



A Modern Approach to Genotyping

Jesse Hoff, PhD | Agrigenomics Product Manager, Gencove
BIF | June 2nd, 2022



Outline

1

Who is Gencove?

2

What is low-pass sequencing?

3

Application of Gencove technology:
An EPD for Bovine Congestive Heart Failure

4

Application of Gencove technology:
Low-pass + target capture with InfiniSEEK



Who we are



Joe Pickrell

CEO

PhD, Human Genetics
University of Chicago
Professor, Columbia/
New York Genome
Center



Tomaz Berisa

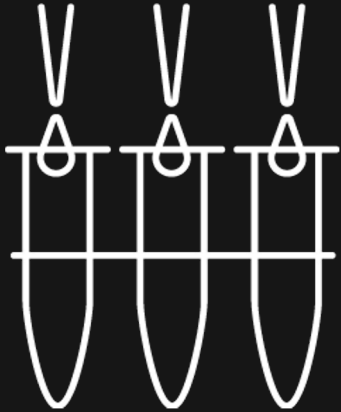
CTO

PhD, EE , Computer
Science University of
Zagreb
New York Genome Center

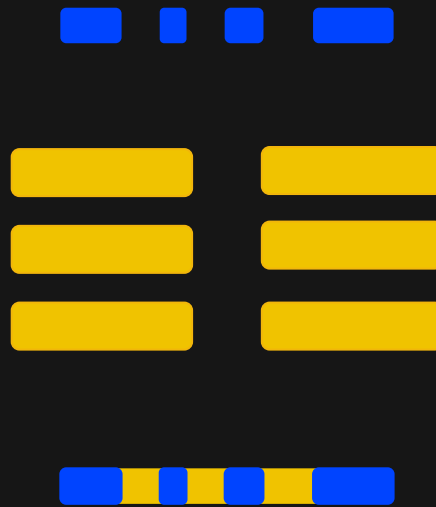


What we do

High-throughput,
cost-effective



Genome-wide
information



From data to
analytics to insights



Flexible end-to-end solution



Evidence of the value

Investors



versionone



Partners



Animal Customers





Enabling advanced genomics at scale

250K

Samples
processed
to date

25

Species
sequenced

4K

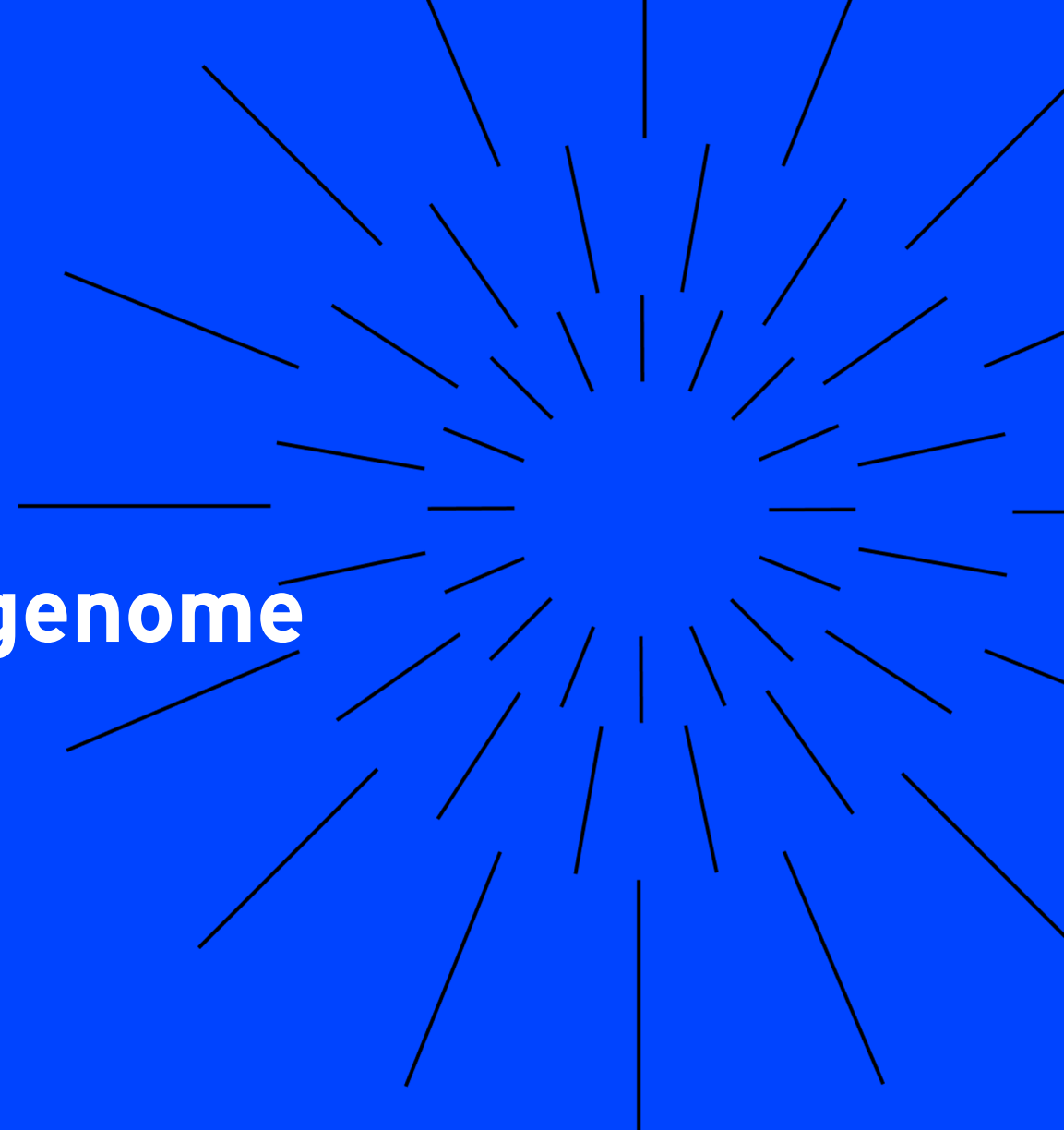
Samples
processed/
day

2W

Weeks to
build a
custom
panel



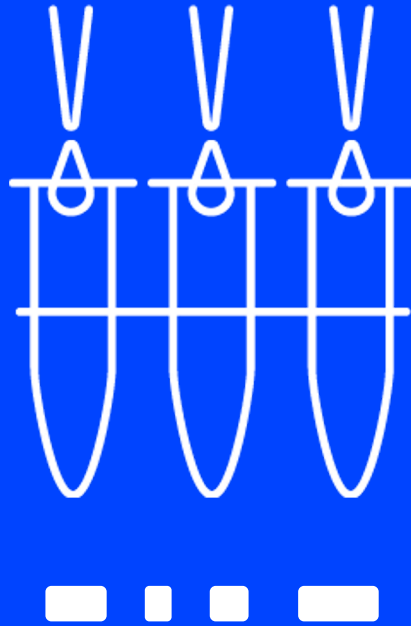
What is low-pass whole genome sequencing?



