

**EAAP 2016 Belfast**

Session 36  
Theatre presentation 10

***Fine mapping of mastitis resistance and milk production QTL in cattle by meta-analysis utilizing *Bos taurus* and *Bos indicus* data***

Bi-lateral Danish-Brazilian partnership

**Julia Gazzoni Jardim**

Corresponding author:

juliagazzoni@mbg.au.dk



# Background



- ↳ Guzerá -> Dual aptitude breed (Beef and Milk)
- ↳ 1992: Official dairy control
- ↳ 1994: Genetic evaluations of Guzerá bulls
  - ↳ Progeny test and MOET nucleus schemes
- ↳ 2000: Publishing breeding program
  - ↳ Milk and fat yields
  - ↳ Protein, lactose, total solids, type and functional traits




# Background



- 🐄 Famous about fat content in milk, robust and good maternal ability
- 🐄 QTL for mastitis in Nordic Holstein cattle was found, but not yet analyzed in Guzera Breed
- 🐄 New and important information for Indicus breed and cross animals

# Objectives

---

-  To carry out genome-wide association study milk component and mastitis in Guzera cattle using imputed high-density markers (770k)
-  To examine if QTL identified in Nordic dairy breeds are segregating in Guzera breed (*In progress*).
-  To fine-map the QTL segregating in both *taurus* and *indicus* population by meta-analysis (*In progress*).



## *Genotypic Data*

---

- 🐄 Illumina Chip data
  - 🐄 25 genotyped Guzera sires high-density (HD) marker data (770k)
  - 🐄 900 genotyped cows with medium-density marker data (50k)
  - 🐄 Reference population sequences from EMBRAPA

# *Workflow of the imputation*

---

Data set consisting of measured genotypes



Quality control: MAF, HWE, etc.



The reference set was used



Create the input files for imputation







Imputation using BEAGLE



Data set with imputed genotypes

# Phenotypic Data – Summary of data structure

-  Test day records – monthly sampling regimes
-  Milk, Fat, Protein yields and SSC
-  3 records for Milk and 1 record for other traits in each month
-  Lactation was defined as 5 to 305 Days-in-milk

Information	Milk	Fat	Protein	SSC
<b>RRM - EBV</b>				
N Records	6107	2161	1645	1644
<b>Association analysis</b>				
N animals with records	831	651	643	646
N animals pedigree	9988			
N herds	122			
Number of lactations (mean)	3			
Age at Calving (mean)	6			



# Random Regression Models

---

$$Y = X_b + \sum_{j=0}^{k-1} Z_j a_j + \sum_{j=0}^{k-1} Z_j p_j + \varepsilon$$

$Z_j$  :  $n \times q$  matrices for the 3th order of Legendre polynomial( $\phi$ )

$a_j$ : Vectors with random regression coefficients for all animals for additive genetic effects

$p_j$ : Vectors with random regression coefficients for all animals for permanent environmental effects

$\varepsilon$ : Residual effects

Fixed effects: Id, Calve age, Herd-Year, Year-Season, Parity, Location

Permanent environmental effects: Milk, Protein, SSC, Fat

Analysis was procedure in DMU using DMUai function

# *A Linear Mixed Model Analysis*

---

*Genome-wide association mapping*

$$Y = Xb + Za_i + Zp_i + e$$

*Y* Vector of vector of quantitative traits (phenotypes)

*b* is the effect size of the markers

*a<sub>i</sub>* Random regression coefficients for animal *i* for the additive genetic effects

*p<sub>i</sub>* Random regression coefficients for animal *i* for the permanent environmental effects

