



Berner Fachhochschule  
Haute école spécialisée bernoise  
Bern University of Applied Sciences

QUALITAS<sup>+</sup>



# Accurate sequence imputation enables precise QTL mapping in Brown Swiss cattle

Mirjam Frischknecht

29.08.2016

EAAP 2016 Belfast

# Imputation to sequence data

Test panel: SNP chip data

1	?	?	?	1	?	1	?	0	2	2	?	?	2	?	0
0	?	?	?	2	?	2	?	0	2	2	?	?	2	?	0
1	?	?	?	2	?	2	?	0	2	1	?	?	2	?	0
1	?	?	?	2	?	1	?	1	2	2	?	?	2	?	0
2	?	?	?	2	?	2	?	1	2	1	?	?	2	?	0
1	?	?	?	1	?	1	?	1	2	2	?	?	2	?	0
1	?	?	?	2	?	2	?	0	2	1	?	?	2	?	1
2	?	?	?	1	?	1	?	1	2	1	?	?	2	?	1
1	?	?	?	0	?	0	?	2	2	2	?	?	2	?	0

Reference panel: sequence-level data

0	0	0	0	1	1	1	0	0	1	1	1	1	1	1	0
1	1	1	1	1	1	0	0	1	0	0	1	1	1	0	0
1	1	1	1	1	0	1	0	0	1	0	0	0	1	0	1
0	0	1	0	1	1	1	0	0	1	1	1	1	1	1	0
1	1	1	0	1	1	0	0	1	1	1	0	1	1	1	0
0	0	1	0	1	1	1	0	0	1	1	1	1	1	1	0
1	1	1	1	1	0	1	0	0	1	0	0	0	1	0	1
1	1	1	0	0	1	0	0	1	1	1	0	1	1	1	0
0	0	0	0	1	1	1	0	0	1	1	1	1	1	1	0
1	1	1	0	0	1	0	0	1	1	1	0	1	1	1	0

1	1	1	1	1	2	1	0	0	2	2	0	2	2	2	0
0	0	1	0	2	2	2	0	0	2	2	2	2	2	2	0
1	1	1	1	2	2	2	0	0	2	1	1	2	2	2	0
1	1	2	0	2	2	1	0	1	2	2	1	2	2	2	0
2	2	2	2	2	1	2	0	1	2	1	1	2	2	2	0
1	1	1	0	1	2	1	0	1	2	2	1	2	2	2	0
1	1	2	1	2	1	2	0	0	2	1	1	1	2	1	1
2	2	2	1	1	1	1	0	1	2	1	0	1	2	1	1
1	2	2	0	0	2	0	0	2	2	2	1	2	2	2	0

