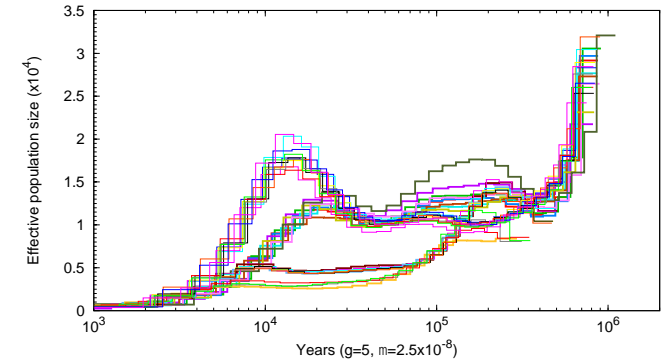


Colored points indicate the mean corrected f or contig length

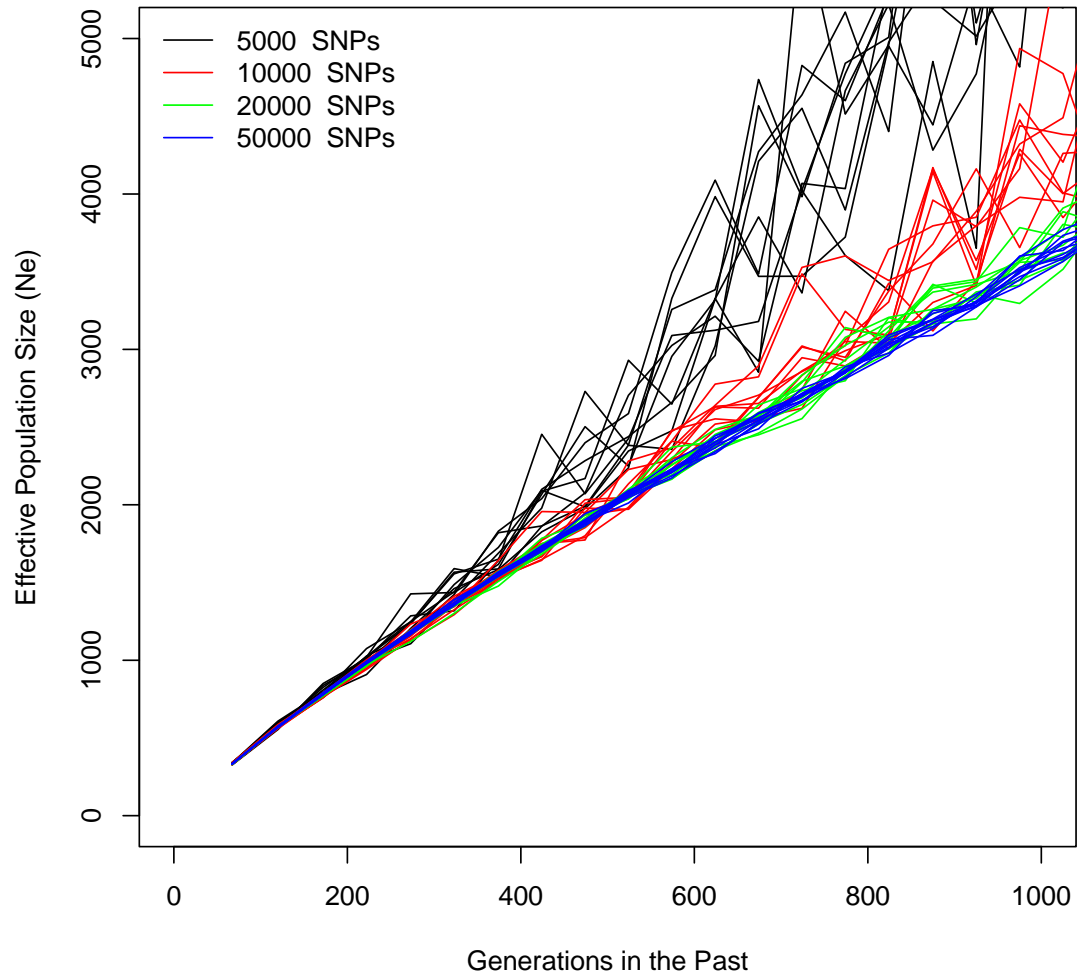
Nucleotide Diversity (π)



