



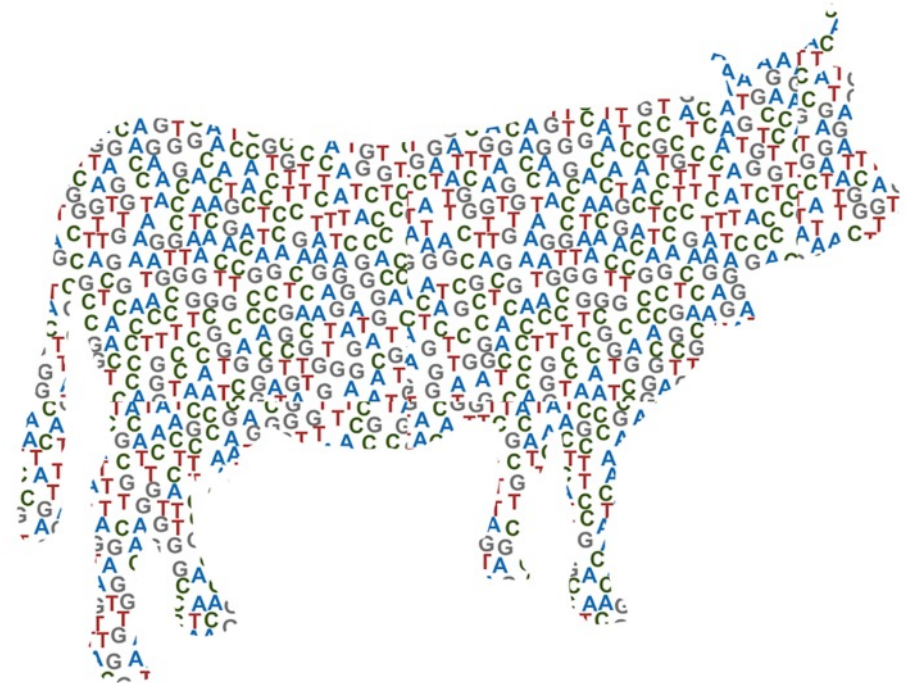
Current projects at the  
ETH Animal Genomics group

**Hubert Pausch**  
Animal Genomics  
11 January 2024



## Topics covered today

- Bovine Pangenomics - Genome assembly and pangenome integration
- Transcriptome-wide association testing - Genome and transcriptome analyses to pinpoint male fertility-associated loci
- Structural variant analyses



A linear reference sequence *fails* to represent diversity.



# Aurochs is the most recent common ancestor of 1.5 billion *taurine* and *indicine* cattle



Aurochs

*Bos taurus primigenius* – *extinct*

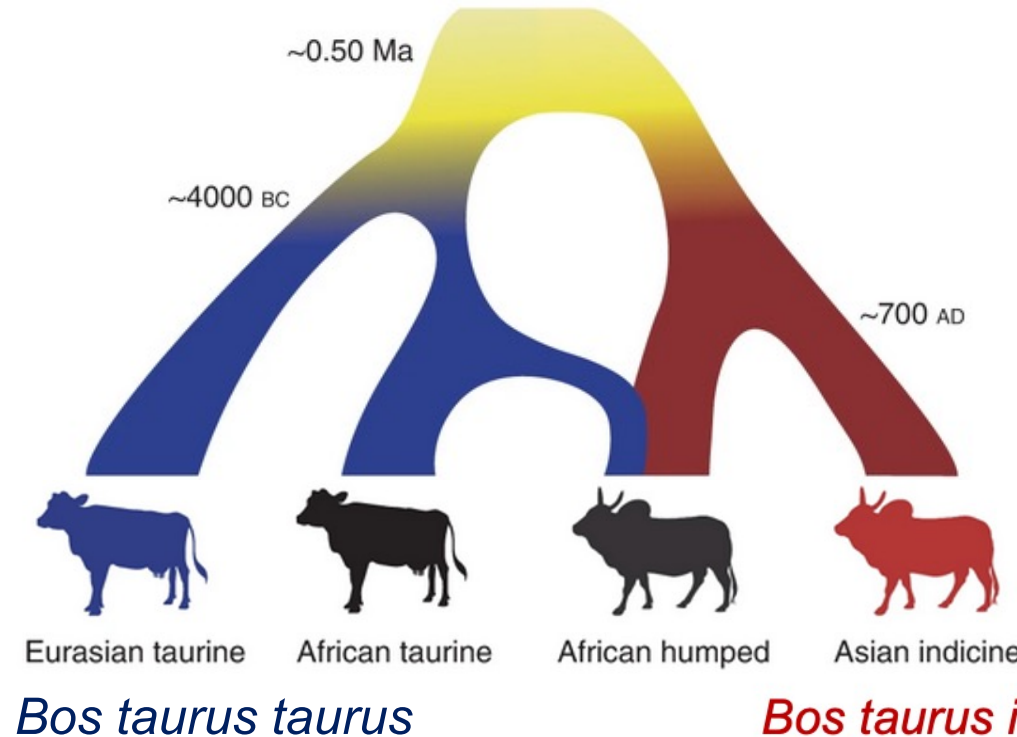


Guzerat



Hereford

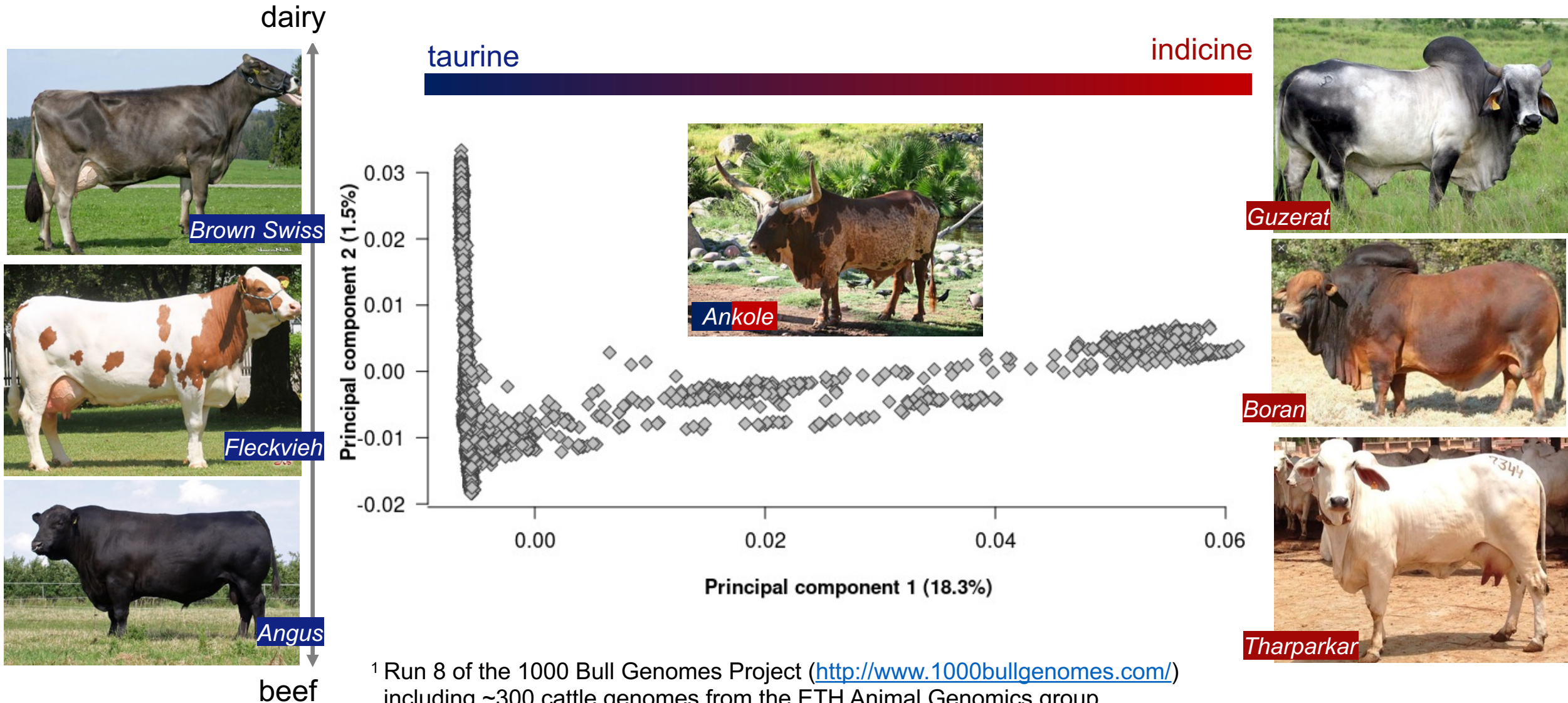
- *Mainly raised under temperate climate conditions*
- *Beef and milk production in both extensive and intensive farming schemes*



- *Mainly raised under tropical conditions*
- *Resistant to diseases and parasites*
- *Well adapted to harsh environments*

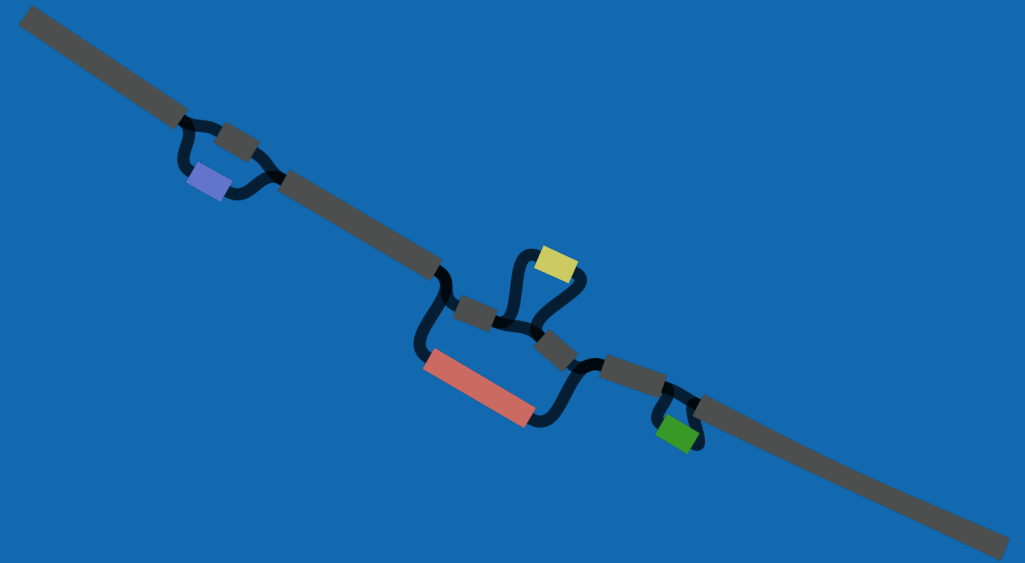


# Genetic diversity in >4700 *taurine* and *indicine* cattle genomes<sup>1</sup>



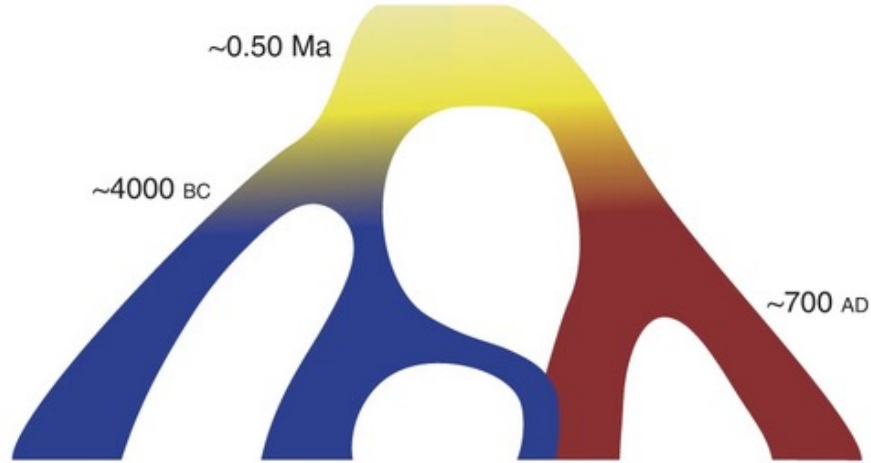
<sup>1</sup> Run 8 of the 1000 Bull Genomes Project (<http://www.1000bullgenomes.com/>) including ~300 cattle genomes from the ETH Animal Genomics group 120 million SNPs

