

# Breed specific reference genomes in cattle

**Bartosz Czech<sup>1</sup>, Magda Mielczarek<sup>1,2</sup>, Magdalena Frąszczak<sup>1</sup>, Joanna Szyda<sup>1,2</sup>**

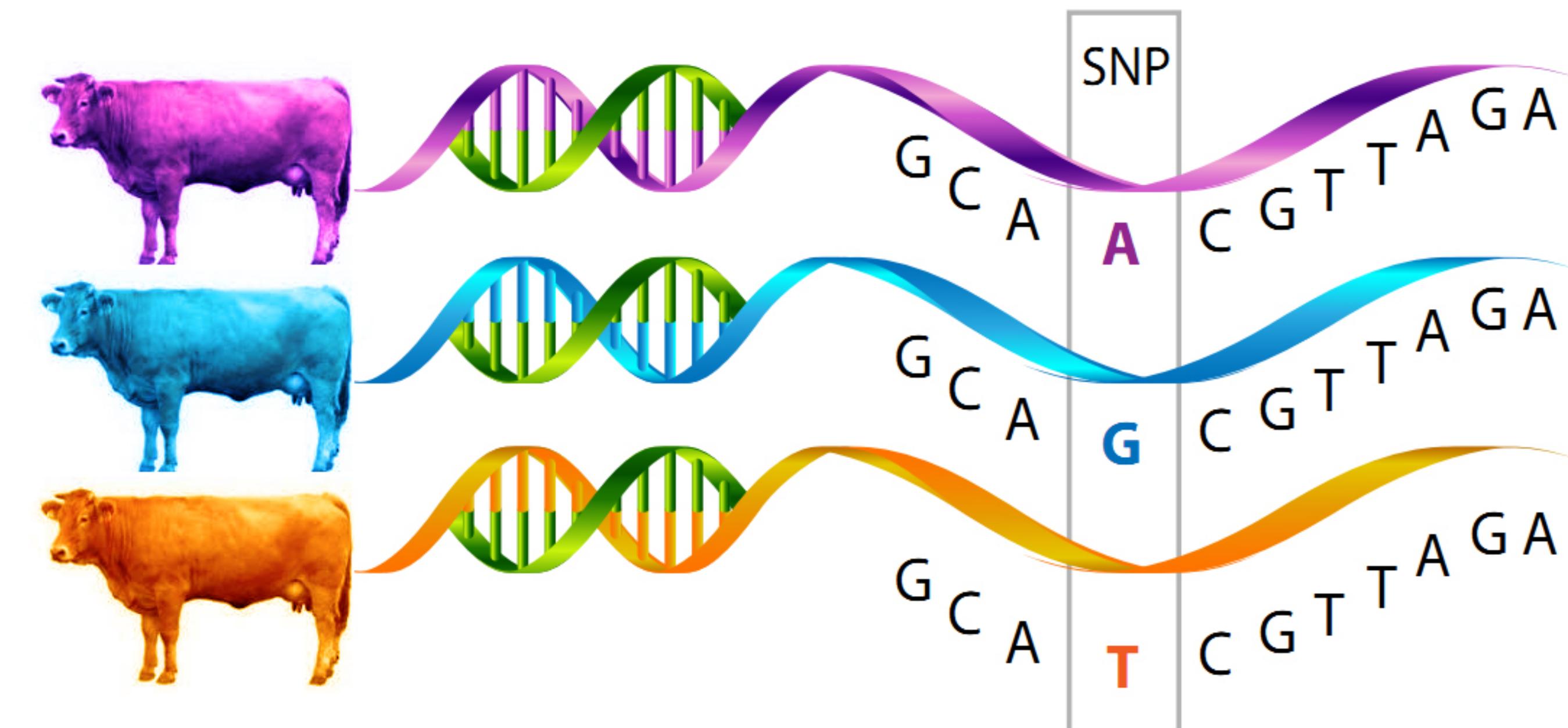
<sup>1</sup>Wrocław University of Environmental and Life Sciences, Biostatistics Group

<sup>2</sup>National Research Institute of Animal Production



# Objectives

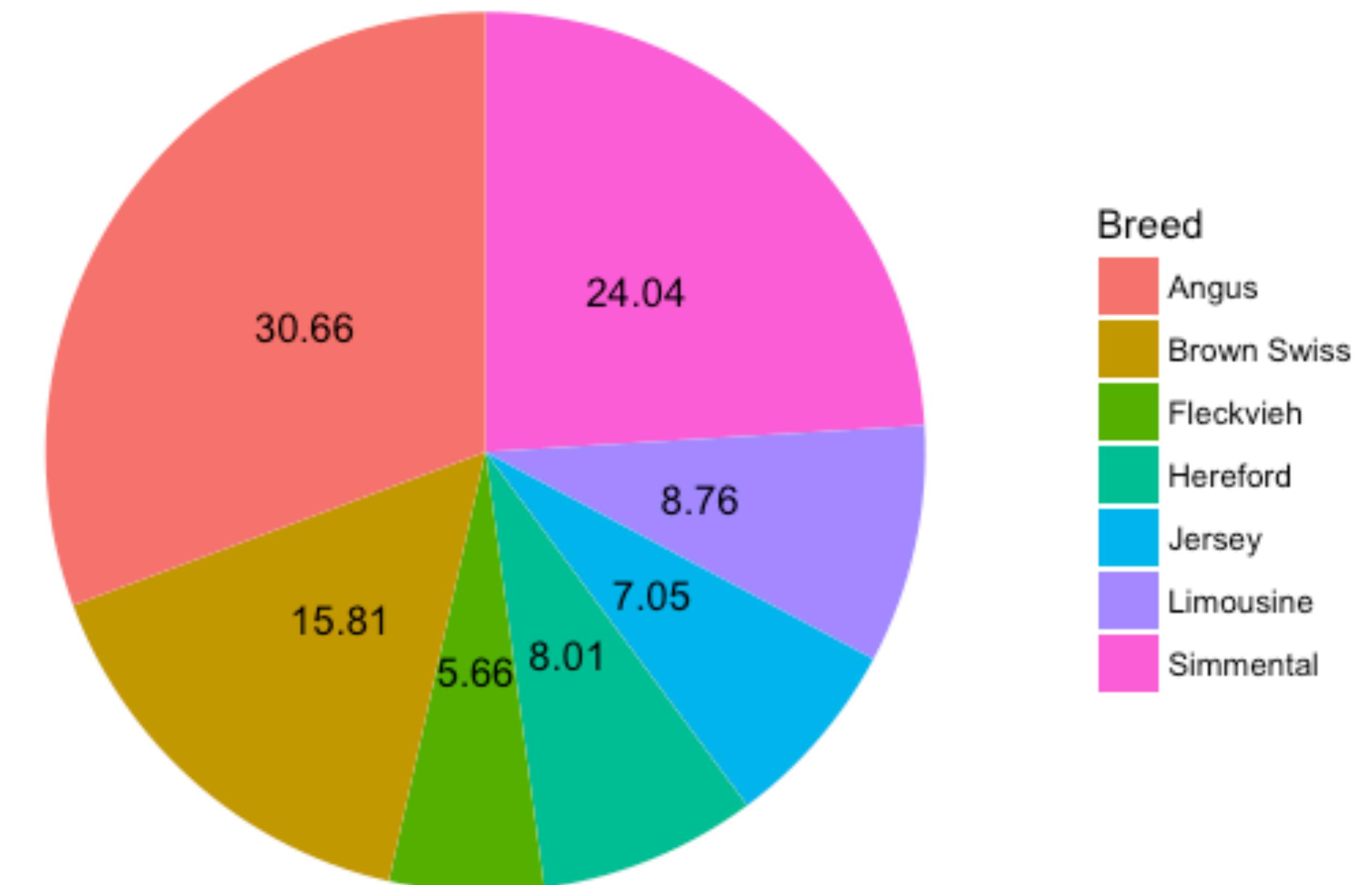
- Identification of breed-specific SNPs
- Creating breed-specific reference genomes
- Annotation of breed-specific SNPs



