

Genome-wide association study and gene network analysis of fertility, retained placenta, and metritis in U.S. Holstein cattle (**Paper 610**)

J.B. Cole,*¹ K.L. Parker Gaddis,² D.J. Null,¹ C. Maltecca,³ and J.S. Clay⁴

¹Animal Genomics and Improvement Laboratory, ARS, USDA, Beltsville, MD, USA

²Council on Dairy Cattle Breeding, Bowie, MD, USA

³Department of Animal Science, College of Agriculture and Life Sciences, North Carolina State University, Raleigh, NC, USA

⁴Dairy Records Management Systems, Raleigh, NC, USA

*john.cole@ars.usda.gov

Introduction

- **3 fertility traits recently dissected using single- and multiple-trait genomewide association studies (GWAS) in all-bull, all-cow, and mixed predictor populations (Parker Gaddis et al., 2016)**
- **Network analysis identified several important genes not identified by GWAS alone**
- **Genetic evaluations for 6 Holstein health traits to be introduced in the U.S. in April 2018 (Parker Gaddis et al., 2017)**
- **Retained placenta (RETP) and metritis (METR) included as measures of reproductive health**

Introduction *(continued)*

- Longer days open and lower conception rates for cows starting lactation with RETP or METR (e.g., Fourichon et al., 2000)
- Influence by common sets of genes not known for susceptibility to reproductive tract diseases and cow fertility
- Analyses with combined analytical approaches used to better understand challenging phenotypes such as bull fertility (e.g., Han & Peñagaricano, 2016)

Objectives

- Identify **genes, genomic regions, and gene networks** associated with **3 fertility traits**
 - Daughter pregnancy rate (**DPR**)
 - Heifer conception rate (**HCR**)
 - Cow conception rate (**CCR**)and **2 reproductive health traits** (METR and RETP) using **producer-reported data** from U.S. Holstein cows
- Determine if common sets of genes influence susceptibility to **reproductive tract diseases and cow fertility**

Phenotypic and genomic data

- December 2016 genomic PTAs for DPR, HCR, and CCR combined with METR and RETP genomic PTAs calculated as by **Parker Gaddis et al. (2014, 2017)**
- Genotypes included **60,671** SNPs from routine U.S. evaluations included in genotypes
- Predictor populations included animals with PTA reliabilities for lifetime net merit > parent average reliability
- Animals required to have PTAs for all traits
- All **35,724** bulls retained and random sample of **35,000** cows drawn from predictor population

- **Model for 5-trait multivariate GWAS (Parker Gaddis et al., 2016):**

$$Y = \mu + x\beta' + U + E$$

- Y is $n \times 5$ matrix of phenotypes for n individuals
- μ is the intercept
- x is n -vector of marker genotypes
- β' is transpose of 5-row vector of marker effect sizes
- U is $n \times 5$ matrix of random effects
- E is $n \times 5$ matrix of errors
- Variance structures as described in **Zhou (2016)**

SNP annotation

- Autosomal markers assigned to closest gene within 25 kbp using BEDTools (v2.21.0) (Quinlan & Hall, 2010)
- Gene information from Bovine UMD3.1.1 genome assembly (Zimin et al., 2009)
- **36,435** markers remained after merging with annotation data
- SNPs from GWAS with Wald P -value $>5 \times 10^{-8}$ selected for further analysis
- Gene function determined by literature review

Enrichment analyses

- All SNPs with P -values of <0.05 compared with all annotated genes in bovine genome
 - Gene ontology (GO) (Ashburner et al., 2000)
 - Medical subject heading (MeSH) (Morota et al., 2015)
- GO and MeSH term analyses carried out in R (v3.4.0) with GOSTATS (v1.5.3) and meshr (v1.12.0) packages as distributed in Bioconductor (v3.5)

Association weight matrices and PCIT

- Association weight matrix (AWM) constructed as implemented by Fortes et al. (2010).
- Row-wise partial correlations computed on AWM using pcit package (v1.5-3) in R (v3.4.0) to produce *m*-symmetric adjacency matrix
- Non-significant values set to 0; significant correlations interpreted as significant gene-gene interactions.
- Only connections with partial correlation of ≥ 0.98 included
- Visualization using Cytoscape (v3.2.1) (Shannon et al., 2003)

GWAS results

- **43** and **11** SNPs significant in bull and cow populations, respectively
- Developmental, cell-signaling, and protein modification processes represented in both populations
- Top SNPs did not overlap between populations

