11WCGALP
Auckland, New Zealand,
11–16 February 2018

# Heat stress tolerance indicators to be used as phenotypes in GWAS analyses: a comparison study in dairy cattle.

*M. Ramón[1], M.J. Carabaño[2] & C. Díaz[2]*

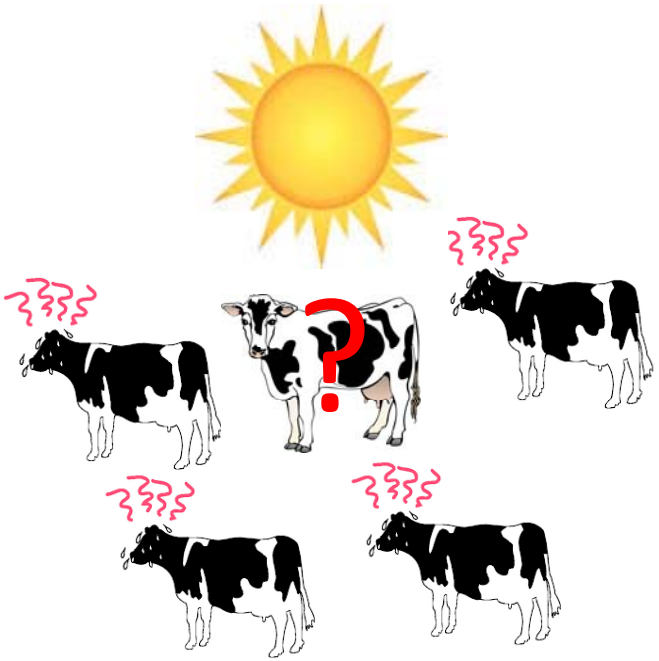
*[1] CERSYRA-IRIAF. Avda. del Vino 10, 13300, Valdepeñas, Ciudad Real, Spain*

*[2] Dpto. Mejora Genética Animal, INIA, Ctra. A Coruña Km. 7,5, Spain*

INIA
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria

CERSYRA
CENTRO REGIONAL DE SELECCIÓN Y REPRODUCCIÓN ANIMAL DE CASTILLA-LA MANCHA

# Justification




- Characterization of individual heat-tolerance is a key factor for the genomic characterization of heat tolerance (GWAS)
- Most common approach: use of routine milk recording information merged with meteorological information (Misztal, 1999).
- Problem: antagonistic relationship between production level and heat tolerance ➔ Challenge: Obtain measures of tolerance independent from mil production

Goal(s): **1. Compare alternative ways of measuring heat tolerance**

**(2. Compare the use of different methods of detection of genomic signals)**

# Material and Methods

# Data-Phenotypes

- Test day fat and protein yields from Southern Spain

| | μ | [1%-99%] | Cows | Records |
|---|---|---|---|---|
| **Fat (kg/d)** | **1.07** | **[0.4-2.0]** | **128,003** | **2,466,055** |
| **Protein (kg/d)** | **0.99** | **[0.4-1.6]** | **127,977** | **2,475,092** |
| Milk (kg/d) | 30.7 | [11 -52.1] | 128,112 | 2,514,762 |
| Fat (%) | 3.56 | [1.8-5.5] | 128,003 | 2,466,055 |
| Protein (%) | 3.27 | [2.6-4 ] | 127,977 | 2,475,092 |

- Daily temperature and humidity from nearest weather stations

| Heat load | Percentiles | | | | | |
|---|---|---|---|---|---|---|
| | 0% | 50% | 75% | 90% | 99% | 100% |
| **THIavg** | **19** | **59** | **68** | **72** | **76** | **81** |
| THImax | 32 | 66 | 74 | 78 | 82 | 89 |
| **Tavg (ºC)** | **-5** | **15** | **22** | **26** | **30** | **36** |
| Tmax (ºC) | -1 | 21 | 29 | 34 | 40 | 46 |

# Data: Pseudo_phenotypes (PF)

Random regression model (RRM) on heat load (Tave, THIave):

$$\mathbf{y} = \text{HTD} + \text{Lac-DIM} + LP_m(2) + \mathbf{LP_{animal}(2)} + \mathbf{e}$$