# Material

---

- Whole genome DNA-seq – 936 individuals
- Reference genome UMD 3.1



# Methods

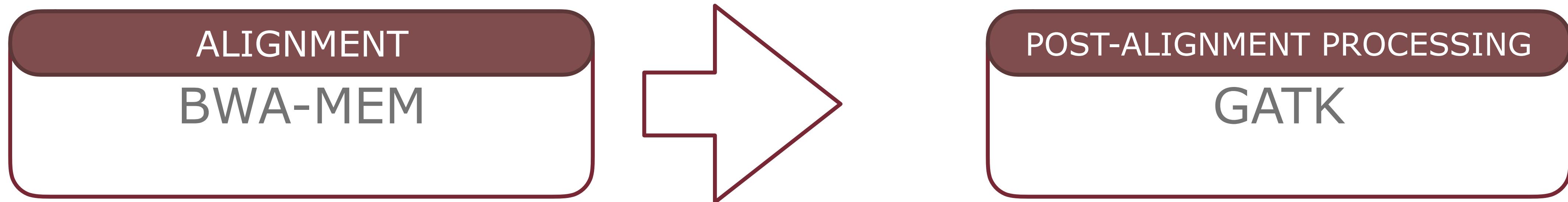
---

ALIGNMENT

BWA-MEM

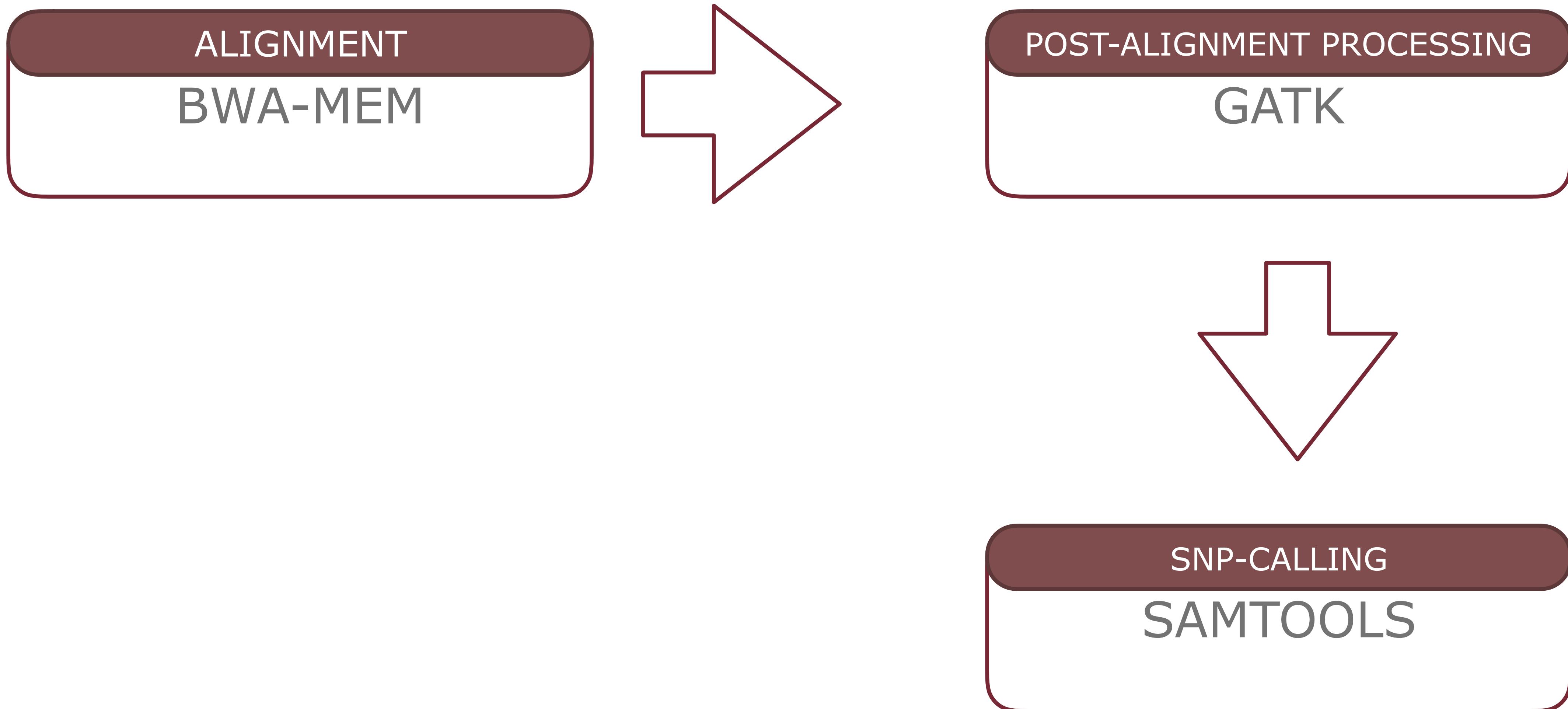
# Methods

---