GWAS results *(continued)*

Group	SNP	Chr	Location	Gene	Function	$-\log_{10}(P)$
Bulls	BTB-00790451	20	57,373,160	<i>FBXL7</i>	Ubiquitination	44.67
	ARS-BFGL-NGS-64415	18	48,486,442	<i>ECH1</i>	Fatty acid degradation	41.43
	ARS-BFGL-NGS-72630	6	118,871,663	<i>SORCS2</i>	Nervous system development	21.88
	BTB-00259343	6	62,642,435	<i>BEND4</i>	Longevity	15.02
	Hapmap55409-rs29022997	4	33,236,485	<i>CROT</i>	Lipid metabolism	12.77
Cows	ARS-BFGL-NGS-23066	6	92,153,394	<i>CDKL2</i>	Sex differentiation	13.26
	BTB-00062715	1	135,269,426	<i>EPHB1</i>	Cell signaling	9.08
	BTB-00176697	4	40,934,520	<i>SEMA3C</i>	Embryonic development	8.02
	ARS-BFGL-NGS-111133	4	119,341,142	<i>UBE3C</i>	Ubiquitination	7.96
	ARS-BFGL-NGS-36082	17	55,916,203	<i>KDM2B</i>	Ubiquitination	7.45

GO and MeSH results

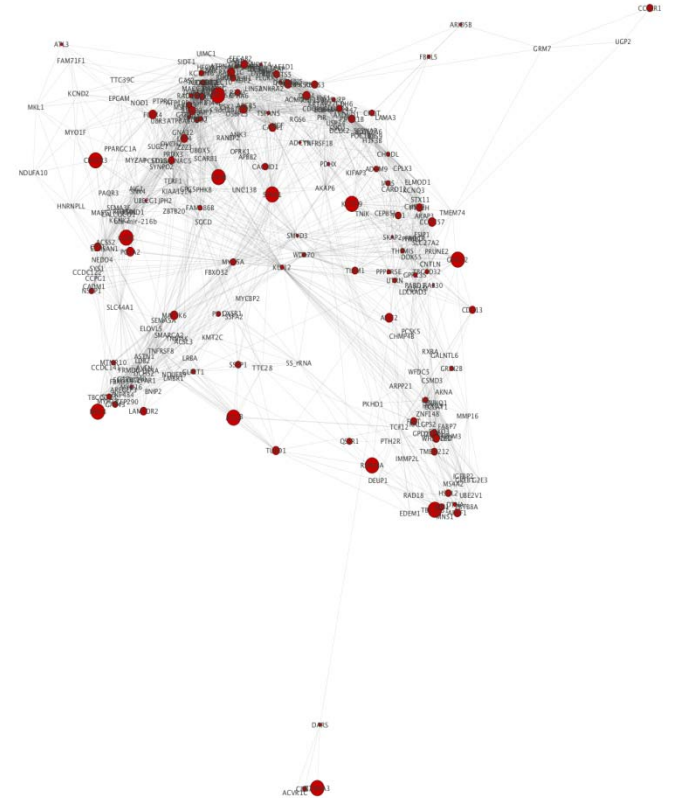
- **GO terms from Biological Processes category**
- **Enriched bull processes included spermatogenesis and DNA processing**
- **Enriched cow pathways included embryonic development and gene expression**

GO and MeSH results *(continued)*

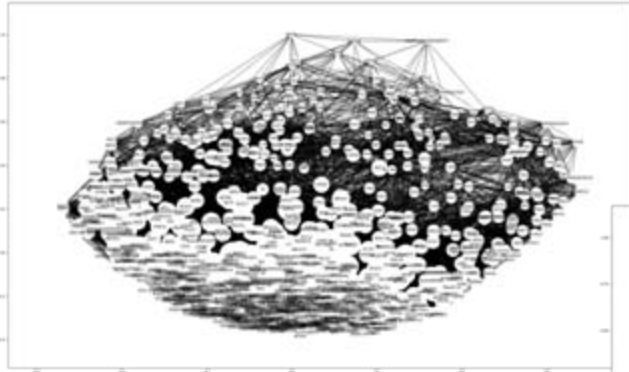
Group	GO			MeSH		
	GO ID	Term	P-value	MeSH ID	Term	P-value
Bulls	0006270	DNA replication initiation	0.005	D002970	Cleavage stage, ovum	0.004
	0007288	Sperm axoneme assembly	0.014	D003599	Cytoskeleton	0.032
	0051661	Maintenance of centrosome location	0.014	D009210	Myofibrils	0.035
	1902979	Mitotic DNA replication termination	0.014	D013116	Spinal cord	0.035
	0007283	Spermatogenesis	0.036	D042541	Intracellular space	0.036
Cows	2000738	Positive regulation of stem cell differentiation	0.016	D002823	Chorion	0.034
	0070126	Mitochondrial translational termination	0.024	D009092	Mucous membrane	0.043
	2000637	Positive regulation of gene silencing by miRNA	0.024	—	—	—
	0048701	Embryonic cranial skeleton morphogenesis	0.039	—	—	—
	0060147	Regulation of posttranscriptional gene silencing	0.039	—	—	—