Trio-binned *haplotype* assemblies built from ONT and HiFi sequencing





# Embrace diversity!



Brown Swiss  
*Bos taurus taurus*

X



Nellore  
*Bos taurus indicus*

F1 (\*25.02.2020)  
12.2 M heterozygous SNP

PacBio HiFi: 6 SMRT cells, 140 Gb CCS (20 kb library): **46-fold**

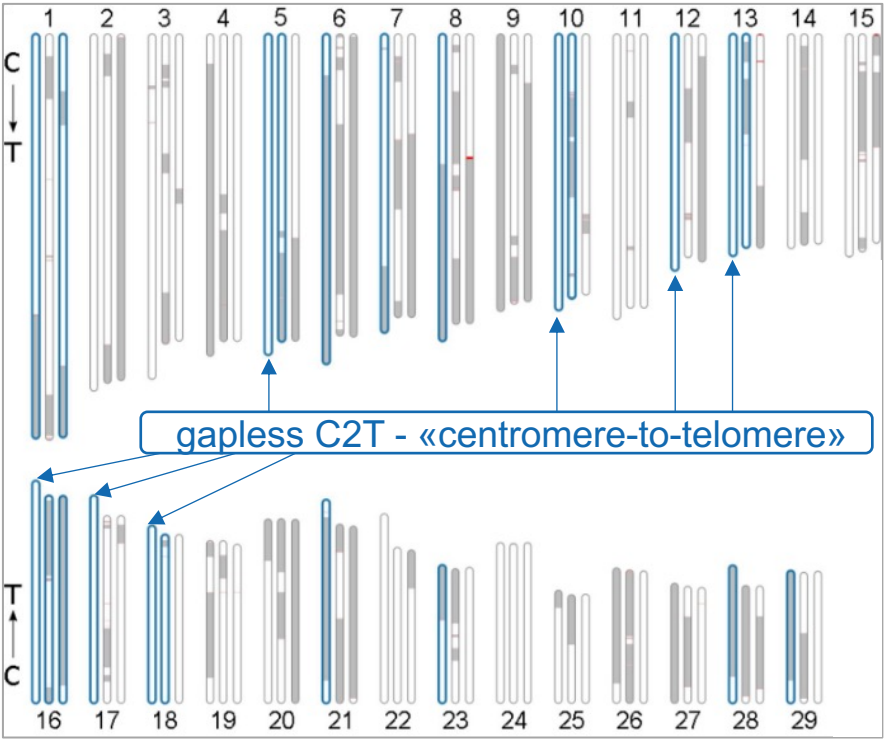
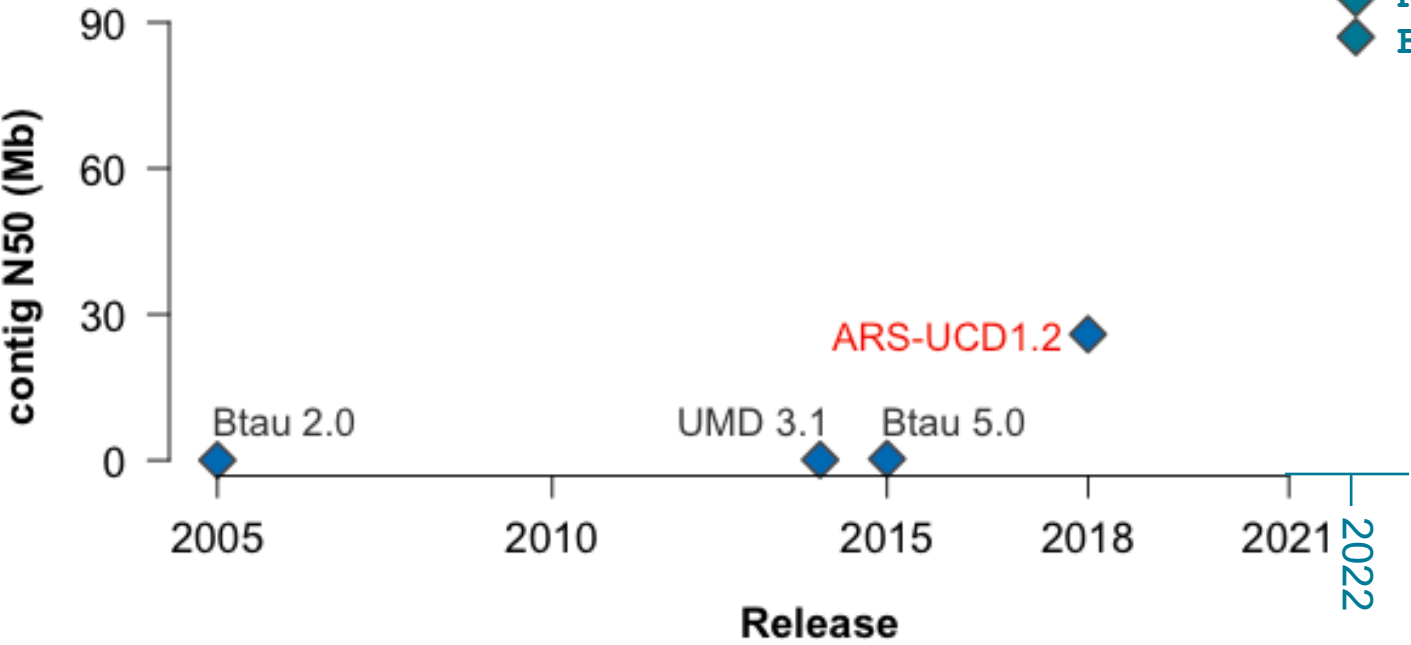
AgroVet  
Strickhof



# Two near-perfect assemblies from one F1



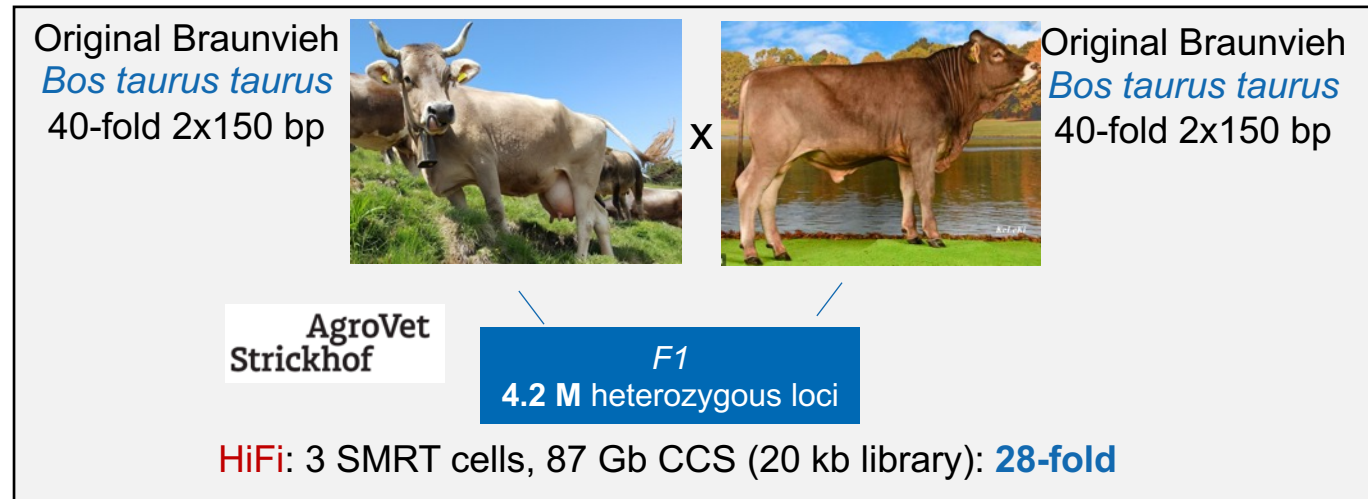
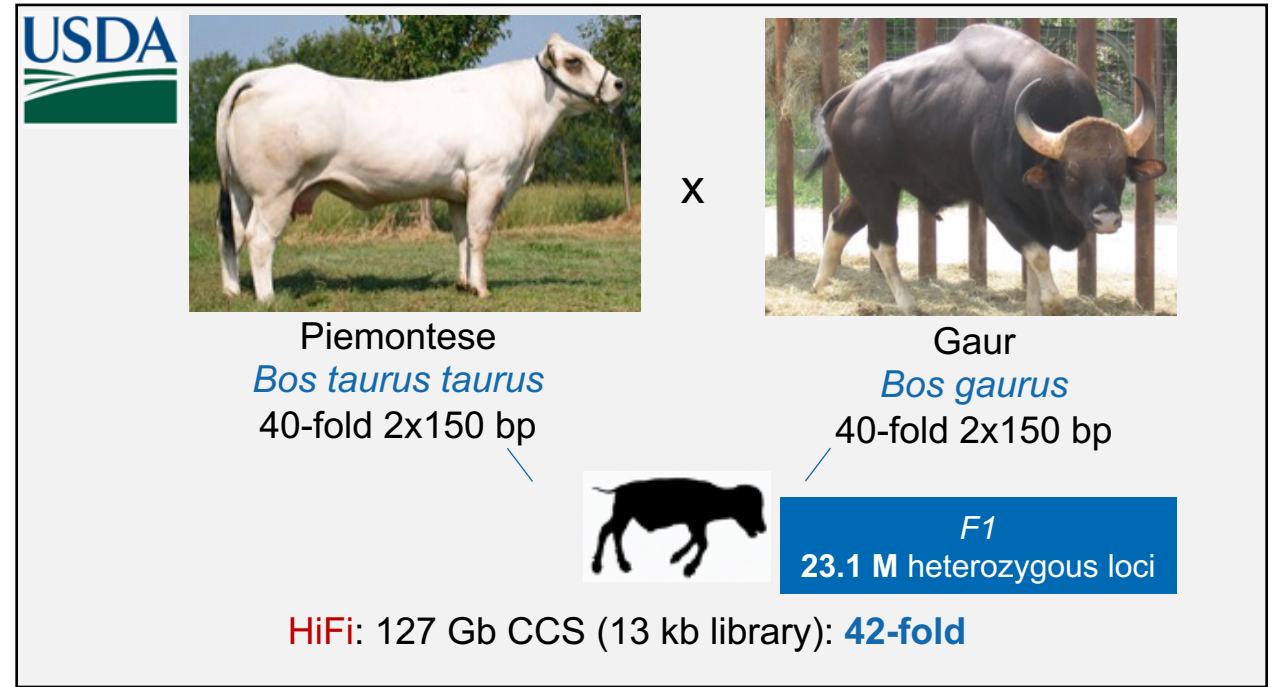
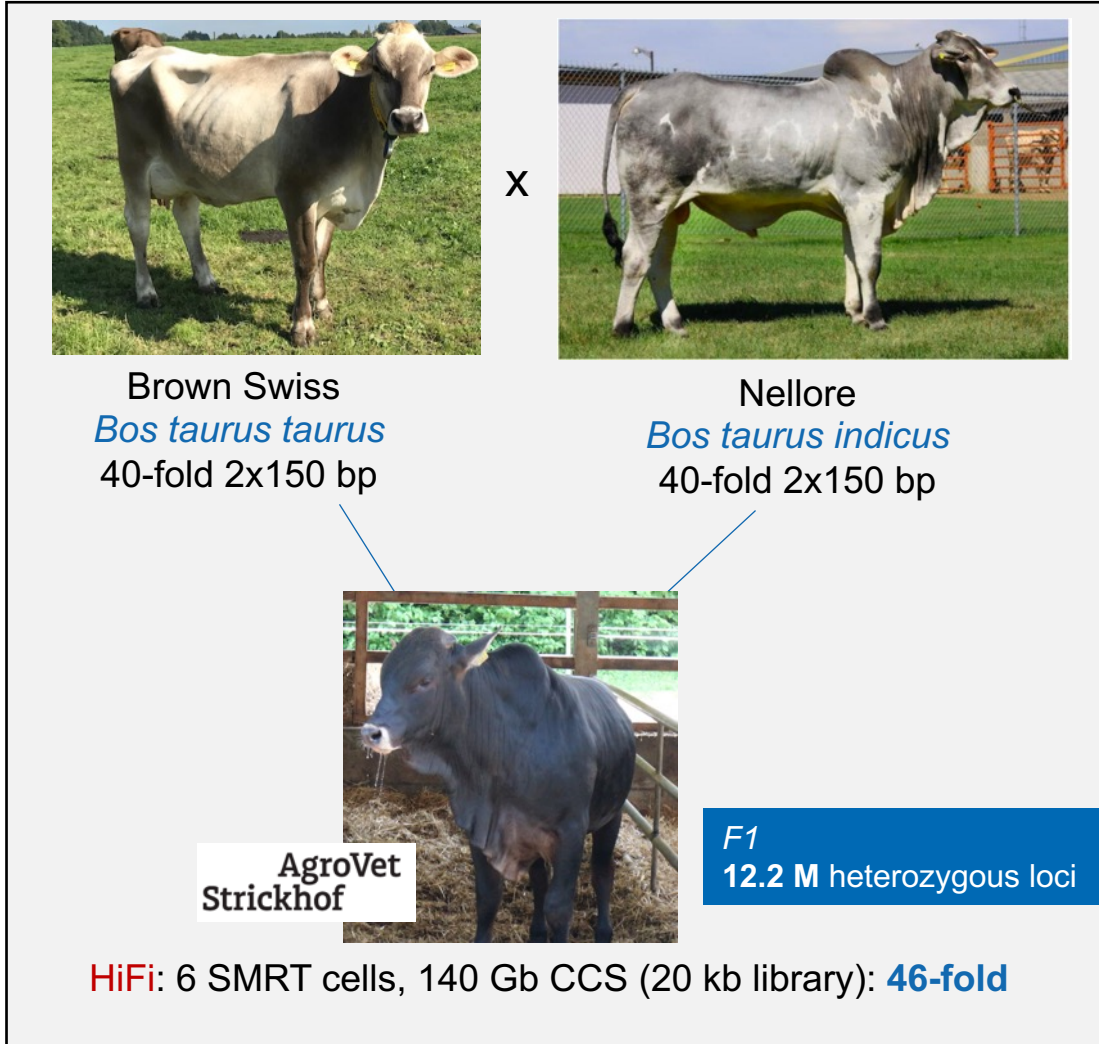
◆ Nelore @ETH  
 ◆ BSW @ETH