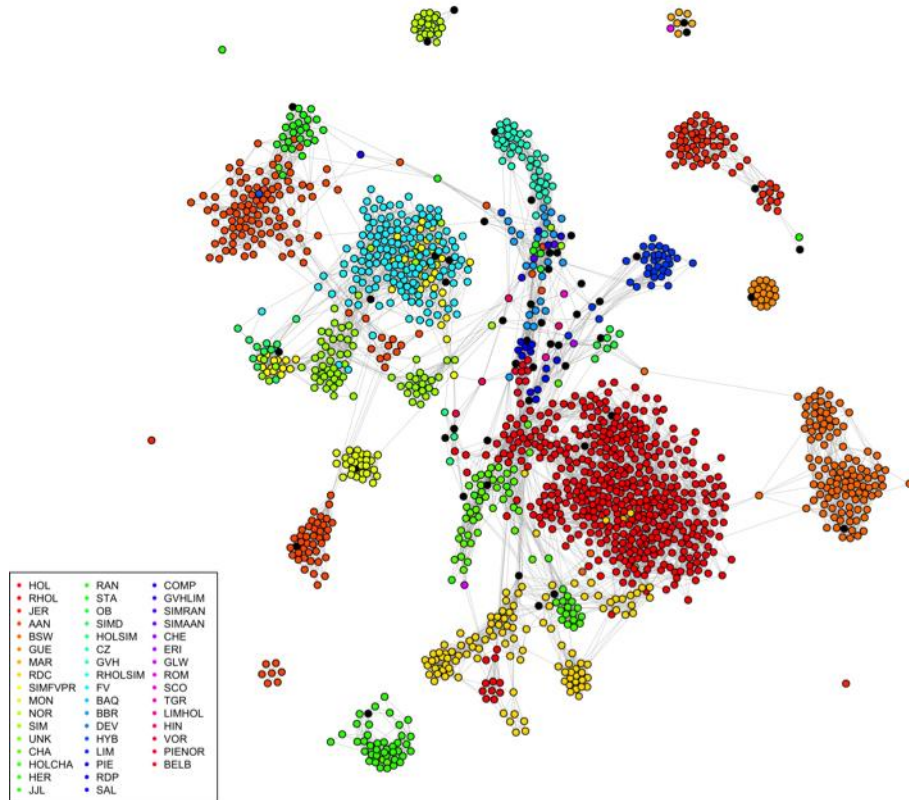
# Objectives

- ▶ **Assess accuracy of sequence imputation**
  - ▶ **Impact on GWAS for fat content**

# Imputation software

	Beagle v4	FImpute	Impute2	Minimac v3
Algorithm	LD-based	Family / LD-based	LD-based	LD-based
Phasing	yes	yes	yes (but often done in SHAPEIT)	No
Dosage data	yes	no	Genotype probability	yes

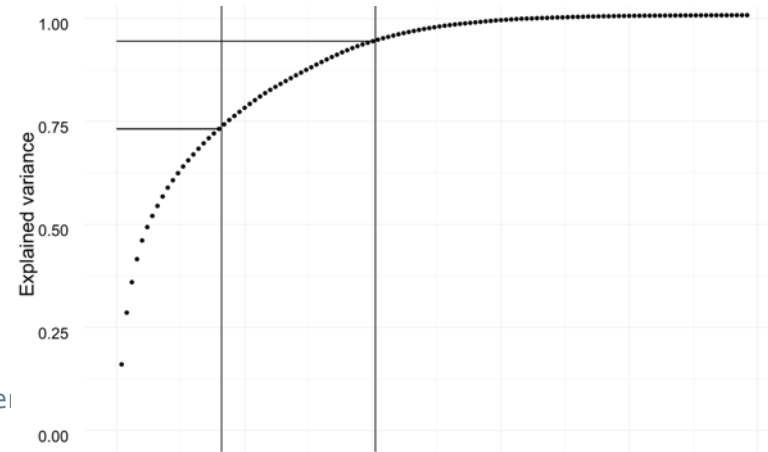
# Data: 1000 bulls genome project Run 5



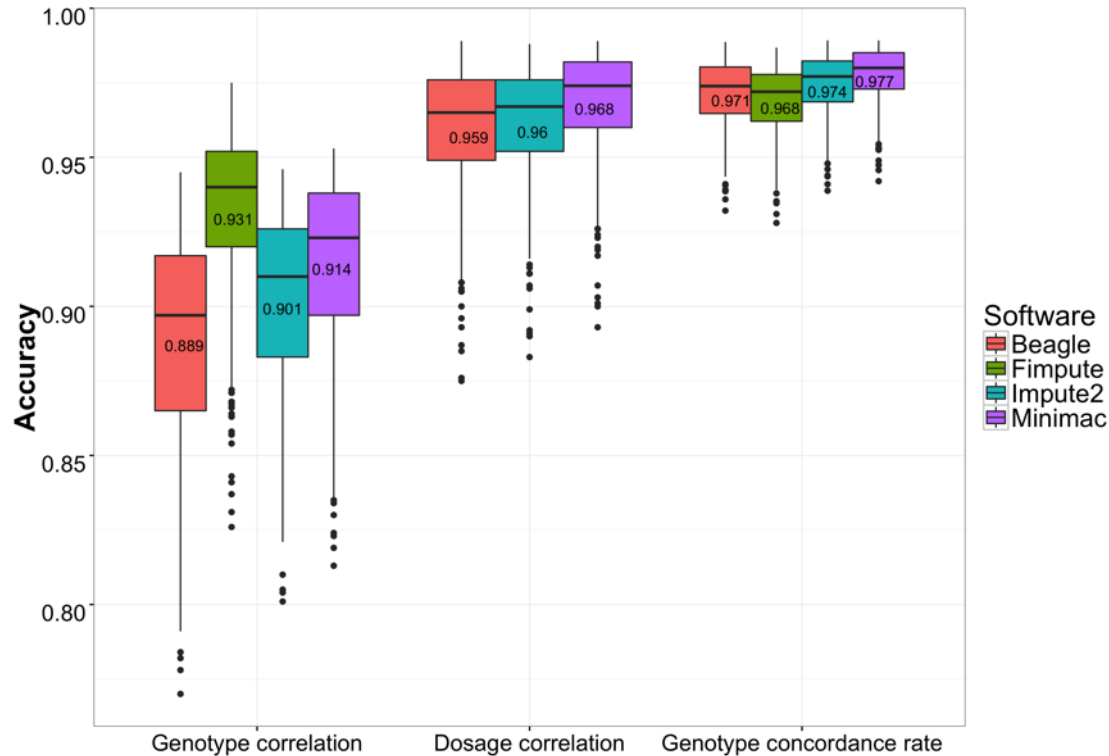
- ▶ 1577 *Bos taurus*
- ▶ 123 Brown Swiss including 8 original Braunvieh
- ▶ Imputation from HD to sequence data