Colored points indicate the mean corrected for contig length



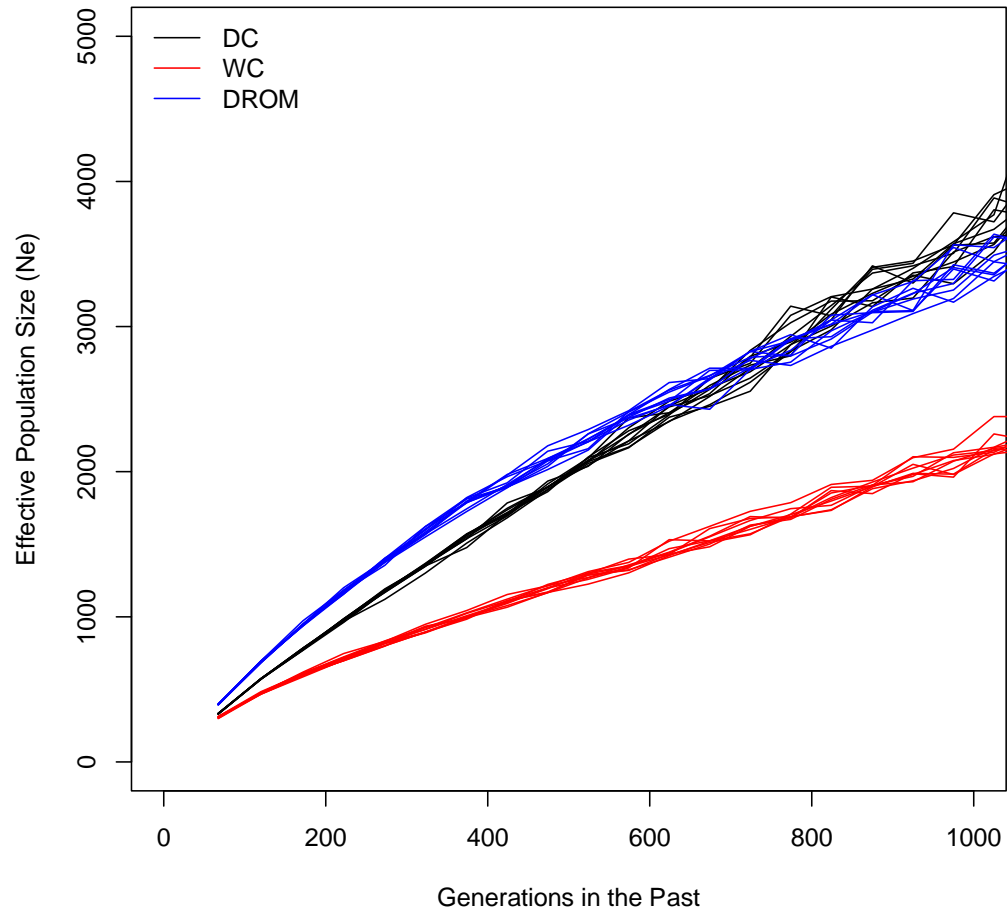
SNeP



SNeP



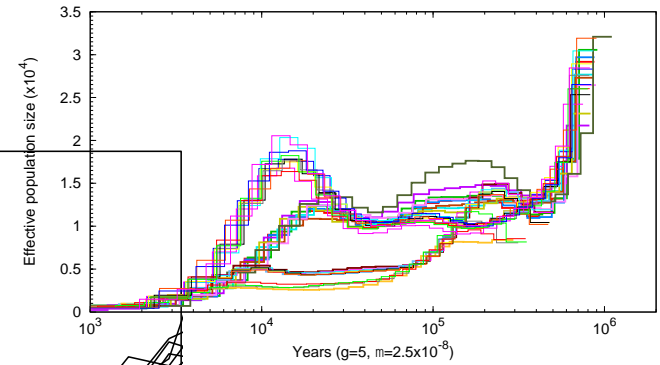