Alexander Leonard



# Sequenced trios



# HiFi and ONT trio-binned assemblies



Alexander Leonard

Breed or species	Haplotype	Read technology	Size (autosomal size)	Contigs (autosomal contigs)	NG50	PG50	QV	BUSCO (single copy)	Repeat
Original Braunvieh	Maternal	HiFi	3.11 (2.57)	1706 (105)	47.0	23.6	49.7	95.7 (93.9)	48.95
	Maternal	ONT	2.70 (2.48)	2622 (109)	71.6	2.8	40.7	95.1 (93.4)	43.19
Nellore	Paternal (Y)	HiFi	2.95 (2.60)	1217 (52)	94.4	79.1	46.1	93.3 (91.8)	47.81
	Paternal (Y)	ONT	2.57 (2.49)	1457 (67)	68.5	64.9	42.4	92.8 (91.3)	42.64
Brown Swiss	Maternal	HiFi	3.07 (2.62)	1045 (58)	86.7	81.1	45.6	95.9 (94.2)	48.43
	Maternal	ONT	2.67 (2.48)	1268 (71)	64.0	53.0	42.5	95.3 (93.7)	42.85
gaur	Paternal (X)	HiFi	3.02 (2.52)	1352 (75)	73.5	61.2	48.4	95.7 (94.1)	47.73
	Paternal (X)	ONT	2.64 (2.48)	532 (89)	68.1	68.1	41.2	95.1 (93.3)	42.26
Piedmontese	Maternal	HiFi	3.10 (2.56)	1427 (90)	52.0	47.6	48.3	95.8 (94.1)	48.43
	Maternal	ONT	2.66 (2.48)	782 (64)	82.8	82.8	40.9	95.3 (93.6)	43.06
Hereford (ARS-UCD1.2)	(N/A)	CLR	2.72 (2.49)	2597 (289)	25.9	N/A	35.8	95.7 (93.9)	42.96
<i>VGP Standards</i>					1	0.1	40	90	N/A



ARTICLE



<https://doi.org/10.1038/s41467-022-30680-2> OPEN

## Structural variant-based pangenome construction has low sensitivity to variability of haplotype-resolved bovine assemblies

Alexander S. Leonard<sup>1</sup>, Danang Crysanto<sup>1</sup>, Zih-Hua Fang<sup>1</sup>, Michael P. Heaton<sup>2</sup>, Brian L. Vander Ley<sup>3</sup>, Carolina Herrera<sup>4</sup>, Heinrich Bollwein<sup>4</sup>, Derek M. Bickhart<sup>5</sup>, Kristen L. Kuhn<sup>2</sup>, Timothy P. L. Smith<sup>2</sup>, Benjamin D. Rosen<sup>6</sup> & Hubert Pausch<sup>1</sup>



RESEARCH

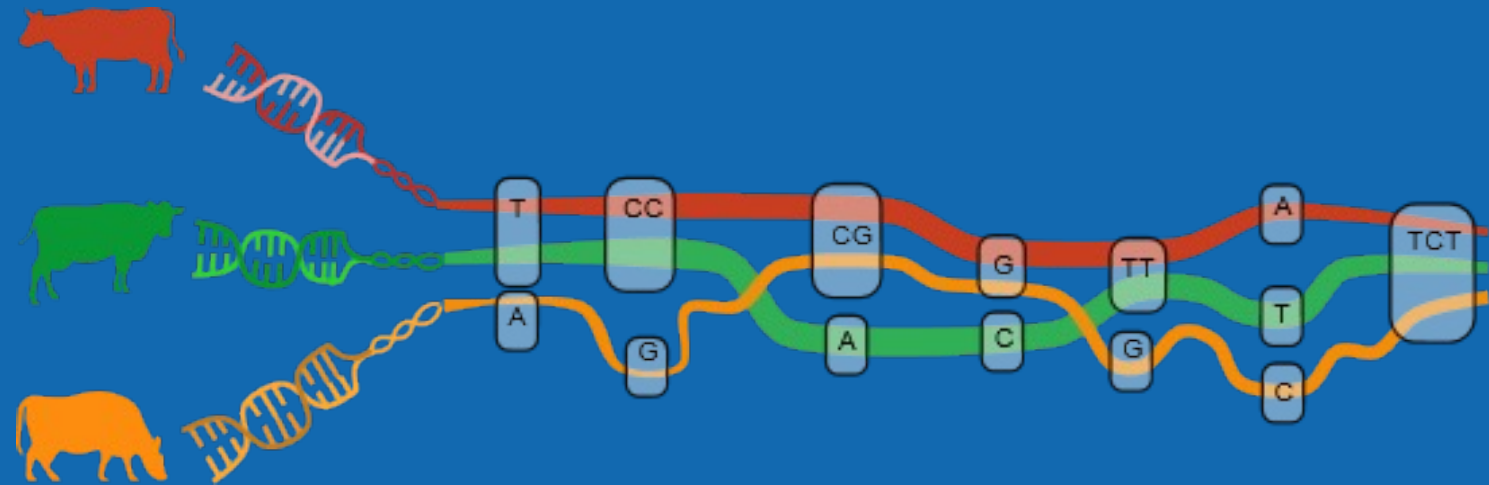
Open Access

## Graph construction method impacts variation representation and analyses in a bovine super-pangenome



Alexander S. Leonard<sup>1\*</sup>, Danang Crysanto<sup>1</sup>, Xena M. Mapel<sup>1</sup>, Meenu Bhati<sup>1</sup> and Hubert Pausch<sup>1\*</sup>

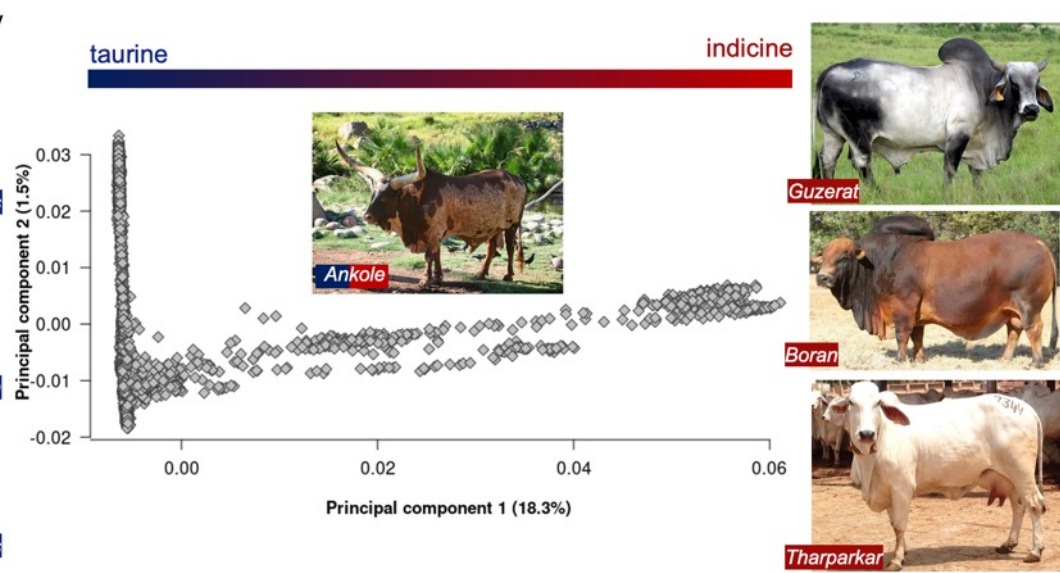
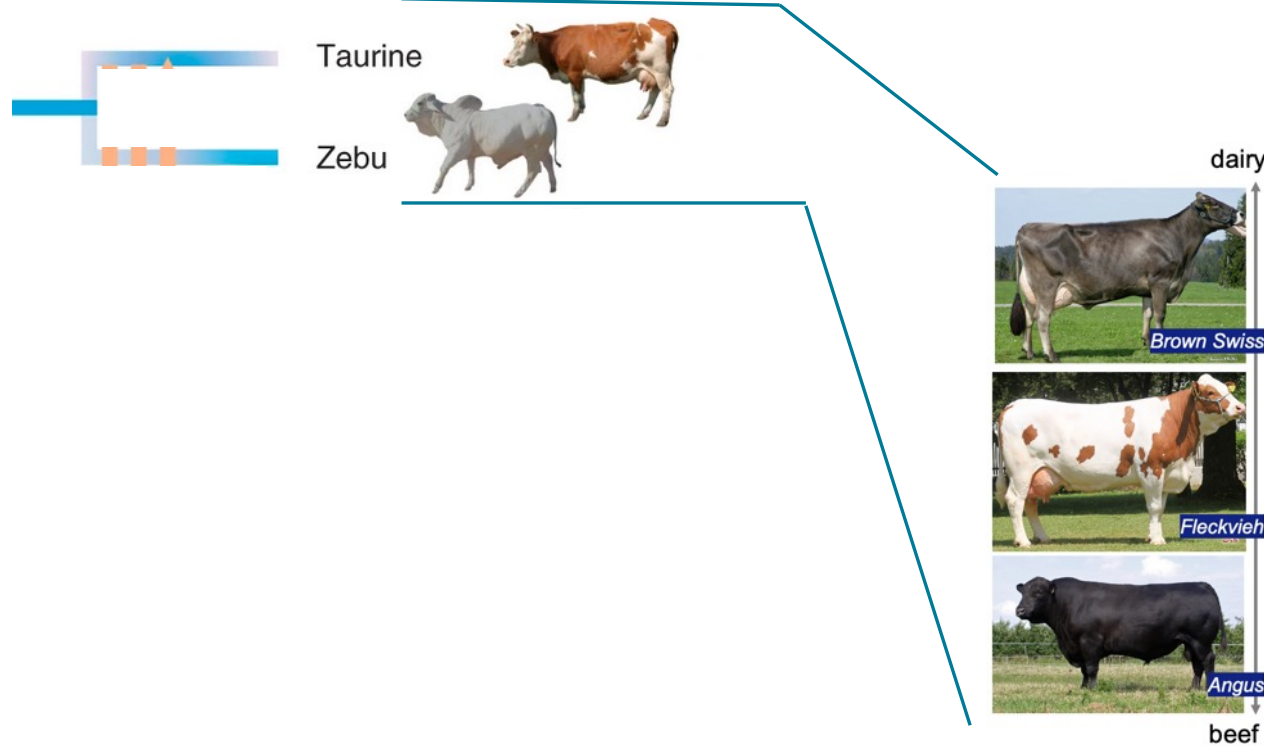
En route to a bovine *super-pangenome*\*.



