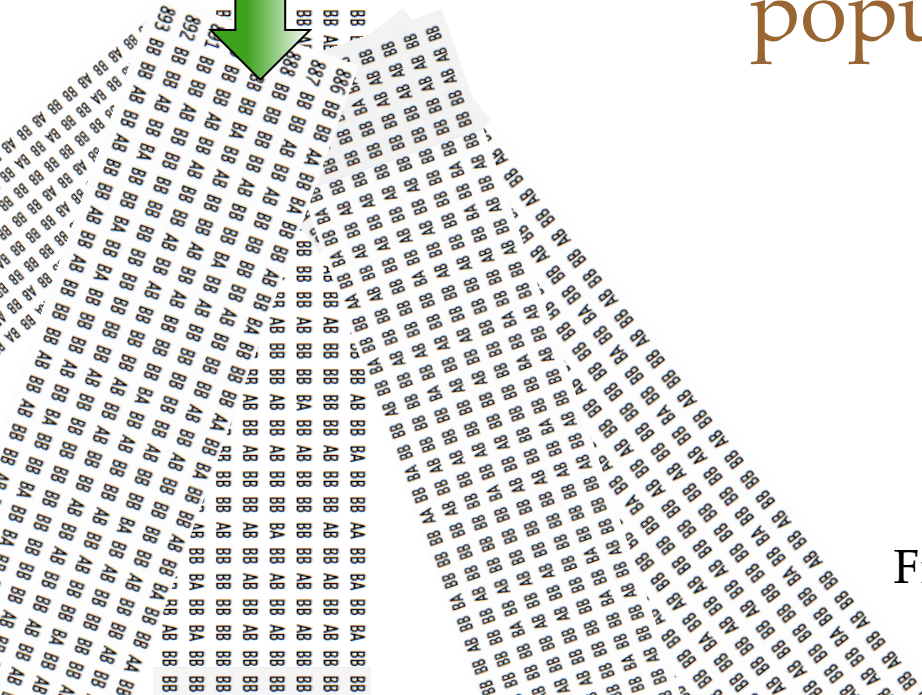




Analysis of next-generation sequence data of Swiss dairy populations



Christine Baes, Marlies Dolezal, Beat Bapst,
Christine Flury, Heidi Signer-Hasler, Rohan
Fernando, Dorian Garrick, Christian Stricker,
Fritz Schmitz-Hsu, Sandra Jansen, Ruedi Fries,
Birgit Gredler

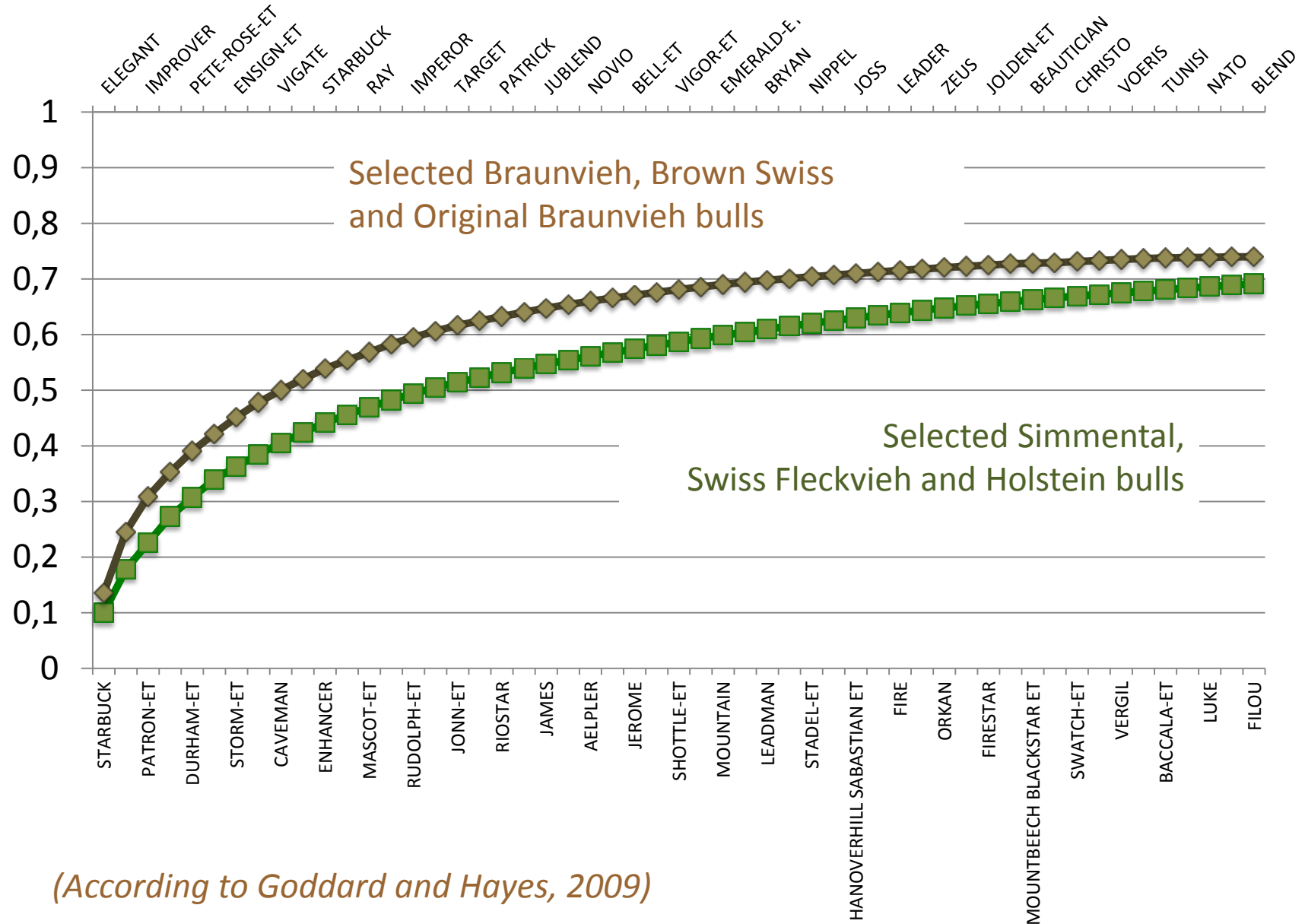
Overview

- Bull Selection
- Sequence Data
- Pipeline
- Concordance analysis
- Summary

Bull Selection

Bull selection

Proportion of genetic variance explained



(According to Goddard and Hayes, 2009)

