The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

School of Veterinary Medicine and Science

A Multidisciplinary Approach Involving Sensor Dynamic Systems & Bioinformatics to Predict Cattle Health

NAME: Emily Kate Walton

STUDENT ID: 4330266

PROJECT SUPERVISOR: Tania Dottorini

ADDITIONAL SUPERVISORS: Jasmeet Kaler

Abstract

Dairy herds contribute largely to the agricultural economy of the UK. However, with the demand for cheap dairy products and pressure to reduce the environmental impact of farming practices, there is a growing interest in increasing production efficiency. One way to achieve this is by understanding the relationship between disease and growth. This pilot study looked at the trend and correlation of weight, metagenome and core temperature of calves along time on UK farms. The aim of the study was to assess the use of weight, metagenome and core temperature in combination to aid in monitoring calf health. A total of 22 calves were used in the study that spanned over 6 weeks. 20 calves aged between 7 and 8 months old were followed over 6 weeks. A further 12 animals around 14 weeks of age were included in the final week of the study. Core temperature was measured via ruminal boluses while weight was recorded manually using a weigh band. The metagenome was determined via shotgun sequencing of DNA extracted from rectal swab samples.

From our preliminary results, it appears that the metagenome is highly enriched with organisms that are taxonomically unclassified (~50% of reads). The classified metagenomics component reflects a more dynamic and constantly changing pattern at an individual level than at a farm level.

No major differences could be identified between either healthy and unhealthy or young and old classified calves using the taxonomically classified reads representing over half of the overall data. However, interestingly it was found that the Lactobacillales order was only found in detectable abundances in the older calves. Lactobacillalles have been an important order of bacteria in probiotic research in recent years and one found to be very dependent on diet. Similarly, only a weak pattern was observed linking the calves demonstrating abnormal or poor weight gain.

Our approach and preliminary results demonstrate a novel approach for the study of health indicators integrating large sets of heterogeneous data such as field, clinical, sensor and genomic data. Likewise, the findings highlight the importance of discovering dedicated pipelines to analyse the unclassified reads to provide a more complete and comprehensive overview of the metagenomics data.

Contents

1. Introduction – Literature Review

1.1. Introduction

The dairy industry is an important business in the UK. According to a 2016 report (Bate, 2016), milk accounted for 17.8% of total agriculture output in the UK in 2014. This is equivalent to £4.6 billion. The same report stated that the UK was the 3rd largest milk producer in the EU and 10th largest in the world. Many studies have reported that the average dairy herd size in the UK is increasing (Bate, 2016, DairyCo, 2016).

Despite this increase in herd size, DairyCo, 2016 found that the actual number of dairy holdings in the UK has declined from 23,682 in 2010 to 20,804 in 2015 (DairyCo, 2016). A similar trend was noted in dairy farms globally (Barkema et al., 2015, McDonald et al., 2007, Noordhuizen et al., 2008). Specifically, Barkema and collaborators proposed that these shifts in herd size are due to pressures from areas such as economics, technology, shifting demographics, consumer expectations and regulations (Barkema et al., 2015). A major motivator for farms to expand for the increased profits and efficacy that larger herds produce (McDonald et al., 2007).

However growing herds do not come without problems. Expanding herds are at a higher risk of many diseases such as; bovine viral diarrhoea (BVD), infectious bovine rhinotracheitis (IBR), Q fever, bovine respiratory diseases and Johne's disease (Faust et al., 2001, Anastácio et al., 2016, Doyle et al., 2014, Wolf et al., 2014). In addition to disease risk, growing herds are at a higher risk of mortality (Shahid et al., 2015). The increase in individual cow milk yields is one factor that has made larger herds more profitable (Bate, 2016). However increased milk yield has been associated with an increase in conditions such as mastitis and cystic ovaries as well as a general decrease in cow health (Koeck et al., 2014, Barkema et al., 2015).

1.2. Calf Health

1.2.1. Calf Weight

In order for dairy farms to be profitable, they need to be able to raise high-quality replacement heifers with minimal cost to the farmer (Heinrichs, 1993). Farms aim to keep the age at first calving below 22 to 24 months in order to minimize the investment required (Hoffman and Funk, 1992). Diet and the resulting rate of growth directly affect the age at first calving and therefore increase the productivity per day of life (Heinrichs, 1993). Not only does increased weight gain decreased the age at first calving, but it has also been found that the pre-weaning growth rate has a positive correlation with milk yield during first lactation (Soberon et al., 2012). Due to these previous attempts to increase productivity have revolved around changing the diet (Gardner et al., 1988). However, adjustments to the diet can often be costly and diet is not the only factor affecting growth as Berry and collaborators found that different breeds of cattle vary in growth rate, birth weight and final weight as well as final height and length (Berry et al., 2005). There is also the consideration that even if a sufficient diet is provided is relies on the calves intake in order to influence daily weight gain. Studies have found that factors such as housing location, separation from the dam, weaning, season and

individual farm variation can have a significant effect on the average daily intake of calves (Place et al., 1998). Increasing dry matter intake is important for weight gain without having to increase dietary provisions.
Weight gain may also be influenced by factors outside of daily intake. Donovan and collaborators found that the season of birth, as well as the occurrence of disease, could decrease weight gain in the first 14 months of a calf's life (Donovan et al., 1998). Other research has also indicated that there could be a genetic influence on feed efficiency resulting in better weight gain when fed the same diet in the same volumes (Koch et al., 1963). These studies highlight the role of calf health in ensuring optimal weight gain and minimizing production costs.
Once at a mature weight calves should be calved as soon as possible to avoid additional production costs associated with additional weight gain as well as avoiding the drop in first lactation milk yield associated with excessive weight before first calving (Miller and McGilliard, 1959). Bodyweight in dairy calves or cattle is not always easy or able to be measured but traits such as heart girth circumference have been found to be accurate in giving an estimate of weight that can be used for monitoring growth rates and identifying when a mature weight has been reached (Heinrichs and Hargrove, 1987).

1.2.2. Calf Gut Metagenome

Another factor recently being recognized for its importance in weight gain in the gut metagenome. A metagenome is the collective genome of microorganisms from an environmental sample, so the gut metagenome is the genome of all the microorganisms present in the gut.
In a study comparing the weight gain of normal honeybees to germ-free honeybees, the presence of a microbiome appeared to be essential in promoting healthy weight in their early life (Zheng et al., 2017). It is believed that the microbiome could be similarly essential in the growth rate of other animals too. In humans, relationships have been found between the gut microbiota and the health of the elderly. Different groups of gut bacteria appeared to form clusters that reflected residence (inpatient, outpatient or community dweller), diet and health (Claesson et al., 2012). A decrease in gut microbiota appeared to indicate increased fragility and vulnerability to poor health.
In adult dairy cattle research has already identified associations between feed intake microbiota and feed utilization. Delgado and collaborators found this association in their experiment when they looked at the feed intake, feed efficiency and rumen metagenome based on metagenome sequencing from rumen fluid collection (Delgado et al., 2019). Other research has also backed up the importance of the gastrointestinal bacterial community and functioning for milk production and host health in dairy cattle (Mao et al., 2015). During their study, Mao and collaborators also found there to be significant spatial heterogeneity between different sites in the gastrointestinal tract (Mao et al., 2015). Additionally, one study found that milk production of a lactating dairy cow could be predicted based on rumen microbiome demonstrating the importance of the microbiome in dairy cow productivity (Lima et al., 2015).
Previous research has studied the perinatal microbiome of dairy calves and its subsequent development in early pre-weaning life (Alipour et al., 2018). Using rectal swabs to extract DNA and perform 16S rRNA sequencing it was found that the microbiota of new-born rectal mucosa was diverse but low in abundance, resembling the oral microbiota of the dam. After the first day of

life diversity of microbiota in rectal mucosa collapsed before starting to increase again at 7 days old. Oikonomou and collaborators noted that during the first few weeks of life some bacteria increased in abundance in the faecal microbiota while others decreased (Oikonomou et al., 2013). These changes were thought to be firstly an adaption to consuming milk followed by a further adaption to consuming the solid feed. However, overall the diversity of the faecal microbiota increased in the first 4 weeks of life. Other studies also support the findings that diversity is seen to increase from day 7 until day 21 of life (Yeoman et al., 2018). It was also noted by Yeoman and co-workers that the vaginal microbiome of the dam influenced the gut microbiome seen in the calves in the first 21 days of life (Yeoman et al., 2018).

In human infants, administration of antibiotics, birth by caesarean section and feeding of milk replacers can cause disturbances in the microbiome that result in metabolic and immunological consequences in later life (Bokulich et al., 2016). Li and collaborators found in a study that the administration of milk containing antibiotics as a cost-effective milk replacer can impact the normal rumen microbiota (Li et al., 2017). The changes in abundances observed not only made bacteria more prone to the development of antibiotic resistance but also altered the fat metabolism of the milk since microbiota plays an important role in digestion in the ruminant.

The presence of disease also influences the microbiome. Calves experiencing diarrhoea, a common occurrence in early life, have been found to vary in bacterial community membership and structure found in the faeces when compared to faeces of normal healthy calves (Gomez et al., 2017, Oikonomou et al., 2013). From the 16S rRNA sequencing of faeces, it was found that calves experiencing diarrhoea had a lower abundance of bacteria responsible for the metabolism of vitamins, amino acids, and carbohydrates. Pneumonia was also found to decrease the microbial diversity found in the faecal microbiota in the weeks following infection (Oikonomou et al., 2013). However, since pneumonia cases in the study were promptly treated with antibiotics it was predicted that any changes in the microbiota were due to the antibiotic treatment rather than the disease process.

Another factor explored in the study by Oikonomou and co-workers were the differences in faecal microbiota between calves exhibiting good weight gain in comparison to those exhibiting poor weight gain (Oikonomou et al., 2013). Calves that demonstrated good weight gain up until weaning were found to have a larger increase in faecal microbiota diversity between week one and week four as well as a higher final microbial diversity by week four. The high presence of *Faecalibacterium spp* in the first week of life in calves that showed good weight gain was thought to contribute firstly to the lower diversity in the first week. This research demonstrates the importance of the calf gut microbiome in influencing the calves' early life productivity.

1.2.3. Cryptosporidium in Calves' Guts and its Impact on Health

The metagenome is made up of many organisms many of which are parasitic. Protozoa are common parasites found in the guts of mammals with varying impact on the host's health including weight loss. Cryptosporidium a commonly found protozoa in calves' guts that cause gastrointestinal and respiratory issues that can lead to a decrease in appetite in addition to reduced growth and poor weight gain (Chalmers and Giles, 2010). These factors lead to poorly performing calves.

Cryptosporidium is a protozoan parasite that infects animals and humans across the globe (Mosier and Oberst, 2000, Manyazewal et al., 2018, Mirzai et al., 2014, Paul et al., 2009, Razakandrainibe et al., 2018). The parasite is very prevalent in the UK with a 2014 study finding at least one positive sample on 92.5% of farms (Smith et al., 2014). They also found a prevalence of 10.2% in UK cattle with higher prevalence in calves. The infective stage of the lifecycle of the parasite is an oocyst. These are ingested resulting in infection as the parasite completes its lifecycle inside the host. Once the lifecycle is complete infective oocysts are shed in the host's faeces (https://www.cdc.gov/dpdx/cryptosporidiosis/index.html). In the faeces, the infective oocysts can end up contaminating water and feed supplies as well as bedding materials. Bodley-Tickell and collaborators looked at the presence of Cryptosporidium oocysts in surface water on a UK farm over 17 months. 66% of samples were found to be positive for the parasite (Bodley-Tickell et al., 2002). Although, the highest concentrations and frequency of oocysts were found in the autumn and winter they were detected all year round. Calving time also saw a rise in frequency and concentration. Once cryptosporidium infection is present on a farm it can be very hard to control and since it is zoonotic it poses a risk to human health. Many methods are available to detect Cryptosporidium oocysts in faecal samples. Brook and co-workers compared three different diagnostic tests on fresh and frozen faecal specimens with PCR as a gold standard control (Brook et al., 2008b). The methods compared were; modified Ziehl-Neelsen; auramine phenol stains; and a commercial enzyme immunoassay. All three methods were found to be effective in detecting Cryptosporidium oocysts however there were disagreements between some samples. However, PCR is still the preferred method of detection for the parasite due to its ease of ability to perform on batches of samples, simple interpretation and potential to further differentiate between different species and genotypes. Other studies examining the presence of Cryptosporidium via detection of oocysts in faecal samples have successfully used PCR (Brook et al., 2008a, Sturdee et al., 2003).
Determining the presence of cryptosporidium is an important factor in calf health due to its potential effects on calf performance.

1.3. Precision Livestock Farming

1.3.1. Background

As dairy herds grow, they require not only more resources in the forms of feed, housing etc. but also more labour is required in order to effectively monitor for ill health. Not only does disease need to be identified early but also key events such as oestrus need to be accurately identified in order to continue to produce from cows in the most cost-effective manner. For this reason, technology has started to play a bigger role in farming. In a study on 217 dairy farms in Holland research found that most farms using an automatic milking system chose to use built-in sensors as they came at no extra cost to the producer (Steeneveld and Hogeveen, 2015). However, many farms still using a conventional milking system still chose to invest in sensors with mastitis and oestrus detection systems being the most common. A later study on the same Dutch dairy farms investigated the economic consequences of investing in sensor technology and found surprisingly that most farms did not experience an increase in productivity or decrease in labour costs(Steeneveld et al., 2015).

This highlights the importance of farmer education and continual guidance required when introducing new technology to ensure it helps those it is designed to help.

1.3.2. Sensors

Oestrus detection is one area where sensors are already being used in dairy cows. One study compared pedometers, neck collars, heat detectors and human observation (Holman et al., 2011). It was observed that the neck collars and human observation had similarly high positive predictive values, but the best accuracy was achieved when these two methods were combined. This does not help reduce the labour requirements but may reduce the number of false positives and improve oestrus detection.

Accelerometers have also been widely researched for their potential in disease detection. In sheep it was demonstrated that tri-axial accelerometers attached ear tags were 95% accurate in identifying lying, standing and walking behaviours (Walton et al., 2018). This makes them a potential tool in monitoring for lameness that alters walking patterns as well as reviewing for behavioural changes such as increased lying that may be indicative of poor health. An earlier study found a similar potential for 3D accelerometers in cows (Bikker et al., 2014). Sensors on ear tags were used to classify data as ruminating, eating, resting or active. Eating was the most accurately classified activity as a useful metric as the disease may reduce appetite but could go unnoticed in a large herd that shares a feed trough. Accelerometers have also been investigated in combination with temperature to predict calving. Sensor data from an ear tag that integrated cumulative activity, ruminating activity, feeding activity, and the temperature was found to predict calving with a low sensitivity within an hour of calving which a much higher sensitivity for calving within 3 hours (Rutten et al., 2017).

1.3.3. Core Temperature

Core temperature is an important parameter used to monitor health in humans and animals. The presence of abnormally high temperature over a prolonged period (fever) can indicate the presence of disease or illness. The traditional method of core temperature recording in an individual is by using a rectal thermometer. However, this procedure requires manual labour, good animal handling and can be time-consuming. Automatic detection of temperature changes would offer a much more time-efficient and easier way to detect temperature abnormalities in large herds.

1.3.3.1. Ruminal Temperature Boluses

In 2008 a pilot study was carried to look at the feasibility of an intra-ruminal bolus sensor in order to continuously monitor and transmit deep core temperature information(Ipema et al., 2008). A diurnal rhythm in the cow's core temperature was observed as well as a significant reduction in temperature when the cows drank water. The sensors also had issues with the transmission of data particularly seen when cows were lying down. Some studies reported that the ruminal temperature was higher than the rectal temperature (Prendiville et al., 2002, Bewley et al., 2008a) while another reported them to be very similar (C. Hicks et al.) and has even been reported that the rumen temperature was lower than the rectal temperature (Rose-Dye et al., 2011). However, they all agree that the two measurements are closely correlated. Tympanic temperature was also compared to ruminal and rectal temperature in one study (Prendiville et al., 2002). Although the tympanic sensor provided a

good alternative to rectal or ruminal temperature measurement as it was closely correlated to both it was found that the proximity to the tympanic membrane resulted in the development of ear problems in the cows. Factors such as ambient temperature, housing system, parity and milking were found to influence the core temperature as measured by a ruminal bolus so needs to be considered when processing any temperature data (Bewley et al., 2008a). As previously mentioned, water intake has a significant effect on ruminal temperature readings. One study investigated this effect further looking at the data produced from cows with ruminal boluses when given hot water, warm water, cold water or body temperature water in comparison to controls who consumed no water during the sampling period (Bewley et al., 2008b). Control cows remained within the boundaries of a normal baseline temperature throughout. Cows that consumed hot, warm or cold water experienced a significant drop in ruminal temperature that did not return to baseline (above 37.7°C) until over 3 hours after the drinking event. However, cows that consumed body temperature water only experience a moderate decrease in ruminal temperature that returned to normal with 15 minutes from the drinking event. Hence when recording ruminal temperature over time drinking events need to be considered when interpreting the data.

The potential of temperature boluses to detect changes in temperature associated with disease has already started to be investigated. One study looked at data from ruminal boluses in beef steers when challenged with either Bovine Viral Diarrhoea Virus (BVDV), Mannheimia Haemolytica (MH) or both in comparison to controls (no challenge) (Rose-Dye et al., 2011). They found temperatures efficacious in detecting infection with the pathogens of interest. The sensors in this study had a 90% rate of transmission of data a value highly improved from the 50% reported in earlier studies (Ipema et al., 2008). Another study also found that rumen hyperthermia detected by rumen boluses was associated with infection with naturally occurring bovine respiratory disease (Timsit et al., 2011). There was often a lag of 12 to 136 hours between the start of the RH episode and the manifestation of clinical signs of BRD. The lag between the onset of RH and clinical signs demonstrates the potential of the boluses to enable detection of disease before the onset of clinical signs, therefore, making treatment and prevention of spread more effective. However, since the rise in rumen temperature is not specific for BRD the clinical exam is still required for definitive diagnosis.

Studies have also looked at the use of rumen boluses to measure core temperature in adult dairy cattle. One study not only looked at the changes in rumen temperature observed when cattle were challenged with *E. Coli, but* they also studied the effect of sub-acute ruminal acidosis (SARA) due to diet on the core temperature (AlZahal et al., 2011). In this study vaginal temperature was also used in order to validate the rumen temperatures recorded. Increases in ruminal temperature were seen in both cows with SARA due to diet and *E. coli* challenge. The increases seen in the rumen temperature were smaller but correlated to a similar pattern in the vaginal temperature. Fever was defined as a core temperature between 40.5 and 41°C for 2 hours or more. Fever was only seen in cows challenged with E. coli.

A later study also looked at the ability of ruminal temperature sensors to detect disease in adult dairy cattle (Adams et al., 2013). In this case, the temperature was recorded 3 times daily as they exited the milking parlour rather than

throughout the day. This method of recording increased the battery life and successful transmission percentage but may mean that rumen hyperthermia events may be missed. The times when the temperature was recorded as increased (0.8℃ over previous 10 days average for individual cow) were noted as well as any health events such as mastitis, metritis, lameness or pneumonia. Specificity and sensitivity for an increase of 0.8°C above baseline RT within 4 d of disease diagnosis were 76.85 and 66.97% for mastitis, and 69.23 and 76.92% for pneumonia, respectively. Although these numbers are not perfect the sensors still detect illness in many the calves. However, the use of sensors was considerably less time consuming and required less personnel than measurement of rectal temperature for each cow. It was noted that cows diagnosed with mastitis or pneumonia were more likely than a healthy cow to have experienced an increased temperature recording 4 days before the clinical signs of the health event manifested. There was no significant difference in core temperature for cows identified as being lame or having metritis in comparison to healthy cows.

1.4.Sampling and Methods

1.4.1. Sample Collection

In order to study the microbiome, samples need to be collected and DNA extracted prior to sequencing. Many studies in both humans and animals have used faecal samples to examine the microbiota contained in faeces from various species (Claassen et al., 2013, Ingala et al., 2018, Kumar et al., 2016). Although rumen fluid would be the ideal fluid to examine gut microbiome in calves it requires much more invasive sampling. The invasive sampling is hard to justify ethically and puts the calves at a much greater risk of infection and ill-health. Therefore, a less invasive sample such as faeces is more suitable for a study in which we are interested in the health of the calves alongside the metagenome. The collection method is also important. While some studies preferred to obtain faeces via evacuation of the rectum with a sterile glove (Kuhnert et al., 2005), others preferred to use rectal swabs (Gioffré et al., 2004, Rapp, 2010, Jee et al., 2004). One study found that the DNA extraction and later amplification was more successful in samples containing small volumes of faeces such as swabs (Rapp, 2010). There has been some research on the effects of different swab types on the success of the DNA extracted from the sample. Brownlow and collaborators compared the use of traditional cotton swabs against newly developed nylon flocked swabs (Brownlow et al., 2012). Although there were pros and cons identified for both swabs types, it was concluded that there was not enough evidence to prove that either swab was superior to the other so most research still uses the cheaper and more available cotton swabs.

1.4.2. Sample Storage

Once collected the samples must be stored correctly to ensure that the sample remains reflective of the faecal microbiome. In 2012 a study looked at the stability of the microbiota in human faecal samples that were stored for 6 months at either -80℃ or at room temperature (Carroll et al., 2012). Samples were examined by extracting the DNA and performed 454 pyrosequencing at set time points throughout the 6 months. It was found that the faecal microbiome was altered after more than 24 hours at room temperature but that storage at -80℃ for 6 months had no significant effect on the microbiota. Other

studies have also found that degradation of samples occurs when stored at room temperatures for over 24 hours, with freezing or cooling having a protective effect on the microbiota (Ott et al., 2004, Cardona et al., 2012). However, it is not always possible to freeze samples within a 24 hours window after collection, so some type of preservation is often required for samples collected in the field. RNAlater is one potential solution for the preservation of samples that can't be frozen within 24 hours. One piece of research looked at the differences between immediate freezing, storage in RNAlater or storage in 96% ethanol on the bacterial composition of gorilla faeces (Vlčková et al., 2012). It was concluded that there was no significant difference between freezing and RNAlater however if required 96% ethanol would provide a suitable alternative. One piece of research also looked at the use of RNAlater for the preservation of faecal samples, this time in humans (Dominianni et al., 2014). RNAlater was compared to faecal occult blood test cards, sterile Eppendorf tubes and freezing. It was noted that there was more variation between individual's samples than between storage methods. A decrease in purity and diversity of samples stored in RNAlater was seen however the storage methods used did not vary in bacterial composition or taxon distributions.

1.4.3. DNA Extraction

Faeces can be complicated to extract bacterial DNA for sequencing from due to the high presence of host DNA and PCR inhibitors as well as other organic material. This means that the selection of the correct extraction kit or method is essential. Many studies have look at the differences between various kits and extraction methods for use on human faeces (Mirsepasi et al., 2014, Claassen et al., 2013, Kumar et al., 2016, McOrist et al., 2002, Salonen et al., 2010, Smith et al., 2011, Nechvatal et al., 2008, Lim et al., 2018, Hosomi et al., 2017, Zhang et al., 2006). A wide variety of automated and manual kits have been compared in these studies on human faeces including easyMag semi-automated DNA extraction kit; QIAamp DNA stool mini kit; QIAsymphony viral/bacterial midi kit; ZR faecal DNA miniprep; Ultraclean faecal DNA miniprep; PowerSoil DNA isolation kit; Extract Master faecal DNA extraction kit; Favor Prep stool DNA isolation kit; FastDNA kit; Nucleospin C + T kit; Quantum Prep Aquapure genomic DNA isolation kit and Promega Genomic Wizard DNA purification kit. One paper looked at both the storage and extraction method used (Nechvatal et al., 2008). In this instance, it was found that the storage of samples at ambient temperature in RNAlater followed by DNA extraction with QIAamp DNA stool mini kit yielded the highest amount of amplifiable DNA with little PCR inhibition. A study found that the QIAamp DNA stool mini kit yielded a higher amount of DNA from human faeces than the easyMag DNA extraction kit (Mirsepasi et al., 2014). They also found that the easyMag DNA extraction kit failed to remove some inhibitory compounds that reduced the success of PCR reactions. Another paper compared a wide range of commercial kits and found that they were all reproducible but that out of the kits tested the ZR faecal DNA miniprep kit had the best yield and diversity of DNA isolated (Claassen et al., 2013). Some papers also investigated the inclusion of pre-treatment steps such as bead beating. All the of the studies concluded that the inclusion of a mechanical disruption step such as bead beating improved the yield from commercial DNA extraction kits in comparison to using the kits with no pre-treatment (Smith et al., 2011, Salonen

et al., 2010, Lim et al., 2018). Two studies compared some of the most popular commercial techniques with newly developed non-commercial methods (Kumar et al., 2016, Zhang et al., 2006). Both found that the non-commercial methods were just as effective in extracting DNA and cheaper, however, the process was longer.
However, all the research above was performed on human faecal samples. A 2015 study compared five different extraction kits on faecal samples from five different species (Hart et al., 2015). They demonstrated that the success of the extraction kit depends on the species the sample is obtained from. Different kits had more success in different species. This demonstrates the importance of finding a kit that will be effective in obtaining DNA from bovine faecal samples. Multiple studies both found that the QIAamp stool DNA mini kit was effective in isolating DNA from cattle faecal samples to investigate the presence of certain bacterial pathogens (Inglis and Kalischuk, 2003, Kuhnert et al., 2005). A comparison of QIAamp stool DNA kit with traditional culture methods and a heat lysis non-commercial extraction protocol for identification of E. coli in cow faeces found that the QIAamp stool DNA kit and traditional culture method identified the same number of positives (Gioffré et al., 2004). However, fewer positives were identified with the non-commercial heat lysis extraction protocol.

1.4.4. Metagenome Sequencing

Sequencing is one tool that can be used to explore the calve metagenome. Traditionally culture-based methods have been used to detect the presence of bacterial populations. Both these methods allow detection of a large range of different microorganisms. However, not all bacteria are able to be grown easily or at all on culture so will remain undetected if culture-based methods are used (Yarza et al., 2014). Originally sequencing was performed by Sanger methods, but these were time-consuming taking years to carry out (Ansorge, 2009). Next-generation sequencing (NGS) methods have proven to be much quicker than Sanger sequencing, taking days or weeks instead of years to complete (Behjati and Tarpey, 2013). By determining sequence data from amplified single DNA fragments, NGS avoids the need to clone DNA fragments. In addition to this NGS has a higher throughput, is cheaper per base, captures a broader spectrum of gene mutations and reduces sequence errors in comparison to Sanger sequencing. The limitations to NGS are the computer capacity and storage required as well as a need for experienced personnel to carry out the sequencing and a large initial start-up cost.
NGS can be carried out in the form of 16S rRNA sequencing or whole-genome sequencing (WGS). Both have their own advantages and disadvantages as described in the literature (Petrosino et al., 2009, Ranjan et al., 2016). 16S rRNA sequencing is a cost-effective method that is useful for the identification of microbial composition and phylogenetic diversity of complex communities. With a large body of archived data and established pipelines for data analysis 16S rRNA sequencing is commonly used in microbial community characterisation. However, it lacks accuracy at a species level and is not sufficiently sensitive for comprehensive microbiome studies due to its inability to detect rare microorganisms in the community.
Whole-genome sequencing is a good alternative to 16S rRNA sequencing. WGS can assess microbial diversity and richness in mixed populations with more comprehensive coverage and higher throughput. WGS allows

identification and annotation of diverse arrays of microbial genes in order to investigate biochemical and metabolic functions. Additionally, WGS holds the potential to detect microorganisms from other kingdoms such as viruses, fungi and protozoa. Unfortunately, WGS is not without its disadvantages. A larger amount of starting material is required with a higher risk of contamination from the host material. The high number of unknown genes can make interpretation of WGS data increasingly challenging.

Many different platforms exist to carry out NGS. (Quail et al., 2012) compared three of the most widely used platforms, the Ion Torrent's PGM, Pacific Biosciences' RS and Illumina's MiSeq. A set of four microbial genomes were sequenced on each platform and the data acquired compared. All three platforms had almost perfect coverage on genomes that were GC rich, neutral and moderately AT-rich. However, the Ion Torrent's PGM had a bias in AT-rich genomes resulting in no coverage for 30% of the genome. Despite the Ion Torrent's PGM having the ability to call slightly more variants than Illumina's MiSeq it also had a higher false-positive rate. Pacific Bioscience's RS had a limited yield and cost per base in comparison to the other two platforms. The other two platforms (Ion Torrent and Illumina) were however closely matched in user utility and ease of workflow. Regardless of their differences all platforms produced usable sequences with a fast turnaround. One study noted that combining data after sequencing on two different platforms resulted in the best results (Lam et al., 2011). However, in most research, this technique will prove to be too cost-prohibitive and selection of a single platform on which to perform the sequencing needs to be done. Illumina is the most widely used and available platform with high throughput, low cost per base and broad utility present an attractive choice for NGS (Méndez-García et al., 2018).

1.5.Aims and Hypothesis

Since the metagenome in other species has been shown to influence health, including weight gain, it is reasonable to expect a similar pattern to be present in calves. However, the disease is also a major component in predicting lack of live weight gain in production animals. Since core temperature can be a very useful indicator of disease it would be interesting to see if it bears any relationship to weight gain in calves. Although a lot of work has been done on the metagenome of pre-weaning and post-weaning calves no other studies to our knowledge have examined the differences in metagenome that continue to develop post-weaning. This could be important in understanding why early weaning calves are more susceptible to disease and poor weight gain than their older counterparts.

The aims of this pilot study were:

1. Study the relation of the metagenome of post-weaning calves to their resulting weight gain over time.
2. Investigate the link between temperature and weight gain by monitoring both over time.
3. Compare the metagenome for 7 to 8-month-old calves and 14-week-old calves in order to understand more about how the calves' metagenome continues to mature post-weaning during the first year of life.

The hypothesis of this study is that calves showing poor weight gain will have altered abundances of core taxa in their metagenome. In addition to this, it is predicted that the calves with increased episodes of fever measured will show a

stunted weight gain. Also expected is that there will be a difference between the metagenomes of older and younger calves.

