1 **Identification of an immune modulation locus utilising a bovine mammary gland**
2 **infection challenge model**

3

4 Mathew D. Littlejohn[1], Sally-Anne Turner[2*], Caroline G Walker[2,3], Sarah D Berry[3], Kathryn
5 Tiplady[1], Ric G Sherlock[1], Greg Sutherland[4], Simon Swift[5], Dorian Garrick[6], S. Jane Lacy-
6 Hulbert[2], Scott McDougall[7], Richard J Spelman[1], Russell G Snell[8], and J. Eric Hillerton[2]

7

8 [1]LIC, Hamilton, New Zealand
9 [2]DairyNZ, Hamilton, New Zealand
10 [3]Growing up in New Zealand, University of Auckland, Auckland, New Zealand
11 [4]School of Medical Sciences, University of Sydney, Sydney, Australia
12 [5]School of Medical Sciences, University of Auckland, New Zealand
13 [6]Massey University, New Zealand
14 [7]Cognsoco, Anexa Animal Health, Morrinsville, New Zealand
15 [8]School of Biological Sciences, University of Auckland, Auckland, New Zealand

16

17 **Supplementary file**

18

19

20

21

Supplementary Materials and Methods

Animals

Animal experimentation was approved by the Ruakura Animal Ethics Committee (Hamilton, New Zealand) in accordance with the New Zealand Animal Welfare Act (1999). The trial animals consisted of a subset from a population of 864 F_2 Holstein-Friesian x Jersey cross breed dairy cows of known pedigree, representing half-sib offspring of six F_1 sires (Spelman *et al.*, 2001). This population comprised two cohorts, bred over successive years and located on the same research farm.

Pre-trial microbiology and treatments

Five weeks prior to commencement of the trial, the bacteriological status of milk from all animals was assessed to identify and treat pre-existing infections. Milk samples from each mammary gland were collected aseptically, and tested for the presence of mastitis pathogens, according to National Mastitis Council guidelines (Hogan *et al.*, 1999). Glands with a foremilk somatic cell count > 1×10^6 cells/mL and the presence of >100 cfu/mL of the mastitis pathogens *Staphylococcus aureus*, *Streptococcus uberis*, or *Streptococcus dysgalactiae*¸ or more than 1000 cfu/mL of coagulase negative staphylococci (CNS) spp., were treated with intramammary antibiotics. For cows considered to be infected with *Str. uberis* or CNS, the treatment was 1,000,000 IU of penicillin (Penalone, MSD, Upper Hutt, New Zealand) on three occasions at 24 h intervals; for cows considered infected with *S. aureus or Str. dysgalactiae*, the treatment was 200 mg of cloxacillin (Orbenin LA, Zoetis, Auckland, New Zealand) on five occasions at 24 h intervals. Of the 96 cows treated, 46 were treated for *S. aureus*, 24 for a combination of *S. aureus* and *Str. dysgalactiae*, 9 for *Str. uberis*, 8 for CNS and 4 for a combination of *S. aureus* and *Str. uberis*.

46 Milk samples were also collected for bacteriology three days prior to the experimental *Str.*
47 *uberis* challenge. Some 134 (22%) of all cows were found infected with a mastitis pathogen
48 (*S. aureus*, *Str. uberis*, CNS or *Corynebacterium bovis*) in one of more glands. Of these 134
49 cows, 87 had not been treated five weeks previously whilst 47 had been treated previously.
50 The observed prevalence of intramammary infection was similar to previous estimates for
51 mid-lactation cows, managed on pasture in New Zealand (Lopez-Benavides *et al.*, 2006). To
52 account for potential effects of previous infection and antibiotic use on experimental
53 infection, appropriate sub-classifications were incorporated into the phenotypic model for
54 QTL analysis.