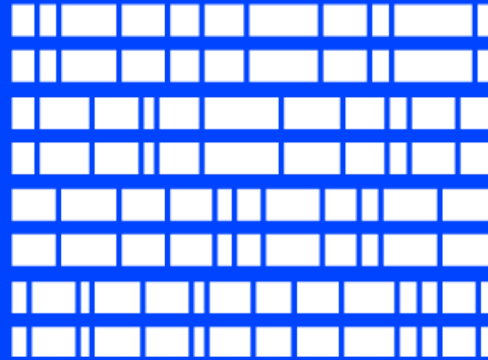


Low-pass whole genome sequencing

Shotgun sequencing on a target individual (typically less than 1x)



Imputation to a haplotype reference panel



Advanced analytics options via Gencove platform





How it helps you

The Gencove platform powers your breeding programs by supporting:



**Trait
Detection**



**Genomic
Selection**



**QTL
Mapping**

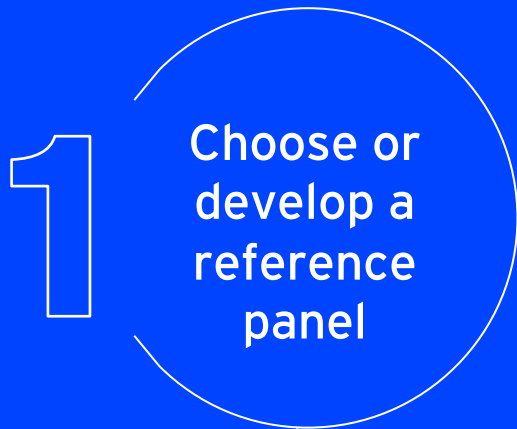


Breed Analysis



How to use the Gencove platform

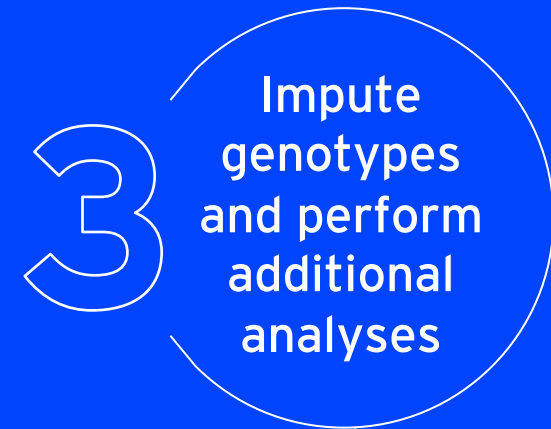
User friendly interface, standard file formats, powerful analyses



Update as needed with new markers or reference genomes



Work with Gencove partner labs or let us help you set up an automated, low-pass optimized sequencing workflow in-house



Receive high-quality genotypes and explore other analysis offerings, such as trait scores, CNV analysis, and microbiome characterization



Advantages of low-pass sequencing and imputation with Gencove

- Whole genome sequencing
- Marker selection flexibility
- Faster, easier setup
- Accelerated turn-around time
- Integrate new markers and reference genomes
- Improved return on investment