**y** is the test day fat or protein yield;

HTD, Lac-DIM are environmental factors that affect yield

$LP_m(2)$ is a quadratic curve representing the average reaction to heat load

**$LP_{animal}(2)$**: quadratic Legendre polynomial (bo, b1, b2) on animal effects (no relationships) representing individual response to heat load; var(**b**)=**B**

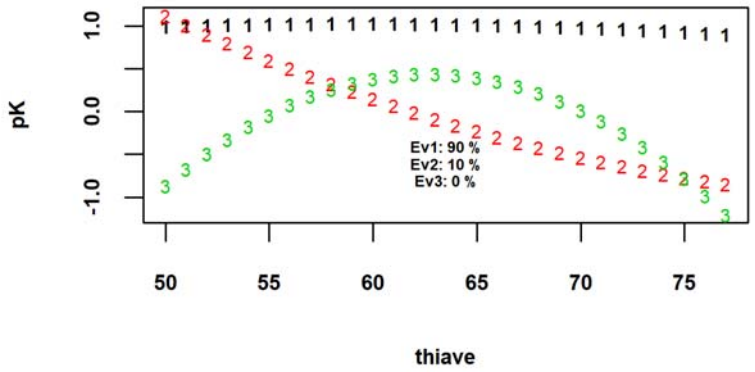
# Data: Pseudo_phenotypes (PF)

**Pseudo-phenotypes:**

A) Production level: **bo**

B) Tolerance:

- **b1** (linear coefficient)
- **Slopes** of the individual response curve at Tavg=26, THI=72,
- **Eigenvariables** eigen2, eigen3 (from **B**, covariance matrix from the RRM for animal effects)



# Data: Genotypes

- 50k Illumina bovine bead chip genotypes of CONAFE data base

- Quality control (using Plink 1.9)
  - Only autosomes (Chr1 to Chr29),
  - call rate animal=0.90,
  - call rate genotypes=0.90
  - MAF : 0.00001*

* to supress fixed markers. For others, SNP frequencies and HWE were checked after GWAS analysis

# Methods: GWAS

**1. Based on Pseudophenotypes (PF_GAbel)**

**p = marker + g + e**

p: pseudo-phenotype for production level and tolerance **for genotyped bulls** with > 6 daughters with phenotypes (**1380 Bulls**)

marker: SNP effect on p (one at a time)

g: polygenic effect on p (covariances=genomic relatioship)

e: residual term

Solved with **GenABEL package** (R Project)

**Relevant SNPs if FDR < 10%**

**2. Single-step GWAS (ssGWAS)**

**p = μ + g + e**

p: pseudo-phenotype for production level and tolerance **for cows** (**127,000 cows**)

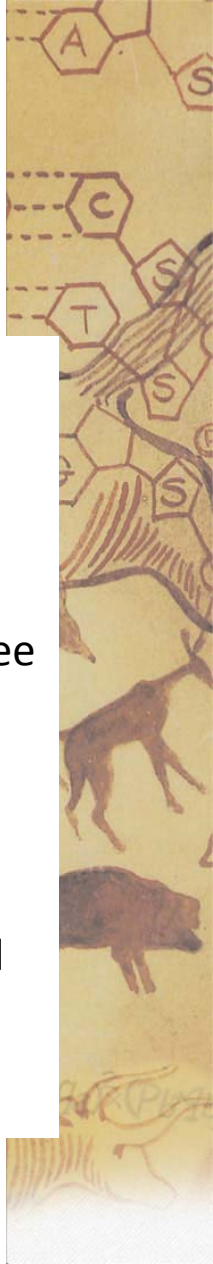
g: polygenic effect on p ( covariances = pedigree and genomic relationship)

e: residual term

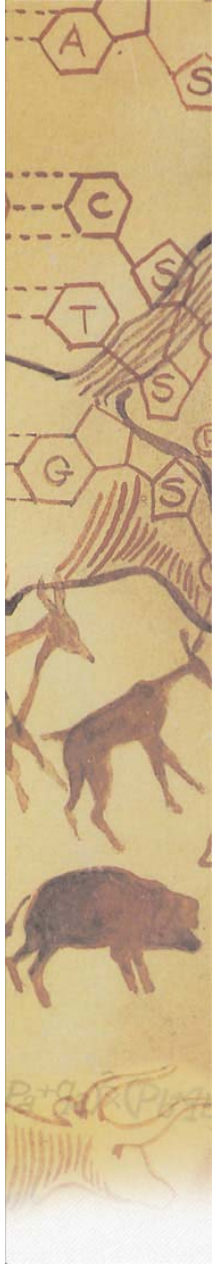
SNP effects and their associated variance is obtained from solutions for **g** and genotypes information.