55

56 *Assignment of challenge, mastitis diagnosis and treatment*

57 The infection challenge was staged over two successive years such that animals within each
58 cohort were the same age and at approximately the same stage of lactation when inoculations
59 were administered (i.e. 2nd to 4th month of lactation). Cows were assigned to one of four
60 groups within each year, with one group treated each week during a 4-week period between
61 November and December when cows had peaked in milk yield. Cows were assigned
62 randomly to a challenge week group, after identification and balancing for expected oestrus
63 dates and previous mastitis treatments (Sanders *et al.*, 2006). Cows that had been treated for
64 mastitis previously during the season were assigned to the third or fourth week, where
65 possible, to allow more opportunity to clear the infection and minimize the risk of any
66 residual antibiotic effects. This resulted in a seven to 10-week interval between treatment
67 and challenge.

68 Intramammary exposure was by infusing *Str. uberis* (approximately 100 cfu, suspended in 1
69 mL of 0.85% w/v saline and 10% v/v glycerol) through the teat canal using an 18 g x 25 mm

70 intravenous catheter, following aseptic preparation of the teat end, as described previously
71 (Sanders *et al.*, 2006). The strain had been isolated from a clinical case of mastitis from a cow
72 in a nearby herd (McDougall *et al.*, 2004). The right rear mammary gland was infused unless
73 this gland had been treated for a pre-existing infection at the pre-trial screening, in which case
74 the left rear gland was used.

75 Inoculations occurred following the Monday morning milking. Immediately prior to infusion,
76 foremilk samples from all quarters were assessed visually for the presence of flecks or clots
77 and California Mastitis Test (CMT) score. The same assessments, as well as collection of
78 milk samples for assessment of bacteriology and SCC, were conducted on the infused gland
79 for the subsequent 13 milkings, or until clinical mastitis was diagnosed. Glands were defined
80 as having clinical mastitis when clots were detected and the CMT score was $\geq$ 2. Upon
81 diagnosis, cows were infused 3 or 4 times at 12 h intervals with 250mg cefuroxime sodium
82 (Spectrazol® Milking Cow; MSD, Upper Hutt, New Zealand).

83 For a random subset of 55 clinically infected animals, the bacterial strain responsible for
84 infection was isolated and confirmed as being the infusion strain by pulsed field gel
85 electrophoresis.

86

87 *Genetic Analysis*

88 Genetic investigations were undertaken in four distinct phases. The first phase consisted of
89 genome-wide linkage analysis to identify QTL for resistance to infection. For the second
90 phase, the most significant of these QTL was prioritised for further analysis, by exome
91 capture and sequencing of the six F_1 sires. Variants discovered within the QTL interval of the
92 six sires were genotyped in the F_2 population, and preliminary association analysis revealed a
93 peak association at a known immune locus. The third phase consisted of genotyping and

94 association analysis of all known variants within this immune locus. The fourth phase
95 consisted of choosing new variations from whole genome sequence data, with all custom
96 variants augmenting a panel of high density SNP genotypes for chromosome-wide
97 association analysis.

98

99 *Exome and Whole-Genome Sequencing and Variant Informatics*

100 DNA sequencing of the six F_1 sire animals of the genetic trial was conducted in two stages,
101 consisting of genomic enrichment and subsequently whole genome sequencing. The Agilent
102 SureSelect system (Agilent Technologies, CA, USA) was used for genomic capture and
103 target a 36.3-68.8 Mbp interval on chromosome 11 for enrichment, with samples then
104 sequenced on the Illumina Genome Analyzer IIx platform (Illumina Inc., San Diego, USA).
105 High-depth whole genome sequencing was conducted on the Illumina HiSeq platform.

106 Exome and whole-genome sequence data were analysed to identify potential functional
107 polymorphisms for custom genotyping. Some 3605 SNPs were discovered in the QTL target
108 region by exome sequencing. Variants were then filtered based on sire family segregation
109 information from half sib QTL analysis i.e. variants not fitting the six-sire zygosity pattern
110 predicted by half sib analysis were filtered out of the target variant set. Based on these data,
111 the families of Sires 2 and 3 appeared to be segregating in their response to the mammary
112 infection challenge at the QTL location, suggesting a heterozygous genotype for the causal
113 variant(s) in these two animals. Applying a filter to remove all variations not heterozygous in
114 Sires 2 and 3, and not homozygous in Sires 1, 4 and 6, reduced the number of target
115 variations to 46. Zygosity rules were not applied to Sire 5 due to suggestive (although non-
116 significant) segregation in this sire family.