2. Materials and Methods

2.1. Farm Background and Subjects

The farm was an intensive dairy farm in Holsworthy, North West Devon. There was a total of 550 milking cows and 570 replacements. 20 Holstein-Friesian dairy calves aged between 7 and 8 months at the start of sampling were selected. The calves were kept in groups of 5 at birth and provided at least 4L colostrum within 4 hours of life. Until 8 weeks old the calves are provided with adlib NWF Ultra Milk Yellow Powder milk replacer (https://www.nwfagriculture.co.uk/product/nwf-ultra-milk-yellow/). The milk replacer is made up by mixing 125g milk powder with 875ml water, at 42°C to make 1L. The milk was fed at 39-40°C. Enough milk was provided to allow for 4 to 6 litres of milk per calf per day for the first 8 weeks. At 8 weeks of age, the calves were started to be weaned off the milk replacer. Between week 8 and week 11 of life milk replacer rations were gradually reduced. Weaning was completed between 11 and 12 weeks of age. Harper's Feeds (http://www.hhmix.co.uk/) supplied a specially made starter, with 16% protein content, called Calf Nut no.2. This was fed adlib from 7 days old up until 16 weeks old. After 16 weeks old, Calf Nut no. 2 was continued to be provided as advised on the bag. Straw was also provided ad-lib from 7 days old. Calves were weaned at 10 weeks old. From weaning to 12months old a mineral lick was provided. Water was provided ad lib throughout.

Once weaned calves were housed in a pen in an old free-standing barn. The barn was open at both ends of the concrete walkway. The sides of the barn had timber sidewalls with space boarding filling the space between the sidewalls and the roof. The roof was a tin apex roof with translucent panels to allow natural light during the daytime. There were 20 calves in each pen, and these were bedded down thickly with straw. The pens in each barn were separated from neighbouring pens by a feed alleyway and thick metal rails. Fresh silage was provided for the calves in the feed alleyways daily and pushed up by farm staff twice a day.

Figure 1: Calves in the barn on the dairy farm in Holsworthy

Figure 2: Calves lying in the barn on the dairy farm in Holsworthy, Devon

The calves were vaccinated against pneumonia, lungworm and infectious bovine rhinotracheitis (IBR) at 3 months old with Rispoval 4 (http://www.noahcompendium.co.uk/?id=-458198). The vaccine was administered as two intramuscular injections of 5ml per calf given 3 to 4 weeks apart. They were also covered against worms with Cydectin 10% LA solution for injection for cattle (http://www.noahcompendium.co.uk/?id=-456817). The wormers were given as a 0.5ml/50kg bodyweight subcutaneous injection in the ear.
In addition to the 20 calves aged between 7 and 8 months, an additional 12 younger calves were also selected on the final sampling date (12.12.2018).

These calves were around 14 weeks old at the time of sampling and underwent the same management as the older calves. These 14-week old calves were housed in a pen of 20 calves in a different barn with a similar layout. In total, we collected samples (rectal and nasal swabs, see chapter 2.3) and monitored health indicators (Wisconsin score, temperature, and weight, see chapter 2.3) from 32 calves. However, the metagenome analysis (see chapter 2.9) and presence of parasite (see chapter 2.8) was done on a subset of calves selected based on the results obtained by mining the data to find associations between the health indicator patterns.
Ethical permission was obtained from the School of Veterinary Medicine and Science, University of Nottingham (Approval Number: 2227 180226).

2.2. Measurement of Core Temperature and Detection of Drinking Events via Ruminal Boluses

2.2.1. Ruminal Boluses

The 20 calves aged between 7 and 8 months had temperature boluses put in on the 11th July using a balling gun operated by a trained professional from prognostiX (https://www.prognostix-uk.com/). PrognostiX is a company, based in Devon, UK. They aim to create collaborations between veterinary surgeons, hardware engineers, software engineers, and data analysts in order to develop easy to use and cost-effective software that farmers can utilise as part of a management plan to improve health and welfare on farms. By using smart technology, they want to predict health events in order to prevent disease. The boluses (https://www.prognostix-uk.com/remote-weighing-and-temperature-sensing-to-transform-calf-production/, PrognostiX, UK) worked via radio-telemeter with a radio transmitter, receiver, and antenna. The temperature was measured and transmitted every 2 minutes up to a range of 30m. The boluses were 2.5cm in diameter x 9.5cm in length (https://www.y-ware.org/).

Figure 3: Bolus against a standard ruler

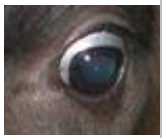
The younger calves were around 14 weeks old when an updated version of the boluses was administered on the 28.11.2018. The updated version of the bolus was smaller in size so that it was able to be applied to calves as young as these. As with the older calves, this procedure was carried out by a trained professional from prognostiX.

2.3. Sampling Procedure – Health Indicators (Wisconsin Score and Weight) and Swabs

The 20 older calves were sampled every 2 weeks for 3 consecutive sampling dates (14.11.2019, 28.11.2019 and 12.12.2019). On the last sampling date (12.12.2019) 12 younger calves were also sampled. The calves due to be sampled on each date were run one by one into a mechanical crush for restraint during the sampling procedure. Figure 4 shows the weekly sampling process. Sampling was carried out as follows; the calf ID was read off the ear tag and recorded on the sampling sheet along with the time at the point of sampling.

Figure 4: Sampling process

The Wisconsin clinical scoring system (Table 1) was used to predict the calves' health (McGuirk, 2008). In this system, five clinical signs (nasal discharge, ocular discharge, coughing, faecal consistency, ear position, naval and joint swelling) are assessed and given a score between 0 and 3 depending on severity. The system is displayed clearly in Table 1 along with images showing examples of each of the scores for each parameter. The nasal score was removed for the purposes of this study as all calves were old enough to have completely healed navels. This system was chosen as it is a quick and simple method that has been shown to be sensitive, specific and accurate in identifying respiratory disease in calves (Love et al., 2014, Buczinski et al., 2015). Calves with a total clinical score ≥5 are classed as clinically ill while calves scoring ≤ 3 are considered healthy (Buczinski et al., 2014). Since the calves were being handled in order to perform the sampling less labour-intensive scoring systems wasn't required (Aly et al., 2014)

Clinical Parameter	0	1	2	3
Nasal Discharge	Normal clear discharge 	A small amount of unilateral cloudy discharge 	Bilateral cloudy or excessive mucous 	Copious bilateral mucopurulent nasal discharge 
Ocular Discharge	Normal 	Mild ocular discharge 	Moderate bilateral ocular discharge 	Heavy ocular discharge 
Ear Position	Normal	Ear flicking	Slight unilateral ear droop	Severe head tilt or bilateral ear droop 

				
Cough	No cough	Induced single cough	Induce repeated coughs or occasional spontaneous cough	Repeated spontaneous coughing
Faeces	Normal	Semi-formed, pasty	Loose but stays on top of the bedding	Watery, sifts through bedding
Navel	Normal	Slightly enlarged, not warm or painful	Slightly enlarged with slight pain and moisture	Enlarged with pain, heat or malodourous discharge
Joints	Normal	Slight swelling, not warm or painful	Swelling with pain or heat, slight lameness	Swelling with severe heat, pain and lameness

Table 1: Wisconsin Scoring Criteria (http://www.vetmed.wisc.edu/dms/fapm/fapmtools/8calf/calf_respiratory_scoring_chart.pdf)

Eppendorf tubes (Eppendorf UK Limited, UK) were labelled for each animal and sample type. For example, animal 7038 on the 14th November 2019 had 3 Eppendorf tubes labelled 1A to place nasal swab tips in and 3 tubes labelled 1B to place rectal swab tips in. This was the sample ID and was recorded in the sampling table along with the clinical scoring, time, date and calf ID. Each Eppendorf tube contained 3ml RNAlater solution from Qiagen (https://www.qiagen.com/gb/shop/sample-technologies/rna/total-rna/rnalater-rna-stabilization-reagent/?cmpid=Qven10GARNAstabilizationGA_SEM_&kwid=rnalater&gclid=CjwKCAjw1KLkBRBZEiwARzyE77wus3yA9wonom7SKwn-5UimA8EvxLl3O5TFlnG7y_qCxI8TFbKctxoCzS4QAvD_BwE&clear=true#orderinginformation, QIAGEN, UK) in order to preserve the samples. This allowed the samples to be kept at room temperature for up to 12 hours before being transferred to 4-12℃ before later freezing at -20℃.
Next, a clean pair of gloves (Kimberly-Clark Limited, UK) were put on before taking 3 nasal swabs (Alpha Laboratories, UK). The swabs used were 15cm long with wooden handles and cotton tips. Swabs were taken by restraining the animals head with one arm and using the other hand to carefully insert the swab about 5cm into the nostril before rotating against the mucosa for a few seconds. The swab was then removed from the nostril and the cotton tip broken off into a labelled Eppendorf tube containing 3ml of RNAlater (Figure 5). Care was taken to ensure the swab tips did not make contact with the external nares during the sampling process. Each labelled Eppendorf tube contained 1 swab. The lids were kept on the Eppendorf tubes at all other times to prevent contamination and care was taken not to spill any of the RNAlater solutions.

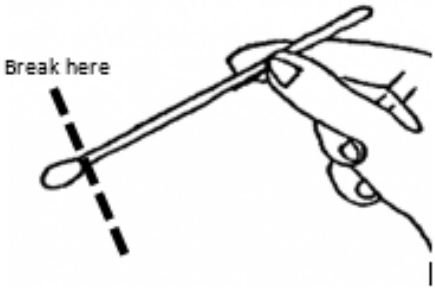


Figure 5: Swabbing Protocol

After changing to a clean pair of gloves again 3 rectal swabs were taken. A similar procedure was followed inserting the swab about 5cm into the rectum before rotating against the mucosa for a few seconds. The cotton tips were

again broken off into labelled Eppendorf tubes containing 3ml RNAlater (Figure 5). Again, care was taken to ensure the tips of the swabs did not make contact with the skin surrounding the anus.
A blank swab was also placed in an Eppendorf tube containing 3ml RNAlater. This was labelled "C" and acted as a control sample.
The samples were stored at 4-12℃ for up to 5 days. The tubes were always kept upright and out of direct sunlight to ensure the samples were completely covered and preserved correctly in the RNAlater solution. After this time the samples were transported to the laboratory and frozen at -20℃ until DNA or RNA extraction.
Finally, the calf's weight was measured and recorded using an MSD animal health cattle band (MSD animal health, UK). The band was suitable for calves ranging from 34kg to 902kg. The measurement was taken by placing the weigh band around the calf's chest behind the front legs and shoulder blades. It was held flat against the skin with no twists in the band. The measurement was carried out by the same person every time to ensure consistency of the tautness of the band and therefore reliability of the measurements.
Once all the samples and measurements had been taken and recorded the calves were released from the crush and returned to their pen.

2.4. DNA Extraction – Pilot Analysis to Select the Best Protocols for DNA Extraction

Firstly, a pilot experimental analysis was done to identify which DNA kit was better performing at extracting the DNA from the samples and to select the best performing kit for subsequent large-scale DNA extraction. Two kits were considered, the QIAamp DNA Stool Mini kit (https://www.qiagen.com/gb/shop/sample-technologies/dna/genomic-dna/qiaamp-dna-stool-mini-kit/#orderinginformation, QIAGEN, UK) and the PowerFaecal DNA Isolation Kit (https://mobio.com/us/shop/dna-isolation/di-fecal-microbiome/qiaamp-powerfecal-dna-kit/, MoBio, US). Four samples were selected (C, 1B, 2B, and 3B) to be compared. The QIAamp DNA Stool Mini Kit was tested first. A few modifications were made, including the addition of a bead-beating step prior to extraction. 700µl Buffer AS was added to a labelled sterile 2ml microcentrifuge tube along with a sterilised 0.5mm steel bead. The swab tip was placed in the tube and then a 2-minute bead beating cycle carried out on a TissueLyser II (https://www.qiagen.com/us/shop/automated-solutions/sample-disruption/tissuelyser-ii/#orderinginformation, QIAGEN, UK) with 0.5mm steel beads (QIAGEN, UK) at a frequency of 30Hz. The remaining 700µl Buffer ASL was added to each tube before continuing according to the manufacturer's instructions. The final elution volume was reduced by adding only 60µl Buffer AE to the QIAamp spin column membrane. The eluted DNA was stored at -20℃ until further processing.
The second kit used to extract DNA from the rectal swabs was the QIAamp PowerFaecal DNA Isolation Kit (https://mobio.com/us/shop/dna-isolation/di-fecal-microbiome/qiaamp-powerfecal-dna-kit/, MoBio, US). This was done again on technical replicates of C, 1B, 2B and 3B. The DNA was isolated with a few modifications to the manufacturer's instructions. Briefly, the cotton tip of the swab was removed from the RNAlater and placed in the Bead Tube using a pair of sterilised forceps. 750µl PowerBead solution was then added along with 60µl Solution C1 before vortexing briefly. The procedure was then

followed according to the manufacturers' instructions. The eluted DNA was then frozen at -20℃ until further processing.
The eluted DNA from the two kits was compared via both 16S bacterial PCR and nanodrop quantification (see chapter 2.6). From these tests, we selected the QIAamp PowerFaecal DNA Isolation kit to extract DNA from the remainder of the samples.

2.5. RNA Extraction

To identify if RNA viruses were also present in the samples, we also extracted the RNA was from the same samples from which we extracted the DNA (see chapter 2.419).
The RNA was extracted from the rectal swabs using the QIAamp Viral RNA Mini Kit (https://www.qiagen.com/gb/shop/sample-technologies/rna/viral-rna/qiaamp-viral-rna-mini-kit/#orderinginformation, QIAGEN, UK) according to the manufacturer's protocol with a few modifications. Briefly, buffer AVL containing carrier RNA mix was pipetted into labelled sterile 1.5ml microcentrifuge tubes and the swab tip added. The tubes were vortexed for 30 seconds then incubated at room temperature for 10 minutes. The manufacturer's instructions were then followed until the final elution step. In order to elute the RNA 60µl Buffer AVE was applied to each column membrane before centrifuging at 6,000 x g for 1 minute. The eluted RNA was labelled and stored at -20℃ until further processing.

2.6. Visualization and Quantification of Extracted DNA and RNA

A 16S rRNA universal PCR was performed, on the DNA extracted from the 5 samples previously described, in order to show that bacterial DNA had been isolated. The samples selected were two eluted with the PowerFaecal kit plus a negative control eluted with the PowerFaecal kit and 2 eluted with the QIAGEN DNA kit. The surfaces and pipettes were cleaned with ethanol prior to performing the PCR. The eluted DNA samples were removed from the freezer and defrosted at room temperature. Sterile 0.2ml PCR tubes were prepared and labelled. The master mix used was QIAGEN Multiplex Master Mix (https://www.qiagen.com/gb/shop/pcr/end-point-pcr-enzymes-and-kits/regular-pcr/qiagen-multiplex-pcr-kit/#orderinginformation, QIAGEN, UK). 5'- ACG GCC CAG ACT CCT ACG GGA GGC -3' was used as the forward primer while 5'- GGC TTG CGC TCG TTG CGG CAC TTA -3' was used as the reverse primer (Wilson et al., 1990, Franciosa et al., 1994). In each tube, 5µl master mix, 1µl forward primer, 1µl reverse primer, 1µl of water and 2µl of the DNA sample was pipetted in order to make up a 10µl reaction volume. Each tube was then vortexed briefly before spinning at full speed for 10 seconds. The tubes were then amplified in an AppliedBiosystems 2720 Thermal Cycler (https://www.fishersci.co.uk/shop/products/applied-biosystems-2720-thermal-cycler-range-10-100-l/12313653, Fisher Scientific, UK) for 30 cycles of. The cycling conditions were as follows: initial denaturation at 94℃ for 5 minutes, denaturation at 94℃ for 30 seconds, annealing at 55℃ for 30 seconds, extension at 72℃ for 45 seconds then final extension at 72℃ for 7 minutes. Once the reaction was complete the samples were held at 4℃ before being frozen at -20℃ prior to running the gel.

The gel mould was then assembled. 0.80g agarose (https://www.sigmaaldrich.com/catalog/product/sigma/a9539?lang=en&region=GB, Sigma Aldrich, US) and 80ml TAE x 1 were microwaved in a conical flask for 2 minutes. After cooling briefly under a running tap 3µl Nancy-520 (https://www.sigmaaldrich.com/catalog/product/sigma/01494?lang=en&region=GB), Sigma Aldrich, US) was added before pouring into the gel mould and leaving to set. Any bubbles in the poured agarose gel should be popped with a sterile pipette tip. The amplified DNA samples were prepared by adding 1µl BioLab's gel loading dye in blue (https://international.neb.com/products/b7021-gel-loading-dye-blue-6x#Product%20Information, Bio Laboratories, UK), pipetting up and down to mix, and spinning in a microcentrifuge for a few seconds. 500ml TAE was poured into a horizontal submerged electrophoresis cell (http://www.bio-rad.com/en-uk/sku/1704468-wide-mini-sub-cell-gt-horizontal-electrophoresis-system-15-x-10-cm-tray?ID=1704468, Bio-Rad Laboratories, US) connected to a BioRad PowerPac Basic Power Supply (http://www.bio-rad.com/en-uk/product/powerpac-basic-power-supply?ID=bea5dea1-cef0-43ad-8af5-b2c0287f6e07, Bio-Rad Laboratories, US). The agarose gel was removed from the mould and placed in the tank of TAE. 7µl of 100bp BioLabs Quick Load 100bp DNA ladder (https://international.neb.com/products/n0467-quick-load-100-bp-dna-ladder#Product%20Information, Bio Laboratories, UK) was pipetted into the first well followed by 7µl of negative control in the second well. The eluted DNA samples were then pipetted in the remaining wells and their position noted. The gel was set to run at 80V (400mA) for 111 minutes. Once the run was complete the agarose was removed from the tank and placed in the ImageQuant 300 Imager (http://www.promix.ru/manuf/ge/analise/imagers/ImageQuant300.pdf, GE healthcare life sciences, UK), with the filter-wheel set to 2 and the IQuant capture set to 300, to visualise the PCR bands.

The concentration of DNA or RNA in the eluted samples were quantified using a nanodrop 8000 spectrophotometer (https://assets.thermofisher.com/TFS-Assets/CAD/Product-Guides/User-Manual-nd-8000-v2.2-users-manual-8.5-x-11.pdf, Thermo Fisher Scientific, UK). The samples were defrosted at room temperature and placed on ice. The elution buffers were also placed on ice. The nanodrop machine was cleaned and wiped with reverse osmosis water before logging into the program on the computer, ND-8000 V2.2.0. The 8 samples nucleic acid option on the program was selected before running the nanodrop with RNAase free water. Next, the nanodrop was blanked with the elution buffer specific to each extraction method (Solution 6 for DNA) and set to measure DNA. The nanodrop was wiped with a clean tissue before loading the DNA samples then measured. The well in which each sample was in was noted in order to record the concentration of each sample. Samples continued to be loaded and measured, ensuring to wipe the nanodrop between each load until all the DNA samples were measured. The procedure was then repeated with RNA starting with blanking using the AVL elution buffer.

2.7. Sample Selection

Given the high cost of the metagenomic analysis, we have selected a subset of animals to be further analysed for the presence of parasites (see chapter 2.8) and for the presence of specific microorganism (metagenomic analysis, see chapter 2.10.3). In total, we have decided to investigate only samples collected from 11 calves out of the 32 calves under analysis due to the high cost of shotgun metagenome sequencing as well as the relatively short time frame available for analysis of the sequences produced. The selection included 8 of the older calves and 3 of the younger ones. To select which calves had more interesting health patterns we considered the following procedure: Firstly, we considered the weight and Wisconsin scores collected from the calves. The calves' weight over the 4 weeks as well as weight gain every 2 weeks was imported into a table in Excel. The table was then sorted by total weight gain in order to identify the calves that displayed poor weight gain. The median, 1st quartile and 3rd quartile of the total weight gain was also calculated. A ratio between weight gain in week 0 to week 2 and weight gain in week 2 to week 4 was worked out. Using the ratio and median calves with a normal consistent weight gain or above average but still consistent weight gain were identified. The ratio was also used to pick out calves that showed very uneven weight gain across the two timeframes. In total 8 older calves were selected. Samples from the 14.11.2018 and 12.12.2018 were used for these 8 calves.
The younger calves were sorted by weight estimated on the 12.12.2018 in excel. The median weight was calculated as well as the 1st and 3rd quartiles. The calves with a Wisconsin score of >0 were filtered out. From the remaining calves, 4 were selected based on the proximity of weight to the median.
The 16S PCR was repeated on all the selected samples to confirm the presence of bacterial DNA.
From this selection process, we hoped to sequence samples from 4 older calves who were clinically healthy according to the Wisconsin scoring (0) and had an even weight gain close to the desirable daily live weight gain. These 4 calves acted as our normal controls. The remaining 4 older calves were selected because they demonstrated a weight gain of uneven distribution or very poor weight gain. These calves were considered examples of unhealthy calves. The 3 younger calves that were selected were again clinically normal (0) according to Wisconsin scoring with weights close to the median for their age. These 3 "Healthy" younger calf samples were for comparison with the 4 older "normal controls" to investigate any differences between the two age groups.

2.8. Testing for Cryptosporidium Parasite Using PCR

The presence of the Cryptosporidium parasite was tested using a nested PCR following the protocols described in (Santın et al., 2004, Xiao et al., 1999). The specific nested PCR is designed to target and amplify a18S rDNA from the Cryptosporidium parasite.

2.8.1. Selection of Positive Controls

We first tested, via nested PCR (Santın et al., 2004, Xiao et al., 1999) and sequencing, for the quality of our in-house samples of DNA to act as positive controls. Specifically, several DNA samples extracted from Cryptosporidium isolated from a clinically affected animal that was donated by Dr A. Zintl from the School of Veterinary Medicine, University College Dublin – Ireland were analysed.

For the Nested PCR, the first PCR was done using the forward primer Crypto_1F (5'- TTC TAG AGC TAA TAC ATG CG -3') and the reverse primer Crypto_1R (5'- CCC TAA TCC TTC GAA ACA GGA -3'). The second PCR reaction then used Crypto_2F (5'- GGA AGG GTT GTA TTT ATT AGA TAA AG -3') as the forward primer and Crypto_2R (5'- AAG GAG TAA GGA ACA ACC TCCA -3') as the reverse primer. Each primer was used at a concentration of 10µM solution for the PCR reaction. A 10µl PCR reaction was prepared for the first PCR. 15µl of 2x QIAGEN Multiplex PCR Master Mix was pipetted into a 1.5ml Eppendorf tube along with 3µl nuclease-free water, 3µl Forward primer 1(Crypto_F1) and 3µl Reverse primer 1 (Crypto_R1). Each addition to the tube was pipetted up and down multiple times to mix the solution. 8µl of the mix produced was then pipetted into clean labelled 0.2ml PCR tubes. The tubes were then placed in an AppliedBiosystems 2720 Thermal Cycler (https://www.fishersci.co.uk/shop/products/applied-biosystems-2720-thermal-cycler-range-10-100-l/12313653, Fisher Scientific, UK) under the following conditions for the first PCR reaction: initial denaturation at 94℃ for 3 minutes, denaturation at 94℃ for 45 seconds, annealing at 55℃ for 45 seconds, extension at 72℃ for 1 minute and then final extension at 72℃ for 7 minutes. The first PCR reaction produced a product with a length of 1325 base pairs that was visualized on a 0.9% TAE agarose gel and using the BioLabs Quick Load 100bp DNA Ladder as a ladder.
Once completed a mix was prepared for the second reaction. This time the reaction volume was 50µl. Into a clean 1.5ml Eppendorf tube 75µl of 2x QIAGEN Multiplex PCR Master Mix (https://www.qiagen.com/gb/shop/pcr/end-point-pcr-enzymes-and-kits/regular-pcr/qiagen-multiplex-pcr-kit/#orderinginformation, QIAGEN, US), 39µl nuclease-free water, 15µl forward primer 2 (Crypto_F2) and 15µl reverse primer 2 (Crypto_R2) were pipetted. After each addition, the mix was pipetted up and down multiple times to mix. The mix was then transferred into clean labelled 0.2ml PCR tubes with 48µl of the mix in each tube. 2µl of the PCR product from the initial PCR reaction was pipetted into each PCR tube. The tubes were then placed in an AppliedBiosystems 2720 Thermal Cycler under the same conditions as the initial reaction (initial denaturation at 94℃ for 3 minutes, denaturation at 94℃ for 45 seconds, annealing at 55℃ for 45 seconds, extension at 72℃ for 1 minute and then final extension at 72℃ for 7 minutes). The second reaction produced a product with a length of 830 base pairs that was visualized on a TAE 0.9% and using the BioLabs Quick Load 100bp DNA Ladder as a ladder.
The electrophoresis gel was then run for 120 minutes at 80V and 400mA. Once finished the gel was removed and visualised in an ImageQuant 300 Imager (http://www.promix.ru/manuf/ge/analise/imagers/ImageQuant300.pdf, GE healthcare life sciences, UK) with the filter-wheel set to 2 and the IQuant capture set to 300.

Finally, in order to confirm that the 830bp band seen on the gel of the positive control sample was Cryptosporidium the 830bp PCR product was sent for sequencing. To do so we first purified the PCR product using the QIAquick PCR purification kit (https://www.qiagen.com/gb/shop/sample-technologies/dna/dna-clean-up/qiaquick-pcr-purification-

kit/#orderinginformation, QIAGEN, US) according to the manufacturer's instructions.
Then, the samples were diluted to 10ng/µl with nuclease-free water and pipetting 5µl into a labelled PCR tube. 5µl of each of the primers diluted to 3.2pmol/µl were also pipetted into labelled PCR tubes. The primers and the samples were submitted to Source BioScience for overnight Sanger Sequencing (https://www.sourcebioscience.com/services/genomics/sanger-sequencing-services/, Source BioScience, UK) in order to determine whether the PCR product was 18S rDNA from the Cryptosporidium parasite. The sequencing confirmed the PCR being 18S rDNA from Cryptosporidium.

2.8.2. PCR Screening for Detection of Cryptosporidium in the collected DNA Samples

The 20 selected DNA samples were then tested for the Cryptosporidium parasite using a PCR alongside the positive control and negative control (nuclease-free water instead of DNA). Since during our experimental analysis of the controls, we noticed that the second PCR was also able to produce a band of 830bp without performing the first PCR we decided to analyse the DNA samples using only the second PCR. All reactions and PCR settings have been done as previously described for the positive control.

Calf ID	Age	Sample Collection Date	Sample ID	PCR Tube Label
NA	NA	NA	Nuclease free water	N (negative)
NA	NA	NA	Cryptosporidium DNA	P (positive)
NA	NA	NA	C	1
7212	14 weeks	12.12.2018	2E	2
7231	14 weeks	12.12.2018	11E	3
7189	14 weeks	12.12.2018	12E	4
7115	7-8 months	14.11.2018	3B	5
7099	7-8 months	14.11.2018	6B	6
7005	7-8 months	14.11.2018	8B	7
7129	7-8 months	14.11.2018	9B	8
7095	7-8 months	14.11.2018	11B	9
7069	7-8 months	14.11.2018	14B	10
7118	7-8 months	14.11.2018	17B	11
7124	7-8 months	14.11.2018	18B	12
7115	7-8 months	12.12.2018	45B	13
7099	7-8 months	12.12.2018	48B	14
7005	7-8 months	12.12.2018	50B	15
7129	7-8 months	12.12.2018	51B	16
7095	7-8 months	12.12.2018	53B	17
7069	7-8 months	12.12.2018	56B	18
7118	7-8 months	12.12.2018	59B	19
7124	7-8 months	12.12.2018	60B	20

Table 2: Labelling of PCR Tubes

2.9. Metagenome analysis

2.9.1. Sample Preparation

The samples selected for sequencing were removed from the freezer and defrosted on ice. Each sample was allocated a new number from 1 to 20 in order. 20µl of each sample were shipped on dry ice to The University of Leeds Sequencing Centre.

2.9.2. Library Preparation and Metagenome Sequencing

Metagenome sequencing and library preparation were done by the Leeds Institute of Molecular Medicine at the University of Leeds. Prior to library preparation quantitation of the DNA in each sample was carried out using Quant IT DNA BR reagent (https://www.thermofisher.com/order/catalog/product/Q33130, Thermo Fisher, UK).

2.9.2.1. DNA Shearing

The Covaris E220 (https://covaris.com/products/afa-ultrasonication/e-series/, Covaris, UK) and the Covaris MicroTUBE Plate (cat no 520078) (https://covaris.com/products/afa-tubes-and-vials/microtube/, Covaris, UK) were used to shear the DNA to an average of 200bp with the following parameters:

- Duty Factor – 10
- Peak Power – 175
- Cycles/Burst – 200
- Time – 120sec.

This was carried out in a volume of 50µl made up of DNA sample and nuclease-free water. The volume of DNA sample and nuclease-free water used for each sample varied due to variations in the concentration of DNA in each sample. 2µl of the sheared DNA was used to perform a quality control check on the HSD1000 tapes from Agilent (https://www.agilent.com/en/product/tapestation-automated-electrophoresis/tapestation-dna-screentape-reagents/high-sensitivity-dna-screentape-analysis-228262, Agilent, US). The samples were then transferred to a 96 well plate and 7.5µl nuclease-free water added to make the volume up to 55.5µl ready for the library prep.

2.9.2.2. NEBNextUltra DNA Library Preparation

The NEBNextUltra DNA Library Prep Kit for Illumina (https://www.neb.com/products/e7370-nebnext-ultra-dna-library-prep-kit-for-illumina, NEB, US) was used in order to convert the DNA into indexed libraries ready to perform next-generation sequencing on an Illumina platform. The procedure was performed according to the kit instructions with a few modifications as described below.

2.9.2.2.1. NEBNext End Prep

3µl End Prep Enzyme Mix, 6.5µl End Repair Reaction Buffer (x10) and 55.5µl of fragmented DNA sample were added to a sterile nuclease-free tube and pipetted up and down at least 10 times to thoroughly mix. The tubes were then spun quickly and placed in a thermocycler, with the heated lid set to ≥75°C, to

run for 30 minutes at 20℃, followed by 30 minutes at 65℃ then held at 4℃ until the next step.

2.9.2.2.2. Adaptor Ligation

15µl of Blunt/TA Ligase Master Mix, 2.5 µl NEBNext Adaptor for Illumina and 1 µl of Ligation Enhancer were added to the End Prep reaction mixture and mixed well by pipetting up and down. The samples were quickly spun then incubated for 15 minutes at 20℃ in a thermal cycler. 3 µl USER Enzyme was added to the ligation mixture, mixed well and incubated at 37℃ for 15 minutes with the heated lid set to ≥47℃.