# Scenarios

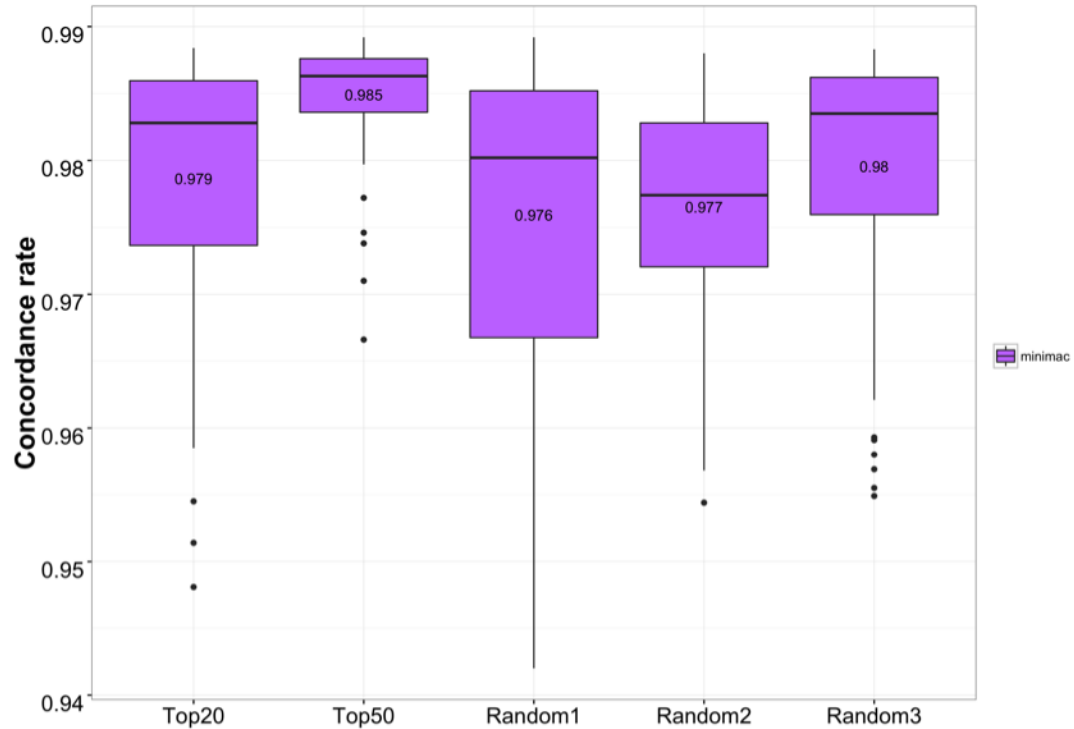
- ▶ Variants:
  - ▶ SNPs on BTA25 (~643,000 Sequence including 12,222 HD)
- ▶ Scenario 1: 20 random BSW for reference (3 times), 103 BSW as validation
  - ▶ Comparison of programs and comparison of measure type
- ▶ Scenario 2: Top20 or 50 BSW key ancestors as reference, 103 or 73 BSW as validation



