Royal Veterinary College
University of London

# Searching for genotype-by-environment interactions at the SNP level in Holstein dairy traits recorded in both China and UK

G. E. Pollott[1], J. J. Wu [2], J. S. Cooke[1], S. J. Zhang [2] and D. C. Wathes[1].

[1]Royal Veterinary College, Royal College Street, London, NW1 0TU, UK.

[2] College of Animal Sciences and Technology, Huazhong Agricultural University, Wuhan, Hubei Province, P. R. China.

# Motivation

- Genotype-by-environment interactions (GxE) have been widely studied in recent years with mature methodology
- Basic question – how does the same genotype perform in differing environments? Can we use this information to improve selection, management , etc.?
- Often thought of as between breed choices in different climates.
- More recent developments looked at random regression approach of individual genotypes (sires) on measureable differences between environments
- The use of SNP chips allows us to ask the question – how do animals with a particular SNP genotype perform in different environments – Equivalent of different QTL polymorphisms.

RVC

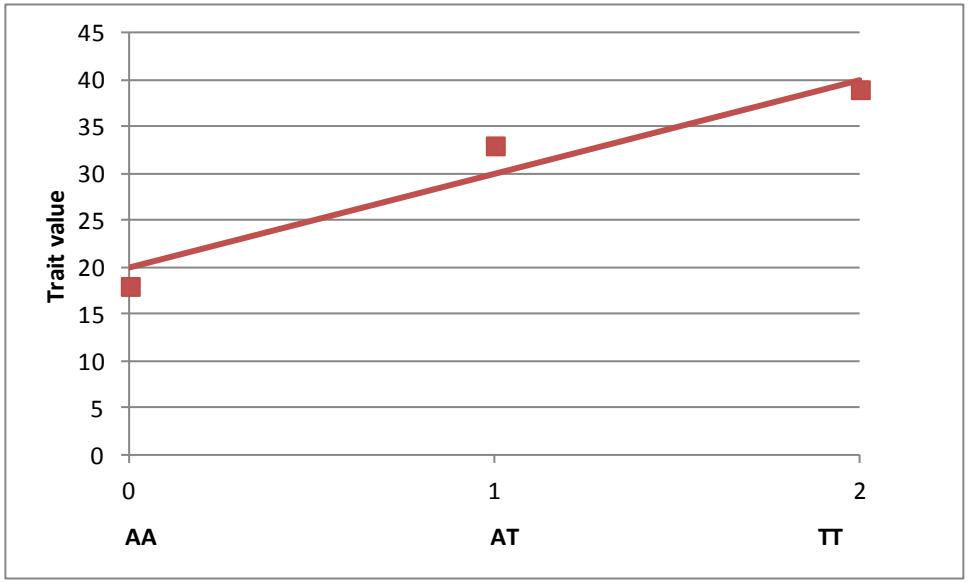
# GxE and SNP genotypes

- At each SNP on a chip we expect to see one of three possible genotypes in a population (say AA, AT, TT)
- Each allele may be in LD with an alternative form of a QTL
- Basic question – do the phenotypes of the 3 genotypes differ?  GWAS, GRAMMAR etc.
- GxE question -  do the phenotypes of the 3 genotypes perform differently in 2 or more environments?

# GxE and SNP genotypes - 2

- Under an additive model we would expect the phenotypic performance of AT to be (AA + TT)/2

- Using a regression approach, substituting 0 for AA, 1 for AT and 2 for TT, the slope of the regression line of phenotype on SNP genotype reflects the differences between the animals

