Detection of zoonotic bacterial pathogens in multiple hosts using a microbiome approach and molecular characterization of *Anaplasma phagocytophilum*

A thesis
submitted in fulfilment of the requirements for the degree Doctor Philosophiae
in the Department of Veterinary Tropical Diseases, Faculty of Veterinary Science,
University of Pretoria

By

AGATHA ONYEMOWO KOLO

SUPERVISOR:
PROF MARINDA OOSTHUIZEN

CO-SUPERVISORS:
PROF KELLY BRAYTON
DR NICOLA COLLINS

July 2019
Declaration

I hereby declare that this thesis which I submit for the degree Philosophiae Doctor at the Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria, South Africa is my original research work and has not been submitted by me to any institution or University for the conferment of an award.

____________________
Agatha Onyemowo Kolo
July 2019
Dedication

This thesis is dedicated to the memory of my Father, Sir Augustine Ola Abah who passed away to be with our Lord last year, thank you Daddy for all your selfless love and support, and all the sacrifice you went through to make my life better, thank you for believing in the education of the girl child. Thank you for instilling in me the belief that nothing is impossible, if you can dream it, you can achieve it!
Acknowledgements

I wish to extend my heartfelt gratitude and appreciation to the following people for the dedication, service and impact they have made upon my life and career during the course of this project.

First to God Almighty, my rock, my fortress, my help in time of need. You are the vine and I am your branch, cut off from you I wither and dry.

To my Husband Francis Babaman, the love of my life, the one who holds my heart, my greatest cheerleader, the General! Kolo thank you for loving me deeply, madly, passionately. Thank you for dedicated hours drawing maps upon maps on QGIS.

To my Kids, Daisy Kaka, Sophie Hauwa and Daniel Nmadu, my precious ones whom I adore, my not so little angels anymore thank you for loving mummy unconditionally and unreservedly and being a great blessing to my life.

To my Supervisor, Prof Marinda Oosthuizen, a true description of a mentor: kind, supportive and selfless who has guided and moulded me to be a better scientist, to think outside the box, I am so grateful for all the support you have rendered to me. I have really learnt a lot from you. When I grow up, I want to be just like you!

To my Co-supervisor, Prof Kelly Brayton, who challenges me to critical thinking whose wry humour brings out the best of every situation, whose kindness, assistance and support enveloped me during the course of this journey, thank you for the guidance.

To my Co-supervisor, Dr Nicola Collins, thank you for the being the thorough researcher that you are. For the expertise and keenness, you displayed during the duration of this project, I am deeply grateful.

To my darling Mom Christiana Ada Abah, my siblings: Angela Ogodo, Constance Onyego, Augustine Ola, James Ode, Francis Adah, Fatima Ene, John Ejembi and Emmanuella Anyaoyi, thank you for loving me from my beginning through thick and thin, through lack and plenty, through valleys and mountain tops, you were all there.

To my Mum-in-law Mrs Christiana Kaka Kolo, Brothers in law: Emmanuel, Moses, and Nathaniel Kolo, Sisters in law, nieces and nephews thank you for your support, Indeed Family is everything.
To the support staff at the Department of Veterinary Tropical Diseases, Mrs Milana Troskie, Ms Ilse Vorster and Ms Anna-Mari Bosman. Thank you for lending a helping hand anytime I popped by your office. To Dr Charles Byaruhanga and Dr Zama Khumalo thank you for your help in sample collection from rodents and all the adventures in the Kruger and Dr Charles for running models used for phylogenetic analysis.

To the University of Pretoria, for the PhD Research Support bursary and the travel abroad exchange bursary to undergo training at the Washington State University, Pullman, USA. The expertise I have garnered I will treasure for a lifetime.

To the South African National Research Foundation (NRF) (grants 92739, 110448 and 109350 to Marinda Oosthuizen), the University of Pretoria Institutional Research Theme on Animal and Zoonotic Diseases grant (awarded to Marinda Oosthuizen), and the Belgian Directorate General for Development Co-operation Framework (ITM/DGCD): Thank you for the research funds allocated to this project.

I acknowledge Dr Erich Zweygarth (Freie Universitat, Berlin, Germany) for providing control DNA, and Prof Sonja Matthee (Stellenbosch University, South Africa) and Prof Luis Neves (University of Pretoria) for expertise provided with wild rodent trapping. I acknowledge Ms. Jeanette Wentzel of the Hans Hoheisen Wildlife research station, University of Pretoria for the logistic support she rendered throughout my research visits to the Mnisi community area. I also acknowledge the technical assistance and support of Derek Pouchnik and Mark Wildung of the Genomics Sequencing core at Washington State University, Pullman, USA. I acknowledge Prof Lucille Blumberg, Prof John Frean and Dr Jennifer Rossouw of the National Institute of Communicable Diseases (NICD) Johannesburg, South Africa for the blood samples collected from febrile patients that was used in this study. I thank Prof Celia Abolnik of the Faculty of Veterinary Science, University of Pretoria for the use of the BSL-3 laboratory at the Poultry Research unit. I also acknowledge Prof Armanda Bastos of the Department of Zoology & Entomology, University of Pretoria for the molecular profiling of the rodents used in this study. I thank Prof Eric Etter of the Department of Production Animal Studies, Faculty of Veterinary Science, University of Pretoria for the assistance and insights he provided into the statistical analysis of microbiome data. I acknowledge the environmental monitors of the Mnisi community Programme University of Pretoria who assisted with sample and ectoparasite collection from dogs and cattle. I acknowledge Estelle Mayhew for the production of the map of the study area.

To the staff and all postgraduate students of the Department of Veterinary Tropical Diseases, thank you for the conducive environment, the support and the laughter you have provided during the years of my stay at the Department.
# Table of contents

Declaration........................................................................................................................................... i

Dedication............................................................................................................................................. ii

Acknowledgements .............................................................................................................................. iii

Table of contents ................................................................................................................................. v

List of Tables ........................................................................................................................................ xviii

List of Abbreviations ............................................................................................................................. xix

Thesis Summary ................................................................................................................................ xxi

CHAPTER 1 .............................................................................................................................................. 1

General Introduction ............................................................................................................................. 1

1.1 Introduction .................................................................................................................................. 1

1.2 Research Hypothesis ..................................................................................................................... 2

1.3 Research aims and objectives ......................................................................................................... 2

1.4 Thesis overview ............................................................................................................................. 3

1.5 References ..................................................................................................................................... 5

CHAPTER 2 .............................................................................................................................................. 7

Literature Review ................................................................................................................................. 7

2.1 Microbiome analysis ....................................................................................................................... 7

  2.1.1 Methods to study microbial communities .................................................................................. 7

  2.1.2 Statistical approaches in community analyses .......................................................................... 9

  2.1.3 Tick microbiome studies .......................................................................................................... 9

  2.1.4 Other microbiome studies ....................................................................................................... 10

2.2 Acute febrile illness .................................................................................................................... 11

  2.2.1 *Anaplasma phagocytophilum* ............................................................................................... 11

    2.2.1.1 Biology of *Anaplasma phagocytophilum* ....................................................................... 12

    2.2.1.2 Epidemiology of *Anaplasma phagocytophilum* ............................................................ 12
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>Discussion</td>
<td>56</td>
</tr>
<tr>
<td>3.6</td>
<td>References</td>
<td>62</td>
</tr>
<tr>
<td>4.1</td>
<td>Abstract</td>
<td>70</td>
</tr>
<tr>
<td>4.2</td>
<td>Introduction</td>
<td>71</td>
</tr>
<tr>
<td>4.3</td>
<td>Materials and Methods</td>
<td>71</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Ethics approval</td>
<td>71</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Samples used for the study</td>
<td>71</td>
</tr>
<tr>
<td>4.3.3</td>
<td>PCR amplification and sequencing</td>
<td>72</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Sequence analysis</td>
<td>73</td>
</tr>
<tr>
<td>4.4</td>
<td>Results</td>
<td>74</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Barcoded 16S rRNA gene amplification</td>
<td>74</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Sequence and statistical analysis</td>
<td>74</td>
</tr>
<tr>
<td>4.4.2.1</td>
<td>Dogs</td>
<td>74</td>
</tr>
<tr>
<td>4.4.2.2</td>
<td>Cattle</td>
<td>78</td>
</tr>
<tr>
<td>4.5</td>
<td>Discussion</td>
<td>81</td>
</tr>
<tr>
<td>4.6</td>
<td>References</td>
<td>86</td>
</tr>
<tr>
<td>5.1</td>
<td>Abstract</td>
<td>94</td>
</tr>
<tr>
<td>5.2</td>
<td>Introduction</td>
<td>94</td>
</tr>
<tr>
<td>5.3</td>
<td>Materials and Methods</td>
<td>95</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Ethics approval</td>
<td>95</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Study site and sample collection</td>
<td>96</td>
</tr>
<tr>
<td>5.3.3</td>
<td>DNA extraction and 16S rRNA gene amplification</td>
<td>97</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Sequence analysis of NGS data</td>
<td>98</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Statistical analysis</td>
<td>99</td>
</tr>
</tbody>
</table>
# 7.3 Methods