The power and flexibility of Gencove software



User friendly,
turnkey solution



Enterprise software



Cloud and API enabled



Utilizes and outputs
standard file formats



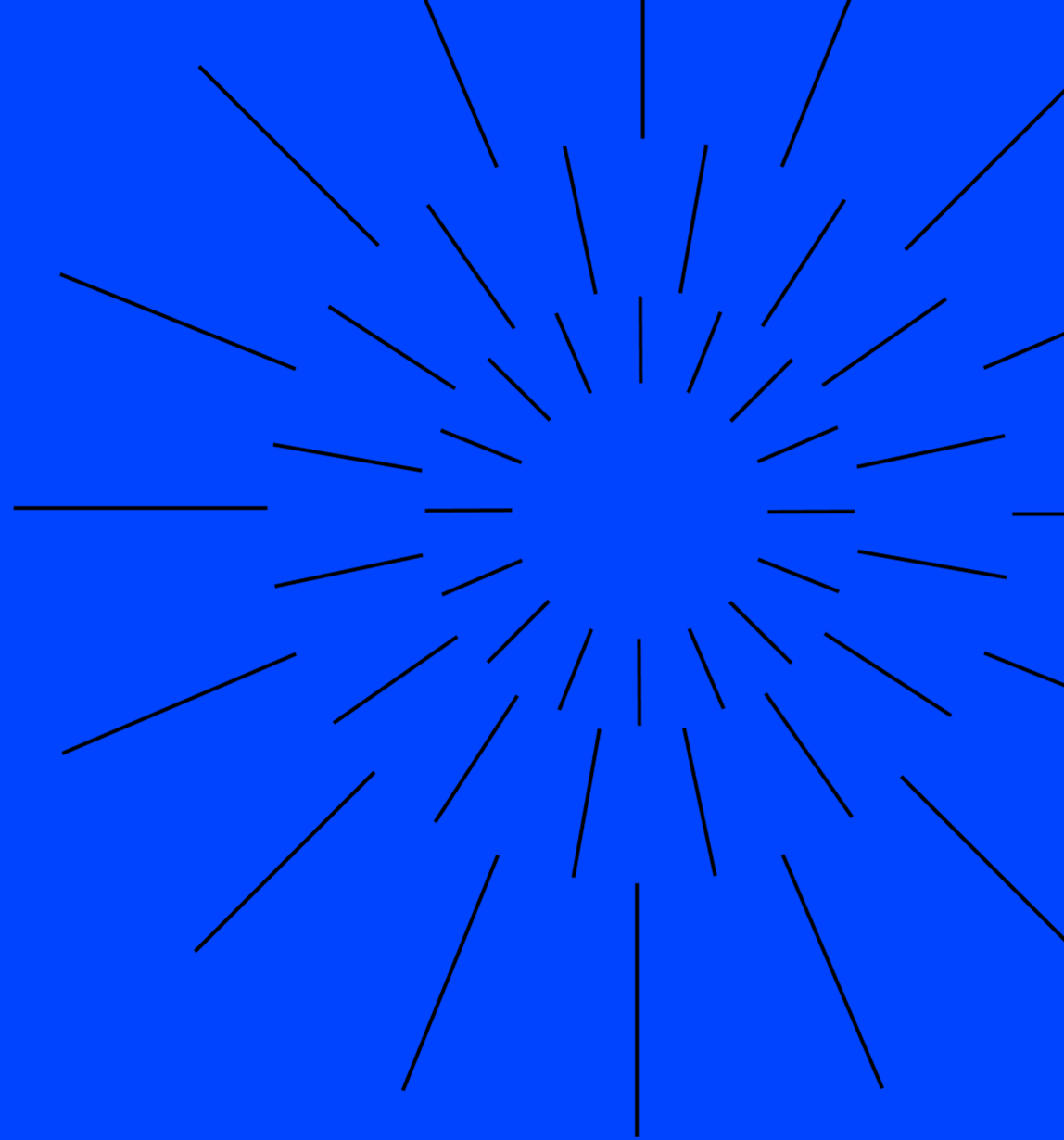
Well documented and version-
controlled



Built by geneticists and optimized
for low-pass sequencing



Comparison to arrays in cattle

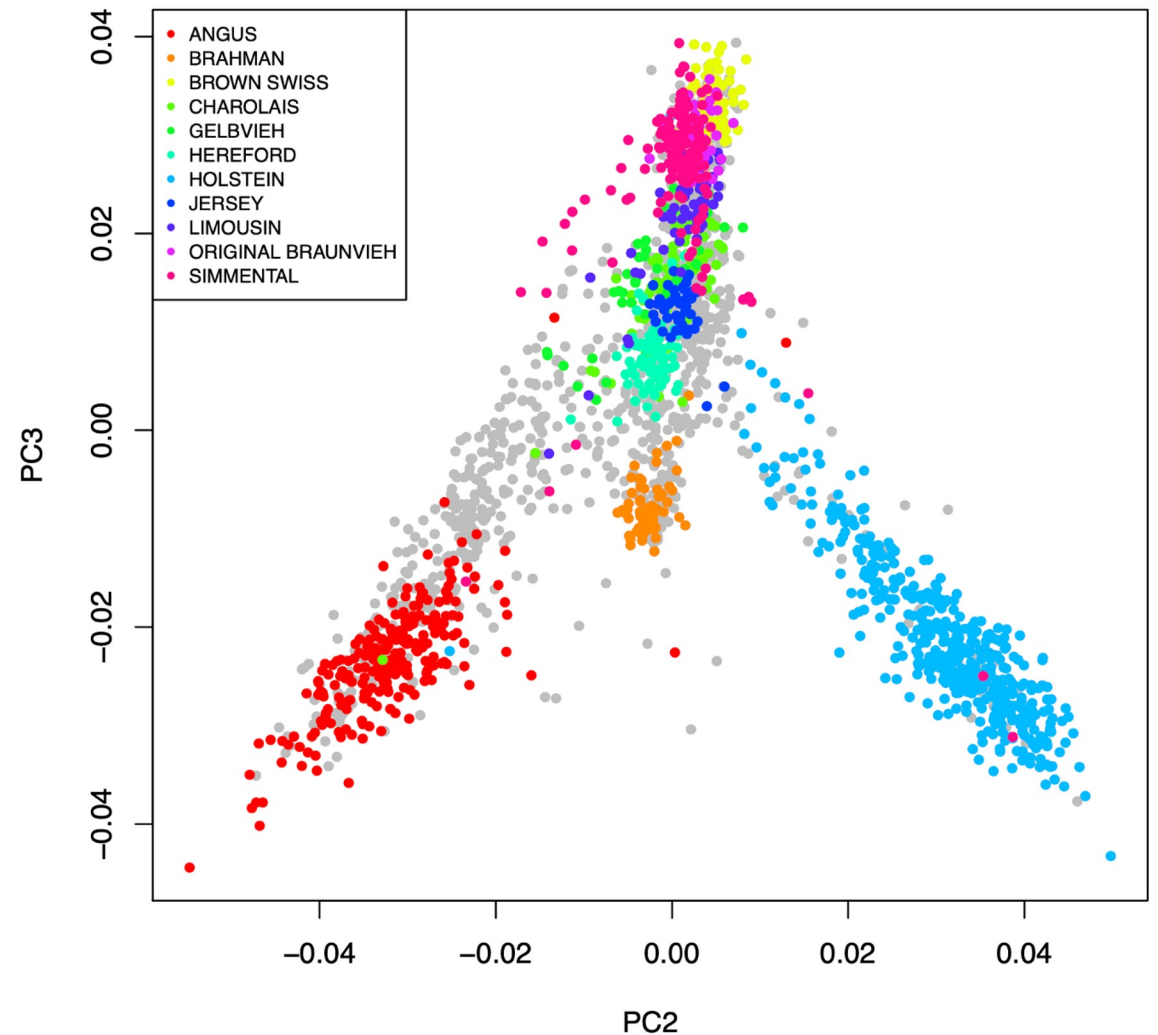




Cattle

Imputation to a haplotype reference panel means:

- 2.2 Million SNPs delivered
- ~786,000 functional variants
- Solution is backwards compatible with over a dozen legacy arrays
- Broad representation of important cattle breeds





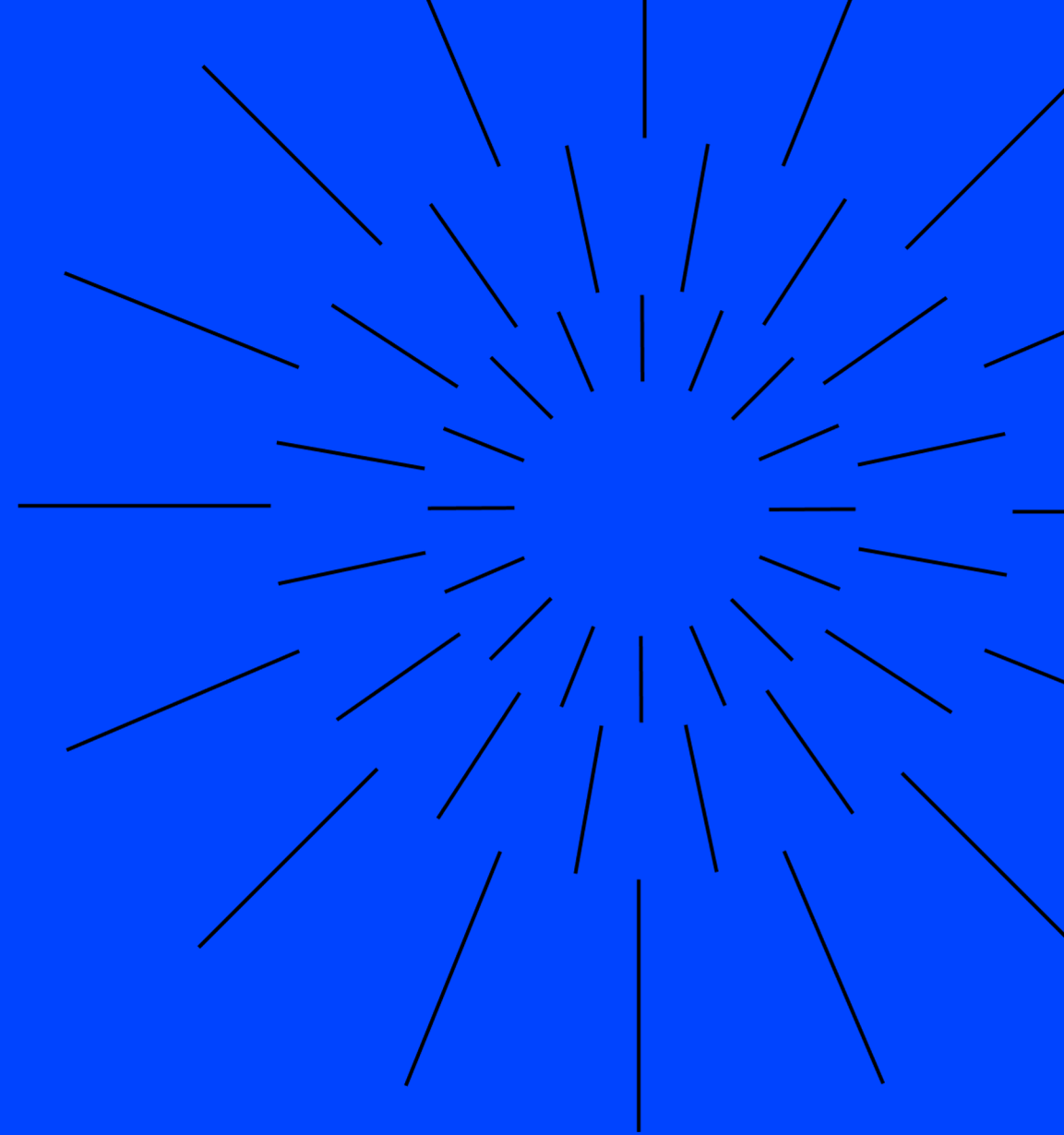
Cattle

Array Concordance		
Angus	Hereford	Simmental
99.3%	99.1%	99.2%

¹ Snelling, W. M. et. al. (2020). Assessment of Imputation from Low-Pass Sequencing to Predict Merit of Beef Steers. *Genes (Basel)*. 2020 Nov; 11(11): 1312

