3.4.2 Milestone 4.1.2 - Archive of Ruminant DNA – Bovine

Description	The objective of this milestone is to build up an archive of DNA samples. The occurrence of Johne's Disease, especially in cattle is sporadic and unpredictable. It is a straight forward exercise to bleed and prepare DNA from 2000 ruminant animals. The long time frame is due to the difficulties envisaged in identifying suitable cattle. Deer, where the disease is more prevalent, and large numbers of clinical Johnes can be found within a single herd, should give large enough numbers of animals within the first year. By 2012 the aim is to have the following DNA's archived from each species: **2000 Dairy Cattle DNA's**; 2000 Deer DNA's.
Alignment with JDRC Strategy	Tool: Gene markers for identification of Johne's disease resistant animals
Status	Ongoing program, ahead of schedule

2011 Science Report

LIC report for the 2011 JDRC Science Review.

This report covers work completed in Milestones 7 – 12, for Year 3 of this project.

Executive Summary

The ultimate objective of this project is to identify genetic markers for susceptibility and resistance to Johne's disease in bovine animals. Year 3 of the project has seen excellent progress made in the current milestones to collect the remaining samples from affected animals. In addition, work has started ahead of schedule on the genotyping phase of the project. LIC is extremely pleased with progress and findings to date and we believe the outcomes of this work will contribute significantly to the achievement or overarching JDRC programme objectives.

Key results

- Herd test screening has identified 160 cows to sample from 33 herds so far – ongoing until May.
- Genotyping phase has commenced early.
- 1650 samples from positive cows had DNA extracted and 1440 sent for genotyping over High Density SNP panel.
- To date: 1324 useable genotypes.
- Preliminary genomic analysis of the data being undertaken.
    - Refining statistical approach
    - Leveraging on LIC genotyping investment

Background:

JDRC full programme milestone 1: An archive of ruminant DNA related to *M. paratuberculosis* infection for genome wide association studies in sheep, cattle and deer.

During the next three years, collect and store a DNA sample of at least 200ug of purified DNA from ruminants that have been phenotyped as accurately as possible regarding their disease status including their immune response and the identity of the *M.paratuberculosis* pathogen causing the disease.

Genomics approach for the Bovine DNA archive:

- Identify 2000 Johne's disease positive cows using ELISA on milk herd test samples, followed by serum confirmation.
- Apply a phenotype breed definition of greater than 13/16th Holstein Friesian or Jersey.
- Use the existing LIC population data (23,000 genotypes) as control population.
- Utilise the Illumina HD and imputed 50k SNP panels to find genes for resistance and susceptibility to Johne's disease.

The timeline that was assigned to this project was:

- Year 1: Dec 08 – June 09 Pilot trial to test screening and collection process (completed – see Appendix).
- Year 2: June 09/10 First season large scale screening (completed – see Appendix).
- *Year 3: June 10/11 Second season large scale screening (in progress).*
- *Year 4: June 11/12 Genomic analysis for markers (in progress).*

Year 3: Herd Test Screening and Genomic Analysis

The following description of work that has been conducted during Year 3 of the project is split into 3 sections for ease of understanding:

1) Introduction
2) Update on current herd test screening
3) Genotyping progress
4) Future direction

For reference, previous reporting of Year 1 & 2 is in the Appendix.

Introduction

Year 3 of the project has seen a combination of work proposed in both Years 3 & 4 being progressed. As a greater number of samples had been collected than originally predicted by the end of Year 2 (1650 vs 1000), permission was sought from JDRC to advance genotyping these animals into Year 3 of the project.

