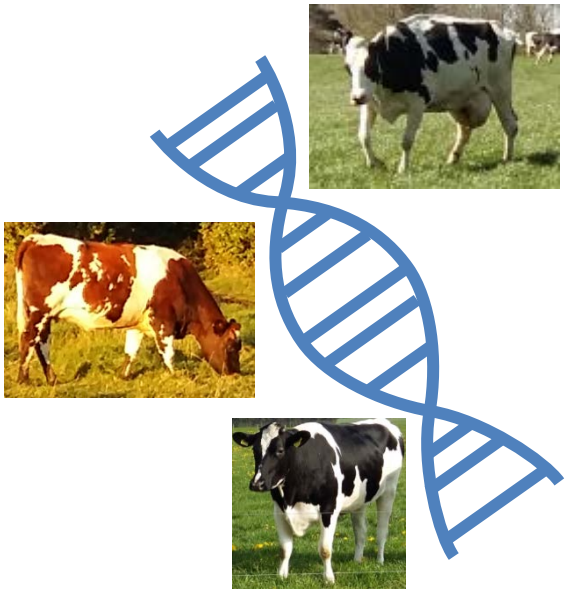
Detection of large chromosomal deletions in cattle and their phenotypic consequences

Md Mesbah Uddin

PhD student



AARHUS UNIVERSITY

GenSAP

EGS ABG

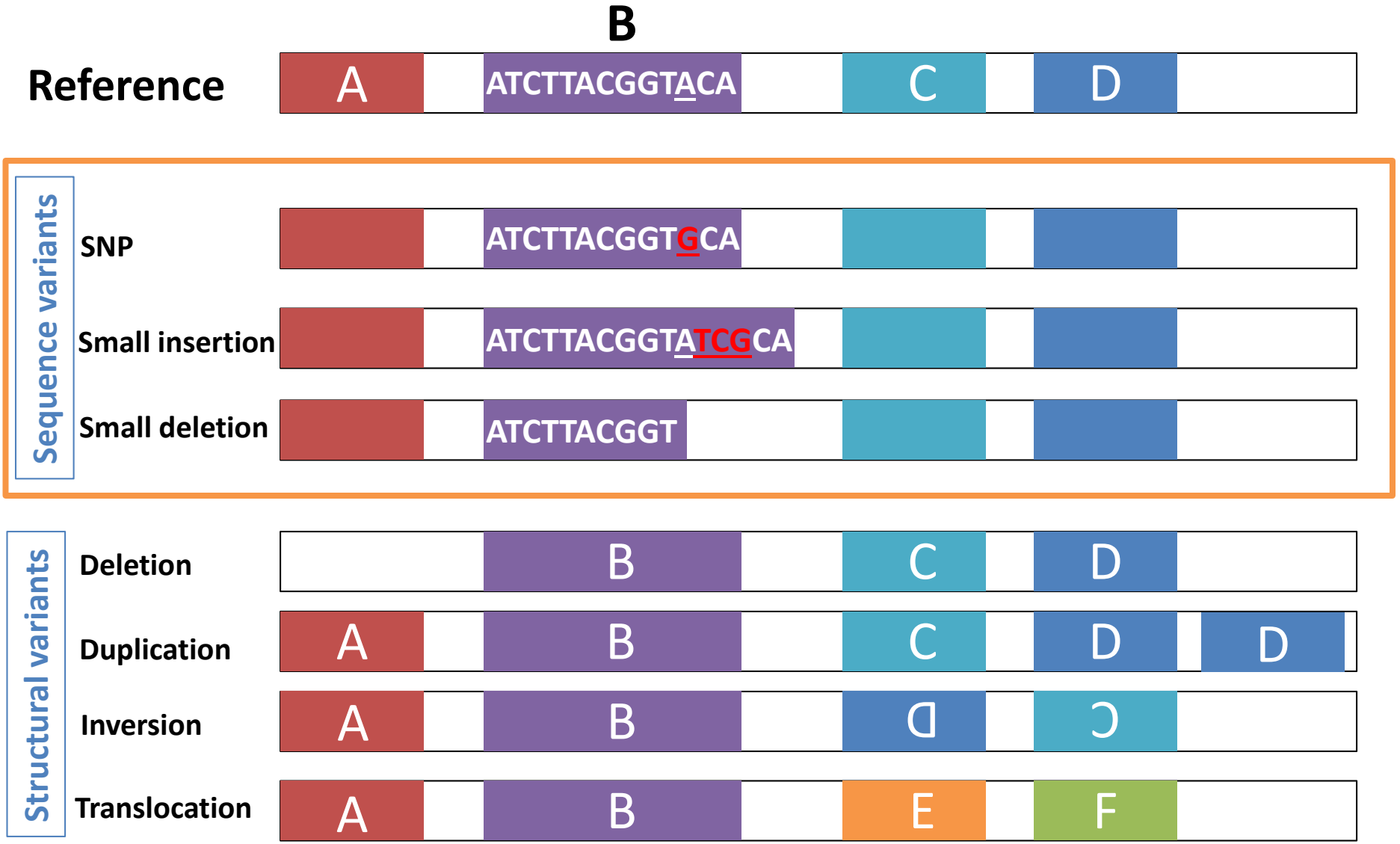
Erasmus Mundus

AgroParisTech

Outline

- Introduction
- Deletion discovery and validation
- Building reference for imputing deletions
  - Extending reference panel using 1KBGP data
  - Imputation accuracy
- Functional impact
- Future prospect
- Conclusion