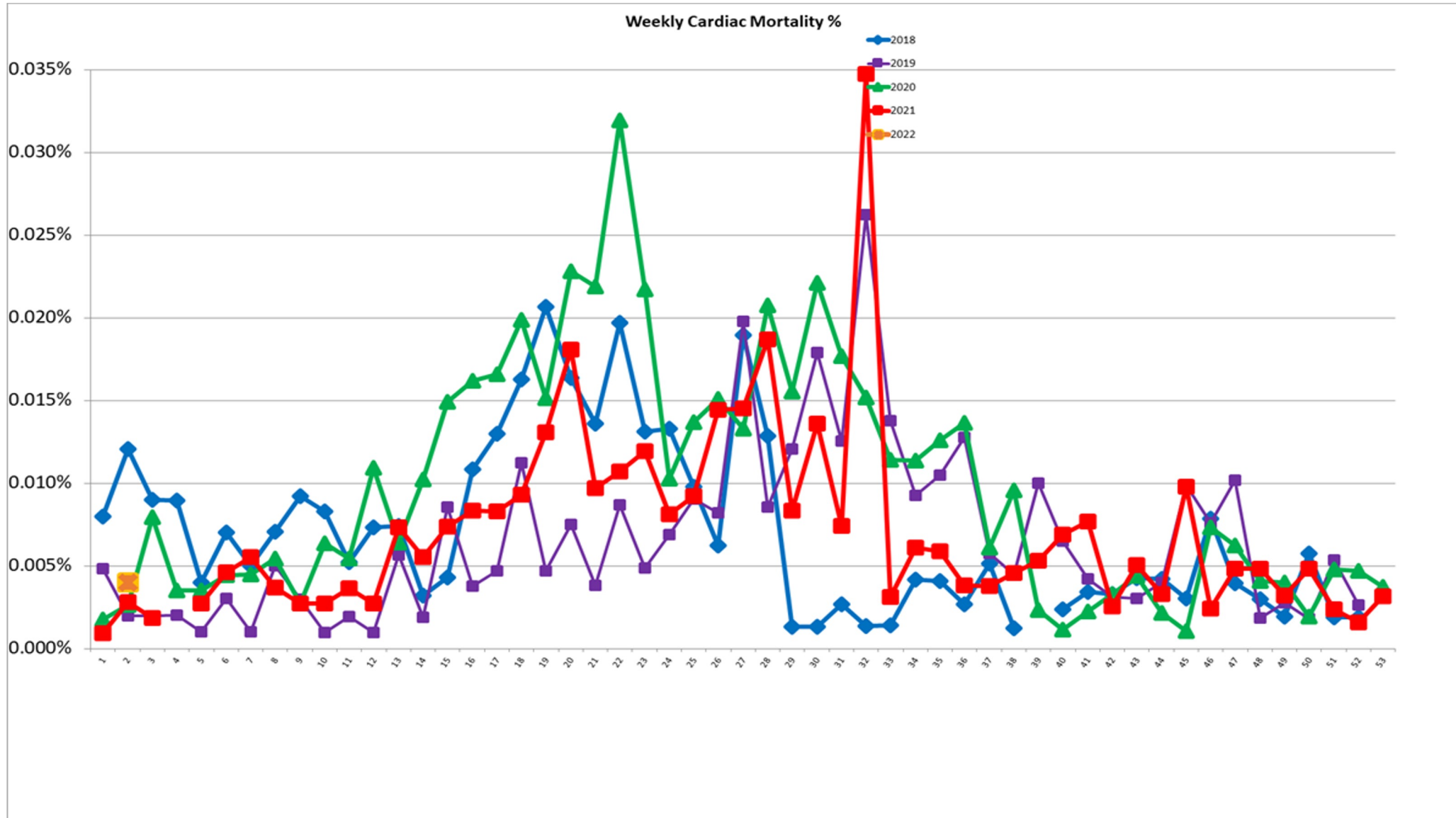


An EPD for Bovine Congestive Heart Failure



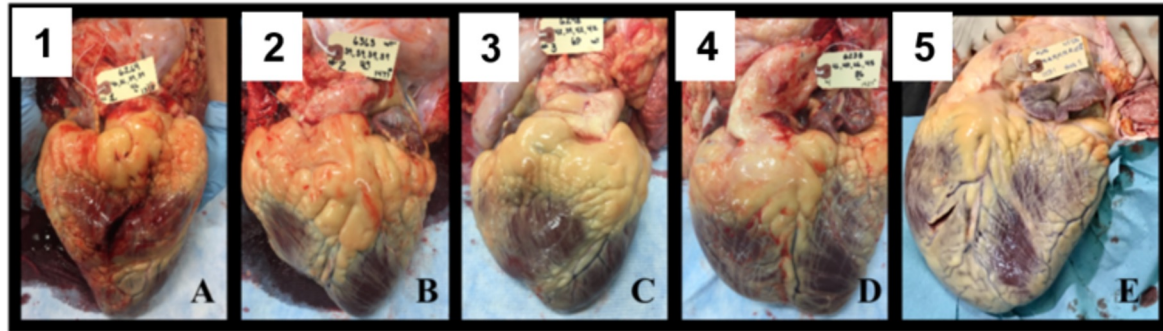


BCHF - a considerable cause of mortality





BCHF phenotype observation & selection for genotyping



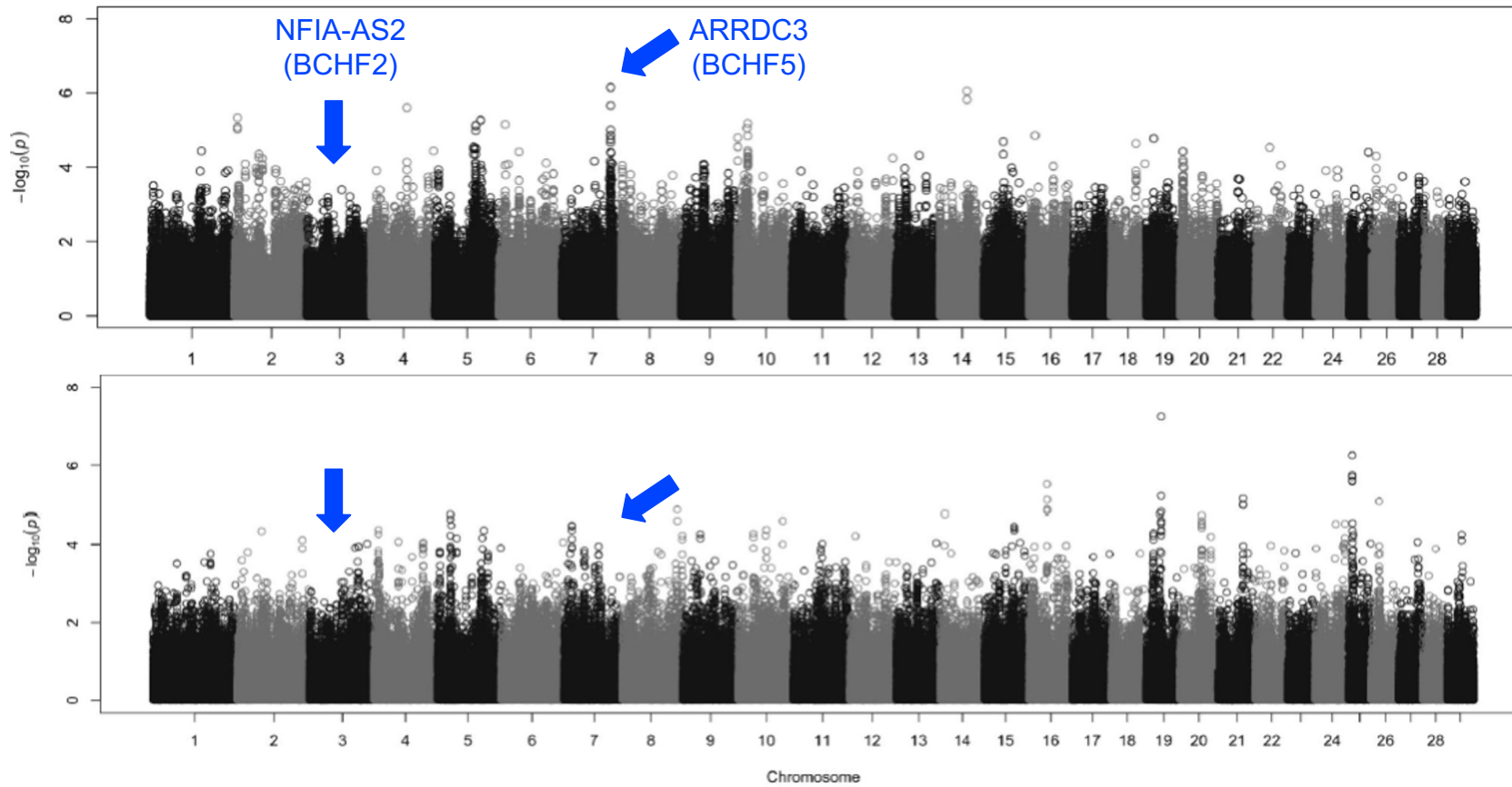
BCHF Score	Count
1 or 2	4,446
4 or 5	555