117 For filtering of variants derived from whole genome sequence-derived data, more
118 sophisticated filters were applied. First, for pragmatic reasons (including genotyping cost),
119 the QTL target interval was set to ~17.5 Mbp (~37.2-54.7 Mbp). Though narrower than the
120 interval targeted for exome capture, this interval was still deemed to be conservative based on
121 the bootstrap distribution of the chromosome 11 linkage QTL, and represented ~8.75 Mbp of
122 sequence either side of the peak association signal obtained from preliminary association
123 analysis of exome-derived data. Other filters included removal of variants based on the
124 physical distance of these to annotated genes (i.e. intergenic variants), redundancy of variants
125 in 'low priority' regions relative to annotated gene structure, variant quality score and class,
126 and filtering of variants not fitting the six-sire zygosity pattern predicted by half sib analysis
127 (as applied to exome sequence data). Starting from a total of 154,152 genome sequence
128 variants in the revised target interval, these filters reduced the number of variants for custom
129 assay design to 542.

130

131 *Genotyping and variant quality filtering*

132 Genomic DNA was extracted from whole blood from the entire herd (Berry *et al.*, 2010). For
133 linkage analysis, pedigree genotyping was conducted using microsatellite markers and the
134 Affymetrix Bovine 10K SNP GeneChip (Affymetrix, Santa Clara, USA), as described
135 previously (Berry *et al.*, 2010). For association analysis, genotyping was conducted by
136 GeneSeek (Lincoln, NE, USA) using the Illumina BovineSNP50 BeadChip platform, and the
137 Sequenom iPLEX system (Sequenom, San Diego, USA) for custom variants. Genotyping
138 data were filtered to remove individuals with less than an 80% call rate across all SNPs, and
139 SNPs with less than an 80% call rate across all individuals. Data were also filtered by minor
140 allele frequency (<0.5%), and a non-conservative Hardy–Weinberg equilibrium threshold (*P*

141 < 0.000001). Applying these filters yielded 485 custom variants, and 37,986 chip-derived
142 variants for downstream analysis.

143

144 *Statistical analysis*

145 For linkage analysis, the clinical infection status of each animal following *Str. uberis*
146 bacterial infusion was treated as a continuous variable, with JMP software (version 8.0.2;
147 SAS Institute Inc., Cary, NC, USA) used to fit cohort year and challenge week group as class
148 effects. Whether a cow received treatment for a pre-existing infection 7 to 10 weeks prior to
149 challenge, or whether the cow was determined to have a sub-clinical infection (defined as the
150 presence of mastitis pathogens, without any clinical signs) prior to infusion were added as
151 fixed effects, with the residuals used for linkage mapping. Linkage analysis was conducted
152 using regression methodology in a half sibling model (Baret *et al.*, 1998; Haley *et al.*, 1994),
153 with a significance threshold of $P < 0.05$. Chromosome-wide significance levels ($P_{\text{chr-wide}}$)
154 were calculated using 1,000 permutations (Churchill and Doerge, 1994). Genome-wide
155 significance level was determined by applying a Bonferroni adjustment to the chromosome-
156 wide threshold using the formula $P_{\text{gen-wide}} = 1 - (1 - P_{\text{chr-wide}})^{1/r}$ where r is calculated as the
157 length of a specific chromosome divided by the autosomal genome length (deKoning *et al.*,
158 1999). Bootstrapping analysis (n=1000) (Visscher *et al.*, 1996) was conducted to guide
159 selection of the target interval for DNA sequencing and identification of candidate functional
160 variations. Following genotyping and SNP quality filtering, BEAGLE software (Versions
161 3.2.1) (Browning and Browning, 2009) was used to impute all missing genotype data for
162 Bayesian association analysis (Habier *et al.*, 2011). This genotype set contained custom
163 variants discovered via DNA sequencing and markers from the Illumina SNP50 panel. For
164 association analysis, markers were fitted simultaneously using Bayes B (Meuwissen *et al.*,