Introduction



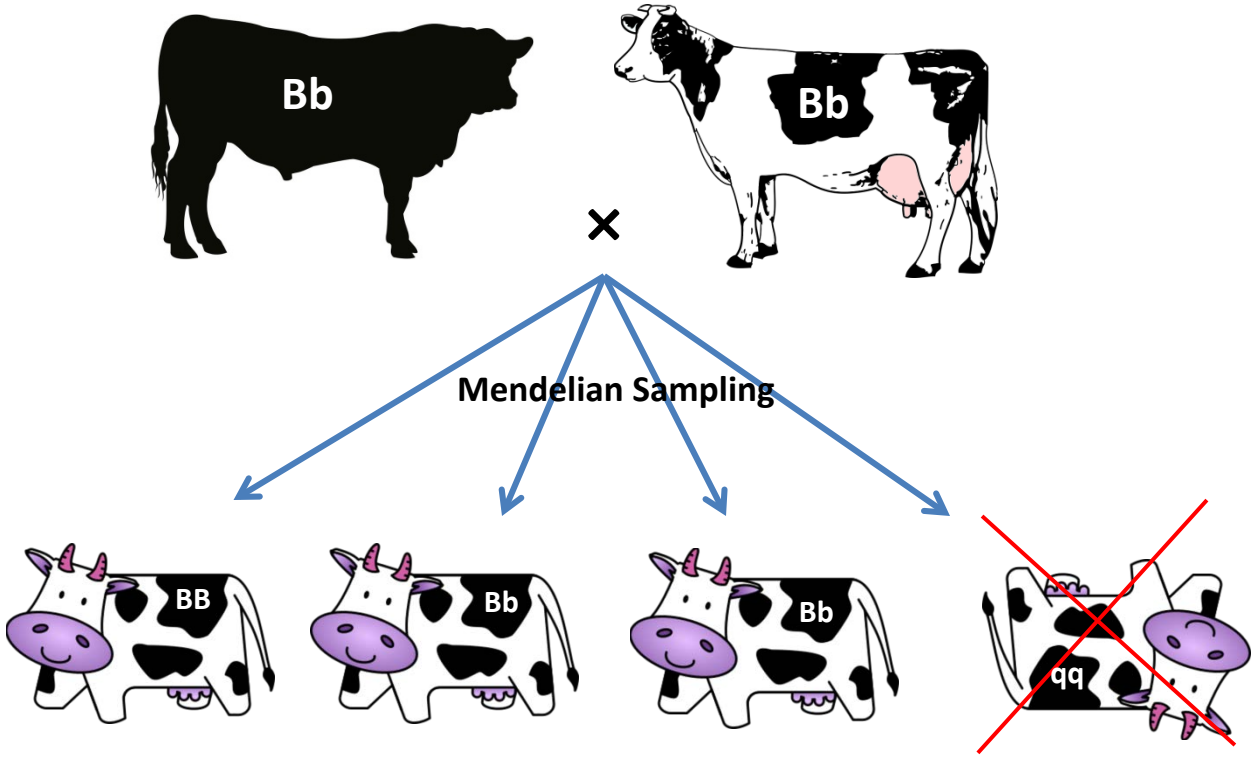
Introduction

- Structural variants: DNA alternations >50bp
  - **least explored polymorphisms** in cattle
  - Phenotypic impact ranges from **beneficial** to **lethal** both in humans [1,2] and animals [3]

1. Zarrei, M., et al. (2015). "A copy number variation map of the human genome." Nat Rev Genet 16(3): 172-183.
2. Weischenfeldt, J., et al. (2013). "Phenotypic impact of genomic structural variation: insights from and for human disease." Nat Rev Genet 14(2): 125-138.
3. Bickhart, D. M. and G. E. Liu (2014). "The challenges and importance of structural variation detection in livestock." Front Genet 5: 37.