Sequence Data

66 key ancestors + 1 calf:

- 8 Brown Swiss
- 18 Braunvieh
- 8 Original Braunvieh
- 17 (Red) Holstein
- 4 Swiss Fleckvieh
- 12 Simmental

Illumina HiSeq2000 data, Paired-end reads

Duplicates marked with Picard

Genome-wide, filtered read depth = 12x

The Pipeline

Step 1: Data Preparation

Raw .bam files

STEP 1a:
Sort & index .bam files

Sorted .bam & .bai
files

Reference (UMD3.1)

STEP 1b:
PrepareReference

Reference.dict
Reference.idx

dbSNP (ensembl)

STEP 1c:
Sort & prepare
dbSNP file

dbSNPwork
dbSNPwork.idx

Reference.dict & .idx
.bam & .bai files

STEP 1d:
ExtractRegion

bam & .bai files

24

Chr24: 200,531

Chr24: 200,581

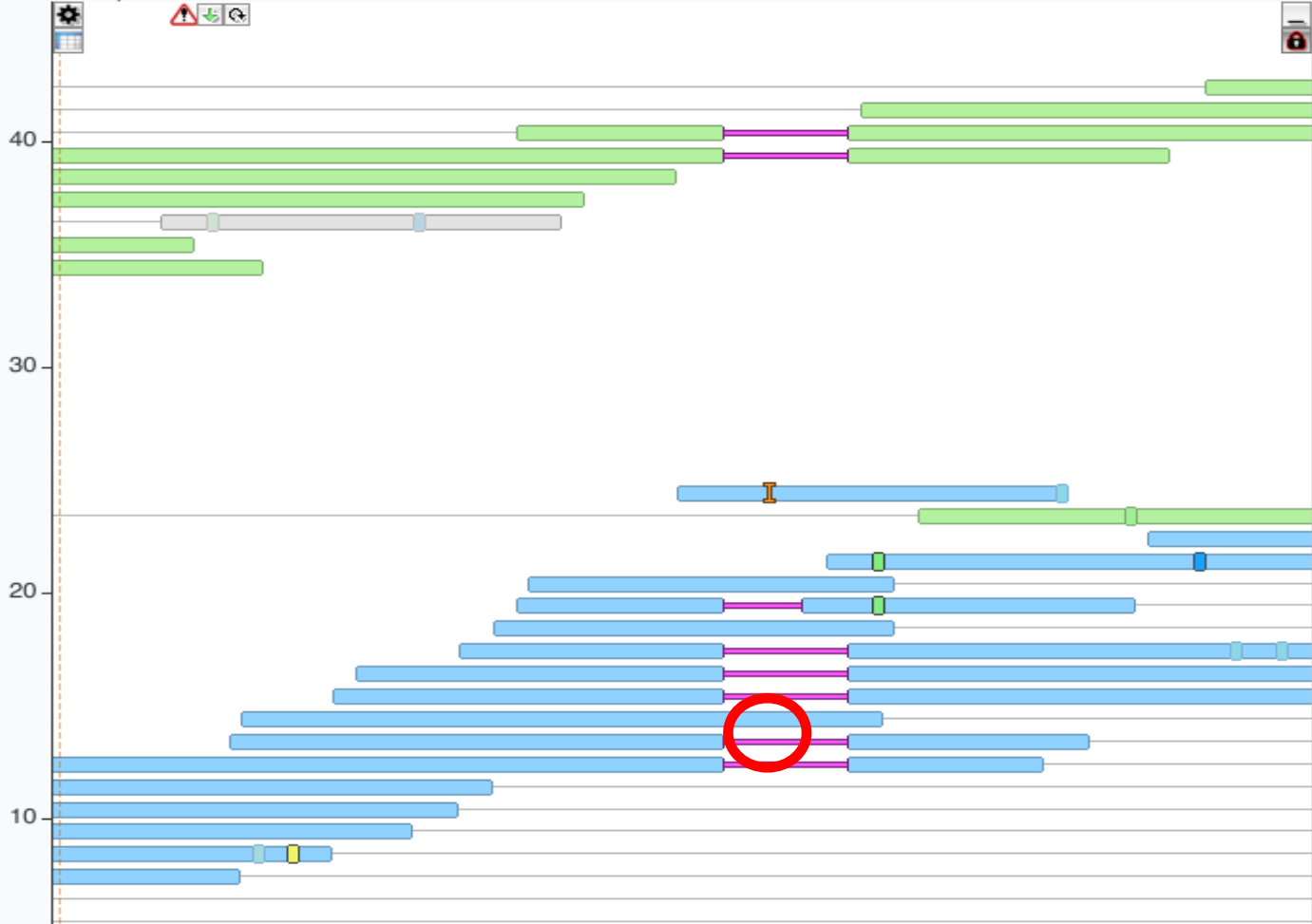
EnsemblGenes, Ensembl (Bos taurus, UMD_3.1)