2.9.2.2.3. Size Selection of Adaptor-ligated DNA

The samples were size selected for a 200bp insert size. SPRIselect beads were vortexed to re-suspend. 13µl of dH_2O was added to the ligation reaction followed by 55µl (0.55X) of re-suspended SPRIselect beads and mixed by pipetting up and down. The samples were then incubated at room temperature for 5 minutes before placing the tube on a magnetic stand to separate the beads from the supernatant. Once the solution was clear the supernatant was transferred to a new tube and the beads discarded. 25µl re-suspended SPRIselect beads were then added to the supernatant and mix thoroughly. The samples were incubated at room temperature for 5 minutes before placing on a magnetic stand to separate the beads from the supernatant. Next, the supernatant that contains unwanted DNA was discarded taking care not to disturb the beads containing the desired DNA. With the tube still in the magnetic stand 200µl of freshly prepared 80% ethanol was added to the tube containing the beads. The solution was incubated at room temperature for 30 seconds then the supernatant was removed and discarded. 200µl freshly prepared 80% ethanol was added to the tube again, incubated for 30 seconds at room temperature and then the supernatant removed and discarded. The beads were air-dried for 5 minutes by opening the tube lid while still on the magnetic stand. The tube was then removed from the magnetic plate and the DNA target was eluted from the beads into 17µl of 10mM Tris-HCl. The solution was mixed well and incubated for 2 minutes at room temperature then placed on a magnetic stand. After 5 minutes 15µl of the solution was transferred to a new PCR tube ready for amplification.

2.9.2.2.4. PCR Enrichment of Adaptor-ligated DNA

15µl Adaptor Ligated DNA Fragments (from the previous step), 25µl NEBNext Q5 Start HiFi PCR Master Mix, 5µl Index Primer and 5µl Universal PCR Primer were added to a sterile strip tube. The solution was mixed thoroughly and quickly spun. The tubes were placed in a thermocycler and PCR amplification performed under the conditions in Table 3.

Cycle Step	Temperature (°C)	Time (seconds)	No. Cycles
Initial denaturation	98	30	1
Denaturation Annealing/Extension	98 65	10 75	4-12 4-12
Final extension	65	5	1
Hold	4	∞	

Table 3: Table showing PCR Conditions for PCR Enrichment of Adaptor Ligated DNA

The Denaturation and Annealing/Extension cycle were performed 8 times.

2.9.2.3. Illumina Sequencing

Sequencing was performed on 2 lanes of HiSeq 3000 (https://emea.illumina.com/systems/sequencing-platforms/hiseq-3000-4000/specifications.html, Illumina, US) with 150bp pair ended. The final library was quantified with qPCR and diluted down to 2nM. SBS reagents, indexing reagents and paired-end reagents were all prepared. Next, the run parameters were entered as prompted by the software. The reagents were then loaded onto the instrument and after confirmation of proper flow, the reagents were primed. The priming flow cell was then removed, the flow cell holder cleaned, and a sequencing flow cell loaded. The proper flow was confirmed again before starting the sequencing run. From this 30M 150bp pair ended inserts per sample were produced in a FastQ format.

2.10. Data Analysis

2.10.1. Analysis of Temperature Data from Ruminal Boluses

2.10.1.1. Drinking Behaviour

Drinking behaviour of the calves was obtained via the temperature data. Reported in the literature that during drinking events the rumen temperature of cattle drops. Most studies report that a drinking event can be presumed whenever the temperature drops below 37.7°C. This was validated recently and confirmed to occur in calves as well as mature cattle, by the University of Nottingham. Using this information, the temperature data from the boluses were examined using Matlab (Matlab and Statistics Toolbox Release. 2017a. The MathWorks, Inc., Natick, Massachusetts, United States). All points that dropped below 37.7°C were classed as drinking events and removed from the data.

2.10.1.2. Core Temperature

The data collected on the farm was imported into Excel. Using the plotting function in Excel pie graphs and line graphs were created in order to visualise the Wisconsin Scores and weight data.
The visualisations and codes used below were obtained using a custom-made script in Matlab (Matlab and Statistics Toolbox Release. 2017a. The MathWorks, Inc., Natick, Massachusetts, United States).
Histogram plots were generated from the temperature data recorded on the 14th of November 2018. These plots were done in order to understand the underlying shape of the reticular temperature for the individual calves on this date since sampling also occurred on this date. Histograms were normalised and temperature data from drinking events removed to allow only the baseline temperature distribution of the calves to be computed. Drops in temperature

were classed as drinking events if they were recorded below 37.7°C (Bewley et al., 2008a, Bewley et al., 2008b). A thick red line at 40.5°C indicated the temperature threshold at 1°C above the upper limit of a calf's normal core temperature of 39.5°C. Values measured above this line could indicate the presence of a fever. Temperature histograms were also plotted for the 7 days preceding the 14th November 2018. The purpose of these plots was in order to visualise the temperature distribution per day over time and identify any variation between different calves. Additionally, comparisons were made for individual calves between different days using an ANOVA. This allowed us to see if there was any temperature variation between different days for the same calf.
A t-test between core temperature on the 7th November 2018 and 14th November 2018 was performed for each to investigate if there was any significant difference between core temperatures between these dates.

2.10.2. Wisconsin Scores and weight data

The data collected on the farm was imported into Excel. Using the plotting function in Excel pie graphs and line graphs were created in order to visualise the Wisconsin Scores and weight data.

2.10.3. Bioinformatics - Metagenomics

Sequences were analysed and processed with the assistance and advice of Andrew Warry from the University of Nottingham's Advanced Data Analysis Centre (https://www.nottingham.ac.uk/adac/index.aspx). The first quality was assessed, and quality control measures are undertaken using the fastp pre-processor tool (Zhou et al., 2018). Quality control steps included trimming adaptor and filler sequences, removing low complexity reads and filtering out sequences composed of >10% N bases. The data was also aligned and filter to remove sequences coding for a bovine genome that represents host contamination. The success of the removal process was evaluated using Centrifuge(Kim et al., 2016). Taxonomic classification was carried out using a program called, Kaiju (Menzel et al., 2016) in order to determine the micro-organisms, present and in which proportions. Unclassified reads were sequences that the software was unable to identify at any taxonomic level. Whereas unassigned reads were sequences that the software could identify at some taxonomic level but could not be identified at the specific taxonomic levels this study focused on (Family and Order). For example, sequences that the software identified at the phylum level (classified) may not be identifiable to the genus level (unassigned).
Sourmash (Brown and Irber, 2016) was used to compare the similarity of the sequences obtained from the different samples. From these comparisons, heat maps were created to visualise the results.
The creation of taxonomic classification graphs was done with custom R code using the Plotly graphing library (https://www.plot.ly/r/bar-charts/). Excel was also used to create stacked bar charts of the top 10 most abundant classes, families, orders, and genera of microorganisms in each sample.

3. Results

3.1. Sampling Data

In total 20 older calves were sampled on 14.11.2018, 28.11.2018 and 12.12.2018. An additional 12 younger calves were samples on the 12.12.2018. 3 nasal swabs and 3 rectal swabs were taken from each calf on each of the dates as well as a Wisconsin score and weight estimate. This means that in total over the three sampling dates 72 Wisconsin scores and weights were recorded, 60 for the older calves and 12 for the younger calves. The raw data collected is shown in appendix 1. The columns for faecal consistency, joint swelling, and ear position were omitted as all calves on all dates scored 0 for these criteria. All the sampling data is shown in the following tables and figures.

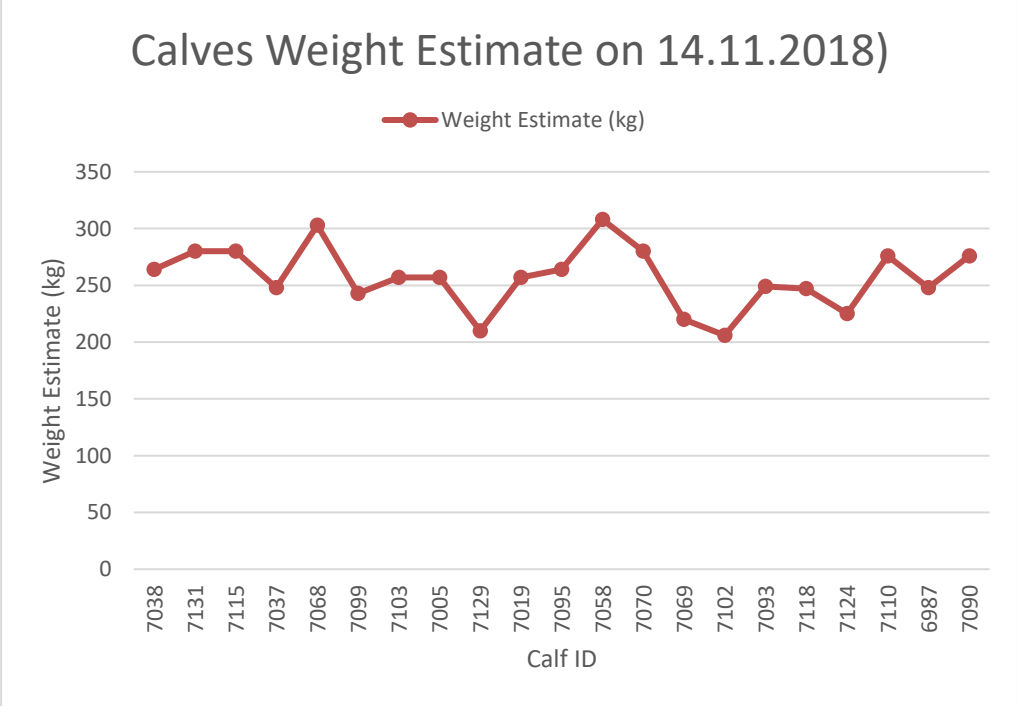


Figure 6: Calves weight estimate according to weigh band on 14.11.2018

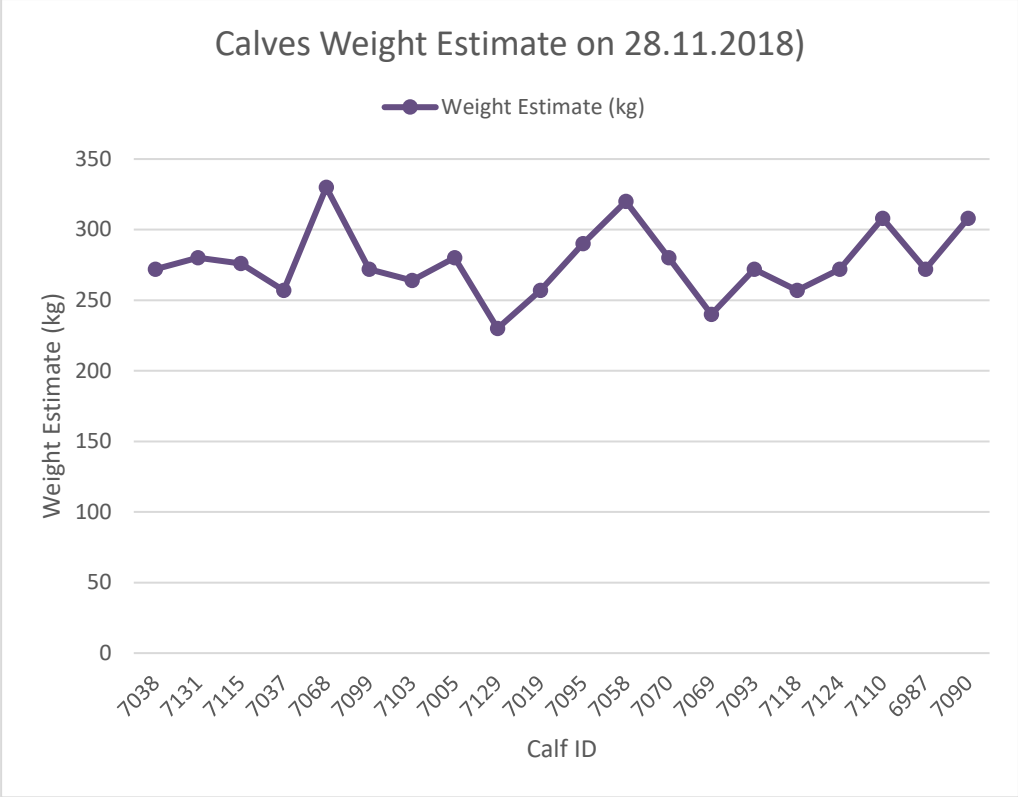


Figure 7: Calves weight estimate according to weigh band on 28.11.2018

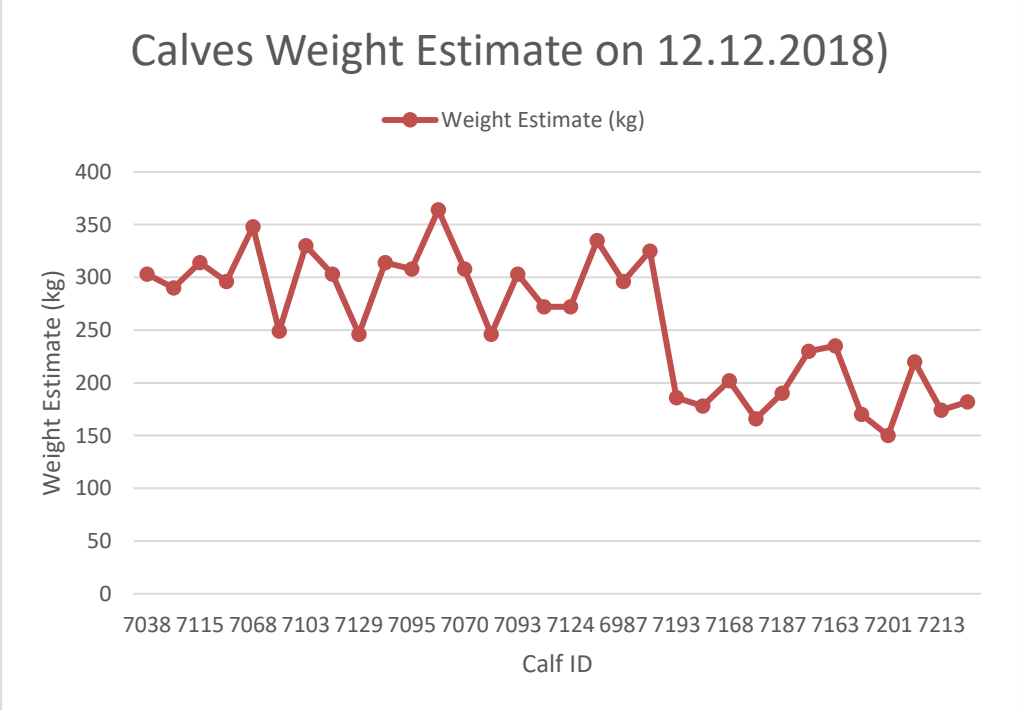


Figure 8: Calves weight estimate according to weigh band on 12.12.2018

Date	No. Calves Scoring Nasal Discharge				
	0	1	2	3	4

14.11.2018	7	11	3	0	0
28.11.2018	8	10	2	0	0
12.12.2018	24	8	0	0	0

Table 4: Number of calves receiving each score for nasal discharge

Date	No. Calves Scoring Ocular Discharge				
	0	**1**	**2**	**3**	**4**
14.11.2018	11	7	3	0	0
28.11.2018	15	5	0	0	0
12.12.2018	27	2	3	0	0

Table 5: Number of calves receiving each score for ocular discharge

Date	No. Calves Scoring for Coughing				
	0	**1**	**2**	**3**	**4**
14.11.2018	18	3	0	0	0
28.11.2018	20	0	0	0	0
12.12.2018	32	0	0	0	0

Table 6: Number of calves receiving each score for coughing

Date	No. Calves Total Wisconsin Score				
	0	**1**	**2**	**3**	**4**
14.11.2018	3	8	6	3	1
28.11.2018	7	9	5	0	0
12.12.2018	21	9	2	1	0

Table 7: Number of calves receiving each total Wisconsin score

3.2. Calf Wisconsin Clinical Score

The Wisconsin scores for the older calves over all three sampling dates as shown in Figure 9 varied from 0 to 4. The most common score was 1 with 41% scoring 1 on one or more sampling dates and the least common score was 4 with only 2% scoring 4 on one or more sampling dates.

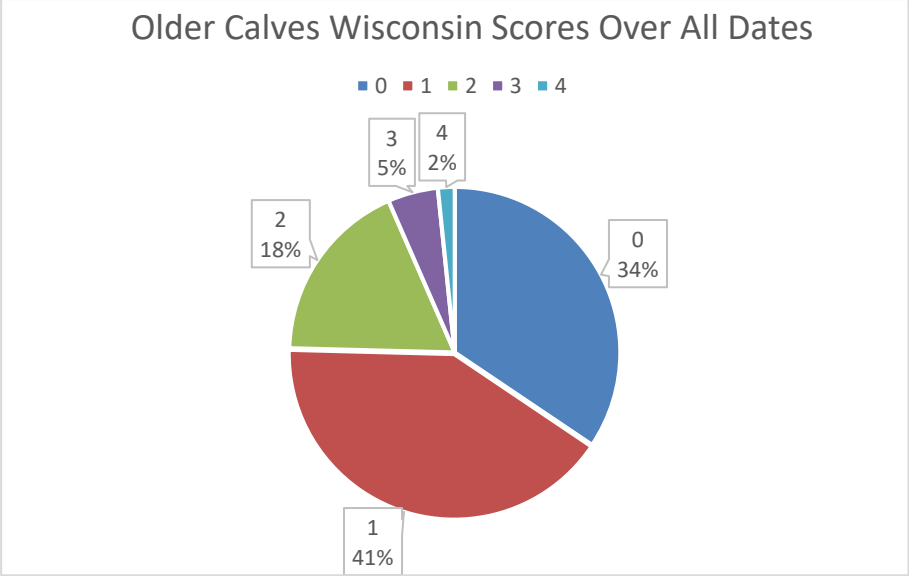


Figure 9 Wisconsin Scores of Older Calves across All Dates

The Wisconsin scores on 14.11.2018 showed the greatest variation with scores of 0, 1, 2, 3 and 4 as shown in Figure 10. In comparison as seen in Figure 12, the Wisconsin scores recorded on 12.12.2018 showed the smallest variation in scores with only scores of 1 or 0. The 28.11.2018 (see Figure 11) had a moderate variation in scores with calves scoring 0, 1 or 2.

On 14.11.2018 (Figure 10) the most common score was 1 with 38% of calves scoring 1 on this day. The least common score on this date was 4 with only 5% of calves scoring 4.

The most common score was also 1 on 28.11.2018 with 40% of calves scoring 1. However, the least common score on the 28.11.2018 was 2 (27%). 0 was the most common score (60%) on 12.12.2018 with the remaining 40% of calves scoring 1.

The proportion of calves scoring 0 increases across the sampling dates from 14% to 33% and eventually 60%. Meanwhile, the proportion of calves scoring 1 increased between the 14.11.2018 and 28.11.2018 but remained the same between the 28.11.2018 and 12.12.2018. The remaining scores (2, 3, and 4) saw a decrease over the sample dates. Scores of 3 and 4 decreased from 14% and 5% respectively to zero between 14.11.2018 and 28.11.2018, remaining at zero on the 12.12.2018. 29% of calves scored 2 in the first sampling week. This then decreased to 27% on the 28.11.2018 before decreasing further to zero on the 12.12.2018.

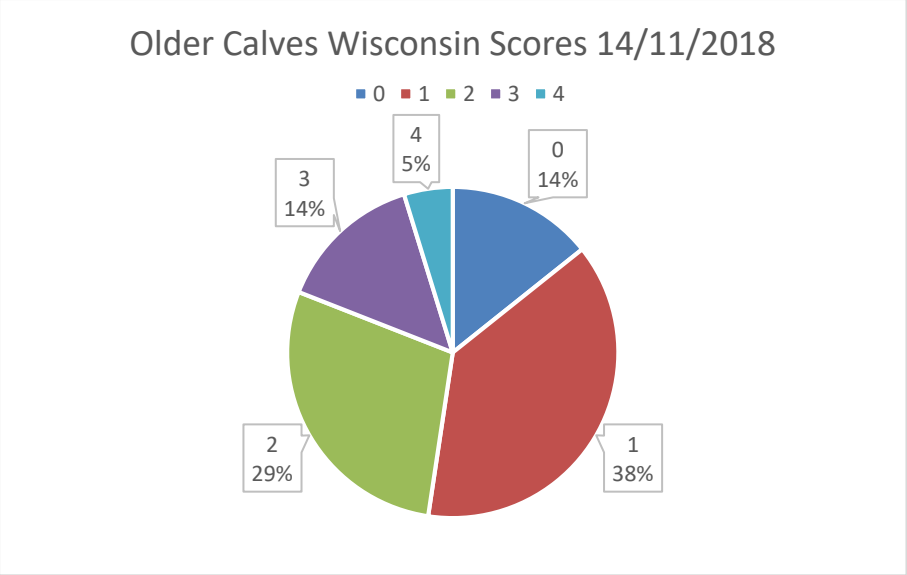


Figure 10: Wisconsin Scores of Older Calves on 14/11/2018

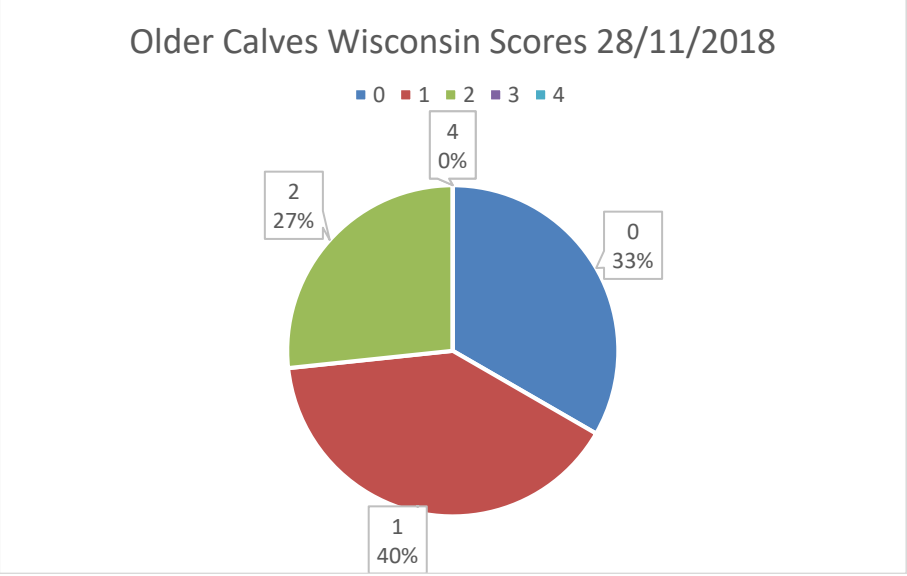


Figure 11: Wisconsin Scores of Older Calves on 28/11/2018

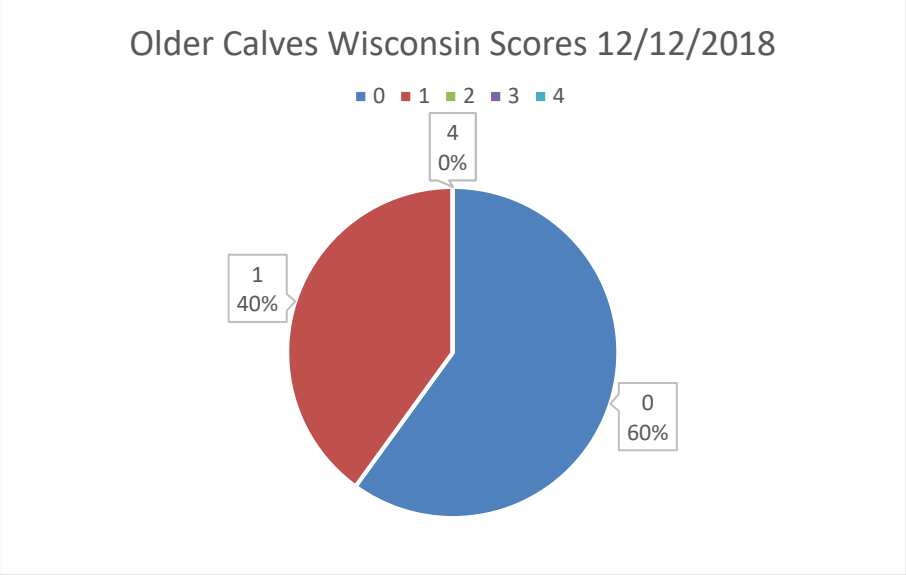


Figure 12: Wisconsin Scores of Older Calves on 12/12/2018

The young calves' Wisconsin scores on the 12.12.2018 (Figure 13) showed more variety than the older calves Wisconsin scores recorded on the same date (Figure 12) with scores of 0, 1, 2 and 3 while the older calves only had scores of 0 or 1. Similarly to the older calves, the younger calves' most common score on the 12.12.2018 was 0 (67%). However, the least common scores in the younger calves were joint between 1 and 3, both in 8% of calves.

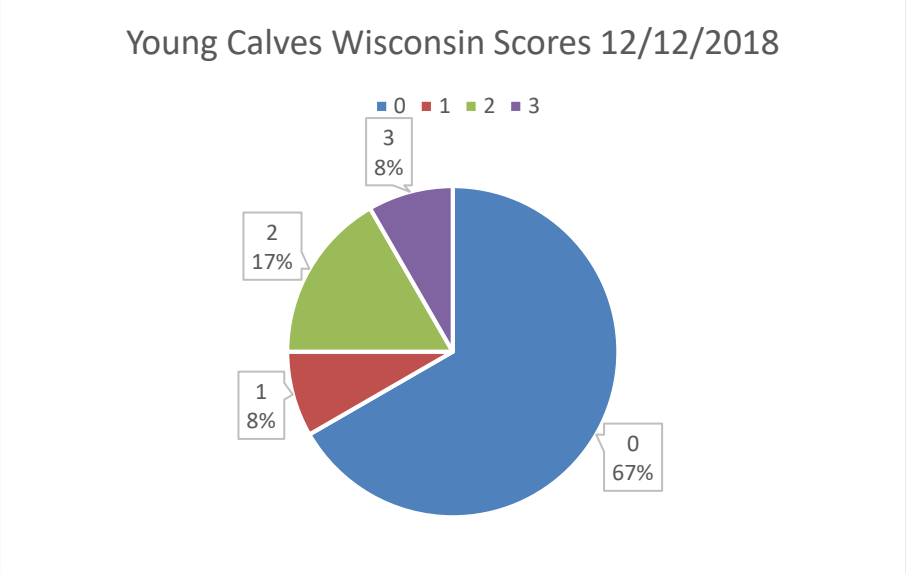


Figure 13: Wisconsin Scores of Younger Calves on 12/12/2018

3.3. Calf Weight

As seen in Figure 14, the weights of the older calves varied on all dates. The pattern of weight gain or loss also varied across the sampling cohort. Some calves gained weight over the sampling period while others lost or maintained their weight between sampling intervals. The rate of gain is also different between the calves with some gaining at a faster rate than others between certain weeks. The lowest weight measured on 14.11.2018 was 206kg while

the heaviest calf on this date was 308kg. On the 12.12.2018 the lightest calf weighed 246kg and the heaviest calf 364kg.

The younger calves measured on 12.12.2018 as seen in Figure 15 varied in weight on this date. The lightest calf weighed in at 150kg while the heaviest calf was 235kg.

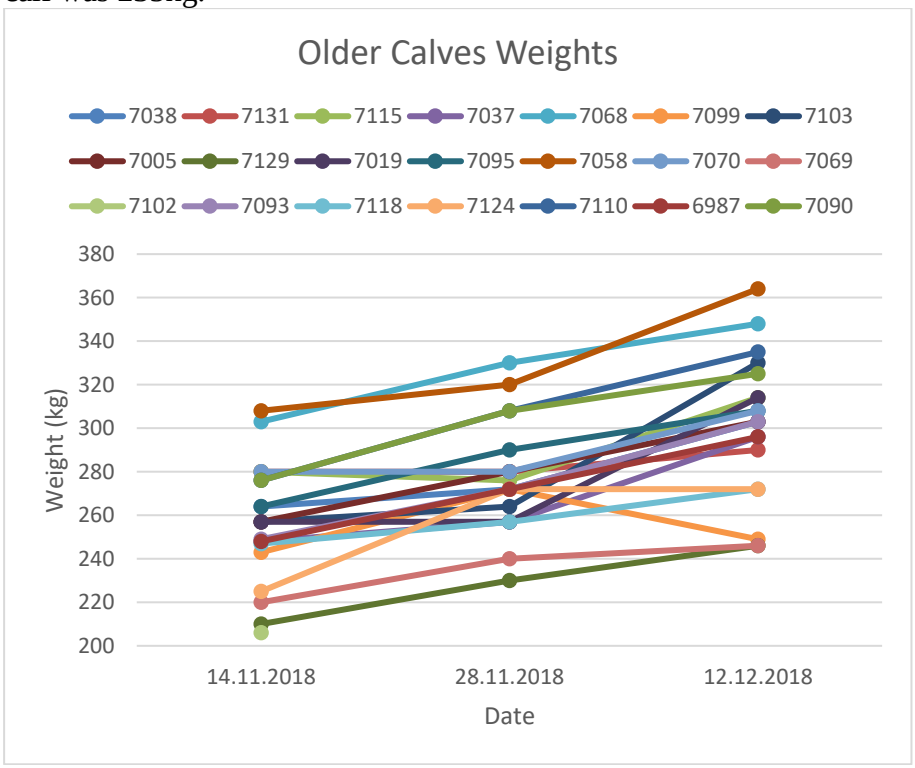


Figure 14: Weight Change in Older Calves over Sampling Dates

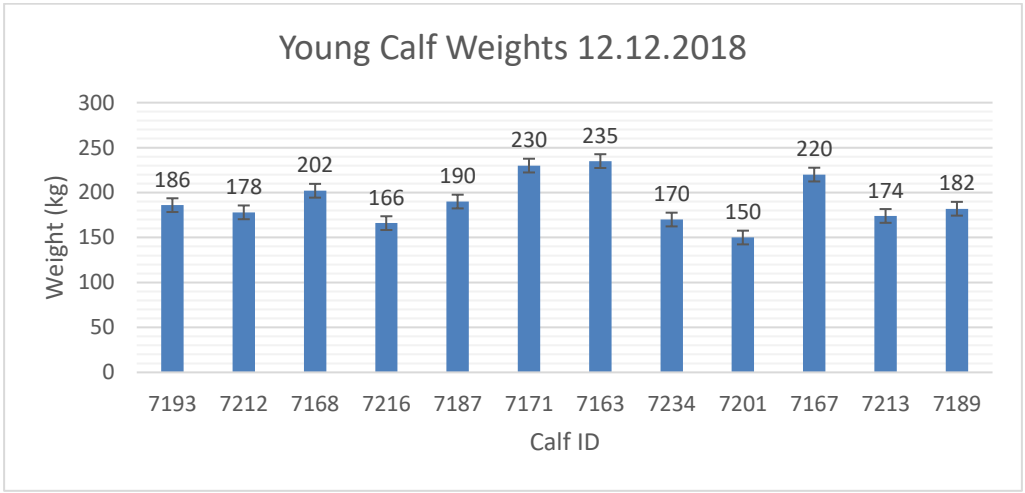


Figure 15: Bar chart showing weight estimates for the young calves on the 12.12.2018

3.4. Wisconsin Score Vs Weight

From Figure 16, Figure 17 and Figure 18 there doesn't appear to be any clear correlation between weight and Wisconsin score for the older calves on any of the sampling dates.