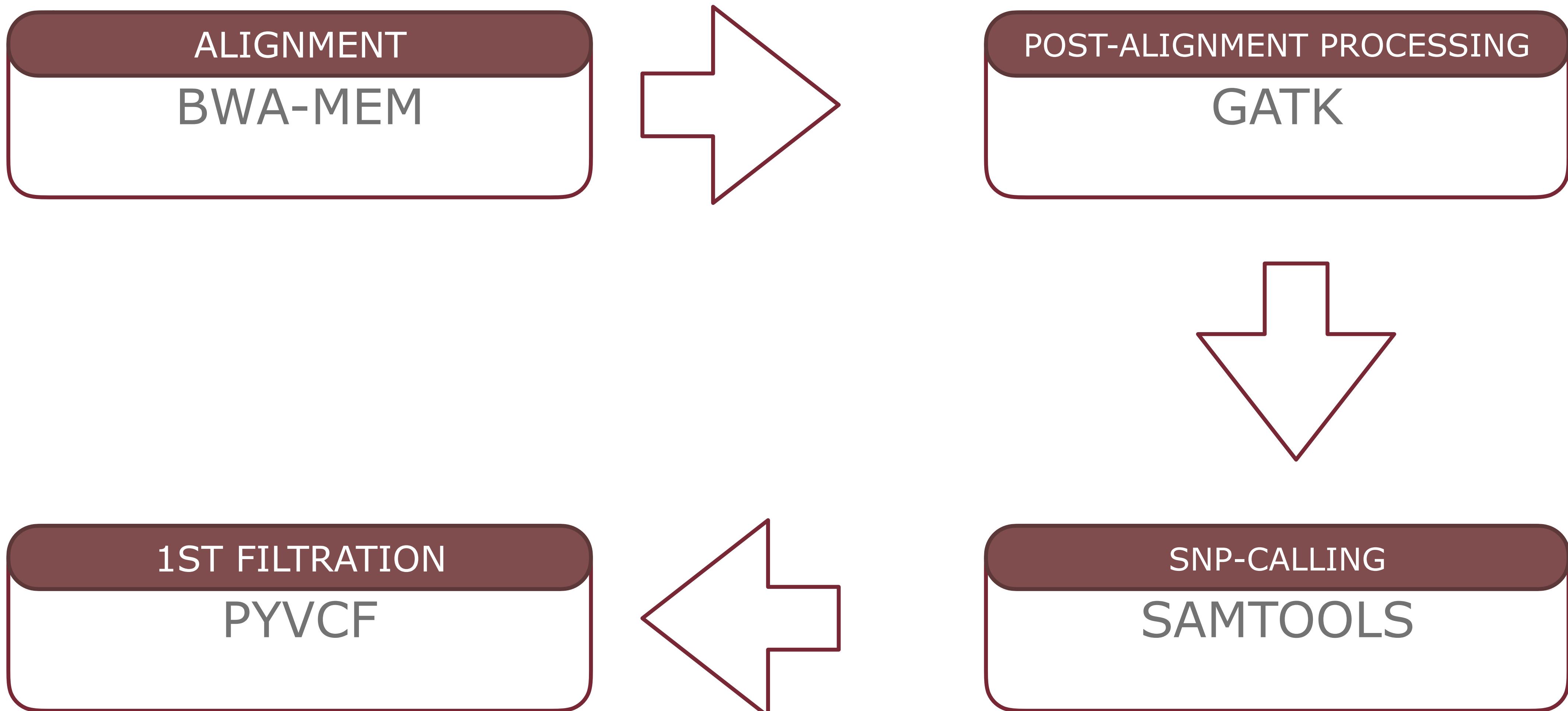


# Methods

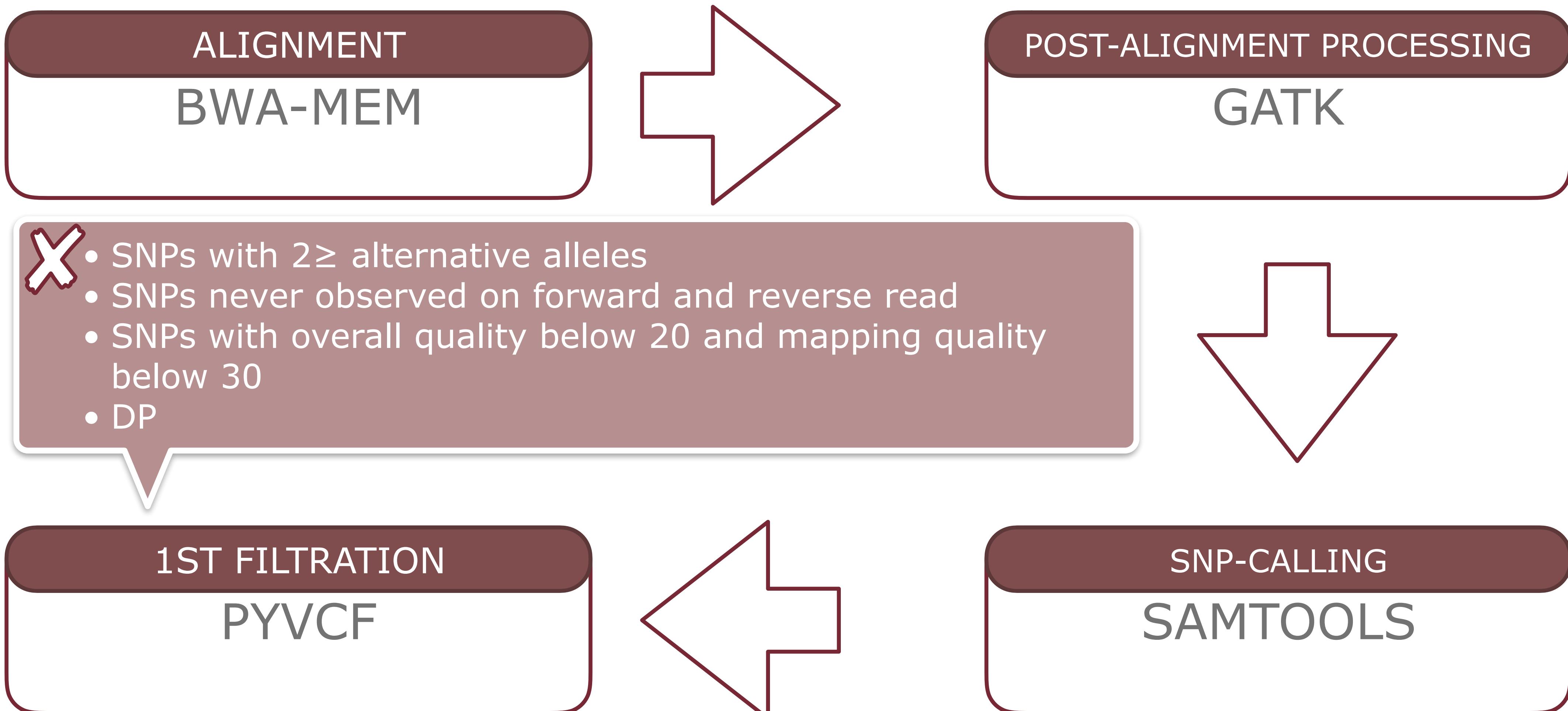
---



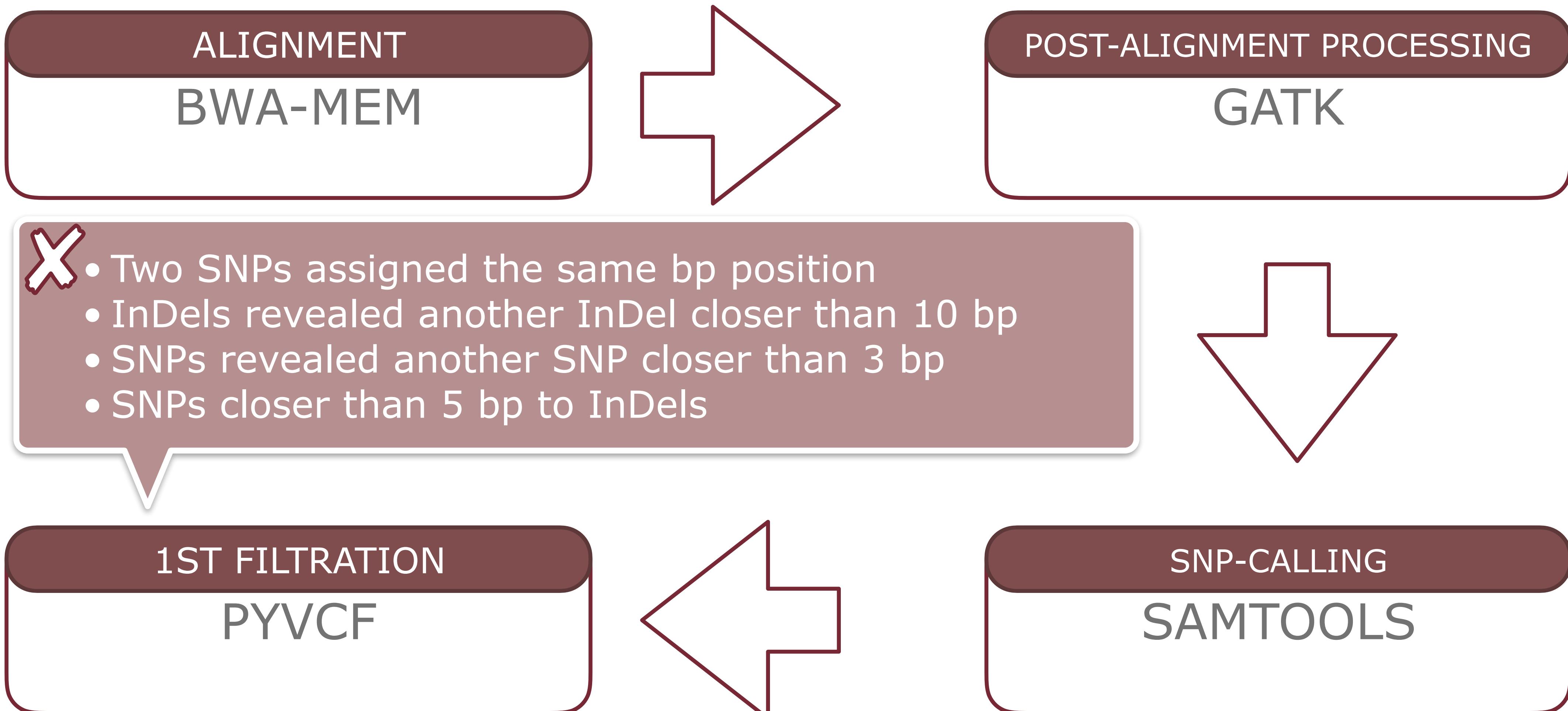
# Methods



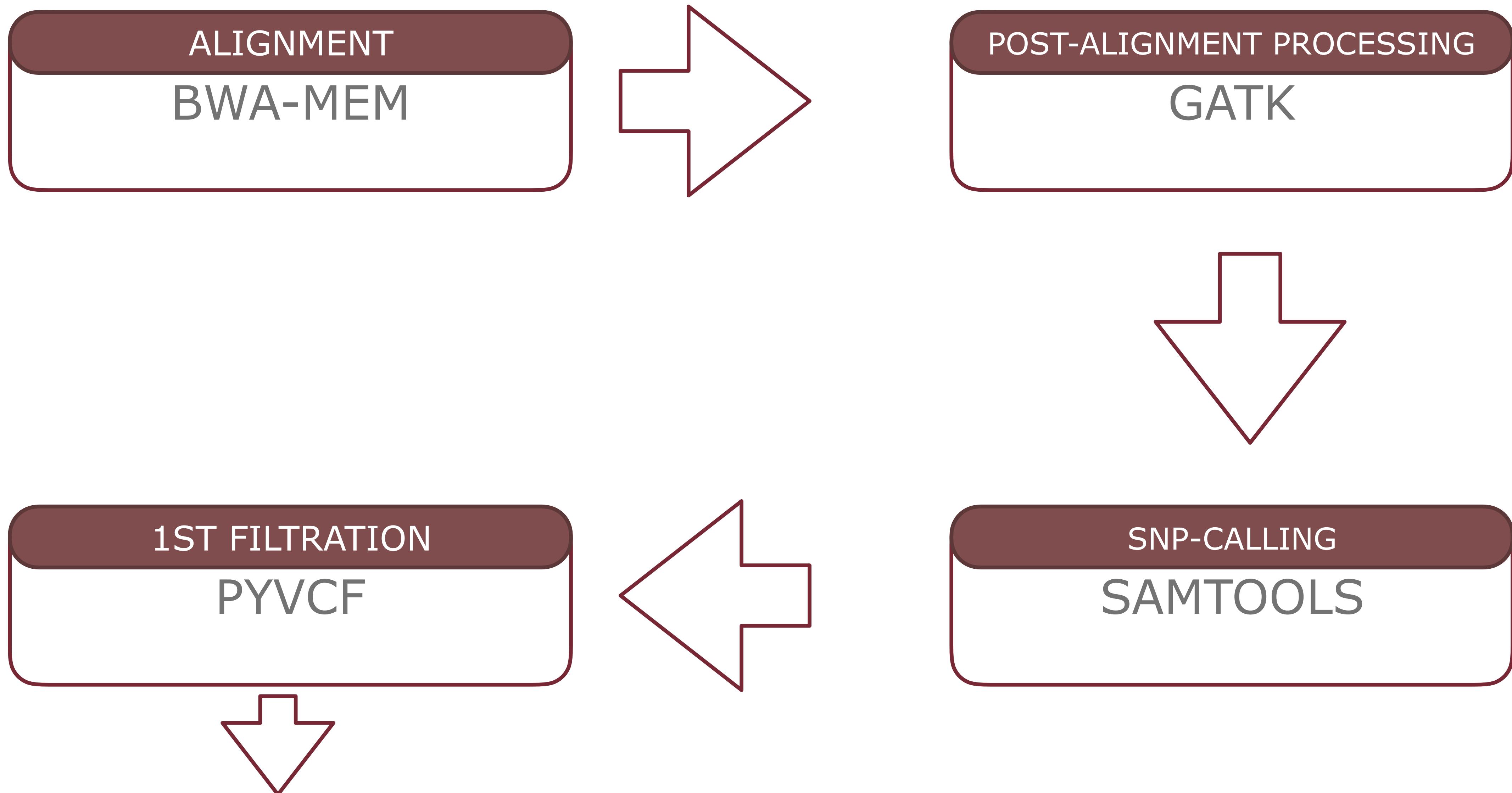
# Methods



# Methods



# Methods



# Methods

---