data.goldenhelix.com

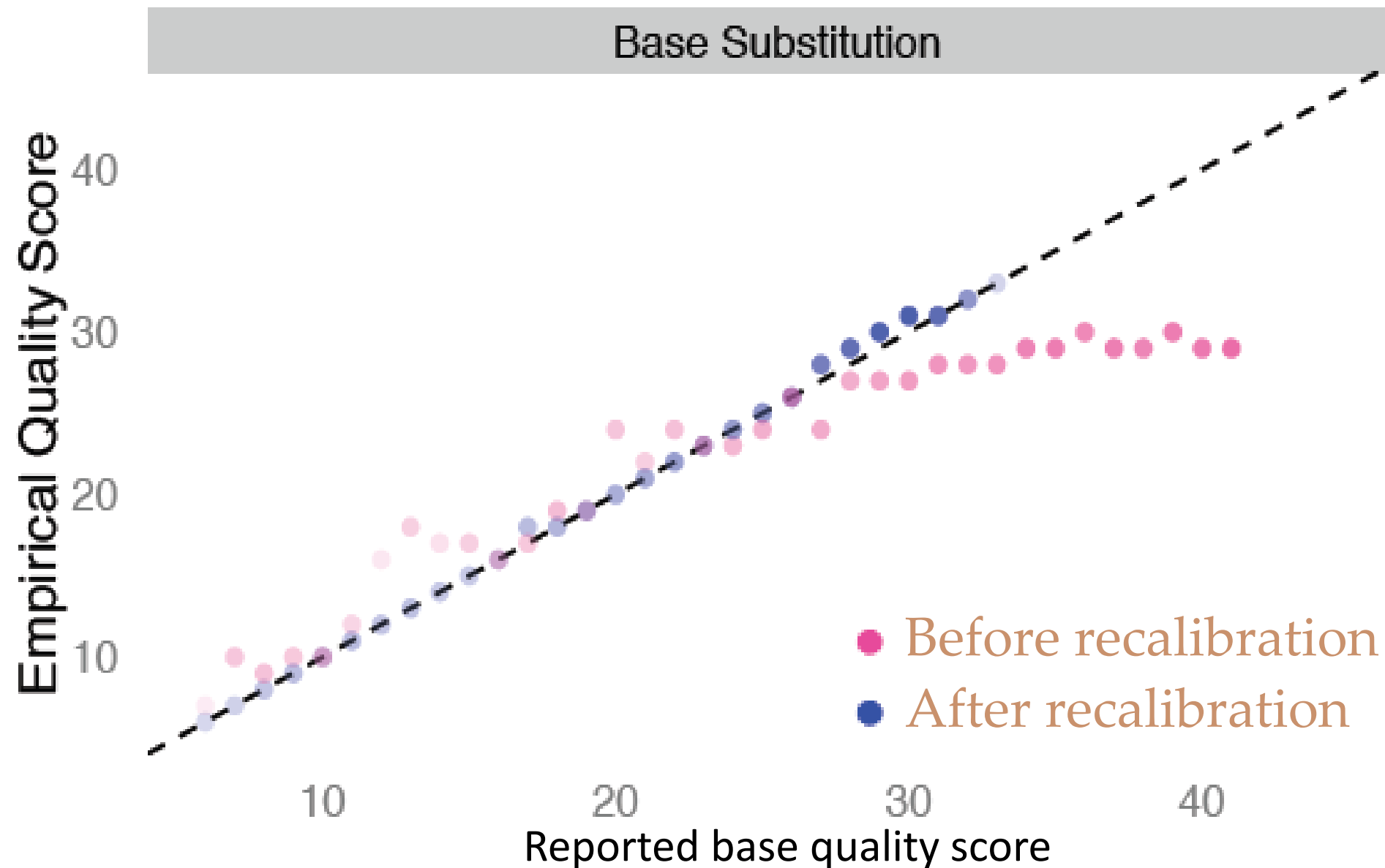
68295.realigned.recal (Bos taurus, UMD_3.1)