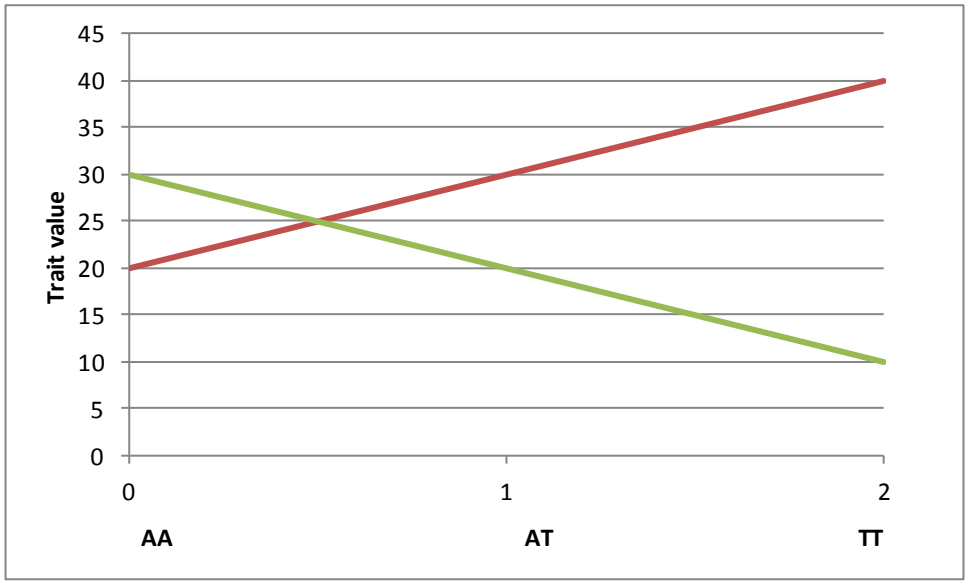



RVC

# GxE and SNP genotypes - 3

- In two environments, a comparison of the two within-environment regression slopes will tells us if the genotypes perform differently in each environment

- If they differ then this implies a GxE




RVC

# Data

- Holstein cows recorded for milk, fertility and other performance characteristics in both UK and China (Wu et al., 2012; Brickell et al., 2009)

- DNA taken from ~250 cows per country and genotyped with Illumina 50K SNP chip – after QC, 36,025 SNP analysed on chromosomes 1 – 29.

- Three traits selected – heifer 305-d milk yield (MY305), first calving interval (CI) and age at first calving (AFC).

RVC

# Analysis

- Environmental residuals derived for each trait (Amin et al., 2007) – residuals from fitting mixed polygenic model to data using ASREML
- Environmental residuals analysed by within-country regression of phenotype on coded genotype for each SNP
- Within-country regression slopes compared using Z score in PLINK ($[b_1-b_2]/s_{b1-b2}$; t test probability)

$b_1$ and $b_2$ are the two regression coefficients

$s_{b1-b2}$ is the standard error of the difference between the slopes

RVC

# Trait statistics

| Trait | Heifer MY305 (kg) | First calving interval (d) | Age at 1st calving (mo) |
|---|---|---|---|
| UK | 8,631± 99.4 | 414 ± 8.64 | 26.5 ± 0.21 |
| China | 6,299 ± 97.2 | 526 ± 6.96 | 28.6 ± 0.26 |

RVC

# Mixed models and polygenic results

All models (MY305, CI, AFC) included:

- Herd-year-season
- Principal components of the kinship matrix
- Polygenic animal term