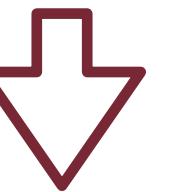
- AF ≠ 1
- Missing genotypes >7%

2ND FILTRATION  
VCF-TOOLS

# Methods

---

2ND FILTRATION  
VCF-TOOLS

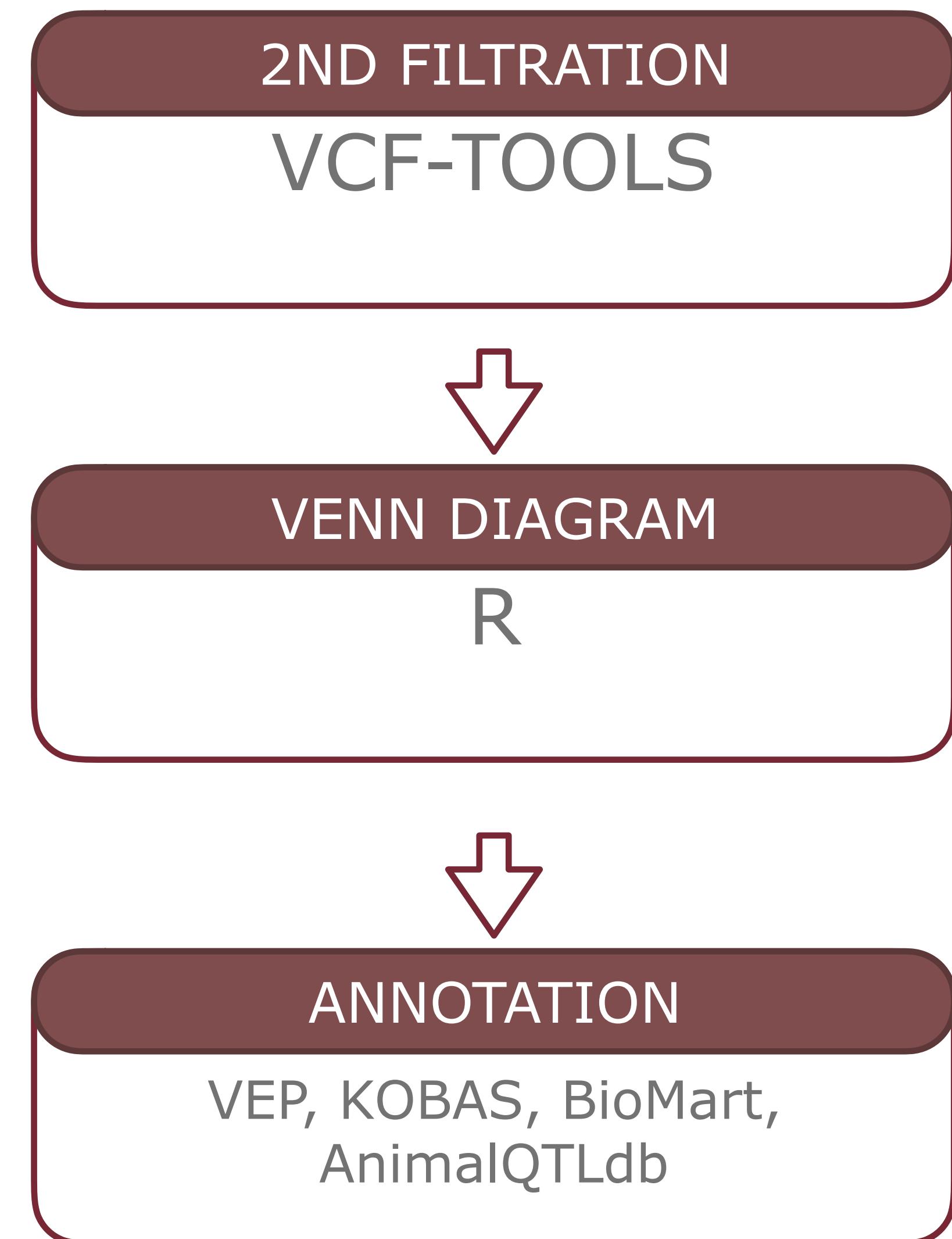


VENN DIAGRAM