Gene network results

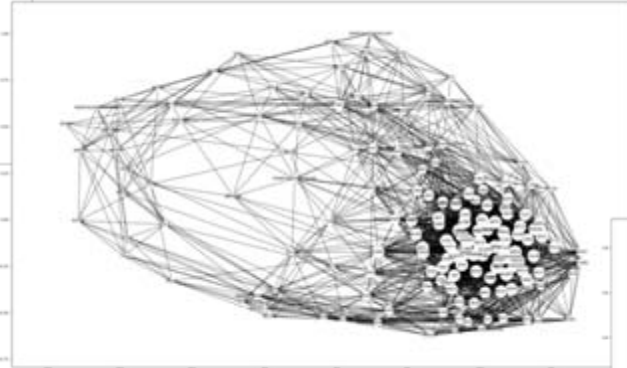
- Gene networks included **824** genes for bulls and **856** genes for cows, with **139** shared genes
- Many highly connected genes associated with male or female fertility and embryo size and morphology
- None of 100 SNPs explaining largest amount of GWAS variance were among most connected network genes



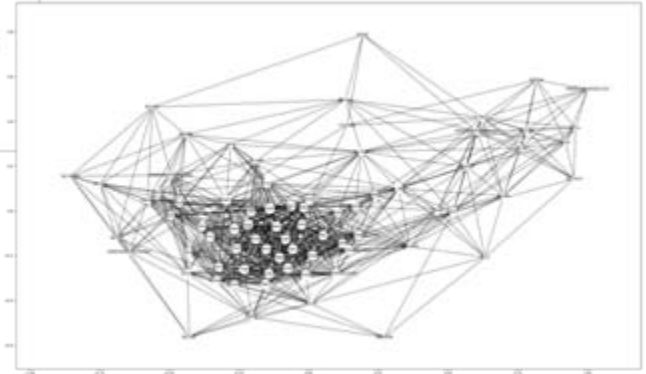
Thinning dense networks (**cows**)



$P \leq 0.05$
980 nodes
71,315 edges



$P \leq 0.01$
158 nodes
2,730 edges



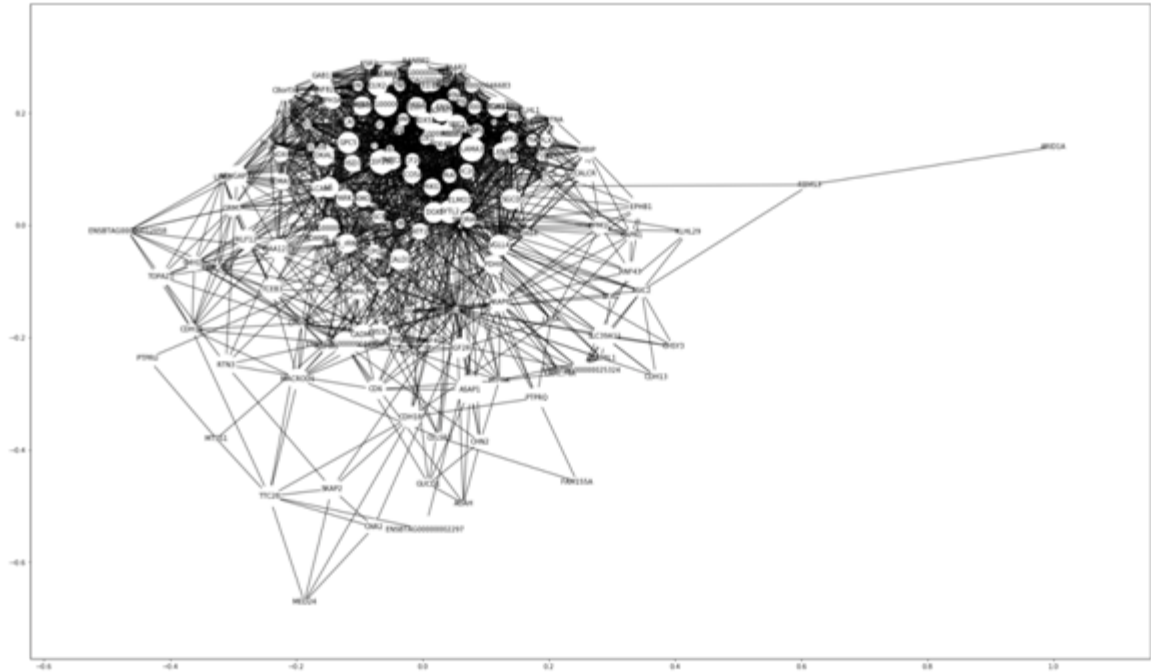
$P \leq 0.005$
75 nodes
806 edges

Key genes in thinned networks (**cows**)