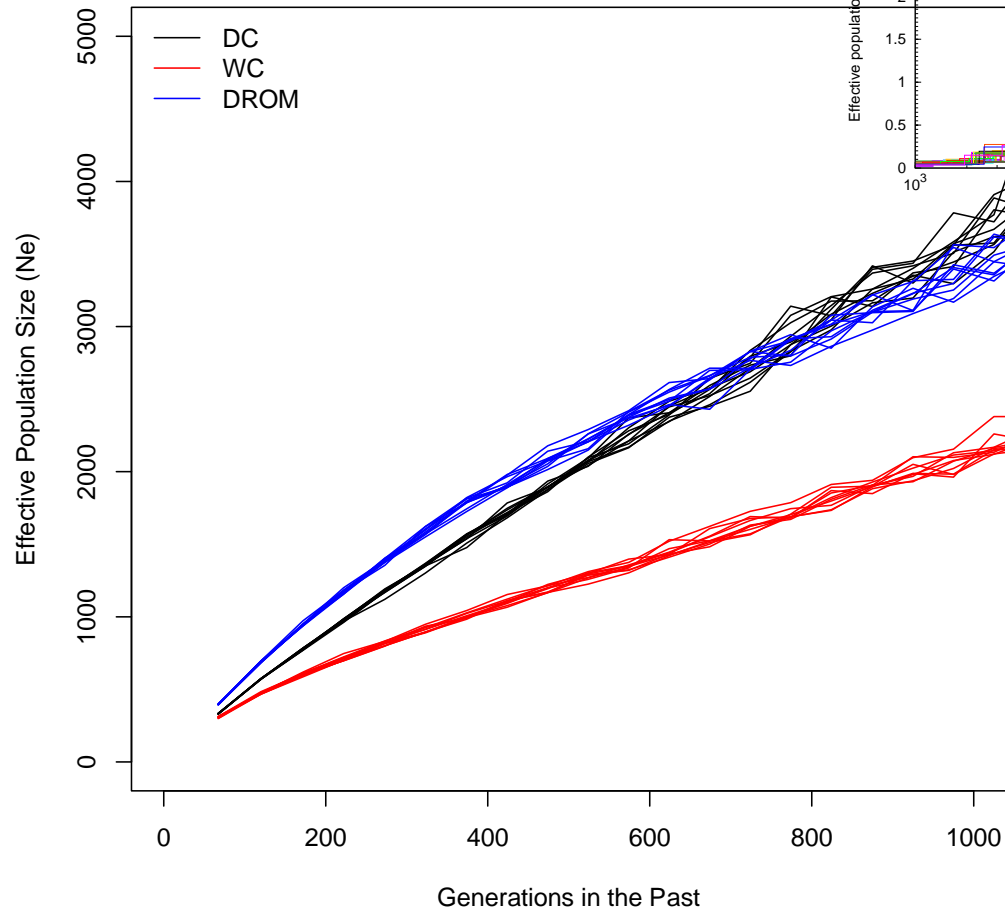
10 permutations of 20K SNPs
(40 Longest Scaffolds)



SNeP



10 permutations of 20K SNPs
(40 Longest Scaffolds)