BTA24_SING_NT

Pile-up



Step 3: Base (Quality Score) Recalibration



Step 4: Variant Calling

Reference.dict & .idx
Real. recal .bam & .bai files
Realigned.grp

GATK:
Haplotype Caller

HC.vcf

Reference.dict & .idx
Real. .bam & .bai files
Realigned.grp

GATK:
Unified Genotyper

UG.vcf

Reference.dict & .idx
Real. .bam & .bai files
Realigned.grp

SAMTOOLS:
Mpileup

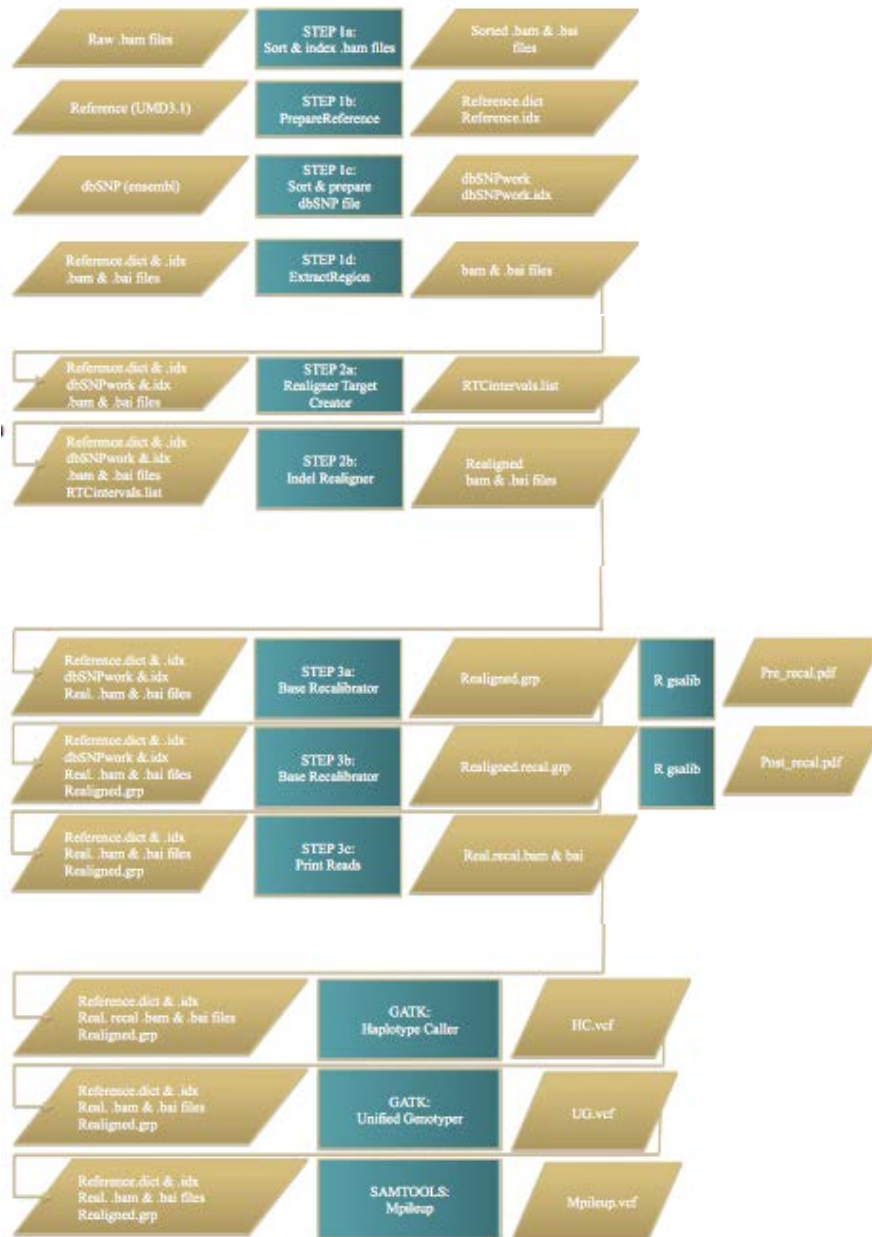
Mpileup.vcf

Data Preparation

Indel Realignment

Base Recalibration

Variant Calling



Variant Calling

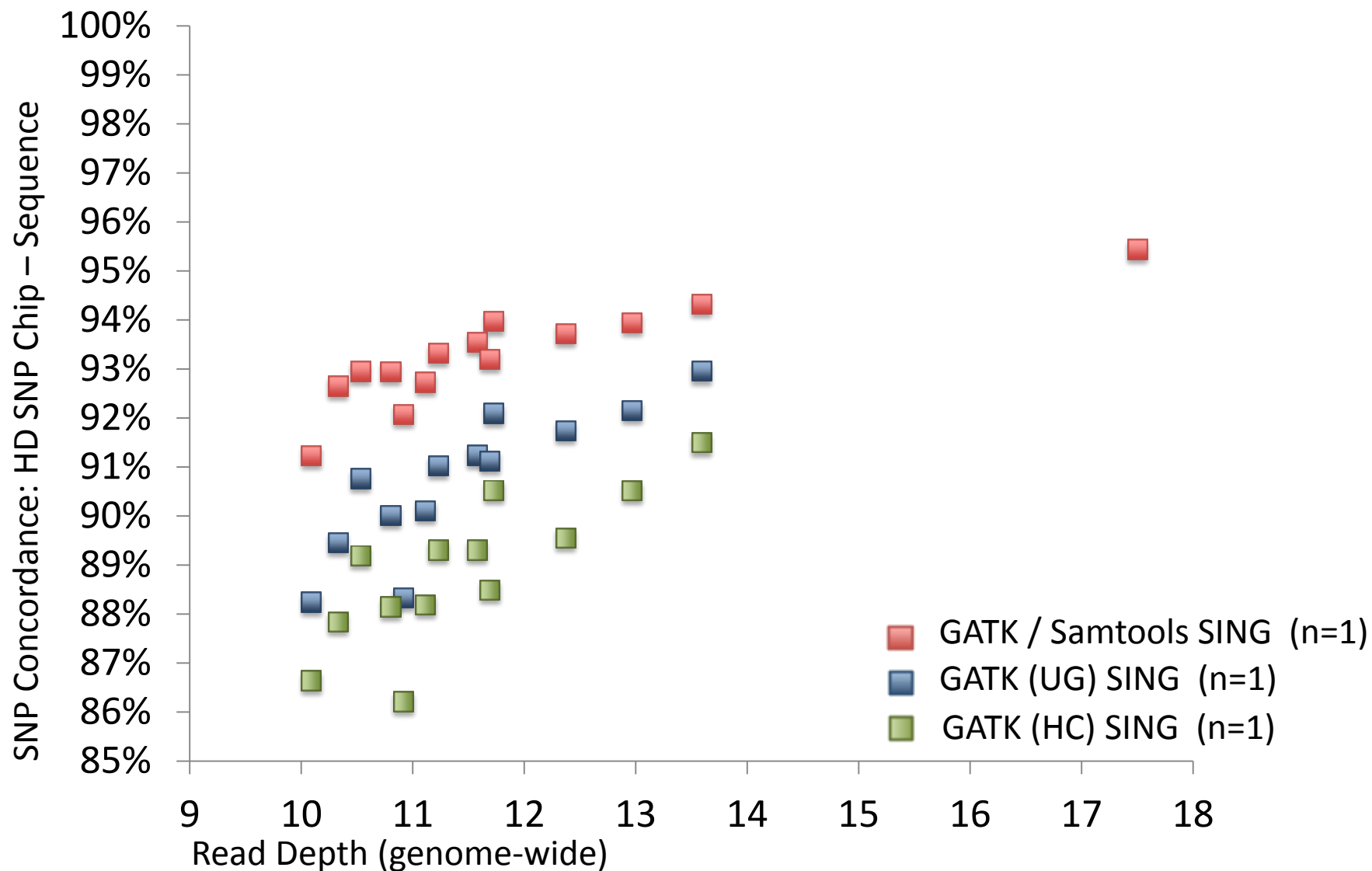
- 14 animals with Illumina HD information
- 1 chromosome (BTA24)
- SNP concordance