Introduction

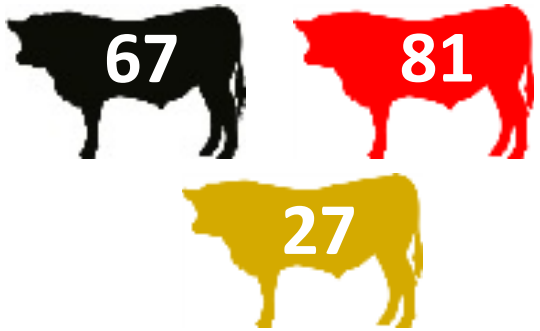
- **Deletions** are potential **candidate for loss-of-function**, which could be **lethal as homozygote**



Discovery of large deletions

- **Objective:** To map large chromosomal deletions

- **Discovery and genotyping**
  - **Samples:** **WGSs of 175 animals**



  - **GenomeSTRiP-2 [1,2] –deletion discovery and genotyping**

1. Handsaker, et al. 2011, Discovery and genotyping of genome structural polymorphism by sequencing on a population scale. Nat Genet, 43, 269-276.
2. Handsaker, et al. 2015, Large multiallelic copy number variations in humans. Nat Genet, 47, 296-303.

Validation of Deletion call

- BovineHD Chip-intensity
- Breakpoint-assembly
- PCR+Sequencing

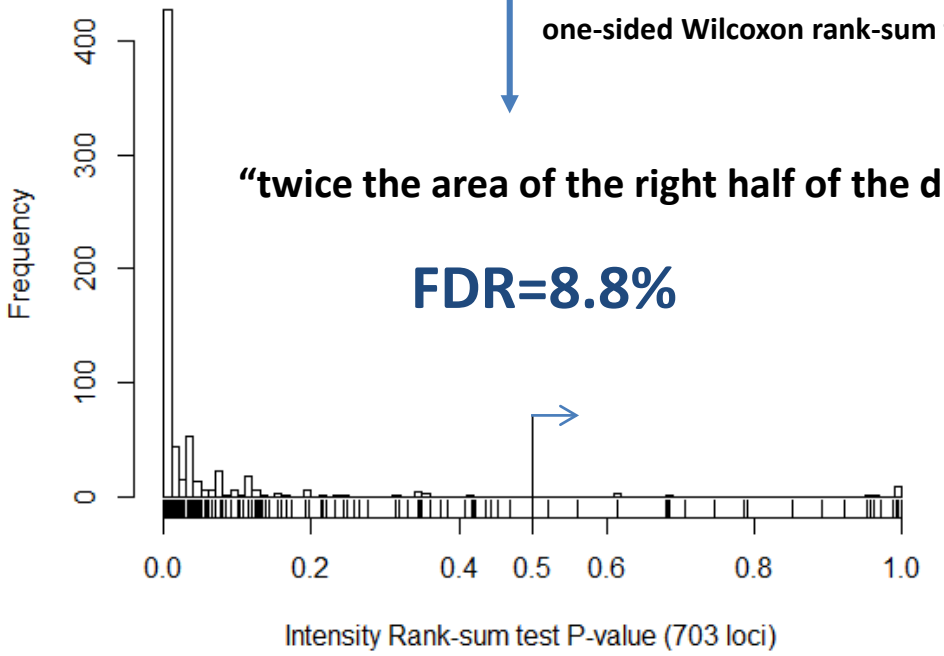
777K BovineHD chip-intensity

Intensity DEL (HET/HOM) < **Intensity REF (HOM)**

one-sided Wilcoxon rank-sum test

"twice the area of the right half of the distribution" [1]