\*full set of genomic elements from all species within a genus

# There's huge diversity within the genus *Bos*

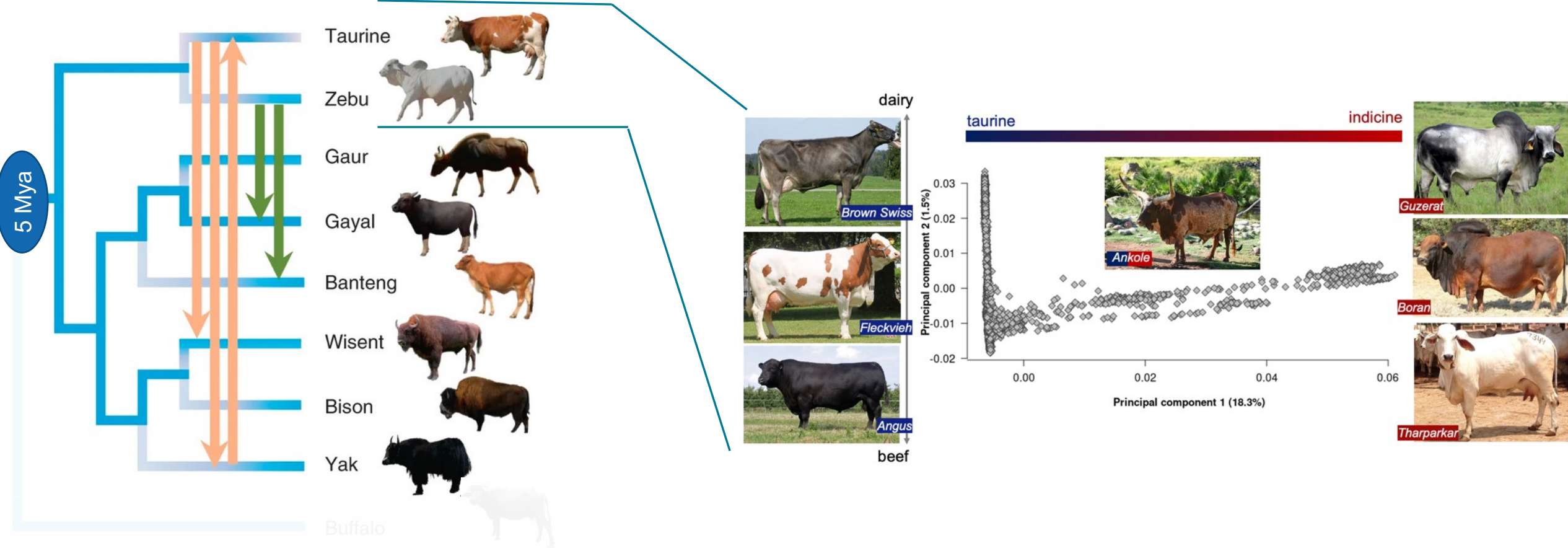
## Species within the genus *Bos*





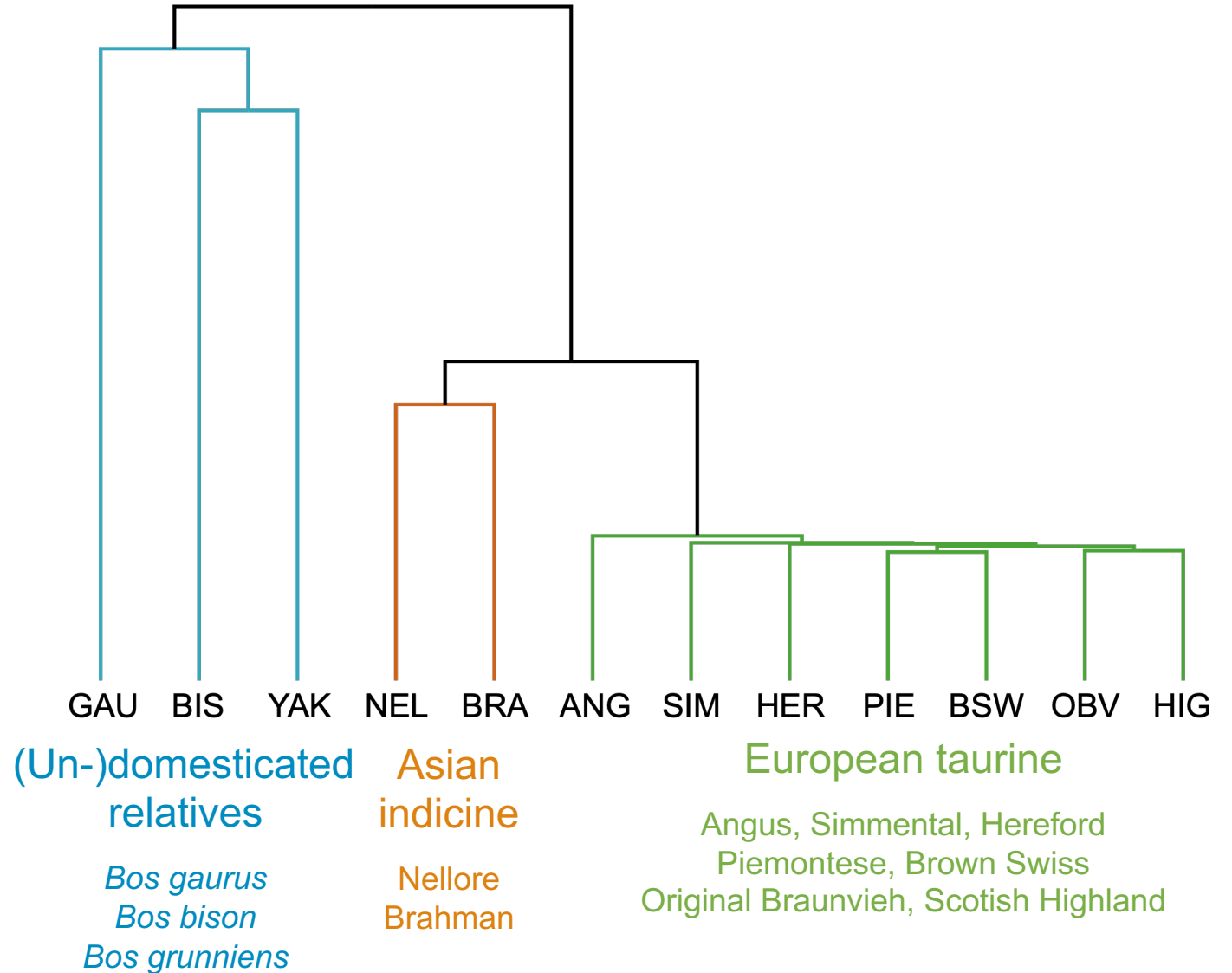
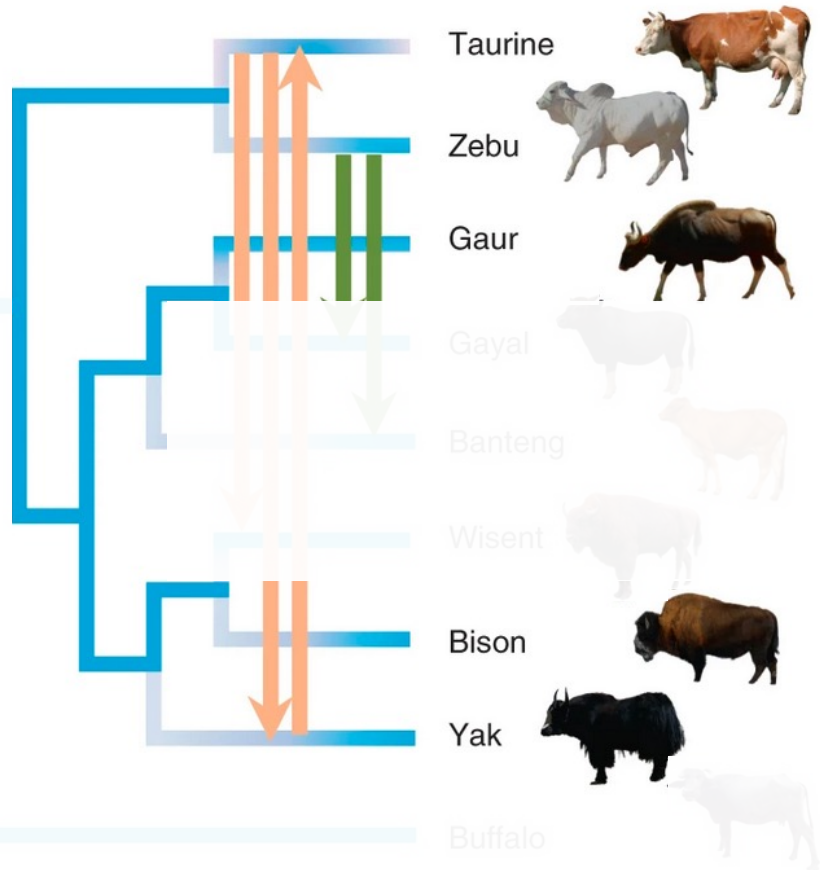
# There's huge diversity within the genus *Bos*