GATK HaplotypeCaller
GATK Unified Genotyper
GATK / Samtools Mpileup

Concordance

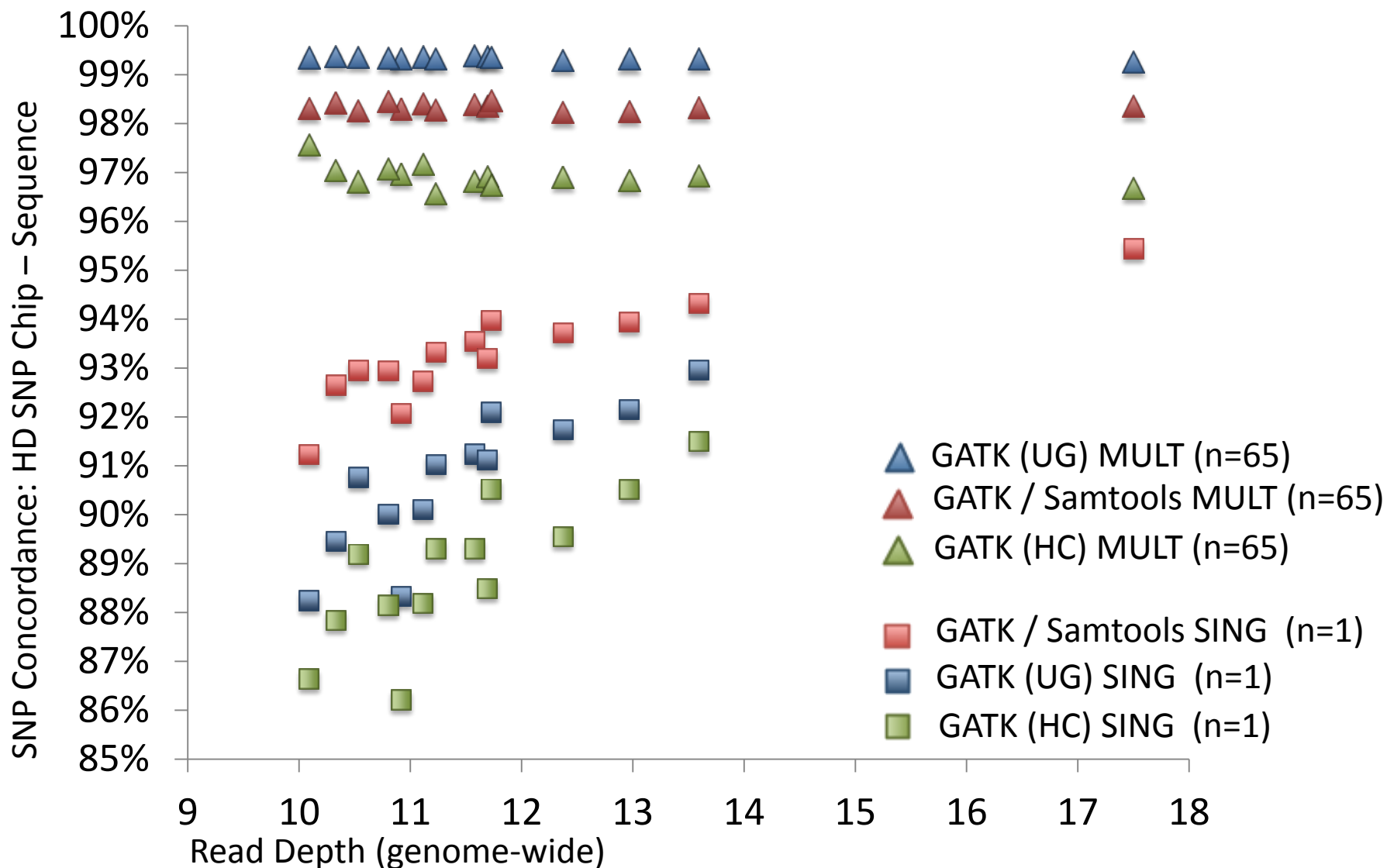
SNP Concordance

HD Chip – Sequence (BTA24)