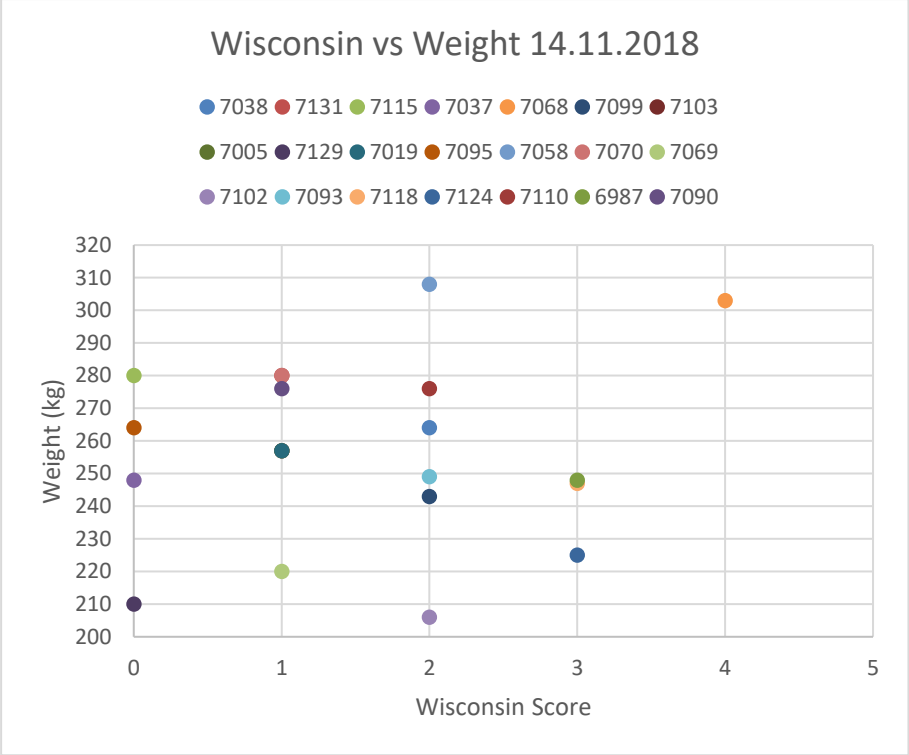


Figure 16: Weight and Wisconsin Score of Older Calves on 14/11/2018

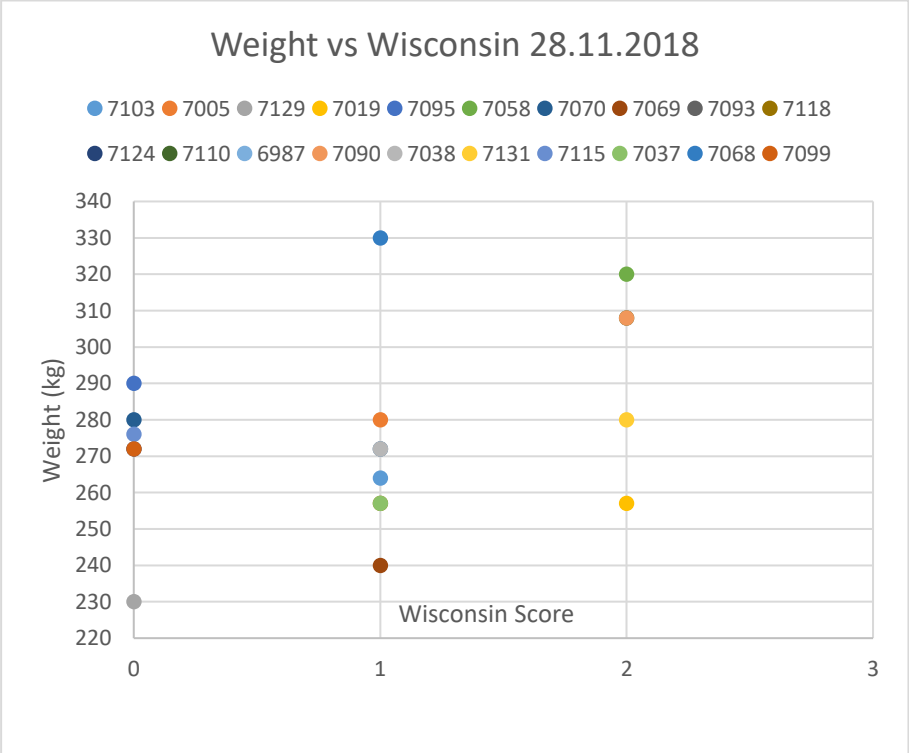


Figure 17: Weight and Wisconsin Score of Older Calves on 28/11/2018

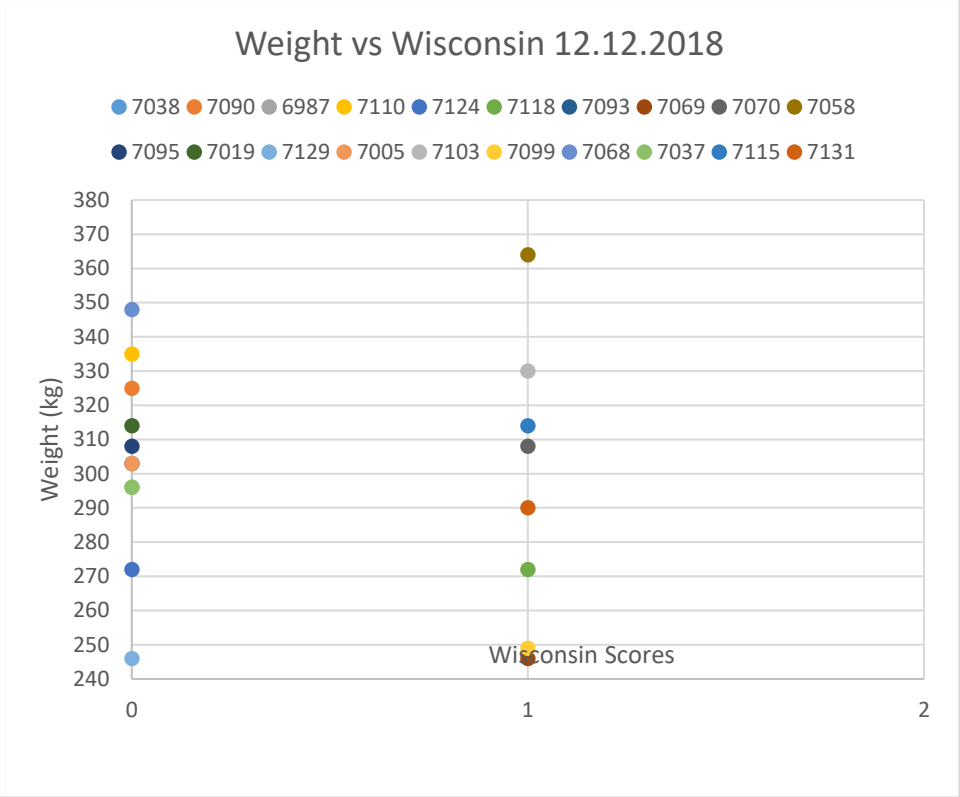


Figure 18: Weight and Wisconsin Score of Older Calves on 12/12/2018

It is noticeable that no calves lost weight linearly between all three of the sampling dates. However, some calves did lose or maintain the same weight between two of the sample dates. Calf 7115 (Figure 19) saw a decrease in weight between 14.11.2018 and 28.11.2018 but then an increase in weight between the 28.11.2018 and 12.12.2018. This calf scored 0 on the first two sampling dates but then increased to 1 on the third sampling date. In Figure 20 we can see that calf 7099 did the opposite of calf 7115 in terms of weight. An increase in weight between the first two sampling dates was followed by a decrease in weight between the last two sampling dates. The only calf that gained weight in the first sampling interval but maintained the same weight in the second sampling interval was calf 7124 as seen in Figure 21. Figure 21 also shows that calf 7124 had a decrease in Wisconsin score in the first sampling interval, going from a score of 3 to 0. During the second sampling interval calf, 7124 remained as a Wisconsin score of 0. Three calves (7019, 7131 and 7070) maintained their weight between 14.11.2018 and 28.11.2018 then had a gain in weight between 28.11.2018 and 12.12.2018 (Figure 22, Figure 23 and Figure 24 respectively). However, while the Wisconsin score for calf 7070 (Figure 24) decreased between the first two weeks then increased between the last two weeks the scores for both calf 7019 and 7131 (Figure 22 and Figure 23) increased between the first two weeks then decreased between the last two weeks. Calf 7070 (Figure 24) started with a Wisconsin score of 1 that dropped to 0 on the second sampling date before increasing back to 1 on the third sampling date. Both 7019 and 7131 had a Wisconsin score that increased from 1 to 2 between the first two sampling dates. Despite both having a decreased Wisconsin score between the 28.11.2018 and 12.12.2018 calf 7019 had a larger decrease going from 2 to 0 whereas calf 7131 only decreased from 2 to 1.

The remaining calves not mentioned above all saw an increase in weight between all sampling dates. However, the rate of increase varied between them. In Figure 25 and Figure 26, we can see that both calf 7095 and calf 7129 as well as consistently gaining weight also had a Wisconsin score of 0 on all three sampling dates. Calves 7069 and 7103 (Figure 27 and Figure 28) also demonstrated constant Wisconsin scores across all three dates but in this instance, they scored 1 across all the dates. No calves had a Wisconsin score that consistently increased across all three sampling dates. Four calves (6987, 7093, 7038 and 7068) had Wisconsin scores that decreased between each sampling interval as seen in Figure 29, Figure 30, Figure 31 and Figure 32. Both calf 7093 and calf 7038 had Wisconsin score that started at 2 on the 14.11.2018, decreased to 1 on the 28.11.2018 and eventually decreased to 0 on the 12.12.2018. As evident in Figure 33, the Wisconsin score for calf 7090 increased from 1 to 2 during the first sampling interval then decreased from 2 to 0 in the second sampling interval. Calf 7118 (Figure 34) saw a decrease in Wisconsin score, from 3 to 1, between the first two sampling dates then the score remained at 1 on the last two sampling dates. Four calves (7005, 7037, 7058, and 7110) had a Wisconsin score that stayed constant between the 14.11.2018 and 28.11.2018 before decreasing between the 28.11.2018 and 12.12.2018. In Figure 35 and Figure 36 we can see that both calf 7005 and 7037 had a Wisconsin score that stayed at 1 on the first two sampling dates before decreasing to 0 in the final sampling date. Calf 7058 (Figure 37) and calf 7110 (Figure 38) both had Wisconsin scores that remained at 2 for the first two sampling dates but while calf 7058 had a score that decreased to 1 on the final sampling date calf 7110 scored 0 on the last sampling date.
The significance of the correlation between weight and Wisconsin score for the calves was not tested.