Selection of individuals for genotyping

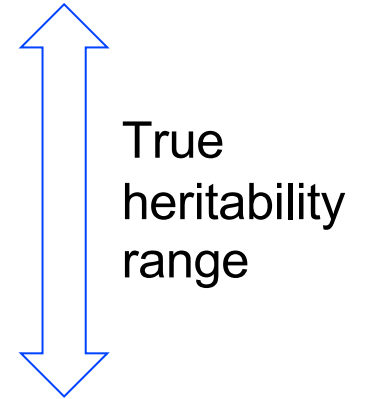
- Total of 19,439 individuals evaluated and categorized into 5 heart scores
- 5,001 contemporary grouped individuals genotyped with Gencove platform with ~1 million markers used for analyses



Sequencing and genetic evaluation



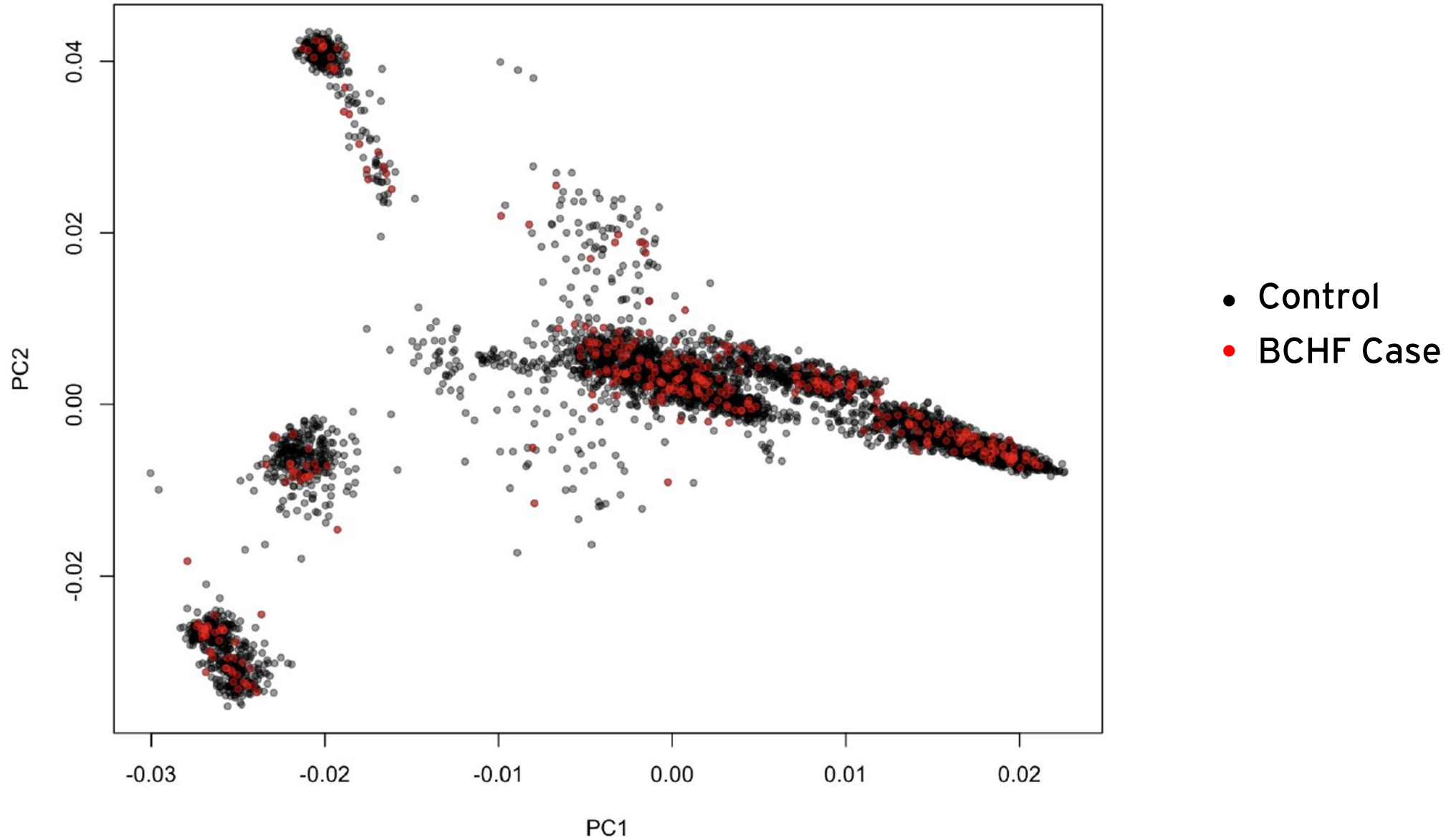
No Contemp Grp
chip $h^2 = 0.423$
case/ctl = 373/4266



Contemp Grp
chip $h^2 = 0.306$
case/ctl = 187/4173



Sequencing and genetic evaluation





Genetic correlations to carcass and growth

Carcass and growth phenotypes captured on individuals with heart score phenotypes

BCHF Genetic Correlation	
HCW	0.399
BACKFAT	-0.120
REA	0.065
MARB	-0.112
ADG	0.266
DMI	0.238

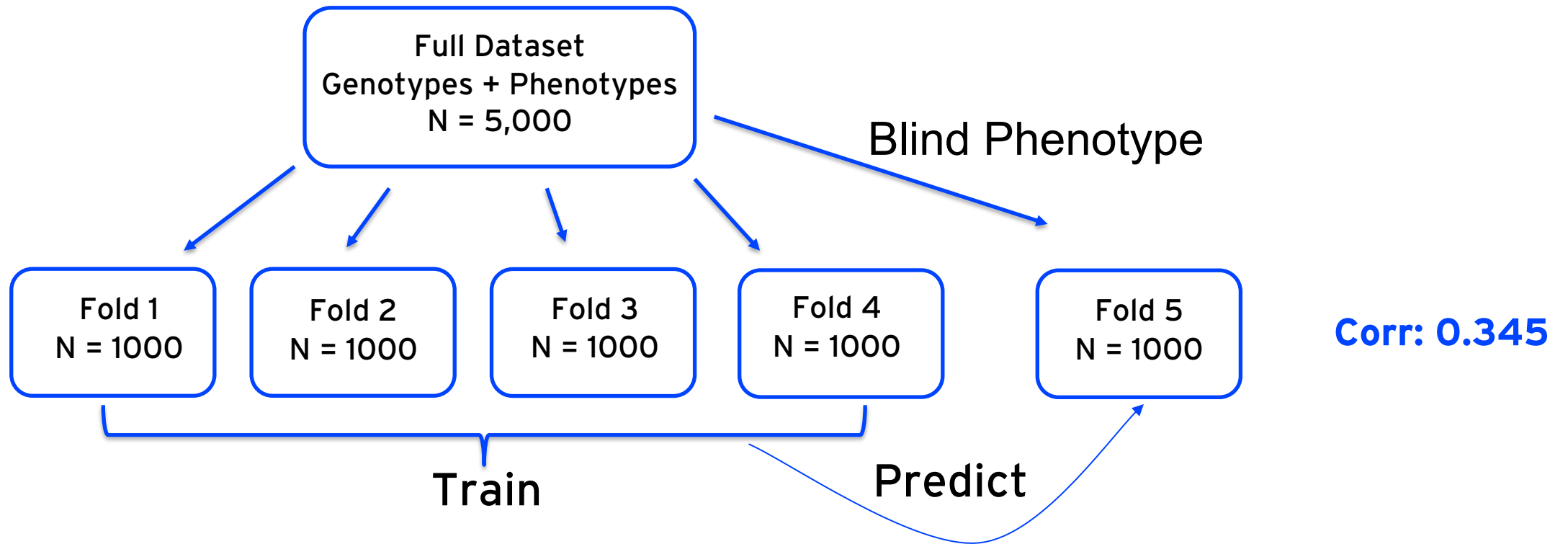
Trait	Heritability
BCHF	0.343
HCW	0.391
BACKFAT	0.297
REA	0.295
MARB	0.373
ADG	0.216
DMI	0.304

