Genome-wide associations for fertility using data from experimental herds in four countries

D.P. Berry[1] J.W.M. Bastiaansen[2], R.F. Veerkamp[2], S. Wijga[2], E. Wall[3], E. Strandberg[4] and M.P.L. Calus[2].
[1]*Animal & Grassland Research and Innovation Centre, Teagasc, Moorepark, Fermoy, Co. Cork.* [2]*Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, The Netherlands.* [3]*Scottish Agricultural College, Penicuik, The UK.* [4]*Swedish University of Agricultural Sciences, Uppsala, Sweden.*

Introduction
Genome-wide associations for difficult to measure traits are limited by sample population size with accurate phenotypic data. Fertility phenotypes using information on hormonal profiles are more heritable (Veerkamp et al., 2000) than traditional fertility measures thereby increasing the power of genome-wide association studies. The objective of this study was to use data on primiparous Holstein-Friesian cows from experimental farms in Ireland, the UK, The Netherlands and Sweden to identify genomic regions associated with fertility including a fertility phenotype derived from milk progesterone profiles.

Materials and Methods
Phenotypic data were available on 2,031 primiparous Holstein-Friesian cows from Ireland, 1,018 cows from the UK, 725 cows from The Netherlands, and 225 cows from Sweden. Sampling and determination of milk progesterone concentration have previously been described in detail for the data originating from Ireland (Horan et al., 2005), the UK (Pollot and Coffey, 2008), The Netherlands (Veerkamp et al., 2000) and Sweden (Petersson et al., 2006). Milk sampling was undertaken two to three times weekly between the years 1991 and 2005. The traditional fertility traits investigated were days from calving to first observed heat (CFH) or first service (CFS), calving interval (CIV), number of services (NS), and pregnancy rate to first service (PRFS). Post-partum interval to the commencement of luteal activity (PPCLA) was defined as the number of days from calving to the first occurrence of two consecutive test-day records with a milk progesterone concentration of $\geq$3 ng/ml. Genetic and residual (co)variances for the fertility traits were estimated using animal linear mixed models. Fixed effects were country-experimental treatment-year and country-year-season of calving. For PRFS, CFS was also included as a fixed effect.
Following the removal of animals that did not pass parentage verification using the genomic information, as well as the removal of single nucleotide polymorphisms (SNPs) that had a minor allele frequency of <0.01 in each country, deviated from Hardy-Weinberg equilibrium, or there was poor quality in calling the genotypes, a total of 37,590 SNPs from the Illumina Bovine50 Beadchip on 1,570 cows from Ireland (n=319), The UK (n=461), The Netherlands (n=583), and Sweden (n=207) remained. The genome-wide association analysis was conducted using a Bayesian Stochastic Search Variable Selection (BSSVS) model that estimates effects for all SNPs simultaneously. All univariate BSSVS models were run for 50,000 cycles, discarding the initial 10,000 cycles for burn-in (i.e., to remove the uncertainty of starting values provided). All bivariate models were run for 100,000 cycles, discarding 20,000 for burn-in.

Results and Discussion
Heritability estimates for the traditional fertility traits varied from 0.03 (PRFS) to 0.16 (CFH). The heritability of PPCLA was 0.13. The interval traits (i.e., CFH, CFS, CIV and PPCLA) were all strongly genetically correlated (0.37 to 0.99) with each other. The posterior QTL probabilities for the traditional fertility traits were all less than 0.021. Posterior probabilities of >0.04 were observed for PPCLA on BTA2 (BTA-49769-no-rs; probability of 0.060) and BTA21 (BTA-12468-no-rs; probability of 0.045). The SNP on BTA2 explained 0.51% of the genetic variance in PPCLA while the SNP on BTA20 explained 0.35% of the genetic variance in PPCLA. The Bayes factors of BTA-49769-no-rs and BTA-12468-no-rs were 24 and 18, respectively. The posterior QTL probability of 0.060 for PPCLA at SNP BTA-49769-no-rs estimated in the univariate analysis increased to 0.094, 0.121, 0.162, 0.662 and 0.162 when included in a bivariate analysis with CFH, CFS, NS, CIV and PRFS, respectively. The posterior probability of 0.045 for PPCLA at SNP BTA-12468-no-rs on BTA20 when estimated in the univariate analysis increased to 0.052, 0.152, 0.072, 0.123 and 0.135 when included in a bivariate analysis with CFH, CFS, NS, CIV and PRFS, respectively

Conclusions
Regions of the genome associated with PPCLA were identified although no obvious region was associated with the traditional fertility measures. This suggests that genome wide associations may be more successful if phenotypes derived from physiological measures, less influenced by management, are used. The inclusion of additional information through bivariate analysis increased the QTL probabilities.

Acknowledgements
This study is part of the RobustMilk project (which is financially supported by the European Commission under the Seventh Research Framework Programme, Grant Agreement KBBE-211708.

References
Horan, B, Mee, J.F., O'Connor, P., Rath M., & Dillon, P. (2005) *Therio. 63: 950-961*
Petersson, K.-J., Strandberg, E., Gustafsson, H., & Berglund, B. (2006). *Anim. Reprod. Sci. 91:201-214*
Pollot, G.E. & Coffey, M.P. (2008). *J. Dairy Sci. 91: 3649-3660*
Veerkamp, R.F., Oldenbroek, J.K., Van Der Gaast, H.J. & Van Der Werf J.H.J. (2000). *J. Dairy Sci. 83:577–583*