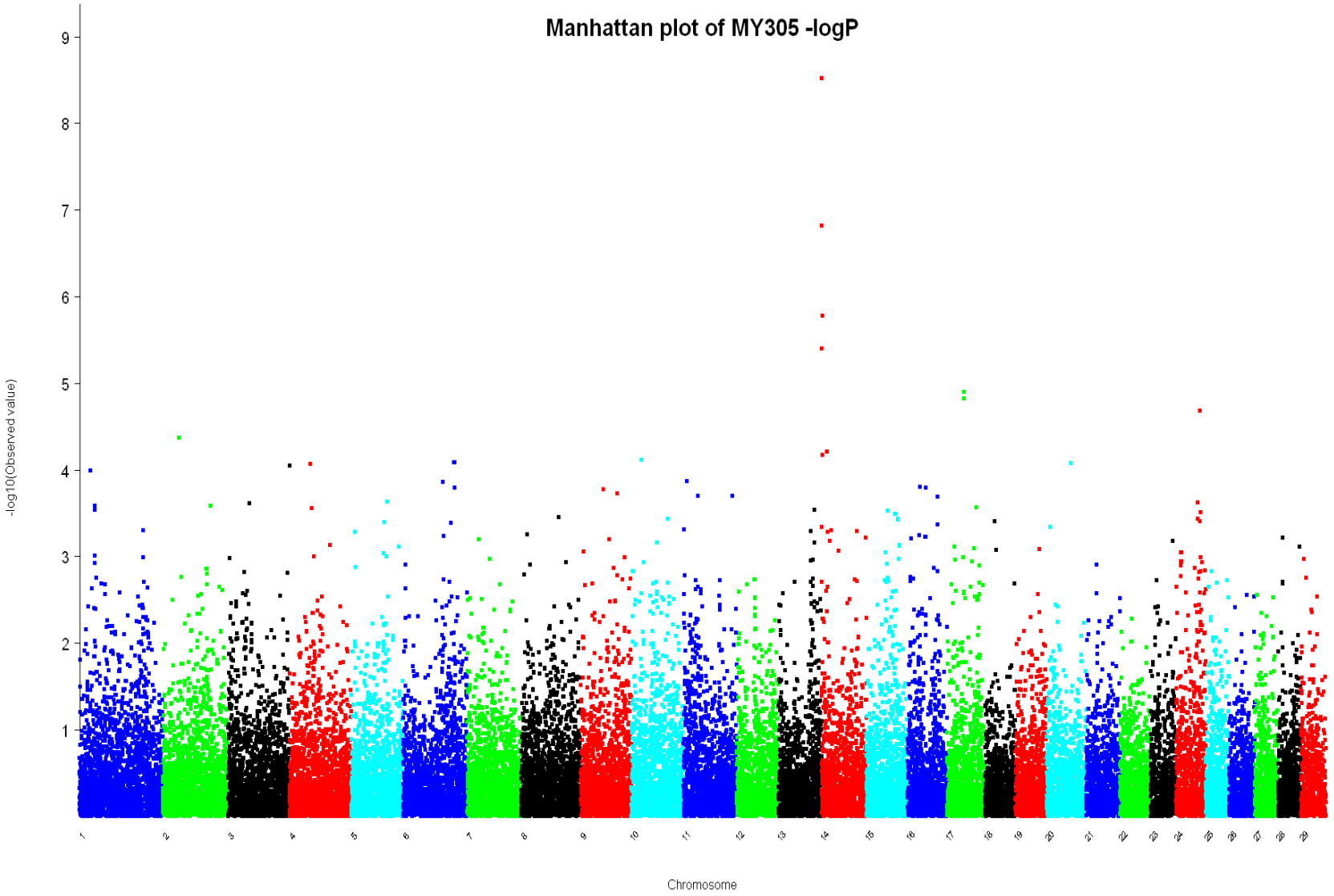
Model for MY305 included AFC

***Heritabilities***

| | |
|---|---|
| Heifer 305-d milk yield | 0.27 ± 0.14 |
| Calving interval | 0.08 ± 0.07 |
| Age at 1st calving | 0.21 ± 0.15 |