- 7.3.1 Ethics approval ........................................................................................................140
- 7.3.2 Study area ..................................................................................................................140
- 7.3.3 Collection of blood samples ......................................................................................141
- 7.3.4 Collection of ticks ......................................................................................................142
- 7.3.5 DNA extraction, quantitative real-time PCR and assay specificity .........................142
- 7.3.6 Characterization of *A. phagocytophilum* by multi-locus gene sequencing ..............143
- 7.3.7 Sequence and phylogenetic analysis ........................................................................144

# 7.4 Results ..........................................................................................................................144

- 7.4.1 Specificity of the qPCR assay ..................................................................................144
- 7.4.2 *Anaplasma phagocytophilum* and/or *Anaplasma* sp. ZAM dog occurred in all hosts tested .145
- 7.4.3 PCR assays for sequencing ......................................................................................146
- 7.4.4 Sequence analysis of the 16S rRNA, *gltA*, *msp4* and *ankA* genes .......................148
  - 7.4.4.1 16S rRNA ...........................................................................................................149
  - 7.4.4.2 GltA ......................................................................................................................151
  - 7.4.4.3 Msp4 ...................................................................................................................152
  - 7.4.4.4 AnkA ...................................................................................................................152
- 7.4.5 Phylogenetic analyses ...............................................................................................153

# 7.5 Discussion .....................................................................................................................155

# 7.6 References ....................................................................................................................159

# CHAPTER 8

General Discussion, Recommendations and Conclusions .................................................164

# 8.1 References .....................................................................................................................172

# Appendix 1 .........................................................................................................................178

# Appendix 2 ........................................................................................................................182
List of Figures

Figure 3-1: Map of the Mnisi community study area in Bushbuckridge Municipality Mpumalanga Province, South Africa. Red stars show sites of rodent capture. Gottenburg and Hlalakahle are urban or periurban areas, while Tlhavekisa is located in communal rangelands. Manyeleti is a protected wildlife reserve. Dark green represents protected areas where wildlife roam freely (map produced by Estelle Mayhew). ................................................................. 45

Figure 3-2: Example of 16S rRNA gene amplicons obtained from rodent samples. PCR amplicons of the 16S rRNA gene were amplified using barcoded Primers 27F and 1435R and separated on a 1.5% agarose gel (stained with ethidium bromide). Lanes c-q contain rodent PCR samples. Lane b contains the positive control *Anaplasma centrale* and lane r contains the negative water only control. Lane a contains the 1kb DNA molecular ladder (Invitrogen, ThermoFisher® Scientific, South Africa) with sizes in bp indicated on the left. ................................................................. 49

Figure 3-3: Rarefaction curves of rodents from the Mnisi community. The mean species diversity of bacterial populations in rodent blood were plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied. ................................................................. 49

Figure 3-4: Relative abundance of major taxa of bacteria in the blood of 25 *Mastomys* spp. rodents. Six to seven rodents from each habitat area were sampled. The village name is indicated under the rodent number. ........................................................................................................ 50

Figure 3-5: Relative abundance of major taxa of bacteria in the blood of *Mastomys* spp. from different habitat areas: Hlalakahle and Gottenburg (urban/periurban area), Tlhavekisa (communal rangeland), and Manyeleti Game Reserve (protected wild reserve). ........................................................................ 51

Figure 3-6: PCA plot of the individuals factor map of rodents from the Mnisi communal area and the nearby wildlife area of Manyeleti. Sample ids in black represent samples from Gottenburg, samples from Hlalakahle are in red, samples from Manyeleti are in green and samples from Tlhavekisa are in blue. Outliers R21, R53 and R159 are from Manyeleti Game Reserve (protected reserve), while R19, R30 and R31 are from Gottenburg (periurban area). Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 32.86% for the first axis and 25.49% for the second axis. ......................... 52

Figure 3-7: PCA plot of the variables factor i.e. bacterial populations from the blood of rodents in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear
combinations of the variables. The percentage of variability for each dimension is given 32.86% for the first axis and 25.49% for the second axis.

**Figure 3-8:** Maximum likelihood tree based on the Jukes-Cantor model showing the phylogenetic relationship between 16S rRNA gene sequences of previously identified *Bartonella* spp. and *Bartonella* spp. identified in the blood of rodents from different habitat areas in the Mnisi community. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 22 nucleotide sequences. There were 1285 positions in the final dataset.

**Figure 3-9:** *Brucella*-specific amplification from rodents. Photograph of a 2% agarose gel (stained with ethidium bromide) of PCR products of the ITS region amplified using primers ITS66 and ITS279. Lanes d and e show amplification products from rodent DNA (R30 and R31). Lane h contained the positive control *B. abortus* strain 544 while lane i contained the negative control. Lane a and j contain the 100 bp DNA molecular ladder with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa). The other lanes contain samples from another study.

**Figure 4-1:** Map of study area with red stars showing locations of blood sampling from dogs and cattle. Dark green represents protected areas where wildlife roam freely (map produced by Estelle Mayhew).

**Figure 4-2:** Example of 16S rRNA gene amplicons from dog and cattle samples. Photograph of a 1.5% agarose gel (stained with ethidium bromide) showing 16S rRNA gene PCR amplicons. Lanes c-u contain dog and cattle PCR samples. Lane b contains the positive control (*A. centrale*) and lane v contains the negative water only control. Lane a and w contain the 1kb DNA molecular ladder, with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

**Figure 4-3:** Rarefaction curves of dogs from the Mnisi community. The mean species diversity of bacterial populations in canine blood were plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied.

**Figure 4-4:** Relative abundance of major taxa of bacteria in the blood of 10 domestic dogs in the Mnisi community.

**Figure 4-5:** PCA plot of individuals factor map of dogs in the Mnisi community. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 29.90% for the first axis and 25.19% for the second axis.
**Figure 4-6:** PCA plot of the variables factor i.e. bacterial populations from the blood of dogs in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 29.90% for the first axis and 25.19% for the second axis.

**Figure 4-7:** Rarefaction curves of species diversity in Mnisi cattle. The mean species diversity of bacterial populations in cattle blood was plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied.

**Figure 4-8:** Relative abundance of major taxa of bacteria in the blood of nine cattle in the Mnisi community.

**Figure 4-9:** PCA plot of the individuals factor map of cattle in the Mnisi community. Sample id in black font represents a sample from Gottenburg, samples from Hlalakahle are in red, samples from Seville A are in green, a sample from Seville B is in blue and samples from Tlhavekisa are in turquoise. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 42.93% for the first axis and 24.20% for the second axis.

**Figure 4-10:** PCA plot of the variables factor i.e. bacterial populations from the blood of cattle in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 42.93% for the first axis and 24.20% for the second axis.

**Figure 5-1:** Map of study site with red stars indicating communal dip tanks where *A. hebraeum* was collected from cattle. Dark green indicates game reserves where wildlife roam freely (map produced by Estelle Mayhew).

**Figure 5-2:** Example of 16S rRNA gene amplicons from *A. hebraeum* salivary gland and midgut pools. Photograph of a 1.5% agarose gel (stained with ethidium bromide) of amplicons. Lanes b-l contain *A. hebraeum* PCR samples. Lane m contains the positive control (*A. centrale*) and lane n contains the negative water only control. Lane a and o contain the 1kb DNA molecular ladder with sizes in bp indicated on the left (Invitrogen, ThermoFisher® Scientific, South Africa).