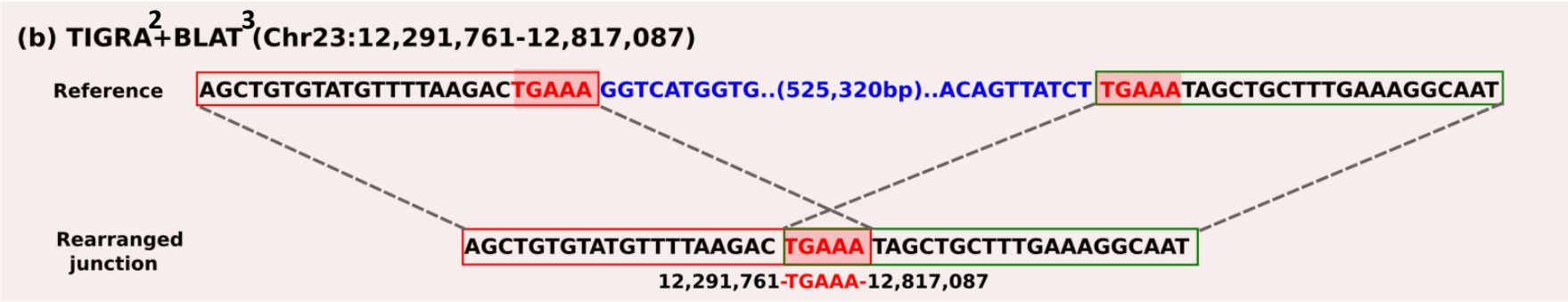
FDR=8.8%



1. Handsaker, et al. (2011). Nat Genet, 43, 269-276.
Mesbah-Uddin, M. et al. (2017). DNA Res. doi:10.1093/dnares/dsx037

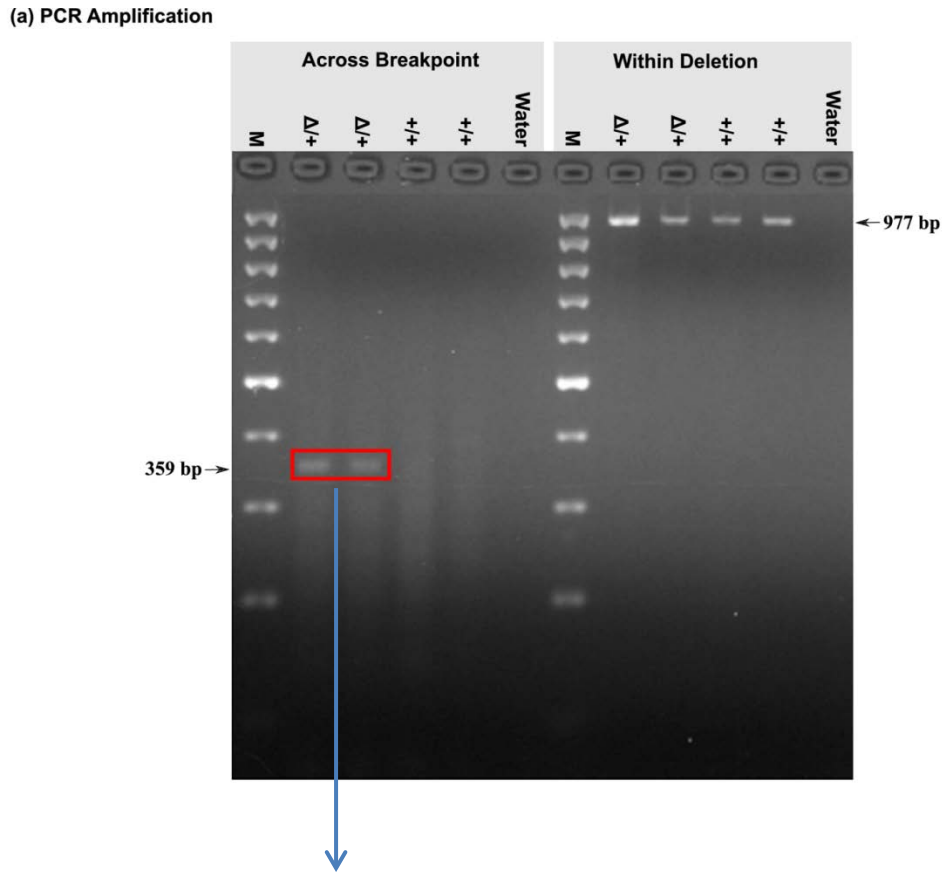
Targeted breakpoint assembly and alignment

- **30** randomly chosen deletions
  - successfully resolved breakpoints of **26** deletions (~**87%** success rate)
- **identified** breakpoint of a **~525 KB deletion**, causing stillbirth[1] in dairy cattle



1. Sahana, et al. (2016). A 0.5-Mbp deletion on bovine chromosome 23 is a strong candidate for stillbirth in Nordic Red cattle. Genet Sel Evol, 48, 35.
2. Chen, et al. 2014, TIGRA: a targeted iterative graph routing assembler for breakpoint assembly. Genome Res, 24, 310-317.
3. Kent, 2002, BLAT--the BLAST-like alignment tool. Genome Res, 12, 656-664.

PCR & amplicon sequencing



Mesbah-Uddin, M. et al. (2017). DNA Res. doi:10.1093/dnares/dsx037

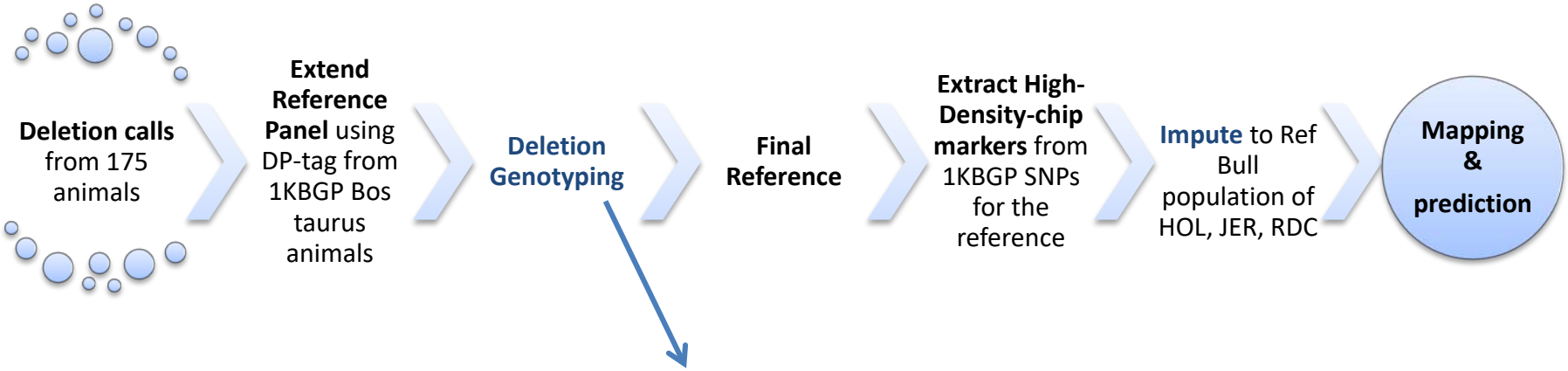
Final deletion-set

- **8,480 large deletions (199bp to 773KB)**
  - Size: mean=4.5KB and median=1KB
  - overall false discovery rate of **8.8%**
  - **82%** are **novel** compared with deletions in the *dbVar* database (last accessed on 27 January 2017)