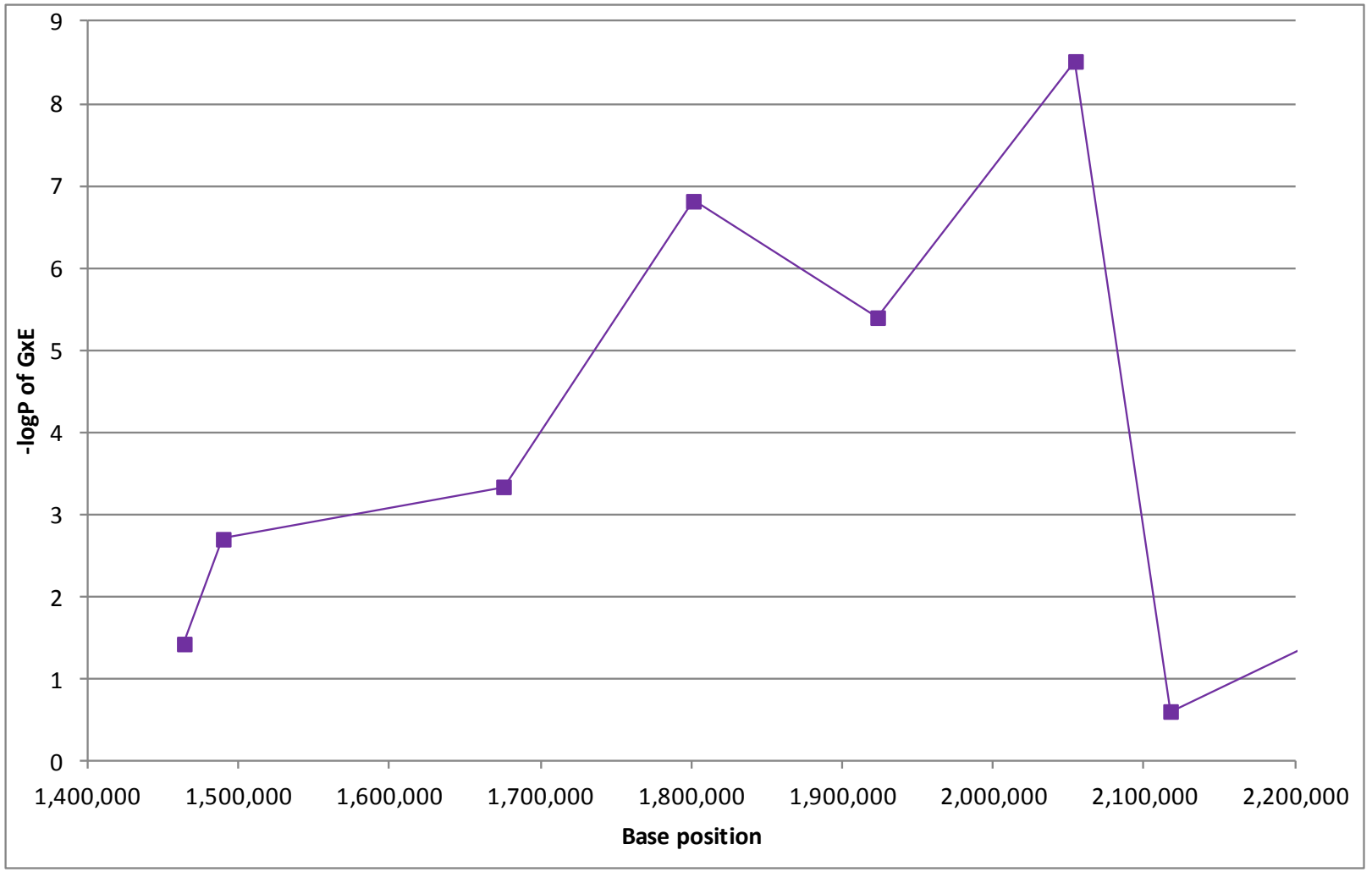
RVC

# Milk yield – SNP showing GxE




RVC

# Area on BTA 14 with high level of GxE




DGAT1 position – 1,795,425 -1,804,838bp

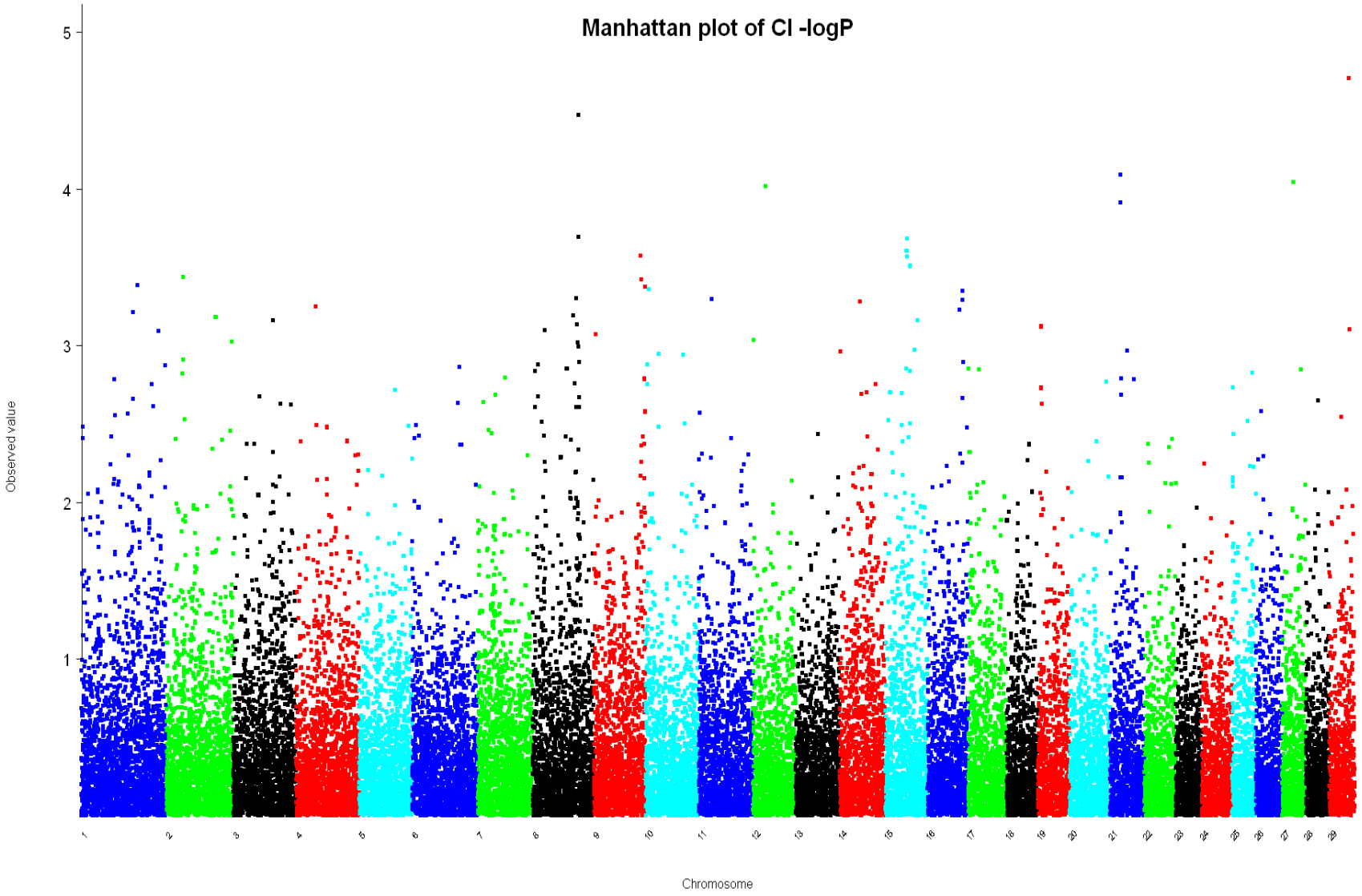
RVC

# Genotype by country means and regression coefficients – MY305 residuals

| Genotype<br>ARS-BFGL-NGS-107379 | 11 | 12 | 22 | Regression coefficient |
|---|---|---|---|---|
| UK | -247.6 ± 232 | 2.1 ± 101 | 160.7 ± 81 | 192.3 ± 96.5 |
| China | -1637.1 ± 308 | -592.9 ±104 | 440.1 ± 71 | 1035.0 ± 104.3 |