R

# Methods

---

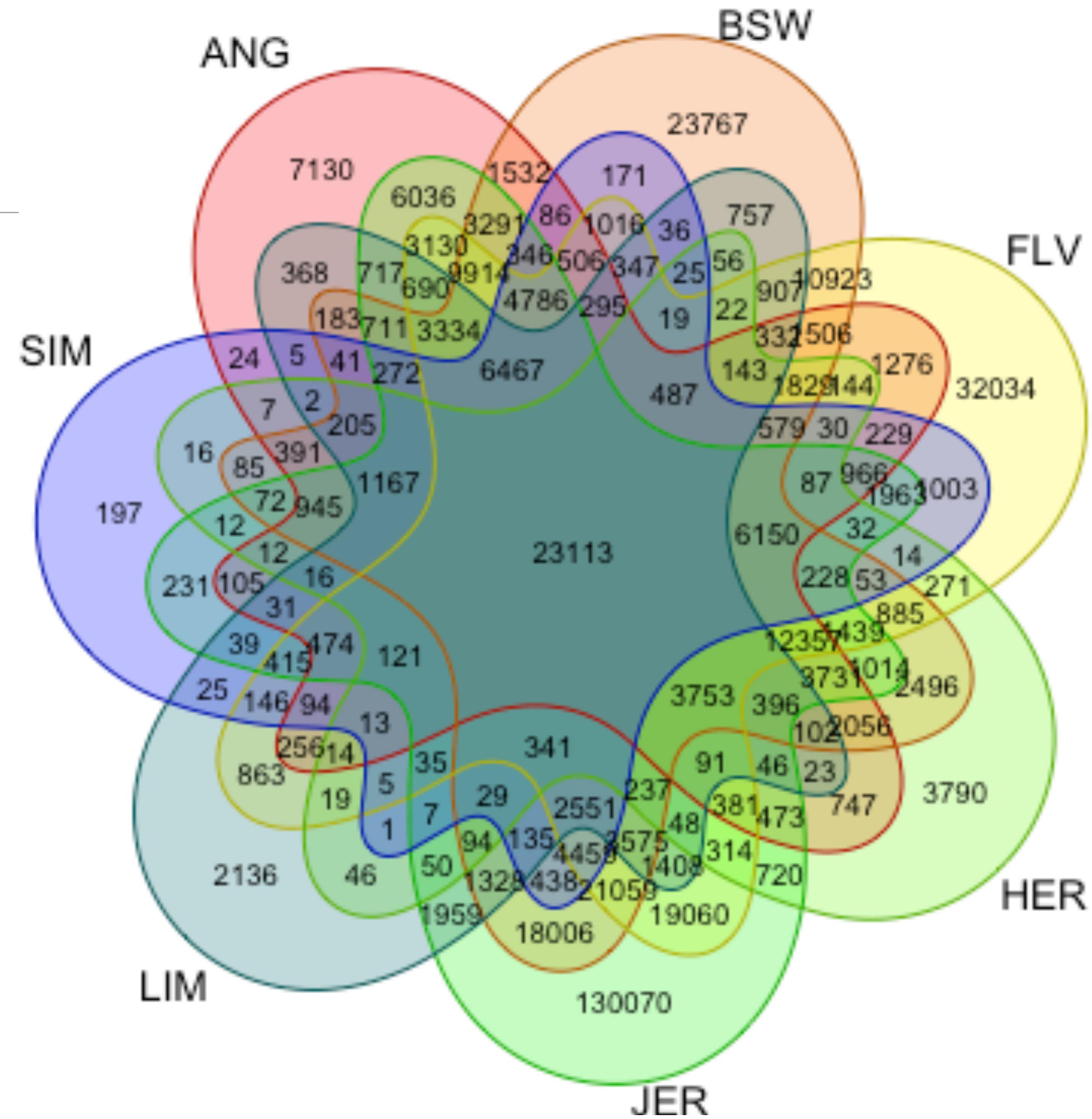


# Results

---

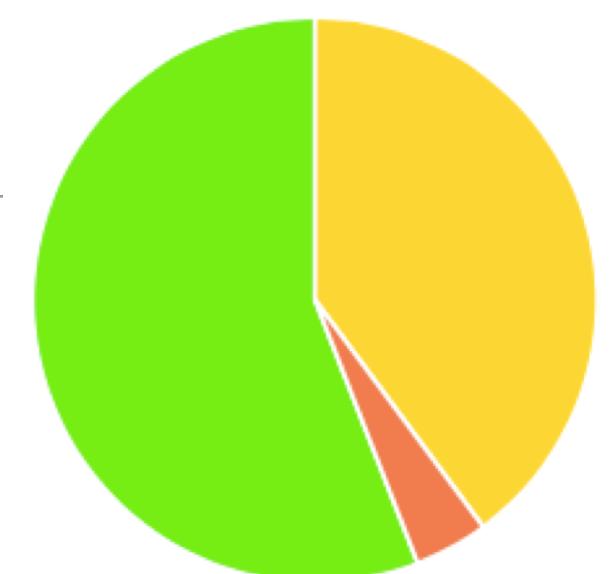
Breed	Total number of SNPs	Number of SNPs after the filtration	% of SNPs after the filtration
Angus	61 823 312	117 744	0.186
Brown Swiss	61 814 873	187 576	0.303
Fleckvieh	61 780 912	189 238	0.306
Hereford	61 813 028	73 772	0.119
Jersey	61 797 641	305 477	0.494
Limousin	61 820 287	61 623	0.100
Simmental	61 824 209	60 582	0.098