**Figure 5-3:** Rarefaction curve of *A. hebraeum* salivary gland pools from the Mnisi community. The mean species diversity of bacterial populations in salivary gland pools was plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied.
Figure 5-4: Relative abundance of important taxa of bacteria in the salivary glands of A. hebraeum ticks collected from dip tanks in the Mnisi community. 1 = Clare A, 2 = Welverdiend A, 3 = Shorty, 4 = Welverdiend B, 5 = Eglington, 6 = Gottenburg, 7 = Utha, 8 = Seville B.

Figure 5-5: Relative abundance of important taxa of bacteria in the salivary glands of A. hebraeum from the dip tanks of origin in the Mnisi community.

Figure 5-6: PCA plot of the individuals factor map of the salivary gland pools of A. hebraeum collected from cattle in the Mnisi community. Sample IDs in black represent samples from Clare A, samples from Eglington are in red, green samples are from Gottenburg, blue samples from Seville B, samples in turquoise are from Shorty, samples in purple are from Utha, samples in grey are from Welverdiend A and samples in brown are from Welverdiend B. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 41.07% for the first axis and 33.51% for the second axis.

Figure 5-7: PCA plot of the variables factor i.e. bacterial populations from the salivary gland pools of A. hebraeum in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 41.07% for the first axis and 33.51% for the second axis.

Figure 5-8: Rarefaction curve of A. hebraeum midgut pools from the Mnisi community. The mean species diversity of bacterial populations in midgut pools was plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied.

Figure 5-9: Relative abundance of important taxa of bacteria in the midgut pools of A. hebraeum ticks collected from dip tanks in the Mnisi community. 1 = Clare A, 2 = Welverdiend A, 3 = Shorty, 4 = Welverdiend B, 5 = Eglington, 6 = Gottenburg, 7 = Utha, 8 = Seville B.

Figure 5-10: Relative abundance of important taxa of bacteria in the midgut pools of A. hebraeum from the dip tank of origin in the Mnisi community.

Figure 5-11: PCA plot of the individuals factor map of the midgut pools of A. hebraeum collected from cattle in the Mnisi community. Sample IDs in black represent samples from Clare A, samples from Eglington are in red, green samples are from Gottenburg, blue samples from Seville B, samples in turquoise are from Shorty, samples in purple are from Utha, samples in grey are from Welverdiend A and samples in brown are from Welverdiend B. Dim 1 and 2 are dimensions or principal components.
produced from the linear combinations of the variables. The percentage of variability for each dimension is given 46.68% for the first axis and 21.31% for the second axis.

**Figure 5-12:** PCA plot of the variables factor map i.e. bacterial populations from the midgut pools of A. hebraeum in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 46.48% for the first axis and 21.31% for the second axis.

**Figure 5-13:** Maximum likelihood tree of 16S rRNA gene sequences showing the phylogenetic relationship between *Rickettsia* species. *Ehrlichia chaffeensis* was used as the outgroup sequence. Representative *R. africae* sequences generated in this study are in bold. The percentage of the bootstrap (1000 replicates) are shown next to the branches. All positions containing gaps and missing data were eliminated. Evolutionary distances based on the Jukes Cantor model were computed in MEGA 7. *M=Midgut pool; S= Salivary gland pool.*

**Figure 5-14:** Representative amplification plot of *Rickettsia* genus-specific qPCR assay targeting the citrate synthase gene.

**Figure 6-1:** Example of 16S rRNA gene amplicons from AFI patients. Photograph of a 1.5% agarose gel (stained with ethidium bromide) of amplicons. Lanes b-m contain human samples. Lane n contains the positive control *Anaplasma centrale* and lane o contains the negative water only control. Lane a and p contain the 1kb DNA molecular ladder, with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

**Figure 6-2:** Rarefaction curves of AFI patients from the Mnisi community. The mean species diversity of bacterial populations in human blood were plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied.

**Figure 6-3:** Relative abundance of major bacterial taxa in the blood of AFI patients from the Mnisi community. *H=AFI patient.*

**Figure 6-4:** PCA plot of the individuals factor map of AFI patients in the Mnisi community. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. Patients from Gottenburg clinic are represented in black, patients from Utha clinic are in red and patients from Welverdiend clinic are in green. The percentage of variability for each dimension is given 59.81% for the first axis and 16.64% for the second axis. H= AFI patient.
Figure 6-5: PCA plot of the variables factor i.e bacterial populations from the blood of AFI patients in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 59.81% for the first axis and 16.64% for the second axis.

Figure 6-6: Maximum likelihood tree of 16S rRNA gene sequences showing the phylogenetic relationship between *Rickettsia* species. *Ehrlichia chaffeensis* was used as the outgroup sequence. Representative *R. africae* sequences from AFI patients generated in this study are in bold. The percentage of the bootstrap (1000 replicates) are shown next to the branches. All positions containing gaps and missing data were eliminated. Evolutionary distances based on the Jukes Cantor model were computed in MEGA 7. H = AFI patient.

Figure 7-1: Map of the study area with red stars indicating locations of sample collection. Dark green indicates protected areas where wildlife roam freely (map produced by Estelle Mayhew).

Figure 7-2: Amplification plot of qPCR assay showing cross-reaction observed with *Anaplasma* sp. ZAM dog DNA. The two curves in between the positive and negative controls indicate amplification observed from *Anaplasma* sp. ZAM dog DNA. DNA from *A. phagocytophilum* dog strain L610 was used as the positive control while PCR grade water was used as the negative control.

Figure 7-3: Occurrence of *A. phagocytophilum* or *Anaplasma* sp. ZAM dog in hosts and ticks as determined by the qPCR assay.

Figure 7-4: Representative amplification plot of qPCR assay targeting the *msp2* gene. DNA from *A. phagocytophilum* dog strain L610 was used as the positive control while PCR grade water was used as the negative control.

Figure 7-5: Gel picture of a 1.5% agarose gel (stained with ethidium bromide) of colony PCR products of the 16S rRNA gene. Primers pJET1.2-F and pJET1.2-R were used for the colony PCR. Lanes 2-22 contain colony PCR products. Lane 23 contained the no template negative control. Empty lanes with no visible bands indicate plasmids that did not have an insert. Lane 1 and 24 contained the 1kb DNA ladder with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

Figure 7-6: Gel electrophoresis picture of a 1.5% agarose gel (stained with ethidium bromide) of nested PCR products of the 16S rRNA gene from positive samples on the qPCR assay. Primers Ap16S-F and Ap16S-R were used for the nested PCR. Lanes 2-21 contained PCR samples. Lanes 2, 11-16 and 18 show non-specific amplification bands. Lane 20 contained the positive control *A. phagocytophilum*
strain L610 and lane 21 contained the negative water only control. Lanes 1 and 22 contain the 1kb DNA ladder with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

Figure 7-7: Gel electrophoresis picture of a 1.5% agarose gel (stained with ethidium bromide) of nested PCR products of the *gltA* gene from positive samples on the qPCR assay. Primers F4B and HG1085R were used for the PCR. Lanes 2-14 contain nested PCR products. Lane 13 contained the positive control *A. phagocytophilum* strain L610 while lane 14 contained the negative control. Lanes 1 and 15 contain the 1kb DNA ladder with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

Figure 7-8: Gel electrophoresis picture of a 1.5% agarose gel (stained with ethidium bromide) of nested PCR products of the *msp4* gene from positive samples on the qPCR assay. Primers AB1692-F and AB1693-R were used for the PCR. Lanes 2 to 22 contained nested PCR samples. Lane 21 contained the positive control *A. phagocytophilum* strain L610 while lane 2 contained the no template negative control. Lane 1 contained the 100 bp DNA ladder with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

Figure 7-9: Gel electrophoresis picture of a 1.5% agarose gel (stained with ethidium bromide) of nested PCR products of the *ankA* gene from positive samples on the qPCR assay. Primers LA6-F and LA1-R were used for the PCR. Lanes 2 to 12 contained nested PCR products. Lane marked 11 contained the positive control *A. phagocytophilum* strain L610 while lane marked 12 contained the no template negative control. Lanes 1 and 13 contain the 100 bp DNA ladder with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

Figure 7-10: Maximum-likelihood phylogenetic tree based on 16S rRNA nucleotide sequences. The tree shows the phylogenetic relationship between the 16S rRNA gene sequence variants obtained in the study and related *Anaplasma* species. The 16S RNA variants were designated as *A. phagocytophilum* Aph1 and Aph2, *Anaplasma* sp. ZAM dog Adog1 and Adog2, *A. platys* Apla1, *Candidatus* Anaplasma boleense Cab1 and *Anaplasma* sp. Mymensingh Asm1. Sequence accession numbers are shown in parentheses. The evolutionary history was inferred by using the Maximum Likelihood method based on the General Time Reversible model (Nei and Kumar, 2000). The tree with the highest log likelihood (-3074.44) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances.
estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.2734)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 39.92% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 31 nucleotide sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There was a total of 1191 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016).