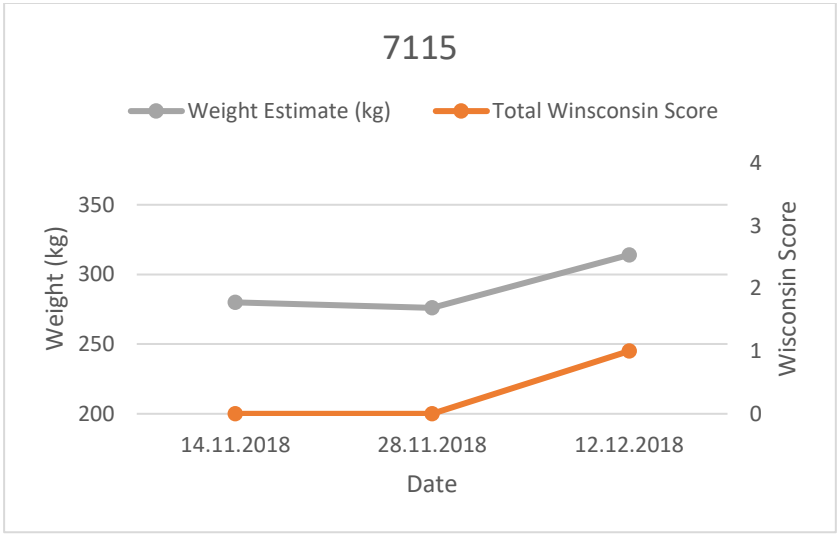


Figure 19: Weight and Wisconsin Score of 7115 over Sampling Dates

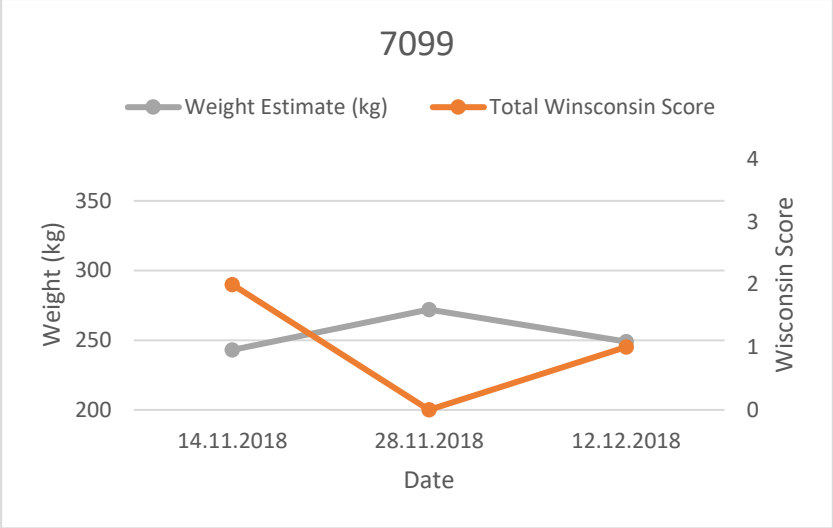


Figure 20: Weight and Wisconsin Score of 7099 over Sampling Dates

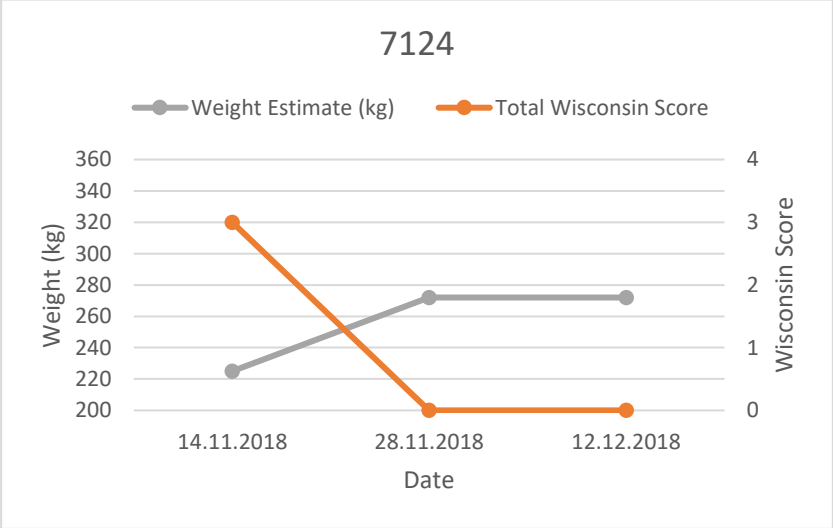


Figure 21: Weight and Wisconsin Score of 7124 over Sampling Dates

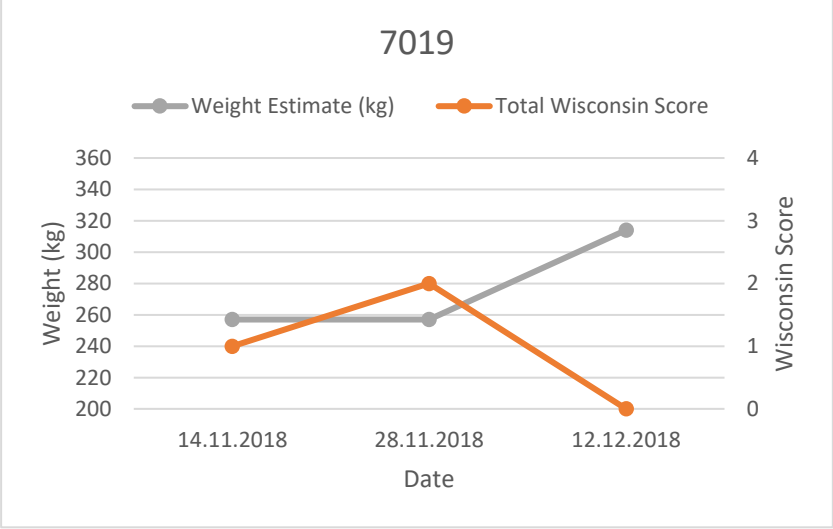


Figure 22: Weight and Wisconsin Score of 7019 over Sampling Dates

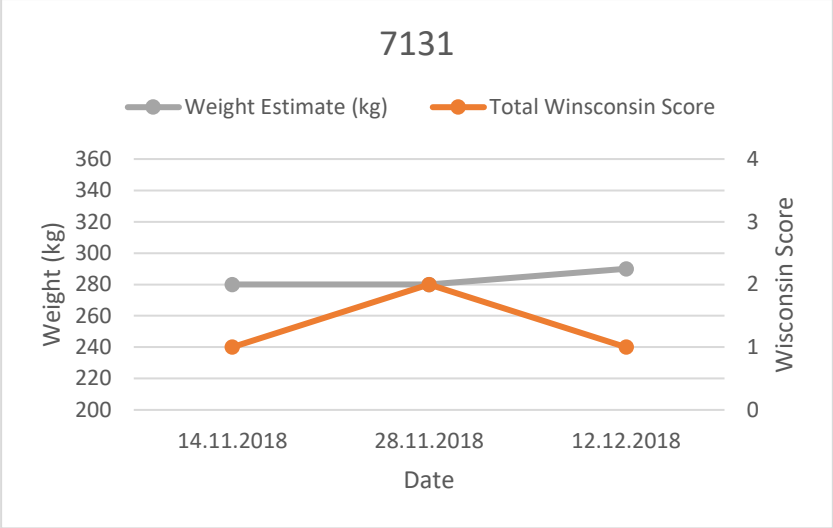


Figure 23: Weight and Wisconsin Score of 7131 over Sampling Dates

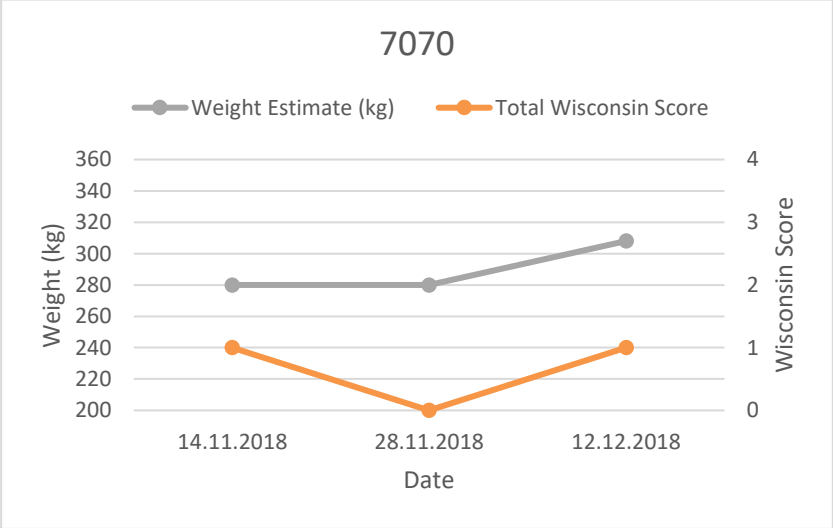


Figure 24: Weight and Wisconsin Score of 7070 over Sampling Dates

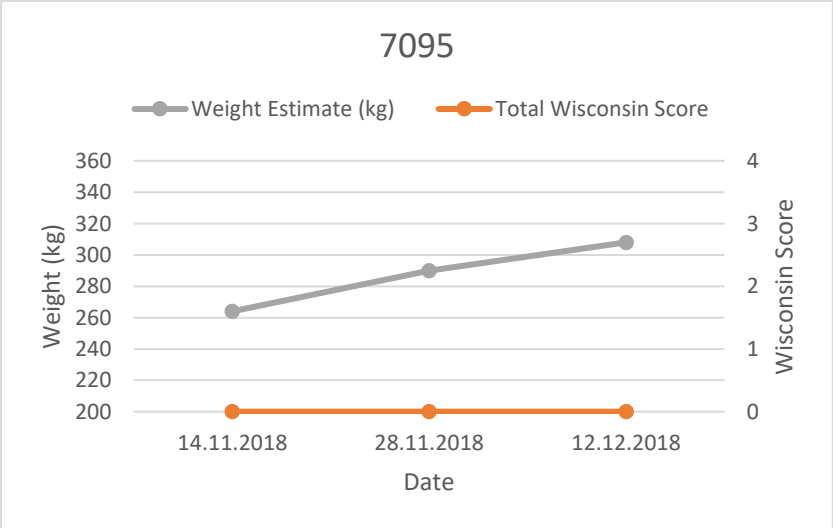


Figure 25: Weight and Wisconsin Score of 7095 over Sampling Dates

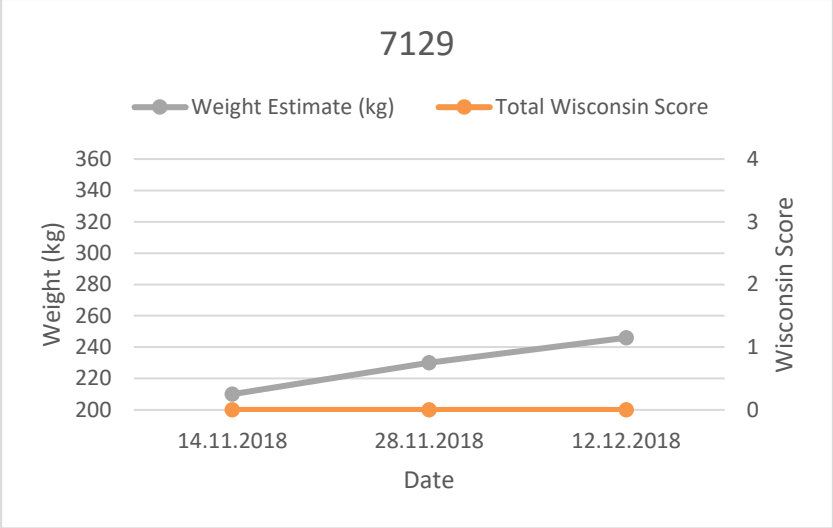


Figure 26: Weight and Wisconsin Score of 7129 over Sampling Dates

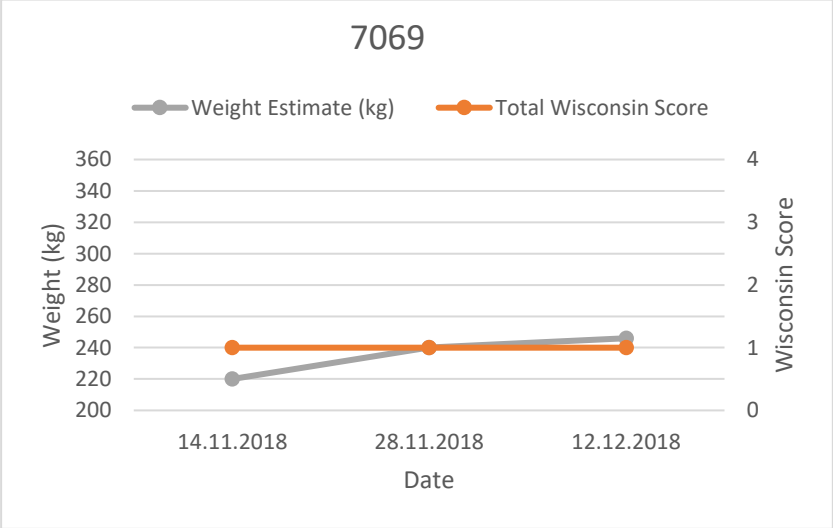


Figure 27: Weight and Wisconsin Score of 7069 over Sampling Dates

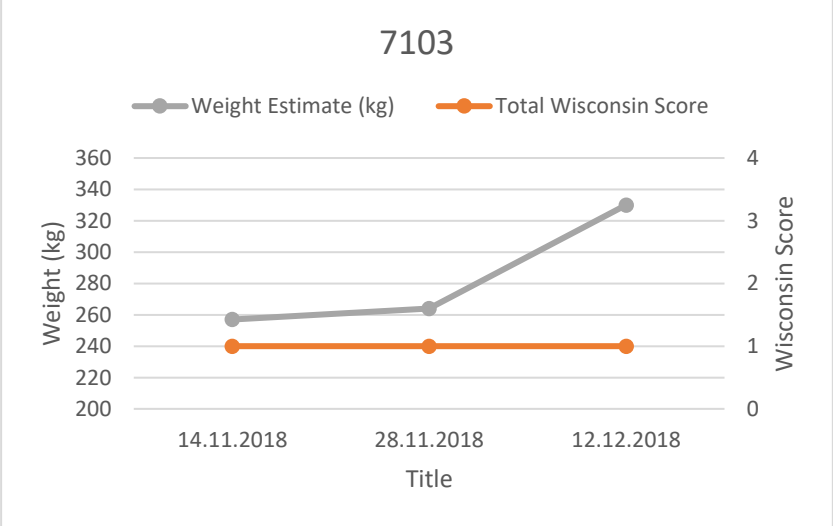


Figure 28: Weight and Wisconsin Score of 7103 over Sampling Dates

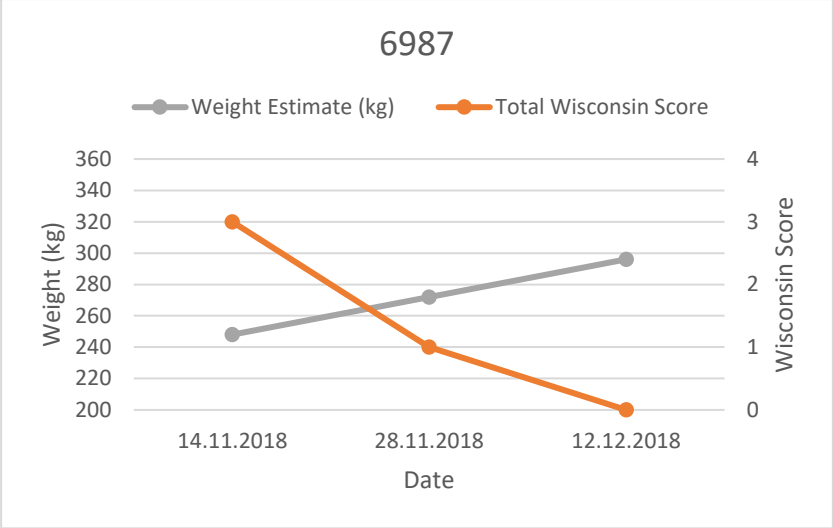


Figure 29: Weight and Wisconsin Score of 6987 over Sampling Dates

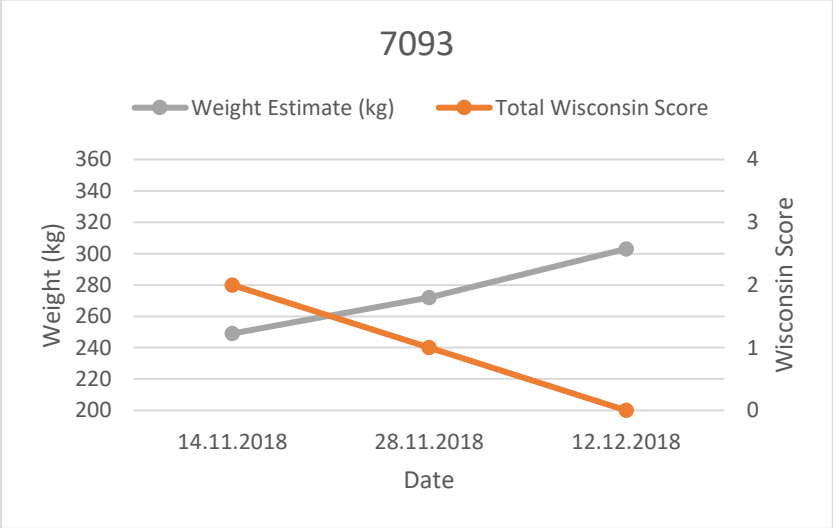


Figure 30: Weight and Wisconsin Score of 7093 over Sampling Dates

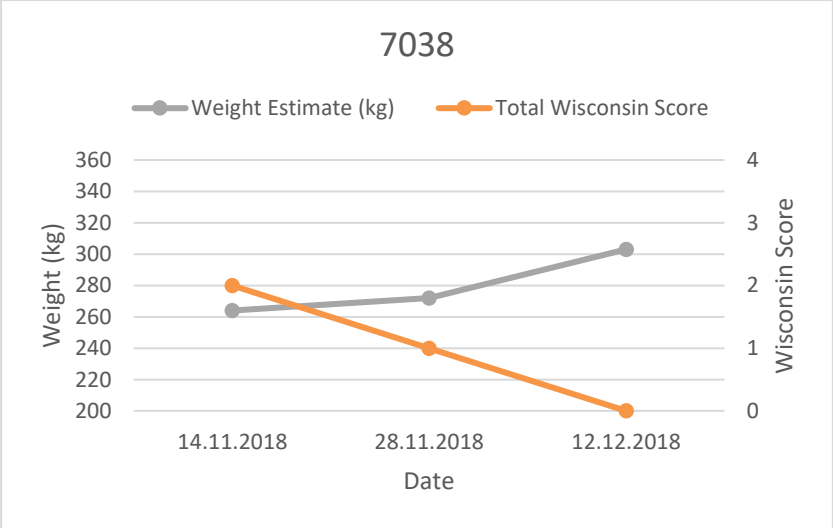


Figure 31: Weight and Wisconsin Score of 7038 over Sampling Dates

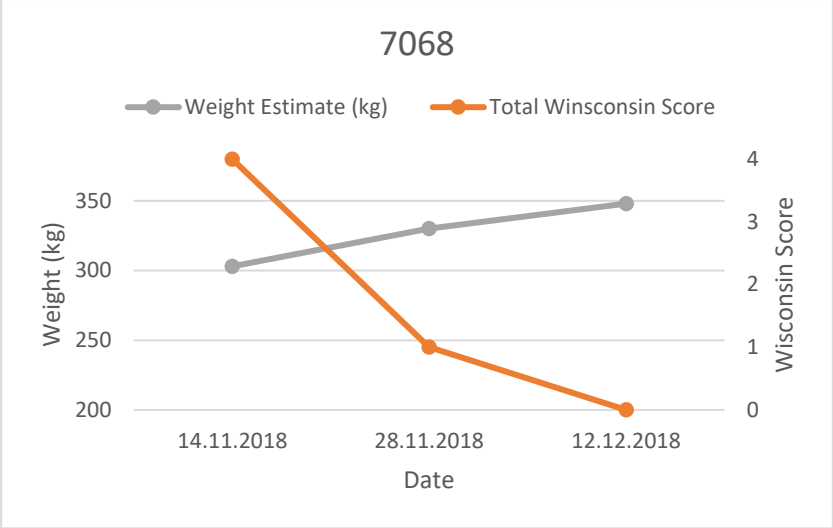


Figure 32: Weight and Wisconsin Score of 7068 over Sampling Dates

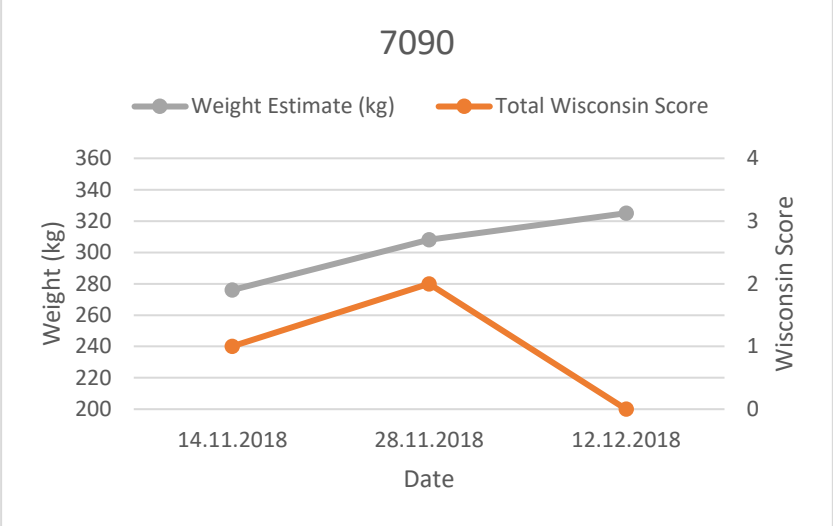


Figure 33: Weight and Wisconsin Score of 7090 over Sampling Dates

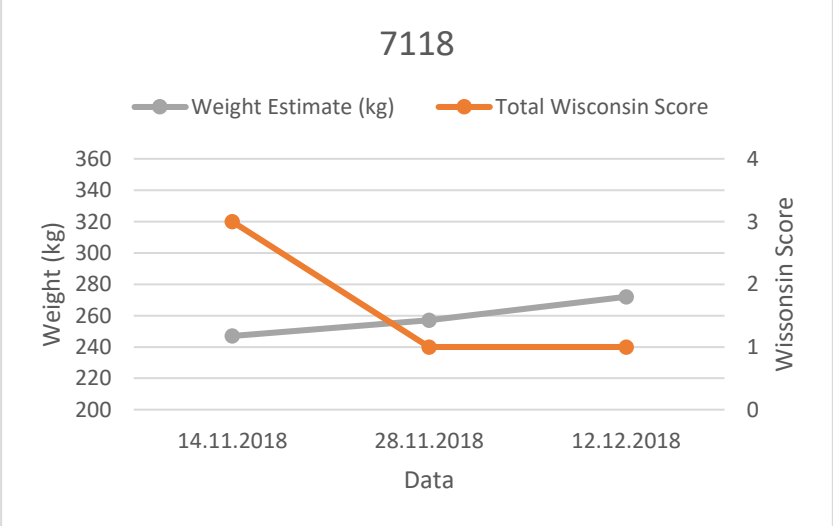


Figure 34: Weight and Wisconsin Score of 7118 over Sampling Dates

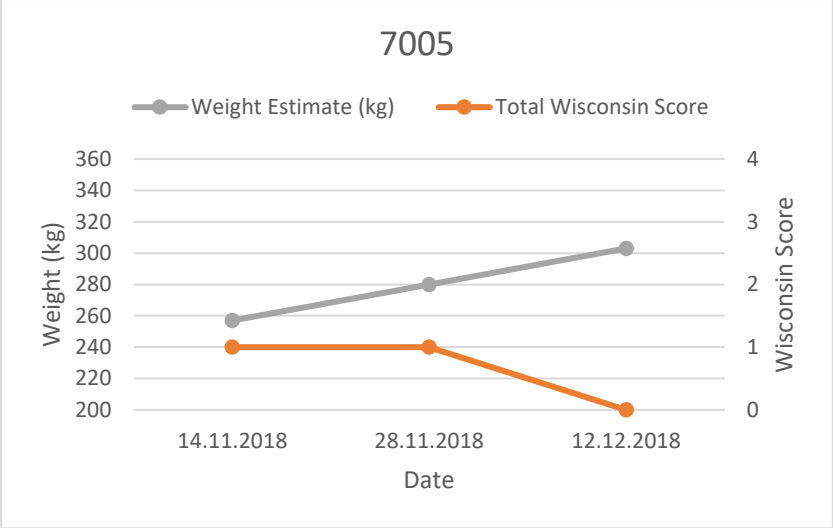


Figure 35: Weight and Wisconsin Score of 7005 over Sampling Dates

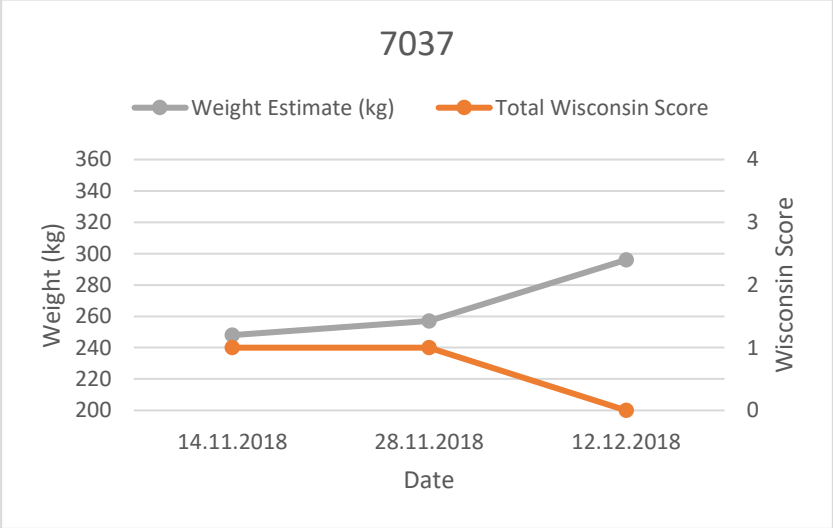


Figure 36: Weight and Wisconsin Score of 7037 over Sampling Dates

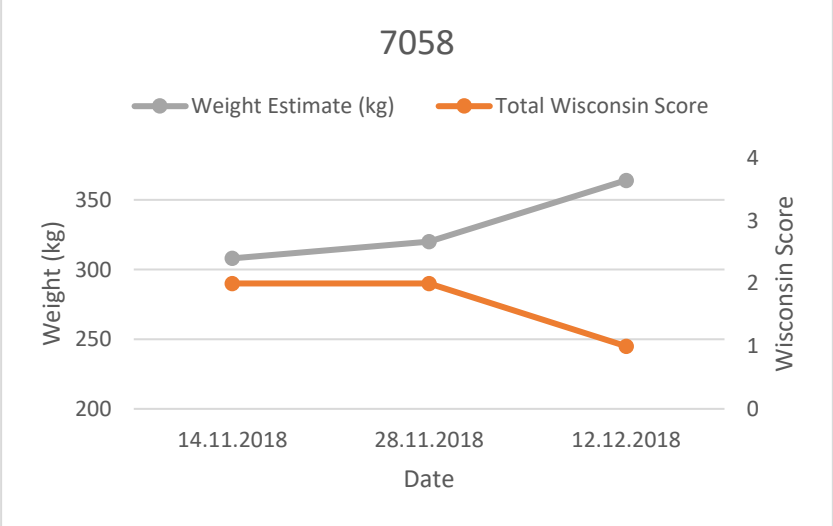


Figure 37: Weight and Wisconsin Score of 7058 over Sampling Dates

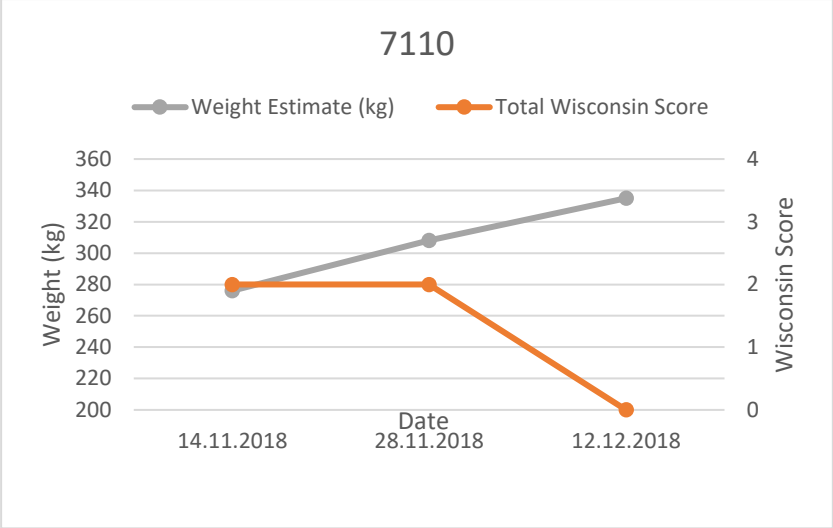


Figure 38: Weight and Wisconsin Score of 7110 over Sampling Dates

3.5. Data Analysis of Core Temperature

3.5.1. Plots of Temperature Distribution for Individual Calves on First Sampling Date and One Week Previous to Sampling

The histograms (Figure 39) show the temperature distribution measured by the ruminal boluses for each of the calves from the 7th November 2018 to the 14th of November 2018. There is no temperature data after the 14th of November as the boluses stopped transmitting data. Figure 39 shows all calves across all of the eight dates. The red crosses indicate values considered to be outliers. The black lines indicate the range of the core temperature values measured for each calf while the blue box indicates the interquartile range from the 1st quartile to the 3rd quartile. The red line within the blue box is the median core temperature value. Calf 7090 and calf 7095 passed the boluses, on the 9th November 2018 and 11th November 2018 respectively, so do not have any temperature data past these dates. However, neither calf had any measurements over the 40.5℃ temperature threshold. Five calves had a temperature distribution that remained below the threshold even at the highest end of the range (7003, 7038, 7068, 7124 and 7129). Many of the calves had a temperature range that extended over the threshold line including calf 7070 on the 7th November 2018, calves 7069, 7103 and 7110 on the 8th November 2018, calf 7115 on the 11th November 2018, calf 7093 on the 12th November 2018 and calf 7099 on both the 10th and 13th November 2018.

Only one calf demonstrated an inter-quartile range that extended over the threshold and that was calf 7102 on the 8th of November 2018, however, the median remained below the threshold. This calf also demonstrated a temperature range that extended over the threshold on the 7th, 8th, 9th, 10th, 11th and 12th November 2018.

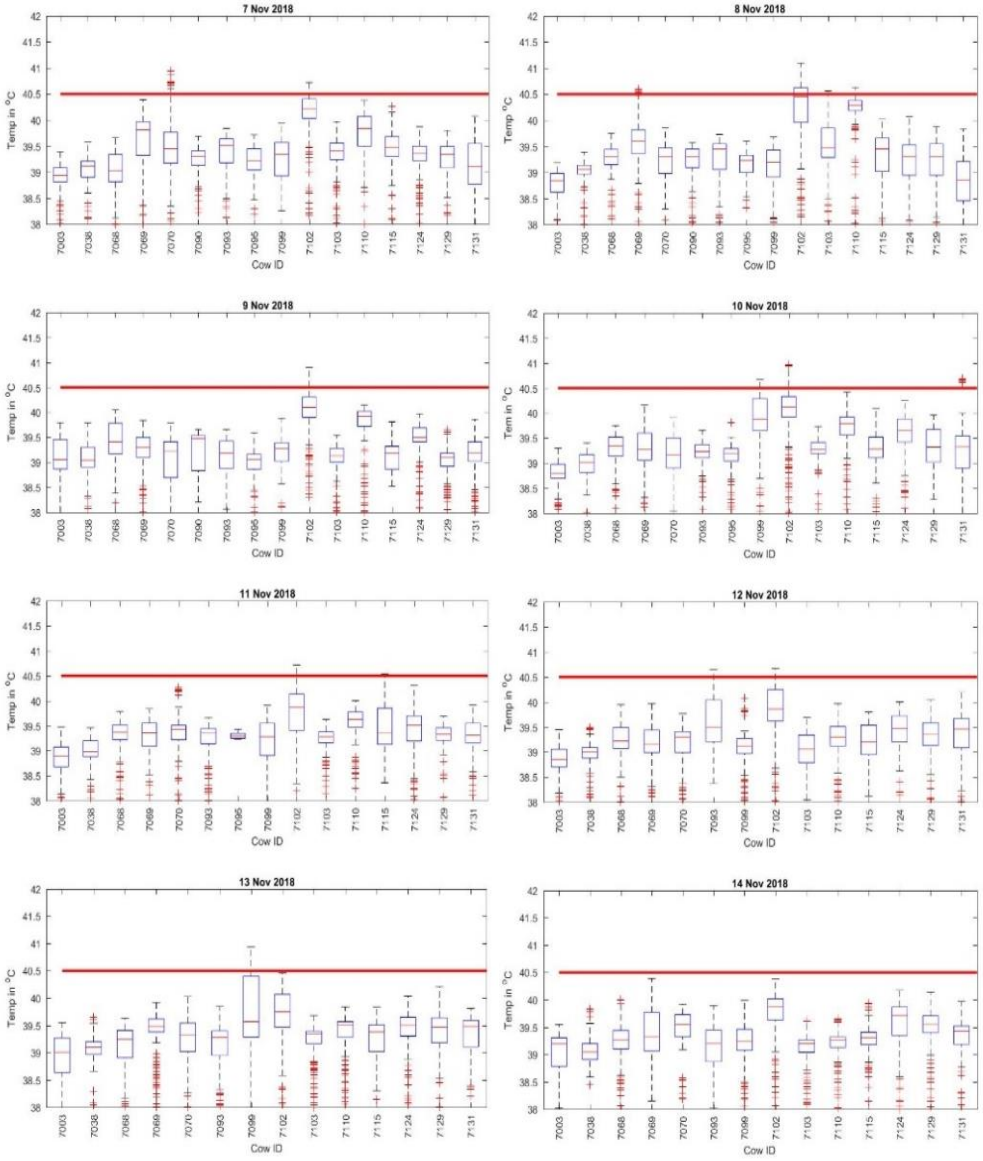


Figure 39: Temperature Distribution for all Calves from 07.11.2018 to 14.11.2018 including Threshold Line at 40.5 ℃

3.5.2. ANOVA Comparison of Temperature between Different Calves

An ANOVA test was performed on the data in order to compare the mean differences in core temperature measurements between individual calves on the 7th of November 2018 and between individual calves on the 14th of November 2018. The results of the ANOVA comparison are shown in Table 8 and Table 9 for the 7th and 14th November 2018 respectively. Calves 7090 and 7095 were not included in the ANOVA for data from the 14th November 2018 as neither had any temperature data from this date. The null hypothesis was that all calves will have a similar core temperature distribution. For both dates P<0.01** suggesting rejection of the null hypothesis. This means that the calves have a statistically significant difference in temperature distribution.

Source	SS	df	MS	F	Prob>F
Groups	193.081	15	12.8721	59.47	1.0666e-149
Error	463.931	2143	0.2164		
Total	656.912	2158			

Table 8: Results of ANOVA comparing temperatures between the different calves on the 7th Nov 2018. The source is the source of variability. SS is the sum of squares due to each source. Df is the degree of freedom associated with each source. MS is the mean squares for each source. F is the ratio of the mean squares. Probs>F is the p-value.

Source	SS	df	MS	F	Prob>F
Groups	101.501	13	7.80777	45.74	5.45354e-104
Error	362.592	2124	0.17071		
Total	464.093	2137			

Table 9: Results of ANOVA comparing temperatures between the different calves on the 14th Nov 2018. The source is the source of variability. SS is the sum of squares due to each source. Df is the degree of freedom associated with each source. MS is the mean squares for each source. F is the ratio of the mean squares. Probs>F is the p-value.

3.5.3. T-Test for Individual Calves Comparing Core Temperatures between Day 7 and Day 14

The t-test was carried out to determine whether there was a significant difference between calf temperatures on the 7th November 2018 and calf temperature on the 14th of November 2018 for each individual calf. The results of the t-test are shown in Table 10. Calves 7090 and 7095 were not included in the t-test as neither had any temperature data on the 14th of November 2018. For all calves, the P values are shown, with P<0.01** indicating rejection of the null hypothesis, that the core temperatures for individual calves are similar on the 7th and 14th November 2018. Calf 7115 had P<0.05* also suggesting rejection of the null hypothesis. However, calves 7038, 7070 and 7099 had P>0.05 supporting the null hypothesis in these individuals.

Cow ID	P-value
7003	4.29e-05
7038	0.536193157
7068	0.000407827
7069	0.000303381
7070	0.87937891
7093	7.93e-08
7099	2.57e-01
7102	2.52e-10
7103	2.31e-06
7110	1.78e-19
7115	2.59e-02
7124	3.38e-05
7129	4.05e-06
7131	9.50e-05

Table 10: T-test comparing core temperature on the 7th Nov 2018 to the 14th Nov 2018 for individual calves

3.6. DNA/RNA Extraction and Quantification

When extracting microbial DNA from a host, a commonly faced problem is that the microbial fraction is not well recorded because of the high content of the host DNA. To test the selected samples for the presence of microbial DNA we analysed them using the PCR. Figure 40, Figure 41 and Figure 42 show the gel electrophoresis results from the PCR for the detection of bacterial 16S rRNA. The columns are labelled as detailed in Table 11. The column labelled "N" in Figure 40 and Figure 41 is blank in both figures demonstrating that the negative control has worked successfully. The columns labelled 1 to 20 in all three figures are the DNA samples. All, except column 1 and 3, demonstrate a band at 763bp. Column 2 and 14 both had very faint bands in comparison to the rest of the samples.

Calf ID	Age	Sample ID	Sample Collection Date	PCR Gel Column Label
NA	NA	NA (nuclease-free water)	NA	N
NA	NA	C	NA	1
7212	14 weeks	2E	12.12.2018	2
7213	14 weeks	11E	12.12.2018	3
7189	14 weeks	12E	12.12.2018	4
7115	7-8 months	3B	14.11.2018	5
7099	7-8 months	6B	14.11.2018	6
7005	7-8 months	8B	14.11.2018	7
7129	7-8 months	9B	14.11.2018	8
7095	7-8 months	11B	14.11.2018	9
7069	7-8 months	14B	14.11.2018	10
7118	7-8 months	17B	14.11.2018	11
7124	7-8 months	18B	14.11.2018	12
7115	7-8 months	45B	12.12.2018	13
7099	7-8 months	48B	12.12.2018	14
7005	7-8 months	50B	12.12.2018	15
7129	7-8 months	51B	12.12.2018	16
7095	7-8 months	53B	12.12.2018	17
7069	7-8 months	56B	12.12.2018	18
7118	7-8 months	59B	12.12.2018	19
7124	7-8 months	60B	12.12.2018	20

Table 11: Table showing the samples tested using the 16S PCR

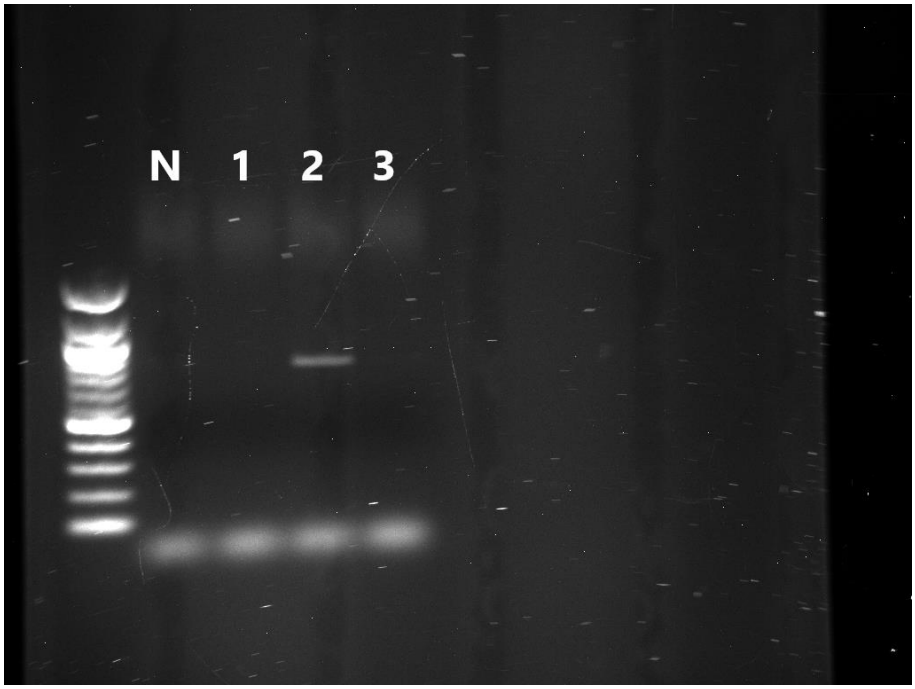


Figure 40: PCR Gel showing PCR for the detection of bacterial 16S rRNA. N - Negative control. 1 to 3 - Samples

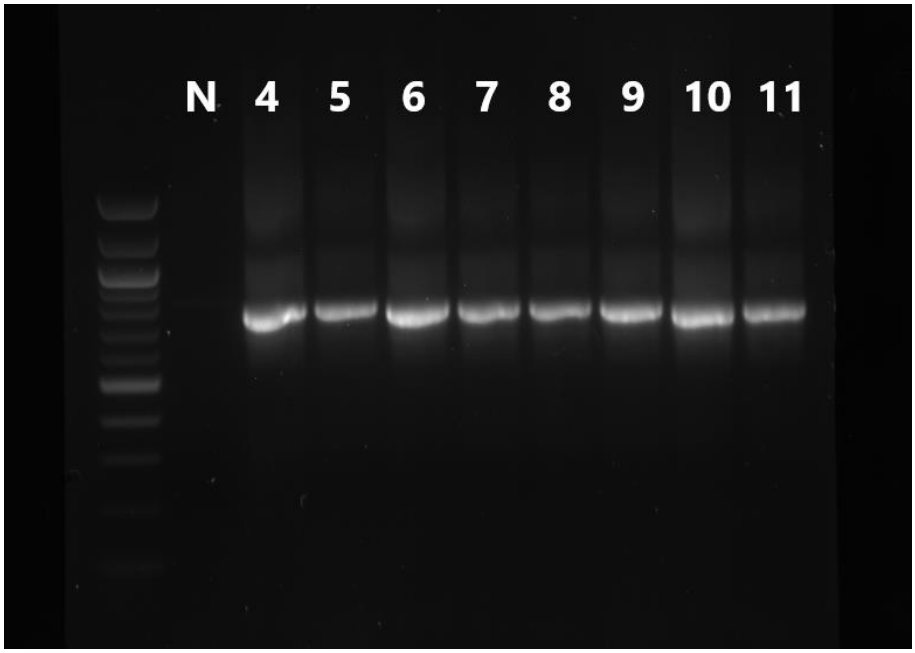


Figure 41: PCR Gel showing PCR for the detection of bacterial 16S rRNA. N - Negative control. 4 to 11 - Samples

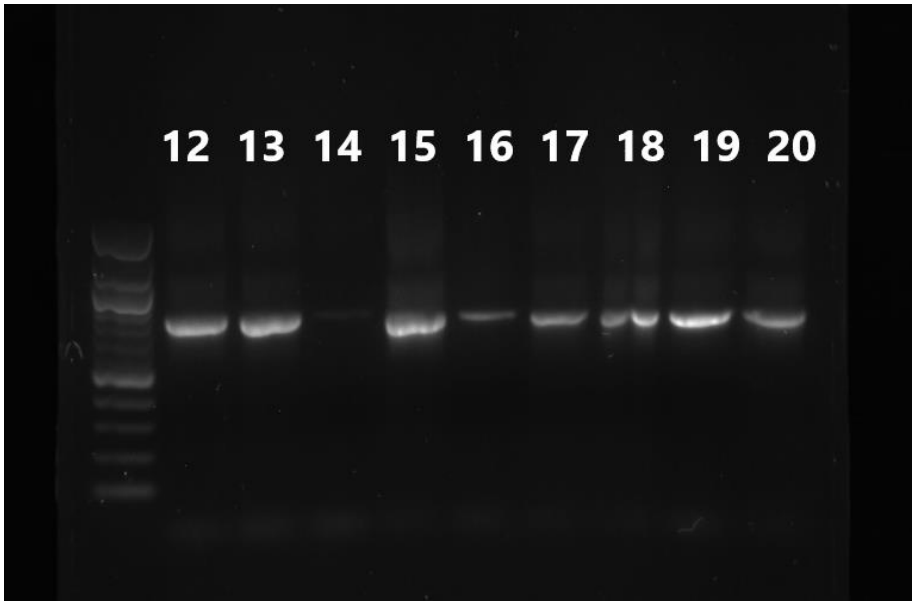


Figure 42: PCR Gel showing PCR for the detection of bacterial 16S rRNA. 12 to 20 - Samples

Ten samples of extracted DNA (Highlighted in orange in Table 12) were below the desired 20ng/µl concentration of DNA but still above the minimum required concentration. Only two DNA samples (62B and 63B) and one RNA sample (3B) of the extracted DNA and RNA failed to meet the minimum concentration of 2.5ng/µl (Highlighted in red in Table 12). One sample (5B) had too little DNA or RNA extracted to quantify on the nanodrop. Sample 10B (calf 7019) had the highest measured concentration of DNA in the extracted DNA sample at 231.5ng/µl. 158.3ng/µl was the highest measured concentration of extracted RNA in the RNA extracted from sample 1B. The average (mean) of the DNA concentration measurements was 53.36ng/µl with a median value of 36.97ng/µl. The RNA concentration measurements had an average (mean) of 78.09ng/µl with a median value of 72.67ng/µl. While the range of the DNA measurements was 230.95ng/µl the RNA measurements had a lower range of 142.06ng/µl. The results from the nanodrop on all of the extracted DNA and RNA samples are shown in Table 12.

Calf ID	Age	Sample Collection Date	Rectal Swab ID	Rectal Swab DNA Nanodrop ng/µl	Rectal Swab RNA Nanodrop ng/µl
7038	7-8 months	14.11.2018	1B	3.889	158.30
7038	7-8 months	28.11.2018	22B	51.97	38.21
7038	7-8 months	12.12.2018	43B	19.22	66.94
7131	7-8 months	14.11.2018	2B	NA	NA
7131	7-8 months	28.11.2018	23B	48.97	98.85
7131	7-8 months	12.12.2018	44B	2.93	113.00
7115	7-8 months	14.11.2018	3B	134.40	16.24
7115	7-8 months	28.11.2018	24B	63.51	61.03
7115	7-8 months	12.12.2018	45B	34.34	116.50
7037	7-8 months	14.11.2018	4B	46.32	81.14
7037	7-8 months	28.11.2018	25B	41.60	55.41

7037	7-8 months	12.12.2018	46B	35.13	39.86
7068	7-8 months	14.11.2018	5B	Sample too small	Sample too small
7068	7-8 months	28.11.2018	26B	30.23	63.78
7068	7-8 months	12.12.2018	47B	81.30	53.83
7099	7-8 months	14.11.2018	6B	32.55	51.48
7099	7-8 months	28.11.2018	27B	51.91	76.04
7099	7-8 months	12.12.2018	48B	34.83	101.90
7103	7-8 months	14.11.2018	7B	92.90	112.60
7103	7-8 months	28.11.2018	28B	36.97	124.30
7103	7-8 months	12.12.2018	49B	39.50	61.87
7005	7-8 months	14.11.2018	8B	153.10	104.70
7005	7-8 months	28.11.2018	29B	41.30	52.59
7005	7-8 months	12.12.2018	50B	24.00	65.20
7129	7-8 months	14.11.2018	9B	212.10	86.58
7129	7-8 months	28.11.2018	30B	39.27	88.03
7129	7-8 months	12.12.2018	51B	27.01	144.60
7019	7-8 months	14.11.2018	10B	231.50	99.57
7019	7-8 months	28.11.2018	31B	116.70	76.61
7019	7-8 months	12.12.2018	52B	18.35	61.48
7095	7-8 months	14.11.2018	11B	30.26	60.07
7095	7-8 months	28.11.2018	32B	27.29	113.50
7095	7-8 months	12.12.2018	53B	22.93	97.56
7058	7-8 months	14.11.2018	12B	85.11	58.85
7058	7-8 months	28.11.2018	33B	84.69	118.70
7058	7-8 months	12.12.2018	54B	63.48	48.59
7070	7-8 months	14.11.2018	13B	80.65	48.57
7070	7-8 months	28.11.2018	34B	37.76	60.61
7070	7-8 months	12.12.2018	55B	28.16	100.40
7069	7-8 months	14.11.2018	14B	52.35	61.93
7069	7-8 months	28.11.2018	35B	35.80	90.51
7069	7-8 months	12.12.2018	56B	61.41	94.14
7102	7-8 months	14.11.2018	15B	23.47	81.23
7093	7-8 months	14.11.2018	16B	17.32	81.10
7093	7-8 months	28.11.2018	37B	25.11	101.20
7093	7-8 months	12.12.2018	58B	29.89	121.70
7118	7-8 months	14.11.2018	17B	30.50	49.29
7118	7-8 months	28.11.2018	38B	86.43	67.73
7118	7-8 months	12.12.2018	59B	40.87	56.52
7124	7-8 months	14.11.2018	18B	12.25	51.48
7124	7-8 months	28.11.2018	39B	27.19	74.52
7124	7-8 months	12.12.2018	60B	19.72	47.10
7110	7-8 months	14.11.2018	19B	46.78	112.60
7110	7-8 months	28.11.2018	40B	113.90	49.48

7110	7-8 months	12.12.2018	61B	2.732	80.84
6987	7-8 months	14.11.2018	20B	11.13	102.70
6987	7-8 months	28.11.2018	41B	23.33	92.39
6987	7-8 months	12.12.2018	62B	1.17	78.24
7090	7-8 months	14.11.2018	21B	28.48	85.52
7090	7-8 months	28.11.2018	42B	34.68	65.98
7090	7-8 months	12.12.2018	63B	0.55	52.24
7193	14 weeks	12.12.2018	1E	89.61	71.69
7212	14 weeks	12.12.2018	2E	135.60	52.78
7168	14 weeks	12.12.2018	3E	88.68	56.38
7216	14 weeks	12.12.2018	4E	14.64	52.61
7187	14 weeks	12.12.2018	5E	22.69	71.55
7171	14 weeks	12.12.2018	6E	110.70	92.61
7163	14 weeks	12.12.2018	7E	109.40	130.20
7234	14 weeks	12.12.2018	8E	44.18	68.91
7201	14 weeks	12.12.2018	9E	24.99	72.67
7167	14 weeks	12.12.2018	10E	77.02	72.52
7213	14 weeks	12.12.2018	11E	109.70	77.77
7189	14 weeks	12.12.2018	12E	31.48	48.81

Table 12: Nanodrop Results for Extracted DNA and RNA

3.7. Sample Selection – Selection of the most interesting individual by combining the different health indicators

3.7.1. Older Calves

Table 13 shows the weight gain of the older calves between week 0 and week 2 as well as between week 2 and week 4. The ratio between these two values is also shown along with the total weight gain over the course of the study. The data is sort from lowest to highest by total weight gain. The lowest total weight gain was 6kg achieved by calf 7099 while the highest total weight gain was 7103 with a gain of 73kg. Calf 7115 had the lowest weight gain (-4kg) between week 0 and week 2 while the lowest weight gain between weeks 2 and 4 was 7099 (-23kg). The highest weight gain between weeks 0 and 2 was 7124 with a gain of 47kg across the two weeks. Calf 7103 as well as having the highest total gain also had the highest weight gain between weeks 2 and 4, gaining 66kg in two weeks. Both calf 7005 and 6987 scored a ratio of 1.00. The furthest ratio from one was calf 7115 with a ratio of -9.50 followed by calf 7103 with a ratio of 9.43.