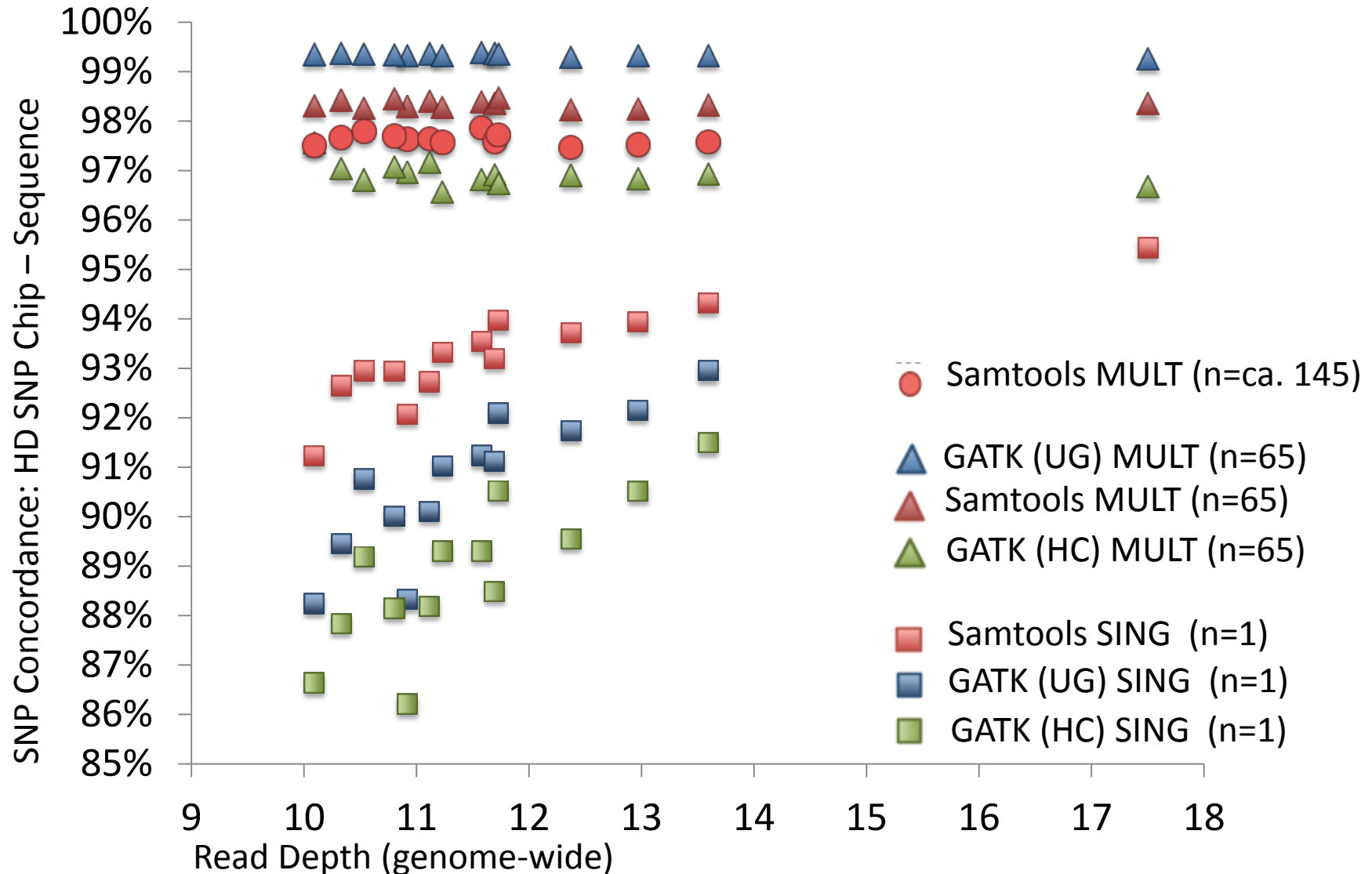
SNP Concordance

HD Chip – Sequence (BTA24)