*X* an  $n \times 1$  vector of marker genotypes

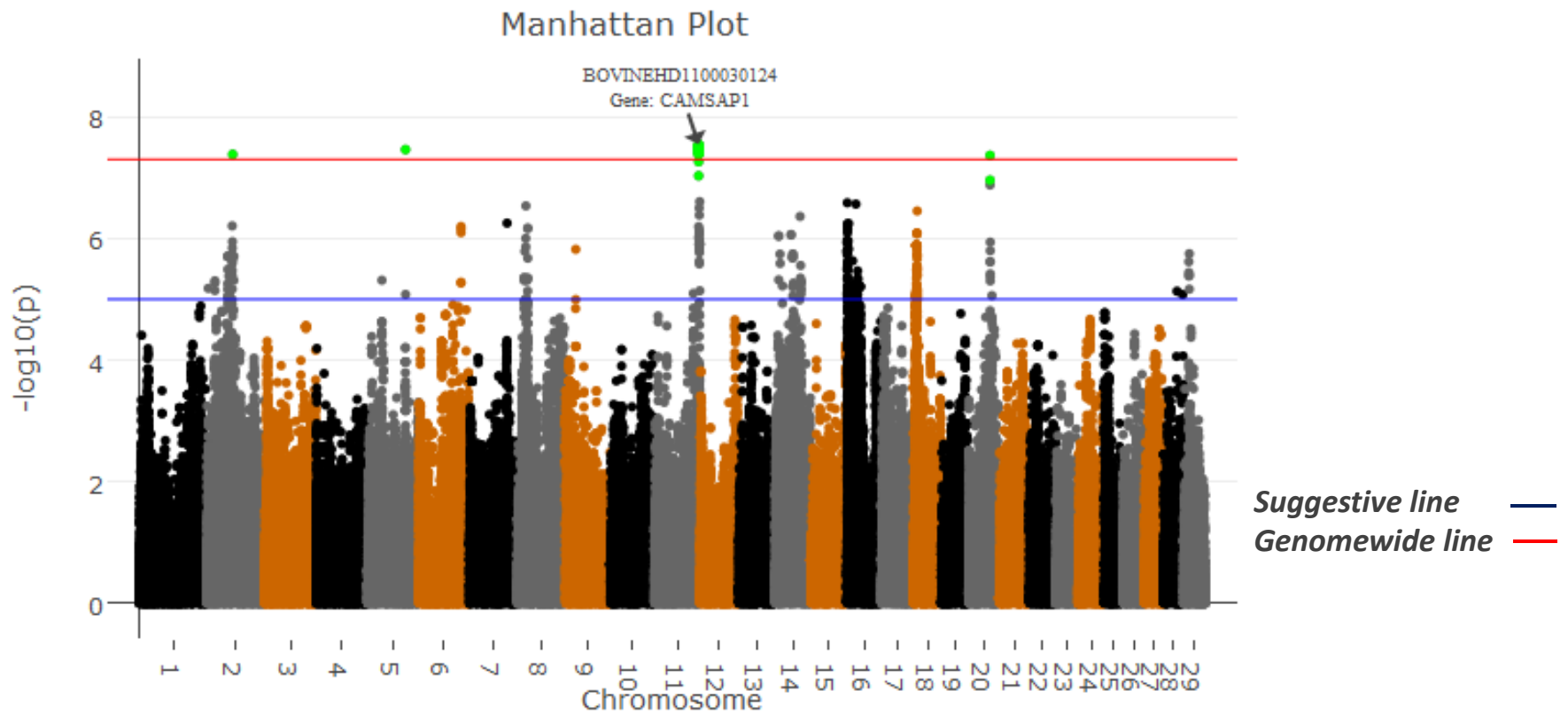
*Z* is an  $n \times m$  loading matrix

*e*: residual effects

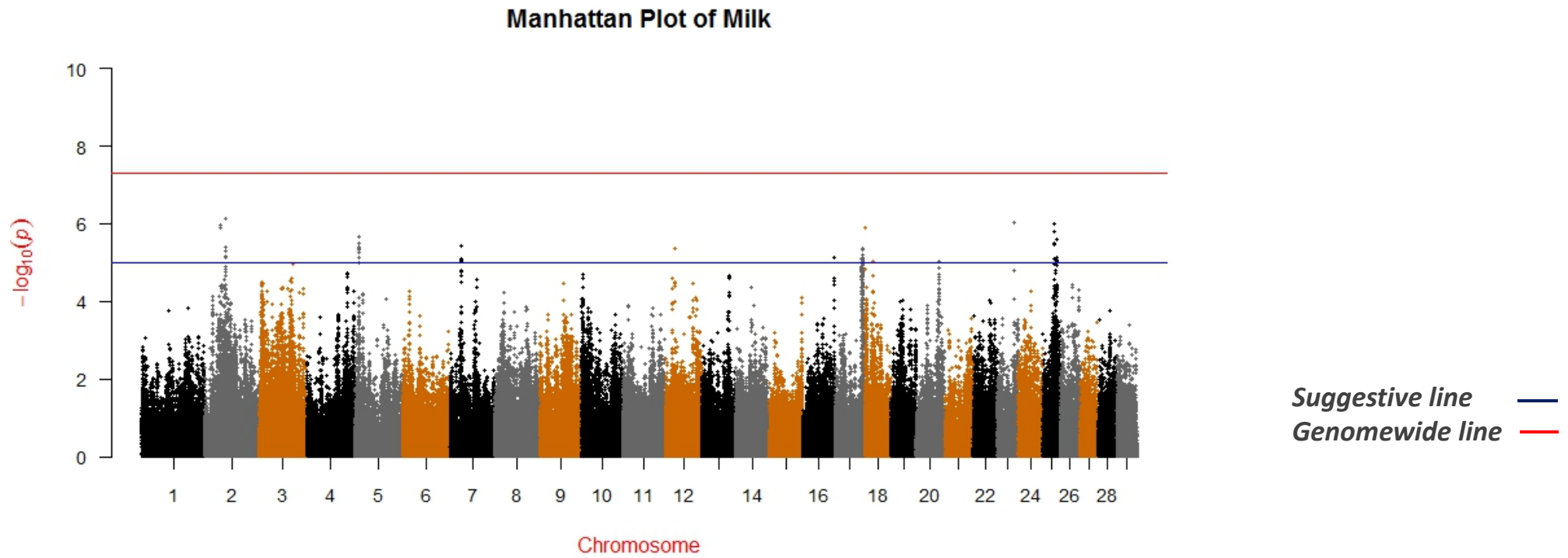
*Analysis was procedure in Rdmu*

# Results

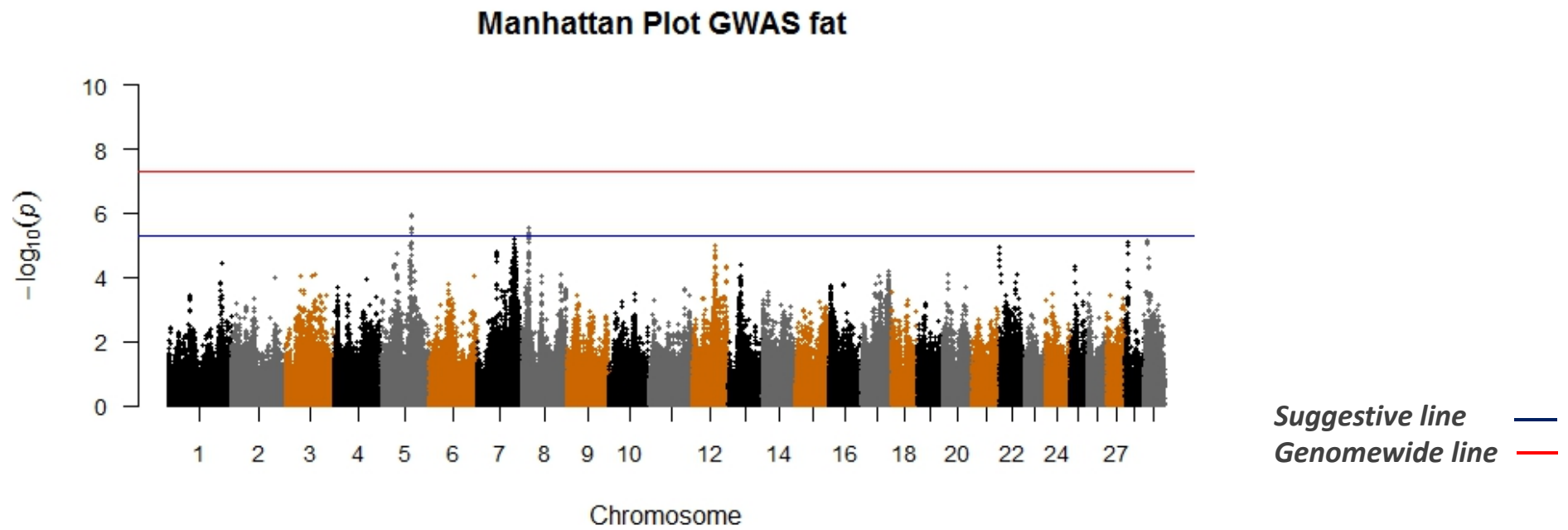