# Results



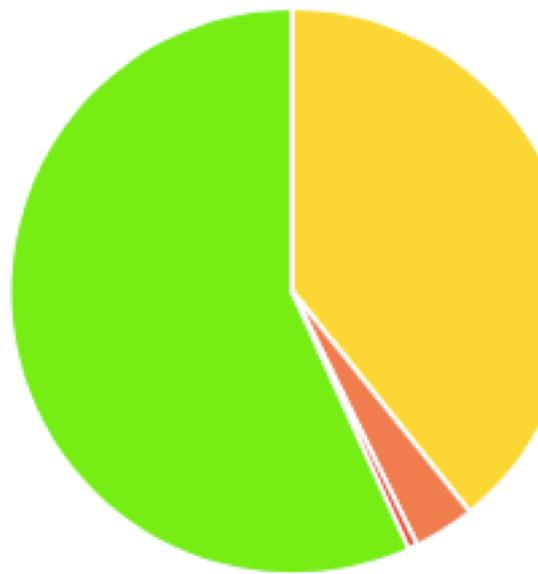
# Results

ANGUS



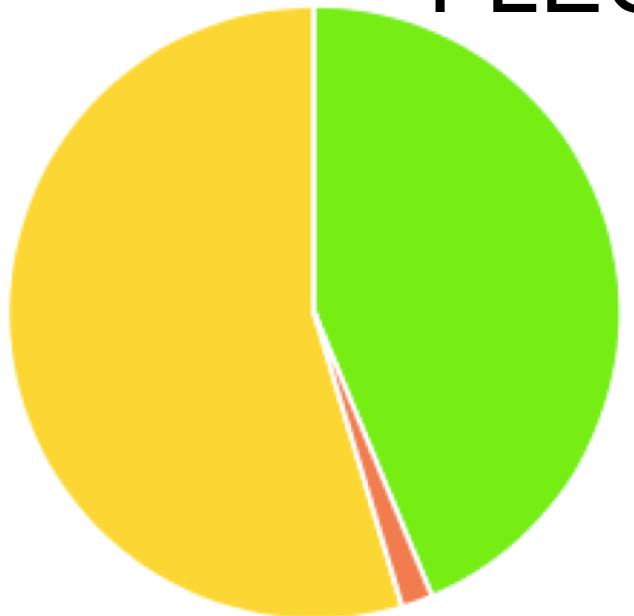
- synonymous\_variant: 56%
- missense\_variant: 40%
- splice\_region\_variant: 4%

BROWN SWISS



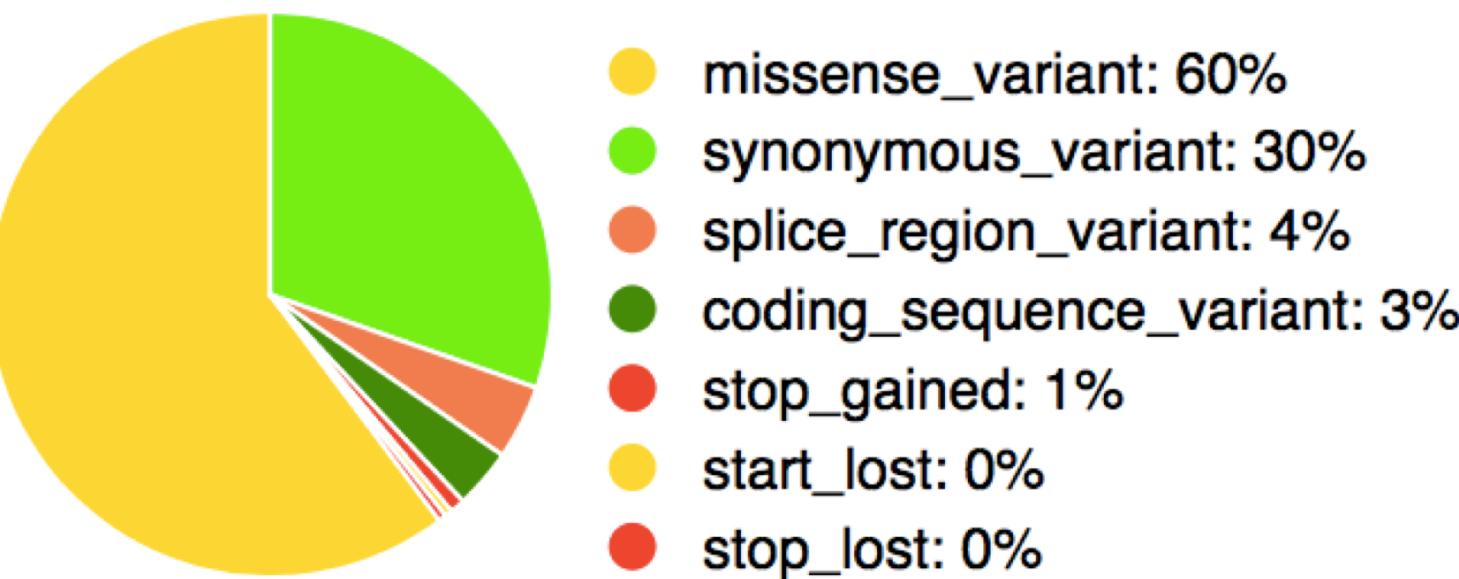
- synonymous\_variant: 57%
- missense\_variant: 39%
- splice\_region\_variant: 4%
- stop\_gained: 1%