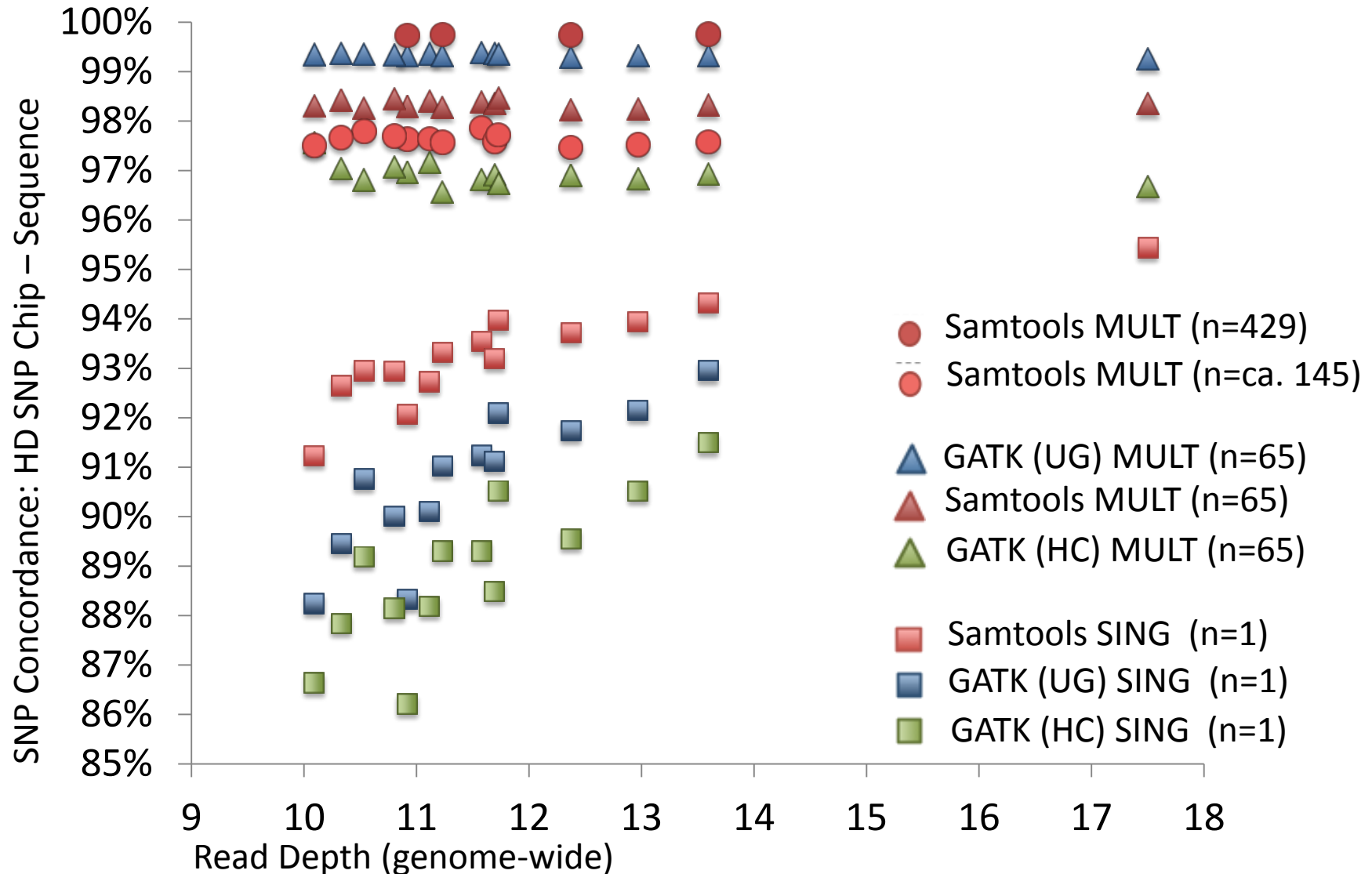
SNP Concordance

HD Chip – Sequence (BTA24)