THRGIBBSF90 threshold-linear model, 100K genotype density, % Angus covariate



Accuracy for potential BCHF EPD

Blinded 5-fold prediction accuracy: $\frac{0.345}{\sqrt{0.3}} = \mathbf{0.63}$ at $h^2 = 0.3$



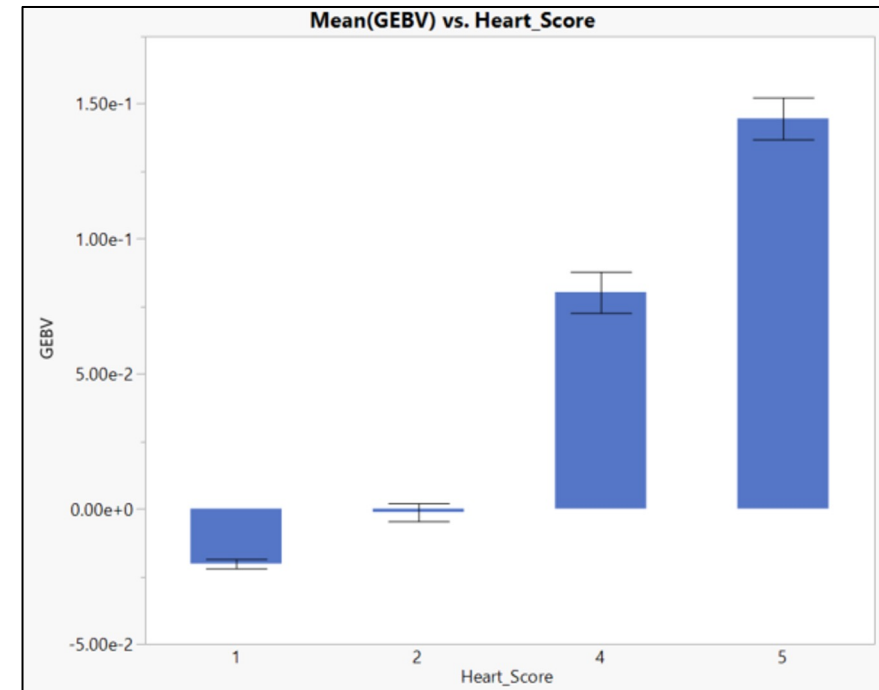
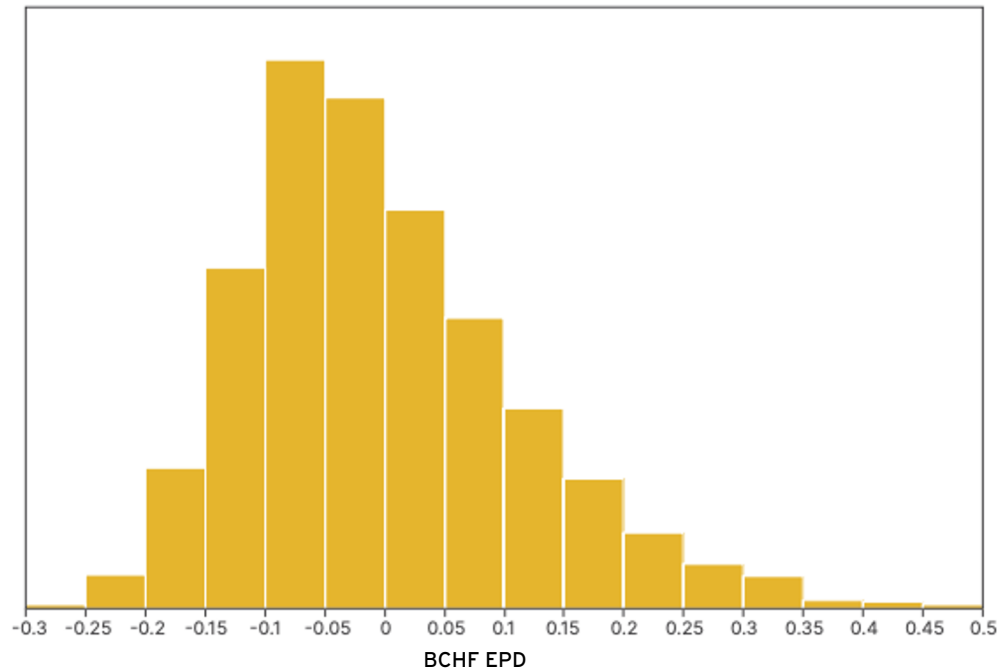
Significant improvement in association with addition of 182 cases: 0.52 to 0.63