## Species within the genus *Bos*



# Construction of bovine *super-pangenomes* from haplotype assemblies

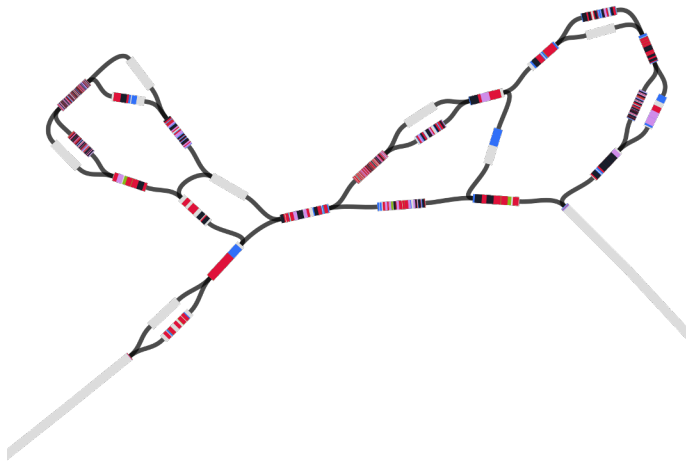
## Twelve haplotype assemblies





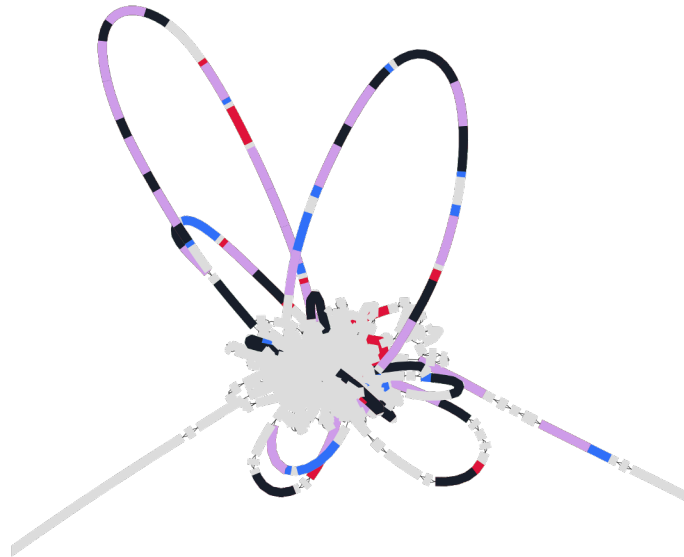
# Three bovine *super-pangenomes* from twelve haplotype assemblies

**minigraph**



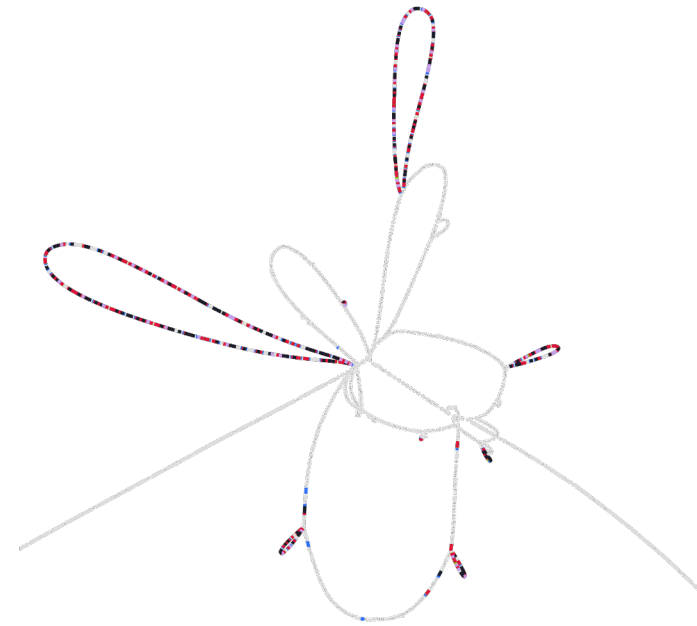
*structural variant pangenomes  
backbone assembly  
includes variants >50bp*

**pggb**



*reference-free methods  
can project onto reference coordinates  
include all types of variants*

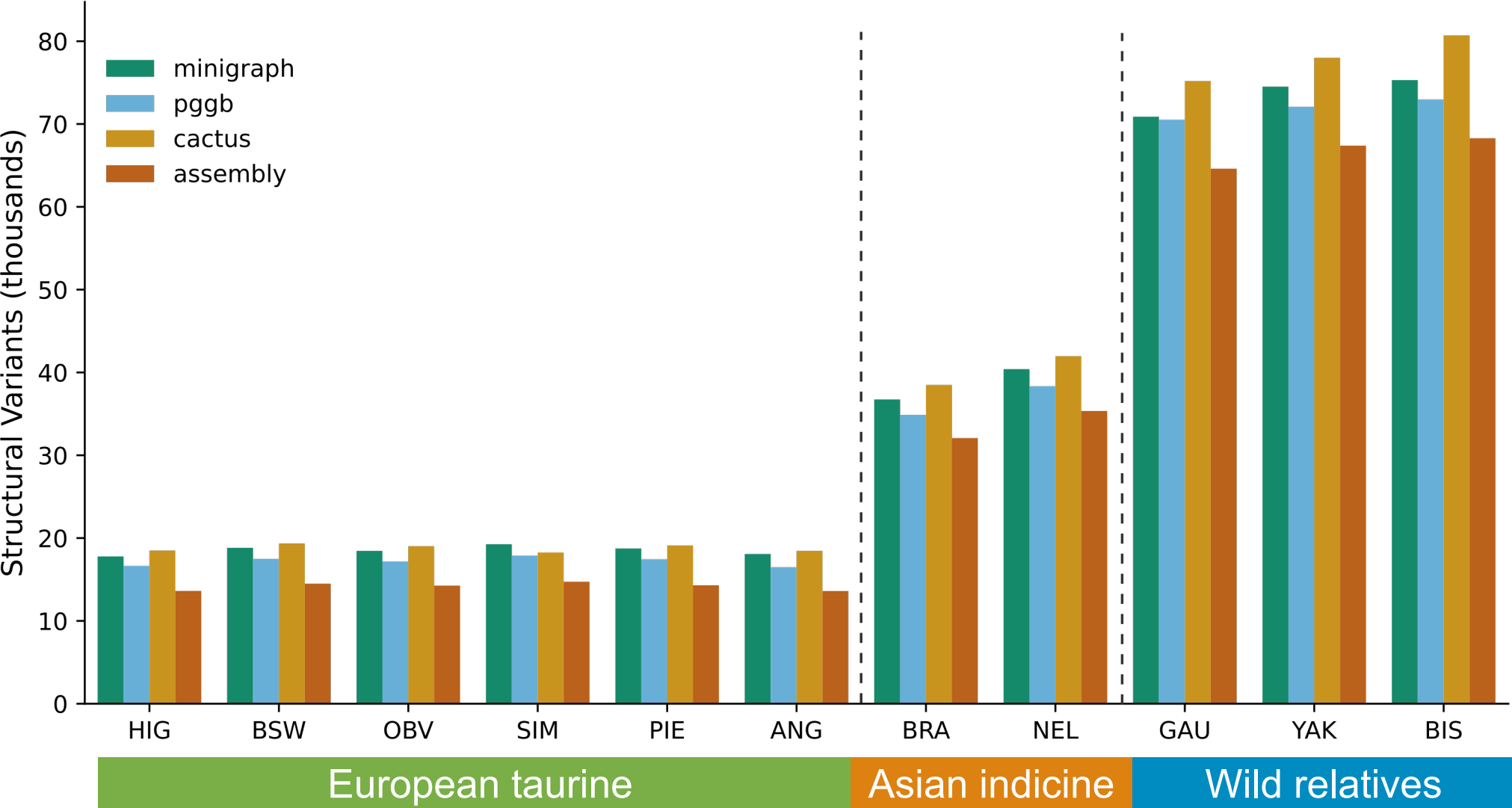
**cactus**



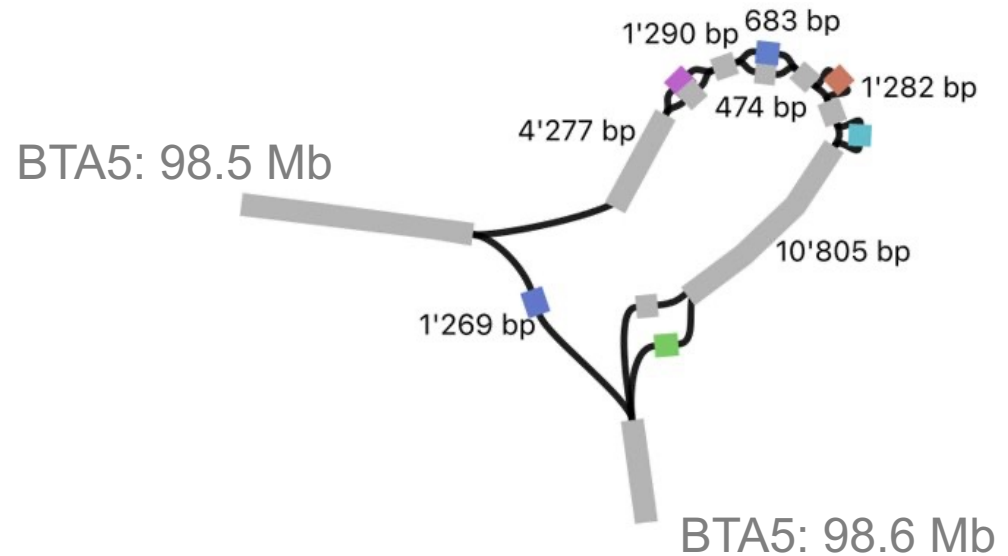
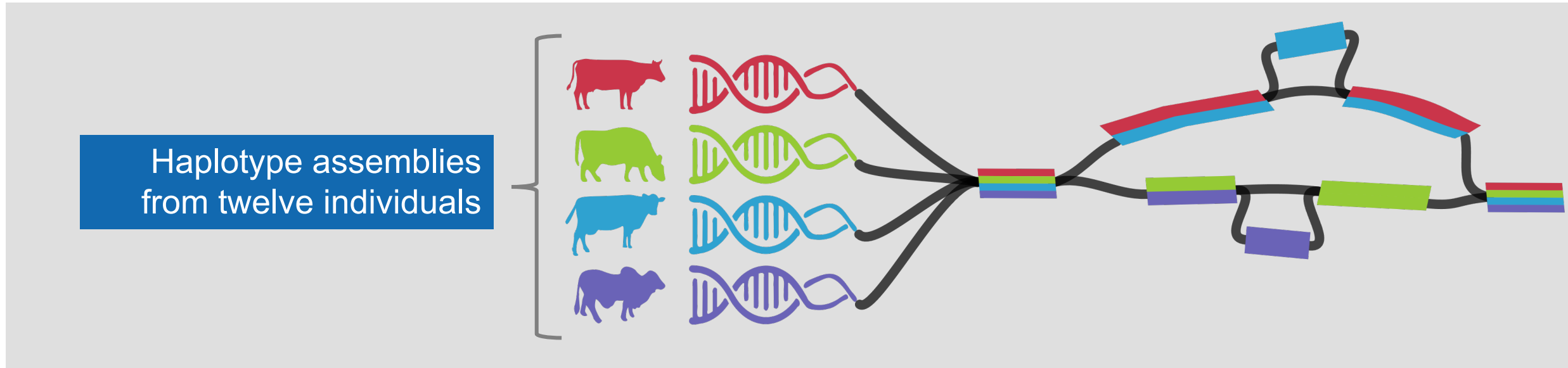