Mesbah-Uddin, M., et al. (2017). **Genome-wide mapping of large deletions and their population-genetic properties in dairy cattle.** *DNA Res*. doi:10.1093/dnares/dsx037

Building reference for imputation of large deletions

Deletion genotyping using read-depth (RD)

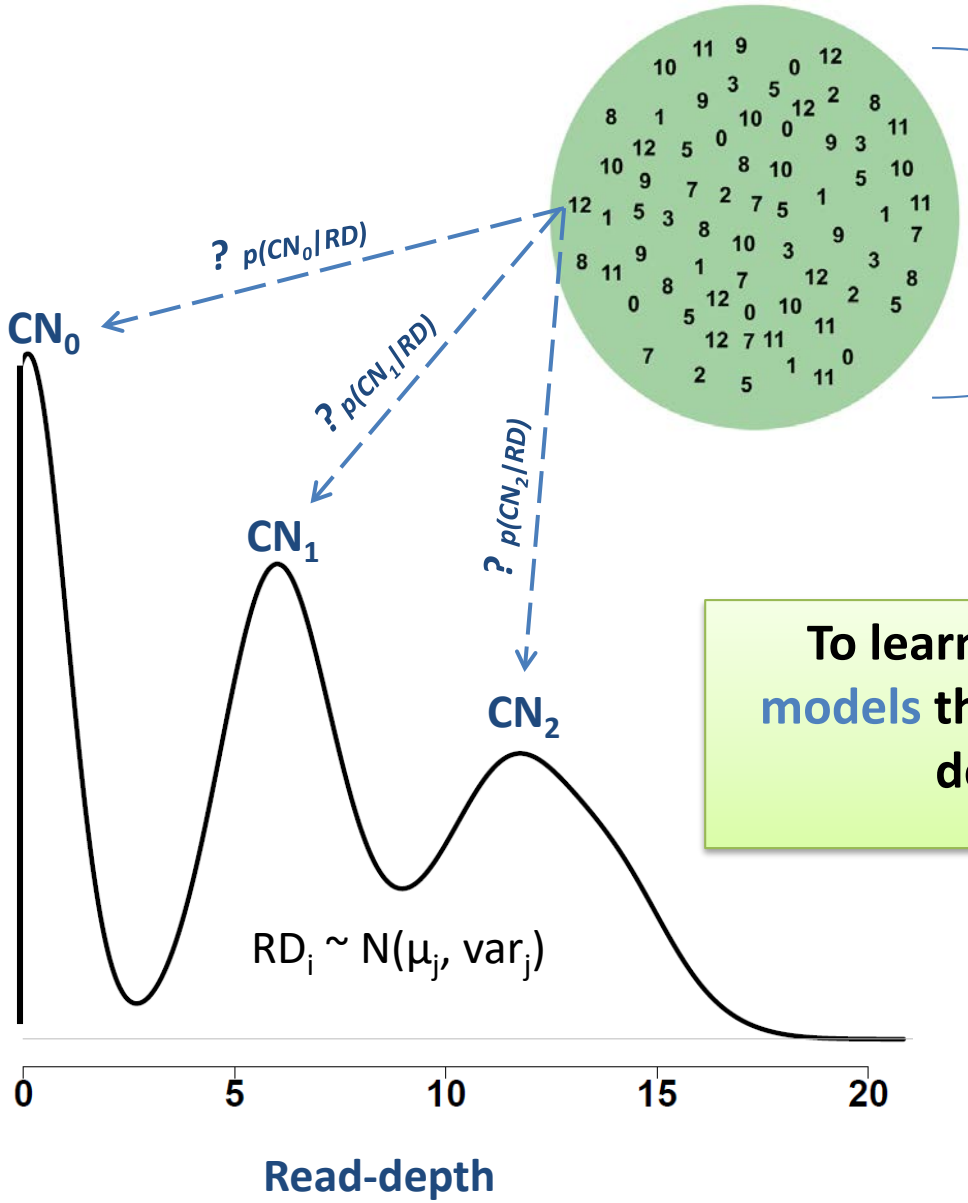


(1,2)

$$\#AlignedReads_{Position} \propto CopyNumber_{Position}$$

$$E(CopyNumber_{Region}) = \frac{Avg_{ReadDepth\_Region}}{(GC-adjusted)\ Avg_{GenomeWideReadDepth}} \times 2$$

1. Yoon, et al. 2009, Sensitive and accurate detection of copy number variants using read depth of coverage. Genome Res, 19, 1586-1592.
2. Miller, et al. 2011, ReadDepth: a parallel R package for detecting copy number alterations from short sequencing reads. PLoS One, 6, e16327.