Solved with **BLUPF90** family programs (Misztal et al, 2002)

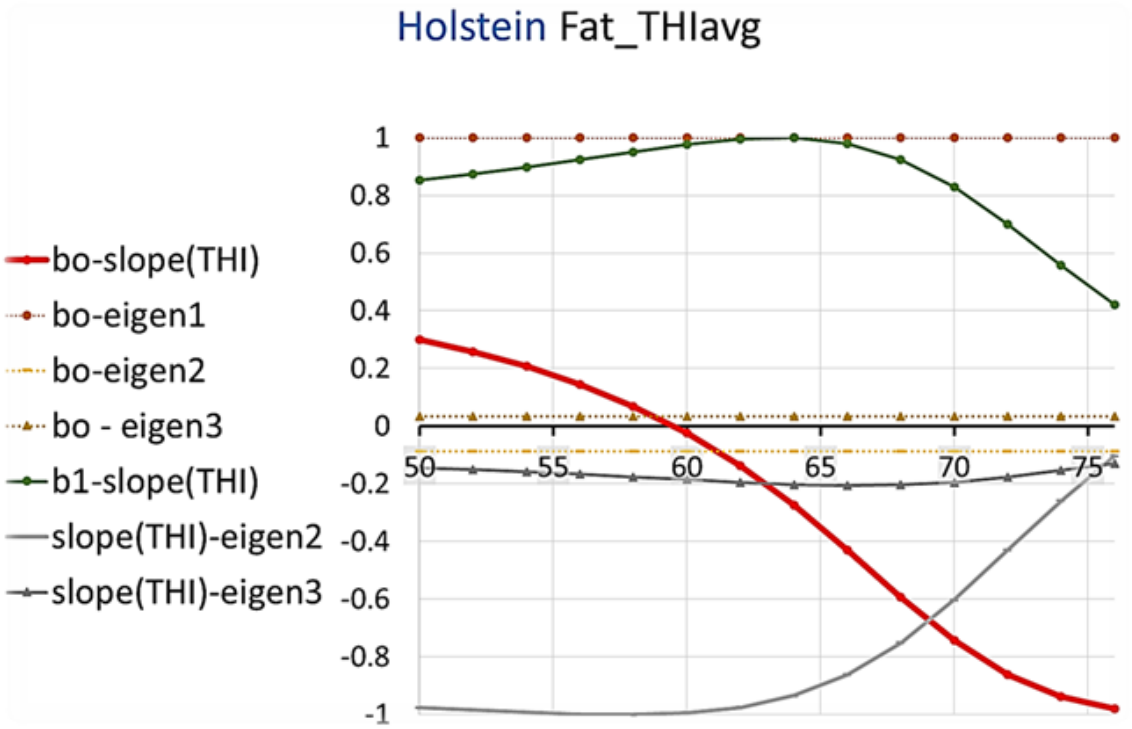
**Relevant SNPs if var(window20)>0.5%**

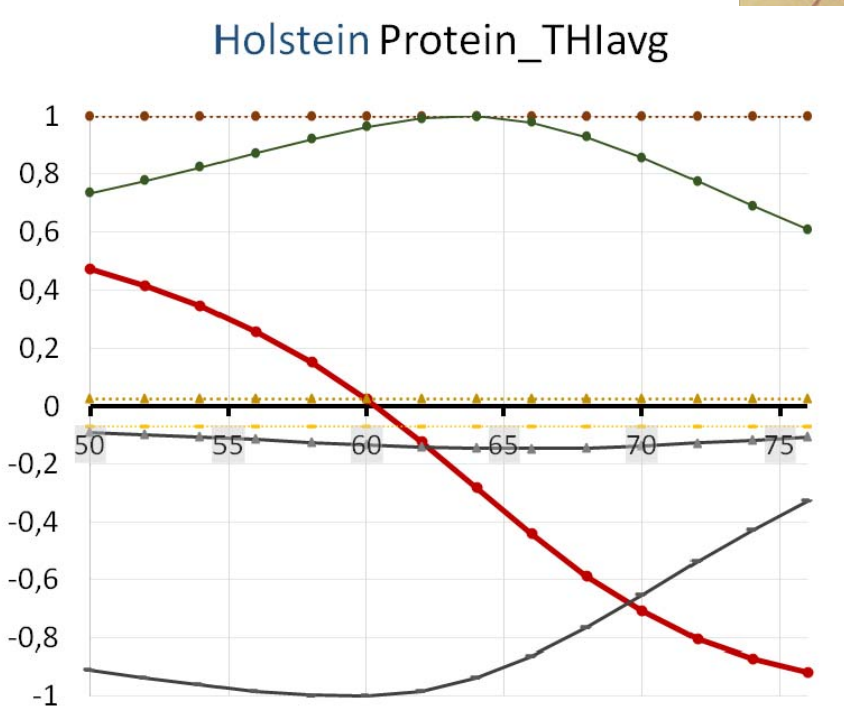


# Results



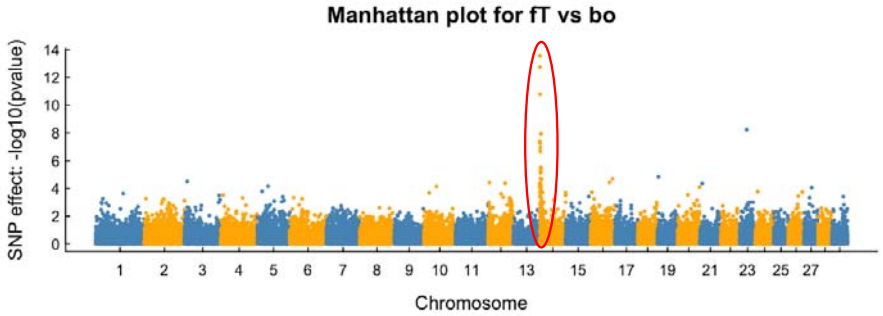