# Amount of structural variation detected from bovine *super-pangenomes*

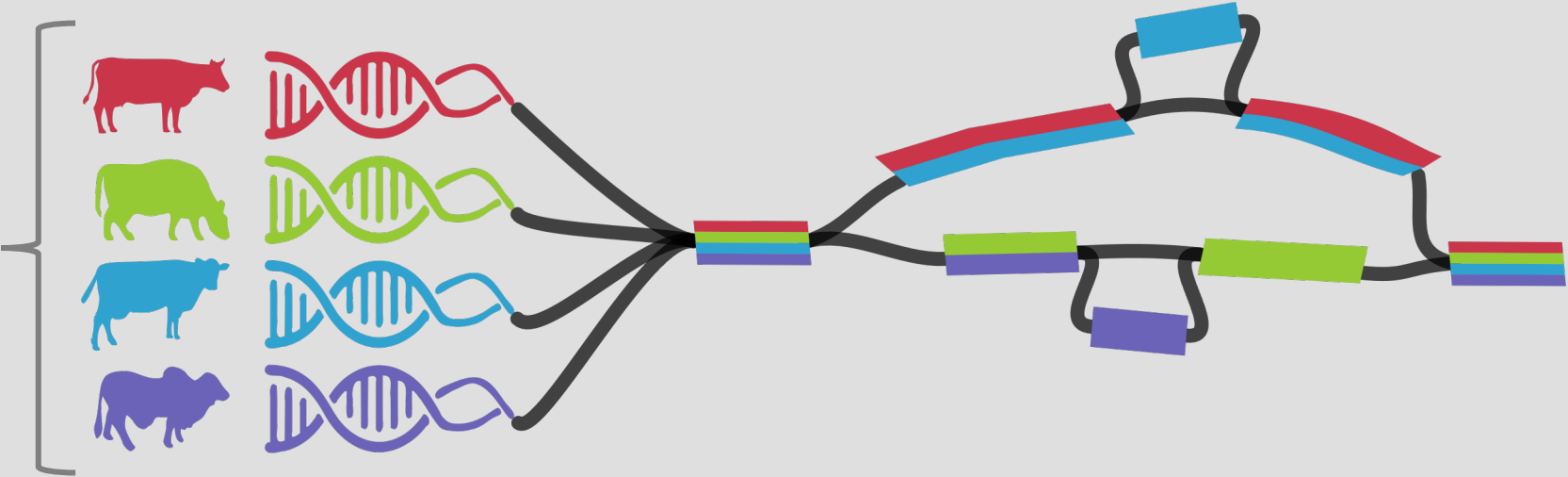


# Large insertion/deletion variants segregate across (sub-)species

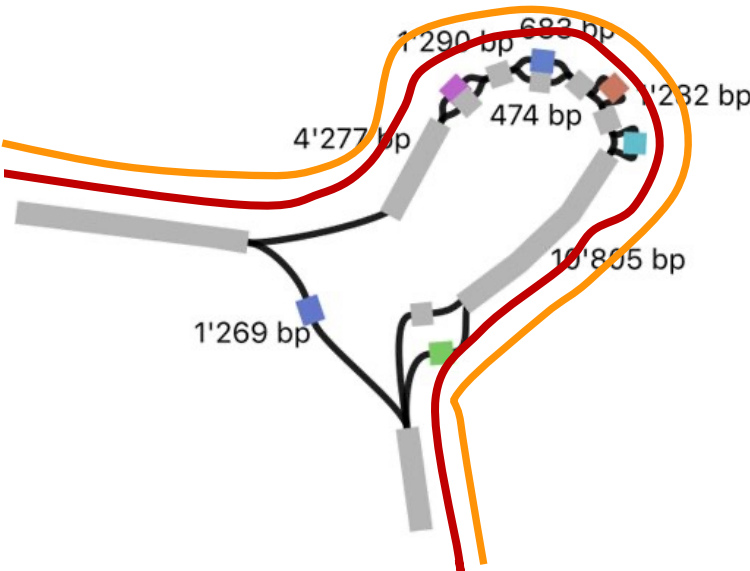


# Large insertion/deletion variants segregate across (sub-)species

Haplotype assemblies from twelve individuals



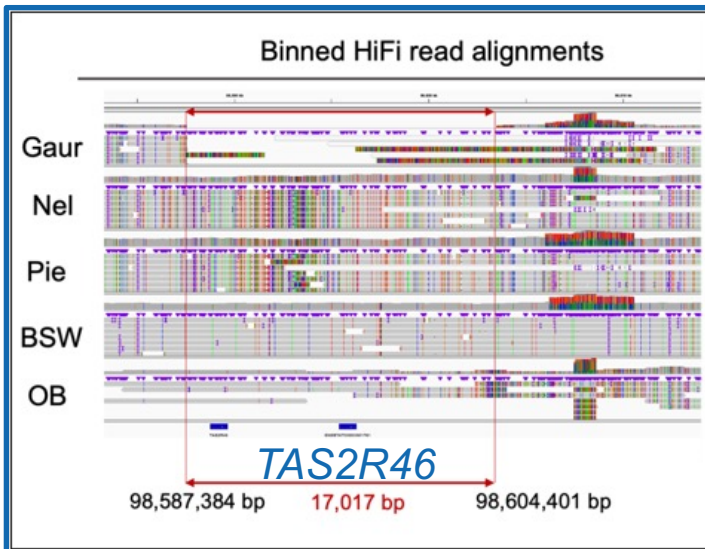
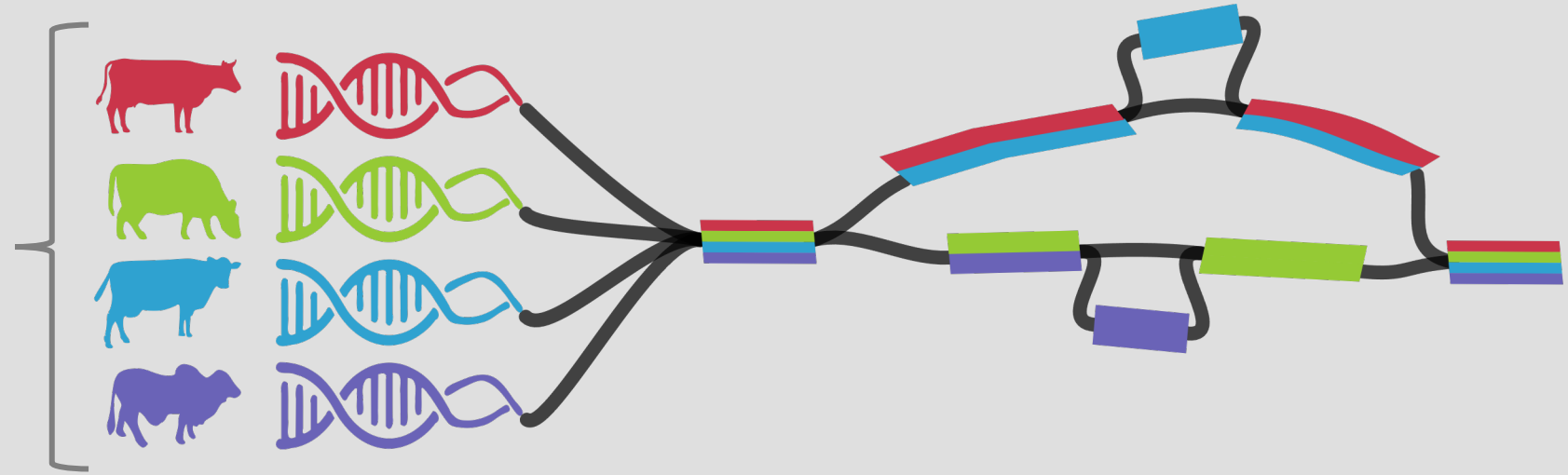
*Bos taurus taurus* / *Bos taurus indicus*



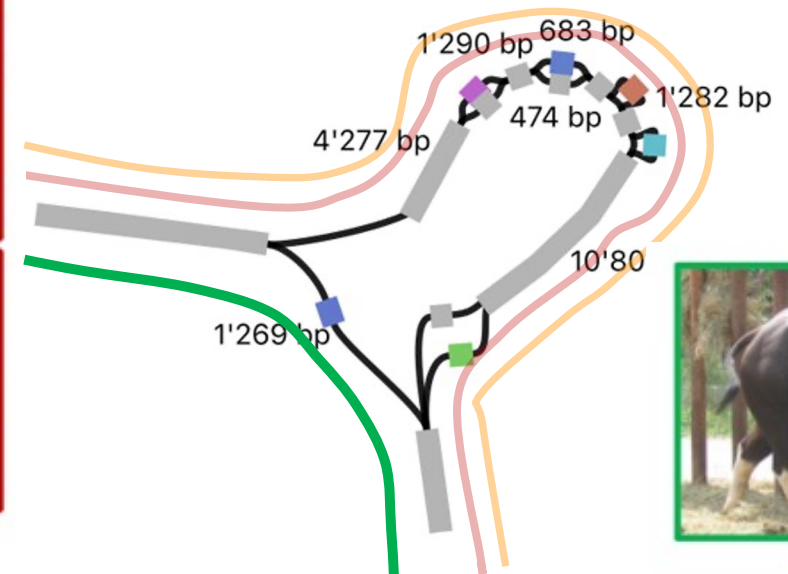


# Large insertion/deletion variants segregate across (sub-)species

Haplotype assemblies from twelve individuals



*Bos taurus taurus* / *Bos taurus indicus*



*Bos gaurus*

*Pangenome-wide association testing.*

A dominant white head is characteristic for Fleckvieh and Hereford cattle



Gir x Simmental



Grauvieh x Simmental



Eringer x Simmental



Future plans



Smith et al. *Genome Biology* (2023) 24:139  
<https://doi.org/10.1186/s13059-023-02975-0>

Genome Biology

CORRESPONDENCE

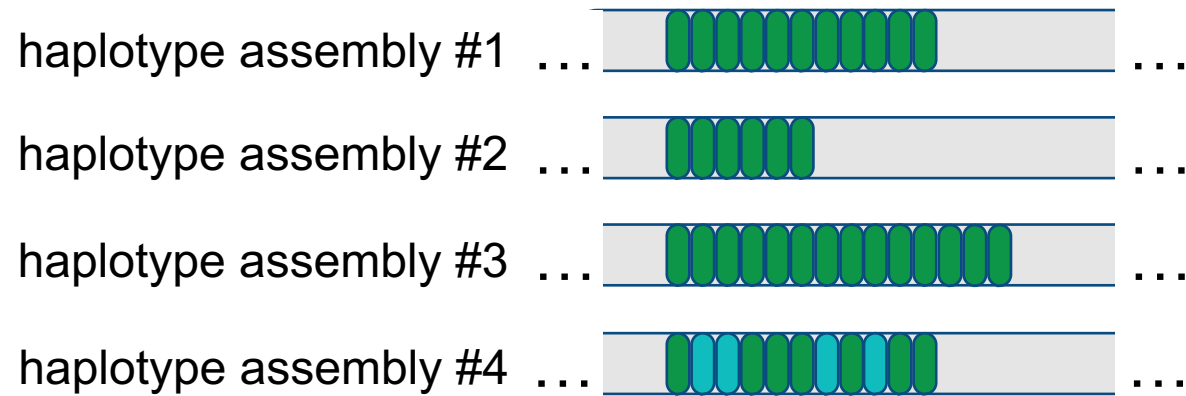
Open Access

## The Bovine Pangenome Consortium: democratizing production and accessibility of genome assemblies for global cattle breeds and other bovine species



Timothy P. L. Smith<sup>1</sup>, Derek M. Bickhart<sup>2</sup>, Didier Boichard<sup>3</sup>, Amanda J. Chamberlain<sup>4,5</sup>, Appolinaire Djikeng<sup>6,7</sup>, Yu Jiang<sup>8</sup>, Wai Y. Low<sup>9</sup>, Hubert Pausch<sup>10</sup>, Sebastian Demyda-Peyrás<sup>11,12</sup>, James Prendergast<sup>7,13</sup>, Robert D. Schnabel<sup>14</sup>, Benjamin D. Rosen<sup>15\*</sup> and Bovine Pangenome Consortium

# Pangenomes make tandem repeat elements (e.g., VNTRs) amenable to association testing



Variable **N**umber **T**andem **R**epeats

This is a **repetitive** motif.