Read-depth (RD) data from SNPs within a Deletion:
RD = {x_1, x_2, x_3 …,x_n}

To learn most likely combination of models that explain the observed read-depth within a deletion

EM algorithm

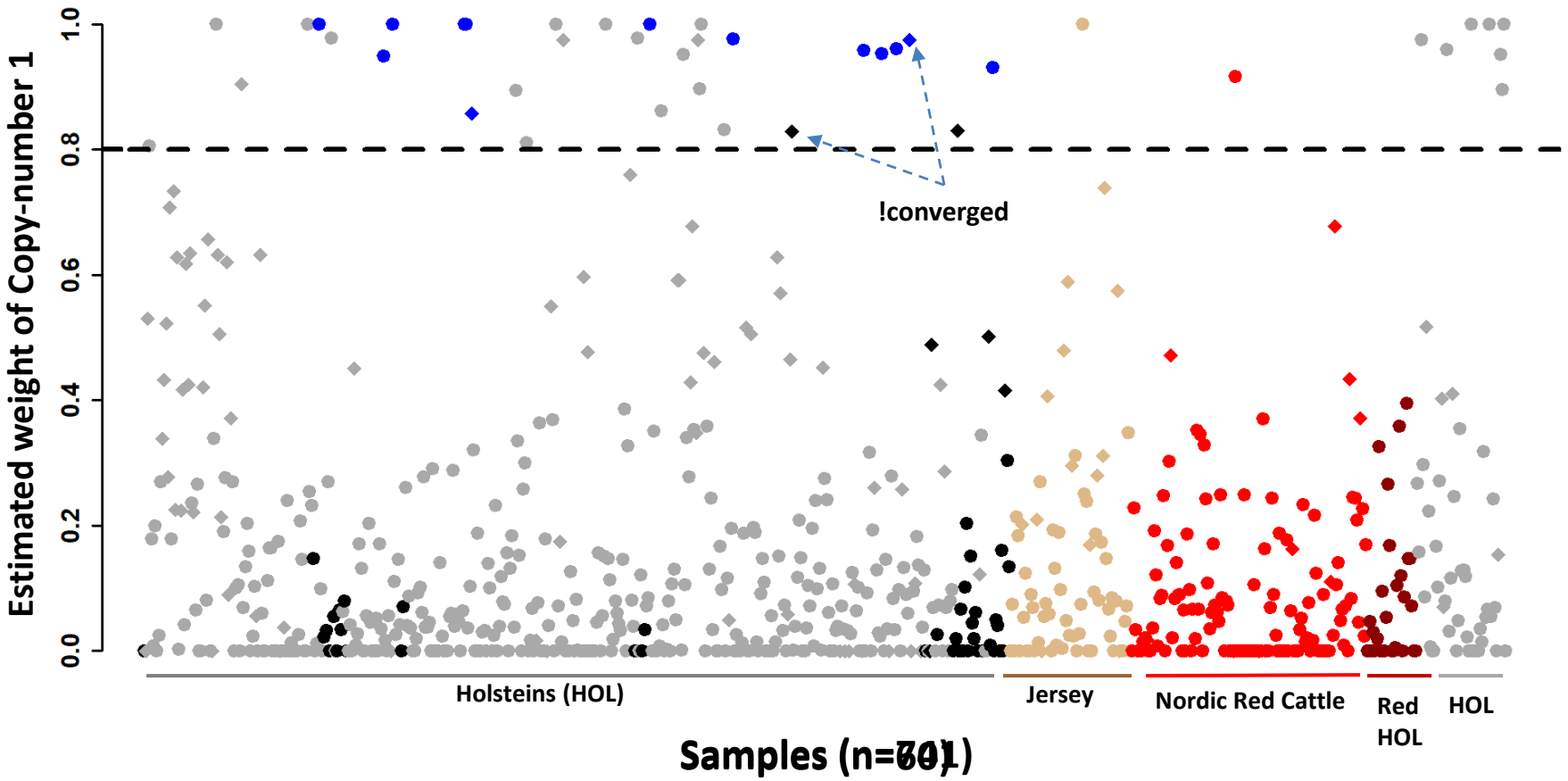
Validation: ~3.3KB deletion on Chromosome 21 (causes Brachyspina[1])

- Total **60 Holsteins** (**13 carriers + 47 non-carriers**) from 1KBGP
  - **55 progeny tested Holstein bulls** with recessive codes **(8 BY + 47 TY)***
  - **5** carriers in deletion-discovery set (confirmed by targeted breakpoint-sequence assembly)

[1] Charlier, C., et al. (2012). A deletion in the bovine FANCI gene compromises fertility by causing fetal death and brachyspina. *PLoS One, 7*(8), e43085