$P \leq 0.05$		$P \leq 0.01$		$P \leq 0.005$	
Gene	Degree	Gene	Degree	Gene	Degree
<i>DOCK1</i>	589	<i>ENSBTAG0000000834</i> 6	72	<i>ENSBTAG0000000834</i> 6	36
<i>U6</i>	573	<i>5S_rRNA</i>	70	<i>EPHA7</i>	35
<i>PRKG1</i>	511	<i>CEP164</i>	66	<i>ENSBTAG0000002698</i> 6	35
<i>NTM</i>	482	<i>HOMER1</i>	66	<i>KMO</i>	34
<i>WDR62</i>	473	<i>UACA</i>	65	<i>BBS4</i>	34
<i>SERGEF</i>	440	<i>TBC1D22A</i>	63	<i>ANTXR1</i>	33
<i>SLC2A13</i>	432	<i>USP34</i>	62	<i>COL8A1</i>	33
<i>TRIO</i>	422	<i>COMMD1</i>	61	<i>COMMD1</i>	33

Combining sex-specific networks

- **Constructed consensus network using only nodes (150) and edges (2,086) shared by both bull and cow networks**
- **Minimal overlap of high-degree nodes with thinned networks**




Conclusions

- Individual SNPs associated with fertility were identified, and enriched pathways included some fertility terms
- Bull- and cow-specific gene networks similarly included genes with known effects on fertility
- No significant loci had any obvious associations with reproductive tract health (as measured by METR or RETP, which have relatively low heritabilities)
- A case-control study with paired animals could provide greater power for identifying SNPs and co-expression networks associated with both reproductive health and fertility

Acknowledgments & disclaimers

- **Dairy Records Management Systems (Raleigh, NC)** provided health data.
Council on Dairy Cattle Breeding (Bowie, MD) provided fertility evaluations.
Cooperative Dairy DNA Repository (Madison, WI) provided genotypes.
- **USDA-ARS project ARS 8042-31000-002-00, “Improving dairy animals by increasing accuracy of genomic prediction, evaluating new traits, and redefining selection goals”**
- **Mention of trade names or commercial products in this presentation is solely for the purpose of providing specific information and does not imply recommendation or endorsement by USDA; USDA is an equal opportunity provider and employer**

Questions?



AIP web site:
<http://aipl.arsusda.gov/>

Holstein and Jersey crossbreds graze on American Farm Land Trust's Cove Mountain Farm in south-central Pennsylvania

Source: ARS Image Gallery, image #K8587-14; photo by Bob Nichols

11th World Congress on Genetics Applied to Livestock Production, Auckland, New Zealand, Feb. 12, 2018 (20)

References

- Ashburner M., et al. 2000. [Gene Ontology: Tool for the unification of biology](#). *Nat. Genet.* 25:25–29.
- Fortes, M.R.S., et al. 2010. [Association weight matrix for the genetic dissection of puberty in beef cattle](#). *Proc. Natl. Acad. Sci. U.S.A.* 107:13642–13647.
- Fourichon, C., et al. 2000. [Effect of disease on reproduction in the dairy cow: A meta-analysis](#). *Theriogenology* 53:1729–1759.
- Han Y., & F. Peñagaricano. 2016. [Unravelling the genomic architecture of bull fertility in Holstein cattle](#). *BMC Genet.* 17:143.
- Morota, G., et al. 2015. [An application of MeSH enrichment analysis in livestock](#). *Anim. Genet.* 46:381–387.
- Parker Gaddis, K.L., et al. 2014. [Genomic selection for producer-recorded health event data in US dairy cattle](#). *J. Dairy Sci.* 97:3190–3199.
- Parker Gaddis, K.L., et al. 2016. [Explorations in genome-wide association studies and network analyses with dairy cattle fertility traits](#). *J. Dairy Sci.* 99:6420–6435.
- Parker Gaddis, et al. 2017. [Development of genomic evaluations for direct measures of health in US Holsteins and their correlations with fitness traits](#). *J. Dairy. Sci.* 100(Suppl. 2):378(abstr. 377).
- Quinlan, A.R., & I.M. Hall. 2010. [BEDTools: A flexible suite of utilities for comparing genomic features](#). *Bioinformatics* 26:841–842.
- Shannon, P., et al. 2003. [Cytoscape: A software environment for integrated models of biomolecular interaction networks](#). *Genome Res.* 13:2498–2504.
- Zhou, X. 2016. [GEMMA User Manual](#). University of Chicago, Chicago, IL.
- Zimin, A.V., et al. 2009. [A whole-genome assembly of the domestic cow, *Bos taurus*](#). *Genome Biol.* 10:R42.