Figure 7-11: Maximum likelihood phylogenetic tree based on GltA deduced amino acid sequences showing the phylogenetic relationship between the obtained Anaplasma GltA sequence variants and related Anaplasma species. The GltA variants were designated as A. phagocytophilum Aph1 and Anaplasma sp. ZAM dog Adog1/GltA. Sequence accession numbers are shown in parentheses. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood (-657.40) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1.7688)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 22.99% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There was a total of 66 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016).
List of Tables

Table 2-1: Studies that have reported the detection of *A. phagocytophilum* in Africa using nucleic acid-based detection methods. ................................................................. 14

Table 3-1: Barcoded primers used for the amplification of the 16S rRNA gene. ................................. 46

Table 3-2: Sample information and primer combinations used to construct the library for circular consensus sequencing (CCS). ................................................................. 47

Table 3-3: Nucleotide positions showing differences discriminating between *Brucella* and *Ochrobactrum* species in an alignment with reference sequences from GenBank. Accession numbers of reference sequences are enclosed in brackets. ................................................................. 56

Table 4-1: Sample and primer information on the library prepared for CCS. ..................................... 73

Table 5-1: Sample and primer information of PCR amplicons for library 1. ..................................... 97

Table 5-2: Sample and primer information for PCR amplicons for library 2. .................................... 98

Table 6-1: Sample and primer information for PCR amplicons from humans for library. .................. 123

Table 7-1: Origin and sample sizes of the specimens used in the study. ......................................... 142

Table 7-2: Primers used for amplification of the *A. phagocytophilum* DNA. ................................. 144

Table 7-3: *Anaplasma* species 16S rRNA, GltA, Msp4 and AnkA sequence variants circulating in the Mnisi community. ................................................................. 151
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S rRNA</td>
<td>16S ribosomal ribonucleic acid</td>
</tr>
<tr>
<td>AFI</td>
<td>Acute febrile illness</td>
</tr>
<tr>
<td>AG</td>
<td>Ancestral group</td>
</tr>
<tr>
<td>ATBF</td>
<td>African tick bite fever</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search tool</td>
</tr>
<tr>
<td>CCS</td>
<td>Circular consensus sequencing</td>
</tr>
<tr>
<td>CGA</td>
<td>Canine granulocytic anaplasmosis</td>
</tr>
<tr>
<td>CME</td>
<td>Canine monocytic ehrlichiosis</td>
</tr>
<tr>
<td>ºC</td>
<td>Degrees Celsius</td>
</tr>
<tr>
<td>dATP</td>
<td>Deoxyadenosine triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>Deoxycytidine triphosphate</td>
</tr>
<tr>
<td>dGTP</td>
<td>Deoxyguanosine triphosphate</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dUTP</td>
<td>Deoxyuridine triphosphate</td>
</tr>
<tr>
<td>EGA</td>
<td>Equine granulocytic anaplasmosis</td>
</tr>
<tr>
<td>Eid</td>
<td>Emerging infectious Diseases</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>FBSF</td>
<td>Flea-borne spotted fever</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrogen chloride</td>
</tr>
<tr>
<td>HDSS</td>
<td>Health and demographic surveillance study</td>
</tr>
<tr>
<td>HGE</td>
<td>Human granulocytic ehrlichiosis</td>
</tr>
<tr>
<td>IFA</td>
<td>Immuno fluorescent assay</td>
</tr>
<tr>
<td>IFAT</td>
<td>Indirect fluorescent antibody test</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G (Gamma)</td>
</tr>
<tr>
<td>IgM</td>
<td>Immunoglobulin M (Mu)</td>
</tr>
<tr>
<td>µl</td>
<td>Microlitre</td>
</tr>
<tr>
<td>µM</td>
<td>Micromolar</td>
</tr>
</tbody>
</table>
min
ml
mm
mM
MEGA
MgCl$_2$
MIF
MSF
NGS
PCR
qPCR
SFG
TBF
TG
TRG
UDG
USA

Minute
Millilitre
Millimetre
Millimolar
Molecular evolutionary genetics analysis
Magnesium chloride
Microimmunofluorescence
Mediterranean spotted fever
Next-generation sequencing
Polymerase chain reaction
Quantitative polymerase chain reaction
Spotted fever group
Tick-borne fever
Typhus group
Transitional group
Uracil DNA glycosylase
United States of America
Detection of zoonotic bacterial pathogens in multiple hosts using a microbiome approach and molecular characterization of *Anaplasma phagocytophilum*

**Candidate:** Dr Agatha Onyemowo Kolo  
**Supervisor:** Prof Marinda Oosthuizen  
**Co-supervisors:** Prof Kelly Brayton (WSU) and Dr Nicola Collins (UP)  
**Department:** Veterinary Tropical Diseases  
**Degree:** PhD

**Thesis Summary**

The Mnisi community, a rural area, is nestled at the cusp of a human/livestock/wildlife interface in Bushbuckridge Municipality, Mpumalanga Province, South Africa. In this area, undifferentiated non-malarial acute febrile illness (AFI) is among the most common presenting signs in patients seeking healthcare at the community clinics. Recent research suggested that zoonotic pathogens either rodent-borne or tick-borne may be common aetiologies of febrile illness in the community. The study had shown that patients presenting with non-malarial AFI had prior exposure to *Bartonella* spp., spotted fever group *Rickettsia*, *Coxiella burnetii* and *Leptospira* spp. Low levels of West Nile and Sindbis, but no Rift Valley fever virus exposure were found. In a separate study, the molecular detection of a bacterium closely related to *Anaplasma phagocytophilum* in dog samples collected in the Mnisi community was also reported. The aim of this study was, therefore to investigate wild rodents, cattle and dogs as well as their associated ticks, as possible sources of zoonotic pathogen infection in the Mnisi community using a microbiome sequencing approach. We also screened AFI patient samples, rodents, dogs, cattle and ticks for the presence of *A. phagocytophilum* using a real-time PCR assay. The *Anaplasma* species detected were subsequently characterized using gene sequencing and phylogenetic analysis.

The sample set consisted of 282 wild rodents trapped across three habitat types, 74 AFI patients, 56 domestic dogs, 100 cattle, 160 *Rhipicephalus sanguineus* ticks collected from dogs and 348 *Amblyomma hebraeum* ticks collected from cattle. Of these, the bacterial blood microbiome of a subset of samples was generated using circular consensus sequencing (CCS) performed on the Pacific Biosciences platform at Washington State University. The full sample set was then also screened for the specific presence of *A. phagocytophilum* using a Taqman real-time PCR assay, followed by the molecular characterization and phylogenetic analysis of *A. phagocytophilum* targeting the 16S rRNA, *gltA, ankA* and *msp4* genes.

The bacterial blood microbiome of 25 *Mastomys* rodent species collected from three habitat areas revealed a total of 65,060 bacterial sequences with 29% of the total sequence reads obtained
corresponding to Bartonella grahamii, 23% to Bartonella sp. strain RF255YX and 12% of sequences to Bartonella spp. Overall, rodents from Hlalakahle (urban/periurban area) and Thhavekisa (communal rangeland) had higher proportions of Bartonella species (~85%), while Gottenburg (urban/periurban area) and Manyeleti (protected area) had lower Bartonella spp. loads (~45%). Other organisms of zoonotic and veterinary significance detected included Ehrlichia sp. (~0.03%), C. burnetii (~0.02%), Anaplasma spp. (~0.5%), and Brucella spp. (~1%).