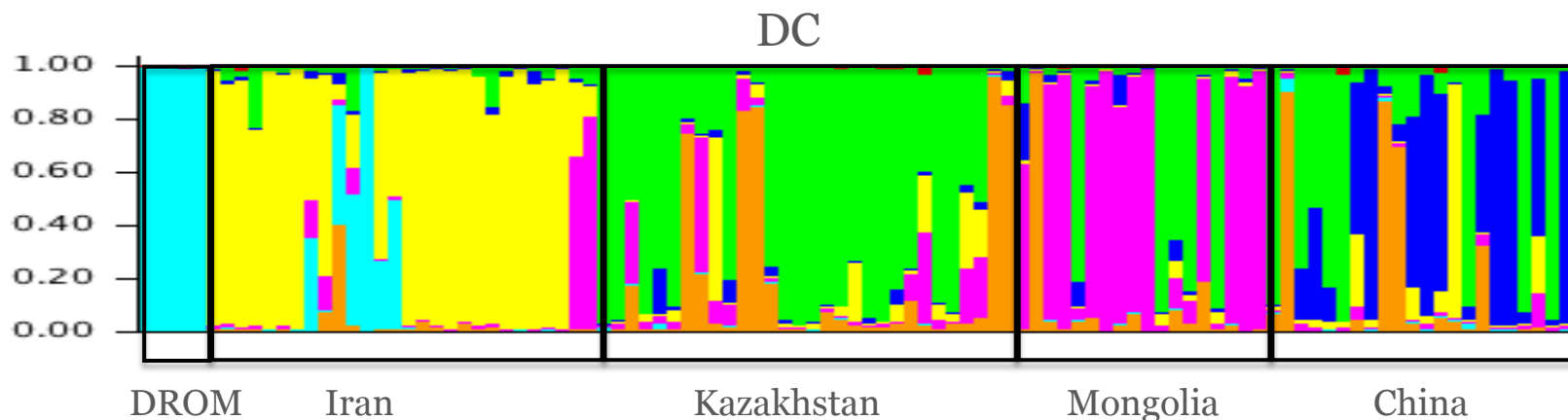
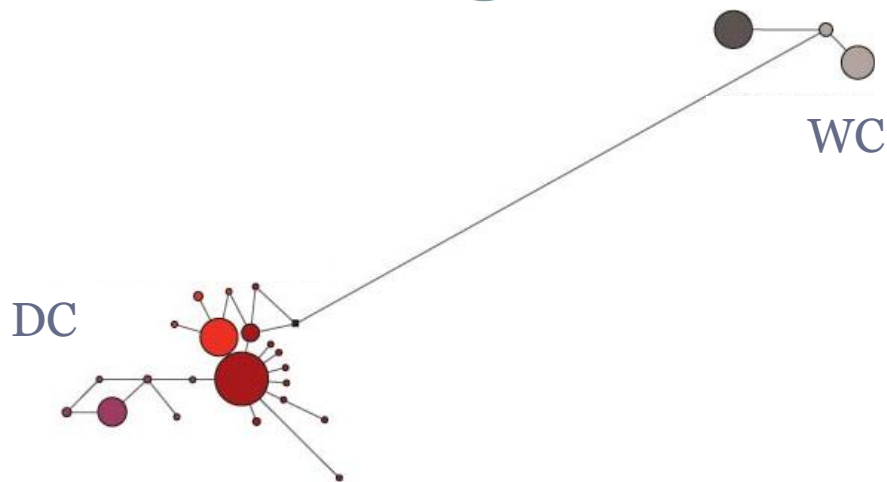


Dromedary Genetics





Wild/Domestic Bactrian Genetics

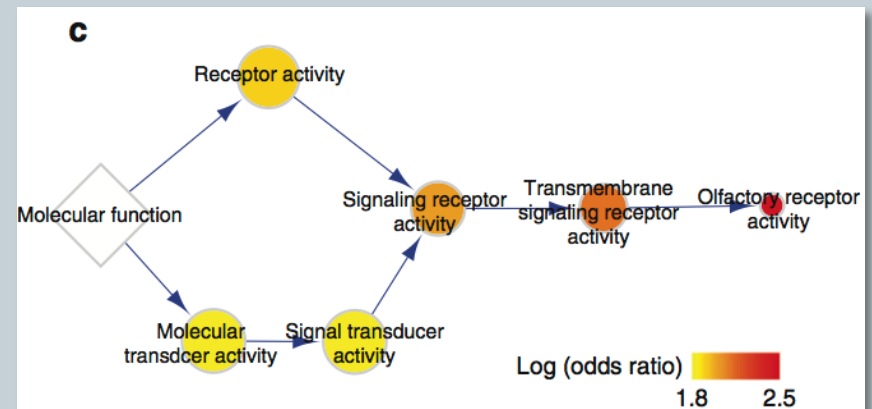


Wild Bactrian Genome Sequence

Table 1 | Statistics of wild camel genome assembly.

Large scaffolds ($\geq 1,000$ bp)		Contigs (≥ 200 bp)	
No. of large scaffolds	13,544	No. of large contigs	73,065
Largest length	15,735,958	Largest length	853,441
Bases in large scaffolds	2,010,007,732	Base in large contigs	1,985,490,767
N50 scaffold length	2,005,940	N50 contig length	85,292
N50 scaffold index	274	N50 contig index	6,136
N90 scaffold length	341,009	N90 contig length	15,299
N90 scaffold index	1,169	N90 contig index	26,984
GC content (%)	41.28	GC content	41.28
N rate (%)	1.18	N rate	0

- Sequenced both wild and domestic individuals
- Estimated genome size: 2.4 GB
- Repetitive element content: 34%
- Protein-coding genes: 20,821



Goals

Demographic history

Natural selection/adaptation

Artificial selection

Applied research on economically valuable traits