# Pseudo-phenotypes (p) correlations

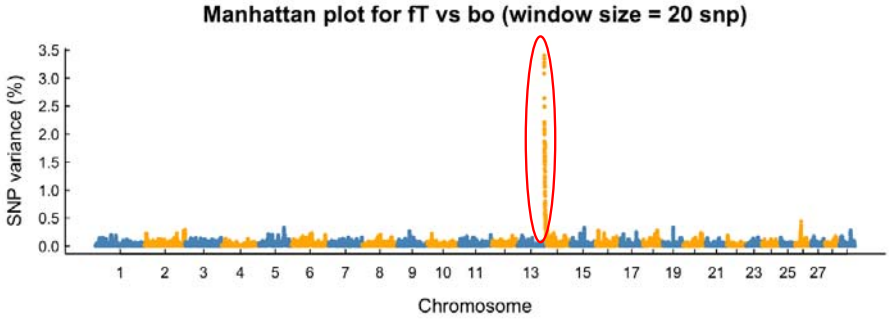


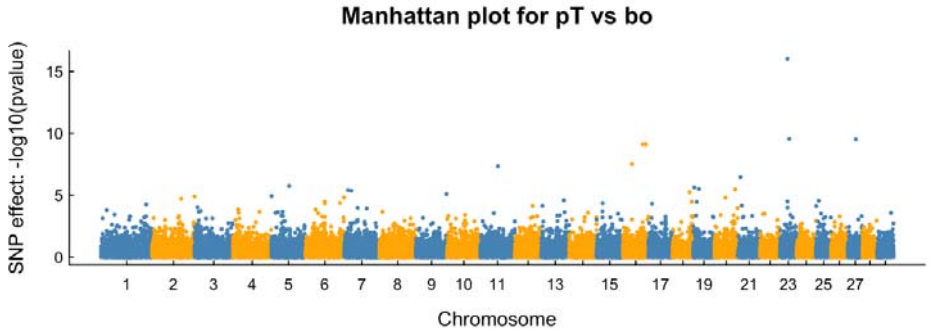



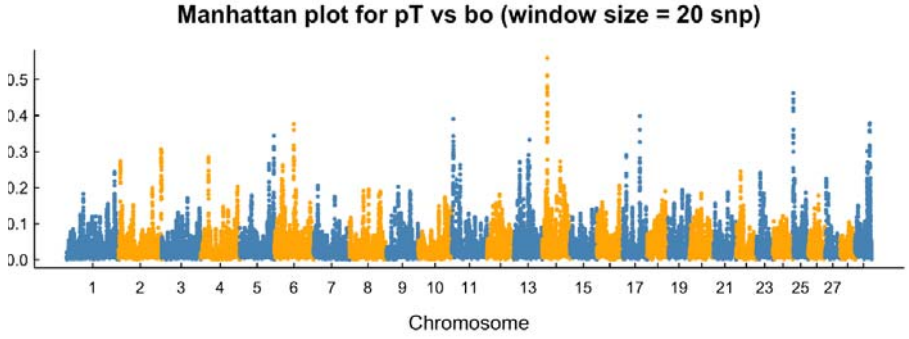
# Manhattan plots – Production level (bo)