Characterization of the blood microbiome of the dogs revealed 30,340 bacterial sequences, with 24% of the total sequences obtained from canine blood corresponding to Ehrlichia canis, 19.3% to A. platys, 14.8% to Anaplasma sp. ZAM dog, while 0.3% of sequences corresponded to A. phagocytophilum. The characterization of the blood microbiome from nine cattle revealed 34,559 bacterial sequences, of which 58% corresponded to A. marginale, 22.2% to Anaplasma sp. Mymensingh, 10.5% to Anaplasma spp. and 5.4% to Anaplasma sp. Dedessa. In addition, these species were detected in the following rates in cattle blood: Anaplasma sp. Hadesa: 2.7%, A. centrale: 1.4%, Bartonella spp.: 0.5%, A. platys: 0.2%, Anaplasma sp. Saso: 0.2% and A. phagocytophilum: 0.01%.

Characterization of the bacterial microbiome from 24 pools of salivary glands and 23 pools of midgut tissues from the 348 A. hebraeum ticks produced a total of 86,308 bacterial sequence reads from the midgut pools as well as 84,420 sequences from the salivary gland pools. Of these, 81.7% of the bacterial sequences from the midgut pools and 83% of the sequences from the salivary gland pools were dominated by the zoonotic pathogen Rickettsia africae, the cause of African tick bite fever (ATBF). A further 6.8% of the sequences from the salivary gland pools and 6.9% of the sequences from the midgut pools corresponded to Rickettsia spp. Six percent of the total sequences from the salivary gland pools and 3.4% of the sequences from the midgut pools corresponded to E. ruminantium, while 1.4% of the sequence reads from the salivary gland pools and 1.2% of the reads from the midgut pools corresponded to Coxiella spp. symbionts. Characterization of the blood microbiome of nine AFI patients revealed 13,726 bacterial sequences. Of significance was the detection of R. africae from three AFI patients and Brucella melitensis from one AFI patient.

DNA from 74 non-malarial acute febrile illness patients, 282 rodents, 100 cattle, 56 dogs, and 160 R. sanguineus ticks were screened using a quantitative real-time PCR designed to target the msp2 gene of A. phagocytophilum. However, the test was found to detect A. phagocytophilum and an Anaplasma sp. recently described from Zambian dogs (Anaplasma sp. ZAM dog). Sequencing of the 16S rRNA and gltA genes confirmed the presence of A. phagocytophilum DNA in humans, dogs and rodents; also highlighting its potential importance as a possible contributing cause of acute febrile illness in humans in this rural community in South Africa. Anaplasma sp. ZAM dog and Anaplasma platys were furthermore identified in dogs, while Candidatus Anaplasma boleense and Anaplasma sp. Mymensingh were identified in cattle. Anaplasma sp. ZAM dog was also identified in R. sanguineus ticks collected from dogs. Phylogenetic analyses grouped Anaplasma sp. ZAM dog into a distinct clade; providing sufficient divergence with the other Anaplasma species to warrant classification as a separate species.
Until appropriate type-material can be deposited and the species can be formally described, we will refer to this novel organism as *Anaplasma* sp. SA dog (for *Anaplasma* sp. southern Africa dog).

In conclusion, the detection of an array of zoonotic bacterial pathogens from wild rodents, domestic dogs, cattle and their associated ticks and humans in this study highlights their significance as possible contributing factors to non-malarial febrile illness in the Mnisi community area. We therefore recommend that healthcare practitioners in the community should consider them in the differential diagnosis of AFI.
Chapter 1
General Introduction

1.1 Introduction

Zoonotic tick-borne diseases have been described as models for emerging diseases worldwide (Raoult and Roux, 1997). It is documented that ticks are ranked second only to mosquitoes in the number of infectious diseases they transmit to humans globally (Parola and Raoult, 2001). It is further estimated that more than six out of ten infectious diseases in humans are spread from animals, often by an arthropod vector (Jongejan and Uilenberg, 2004). In recent years, vector-borne viral, bacterial and parasitic diseases have emerged or re-emerged in many regions, affecting global health and presenting economic concerns (Colwell et al., 2011). A study in Europe has found an increasing number of vector-borne infections in companion animals in recent years (Beugnet and Marié, 2009). Another study found 22.8% of all emerging infectious diseases (EID) to be caused by vector-borne pathogens (Jones et al., 2008). The ecology and epidemiology of these diseases are affected by the interrelations between the pathogen, the host (human, animal or vector) and the environment (Colwell et al., 2011, Parola and Raoult, 2001). Vector-borne infections transmitted from wildlife reservoir hosts to domestic animals or humans are thus of major concern.

In sub-Saharan Africa, a study carried out on patients with febrile illness in East Africa found that malaria was generally over diagnosed and there was a need to improve the reliability of diagnosis to differentiate between true cases of malaria and cases of undifferentiated non-malarial acute febrile illness (AFI) (Amexo et al., 2004, Reyburn et al., 2004). AFI is defined as a rapid fever without an obvious cause, and it poses a challenge in developing countries, as health care workers do not know how best to proceed and treat such cases (Capeding et al., 2013). A study conducted in Kenya found the rates of AFI were grossly underestimated in hospital settings (Feikin et al., 2011). Tick-borne bacterial zoonoses are frequently implicated in cases of AFI, but very little is known about these zoonoses in Africa (Hotez and Kamath, 2009). Studies have thus called for improved diagnostics, pathogen surveillance, and the prompt detection of emerging pathogens (Wolfe et al., 2007).

In Southern Africa, there is a paucity of information on causes of febrile illness in low income communities (Prasad et al., 2015). The Mnisi community, an area of high rural poverty, is situated at the human/livestock/wildlife interface in the northeastern corner of the Bushbuckridge Municipality in Mpumalanga, South Africa and is classified as one of South Africa’s 14 rural poverty nodes. The community is made up of an estimated 29,500 hectares of land (Berrian et al., 2016) and shares 75% of its boundary with wildlife reserves, which are open to the Kruger National Park and the Mpumalanga Tourism and Parks Agency (MTPA) game reserves (Manyeleti and Andover) and private game reserves (Sabi Sands and Timbavati). It is estimated about 40,000 people live in Mnisi in 8,500 households (Statistics, 2012). In the Mnisi community, undifferentiated non-malarial AFI is among the most common
presenting sign in patients seeking healthcare at the community clinics (Quan et al., 2014, Simpson et al., 2018). A study on the one health profile of the community found 72% of all households reported owning an animal, with 68% of households owning livestock (Berrian et al., 2016). From the number of households owning animals: 55% of households owned chickens, 31% dogs, 25% cattle, 16% goats, 9% cats, pigs were owned by 5% of households and less than 1% of the households sampled owned ducks and pigeons (Berrian et al., 2016). Recent research suggests that rodent and vector-borne zoonoses may be common causes of febrile illness in the Mnisi community (Quan et al., 2014). Partial 16S rRNA gene sequences closely related to the zoonotic tick-borne rickettsial pathogen *Anaplasma phagocytophilum* have been detected in domestic dogs in the area and *Rickettsia africae* was found in ticks collected from dogs (Kolo et al., 2016). The aforementioned study in the community (Berrian et al., 2016) also showed that 76% of households sampled reported seeing rodents in and/or around their home. Out of that number, 62% of the respondents said they saw them daily (Berrian et al., 2016). A recent serological survey carried out on non-malarial AFI patients in the area found 63.4% of patients had prior exposure to spotted fever group (SFG) *Rickettsia* spp., 38.3% to *Coxiella burnetii*, the cause of Q fever, while another 9.5% were PCR-positive to zoonotic rodent-borne *Bartonella* species (Simpson et al., 2018). The active surveillance for potential pathogens in febrile patients, wild rodents, domestic dogs, cattle, and their associated ticks is thus of utmost importance in order to predict and combat emerging zoonotic diseases which could impact human health and livestock production in the Mnisi area and beyond.

This study sought to investigate wild rodents, cattle and dogs, as well as their associated ticks, as possible sources of zoonotic pathogen infection in the Mnisi community using a microbiome sequencing approach. Because partial 16S rRNA gene sequences with 99% sequence similarity to the zoonotic pathogen *A. phagocytophilum* have already been found in the region, we characterized this pathogen from samples collected in the area by sequence analysis of several molecular markers.

1.2 Research Hypothesis

A significant proportion of non-malarial acute febrile illness in the Mnisi community area can be attributed to zoonotic bacterial intracellular pathogens.