Initial planning in the study was to use 50K marker panels. It was decided with technology changes to utilise the latest marker technology to derive a better technical output. The new high density (HD) panels are more expensive than 50K panels but the genotyping could continue within the yearly budget due to LIC helping out with cashflow aspects. Below is the an excerpt from the change request that was submitted to the JDRC board (Milestone/Change request: September 2010, Milestone 4.1.2G)

This dataset should be genotyped on the HD panel rather than the 50K panel as the phenotype is very difficult to collect and therefore the best genomic resources should be applied. The reason why we have indicated that the HD panel is the preferred option is that the dairy cattle dataset has a mix of Jersey, Holstein-Friesian and crossbred animals. We (LIC) have shown that genomic predictions with the 50K panel are not transferable across breeds i.e. those that we estimate in HF for milk production do not predict milk production in Jersey. This is due to the markers being too far away from the causative mutations and thus are not in consistent linkage disequilibrium across breeds, i.e. in the HF breed the A marker allele co-segregates with the protein increasing allele, whereas in the Jersey population the B marker allele is just as likely to co-segregate with the protein increasing allele. We are still in the process of validating that the HD panel will ensure that the markers will work across breed, but simulated results indicate that this will occur with marker densities greater than 250,000 markers. LIC has committed to the Illumina array being the HD panel of choice and thus our bull population (the control group for this experiment) will have genotypes from this panel.

Permission was granted by JDRC for the HD genotyping of 1440 samples to progress. Results are in the Genotypes and Genomic Analysis section.

Current herd test screening to identify cows to blood sample:

Using the screening process that was successfully validated during Year 1 of the project (see Appendix), we have continued to collect samples from target animals. Using the same predicted prevalence rates as in Year 2,

- 390 herds remaining from the original data extract were vat tested.
- This resulted in 40 herds to target for herd test screening.
- In addition, we had 25 repeat herds and 5 new requests.
- From 33 herds, 160 milk positive cows identified so far
- Blood sampling of identified cows underway

Feed back from farmers participating in the trial would indicate the screening is viewed in several ways:

A) it is an industry-good research trial so they are happy to participate and 'help out'.
B) They would like to reduce incidence of JD in their herd and see this as a good tool to help do this. They acknowledge the benefit of this being a research trial, with nil cost to them at present but also indicate they would be prepared to continue screening and pay for the service. It is seen as a short-term cost; 3 years screening and culling to be on top of the problem. In terms of the cost they would be willing to pay, it would be relative to cost of getting 1-2 first calving heifers into the herd due to cost of raising replacements (`$2000).

There has also been interest expressed from vets with clients who have a herd with a Johne's disease problem in using this approach.

Genotypes and Genomic Analysis

This section covers the progress made with regards to the Genomic part of the project. Any results presented are preliminary and discussion is needed around the refinement of the final methodology to apply to the completed data set.

- By the end of Year 2, 1650 samples from positive cows had DNA extracted.
- 1440 samples were sent for genotyping with the Illumina HD panel in November 2010 – received back in December 2010.
- Data clean-up and quality control processing commenced.
- Good representation of breeds, geographical regions and cow ages (represented in the following Tables 1-3).

Table 1. *Johne's-positive samples by breed*

Breed	Number
Other Cross	41
Holstein-Friesian	258
Jersey	660
Ayrshire	1
KiwiCross	364

Table 2. *Johne's-positive samples by age*

Age	Number
15	1
13	5
12	9
11	14
10	34
9	76
8	127
7	162
6	193
5	233
4	253
3	155
2	62

Table 3. *Johne's–positive samples by region*

Region	Number
Northland	40
Waikato	272
Bay of Plenty	34
Taranaki	322
Wellington/Hawkes Bay	130
South Island	526

At present there are 1324 Johne's-positive HD genotypes being analysed. The sample exclusion criteria (QC) included: call rate <87%, and heterozygous rate of less than 20%, or greater than 40%. Animals that failed parentage (to sire) were not excluded from the dataset as their genotypes remain informative even without pedigree information. After QC, 711955 SNPs were retained. Those included were autosomal, with a known location and had a minor allele frequency of greater than 2%.

Analysis of genotypic data for stratification has begun. The major challenge will be the choice of an appropriate control population (or populations) from the 23,000 HD genotypes available. We are intending to generate the expectation of the allele frequencies for the Johne's positive cows that we have selected based on their pedigree. With most of the ancestors of the cows being genotyped as part of the LIC genomics programme we can estimate the allele frequencies that we would expect to see. Departures from expectation with the observed genotypes will indicate regions of interest.

Table 4. Johne's-positive genotypes by sires and maternal grand-sires (MG sires).