## PF_Gabel vs.ssGWAS

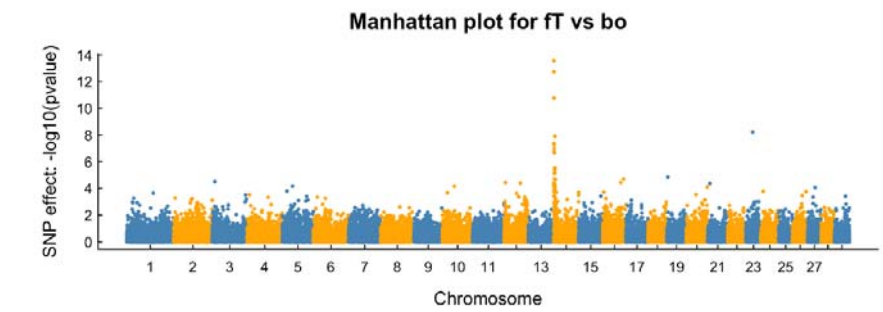


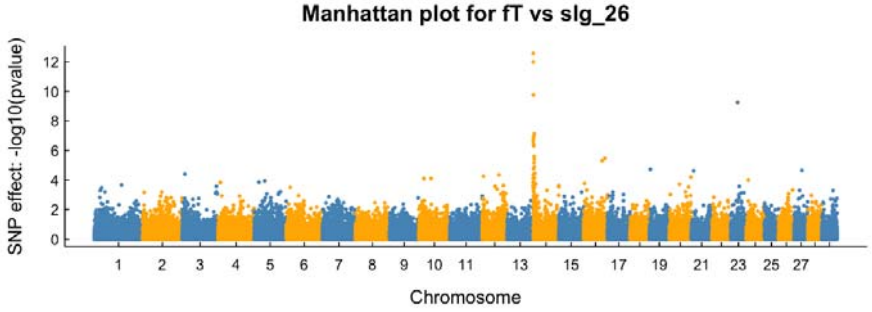



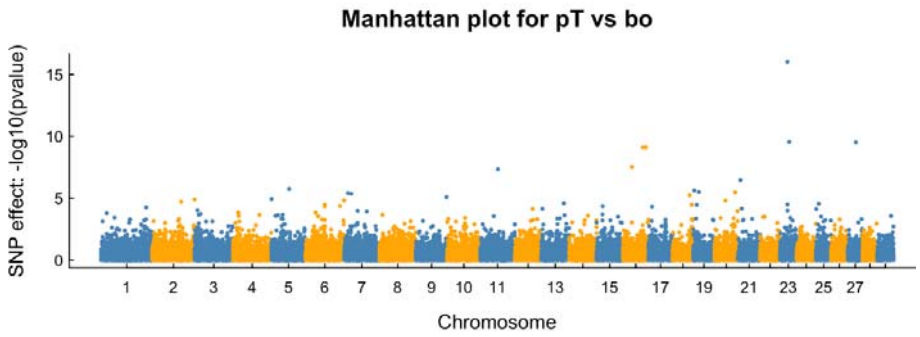



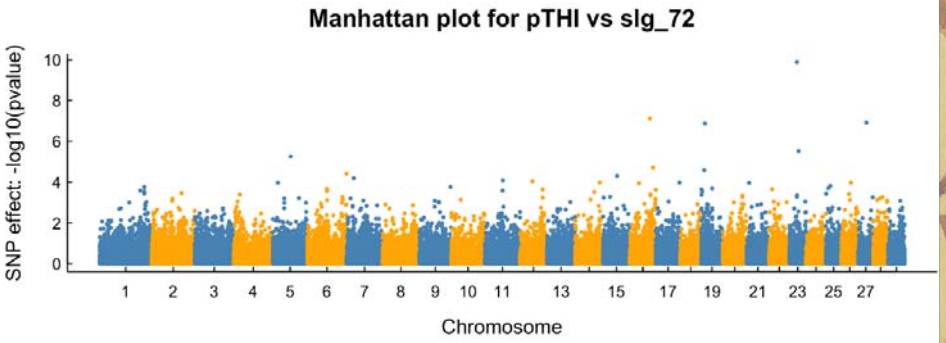



# SNP discovery: Level vs. Slopes of decay



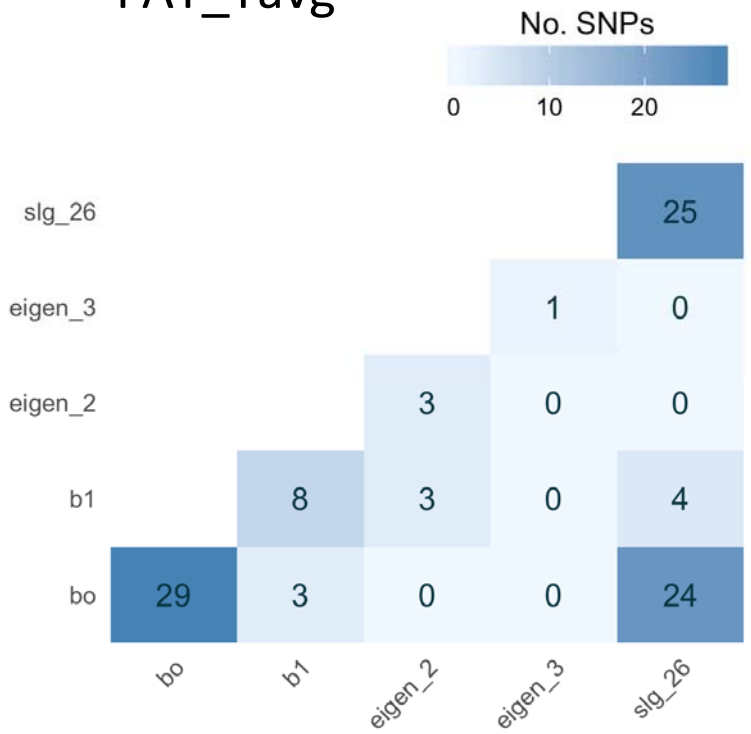



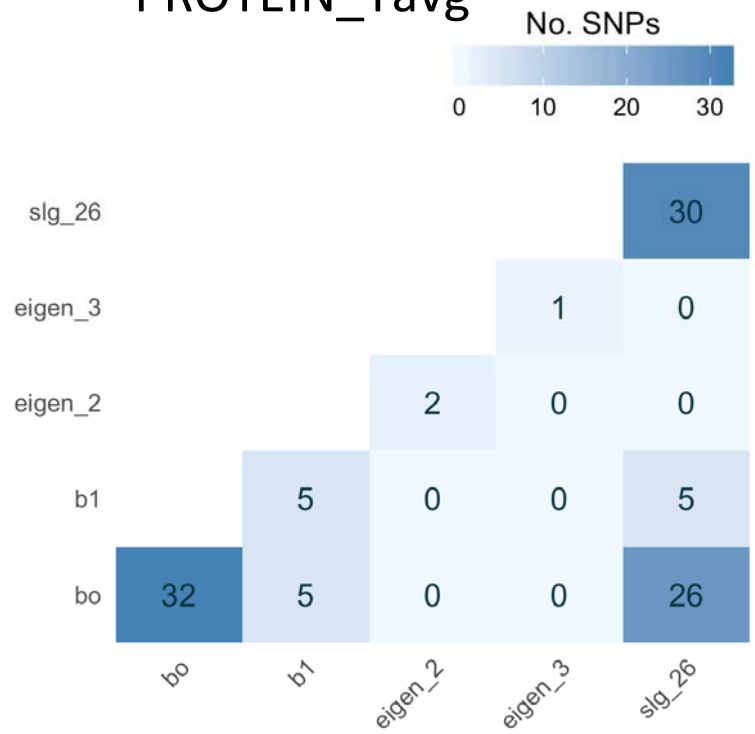



# SNP Discovery  Production vs. Tolerance - PF_GAbel

FAT_Tavg




PROTEIN_Tavg




Many relevant SNPs are shared by level of production and slope under heat stress, but none with eigens

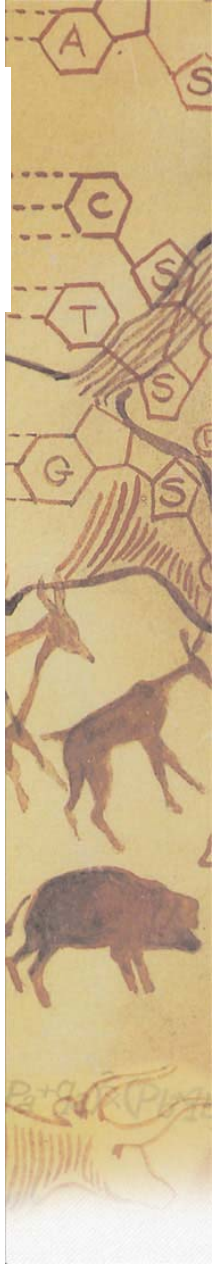