1.3 Research aims and objectives

The aim of this research was to investigate wild rodents, cattle and dogs as well as their associated ticks, as possible sources of zoonotic pathogen infection in the Mnisi community using a microbiome sequencing approach. We also screened AFI patient samples, as well as rodents, dogs, cattle and ticks for the presence of *A. phagocytophilum* using a real-time PCR assay. The specific objectives were:

1. The characterization of the bacterial blood microbiome of wild rodents collected from three different habitat areas in the Mnisi community area using next-generation sequencing.
2. The characterization of the bacterial blood microbiome of domestic dogs and cattle in the Mnisi community using next-generation sequencing.

3. The characterization of the bacterial microbiome of the salivary gland and midgut of *Amblyomma hebraeum* ticks collected from cattle across eight different dip tanks in the Mnisi community area.

4. The characterization of the bacterial blood microbiome of nine AFI patients in the Mnisi community using next-generation sequencing.

5. The occurrence and genetic diversity of *A. phagocytophilum* in AFI patients, rodents, cattle, dogs and *Rhipicephalus sanguineus* ticks collected from dogs using a quantitative real-time PCR assay and targeted sequencing of the 16S ribosomal RNA, *gltA*, *msp4* and *ankA* genes.

1.4 Thesis overview

**Chapter 1:** This gives a general overview of the background to the study and ends with the aims and objectives of the study.

**Chapter 2:** This gives a review of microbiome studies, especially in ticks, rodents, dogs, cattle and humans and a review of zoonotic bacterial pathogens previously incriminated in the aetiology of non-malarial acute febrile illness with a focus in Africa.

**Chapter 3:** The aim of this chapter was to provide a comprehensive insight into bacterial pathogens in the blood of wild rodents captured from different habitat areas in the Mnisi communal area using next-generation sequencing approaches. Barcoded sample-specific primers were used to amplify the 16S ribosomal RNA (rRNA) gene from genomic DNA from 25 wild rodents all previously molecularly identified as *Mastomys* spp. Purified PCR amplicons were submitted for circular consensus sequencing on the Pacific Biosciences platform at the genomic sequencing core of the Washington State University, Pullman, USA. Sequence analysis revealed notable organisms of zoonotic interest. Targeted species of interest from the microbiome analysis were further validated using conventional PCR techniques. Phylogenetic analysis was carried out on 16S rRNA gene sequences from the most prevalent pathogen detected in the microbiome analysis.

**Chapter 4:** The aim of this chapter was to survey the bacterial populations in the blood of domestic dogs and cattle in order to understand the role they play as reservoirs of zoonotic tick-borne pathogens in the Mnisi community area. Barcoded sample-specific primers were used to amplify the 16S ribosomal RNA gene from genomic DNA from 10 domestic dogs and nine cattle. Sequence analysis revealed the detection of *Anaplasma*, *Ehrlichia*, *Bartonella* and *Mycoplasma* species.

**Chapter 5:** The aim of this project was to use next-generation sequencing on the Pacific Biosciences platform to characterize the bacterial microbiome of *A. hebraeum* ticks collected from domestic cattle across eight different cattle dip tanks in the Mnisi community area, Mpumalanga, South Africa. Barcoded sample-specific primers were used to amplify the 16S rRNA gene from genomic DNA of 24
pools of salivary glands and 23 pools of midgut tissue dissected from A. hebraeum. Sequence analysis revealed R. africae as the dominant symbiont in the microbiome. Microbiome analyses were verified using a Rickettsia genus-specific qPCR assay. Phylogenetic analysis was carried out on 16S rRNA gene sequences of Rickettsia species.

Chapter 6: The aim of this chapter was to determine the bacterial population in the blood of nine patients in Mnisi community with AFI. Barcoded sample-specific primers were used to amplify the 16S rRNA gene from genomic DNA of nine AFI patients. Sequence analysis revealed zoonotic bacterial pathogens such as R. africae, Rickettsia spp., Brucella sp. and Stenotrophomonas spp., Herbaspirillum spp. from the blood of AFI patients.

Chapter 7: The aim of this chapter was to determine the occurrence of A. phagocytophilum and other Anaplasma spp. in different hosts in Mnisi Community, Mpumalanga Province, South Africa. The genetic diversity of A. phagocytophilum in AFI patients, rodents, cattle, dogs and R. sanguineus ticks was explored in order to understand its circulation in the study area. Samples were first screened with a quantitative real-time polymerase chain reaction (qPCR) assay that targets the msp2 gene of A. phagocytophilum; however this test was found to cross-react with Anaplasma sp. ZAM dog. Positive samples were characterized using discriminant genetic markers: 16S rRNA, gltA, msp4 and ankA genes. Phylogenetic analysis of the 16S rRNA and the gltA gene sequences obtained in this study were analysed to infer phylogenetic relationships between A. phagocytophilum and other Anaplasma spp.

Chapter 8: This chapter gives a general discussion, conclusion and recommendations on the outcomes of this study.
1.5 References


Chapter 2
Literature Review

2.1 Microbiome analysis

Expansion of the human population and the concurrent movement of domestic animals, together with ecological and environmental changes, including climate change and socio-economic changes, as well as the complex relationships that exist between disease-causing organisms, human populations and reservoirs have led to an increasing number of emerging diseases (Colwell et al., 2011). The epidemiology of known diseases is also constantly changing, especially that of vector-borne diseases (Colwell et al., 2011). Thus, although many causal agents of undifferentiated acute febrile illness (AFI) in humans are known and assays to detect them have been developed and can be used in active surveillance programs, it is possible that AFI could be caused by previously unknown pathogens. Microbiome analysis of bacterial communities in vectors, humans and animal reservoirs can be used as an effective tool in detecting the causes of AFI. The microbiome is simply defined as being made up of the entire community of microorganisms (commensals, symbionts and pathogens) that occupy a particular environment (Peterson, 2009).

2.1.1 Methods to study microbial communities

In classical microbiology, culture-based methods have been used in the study of microbes. Other traditional methods of pathogen identification involve microscopy and serology (Houpikian and Raoult, 2002). Optic microscopy has been used to detect new microorganisms from stained smears from blood and tissue (Burgdorfer, 1984, Goodman et al., 1996). *Bartonella quintana* was first detected in haematoxylin and eosin (H&E) stained tissue sections of patients with bacillary angiomatosis using histopathology (Maurin and Raoult, 1996). Electron microscopy has advantages over traditional optic microscopy in its adaptability and high sensitivity (Curry, 2000) and was first used to detect *Helicobacter pylori* in the gastric epithelium (Warren and Marshall, 1983). It however has the disadvantages of being expensive, not readily available, tedious and requires the use of experienced staff (Curry, 2000). Serology and antigenic detection have also been useful tools to establish indirect evidence to disease exposure (Houpikian and Raoult, 2002). Serology however, has the disadvantage of widespread cross-reaction observed between members of the same bacterial genus (Houpikian and Raoult, 2002). Bacterial culture remains the gold standard in the identification of microbes as it enables antigenic and genetic research, experimental modelling and antibiotic susceptibility tests (Houpikian and Raoult, 2002). However culture based methods have the disadvantage of being laborious and it has been estimated that between 20-60% of microorganisms associated with the human microbiome are uncultivable (Bik et al., 2006, Eckburg et al., 2005).

Next-generation sequencing (NGS) techniques have been shown to offer an alternative to bacterial culture in studying the diversity of bacterial communities in different organs and systems within the
human body (Whittle et al., 2019). Two types of NGS methods have been used in the study of microbiomes; they include shotgun sequencing and amplicon sequencing (Greay et al., 2018). The most common approach to the study of microbiomes has been amplicon sequencing of the 16S rRNA gene (Huttenhower et al., 2012). The 16S rRNA gene is a common tool used in amplicon sequencing because it is made up of both conserved and variable sequence regions that allows for the identification of microorganisms (Barko et al., 2018). Shotgun sequencing encompasses the fields of metagenomics and transcriptomics (Greay et al., 2018). Shotgun sequencing also offers some advantages over amplicon sequencing because it can be used to determine whole or partial genomes, it is also PCR-free therefore eliminating the bias of amplicon sequencing (Kuczynski et al., 2012). On the other hand, it is more expensive than amplicon sequencing and requires considerably more data analysis (Sims et al., 2014).