This is a **repetitive rapetitive** motif.

This is a **repetitive repetitive repetitive** motif.

This is a **repatotive rapetitive repetitive** motif.

This is a **repatotive repetitive** motif.

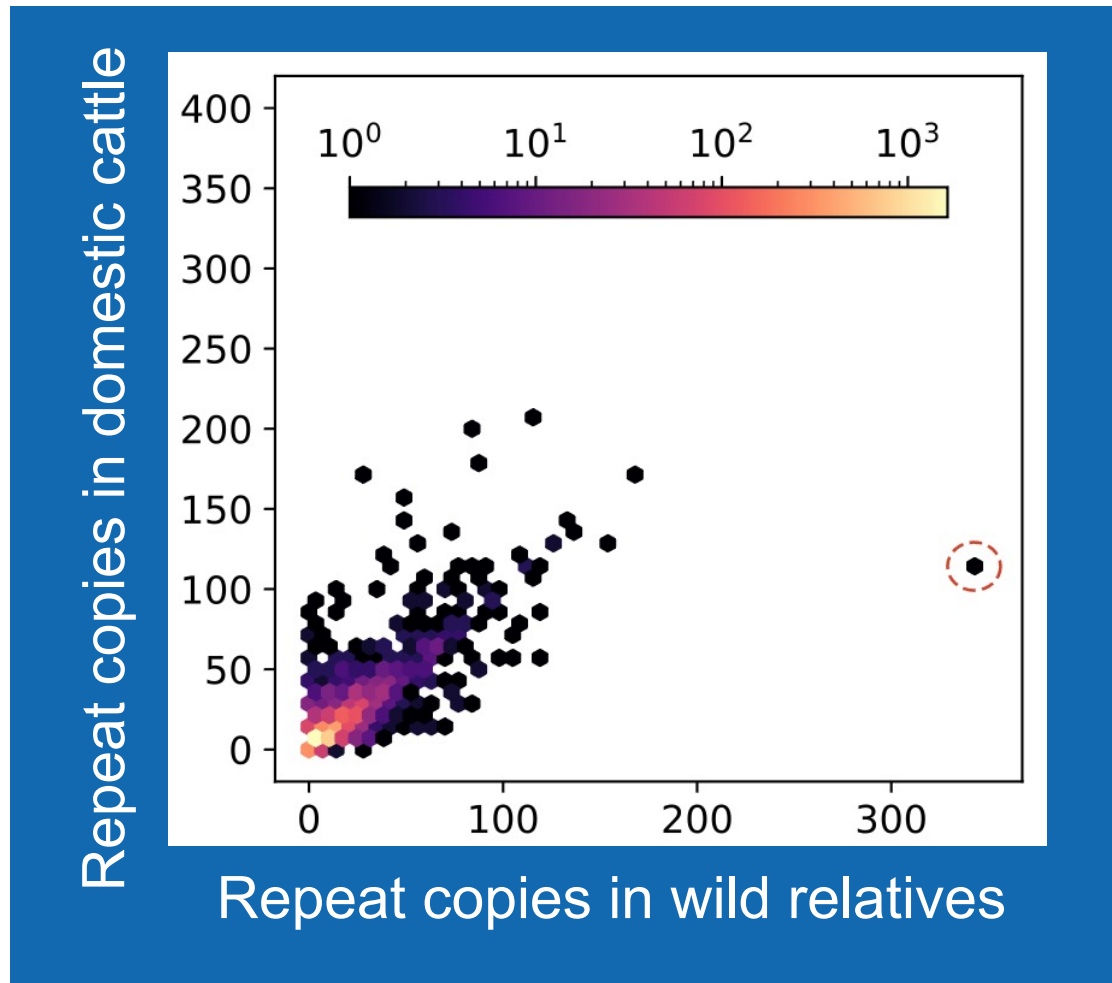
This is a **repetitive rapetitive repetitive rapetitive** motif.

Repeat counts

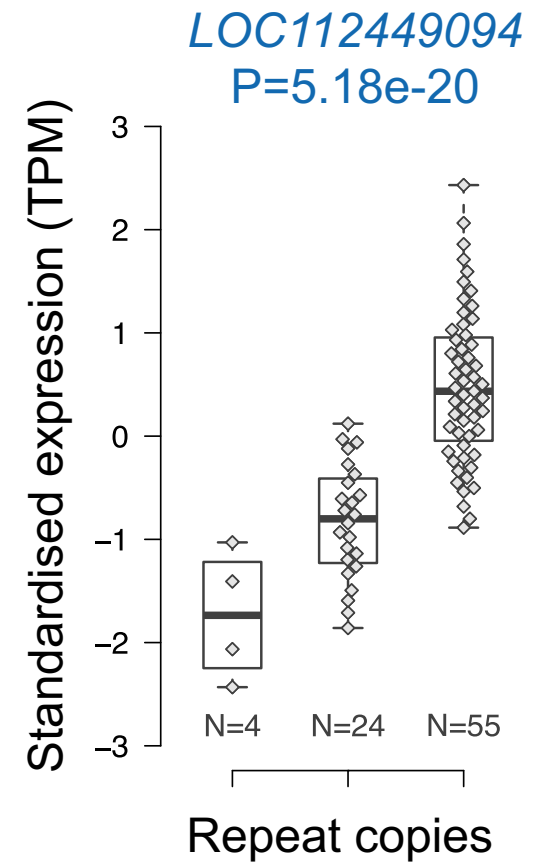
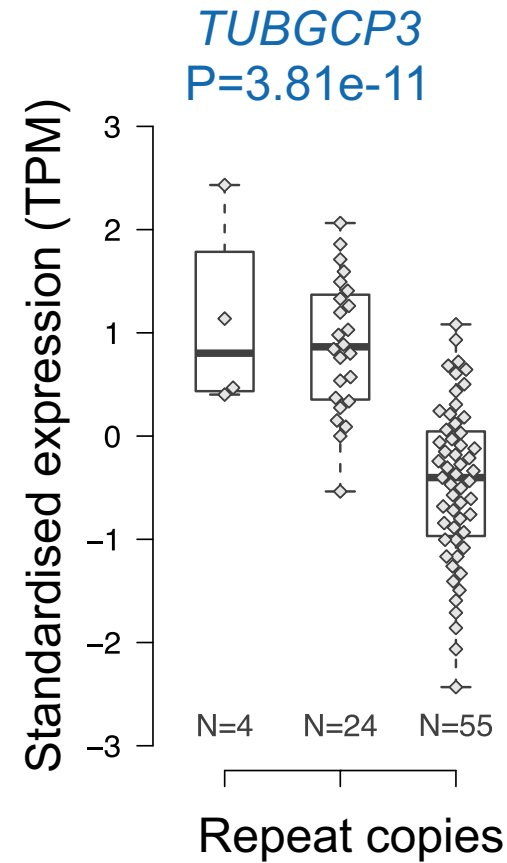
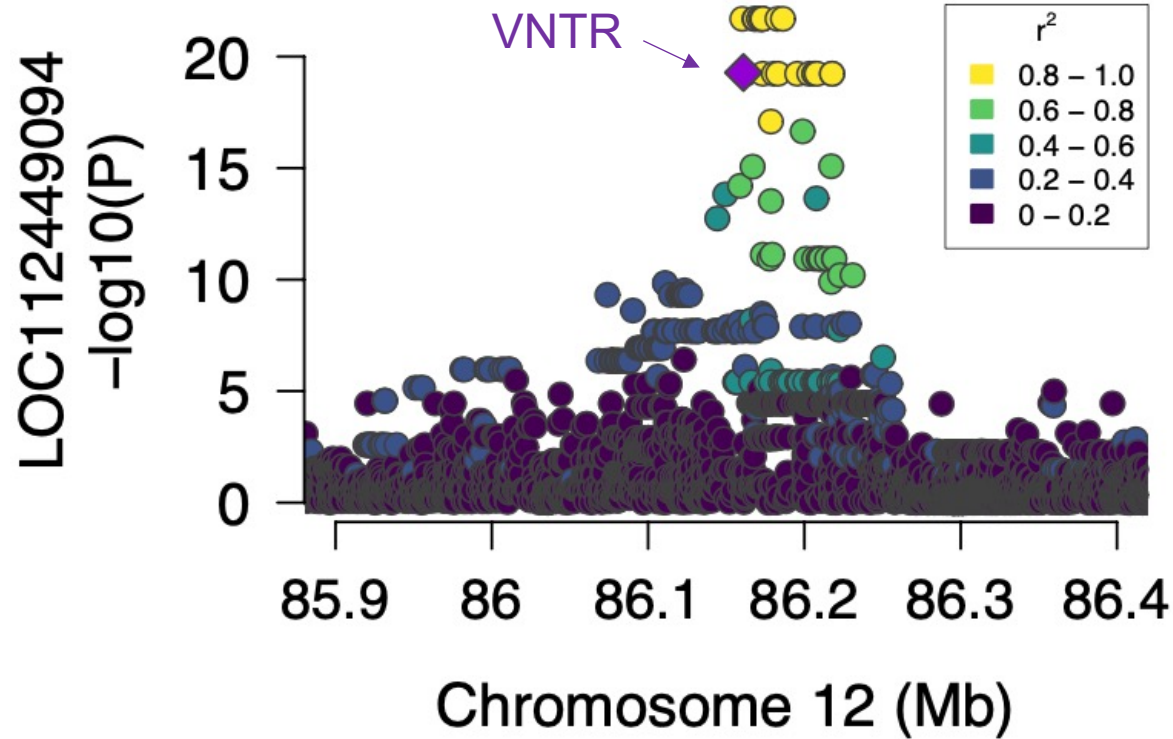
1  
2  
3  
3  
2  
4



# Pangenomes make tandem repeat elements (e.g., VNTRs) amenable to association testing



# The VNTR mediates expression of neighboring genes and non-coding RNAs



# Pangenome variation improves QTL mapping

New Results

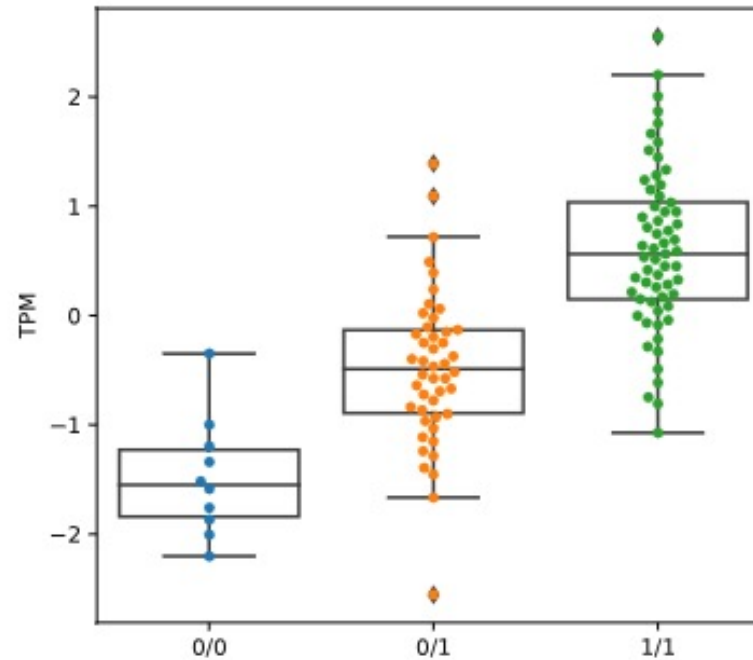
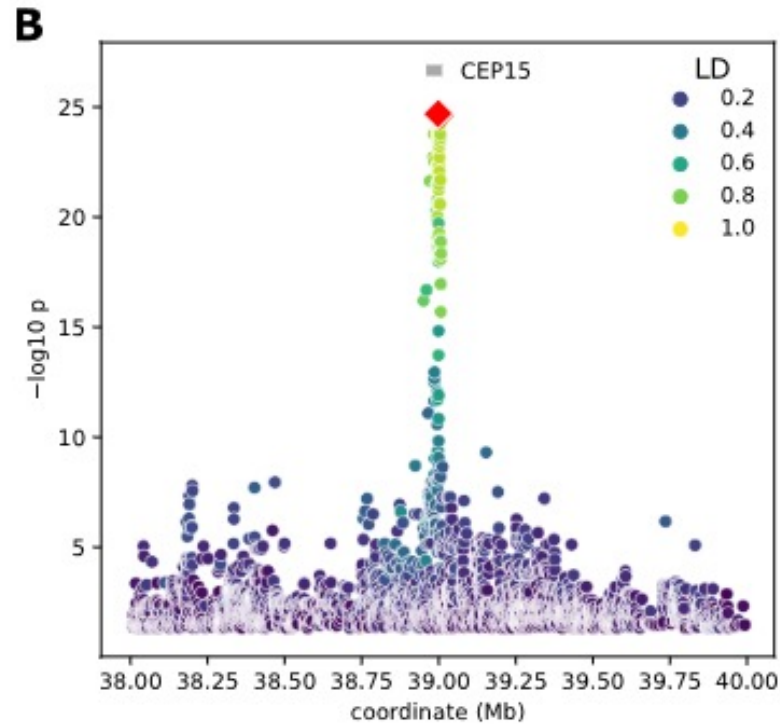
[Follow this preprint](#)