# Number of 'relevant' SNP signals for production level and heat tolerance for fat and protein using PF_Gabel (p-values) or ssGWAS (%variance) and number of common signals between methods

## FAT vs. PROTEIN

| | | Level* | Heat Tolerance* |
|---|---|---|---|
| Fat | PF_GAbel | 30 | 32 |
| | ssGWAS | 95 | 151 |
| | common | 21 | 16 |
| Protein | PF_GAbel | 34 | 33 |
| | ssGWAS | 17 | 68 |
| | common | 0 | 0 |

** Level=bo, Heat Tolerance= (b1, slp_26, slp_72, eigen_2, eigen_3)*

ssGWAS: Much larger number of signals for fat (Chr14-DGAT)



# Number of 'relevant' SNP signals for production level and heat tolerance for fat and protein using PF_Gabel (p-values) or ssGWAS (%variance) and number of common signals between methods

## PF_Gabel vs. ssGWAS

| | | Level* | Heat Tolerance* |
|---|---|---|---|
| Fat | PF_GAbel | 30 | 32 |
| | ssGWAS | 95 | 151 |
| | **common** | **21** | **16** |
| Protein | PF_GAbel | 34 | 33 |
| | ssGWAS | 17 | 68 |
| | **common** | **0** | **0** |

** Level=bo, Heat Tolerance= (b1, slp_26, slp_72, eigen_2, eigen_3)*

No common SNPs except for large signals (fat-Chr14_DGAT)