BCHF EPD as a breeding tool

Goal: an EPD that predicts risk of BCHF in progeny of tested Sire or Dam

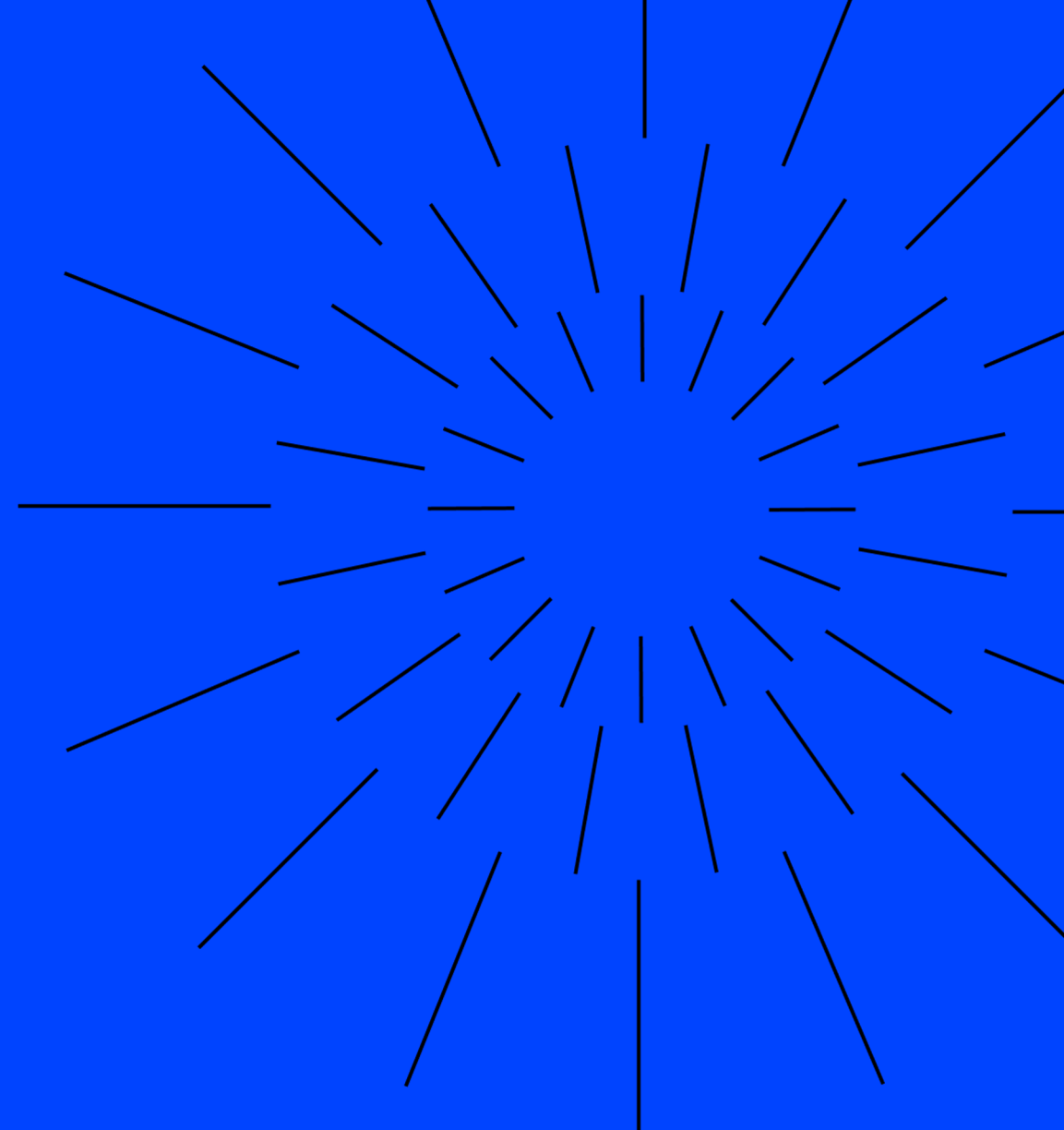
BCHF EPD Distribution



2022 Phenotyping: 1,100 so far



Low-pass + target capture with InifiniSEEK





Low-pass + target capture

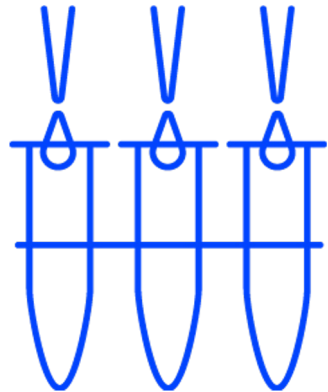
A revolutionary, cost-effective solution for low-pass whole genome sequencing and targeted high-coverage information **in one solution**

- Highly accurate genome wide information from a single technology
- Reliable, simple process easily integrates into your existing workflows

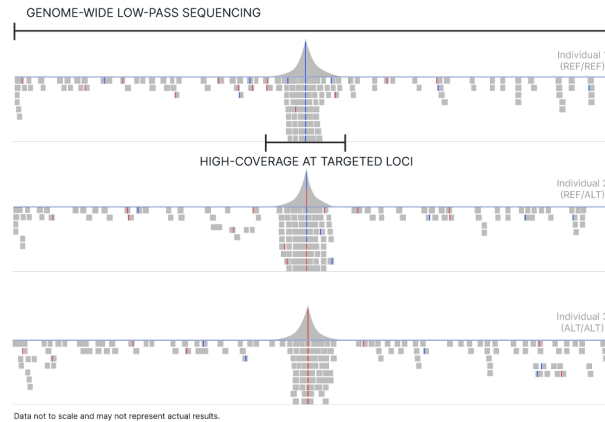


Single, reliable molecular reaction

High- throughput end-to-end reaction

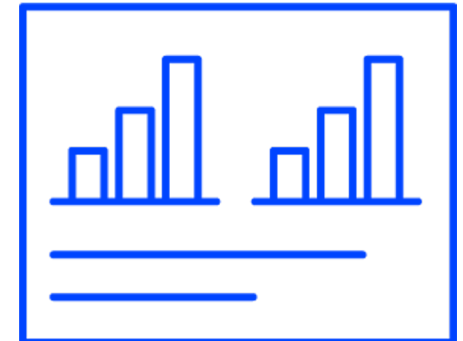


Low-pass whole genome sequencing and target region coverage



+

Imputation and advanced analysis





The power of high coverage sequencing

- **Direct observation** of the known and unknown genetic variation
- **More accurate** than existing genotype tools
- **Flexibility for further analysis** without added cost or complexity



Beyond 99%

- Designed hybrid capture probes at over 800 locations
- Validated in ~900 samples with known carriers of key alleles with chip genotypes
- Successfully designed over 99% of capture targets in v1
 - Proprietary diagnostics
 - Parentage markers