* source: The US Council on Dairy Cattle Breeding database (https://www.uscdcb.com/CF-queries/index.cfm, last accessed on 25 September, 2017)

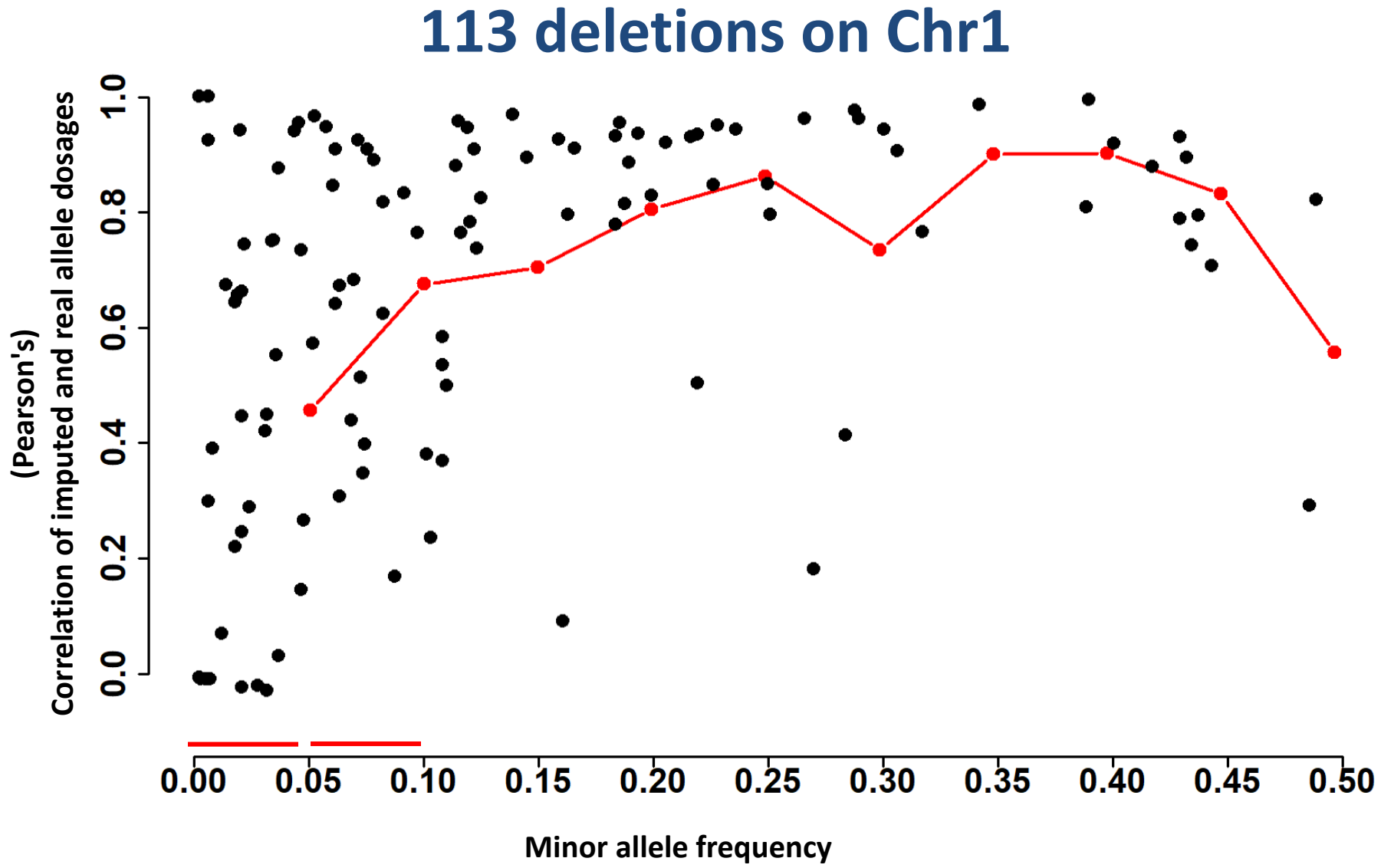
~3.3 KB deletion on Chromosome 21 (causes Brachyspina)



Imputation accuracy

- **leave-out trials: 772 animals** (175 discovery samples + 597 1KBGP)
  - Masked genotype: **25** samples
  - **31** trials (22 samples in last trial)