## Protein Yield



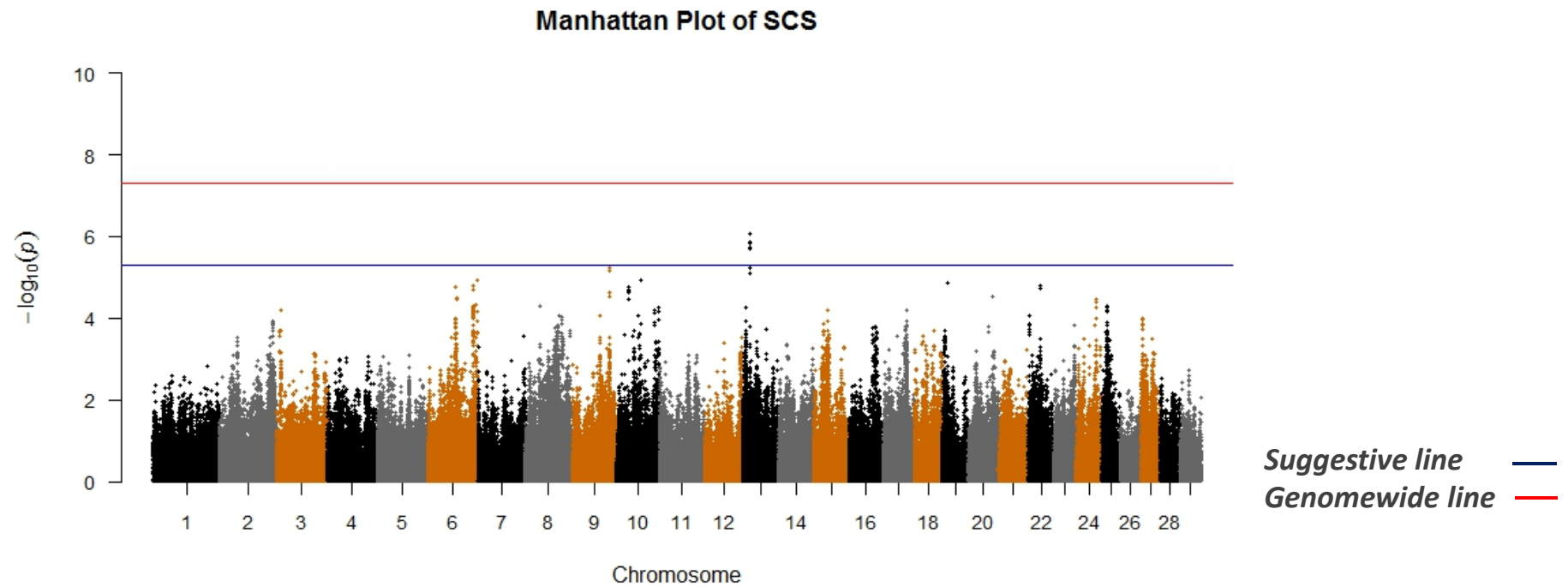
# Milk Yield



# Fat Yield



# Somatic Cell Score



# Conclusions

---



Were identified regions on chromosome 11 significantly associated SNPs that explain the phenotypic variance with protein yield in the studied population.



This genomic region have candidate gene with unknown functions.



Further investigation of these regions can provide greater understanding of the genetic control of milk production and components in Zebu.