	Sires	MG sires
JD (+) daughters	# sires	# sires
>49	3	0
20-49	7	10
10-19	21	20
5-9	23	25
<5	282	345

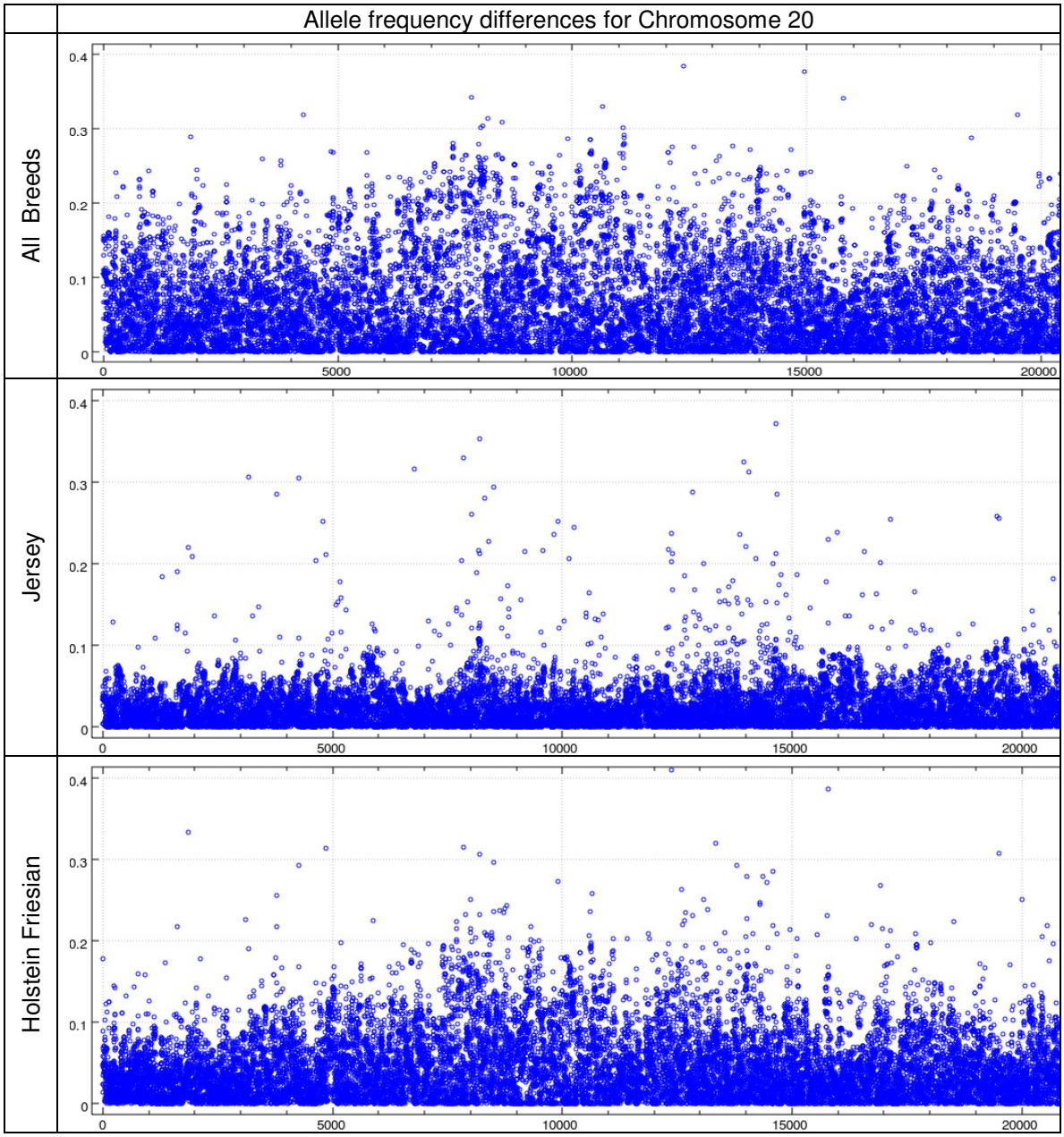
Table 4 shows that for the Johne's-positive cows in the study, there are 336 sires with Johne's-positive daughters and 10 sires accounted for 32.3% of the positive daughters. 9 of the top 10 sires have HD genotypes. Whereas there are 400 maternal grand-sires with Johne's-positive daughters and 10 sires accounted for 23.7% of the positive daughters. 3 of the top 5 MG sires have HD genotypes.