As seen in Table 14 the mean total weight gain was 41.5kg with a median value of 45.5kg. The 1st quartile for total weight gain was 29.5kg while the 3rd quartile was 52.75kg. Ten calves had a total weight gain that fell between the two quartiles (7115, 7129, 7038, 7095, 7068, 7005, 7124, 6987, 7037 and 7090).

The mean for weight gain between week 0 and week 2 was 17.25kg with the 1st quartile of 7.25kg, the median of 20kg and the 3rd quartile of 26.75kg. Twelve calves fall in the range of the 1st and 3rd quartile for weight gain between week 0 and week 2.

For weight gain between week 2 and week 4, the mean was 24.25kg with the 1st quartile of 15.25, the median of 23.5kg and the 3rd quartile of 36.25kg. Ten calves were within the interquartile ranges for weight gain between week 2 and week 4 (7070, 7129, 7038, 7095, 7068, 7005, 6987, 7090, 7093 and 7110). Three calves (7129, 7095 and 7005) were selected for sampling as they have a ratio of 1.00 or close to 1.00 as well as a total weight gain that is within the interquartile range. These calves also had a Wisconsin score of either 0 or 1 on all three sampling dates, with a score of 0 on at least one date. These calves (7129, 7095 and 7005) formed the "healthy" samples (Highlighted in green in Table 13). Both calf 7118 and calf 7069 were selected for sampling as they had poor overall weight gain as well as poor weight gain in one or both of the sampling intervals. Calf 7099 and calf 7124 were selected for sampling as they had a good or higher than average weight gain between weeks 0 and 2 but showed no gain or weight loss between week 2 and week 4 meaning they had ratio values very close to zero. The final calf selected for sampling was 7115. This calf lost weight between weeks 0 and week 2 but gained a reasonable amount of weight between weeks 2 and 4 resulting in an average total weight gain but a ratio value very far from 1.00. These five calves with abnormal or poor weight gain (7118, 7069, 7099, 7124 and 7115) formed our "unhealthy" samples (highlighted in orange in Table 13) and none of them had a Wisconsin score of 0 for all three sampling dates.

Calf ID	Week 0 - 2 Weight Gain (kg)	Week 2 - 4 Weight Gain (kg)	Ratio	Total Weight Gain (kg)
7099	29	-23	-0.79	6
7131	0	10	NA	10
7118	10	15	1.50	25
7069	20	6	0.30	26
7070	0	28	NA	28
7115	-4	38	-9.50	34
7129	20	16	0.80	36
7038	8	31	3.88	39
7095	26	18	0.69	44
7068	27	18	0.67	45
7005	23	23	1.00	46
7124	47	0	0.00	47
6987	24	24	1.00	48
7037	9	39	4.33	48
7090	32	17	0.53	49
7093	23	31	1.35	54
7058	12	44	3.67	56
7019	0	57	NA	57
7110	32	27	0.84	59
7103	7	66	9.43	73

Table 13: Sample Selection of Older Calves by Sorting Weight Gain from Lowest to Highest and Calculating Ratio between Weight Gains over Individual Weeks

	Week 0 - 2 Weight Gain (kg)	Week 2 - 4 Weight Gain (kg)	Total Weight Gain (kg)
MEAN	17.25	24.25	41.5
MEDIAN	20	23.5	45.5
1ST Quart	7.25	15.25	29.5
3rd Quart	26.75	36.25	52.75

Table 14: Mean, median, 1st quartile and 3rd quartile values for weight gain over each of the weeks and in total

3.7.2. Younger Calves

Table 15 shows the weights of the younger calves in order from lowest to highest as well as the Wisconsin score for each calf. For the younger calves, the median weight was 184kg with the 1st quartile of 171kg and the 3rd quartile of 216kg as seen in Table 16. Six of the younger calves had a weight that fell in between the 1st and 3rd quartiles. In total eight of the younger calves scored 0 in the Wisconsin scoring however only four of these calves also had a weight that fell between the 1st and 3rd quartile too. Of these four calves, the three closest to the median (7213, 7212 and 7189) were selected to be sent for sequencing (highlighted in green in Table 11).

Calf ID	Total Wisconsin Score	Weight Estimate (kg)
7201	2	150
7216	1	166
7234	0	170
7213	0	174
7212	0	178
7189	0	182
7193	3	186
7187	2	190
7168	0	202
7167	0	220
7171	0	230
7163	0	235

Table 15: Sample Selection of Younger Calves Using Wisconsin Score and Weight

	Wisconsin Score	Weight Estimate (kg)
AVERAGE	1	190
MEDIAN	0	184
1st QUART	0	171
3rd QUART	2	216

Table 16: Mean, median, 1st quartile and 3rd quartile of Wisconsin Score and Weight for Younger Calves

3.7.3. Older and younger selected calves

Shown in Table 17 are all of the calves selected to send for shotgun metagenome sequencing. Eight of the selected calves are 7 to 8 months old (7099, 7118, 7069, 7115, 7129, 7095, 7005 and 7124). Out of these eight, three calves (7129, 7095 and 7005) are classified as “healthy” according to their

total weight gain, ratio of weight gain between weeks 0 to 2 and weeks 2 to 4, and Wisconsin score. The remaining five calves (7099, 7118, 7069, 7115 and 7124) were classified as "unhealthy" due to showing poor or abnormal weight. Three calves aged 14 weeks old were also selected (7213, 7212 and 7189). These calves were all considered "healthy" as they had a weight close to the median as well as a Wisconsin score of 0.

Calf ID	Age	Sample Collection Date	Sample ID	Wisconsin Score	Weight (kg)	Weight gain (kg)
7099	7-8 months	14.11.2018	6B	2	243	NA
7099	7-8 months	12.12.2018	48B	1	249	6
7118	7-8 months	14.11.2018	17B	3	247	NA
7118	7-8 months	12.12.2018	59B	1	272	25
7069	7-8 months	14.11.2018	14B	1	220	NA
7069	7-8 months	12.12.2018	56B	1	246	26
7115	7-8 months	14.11.2018	3B	0	280	NA
7115	7-8 months	12.12.2018	45B	1	314	34
7129	7-8 months	14.11.2018	9B	0	210	NA
7129	7-8 months	12.12.2018	51B	0	246	36
7095	7-8 months	14.11.2018	11B	0	264	NA
7095	7-8 months	12.12.2018	53B	0	308	44
7005	7-8 months	14.11.2018	8B	1	257	NA
7005	7-8 months	12.12.2018	50B	0	303	46
7124	7-8 months	14.11.2018	18B	3	225	NA
7124	7-8 months	12.12.2018	60B	0	272	47
7213	14 weeks	12.12.2018	11E	0	174	NA
7212	14 weeks	12.12.2018	2E	0	178	NA
7189	14 weeks	12.12.2018	12E	0	182	NA

Table 17: All calves and samples selected for further analysis via shotgun metagenome sequencing

3.8. Metagenome analysis: Sample Preparation

All selected samples had 20µl aliquoted. The highest amount of DNA in an aliquot was sample 9B with 4242ng of DNA. The lowest amount of DNA in an aliquot was 245ng for sample 18B.

3.9. Testing for Cryptosporidium Parasite Using PCR

The gel shown in Figure 43 shows the presence of a band of DNA at approximately the 830bp mark on the DNA ladder. This formed a positive control when performing PCR for the detection of cryptosporidium in the clinical samples.

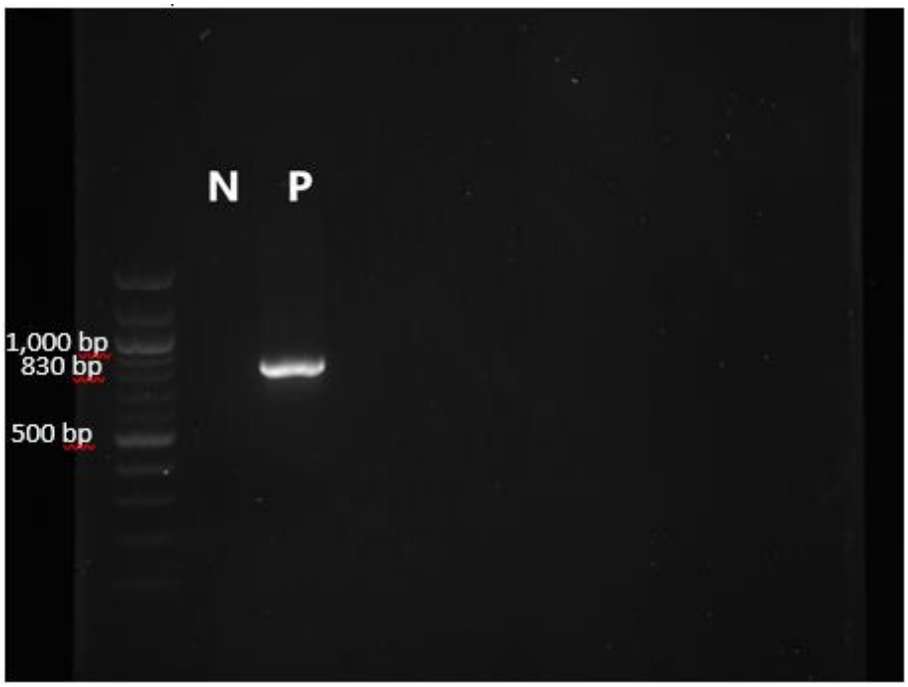


Figure 43: PCR Gel showing positive control sample

Figure 44, Figure 45 and Figure 46 show the results from gel electrophosesis of the PCR for the detection of Cryptosporidium parasite in the samples. The columns labelled "N" in Figure 44 and Figure 45 are the negative controls and neither of these shows any bands. The columns labelled "P1" in Figure 44 and "P" in Figure 45 are the positive controls and both show a clear band at 819-825bp. Columns labelled 1 to 20 in all three figures are the samples and none show any bands.

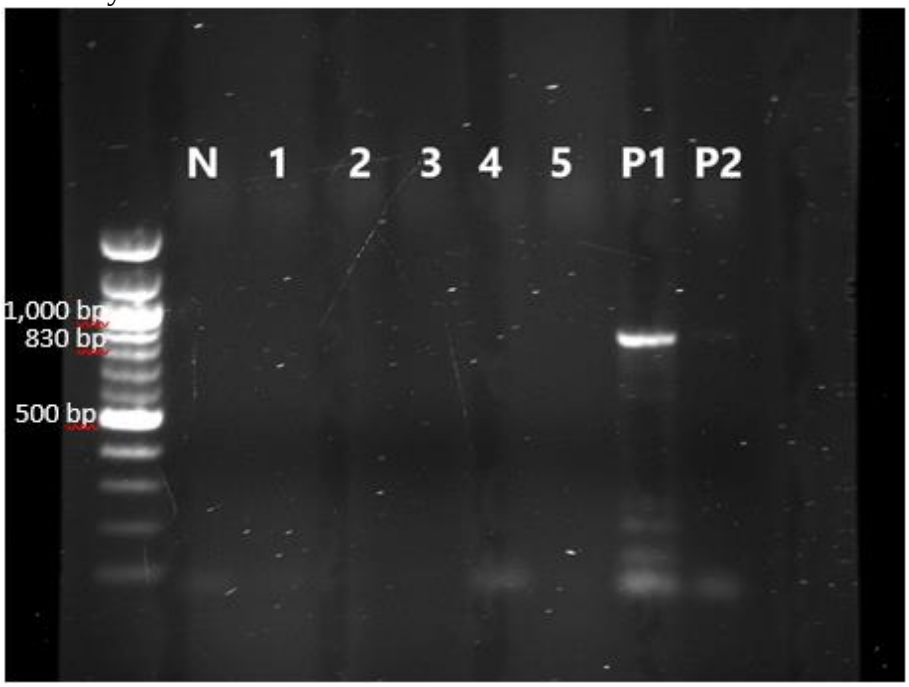


Figure 44: PCR Gel from PCR for the detection of Cryptosporidium. N - Negative control. P1 - Positive control. P2 - Rinsed out positive control. 1 to 5 - Samples

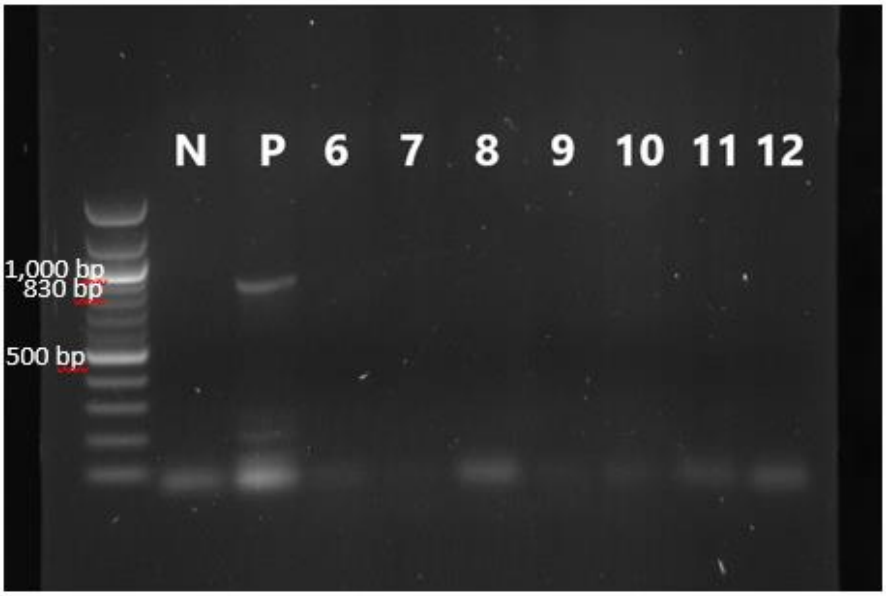


Figure 45: PCR Gel shows PCR for the detection of Cryptosporidium. N - Negative control. P - Positive control. 6 to 12 - Samples

Figure 46: PCR Gel show PCR for detection of Cryptosporidium. 13 to 20 - Samples

3.10. Library Preparation and Metagenome Sequencing

The library preparation went well and produced libraries from all of the samples except sample C. Sample C was a negative control so nothing was able to be amplified during the library preparation and could not be detected by the TapeStation (detection limit 5pg/µL) so it was left out of the final sequencing.

3.11. Bioinformatics – Metagenomics analysis

3.11.1. Heat Maps

The heat maps, shown in Figure 47, Figure 48 and Figure 49, show the sequences sequenced for each sample clustered according to similarity. The darker the square where the samples meet the more similar the two samples are. On the left-hand side of the maps, a colour coded key demonstrates the phenotype of the sample. These heat maps show that the data is characterized by the presence of a high proportion of unclassified reads and suggest that the unclassified half of the data is highly divergent between all of the samples except samples 59B (7118, 7-8 months old, "unhealthy", Wisconsin score 1) and 60B (7124, 7-8 months old, "unhealthy", Wisconsin score 0) . For the definition of unclassified and classified data see the materials and methods. When instead the reads where assigned to microbial groups, taxonomic assignment, and used to build the heat map, moderate clustering (0.8) between samples 48B (7099, 12.12.2018, 7-8 months, "unhealthy"), 45B (7115, 12.12.2018, 7-8 months, "unhealthy") and 17B (7118, 14.11.2018, 7-8 months, "unhealthy") was found see Figure 48. All three of these calves were classified as "unhealthy" due to poor or abnormal weight gain and none of these three samples was from the same calf. There is also some clustering (0.7) visible between samples 59B, 53B, 45B, 17B, 48B, 2E, 18B and 14B. The taxonomically unclassified reads that are an abundant part of the sequencing outputs strongly impact the patterns obtained with the microbial assigned ones (see Figure 47).

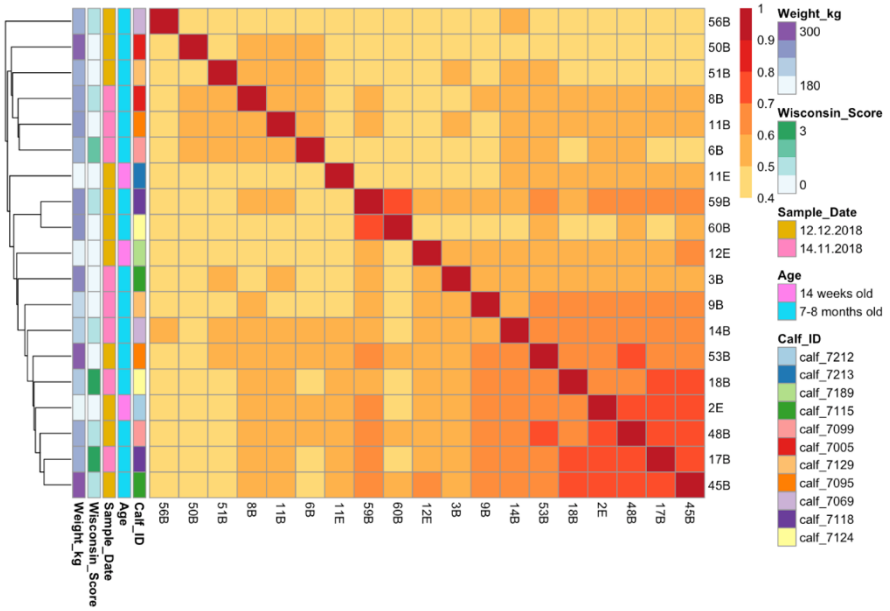


Figure 47: Heat map showing all reads clustered according to the similarity

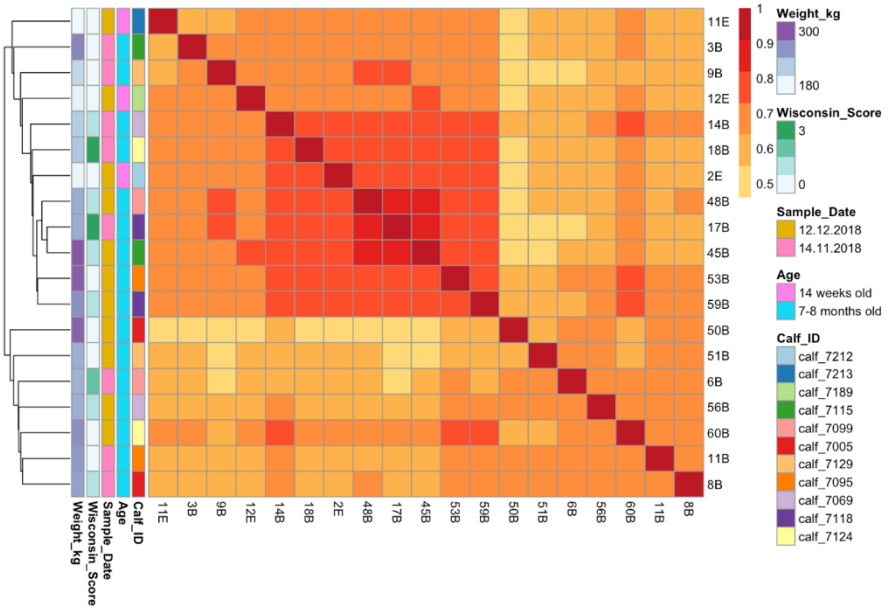


Figure 48: Heat map showing all reads assigned to known microbial groups by kaiju, clustered according to the similarity

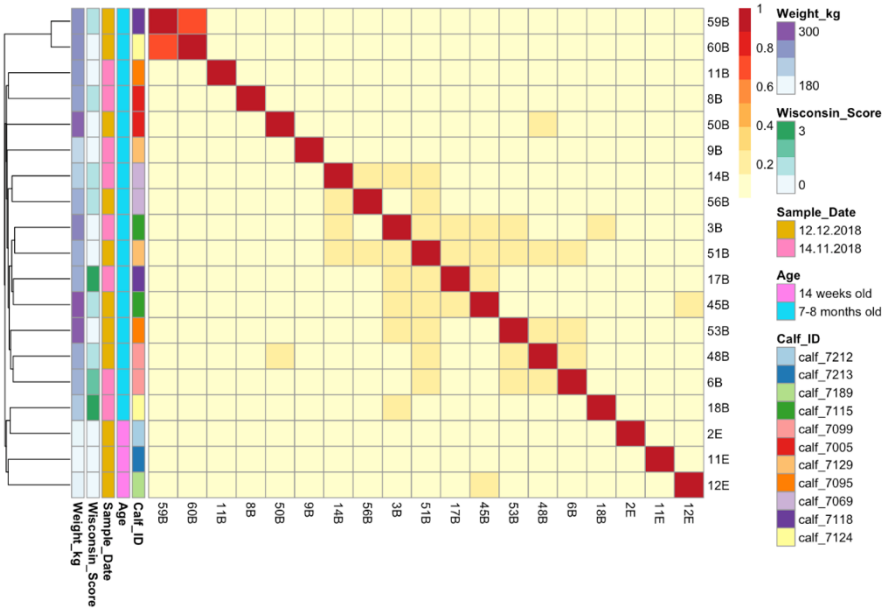


Figure 49: Heat map showing remaining reads with no taxonomic assignment, clustered according to the similarity

3.11.2. Taxonomy

3.11.2.1. Families

At a family level between 14.28 and 18.83% of sequences could not be assigned. Between 7.6% and 14.76% of reads were assigned to a family present in an abundance of less than 0.5%. These can all be seen in Figure 50.

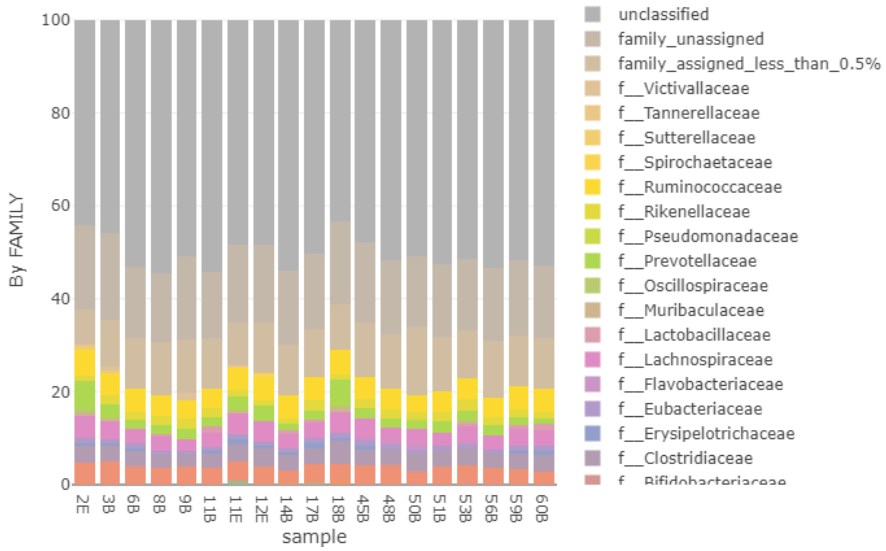


Figure 50: Bar chart showing proportions of families of bacteria in each of the samples

The most abundant family in all of the samples was Ruminococcaceae with between 3.74 (sample 50B) and 5.56 % (sample 2E). Ruminococcaceae, Rikenellacaea, Prevotellaceae, Lachnospriraceae, Eubacteriaceae, Clostridiaceae and Bacteroidaceae were all families present at greater than 0.5% abundance in all samples. Sutterellacaea and Oscillospiraceae were only found in greater than 0.5% abundance in sample 2E (7212, 14 weeks old, 12.12.2018, “healthy”) whereas Spirochaetaceae was only greater than 0.5% abundance in sample 12E (7189, 14 weeks old, 12.12.2018, “healthy”). Sample 50B (7005, 7-8 months, 12.12.2018, “healthy”) was the only sample where Pseudomonadaceae was present in over 0.5% abundance while the same was true for Atopobiacaea in sample 17B and Akkermansiaceae in sample 11E (7213, 14 weeks old, 12.12.2018, “healthy”).

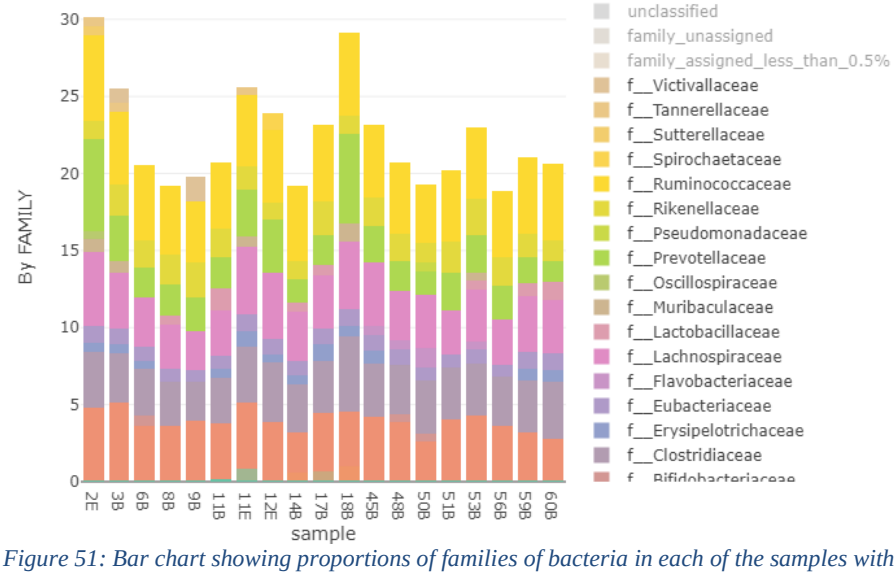


Figure 51: Bar chart showing proportions of families of bacteria in each of the samples with the unclassified, class unassigned and classes assigned at less than 0.5% removed

3.11.2.2. Orders

As shown in Figure 52, at an order level between 8.58 (sample 11B) and 11.89% (sample 9B) of reads were unassigned to an order. Reads that are unassigned to a taxonomic order could be due to many reasons including; ambiguity, poor mapping quality or short fragment length. These result in a read that cannot be assigned at an order level. The abundance of reads assigned to an order with a relative abundance of less than 0.5% was between 5.56 (sample 2E) and 8.13% (sample 50B).

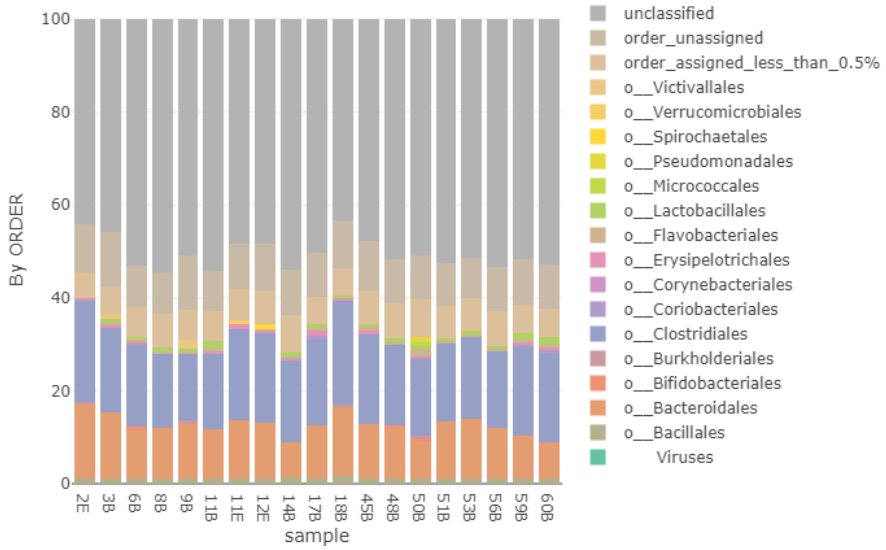


Figure 52: Bar chart showing proportions of orders of bacteria in each of the samples

Figure 53 shows that Clostridiales, Bacteroidales, and Bacillales were found in all of the samples. Clostridales was the most abundant order in all of the samples present in abundances from 14.39% (sample 9B) to 22.23% (sample 18B). Verrucomicrobiales were only present in over 0.5% relative abundance in sample 11E while the same was true for Spirochaetales in sample 12E. Pseudomonadales, Micrococcales and Corynebacteriales were only abundant at levels greater than 0.5% in sample 50B. Lactobacillales were not found at greater than 0.5% in sample 2E (7212, 14 weeks old, 12.12.2018, “healthy”), 11E (7213, 14 weeks old, 12.12.2018, “healthy”) or 12E (7189, 14 weeks old, 12.12.2018, “healthy”).