RVC

# Calving interval – SNP showing GxE



# Calving interval SNPs

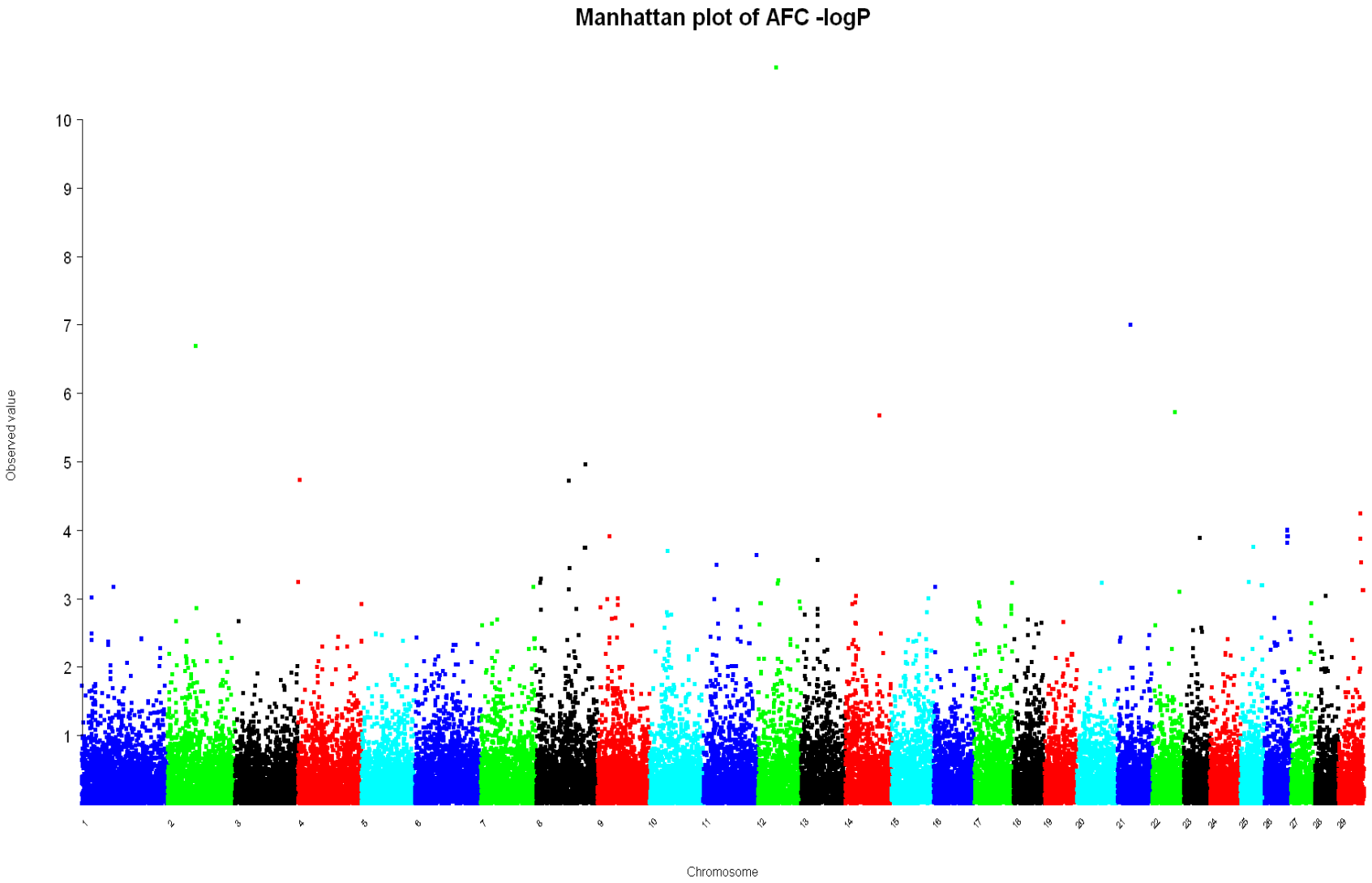
None achieved genome-wide significance (-logP = 5)

BTA 29 – two adjacent SNPs, -logP = 1.45, 4.71

BTA 8 – four adjacent SNPs, -logP = 1.79, 4.47, 2.33,1.93

RVC

# Age at 1st calving – SNP showing GxE



# Age at 1st calving SNPs

BTA 2 – Single SNP P = 6.69

BTA 12 – Single SNP P = 10.75

BTA 21 – Single SNP P = 6.99

BTA 21 – Single SNP P = 5.71

RVC

# Concluding remarks

Method works

Reveals some interesting results DGAT1

Larger numbers

More sophisticated models - dominance effects

- genotype as fixed effect

Functional genomics etc.

RVC

# Acknowledgements

Collaborating farmers for their time, data and cooperation

BBSRC through ICUK programme for funding

Thank you for your attention

RVC