A preliminary analysis of chromosomes 20 and 6 has been conducted using the existing LIC HD genotypes as the control population (Table 5 & Figure 1). The HD control consists of other animals that have HD genotypes and have been used as the control for this exercise as it is the only HD dataset that is currently available for use. This has limitations because of differences in breed representation between the two sample populations (Table 5).

Table 5. Comparison of existing LIC HD genotypes and Johne's-positve genotypes.

Breed	HD control	Johne's-positive
Other cross	0	41
Holstein-Friesian	999	258
Jersey	447	660
Ayrshire	0	1
KiwiCross	11	364
Total	1457	1324

Figure 1 shows that accounting for breed stratification helps reduce the noise in the case/control allele frequency differences potentially enabling the identification of regions of interest.



Future direction

- Completion of collection and genotyping of remaining Johne's positive samples
- A decision needs to be made regarding genotyping the remaining samples over the HD panel
- Add additional 20,000 imputed genotypes to control population
- Examine for evidence of stratification (breeds/sires/age)
- Refine genomic analysis

In addition, there has been ongoing work in our collaborations with:

Des Collins (Milestone 3.2, Molecular Strain Typing.
Faecal culturing of samples from the pilot trial resulted about 200 isolates of *M. avium* subsp. *paratuberculosis.* In addition, approx 100 samples from control cows sampled during Year 2 have been cultured. LIC has a significant amount of epidemiological data from the animals from which these isolates came, and this data could be combined with the typing information generated to improve the value of the results and perhaps indicate possible future directions of collaborative work between our two groups.

Bryce Buddle (Milestone 3.1 Mucosal Immune Responses).
Results from this milestone show promising results in differences in Toll-like Receptor genes for resistance and susceptibility. It may be possible to genotype these mutations using LIC population data. If the LIC work results in a good panel for resistant or susceptible animals, there is also the potential to validate these results by screening animals through the experimental challenge system being developed.

Appendix

Year 1: Screening process to identify affected cows

A large number of herds will need to be screened to collect DNA from the required number of cows for this project. A screening process to efficiently identify affected cows has been devised; using dairy company vat test samples (a bulk milk sample for that herd) to pre-screen herds and the LIC herd testing system to select individual cows for sampling. The Animal Health Lab at LIC has validated a commercial ELISA for use with milk samples (both bulk and individual cow samples) to identify Johne's disease positive cows. The confirmation test is done using a blood sample that for efficiency is also used for DNA extraction and analysis.

A pilot trial tested the robustness of this approach. To test the validity of using the vat sample as a screening tool, 400 herds were screened during January-February 09 from both low prevalence regions, and high risk regions and Johne's disease culling herds. Results showed that the bulk milk ELISA is an efficient screening tool to target high(er) Johne's disease prevalence herds.

To test the validity of the vat sample and understand its relationship with the proportion of Johne's disease reactor cows within the herd, 60 herds were selected across the range of test results. The herd test sample of cows were pooled (10:1) and this sample tested. Samples within pools that show a positive result (indicating potential reactors within that pool) are then tested individually. There was a strong correlation between vat results and proportion of Johne's disease positive cows in the herd ($r^2 = 0.86$). Hence, the vat test is a powerful tool to help target herds with higher Johne's disease risks for the pool – individual cow testing process to identify affected cows for the genomic study. The pooling strategy also resulted in a 75% test reduction compared with testing all cows individually so is efficient and cost effective.

Results from the 50 herds that participated in the pilot trial:

- 18,922 cows were screened, with 2.43% positive for Johne's disease.
- Older cows had a higher occurrence of the disease (range from 1% in 2 yr olds – 3.5% in 8 year olds).
- A higher proportion of positive samples were seen in Jersey cows (3.1%), followed by crossbred (2.7%) and Holstein Friesian cows (0.9%).
- DNA samples were collected from 287 blood positive cows.
- Faecal cultures from these cows showed 216 cultures positive for *M.paratuberculosis* and 71 were negative (32%).
- In addition, these samples and phenotypic information are being used by Geoff DeLise (Milestone 1.1, Microbial Quantitation). We also sourced animals from the pilot trial for Bryce Buddle (Milestone 3.1, Mucosal Immune Responses).