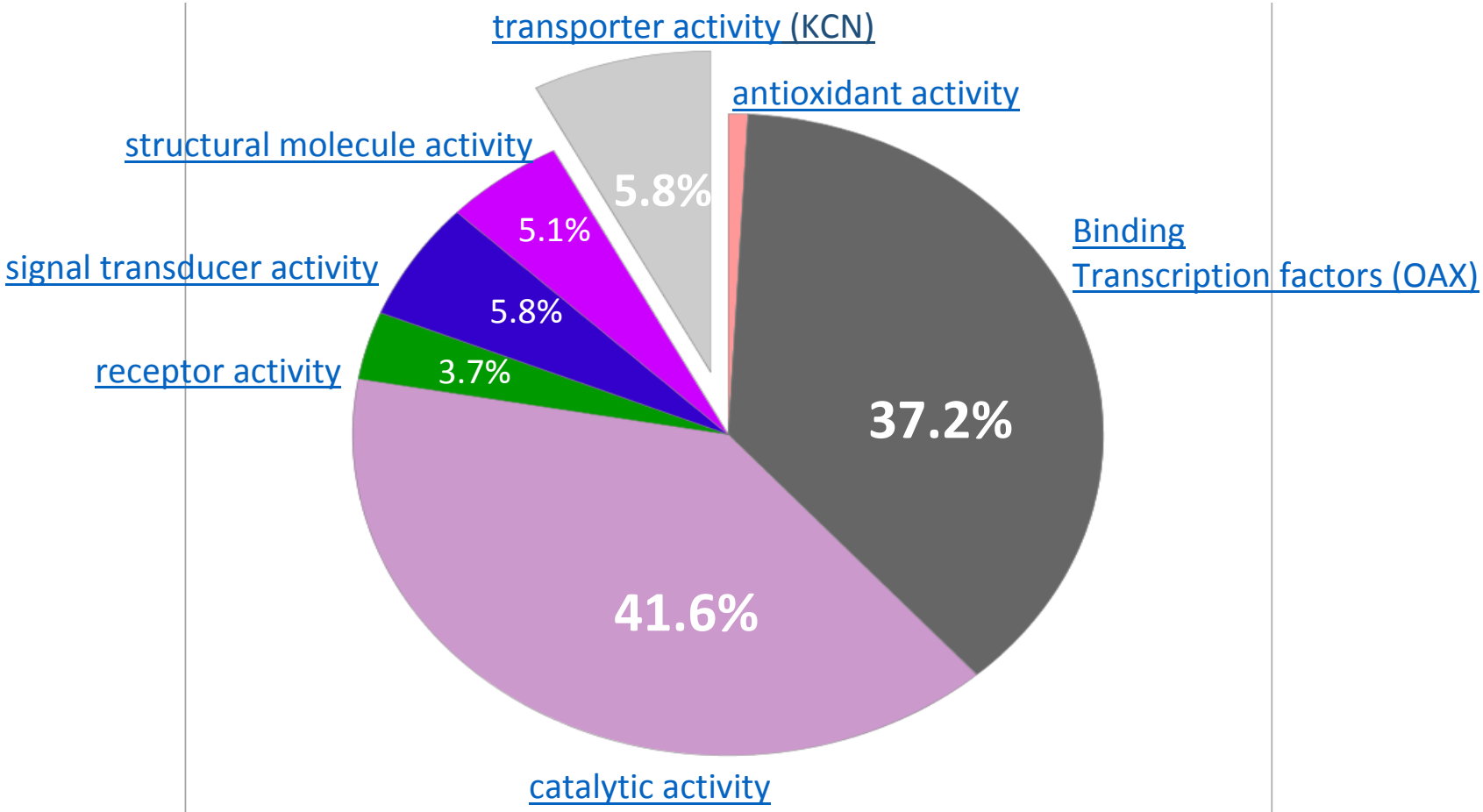
# Number of 'relevant' SNP signals for production level and heat tolerance criteria for fat and protein using PF_Gabel (p-values) or ssGWAS (%variance)

| | | | Heat Tolerance | | |
|---|---|---|---|---|---|
| | | Level | Linear | Slp_heat | **Eigen** |
| Fat | PF_Gabel | 30 | 8 | 25 | **6** |
| | ssGWAS | 95 | 58 | 98 | **40** |
| Protein | ssGWAS | 34 | 5 | 30 | **3** |
| | PF_GAbel | 17 | 25 | 23 | **33** |

** Level=bo, Linear= b1, Slp_heat=slp_26, slp_72, Eigen= eigen_2, eigen_3)*

**188 genes in +- 1Mb window**

# MOLECULAR FUNCTIONS OF POSITIONAL CANDIDATES GENES detected by the EIGEN VALUES as pseudo-phenotypes



# CONCLUSIONS

- Production level is a large component of heat tolerance in highly selected dairy cattle
- Eigenvariables gathering variability in fat and protein yields associated to changes in heat load explained a small proportion of the variability observed (10%), but, could still be used to detect genomic signals and candidate genes that influence response in milk quality to heat load without a relevant effect on production level.
- ssGWAS and PF_GenAbel picked rather different relevang SNP except for strong signals

# Thanks for your attention

- Ackowledgments
  - The study was financed by Project RTA2015-00108 (MINECO, Spain)
  - MEDGAN project