After the completion of the first human genome sequence which was achieved with Sanger sequencing, also known as first-generation sequencing, there was a clamour for faster and cheaper sequencing techniques, and this led to the development of second-generation sequencing or next-generation sequencing (NGS) methods (Grada and Weinbrecht, 2013). Different NGS platforms carry out huge parallel sequencing in the course of which millions of DNA fragments from a single sample are sequenced simultaneously (Grada and Weinbrecht, 2013). NGS has proven useful in the fields of public health and epidemiology through the sequencing of bacterial and viral species and has led to the discovery of novel virulence determinants (Grada and Weinbrecht, 2013).

High throughput sequencing using Roche 454, Pacific Biosciences (PacBio), Ion Torrent PGM and Illumina platforms can be used to determine the microbiome, or the organisms that reside within a particular host, both pathogenic and non-pathogenic (Grada and Weinbrecht, 2013). The Roche 454 platform has an average reading length of between 400-500 base range with paired reads and has been useful in the genome sequencing of bacteria, animals and humans (Ronaghi et al., 1996). The Illumina platform uses a reversible terminator chemistry, has a high-throughput, produces short reads and is based on the detection of fluorescence produced by the addition of fluorescently tagged nucleotides to a growing fragment of DNA with an error rate of <0.4% in the data produced from this platform (Grada and Weinbrecht, 2013). The Roche 454 platform refers to as one of the third-generation sequencing methods, uses an instrument that sequences separate DNA fragments in real-time. The methodology involves single DNA polymerases being fixed to the surface of microscope slides (Glenn, 2011). The sequence of the individual DNA strands is detected when each dNTP gives off a distinctive fluorescent label before being separated during synthesis (Glenn, 2011). PacBio sequencing gives the longest available read length of 2500 bp, produces strobic reads containing numerous sub reads from a single strand of DNA and is relatively inexpensive compared to other NGS platforms (Glenn, 2011). PacBio library preparation however
requires a high DNA input when compared to Illumina and Ion Torrent PGM which both require far less DNA (Quail et al., 2012).

2.1.2 Statistical approaches in community analyses

The diversity of bacterial microbiomes is commonly analysed using alpha and beta diversity indices. The alpha diversity refers to the number of species in a particular sample, i.e. diversity within a sample, while the beta diversity index is a measure of the dissimilarity between samples or diversity seen between different samples (Lozupone et al., 2007, Whittaker, 1972). Alpha diversity rarefaction plots are commonly used in microbiome studies to determine if sufficient sequencing depth was achieved for every sample (Greay et al., 2018). A plateau in the plot indicates that sufficient sequencing depth was achieved while an increasing trend in the plot indicates that more sequencing depth needs to be attained (Greay et al., 2018). Ordination techniques such as principal component analysis (PCA) and correspondence analysis (CA) have been used to break down complex data sets into primary values that show variation within different variables (Skalski et al., 2018). Nonmetric multidimensional scaling (NMDS) ordination has also been used as a statistical tool in microbiome studies to compare samples based on their rarefied bacterial populations while also analysing the phylogenetic relationship between operational taxonomic units (OTUs) (Gall et al., 2017, Rynkiewicz et al., 2015). Statistical tests like the t test, one way analysis of variance (ANOVA), Pearson’s test and the non-parametric ANOVA (Kruskal-Wallis test) have been used to calculate statistical inferences in microbiome studies (Gosiewski et al., 2017, Li et al., 2018, Paisse et al., 2016).

2.1.3 Tick microbiome studies

It is documented that ticks are second to mosquitoes in the transmission of a large number of pathogenic organisms, including protozoa, bacteria and viruses, to humans and animals worldwide. (Colwell et al., 2011, Jongejan and Uilenberg, 2004). Evaluation of the microbial diversity residing in tick vectors is thus of medical and veterinary significance and is important in detecting endemic infections, for assessing new and emerging zoonotic pathogens and to reveal the associated bacteria within the tick hosts (Carpi et al., 2011). Non-pathogenic microbes residing in ticks have also been implicated as important drivers in the transmission of tick-borne pathogens that impact on human and animal health (Bonnet et al., 2017). Numerous papers have been published describing the microbiome of various species of ticks (Bonnet et al., 2014, Budachetri et al., 2014, Carpi et al., 2011, Fryxell and DeBruyn, 2016, Gofton et al., 2015a, Qiu et al., 2014, Vayssier-Taussat et al., 2013). A study using the Illumina platform was able to identify the emerging zoonotic pathogens Candidatus Neoehrlichia mikurensis, Borrelia miyamotoi and Rickettsia felis from questing Ixodes ricinus nymphs collected from the vegetation in Western France, as well as several expected pathogens like A. phagocytophilum, R. helvetica and Borrelia burgdorferi (Vayssier-Taussat et al., 2013). Other studies using the Illumina platform revealed the zoonotic parasites Babesia divergens, B. microti and Babesia sp. EU1 as well as B. major and Theileria species in I. ricinus ticks collected from Eastern France (Bonnet et al., 2014).
Another study revealed how the choice of blood meals alters the pathogen richness and composition of the microbiome of *Ixodes pacificus* in the USA (Swei and Kwan, 2016).

Studies using the Ion Torrent platform have detected *Borrelia* sp., *Bartonella henselae* and *Rickettsia* spp. in *Ixodes* ticks in Australia (Gofton et al., 2015b), *Coxiella* and *Rickettsia* spp. as dominant taxa from *Haemaphysalis* ticks in Malaysia (Khoo et al., 2016) and also microbiome changes following blood feeding in *Amblyomma maculatum* in the US (Menchaca et al., 2013).

A study done using the Pac Bio platform revealed how microbiome disruptions of *Dermacentor andersoni* ticks influenced infection of the ticks with *Anaplasma marginale* (Gall et al., 2016). Another study from the same group also using the Pac Bio platform showed how the bacterial microbiome of *D. andersoni* differed in ticks collected from different geographic locations and was unique to particular tissues, in this case the salivary glands and midgut (Gall et al., 2017).

In general, NGS has been used as a tool to discover tick-borne parasites which could be potentially transmitted by the ticks to susceptible hosts, as well as the discovery of unexpected parasites (Bonnet et al., 2014, Carpi et al., 2011, Vayssier-Taussat et al., 2013).

### 2.1.4 Other microbiome studies

The majority of microbiome studies in humans funded by the US NIH Human Microbiome Project (HMP) and the European Commission Metagenomics of the Human Intestinal Tract (MetaHIT), have focused on describing the microbiota of the gut, nasal cavity, oral cavity, skin and vagina (Deng and Swanson, 2015). A study in humans, conducted in France, assessed the microbial profile in the blood of healthy donors using 16S rRNA gene sequencing in order to ensure that blood supplied to patients in need was not harmful (Paisse et al., 2016). Other studies have used NGS as a diagnostic tool to identify infectious pathogens in the blood of septic patients (Grumaz et al., 2016) as well as comparing the bacterial diversity in the blood of septic patients against blood collected from healthy volunteers (Gosiewski et al., 2017). Another recent study gave a comprehensive description of heterogeneous microbiota in the blood of apparently healthy individuals as well as in patients with severe acute pancreatitis (Li et al., 2018).

Novel pathogens, mostly viruses, have been discovered using NGS techniques worldwide. Examples include the detection of a novel human papillomavirus from patients with febrile respiratory conditions in the United States (Mokili et al., 2013), and the detection of lujo virus, a haemorrhagic fever associated virus from patients in South Africa (Briese et al., 2009). Other studies include the discovery of a new strain of *E. coli* responsible for an outbreak of haemolytic-uremic syndrome in Germany (Rasko et al., 2011) and the discovery of a new species, *Staphylococcus cornubiensis*, from a human skin infection in the United Kingdom (Murray et al., 2018).
A few studies have also been published that elucidated the diversity of bacterial communities in the blood of rodents and their ectoparasites (Cohen et al., 2015, Rynkiewicz et al., 2015). However, the majority of microbiome studies that have been carried out on animals have been primarily focused on the microbiome of the gastrointestinal tract. Gastrointestinal microbiome studies have been carried out in mice (Gu et al., 2013), dogs and cats (Hand et al., 2013, Handl et al., 2011, Ritchie et al., 2008, Suchodolski et al., 2008, Tun et al., 2012), and cattle (Dowd et al., 2008, Myer et al., 2016).