In conclusion, use of the vat pre-screen and targeted pooling of milk herd test samples provides an efficient, cost-effect method to undertake large-scale screening for Johne's disease sample collection.

Year 2 First season large scale screening

The main objective for the current year of the project is to obtain DNA samples from 1000 serum positive cows. To do this we have used the sample collection process validated in the pilot trial.

The vat sample pre-screen was conducted October – December 09. 5000 herds were selected using breed proportion (down to 40% 13/16th HF or J). Table 1 shows the results for the herds tested, with results split into the following categories on predicted prevalence rates:

- Positive herds: S/P ratio > 0.1 prevalence = 3%
- Suspect herds: S/P ratio >0.05 prevalence = 1.5 – 2 %
- Check herds: S/P ratio > 0.04 prevalence = 0.5%

Table 1. Results from the vat test pre-screen.

Breed	Herds	Positive	Suspect	Check	Total
Holstein Friesian	1144	1%	4%	2%	6%
Mixed herds	3040	1%	4%	3%	8%
Jersey	532	3%	9%	6%	17%
Total	4716	1%	5%	3%	9%

The Jersey herd reactor rates are approximately 3x higher than Holstein Friesian herds but there are a smaller number of herds to collect from.

From the herds tested, 400 herds were selected for the targeted herd test screen. Depending on when individual herds have their herd test scheduled, we have tried to screen all positive and suspect herds plus as many check herds as the lab can process (the limit is approx 15,000 cows/week).

Once the potentially positive animals are identified from the milk test, farmers are approached for permission to blood and faecal sample these animals. Negative control animals are also being tested from some herds. Faecal samples from negative control cows are being sent for faecal culture whereas faecal samples from positive cows are being stored for use in future PCR development.

As of early April, we have identified

- approx 1915 potential reactors from 300 herds from the milk herd test samples.
- 340 herds have been screened, with 37 showing no positive pools.
- Have collected 1156 positive blood samples from 270 herds, which will be used for DNA extraction.
- 109 negative results = error rate 9%.

The error rate is determined from cows returning milk positive but blood negative tests. This is due to several factors: cross contamination of the milk sample during herd testing, the animal is in early stage of disease or potentially is negative. Given the method of milk sample collection and limitation of the ELISA detecting animals in the early stage of disease, this error is acceptable.

By the end of the project year (June 2010) we have the potential to have samples from 1500 positive animals.

Future direction

Current progress indicates we are well ahead of the proposed sample collection target (approx 1500 from the current year and 287 from the pilot trial). The original plan was to genotype the animals in Year 4 (2011/12). If the current collection target of samples from 1500 positive cows is met, there is the potential to genotype the animals we have identified to date in Year 3 rather than

Year 4 of the project, along with the additional cows that we would identify in Year 3 to meet the 2000 target.

This has been communicated to JDRC and we are seeking an indication from the consortium if they would like to utilise this opportunity that has arisen from faster than expected phenotypic screening.

In addition, there has been discussions re collaboration with:

Des Collins (Milestone 3.2, Molecular Strain Typing.
Faecal culturing of samples from the pilot trial resulted about 200 isolates of *M. avium* subsp. *paratuberculosis.* Typing of these isolates would show the recent distribution of the different types of *M. avium* subsp. *paratuberculosis* in dairy cattle in New Zealand. These could then be compared with the types occurring in dairy cattle 25 years previously. LIC has a significant amount of epidemiological data from the animals from which these isolates came, and this data could be combined with the typing information generated to improve the value of the results and perhaps indicate possible future directions of collaborative work between our two groups.

Bryce Buddle (Milestone 3.1 Mucosal Immune Responses).
Results from this milestone show promising results in differences in Toll-like Receptor genes for resistance and susceptibility. It may be possible to geneotype these mutations using LIC population data. If the LIC work results in a good panel for resistant or susceptible animals, there is also the potential to validate these results by screening animals through the experimental challenge system being developed.