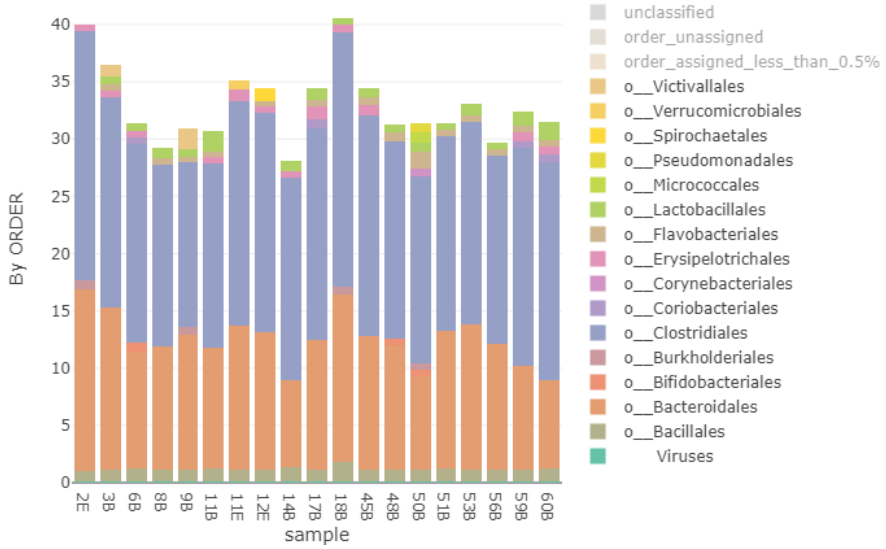


Figure 53: Bar chart showing proportions of orders of bacteria in each of the samples with the unclassified, class unassigned and classes assigned at less than 0.5% removed

4. Discussion

This study represents a novel and ambitious approach to this type of research. It focused on generating preliminary data and investigating a new analytical approach that could be developed and used to monitor cattle health. The aim was to collect a large amount of heterogeneous data, including field information, clinical observations, and sensor data together with information about the animal microbiome. The possibility of a new approach was investigated where the data were integrated and analysed to identify the existence of a correlation between different health indicators such as weight, Wisconsin clinical score, core temperature, and the gut microbiome.

4.1. Wisconsin Scores

Wisconsin scoring allows evaluation of calf health with minimal manual handling and no invasive procedures, such as blood sampling. In addition to this, Wisconsin scoring has been extensively reviewed and used across the literature. The use of a scoring system reduces the risk of subjectivity between assessors. The scoring covers a range of parameters to examine respiratory health (coughing, nasal discharge, ocular discharge) and gut health (faeces) in addition to general health (temperature, ear position, naval, joints). Usually, clinical illness is defined as a total Wisconsin score ≥ 5 (McGuirk, 2008). None of the calves in this study scored higher than 4 indicating that the calves in this study were clinically healthy at the time of the scoring with no signs of clinical disease. This was to be expected for several reasons. Firstly, the calves were vaccinated appropriately at 3 months old for pneumonia, lungworm, and IBR. These vaccinations should prevent most cases in healthy individuals hence the lack of clinical illness identified by the clinical scoring (Makoschey et al., 2008, Jericho and Langford, 1982, Jacobs, 2017). Although, the coughing score could be indicative of respiratory or general illness it could also be affected by factors such as environmental temperature, air quality and quality of bedding materials. However, this was controlled as much as possible

by keeping the calves in the same groups in the same sheds with the same bedding. The environmental temperature, humidity, and air quality were not able to be controlled due to the study being carried out under field conditions. Most clinical respiratory disease occurs between the ages of 1 week old and peaks at 5 to 7 weeks old so it is perhaps not unusual that we should see no clinical disease in our sampling cohort since the youngest calves were already 14 weeks of age (Ames, 1997, Sivula et al., 1996, Waltner-Toews et al., 1986). One issue with the Wisconsin scoring system is that it is focused on detecting respiratory disease. This means that despite including parameters indicative of general health, such as ear position, or intestinal health, such as faecal score, it may not be the most appropriate system to detect gastrointestinal disease, but currently, no other studies appear to validate this. Additionally, the method of clinical scoring could still be subjective depending on the observer. The effect of this was limited by using the same observer scoring under the same conditions for every calf on every sampling date however some observer bias could still be present due to the nature of the method.

4.2. Calf Weight

Another way in which the health of the calves in this study was assessed was via weight. Illness, including subclinical conditions, may influence the weight of an animal causing either a decrease in average weight gain over time or even weight loss in severe cases (Place et al., 1998). We used a weigh band in order to determine weight. Using a weigh band in order to measure heart girth circumference has been described as an accurate method to determine weight when devices such as weighing scales aren't available (Davis et al., 1961). The breed of the animal can have an impact on the relationship between heart girth circumference and weight so when selecting the correct circumference to weight conversion we took this into account and used one designed for Holstein-Friesian dairy calves. Despite the strong correlation between weight and heart girth circumference in cows, the method of using a weigh band is still subject to errors that would not be encountered using digital scales. The exact position of the band around the calves' heart girth circumference can vary enough to greatly affect the predicted weight measured. Additionally, the pressure applied to the band by the user can alter the measurements obtained. In order to reduce the effects of user error of the weigh band, the same operator took all the weigh band measurements. This ensured that band position and pressure applied were as similar as possible across all measurements.
For this study since we were looking more at the trend in weight over time rather than exact weight so the difference between actual weight and weight determined by the weigh band was less important as long as it was consistent. In future studies, weighing scales or a weigh crush would have been preferable due to the accuracy and reliability they provide. Current work is in progress to develop weighing platforms that work wirelessly to obtain the weights of the calves without the need to run them through a weigh crush (Swire, 2017). Despite all the calves in each group being the same age, under the same management, in the same housing and offered the same diet there was variation seen between the calves' weights and the amount of weight they gained. From this, we can conclude that individuals have a natural variation in weight and weight gain over time. One reason for this could be due to social circumstances. Studies have found that cows have a social hierarchy within the

herd that dictates which cows have priority over access to resources such as feed and water troughs (Bøe and Færevik, 2003, Lauwere et al., 1996, Reinhardt and Reinhardt, 1975). Although there is little research on social structures within groups of calves this unevenness in access to the feed offered could be a possible reason for the variation in weight seen across groups of calves (Sato, 1984). Another reason could be varying activity levels across the calves. We did not look at the activity levels in the calves but within herds of cattle, it is not uncommon for some cows to be more or less active than others (Turner et al., 2000, Rutten et al., 2017, Roelofs et al., 2005). A calf with an increased activity level may explain a lower weight in comparison to a less active calf on the same diet. Furthermore, the exact birth dates of each of the calves included in the study were unknown. Since calves should be gaining around 600g per day, a few days difference in birth date could influence some of the minor differences in weight observed but does not explain the larger differences in weights (McNaughton and Lopdell, 2012).
Regardless of the variation in weight between the calves, the majority of them showed a good pattern of healthy weight gain over the study period.

4.3. Comparison of Calf Weight and Wisconsin Clinical Score

One of the objectives of this study was to compare the calf weight and Wisconsin score metrics to see if there is any correlation. To our knowledge, no previous study has looked at the relationship between these factors so by comparing the weight and Wisconsin scores for each calf we hoped to get a better overall idea of calf health. Despite there being little direct correlation between the two scores both agreed that healthy calves were healthy calves. The lack of correlation is perhaps to be expected for a few reasons. Firstly, it would be expected that there would be a time-lapse between the manifestation of the clinical signs measured by the Wisconsin score and the effect of clinical disease on the weight of an animal (Place et al., 1998). Additionally, external factors could be acting on the results. The Wisconsin score could be affected by changes in the weather such as cold or reduced humidity causing an increase in respiratory signs that are related to the environment rather than clinical disease. Weight could be influenced by social factors within a group of calves meaning that certain dominant calves have better access to feed (Sato, 1984).
In order to confirm the conclusion that the Wisconsin score and calf weights can identify whether a calf is healthy the same study would need to include a cohort of unhealthy calves. Also, by extending the duration of the study from 6 weeks to 12 months or longer it would be easier to study the correlation of these factors while accounting for any lag seen between clinical disease and weight.

4.4. Core Temperature

The core temperature readings supported the results of the Wisconsin scoring identifying that the animals in this study were healthy during the recording period. This is promising evidence that the boluses can be used to monitor clinically normal calves. However, as with Wisconsin scoring the further research needs to be done including unhealthy calves in order to confirm that the changes in core temperature trends will be detected and able to differentiate between the unhealthy and healthy animals.

The use of the boluses in this study was not without its challenges since this was an early generation of the technology that is continuing to be improved and developed. The boluses used in this study had no internal memory meaning that if the temperature, recorded every 2 minutes, was not transmitted immediately it would be lost. The boluses were able to transmit data up to 30m which was adequate to cover all areas of the calves' pen. However, numerous gaps in the temperature data indicated that there were problems with the transmission of the data at certain points for most calves. These transmission problems were hypothesised to occur during periods during which the calves were lying down but have not been confirmed. Regardless, the data for these points are missing. Future versions of the boluses would benefit from including internal memory allowing for un-transmitted data to be stored until it is able to be transmitted.

Additionally, there were issues with the retention and battery life of the boluses. For this study that meant that only the 14th November had temperature data from 7 days and 7 days after the sample and clinical information collection. More studies need to be done focusing on the best size, shape, and weight of the boluses in order to achieve the best retention rate for the boluses. Additionally, optimisation of the recording and transmission frequency may allow for boluses with longer battery lives that therefore need replacing less often.

4.5. PCR for Detection of Cryptosporidium Parasite

The negative results from the PCR for Cryptosporidium were unexpected. Although no calves showed clinical signs it would be expected that the protozoa would be present on the farm at least a subclinical level as is common on UK farms (Bodley-Tickell et al., 2002, Razakandrainibe et al., 2018). Two possible conclusions could be drawn from these findings. Either the calves are free from any Cryptosporidium infection or the calves are sub-clinically infected so the numbers of oocysts being shed in faeces are too small to be detected in the rectal swabs.

This information shows that in future studies a larger volume of faecal samples from each calf should be taken to pool and test for parasites. This would have been more likely to identify the presence of any parasites present in too smaller numbers to cause clinical disease and hence shedding a reduced volume of eggs or oocysts (Quilez et al., 1996).

4.6. Calf Metagenome

The sequencing results demonstrate that these calves' metagenomes are populated by the same bacteria as calves in previous studies of a similar kind studies (Jami et al., 2013, Durso et al., 2010, Dowd et al., 2008, Stewart et al., 2018, Özbayram et al., 2018). We were able to assign at genera level for much of the metagenome making our data comparable to other studies. There are some discrepancies between the exact abundances of bacteria between studies. These are to be expected for several reasons. Different studies use different sampling methods, DNA extraction protocols, and sequencing technologies. This means that at each level of the process there is a bias present altering the final abundances of organisms reported in the results. Even the use of calves on a farm that has not previously been studied could alter the observed abundances of organisms due to variation between farms (Gomez et al., 2017).

Our research also supports previous evidence that there is a considerable amount of individual variation in the metagenome between healthy age-matched animals on the same farm (Gomez et al., 2017, Durso et al., 2010). This variation in individual metagenome could perhaps explain why the pattern of weight gain is not identical between calves despite living in the same environment and receiving the same feed. By identifying individuals showing optimal weight gain and comparing their metagenome to calves showing less optimal weight gain the exact bacteria responsible for improved performance could be identified. Identification of these bacterial groups could inform future work into the production of supplements for calves in order to achieve an optimal metagenome for performance.

Interestingly the metagenome of the same calves but on different dates appeared to show little similarity in bacterial abundances. This strongly suggests that the gut metagenome is constantly evolving and changing significantly even over a short period of time. One conclusion that could be drawn from this is that perhaps there is no optimal metagenome for healthy calves but instead that health is achieved from the ability of the gut metagenome to change and adapt.

When comparing the metagenome for older and younger calves the main finding of interest was concerning the Lactobacillales order and its increased abundance in the guts of older calves. Lactobacilli have also been isolated in the guts of human infants, however, unlike in our results, they appeared to decrease with age in humans (Nagpal et al., 2018). The Lactobacillilales order has been shown to have an important role in increasing the gut levels of short-chain fatty acids resulting in immune-modulating and anti-inflammatory effects (Vulevic et al., 2008, Jandhyala et al., 2015, Kim et al., 2011). This order has also been used in many studies to examine their probiotic properties (Frizzo et al., 2010, Lebeer et al., 2008, Signorini et al., 2012). It was also noted in the literature that in humans Lactobacilli populations in the gut, despite being found in relatively high numbers, are in fact very dependent on dietary sources to establish a population in the gut (Walter, 2008). It may be that Lactobacillilales were found in the 7-8-month calves rather than the 14-week old calves due to the prolonged duration of feeding since weaning. Lactobacillilales is a commonly used bacteria in probiotics so if the calves feed contained this as an additive it may explain the presence in the older calves' guts as they had been consuming the feed for a longer duration of time. However, since we were unable to obtain information about the exact formulation of the feed this remains a theory.

The metagenome data in this research featured a lot of sequences that could not be classified at any level. Previous studies do not appear to have quite such a high number of unclassified reads as seen in this current study (Biscarini et al., 2018, Gomez et al., 2017, Meale et al., 2017, Smith et al., 2018, Thomas et al., 2017). Reasons for higher numbers of unclassified reads in our data could be either from contamination in the sampling, as we were unable to undertake more invasive and sterile techniques, or due to the limited databases, we used to classify our data. This means that conclusions were only able to be made based on the classified data so there may be many more organisms present that were not able to be identified given the time restraints to complete this work. However, the sequencing results are currently under investigation with more sophisticated and laborious analytical approaches.

Our findings highlight the importance of finding dedicated pipelines to analyse the unclassified and unassigned reads to provide a comprehensive overview of metagenome data. Since around 50% of the data remained unclassified with more unable to be assigned at lower taxonomic levels it was therefore unable to be analysed so it cannot be said for sure that there is no pattern to be observed in the metagenome of the different phenotypic types of calves. A pipeline able to classify the missing information will not only enable a better evaluation of the taxonomy at a family and order level but would also enable for enough classified data to interpret at a deeper level such as genera or species. Despite the best attempt being made to minimise contamination of swabs and using DNA extraction methods that optimise the extraction of bacterial DNA, there is still the potential for non-bacterial DNA to be present in the reads. Firstly, bovine DNA will be present in the samples as the DNA extraction kit was unable to completely select for bacterial DNA. This was accounted for in the processing steps by removing reads that aligned with a known bovine DNA database. However, there could still be parts of bovine DNA that remain unknown or not included in the database and therefore appear as unclassified. Additionally, with rectal samples, there is a large amount of plant and digestive matter from the faeces present. It is possible that digesta DNA makes up a proportion of the unclassified reads. The use of more databases including possible sources of contamination could help to identify more of the unclassified reads and differentiate between which reads represent contaminants rather than metagenome data.

4.7. Future Work

This study was a pilot aimed at studying the impact of core temperature, weight, Wisconsin score and metagenome on calf health. It was hoped to determine the feasibility of studying multiple factors and identify any improvements to be included in future work.

Firstly, any future study would need to include a larger number of animals studied over a longer period. The inclusion of more animals may require the recruitment of larger farms and the inclusion of multiple farms. However, if multiple farms are studied then differences between the farms would need to be considered for in the interpretation of the results. A longer time period with more regular sampling intervals would allow daily or weekly fluctuations in metagenome and core temperature to be identified and accounted for in the results. Ideally, calves should be sampled from birth up until their first lactation either every two weeks of every month with core temperature data provided daily and weekly weights taken.

In future studies, it would be extremely valuable to have a full medical history for the calves from the day of birth up to 6 weeks after the final samples were collected. As well as being able to correlate Wisconsin scores, core temperature, weight and metagenome to clinical outcomes the treatment data could allow examination of the effects of drugs, such as antibiotics on these parameters.

In addition to the calf data, future studies should look at including metagenome, disease and weight gain data from the dams. This would allow the influence of any genetic or epigenetic links between calf health and the dam to be explored. The limitation with both having disease data for calves and having more data on the dams is that very few farms record to that level of

detail. This means that any study wanting to utilize this data may need to span many years in order to start recording the data themselves from farms.
When recording the core temperature, it would have been useful to have an ambient temperature as a comparison. This could easily be included via an electronic thermometer placed in the calf housing. Variations in ambient temperature could influence many factors such as appetite and disease prevalence making it an essential piece of data to include in an analysis of calf health. Rectal temperature taken during Wisconsin scoring could also be valuable to compare to the core temperature.
For this study, the drinking events were removed using a set threshold that was the same for all calves. However, this doesn't account for any individual variation in core temperature and the reduction in core temperature experienced as a result of drinking. Current research is already investigating how to adapt the threshold used for drinking events in order to better consider individual, farm and seasonal variability (Vázquez-Diosdado et al., 2019). However, these studies only focus on the use of the boluses in adult dairy calves so it would be interesting to repeat the methods on different ages of dairy calves to see if the same results are observed.
Drinking events were removed in this study in order to focus more on the variations in core temperature influenced by disease. However, it would be interesting to use the removed data to further study the duration, frequency, and volumes of fluid consumed by the calves. It would be interesting to investigate the changes in drinking behaviour exhibited by different individuals, different aged calves and calves of different health statuses.

5. Conclusions

In conclusion, no correlation was able to be determined between an individual calf's metagenome and the resulting weight gain. This is likely to be heavily influenced by the high proportion of unclassified reads in the metagenome data. The metagenomes appeared to vary significantly between individuals regardless of the weight gain exhibited so no pattern was able to be observed. There also seemed to be little evidence that the metagenome continued to develop between week 14 and week life. Instead, the metagenome appears to be more dynamic and constantly changing within individuals.
No conclusion was able to be formed regarding the relationship between core temperature and weight gain since no significantly abnormal temperatures were recorded. Future studies may better be able to investigate the correlation between core temperature and weight gain by using a larger sample size over a longer duration in order to more accurately identify patterns and abnormalities in the data. It would also be essential in the future to establish a reliable pipeline for the analysis of the unclassified proportions of metagenome data. This data may hold key differences in the metagenome that could be essential to identify correlations between other phenotypic features.
Additionally, none of the animals in this study were clinically un-healthy or receiving any oral medication over the duration of this study. It would be interesting to see if moderate clinical illness and the subsequent treatment altered the metagenome significantly.

6. Works Cited

ADAMS, A. E., OLEA-POPELKA, F. J. & ROMAN-MUNIZ, I. N. 2013. Using temperature-sensing reticular boluses to aid in the detection of production diseases in dairy cows. *Journal of Dairy Science,* 96**,** 1549-1555.

ALIPOUR, M. J., JALANKA, J., PESSA-MORIKAWA, T., KOKKONEN, T., SATOKARI, R., HYNÖNEN, U., IIVANAINEN, A. & NIKU, M. 2018. The composition of the perinatal intestinal microbiota in cattle. *Scientific reports,* 8**,** 10437.

ALY, S. S., LOVE, W. J., WILLIAMS, D. R., LEHENBAUER, T. W., VAN EENENNAAM, A., DRAKE, C., KASS, P. H. & FARVER, T. B. 2014. Agreement between bovine respiratory disease scoring systems for pre-weaned dairy calves. *Animal Health Research Reviews,* 15**,** 148-150.

ALZAHAL, O., ALZAHAL, H., STEELE, M. A., VAN SCHAIK, M., KYRIAZAKIS, I., DUFFIELD, T. F. & MCBRIDE, B. W. 2011. The use of a radiotelemetric ruminal bolus to detect body temperature changes in lactating dairy cattle. *Journal of Dairy Science,* 94**,** 3568-3574.

AMES, T. R. 1997. Dairy calf pneumonia: the disease and its impact. *Veterinary Clinics: Food Animal Practice,* 13**,** 379-391.

ANASTÁCIO, S., CAROLINO, N., SIDI-BOUMEDINE, K. & SILVA, G. 2016. Q fever dairy herd status determination based on serological and molecular analysis of bulk tank milk. *Transboundary and emerging diseases,* 63.

ANSORGE, W. J. 2009. Next-generation DNA sequencing techniques. *New Biotechnology,* 25**,** 195-203.

BARKEMA, H. W., VON KEYSERLINGK, M. A. G., KASTELIC, J. P., LAM, T. J. G. M., LUBY, C., ROY, J. P., LEBLANC, S. J., KEEFE, G. P. & KELTON, D. F. 2015. Changes in the dairy industry affecting dairy cattle health and welfare. *Journal of Dairy Science,* 98**,** 7426-7445.

BATE, A. 2016. UK Dairy Industry Statistics. UK: House of Commons Library.

BEHJATI, S. & TARPEY, P. S. 2013. What is next generation sequencing? *Archives of disease in childhood - Education & practice edition,* 98**,** 236-238.

BERRY, D. P., HORAN, B. & DILLON, P. 2005. Comparison of growth curves of three strains of female dairy cattle. *Animal Science,* 80**,** 151-160.

BEWLEY, J., EINSTEIN, M., GROTT, M. & SCHUTZ, M. 2008a. Comparison of reticular and rectal core body temperatures in lactating dairy cows. *Journal of dairy science,* 91**,** 4661-4672.

BEWLEY, J., GROTT, M., EINSTEIN, M. & SCHUTZ, M. 2008b. Impact of intake water temperatures on reticular temperatures of lactating dairy cows. *Journal of dairy science,* 91**,** 3880-3887.

BIKKER, J., VAN LAAR, H., RUMP, P., DOORENBOS, J., VAN MEURS, K., GRIFFIOEN, G. & DIJKSTRA, J. 2014. Evaluation of an ear-

attached movement sensor to record cow feeding behavior and activity. *Journal of Dairy Science,* 97**,** 2974-2979.

BISCARINI, F., PALAZZO, F., CASTELLANI, F., MASETTI, G., GROTTA, L., CICHELLI, A. & MARTINO, G. 2018. Rumen microbiome in dairy calves fed copper and grape-pomace dietary supplementations: Composition and predicted functional profile. *PLOS ONE,* 13**,** e0205670.

BODLEY-TICKELL, A. T., KITCHEN, S. E. & STURDEE, A. P. 2002. Occurrence of Cryptosporidium in agricultural surface waters during an annual farming cycle in lowland UK. *Water Research,* 36**,** 1880-1886.

BØE, K. E. & FÆREVIK, G. 2003. Grouping and social preferences in calves, heifers and cows. *Applied Animal Behaviour Science,* 80**,** 175-190.

BOKULICH, N. A., CHUNG, J., BATTAGLIA, T., HENDERSON, N., JAY, M., LI, H., D. LIEBER, A., WU, F., PEREZ-PEREZ, G. I., CHEN, Y., SCHWEIZER, W., ZHENG, X., CONTRERAS, M., DOMINGUEZ-BELLO, M. G. & BLASER, M. J. 2016. Antibiotics, birth mode, and diet shape microbiome maturation during early life. *Science Translational Medicine,* 8**,** 343ra82-343ra82.

BROOK, E., HART, C. A., FRENCH, N. & CHRISTLEY, R. 2008a. Prevalence and risk factors for Cryptosporidium spp. infection in young calves. *Veterinary Parasitology,* 152**,** 46-52.

BROOK, E. J., CHRISTLEY, R. M., FRENCH, N. P. & HART, C. A. 2008b. Detection of Cryptosporidium oocysts in fresh and frozen cattle faeces: comparison of three methods. *Letters in Applied Microbiology,* 46**,** 26-31.

BROWN, C. T. & IRBER, L. 2016. sourmash: a library for MinHash sketching of DNA. *J. Open Source Softw,* 1**,** 27.

BROWNLOW, R. J., DAGNALL, K. E. & AMES, C. E. 2012. A Comparison of DNA Collection and Retrieval from Two Swab Types (Cotton and Nylon Flocked Swab) when Processed Using Three QIAGEN Extraction Methods. *Journal of Forensic Sciences,* 57**,** 713-717.

BUCZINSKI, S., FORTÉ, G., FRANCOZ, D. & BÉLANGER, A. M. 2014. Comparison of thoracic auscultation, clinical score, and ultrasonography as indicators of bovine respiratory disease in preweaned dairy calves. *Journal of veterinary internal medicine,* 28**,** 234-242.

BUCZINSKI, S., L OLLIVETT, T. & DENDUKURI, N. 2015. Bayesian estimation of the accuracy of the calf respiratory scoring chart and ultrasonography for the diagnosis of bovine respiratory disease in pre-weaned dairy calves. *Preventive Veterinary Medicine,* 119**,** 227-231.

C. HICKS, L., S. HICKS, W., A. BUCKLIN, R., K. SHEARER, J., R. BRAY, D., SOTO & V. CARVALHO, P. Comparison of Methods of Measuring Deep Body Temperatures of Dairy Cows. Pp. 432-438 in Livestock Environment VI: Proceedings of the 6th International Symposium (21-23 May 2001, Louisville, Kentucky, USA) ed. Richard R. Stowell, Ray Bucklin, and Robert W. Bottcher., St. Joseph, MI. ASABE.

CARDONA, S., ECK, A., CASSELLAS, M., GALLART, M., ALASTRUE, C., DORE, J., AZPIROZ, F., ROCA, J., GUARNER, F. &

MANICHANH, C. 2012. Storage conditions of intestinal microbiota matter in metagenomic analysis. *BMC microbiology,* 12**,** 158.
CARROLL, I. M., RINGEL-KULKA, T., SIDDLE, J. P., KLAENHAMMER, T. R. & RINGEL, Y. 2012. Characterization of the fecal microbiota using high-throughput sequencing reveals a stable microbial community during storage. *PloS one,* 7**,** e46953.
CHALMERS, R. M. & GILES, M. 2010. Zoonotic cryptosporidiosis in the UK – challenges for control. *Journal of Applied Microbiology,* 109**,** 1487-1497.
CLAASSEN, S., DU TOIT, E., KABA, M., MOODLEY, C., ZAR, H. J. & NICOL, M. P. 2013. A comparison of the efficiency of five different commercial DNA extraction kits for extraction of DNA from faecal samples. *Journal of Microbiological Methods,* 94**,** 103-110.
CLAESSON, M. J., JEFFERY, I. B., CONDE, S., POWER, S. E., O'CONNOR, E. M., CUSACK, S., HARRIS, H. M., COAKLEY, M., LAKSHMINARAYANAN, B. & O'SULLIVAN, O. 2012. Gut microbiota composition correlates with diet and health in the elderly. *Nature,* 488**,** 178.
DAIRYCO. 2016. *Distribution By Size* [Online]. AHDB Dairy. Available: http://dairy.ahdb.org.uk/resources-library/market-information/farming-data/distribution-by-size/#.Wwvl5-4vzIV [Accessed 28/05 2018].
DAVIS, H. P., SWETT, W. W. & HARVEY, W. R. 1961. Relation of heart girth to weight in Holsteins and Jerseys.
DELGADO, B., BACH, A., GUASCH, I., GONZÁLEZ, C., ELCOSO, G., PRYCE, J. E. & GONZALEZ-RECIO, O. 2019. Whole rumen metagenome sequencing allows classifying and predicting feed efficiency and intake levels in cattle. *Scientific Reports,* 9**,** 11.
DOMINIANNI, C., WU, J., HAYES, R. B. & AHN, J. 2014. Comparison of methods for fecal microbiome biospecimen collection. *BMC microbiology,* 14**,** 103-103.
DONOVAN, G. A., DOHOO, I. R., MONTGOMERY, D. M. & BENNETT, F. L. 1998. Calf and disease factors affecting growth in female holstein calves in Florida, USA. *Preventive Veterinary Medicine,* 33**,** 1-10.
DOWD, S. E., CALLAWAY, T. R., WOLCOTT, R. D., SUN, Y., MCKEEHAN, T., HAGEVOORT, R. G. & EDRINGTON, T. S. 2008. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC microbiology,* 8**,** 125.
DOYLE, L., GORDON, A., ABERNETHY, D. & STEVENS, K. 2014. Bovine tuberculosis in Northern Ireland: Risk factors associated with time from post-outbreak test to subsequent herd breakdown. *Preventive veterinary medicine,* 116**,** 47-55.
DURSO, L. M., HARHAY, G. P., SMITH, T. P., BONO, J. L., DESANTIS, T. Z., HARHAY, D. M., ANDERSEN, G. L., KEEN, J. E., LAEGREID, W. W. & CLAWSON, M. L. 2010. Animal-to-animal variation in fecal microbial diversity among beef cattle. *Appl. Environ. Microbiol.,* 76**,** 4858-4862.
FAUST, M. A., KINSEL, M. L. & KIRKPATRICK, M. A. 2001. Characterizing Biosecurity, Health, and Culling During Dairy Herd Expansions1,2. *Journal of Dairy Science,* 84**,** 955-965.

FRANCIOSA, G., FERREIRA, J. L. & HATHEWAY, C. L. 1994. Detection of type A, B, and E botulism neurotoxin genes in Clostridium botulinum and other Clostridium species by PCR: evidence of unexpressed type B toxin genes in type A toxigenic organisms. *Journal of clinical microbiology,* 32**,** 1911-1917.

FRIZZO, L., SOTO, L., ZBRUN, M., BERTOZZI, E., SEQUEIRA, G., ARMESTO, R. R. & ROSMINI, M. 2010. Lactic acid bacteria to improve growth performance in young calves fed milk replacer and spray-dried whey powder. *Animal Feed Science and Technology,* 157**,** 159-167.

GARDNER, R. W., SMITH, L. W. & PARK, R. L. 1988. Feeding and Management of Dairy Heifers for Optimal Lifetime Productivity. *Journal of Dairy Science,* 71**,** 996-999.

GIOFFRÉ, A., MEICHTRI, L., ZUMÁRRAGA, M., RODRIGUEZ, R. & CATALDI, A. 2004. Evaluation of a QIAamp DNA stool purification kit for Shiga-toxigenic Escherichia coli detection in bovine fecal swabs by PCR. *Revista Argentina de microbiologia,* 36**,** 1.

GOMEZ, D. E., ARROYO, L. G., COSTA, M. C., VIEL, L. & WEESE, J. S. 2017. Characterization of the Fecal Bacterial Microbiota of Healthy and Diarrheic Dairy Calves. *Journal of Veterinary Internal Medicine,* 31**,** 928-939.

HART, M. L., MEYER, A., JOHNSON, P. J. & ERICSSON, A. C. 2015. Comparative Evaluation of DNA Extraction Methods from Feces of Multiple Host Species for Downstream Next-Generation Sequencing. *PLOS ONE,* 10**,** e0143334.

HEINRICHS, A. & HARGROVE, G. 1987. Standards of weight and height for Holstein heifers. *Journal of dairy science,* 70**,** 653-660.

HEINRICHS, A. J. 1993. Raising Dairy Replacements to Meet the Needs of the 21st Century. *Journal of Dairy Science,* 76**,** 3179-3187.

HOFFMAN, P. C. & FUNK, D. A. 1992. Applied dynamics of dairy replacement growth and management. *Journal of Dairy Science,* 75**,** 2504-2516.