# Imputation accuracy within Brown Swiss

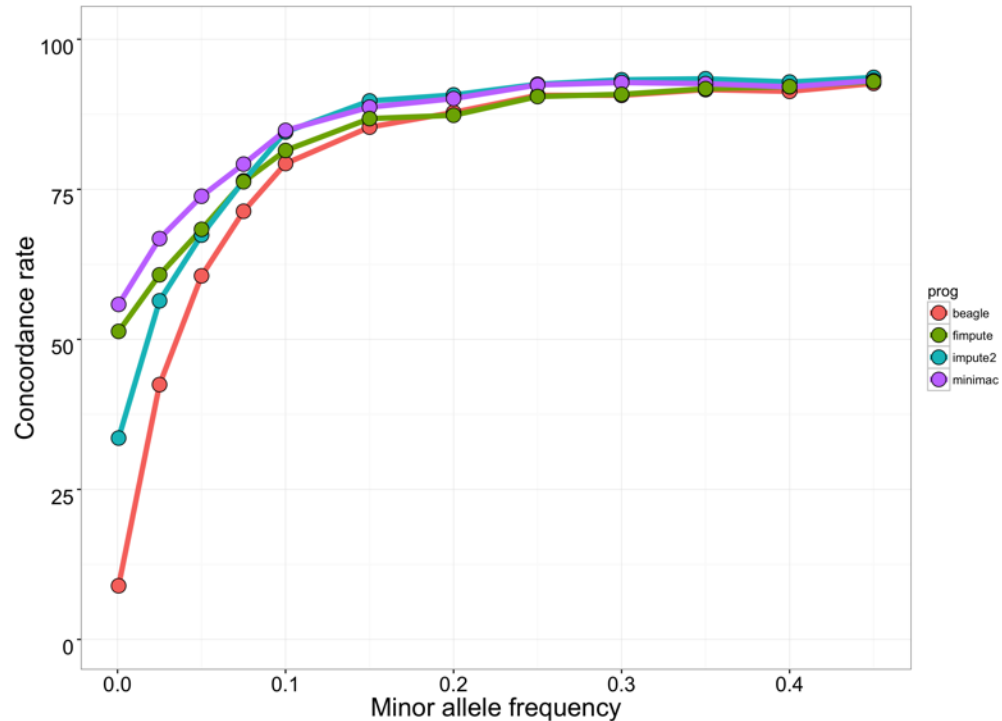


# Selection of Individuals





# Imputation accuracy by MAF classes

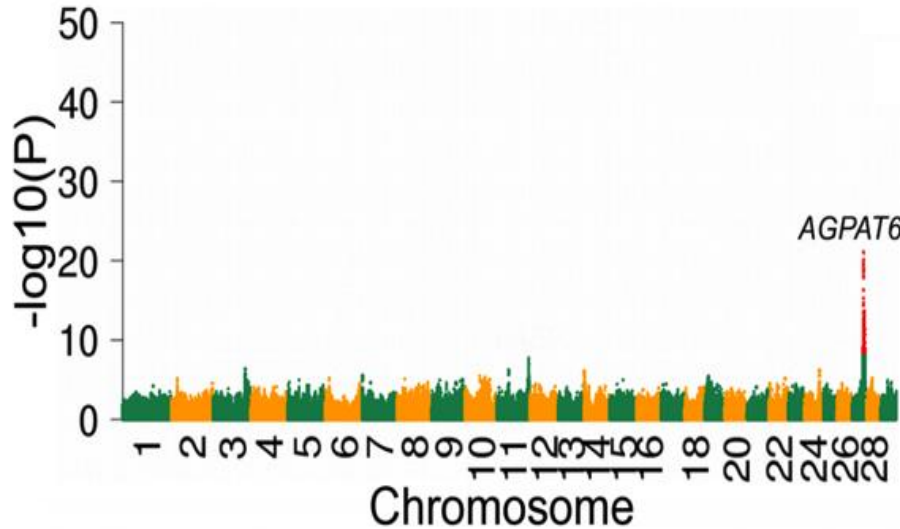


# GWAS

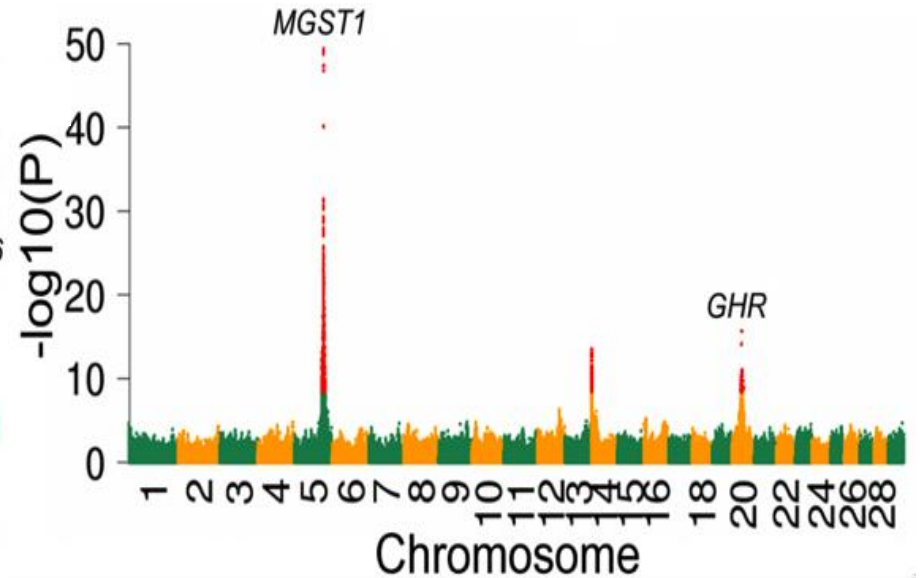
- ▶ Imputation:
  - ▶ FImpute (HD), Beagle (phasing), Minimac (seq)
- ▶ 128 BSW sequenced, ~14 Mio SNPs
- ▶ 2 Phenotypes for fat content (FC) in 1646 bulls:
  - ▶  $FC_{\text{EARLY}}$ : Lactation day 8-12
  - ▶  $FC_{\text{LATE}}$ : Lactation day 298-302
  - ▶ EBVs
- ▶ GWAS with dosage data in EMMAX:  $y = \mu + Xb + Zu + e$