165 2001), together with the same fixed effects used for linkage analysis, treating mammary
166 infection status as a categorical variable. This model assumed a value of Pi=0.99 of variants
167 having zero effect on the trait. A Markov chain of 50,000 iterations was used to sample the
168 posterior distributions, using prior genetic and residual variance estimates obtained by first
169 running Bayes C for 20,000 iterations. In addition to considering estimated effects for
170 individual markers, the combined effect of all SNPs in the 36.3-68.8Mbp interval targeted for
171 exome sequence analysis was estimated. Following discovery of associated markers, marker
172 genotype for those SNPs were added as fixed effects in the models to assess residual variance
173 of the chromosome 11 interval in the linkage and Bayesian association analyses. The point-
174 wise significance of the top four individual markers was also assessed by least squares
175 analysis of variance using JMP software, using the same phenotype residuals assessed for
176 linkage analysis.

177

178 **References**

179 **Baret PV, Knott SA & Visscher PM** 1998 On the use of linear regression and maximum
180 likelihood for QTL mapping in half-sib designs. *Genetics Research* **72** 149-158

181 **Berry SD, Lopez-Villalobos N, Beattie EM, Davis SR, Adams LF, Thomas NL,**
182 **Ankersmit-Udy AE, Stanfield AM, Lehnert K, Ward HE, Arias JA, Spelman RJ**
183 **& RG Snell** 2010 Mapping a quantitative trait locus for the concentration of β-
184 lactoglobulin in milk, and the effect of β-lactoglobulin genetic variants on the
185 composition of milk from Holstein-Friesian x Jersey crossbred cows. *New Zealand*
186 *Veterinary Journal* **58** 1-5

187 **Churchill GA & Doerge RW** 1994 Empirical threshold values for quantitative trait
188 mapping. *Genetics* **138** 963-971

189 **de Koning DJ, Janss LL, Rattink AP, van Oers PA, de Vries BJ, Groenen MAM, van**
190 **der Poel JJ, de Groot PN & JAM van Arendonk** 1999 Detection of quantitative
191 trait loci for backfat thickness and intramuscular fat content in pigs (*Sus scrofa*).
192 *Genetics* **152** 1679-1690

193 **Habier D, Fernando RL, Kizilkaya , & Garrick DJ** 2111 Extension of the bayesian
194 alphabet for genomic selection. *BMC Bioinformatics* **12** 186-197

195 **Haley CS, Knott SA &Elsen JM** 1994 Mapping quantitative trait loci in crosses between
196 outbred lines using least squares. *Genetics* **136** 1195-1207

197 **Hogan J, Gonzalez R, Harmon R, Nickerson S, Oliver S, Pankey JW & KL Smith**
198 1999 *Laboratory Handbook on Bovine Mastitis*, National Mastitis Council Inc,
199 Madison, WI, USA.

200 **Lopez-Benavides MG, Williamson JH, McGowan JE, Lacy-Hulbert SJ, Jago JG,**
201 **Davis KL & MW Woolford** 2006 Mastitis in cows milked in an automated or
202 conventional milking system in New Zealand. *Proceedings of the New Zealand*
203 *Society for Animal Production* **66** 252-257

204 **Meuwissen TH, Hayes BJ & Goddard ME** 2001 Prediction of total genetic value
205 using genome-wide dense marker maps. *Genetics* **157** 1819-1829

206 **Sanders KM, McDougall S, Stanley GE, Johnson DL, Spelman RJ & SJ Harcourt**
207 2006 Responses and factors affecting intramammary infection rates resulting from
208 infusion of a *Streptococcus uberis* strain in Friesian-Jersey crossbred cows.
209 *Proceedings of the New Zealand Society for Animal Production* **66** 70-76

210 **Spelman RJ, Milllar FM, Hooper JD, Thielen M & Garrick DJ** 2001 Experimental

211 design for the QTL trial involving New Zealand Friesian and Jersey breeds.
212 Queenstown, New Zealand. *Proceedingsof the Association for Advancemnet in*
213 *Animal Breeding and Genetics*, Queenstown, New Zealand, pp. 393-396.
214 **Visscher PM, Thompson R & C Haley** 1996 Confidence intervals in QTL mapping
215 by bootstrapping. *Genetics* **143** 1013-1020
216