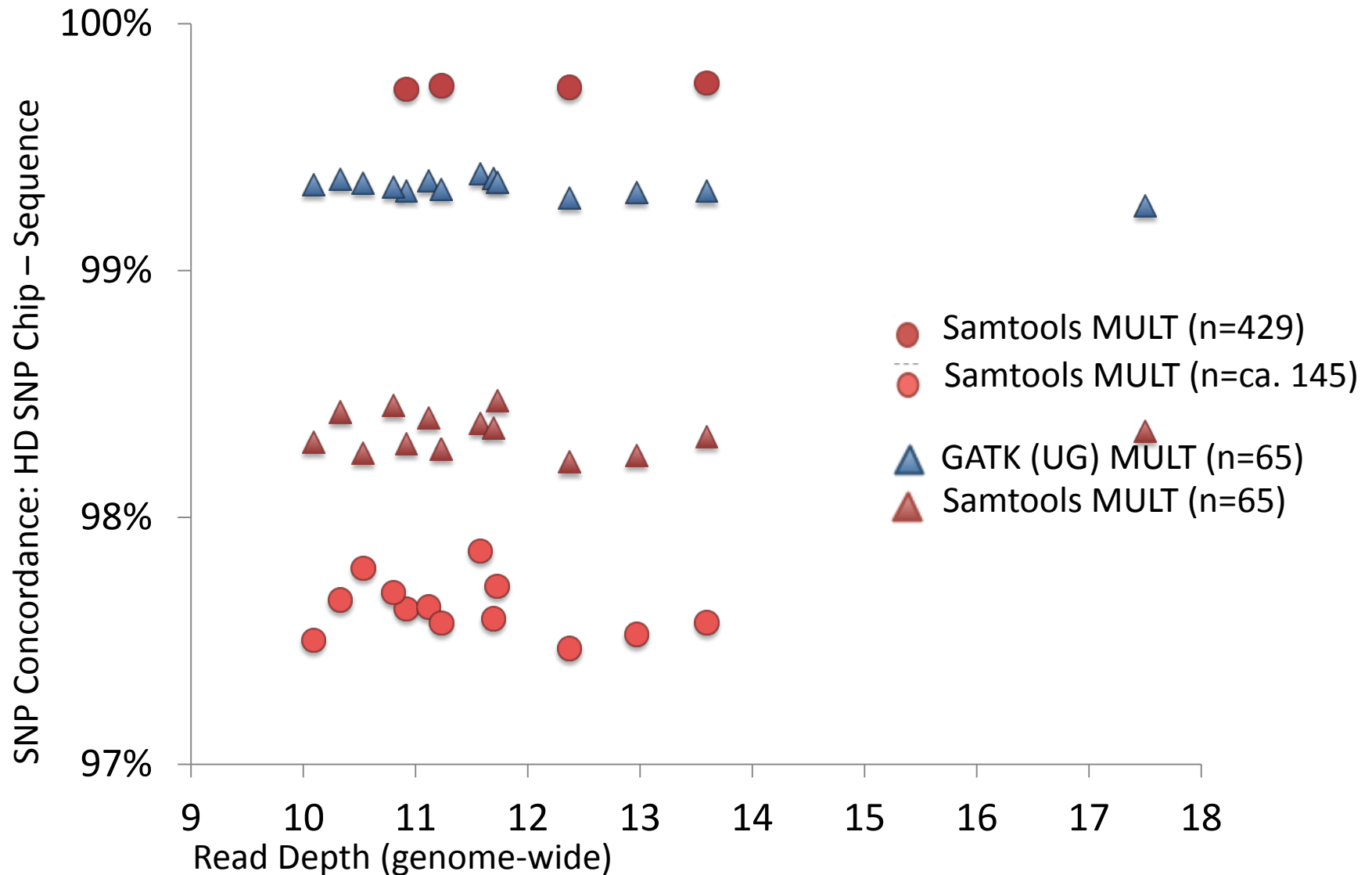
SNP Concordance

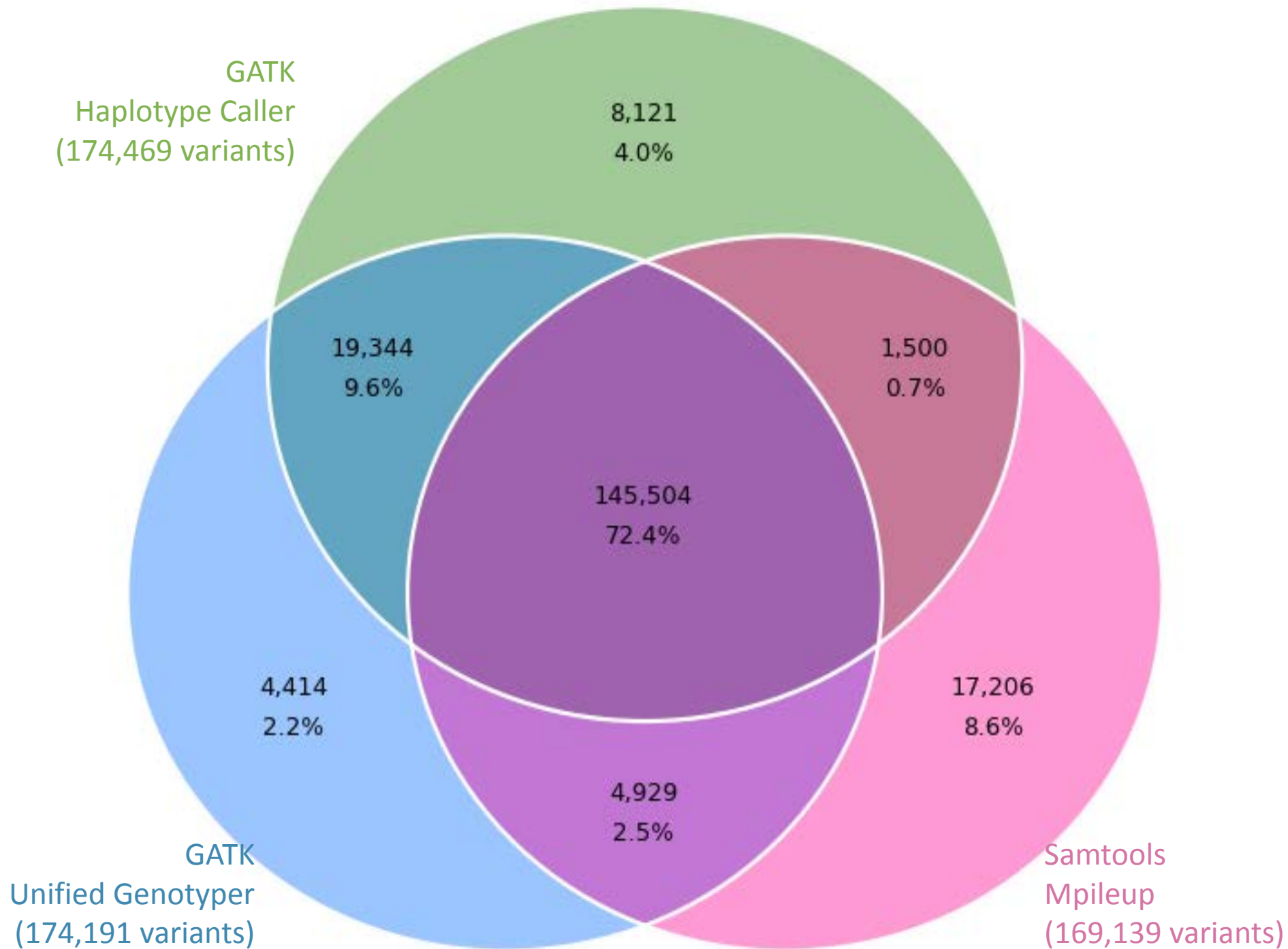
HD Chip – Sequence (BTA24)



SNP Concordance

HD Chip – Sequence (BTA24)





Summary

- Indel Realignment and Base Quality Score Recalibration are important
- A “hybrid” GATK / Samtools and GATK SNP concordance can be as good as stand-alone Samtools results with considerably more animals
- This statement holds for “easily” detected SNPs (still need to analyse genotypes)
- No quantification possible for false positive SNPs or false negatives → especially for rare alleles



Bild: simplyfantasticbooks.com

Thanks to Team Qualitas:

Madeleine Berweger, Andreas Bigler,
Jürg Moll, Franz Seefried, Urs Schuler,
Urs Schnyder

Financial support from the Association
of Swiss Cattle Breeders and the Swiss
Commission for Technology and
Innovation is gratefully acknowledged

Non Reference
Sensitivity

SNP CHIP

		AA	AB	BB
SEQUENCE	AA	a	b	c
	AB	d	e	f
	BB	g	h	i
	--	k	l	m

(NRS)

SNP CHIP

		AA	AB	BB
SEQUENCE	AA	a	b	c
	AB	d	e	f
	BB	g	h	i
	--	k	l	m

Non Reference
Discrepancy

SNP CHIP

		AA	AB	BB
SEQUENCE	AA	a	b	c
	AB	d	e	f
	BB	g	h	i
	--	k	l	m

(NRD)

SNP CHIP

		AA	AB	BB
SEQUENCE	AA	a	b	c
	AB	d	e	f
	BB	g	h	i
	--	k	l	m



created 10 months
ago by [Neilfws](#) ♦♦

37,440 • 1 • 20 • 51

updated 10 months ago by

[Neilfws](#) ♦♦

I think these are the key parts of the error message:

```
ERROR A USER ERROR has occurred (version 2.0-34-g07bda93):  
ERROR The invalid arguments or inputs must be corrected before the GATK can proceed
```

You did something wrong.

```
ERROR Please do not post this error to the GATK forum
```

It's a common mistake and they don't want to hear about it.

```
ERROR See the documentation (rerun with -h) for this tool to view allowable command-line arguments.  
ERROR Visit our website and forum for extensive documentation and answers to commonly asked questions
```

Read the manual.

```
ERROR MESSAGE: Invalid command line: No tribble type was provided on the command line and the type of the file  
ot be determined dynamically. Please add an explicit type tag :NAME listing the correct type from among the su  
ypes
```

This is what you did wrong.