# GWAS results

FC<sub>EARLY</sub>

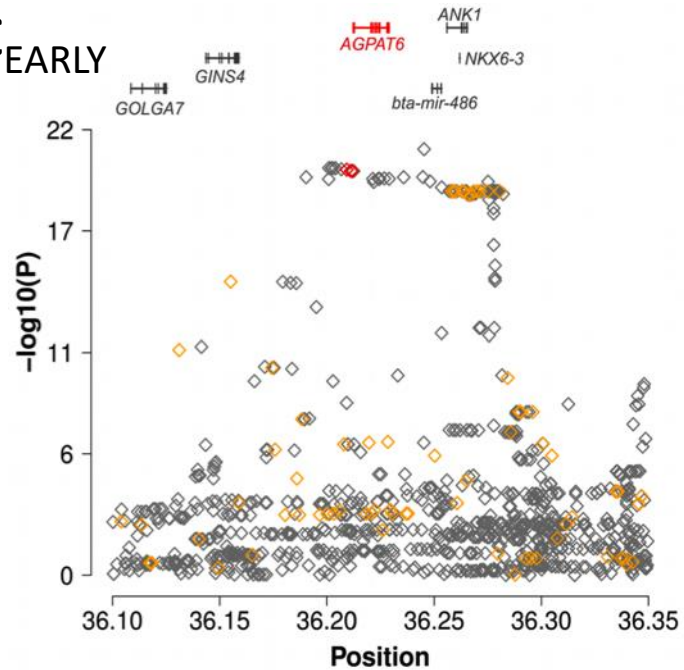


FC<sub>LATE</sub>

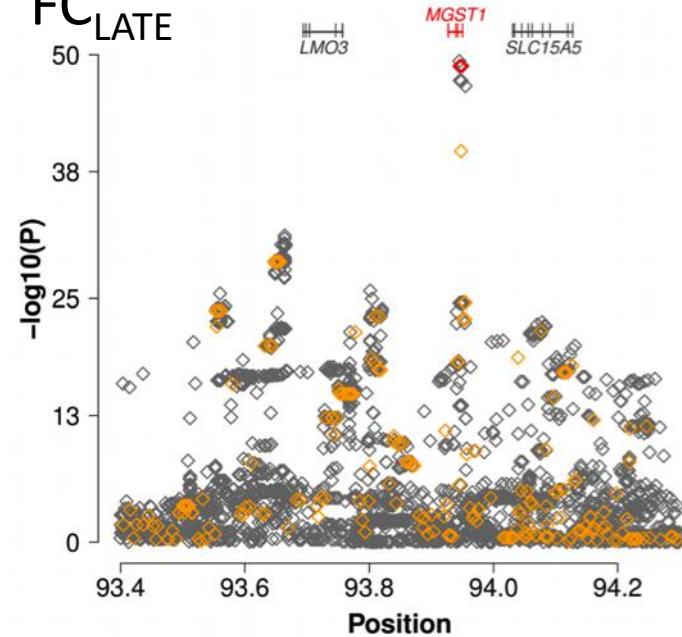


# Zoom

FC<sub>EARLY</sub>



FC<sub>LATE</sub>



# Summary

- ▶ Overall accuracy from HD to sequence data is high
- ▶ The accuracy measurement has an impact on the ranking of the programs
- ▶ MAF impacts accuracy
  
- ▶ Using imputed genotypes causative variants can be identified

# Acknowledgements



**1000 bull genomes project**

**BRAUNVIEH** 

**KTI/CTI**



Thank you for your attention