Imputation accuracy



Functional impact

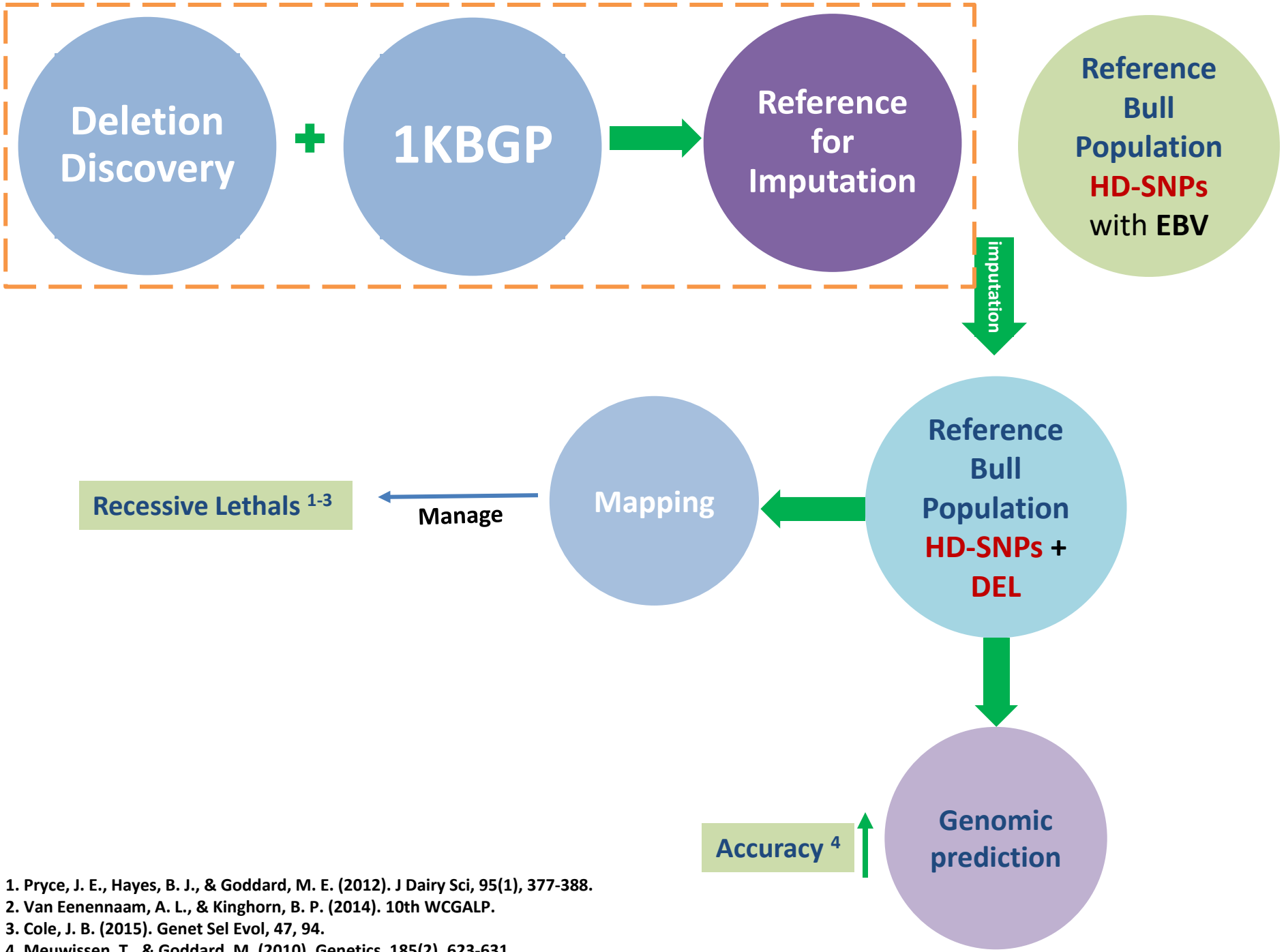
Enrichment of QTL on deletions

Trait Classes[¥]	Fold Enrichment	P value[*]
Health	**2**	8.91×10^{-10}
Reproduction	**1.5**	7.4×10^{-11}
Milk	0.8	2.45×10^{-7}
Exterior	0.5	1.85×10^{-4}
Production	0.5	0.002
Meat and Carcass	0.5	0.058

[¥]Trait classes are from cattleQTLdb. [*]Fisher's exact test (two-sided). QTL from autosomes of Holsteins, Jersey, Nordic Red Cattle, and Ayrshires were considered.

Mesbah-Uddin, M. et al. (2017). DNA Res. doi:10.1093/dnares/dsx037

Future prospect of large deletions



1. Pryce, J. E., Hayes, B. J., & Goddard, M. E. (2012). J Dairy Sci, 95(1), 377-388.
2. Van Eenennaam, A. L., & Kinghorn, B. P. (2014). 10th WCGALP.
3. Cole, J. B. (2015). Genet Sel Evol, 47, 94.
4. Meuwissen, T., & Goddard, M. (2010). Genetics, 185(2), 623-631.

Conclusions

- **Read-depth (DP-tag) from 1KBGP data** could be used for extending **deletion reference panel**
- **Common deletions could be imputed with high accurately**
- **Enrichment of QTL on Deletions** – potential for inclusion in genomic studies

Thanks to All

Conclusions

- **Read-depth (DP-tag) from 1KBGP data** could be used for extending **deletion reference panel**
- **Common deletions could be imputed with high accurately**
- **Enrichment of QTL on Deletions** – potential for inclusion in genomic studies