HOLMAN, A., THOMPSON, J., ROUTLY, J., CAMERON, J., JONES, D., GROVE-WHITE, D., SMITH, R. & DOBSON, H. 2011. Comparison of oestrus detection methods in dairy cattle. *Veterinary Record-English Edition,* 169**,** 47.

HOSOMI, K., OHNO, H., MURAKAMI, H., NATSUME-KITATANI, Y., TANISAWA, K., HIRATA, S., SUZUKI, H., NAGATAKE, T., NISHINO, T., MIZUGUCHI, K., MIYACHI, M. & KUNISAWA, J. 2017. Method for preparing DNA from feces in guanidine thiocyanate solution affects 16S rRNA-based profiling of human microbiota diversity. *Scientific Reports,* 7**,** 4339.

INGALA, M. R., SIMMONS, N. B., WULTSCH, C., KRAMPIS, K., SPEER, K. A. & PERKINS, S. L. 2018. Comparing Microbiome Sampling Methods in a Wild Mammal: Fecal and Intestinal Samples Record Different Signals of Host Ecology, Evolution. *Frontiers in microbiology,* 9**,** 803-803.

INGLIS, G. D. & KALISCHUK, L. D. 2003. Use of PCR for direct detection of Campylobacter species in bovine feces. *Applied and environmental microbiology,* 69**,** 3435-3447.

IPEMA, A. H., GOENSE, D., HOGEWERF, P. H., HOUWERS, H. W. J. & VAN ROEST, H. 2008. Pilot study to monitor body temperature of dairy cows with a rumen bolus. *Computers and Electronics in Agriculture,* 64**,** 49-52.

JACOBS, A. A. C. 2017. Vaccine to protect a ruminant against pneumonia caused by pasteurella multocida. Google Patents.

JAMI, E., ISRAEL, A., KOTSER, A. & MIZRAHI, I. 2013. Exploring the bovine rumen bacterial community from birth to adulthood. *The ISME journal,* 7**,** 1069.

JANDHYALA, S. M., TALUKDAR, R., SUBRAMANYAM, C., VUYYURU, H., SASIKALA, M. & NAGESHWAR REDDY, D. 2015. Role of the normal gut microbiota. *World journal of gastroenterology,* 21**,** 8787-8803.

JEE, J., DEGRAVES, F. J., KIM, T. & KALTENBOECK, B. 2004. High prevalence of natural Chlamydophila species infection in calves. *Journal of clinical microbiology,* 42**,** 5664-5672.

JERICHO, K. & LANGFORD, E. 1982. Aerosol vaccination of calves with pasteurella haemolytica against experimental respiratory disease. *Canadian Journal of Comparative Medicine,* 46**,** 287.

KIM, D., SONG, L., BREITWIESER, F. P. & SALZBERG, S. L. 2016. Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Research*.

KIM, M.-K., LEE, H.-G., PARK, J.-A., KANG, S.-K. & CHOI, Y.-J. 2011. Effect of feeding direct-fed microbial as an alternative to antibiotics for the prophylaxis of calf diarrhea in Holstein calves. *Asian-Australasian Journal of Animal Sciences,* 24**,** 643-649.

KOCH, R. M., SWIGER, L. A., CHAMBERS, D. & GREGORY, K. E. 1963. Efficiency of Feed Use in Beef Cattle. *Journal of Animal Science,* 22**,** 486-494.

KOECK, A., LOKER, S., MIGLIOR, F., KELTON, D., JAMROZIK, J. & SCHENKEL, F. 2014. Genetic relationships of clinical mastitis, cystic ovaries, and lameness with milk yield and somatic cell score in first-lactation Canadian Holsteins. *Journal of dairy science,* 97**,** 5806-5813.

KUHNERT, P., DUBOSSON, C. R., ROESCH, M., HOMFELD, E., DOHERR, M. G. & BLUM, J. W. 2005. Prevalence and risk-factor analysis of Shiga toxigenic Escherichia coli in faecal samples of organically and conventionally farmed dairy cattle. *Veterinary Microbiology,* 109**,** 37-45.

KUMAR, J., KUMAR, M., GUPTA, S., AHMED, V., BHAMBI, M., PANDEY, R. & CHAUHAN, N. S. 2016. An Improved Methodology to Overcome Key Issues in Human Fecal Metagenomic DNA Extraction. *Genomics, Proteomics & Bioinformatics,* 14**,** 371-378.

LAM, H. Y. K., CLARK, M. J., CHEN, R., CHEN, R., NATSOULIS, G., O'HUALLACHAIN, M., DEWEY, F. E., HABEGGER, L., ASHLEY, E. A., GERSTEIN, M. B., BUTTE, A. J., JI, H. P. & SNYDER, M. 2011. Performance comparison of whole-genome sequencing platforms. *Nature Biotechnology,* 30**,** 78.

LAUWERE, C. C. K.-D., DEVIR, S. & METZ, J. H. M. 1996. The influence of social hierarchy on the time budget of cows and their visits to an

automatic milking system. *Applied Animal Behaviour Science,* 49**,** 199-211.
LEBEER, S., VANDERLEYDEN, J. & DE KEERSMAECKER, S. C. 2008. Genes and molecules of lactobacilli supporting probiotic action. *Microbiol. Mol. Biol. Rev.,* 72**,** 728-764.
LI, W., HAN, Y., YUAN, X., WANG, G., WANG, Z., PAN, Q., GAO, Y. & QU, Y. 2017. Metagenomic analysis reveals the influences of milk containing antibiotics on the rumen microbes of calves. *Archives of Microbiology,* 199**,** 433-443.
LIM, M. Y., SONG, E.-J., KIM, S. H., LEE, J. & NAM, Y.-D. 2018. Comparison of DNA extraction methods for human gut microbial community profiling. *Systematic and Applied Microbiology,* 41**,** 151-157.
LIMA, F. S., OIKONOMOU, G., LIMA, S. F., BICALHO, M. L. S., GANDA, E. K., DE OLIVEIRA FILHO, J. C., LORENZO, G., TROJACANEC, P. & BICALHO, R. C. 2015. Prepartum and Postpartum Rumen Fluid Microbiomes: Characterization and Correlation with Production Traits in Dairy Cows. *Applied and Environmental Microbiology,* 81**,** 1327-1337.
LOVE, W. J., LEHENBAUER, T. W., KASS, P. H., VAN EENENNAAM, A. L. & ALY, S. S. 2014. Development of a novel clinical scoring system for on-farm diagnosis of bovine respiratory disease in pre-weaned dairy calves. *PeerJ,* 2**,** e238.
MAKOSCHEY, B., MUÑOZ BIELSA, J., OLIVIERO, L., ROY, O., PILLET, F., VALLA, G. & CAVIRANI, S. 2008. Field efficacy of combination vaccines against bovine respiratory pathogens in calves. *Acta Veterinaria Hungarica,* 56**,** 485-493.
MANYAZEWAL, A., FRANCESCA, S., PAL, M., GEZAHEGN, M., TESFAYE, M., LUCY, M., TEKLU, W. & GETACHEW, T. 2018. Prevalence, risk factors and molecular characterization of Cryptosporidium infection in cattle in Addis Ababa and its environs, Ethiopia. *Veterinary parasitology, regional studies and reports,* 13**,** 79-84.
MAO, S., ZHANG, M., LIU, J. & ZHU, W. 2015. Characterising the bacterial microbiota across the gastrointestinal tracts of dairy cattle: membership and potential function. *Scientific reports,* 5**,** 16116.
MCDONALD, J., MCBRIDE, W. & O'DONOGHUE, E. 2007. Low costs drive production to large dairy farms.
MCGUIRK, S. M. 2008. Disease Management of Dairy Calves and Heifers. *Veterinary Clinics of North America: Food Animal Practice,* 24**,** 139-153.
MCNAUGHTON, L. & LOPDELL, T. Brief Communication: Are dairy heifers achieving live weight targets? Proceedings of the New Zealand Society of Animal Production, 2012. 120-122.
MCORIST, A. L., JACKSON, M. & BIRD, A. R. 2002. A comparison of five methods for extraction of bacterial DNA from human faecal samples. *Journal of Microbiological Methods,* 50**,** 131-139.
MEALE, S. J., LI, S. C., AZEVEDO, P., DERAKHSHANI, H., DEVRIES, T. J., PLAIZIER, J. C., STEELE, M. A. & KHAFIPOUR, E. 2017.

Weaning age influences the severity of gastrointestinal microbiome shifts in dairy calves. *Scientific Reports,* 7**,** 198.
MÉNDEZ-GARCÍA, C., BARGIELA, R., MARTÍNEZ-MARTÍNEZ, M. & FERRER, M. 2018. Metagenomic Protocols and Strategies. *Metagenomics.* Elsevier.
MENZEL, P., NG, K. L. & KROGH, A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications,* 7**,** 11257.
MILLER, R. H. & MCGILLIARD, L. 1959. Relations between weight at first calving and milk production during the first lactation. *Journal of Dairy Science,* 42**,** 1932-1943.
MIRSEPASI, H., PERSSON, S., STRUVE, C., ANDERSEN, L. O. B., PETERSEN, A. M. & KROGFELT, K. A. 2014. Microbial diversity in fecal samples depends on DNA extraction method: easyMag DNA extraction compared to QIAamp DNA stool mini kit extraction. *BMC Research Notes,* 7**,** 50.
MIRZAI, Y., YAKHCHALI, M. & MARDANI, K. 2014. Cryptosporidium parvum and Cryptosporidium andersoni infection in naturally infected cattle of northwest Iran. *Veterinary research forum : an international quarterly journal,* 5**,** 55-60.
MOSIER, D. A. & OBERST, R. D. 2000. Cryptosporidiosis: a global challenge. *Annals of the New York Academy of Sciences,* 916**,** 102-111.
NAGPAL, R., MAINALI, R., AHMADI, S., WANG, S., SINGH, R., KAVANAGH, K., KITZMAN, D. W., KUSHUGULOVA, A., MAROTTA, F. & YADAV, H. 2018. Gut microbiome and aging: Physiological and mechanistic insights. *Nutrition and healthy aging,* 4**,** 267-285.
NECHVATAL, J. M., RAM, J. L., BASSON, M. D., NAMPRACHAN, P., NIEC, S. R., BADSHA, K. Z., MATHERLY, L. H., MAJUMDAR, A. P. & KATO, I. 2008. Fecal collection, ambient preservation, and DNA extraction for PCR amplification of bacterial and human markers from human feces. *Journal of microbiological methods,* 72**,** 124-132.
NOORDHUIZEN, J., VAN EGMOND, M., JORRITSMA, R., HOGEVEEN, H., VAN WERVEN, T., VOS, P. & LIEVAART, J. 2008. Veterinary advice for entrepreneurial Dutch dairy farmers: from curative practice to coach-consultant: what needs to be changed? *Tijdschrift voor diergeneeskunde,* 133**,** 1-5.
OIKONOMOU, G., TEIXEIRA, A. G. V., FODITSCH, C., BICALHO, M. L., MACHADO, V. S. & BICALHO, R. C. 2013. Fecal Microbial Diversity in Pre-Weaned Dairy Calves as Described by Pyrosequencing of Metagenomic 16S rDNA. Associations of Faecalibacterium Species with Health and Growth. *PLOS ONE,* 8**,** e63157.
OTT, S. J., MUSFELDT, M., TIMMIS, K. N., HAMPE, J., WENDEROTH, D. F. & SCHREIBER, S. 2004. In vitro alterations of intestinal bacterial microbiota in fecal samples during storage. *Diagnostic microbiology and infectious disease,* 50**,** 237-245.
ÖZBAYRAM, G., INCE, O., INCE, B., HARMS, H. & KLEINSTEUBER, S. 2018. *Comparison of Rumen and Manure Microbiomes and Implications for the Inoculation of Anaerobic Digesters*.

PAUL, S., CHANDRA, D., TEWARI, A. K., BANERJEE, P. S., RAY, D. D., RAINA, O. K. & RAO, J. R. 2009. Prevalence of Cryptosporidium andersoni: A molecular epidemiological survey among cattle in India. *Veterinary Parasitology,* 161**,** 31-35.

PETROSINO, J. F., HIGHLANDER, S., LUNA, R. A., GIBBS, R. A. & VERSALOVIC, J. 2009. Metagenomic Pyrosequencing and Microbial Identification. *Clinical Chemistry,* 55**,** 856-866.

PLACE, N. T., HEINRICHS, A. J. & ERB, H. N. 1998. The Effects of Disease, Management, and Nutrition on Average Daily Gain of Dairy Heifers from Birth to Four Months. *Journal of Dairy Science,* 81**,** 1004-1009.

PRENDIVILLE, D. J., LOWE, J., EARLEY, B., SPAHR, C. & KETTLEWELL, P. 2002. Radiotelemetry systems for measuring body temperature.

QUAIL, M. A., SMITH, M., COUPLAND, P., OTTO, T. D., HARRIS, S. R., CONNOR, T. R., BERTONI, A., SWERDLOW, H. P. & GU, Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics,* 13**,** 341.

QUILEZ, J., SANCHEZ-ACEDO, C., DEL CACHO, E., CLAVEL, A. & CAUSAPE, A. 1996. Prevalence of Cryptosporidium and Giardia infections in cattle in Aragon (northeastern Spain). *Veterinary parasitology,* 66**,** 139-146.

RANJAN, R., RANI, A., METWALLY, A., MCGEE, H. S. & PERKINS, D. L. 2016. Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochemical and biophysical research communications,* 469**,** 967-977.

RAPP, D. 2010. DNA extraction from bovine faeces: current status and future trends. *Journal of Applied Microbiology,* 108**,** 1485-1493.

RAZAKANDRAINIBE, R., DIAWARA, E. H. I., COSTA, D., LE GOFF, L., LEMETEIL, D., BALLET, J. J., GARGALA, G. & FAVENNEC, L. 2018. Common occurrence of Cryptosporidium hominis in asymptomatic and symptomatic calves in France. *PLoS neglected tropical diseases,* 12**,** e0006355-e0006355.

REINHARDT, V. & REINHARDT, A. 1975. Dynamics of social hierarchy in a dairy herd. *Zeitschrift für Tierpsychologie,* 38**,** 315-323.

ROELOFS, J. B., VAN EERDENBURG, F. J. C. M., SOEDE, N. M. & KEMP, B. 2005. Pedometer readings for estrous detection and as predictor for time of ovulation in dairy cattle. *Theriogenology,* 64**,** 1690-1703.

ROSE-DYE, T. K., BURCIAGA-ROBLES, L. O., KREHBIEL, C. R., STEP, D. L., FULTON, R. W., CONFER, A. W. & RICHARDS, C. J. 2011. Rumen temperature change monitored with remote rumen temperature boluses after challenges with bovine viral diarrhea virus and Mannheimia haemolytica12. *Journal of Animal Science,* 89**,** 1193-1200.

RUTTEN, C., KAMPHUIS, C., HOGEVEEN, H., HUIJPS, K., NIELEN, M. & STEENEVELD, W. 2017. Sensor data on cow activity, rumination, and ear temperature improve prediction of the start of calving in dairy cows. *Computers and Electronics in Agriculture,* 132**,** 108-118.

SALONEN, A., NIKKILÄ, J., JALANKA-TUOVINEN, J., IMMONEN, O., RAJILIĆ-STOJANOVIĆ, M., KEKKONEN, R. A., PALVA, A. & DE VOS, W. M. 2010. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: Effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *Journal of Microbiological Methods,* 81**,** 127-134.

SANTıN, M., TROUT, J. M., XIAO, L., ZHOU, L., GREINER, E. & FAYER, R. 2004. Prevalence and age-related variation of Cryptosporidium species and genotypes in dairy calves. *Veterinary parasitology,* 122**,** 103-117.

SATO, S. 1984. Social licking pattern and its relationships to social dominance and live weight gain in weaned calves. *Applied Animal Behaviour Science,* 12**,** 25-32.

SHAHID, M., RENEAU, J., CHESTER-JONES, H., CHEBEL, R. & ENDRES, M. 2015. Cow-and herd-level risk factors for on-farm mortality in Midwest US dairy herds. *Journal of dairy science,* 98**,** 4401-4413.

SIGNORINI, M., SOTO, L., ZBRUN, M., SEQUEIRA, G., ROSMINI, M. & FRIZZO, L. 2012. Impact of probiotic administration on the health and fecal microbiota of young calves: a meta-analysis of randomized controlled trials of lactic acid bacteria. *Research in Veterinary Science,* 93**,** 250-258.

SIVULA, N., AMES, T., MARSH, W. & WERDIN, R. 1996. Descriptive epidemiology of morbidity and mortality in Minnesota dairy heifer calves. *Preventive Veterinary Medicine,* 27**,** 155-171.

SMITH, B., LI, N., ANDERSEN, A. S., SLOTVED, H. C. & KROGFELT, K. A. 2011. Optimising Bacterial DNA Extraction from Faecal Samples: Comparison of Three Methods. *The Open Microbiology Journal,* 5**,** 14-17.

SMITH, K. E., GARZA, A. L., BUTTERFIELD, K. M., DICKEY, A. M., LINDHOLM-PERRY, A. K., WELLS, J. E., FREETLY, H. C. & LODGE-IVEY, S. L. 2018. Succession of ruminal bacterial species and fermentation characteristics in preweaned Brangus calves1. *Translational Animal Science,* 2**,** S48-S52.

SMITH, R. P., CLIFTON-HADLEY, F. A., CHENEY, T. & GILES, M. 2014. Prevalence and molecular typing of Cryptosporidium in dairy cattle in England and Wales and examination of potential on-farm transmission routes. *Veterinary Parasitology,* 204**,** 111-119.

SOBERON, F., RAFFRENATO, E., EVERETT, R. W. & VAN AMBURGH, M. E. 2012. Preweaning milk replacer intake and effects on long-term productivity of dairy calves. *Journal of Dairy Science,* 95**,** 783-793.

STEENEVELD, W. & HOGEVEEN, H. 2015. Characterization of Dutch dairy farms using sensor systems for cow management. *Journal of dairy science,* 98**,** 709-717.

STEENEVELD, W., HOGEVEEN, H. & LANSINK, A. O. 2015. Economic consequences of investing in sensor systems on dairy farms. *Computers and Electronics in Agriculture,* 119**,** 33-39.

STEWART, R. D., AUFFRET, M. D., WARR, A., WISER, A. H., PRESS, M. O., LANGFORD, K. W., LIACHKO, I., SNELLING, T. J., DEWHURST, R. J., WALKER, A. W., ROEHE, R. & WATSON, M.

2018. Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen. *Nature Communications,* 9**,** 870.
STURDEE, A. P., BODLEY-TICKELL, A. T., ARCHER, A. & CHALMERS, R. M. 2003. Long-term study of Cryptosporidium prevalence on a lowland farm in the United Kingdom. *Veterinary Parasitology,* 116**,** 97-113.
SWIRE, J. 2017. *Remote weighing and temperature sensing to transform calf production* [Online]. Farm Business Available: http://www.farmbusiness.co.uk/livestock/beef/remote-weighing-and-temperature-sensing-to-transform-calf-production.html [Accessed 15/11/2019 2019].
THOMAS, M., WEBB, M., GHIMIRE, S., BLAIR, A., OLSON, K., FENSKE, G. J., FONDER, A. T., CHRISTOPHER-HENNINGS, J., BRAKE, D. & SCARIA, J. 2017. Metagenomic characterization of the effect of feed additives on the gut microbiome and antibiotic resistome of feedlot cattle. *Scientific Reports,* 7**,** 12257.
TIMSIT, E., ASSIÉ, S., QUINIOU, R., SEEGERS, H. & BAREILLE, N. 2011. Early detection of bovine respiratory disease in young bulls using reticulo-rumen temperature boluses. *The Veterinary Journal,* 190**,** 136-142.
TURNER, L., UDAL, M., LARSON, B. & SHEARER, S. 2000. Monitoring cattle behavior and pasture use with GPS and GIS. *Canadian Journal of Animal Science,* 80**,** 405-413.
VÁZQUEZ-DIOSDADO, J., MIGUEL-PACHECO, G., PLANT, B., DOTTORINI, T., GREEN, M. & KALER, J. 2019. Developing and evaluating threshold-based algorithms to detect drinking behavior in dairy cows using reticulorumen temperature. *Journal of dairy science,* 102**,** 10471-10482.
VLČKOVÁ, K., MRÁZEK, J., KOPEČNÝ, J. & PETRŽELKOVÁ, K. J. 2012. Evaluation of different storage methods to characterize the fecal bacterial communities of captive western lowland gorillas (Gorilla gorilla gorilla). *Journal of microbiological methods,* 91**,** 45-51.
VULEVIC, J., DRAKOULARAKOU, A., YAQOOB, P., TZORTZIS, G. & GIBSON, G. R. 2008. Modulation of the fecal microflora profile and immune function by a novel trans-galactooligosaccharide mixture (B-GOS) in healthy elderly volunteers. *Am J Clin Nutr,* 88**,** 1438-46.
WALTER, J. 2008. Ecological role of lactobacilli in the gastrointestinal tract: implications for fundamental and biomedical research. *Appl. Environ. Microbiol.,* 74**,** 4985-4996.
WALTNER-TOEWS, D., MARTIN, S. & MEEK, A. 1986. Dairy calf management, morbidity and mortality in Ontario Holstein herds. II. Age and seasonal patterns. *Preventive Veterinary Medicine,* 4**,** 125-135.
WALTON, E., CASEY, C., MITSCH, J., VÁZQUEZ-DIOSDADO, J. A., YAN, J., DOTTORINI, T., ELLIS, K. A., WINTERLICH, A. & KALER, J. 2018. Evaluation of sampling frequency, window size and sensor position for classification of sheep behaviour. *Royal Society open science,* 5**,** 171442.
WILSON, K. H., BLITCHINGTON, R. & GREENE, R. 1990. Amplification of bacterial 16S ribosomal DNA with polymerase chain reaction. *Journal of clinical microbiology,* 28**,** 1942-1946.

WOLF, R., BARKEMA, H., DE BUCK, J., SLOMP, M., FLAIG, J., HAUPSTEIN, D., PICKEL, C. & ORSEL, K. 2014. High herd-level prevalence of Mycobacterium avium subspecies paratuberculosis in Western Canadian dairy farms, based on environmental sampling. *Journal of dairy science,* 97**,** 6250-6259.

XIAO, L., ESCALANTE, L., YANG, C., SULAIMAN, I., ESCALANTE, A. A., MONTALI, R. J., FAYER, R. & LAL, A. A. 1999. Phylogenetic analysis of Cryptosporidiumparasites based on the small-subunit rRNA gene locus. *Appl. Environ. Microbiol.,* 65**,** 1578-1583.

YARZA, P., YILMAZ, P., PRUESSE, E., GLÖCKNER, F. O., LUDWIG, W., SCHLEIFER, K.-H., WHITMAN, W. B., EUZÉBY, J., AMANN, R. & ROSSELLÓ-MÓRA, R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology,* 12**,** 635.

YEOMAN, C. J., ISHAQ, S. L., BICHI, E., OLIVO, S. K., LOWE, J. & ALDRIDGE, B. M. 2018. Biogeographical Differences in the Influence of Maternal Microbial Sources on the Early Successional Development of the Bovine Neonatal Gastrointestinal tract. *Scientific reports,* 8**,** 3197.

ZHANG, B.-W., LI, M., MA, L.-C. & WEI, F.-W. 2006. A Widely Applicable Protocol for DNA Isolation from Fecal Samples. *Biochemical Genetics,* 44**,** 494.

ZHENG, H., POWELL, J. E., STEELE, M. I., DIETRICH, C. & MORAN, N. A. 2017. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proceedings of the National Academy of Sciences,* 114**,** 4775-4780.

ZHOU, Y., CHEN, Y., CHEN, S. & GU, J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics,* 34**,** i884-i890.

7. Appendices

7.1. Appendix 1

Calf	Calf ID	Age	Date	Nasal Discharge	Ocular Discharge	Coughing	Total Wisconsin Score	Weigh Band (cm)	Weight Estimate (kg)	Nasal Swab ID (A for old, D for young)	Rectal Swab ID (B for old, E for young)
1	7038	7-8 months	14.11.2018	1	1	0	2	149	264	1	1
1	7038	7-8 months	28.11.2018	0	1	0	1	150	272	22	22
1	7038	7-8 months	12.12.2018	0	0	0	0	155	303	43	43
2	7131	7-8 months	14.11.2018	1	0	0	1	152	280	2	2
2	7131	7-8 months	28.11.2018	2	0	0	2	152	280	23	23
2	7131	7-8 months	12.12.2018	1	0	0	1	153	290	44	44
3	7115	7-8 months	14.11.2018	0	0	0	0	152	280	3	3
3	7115	7-8 months	28.11.2018	0	0	0	0	151	276	24	24
3	7115	7-8 months	12.12.2018	1	0	0	1	157	314	45	45
4	7037	7-8 months	14.11.2018	1	0	0	1	145	248	4	4
4	7037	7-8 months	28.11.2018	1	0	0	1	148	257	25	25
4	7037	7-8 months	12.12.2018	0	0	0	0	154	296	46	46
5	7068	7-8 months	14.11.2018	2	2	0	4	155	303	5	5

5	7068	7-8 months	28.11.2018	1	0	0	1	160	330	26	26
5	7068	7-8 months	12.12.2018	0	0	0	0	164	348	47	47
6	7099	7-8 months	14.11.2018	2	0	0	2	142	243	6	6
6	7099	7-8 months	28.11.2018	0	0	0	0	150	272	27	27
6	7099	7-8 months	12.12.2018	1	0	0	1	146	249	48	48
7	7103	7-8 months	14.11.2018	1	0	0	1	148	257	7	7
7	7103	7-8 months	28.11.2018	1	0	0	1	149	264	28	28
7	7103	7-8 months	12.12.2018	1	0	0	1	160	330	49	49
8	7005	7-8 months	14.11.2018	1	0	0	1	148	257	8	8
8	7005	7-8 months	28.11.2018	1	0	0	1	152	280	29	29
8	7005	7-8 months	12.12.2018	0	0	0	0	155	303	50	50
9	7129	7-8 months	14.11.2018	0	0	0	0	135	210	9	9
9	7129	7-8 months	28.11.2018	0	0	0	0	139	230	30	30
9	7129	7-8 months	12.12.2018	0	0	0	0	143	246	51	51
10	7019	7-8 months	14.11.2018	0	1	0	1	148	257	10	10
10	7019	7-8 months	28.11.2018	2	0	0	2	148	257	31	31
10	7019	7-8 months	12.12.2018	0	0	0	0	157	314	52	52
11	7095	7-8 months	14.11.2018	0	0	0	0	149	264	11	11
11	7095	7-8 months	28.11.2018	0	0	0	0	153	290	32	32
11	7095	7-8 months	12.12.2018	0	0	0	0	156	308	53	53
12	7058	7-8 months	14.11.2018	0	2	0	2	156	308	12	12
12	7058	7-8 months	28.11.2018	1	1	0	2	158	320	33	33
12	7058	7-8 months	12.12.2018	1	0	0	1	165	364	54	54

13	7070	7-8 months	14.11.2018	1	0	0	1	152	280	13	13
13	7070	7-8 months	28.11.2018	0	0	0	0	152	280	34	34
13	7070	7-8 months	12.12.2018	1	0	0	1	156	308	55	55
14	7069	7-8 months	14.11.2018	0	1	0	1	137	220	14	14
14	7069	7-8 months	28.11.2018	1	0	0	1	141	240	35	35
14	7069	7-8 months	12.12.2018	1	0	0	1	143	246	56	56
15	7102	7-8 months	14.11.2018	1	0	1	2	134	206	15	15
15	7102	7-8 months	28.11.2018	NA	NA	NA	0	NA	NA	NA	NA
15	7102	7-8 months	12.12.2018	NA	NA	NA	0	NA	NA	NA	NA
16	7093	7-8 months	14.11.2018	1	1	0	2	146	249	16	16
16	7093	7-8 months	28.11.2018	1	0	0	1	150	272	37	37
16	7093	7-8 months	12.12.2018	0	0	0	0	155	303	58	58
17	7118	7-8 months	14.11.2018	1	1	1	3	144	247	17	17
17	7118	7-8 months	28.11.2018	0	1	0	1	148	257	38	38
17	7118	7-8 months	12.12.2018	0	1	0	1	150	272	59	59
18	7124	7-8 months	14.11.2018	0	2	1	3	138	225	18	18
18	7124	7-8 months	28.11.2018	0	0	0	0	150	272	39	39
18	7124	7-8 months	12.12.2018	0	0	0	0	150	272	60	60
19	7110	7-8 months	14.11.2018	1	1	0	2	151	276	19	19
19	7110	7-8 months	28.11.2018	1	1	0	2	156	308	40	40
19	7110	7-8 months	12.12.2018	0	0	0	0	161	335	61	61
20	6987	7-8 months	14.11.2018	2	1	0	3	145	248	20	20
20	6987	7-8 months	28.11.2018	1	0	0	1	150	272	41	41

20	6987	7-8 months	12.12.2018	0	0	0	0	154	296	62	62
21	7090	7-8 months	14.11.2018	1	0	0	1	151	276	21	21
21	7090	7-8 months	28.11.2018	1	1	0	2	156	308	42	42
21	7090	7-8 months	12.12.2018	0	0	0	0	159	325	63	63
22	7193	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
22	7193	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
22	7193	14 weeks	12.12.2018	1	2	0	3	129	186	1	1
23	7212	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
23	7212	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
23	7212	14 weeks	12.12.2018	0	0	0	0	127	178	2	2
24	7168	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
24	7168	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
24	7168	14 weeks	12.12.2018	0	0	0	0	133	202	3	3
25	7216	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
25	7216	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
25	7216	14 weeks	12.12.2018	0	1	0	1	124	166	4	4
26	7187	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
26	7187	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
26	7187	14 weeks	12.12.2018	0	2	0	2	130	190	5	5
27	7171	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
27	7171	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
27	7171	14 weeks	12.12.2018	0	0	0	0	139	230	6	6
28	7163	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA

28	7163	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
28	7163	14 weeks	12.12.2018	0	0	0	0	140	235	7	7
29	7234	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
29	7234	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
29	7234	14 weeks	12.12.2018	0	0	0	0	125	170	8	8
30	7201	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
30	7201	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
30	7201	14 weeks	12.12.2018	0	2	0	2	120	150	9	9
31	7167	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
31	7167	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
31	7167	14 weeks	12.12.2018	0	0	0	0	137	220	10	10
32	7213	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
32	7213	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
32	7213	14 weeks	12.12.2018	0	0	0	0	126	174	11	11
33	7189	14 weeks	14.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
33	7189	14 weeks	28.11.2018	NA	NA	NA	NA	NA	NA	NA	NA
33	7189	14 weeks	12.12.2018	0	0	0	0	128	182	12	12

Appendix 1: Table showing data collected during sampling