FLECKVIEH



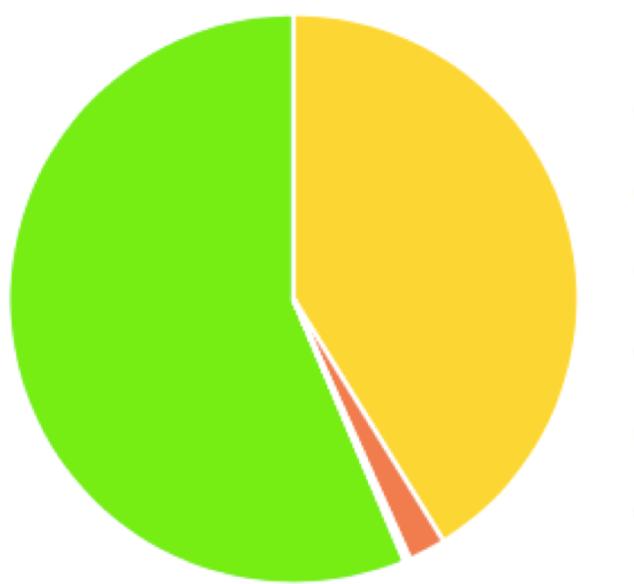
- missense\_variant: 55%
- synonymous\_variant: 44%
- splice\_region\_variant: 2%

HEREFORD



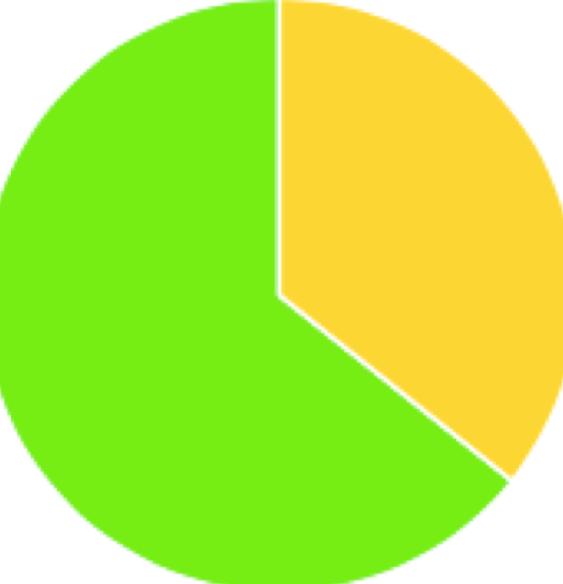
- missense\_variant: 60%
- synonymous\_variant: 30%
- splice\_region\_variant: 4%
- coding\_sequence\_variant: 3%
- stop\_gained: 1%
- start\_lost: 0%
- stop\_lost: 0%

LIMOUSINE



- synonymous\_variant: 56%
- missense\_variant: 41%
- splice\_region\_variant: 2%
- coding\_sequence\_variant: 0%
- start\_lost: 0%
- stop\_gained: 0%

SIMMENTAL



- synonymous\_variant: 64%
- missense\_variant: 36%

# Results

---



## ANGUS

- *SLC39A4* (BTA14) – *Acrodermatitis enteropathica*

## BROWN SWISS

- *SLC4A2* (BTA4) – *Osteopetrosis*
- *LAMA3* (BTA24) – *Epidermolysis bullosa*

## JERSEY

- *GON4L* (BTA3) – dwarfism
- *APOB* (BTA11) – cholesterol deficiency
- *TG* (BTA14) – familial goitre
- *COL7A1* (BTA22) - *Epidermolysis bullosa*

## LIMOUSINE

- *PFAS* (BTA19) – abortion

# Results

---



ANGUS

- iron content in mussels

BROWN SWISS

- Milk characteristics: milk solids, overall proteins, alpha-casein, alpha-lactalbumin, kappa-casein, beta-casein percentages, milk lactose content and yield
- somatic cell score
- growth traits: body weight, longissimus muscle area and gastrointestinal nematode burden

# Conclusions

---

- Differences observed between breeds express the dynamics of changes in the *Bos taurus* genome
- Breed specific reference genomes increase the accuracy of GWAS and SNP genotypes imputation

# Acknowledgement

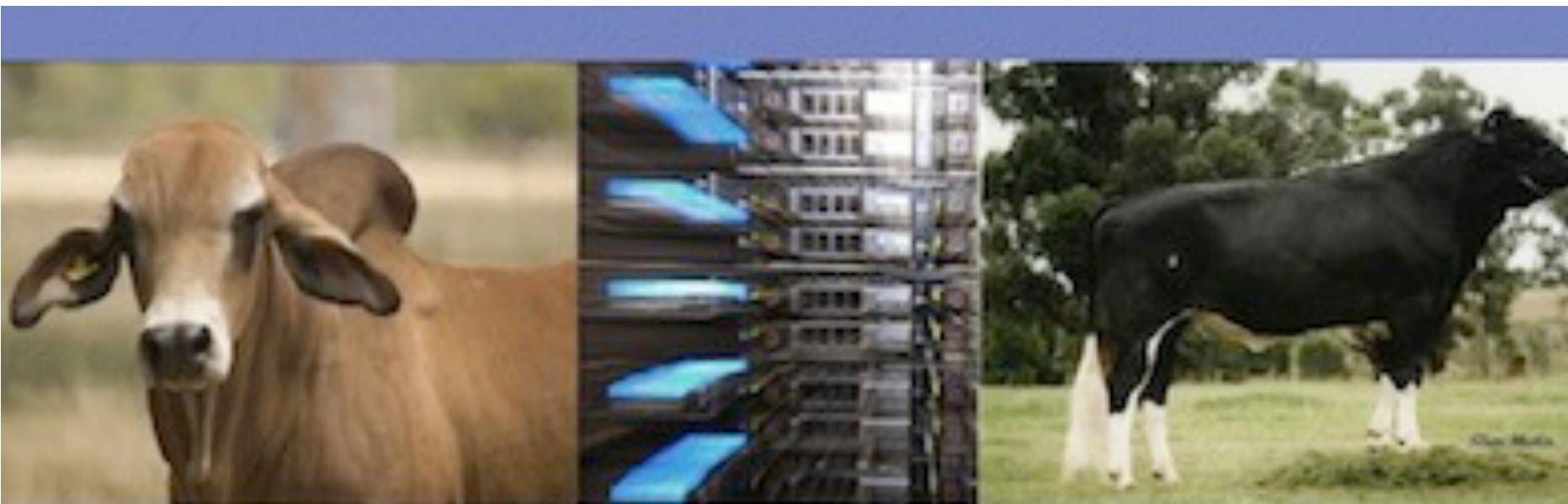


NATIONAL SCIENCE CENTRE  
POLAND



The Leading National  
Research Centre

Wroclaw Centre of Biotechnology 2014-2018



1000 bull genomes project

1000 bull genomes project consortium





Thank you for your attention!