## Pangenome genotyped structural variation improves molecular phenotype mapping in cattle

[ID](#) Alexander S. Leonard, Xena M. Mapel, Hubert Pausch

doi: <https://doi.org/10.1101/2023.06.21.545879>

This article is a preprint and has not been certified by peer review [what does this mean?].

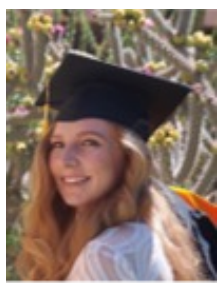




# Transcriptome-wide association testing

# An eQTL cohort to map molecular phenotypes that are associated with male reproductive performance

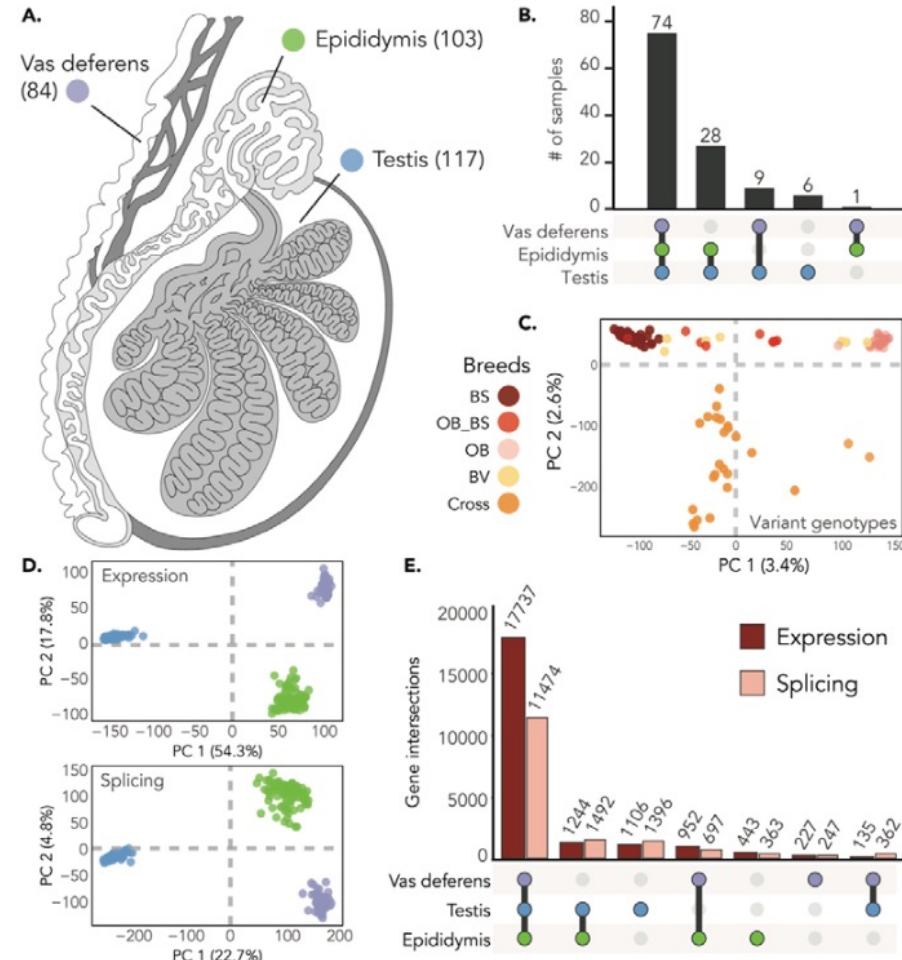
- eQTL cohort
  - Tissue sampled from 128 mature bulls
  - DNA sequencing to 12.6-fold depth
  - Deep RNAseq for male reproductive tissues (testis, epididymis, vas deferens)
  - Used for e/sQTL mapping
  - Reference cohort for TWAS – transcriptome-wide association testing



Xena Marie Mapel



Naveen Kumar Kadri



CSH Cold Spring Harbor Laboratory

**bioRxiv**  
THE PREPRINT SERVER FOR BIOLOGY

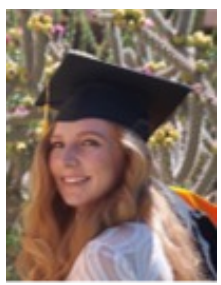
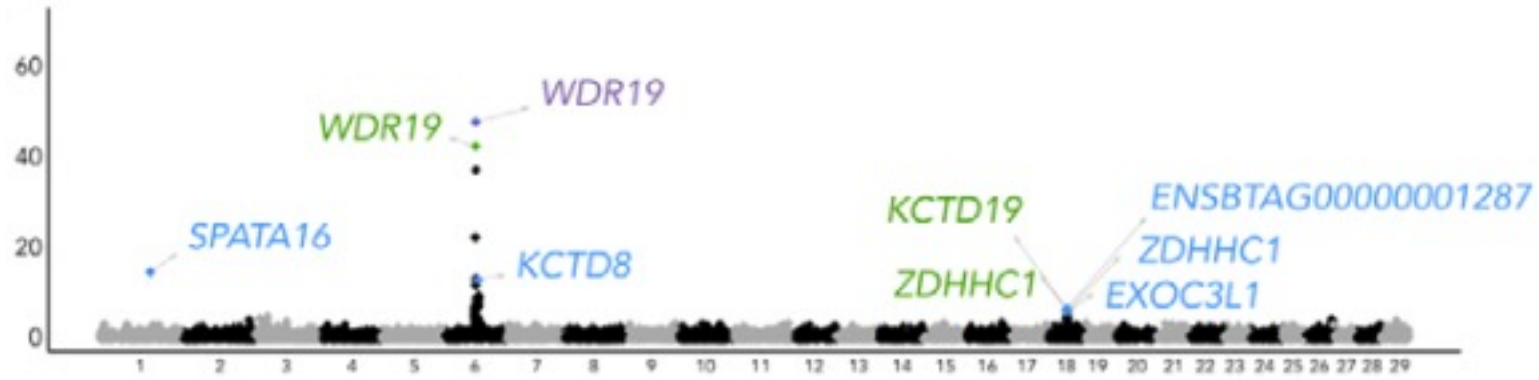
New Results [Follow this preprint](#)

**Molecular quantitative trait loci in reproductive tissues impact male fertility in cattle**

Xena Marie Mapel, Naveen Kumar Kadri, Alexander S. Leonard, Qiongyu He, Audald Lloret-Villas, Meenu Bhati, Maya Hiltbold, Hubert Pausch

doi: <https://doi.org/10.1101/2023.06.29.547066>

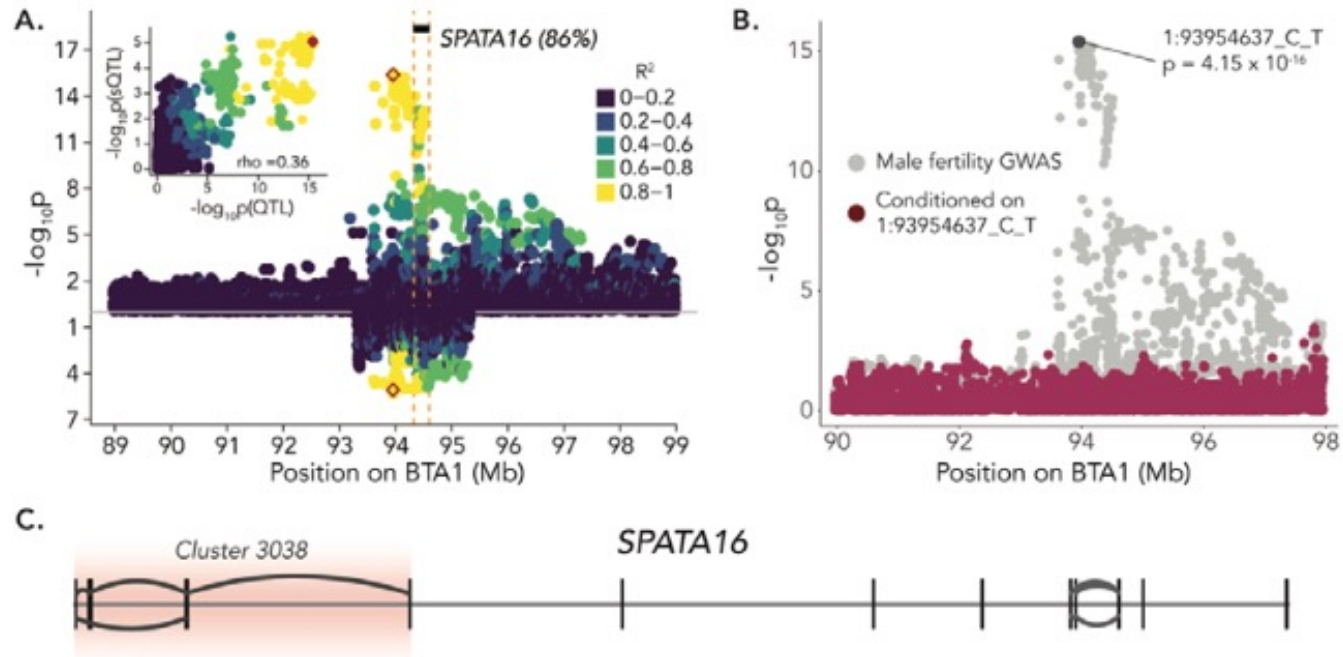
# Transcriptome-wide association testing for male fertility



Xena Mapel



Naveen Kadri





# Contributors & funding



PacBio

BRAUNVIEH 

AgroVet  
Strickhof

Eine Kooperation in Bildung und Forschung:  
  

Qualitas 


Functional Genomics Center Zurich



European  
Commission

Horizon 2020  
European Union funding  
for Research & Innovation

BovReg   
Understanding cattle genomes

 Schweizerische Eidgenossenschaft  
Confédération suisse  
Confederazione Svizzera  
Confederaziun svizra

Federal Office for Agriculture FOAG

Scope 

swissgenetics 

  
SUISAG

FNSNF

FONDS NATIONAL SUISSE  
SCHWEIZERISCHER NATIONALFONDS  
FONDO NAZIONALE SVIZZERO  
SWISS NATIONAL SCIENCE FOUNDATION