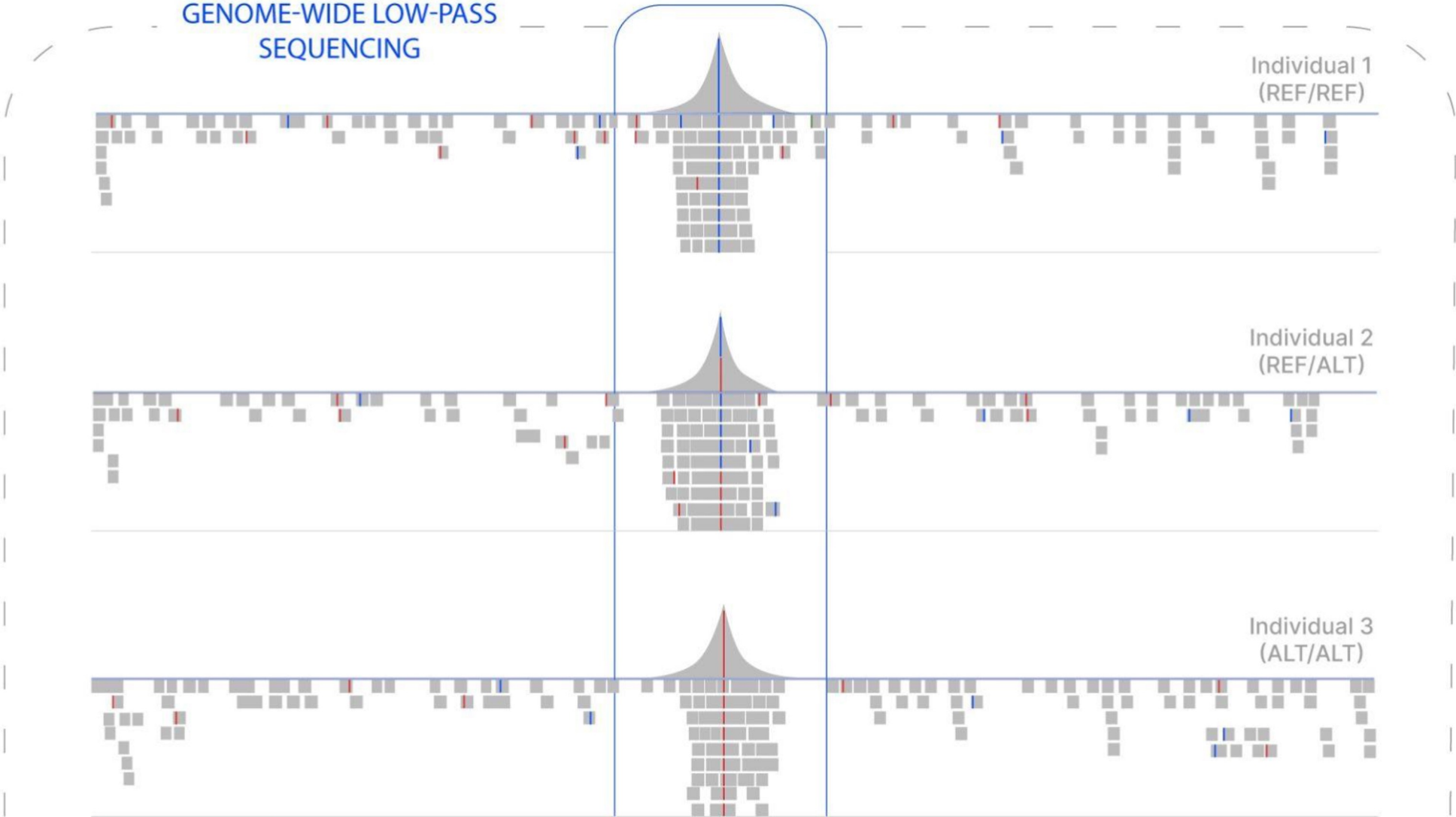
Results in first ~900 animals

Locus	Genotype Match
Polled	100%
MC1R	100%
A2	100%
Calpain	100%



Example InfiniSEEK data (aligned BAM)

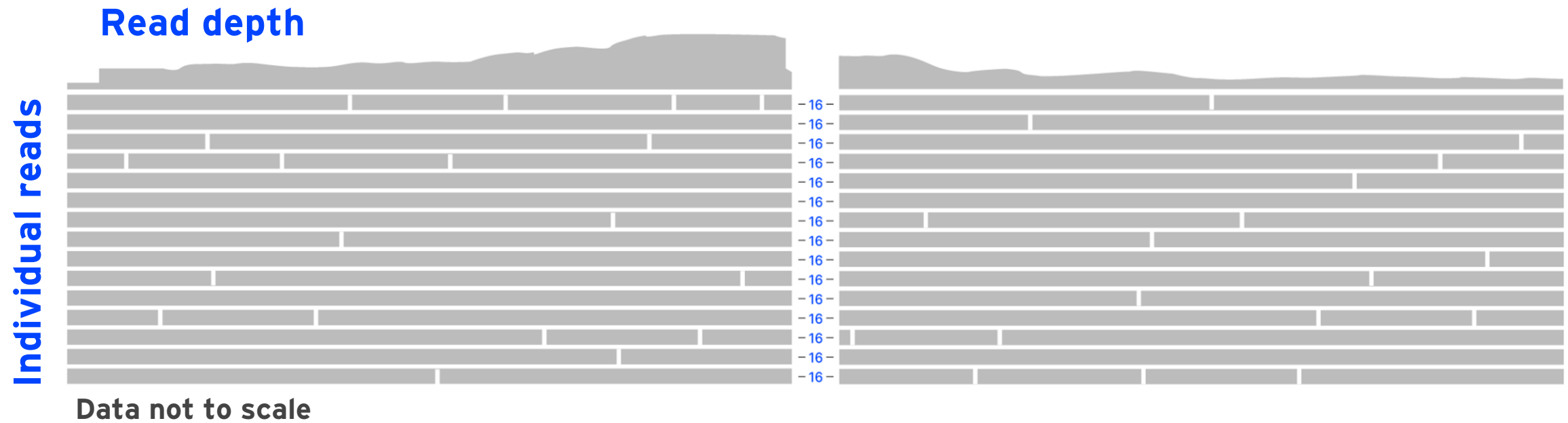
SIMULTANEOUS HIGH COVERAGE AT TARGETED LOCI



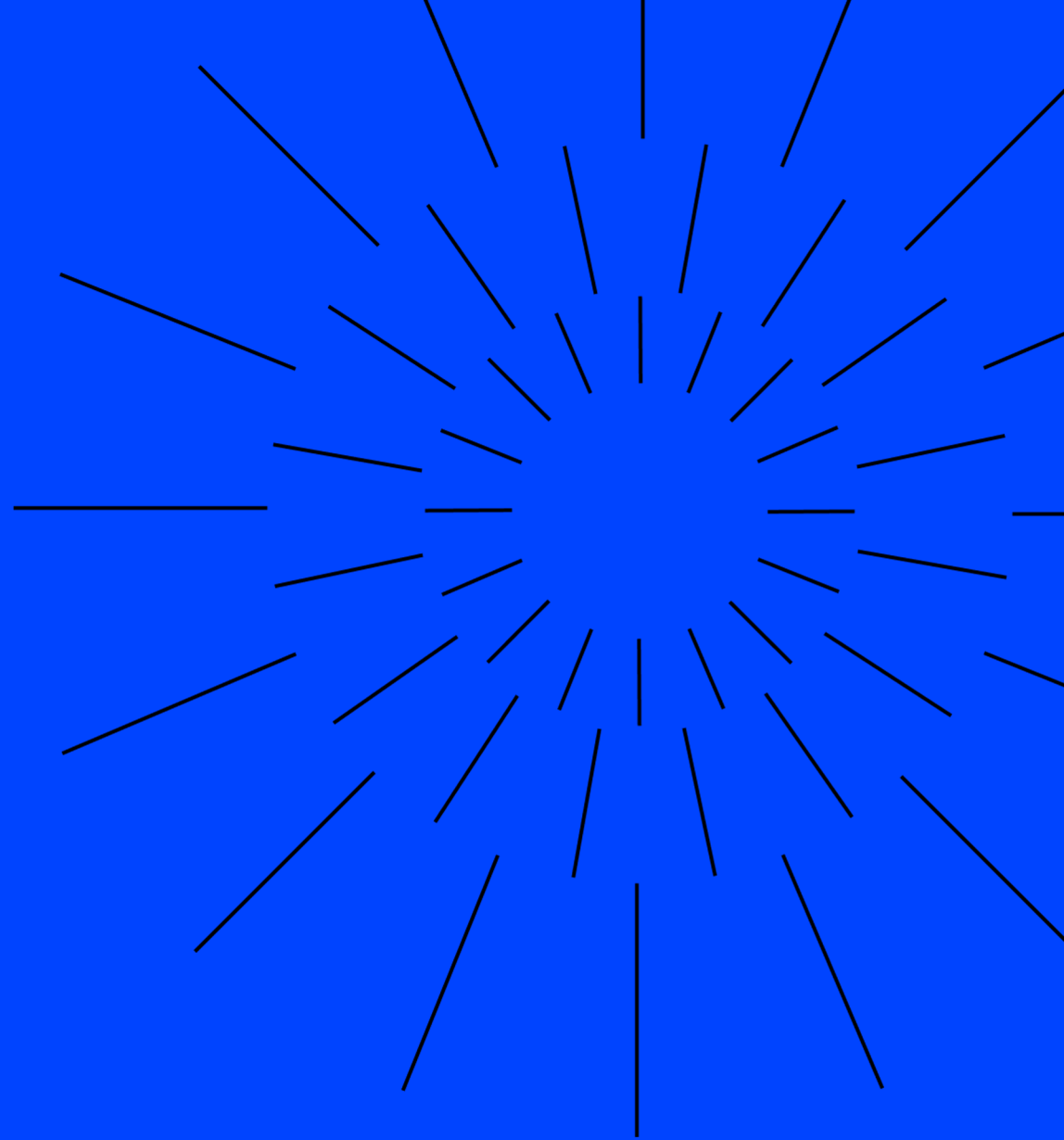
Data not to scale



Polled example - deletion



Summary





Gencove delivers



1

High quality data

Access to high-volume, cost-effective, whole genome sequencing for a wide variety of plant/animal applications

2

Actionable analytics

An easy to implement platform for genotyping and advanced genomic analyses without building another array

3

New insights

More data to enable identification of novel genetic associations to key traits



Thank you!