2.2 Acute febrile illness

In recent years the occurrence of malaria as a febrile illness, and fatalities associated with it, have been on the decline in Africa; this is despite the disease still being a major cause of fever on the continent (D'Acremont et al., 2010). This development has led to research focused on the aetiology of unexplained causes of febrile illness in Africa (Mourembou et al., 2015), and several studies have identified known pathogens as the cause of AFI. The flea-borne pathogen, *R. felis*, the cause of flea-borne spotted fever (FBSF) has been detected in AFI patients in Kenya, Tanzania and Senegal (Maina et al., 2012, Prabhu et al., 2011, Socolovschi et al., 2010b). Serological surveys in Cameroun have detected antibodies to *R. africae*, the cause of African tick-bite fever (ATBF), in the serum of non-malarial AFI patients (Ndip et al., 2004). Other studies conducted in Tanzania have implicated *C. burnetti*, the cause of Q fever, as the leading cause of non-malarial AFI in patients from resource-constrained communities (Crump et al., 2013, Prabhu et al., 2011). *Bartonella* species causing bartonellosis have been detected in AFI patients in the Congo and Tanzania (Hercik et al., 2017, Laudisoit et al., 2011). Studies in East Africa have detected *Brucella abortus* and *B. melitensis* in febrile children in Tanzania (Chipwaza et al., 2015). Brucellosis was also confirmed in 3.5% of hospitalized patients in Northern Tanzania (Bouley et al., 2012), 2.6% of febrile outpatients in Ethiopia (Animut et al., 2009) and in AFI patients in Kenya (Njeru et al., 2016). Leptospirosis has been confirmed in 8.4% of AFI patients in Northern Tanzania (Biggs et al., 2011). In South Africa in the Mnisi community in Mpumalanga Province, recent research carried out on AFI patients found 24.1% of patients to be serologically positive for spotted fever group (SFG) *Rickettsia*, 2.3% to *C. burnetti*, and 6.8% to *Leptospira* spp.; furthermore, 9.5% of patients tested were positive for *Bartonella* species using PCR (Simpson et al., 2018).

The current knowledge of selected zoonotic bacterial pathogens that have been implicated as causes of non-malarial AFI is presented in the following sections.

2.2.1 *Anaplasma phagocytophilum*

*Anaplasma phagocytophilum* is a gram-negative obligate intracellular bacterium which infects neutrophils and causes granulocytic anaplasmosis worldwide (Dumler et al., 2001). Human granulocytic anaplasmosis (HGA) is an emerging tick-borne zoonosis caused by *A. phagocytophilum*. The disease is a febrile illness that can be mild to severe to life threatening (Bakken and Dumler, 2008). It was first reported in the Unites States (Chen et al., 1994) and has occurred in Europe (Petrovec et al., 1997)
and Asia (Ohashi et al., 2013, Zhang et al., 2009). Anaplasma phagocytophilum also causes tick-borne fever (TBF) in ruminants, equine granulocytic anaplasmosis (EGA) in horses and canine granulocytic anaplasmosis (CGA) in dogs (Stuen et al., 2013). Anaplasma phagocytophilum is transmitted to animal hosts by Ixodes scapularis and I. pacificus in the USA and by I. ricinus in Europe and I. persulcatus in Asia (Thomas et al., 2009). Other species of Ixodes ticks reported to maintain enzootic cycles of A. phagocytophilum include I. spinipalpis in America and I. trianguliceps in northwestern England (Bown et al., 2003, Zeidner et al., 2000).

2.2.1.1 Biology of Anaplasma phagocytophilum
In infected granulocytes A. phagocytophilum usually appears as morulae in intracytoplasmic vacuoles (Woldehiwet, 2010). The organism is pleomorphic but can appear as cocci or elliptical in shape and is 0.5-1.5 µm in diameter (Dumler et al., 2005). Anaplasma phagocytophilum divides by binary fission and it lacks a lipopolysaccharide biosynthetic mechanism (Dumler et al., 2005). Different variants of A. phagocytophilum have different host predilections and not all strains can infect humans (Barbet et al., 2013). The Ap-ha strain, which is maintained in the white-footed mouse and which infects humans, does not seem to cause infection in ruminants (Massung et al., 2003), while the Ap-variant 1 maintained in white-tailed deer infects ruminants and horses but does not cause infection in mice and humans (Massung et al., 2006). Other host predilections of A. phagocytophilum include the US human HZ strain and the horse MRK strain. In general, strains infecting humans, rodents and dogs have been found to be more similar to each other and different from strains infecting ruminants (Barbet et al., 2013).

2.2.1.2 Epidemiology of Anaplasma phagocytophilum
(a) Worldwide
In Europe, particularly in the United Kingdom, outbreaks of TBF in ruminants have been reported in domestic goats (Gray et al., 1988), feral goats (Foster and Greig, 1969), sheep (Gordon et al., 1932), and in red, fallow and roe deer (Alberdi et al., 2000, Silaghi et al., 2011). In Norway, Slovenia, Switzerland and Austria, TBF has been reported in red deer, moose and chamois (Jenkins et al., 2001, Liz et al., 2002, Petrovec et al., 2002, Polin et al., 2004, Stuen et al., 2001). CGA was first reported in dogs in the USA (Madewell and Gribble, 1982), subsequently it was reported in Scandinavia (Engervall et al., 1997), Switzerland (Pusterla et al., 1998b), and the United Kingdom (Shaw et al., 2001a, Shaw et al., 2005). HGA variants have also been shown to cause disease in dogs and horses in North America (Pusterla et al., 2001). In Scandinavia and other regions of Europe, infection with A. phagocytophilum in cats has been reported (Bjoersdorff et al., 1999, Lappin et al., 2004, Tarello, 2005). EGA was first reported in the USA (Gribble, 1969). EGA variants have been found to cause infection in horses in Scandinavia (Engvall and Engervall, 2002), Switzerland (Pusterla et al., 1998a), the United Kingdom (Shaw et al., 2001b), Brazil and Croatia (Gotić et al., 2017). In humans, HGA was first reported in the USA (Chen et al., 1994), subsequently infections were reported in different parts of Europe, China and Japan (Lotrič-Furlan et al., 1998, Ohashi et al., 2013, Strle, 2004, Zhang et al., 2009).
Nucleic acid-based techniques have been used to detect *Anaplasma phagocytophilum* infection in horses and ticks in Tunisia (M’Ghirbi et al., 2012), and a serological prevalence study also found horses with antibodies to *A. phagocytophilum* (Said et al., 2014). In Algeria, antibodies to *A. phagocytophilum* were detected in dogs (Azzag et al., 2015) and the pathogen was also detected by qPCR in cattle (Dahmani et al., 2015). In Egypt, 7.5% of farmers tested showed evidence of *A. phagocytophilum* infection (Ghafar and Eltablawy, 2011), and the pathogen was also detected in *R. sanguineus* ticks collected from dogs in that country (Ghafar and Amer, 2012). In Zambia, *A. phagocytophilum* infection was detected in 13.6% of baboons and vervet monkeys tested (Nakayima et al., 2014) and more recently an *Anaplasma* sp. closely related to the *Anaplasma* sp. SA dog strain detected in South Africa (Inokuma et al., 2005) named *Anaplasma* sp. ZAM dog was detected in domestic dogs in Lusaka (Vlahakis et al., 2018). In Zimbabwe, a PCR based analysis of captive wild felids found 7% of lions, 13% of southern African wild cats and 50% of servals to be infected with *A. phagocytophilum* (Kelly et al., 2014).

In South Africa, there has been no official report of the diagnosis of *A. phagocytophilum* infection in humans. The organism has been reported in ticks collected from cattle, sheep and goats in South Africa (Mtshali et al., 2015). However, this finding should be verified since close inspection of the PCR primers used in that study indicates that they could amplify any *Anaplasma* species, and the sequence data presented does not conclusively prove the presence of *A. phagocytophilum*.

There is a published report of the molecular detection of a bacterium closely related to *A. phagocytophilum* designated *Anaplasma* sp. SA dog strain in three dogs presented at the Veterinary Teaching Hospital of the Medical University of South Africa (Inokuma et al., 2005). The same organism was detected in domestic dogs in the Mnisi community area (Kolo et al., 2016). An *A. phagocytophilum*-like bacterium was also reported from ticks collected from dogs and cats from four provinces in South Africa (Mtshali et al., 2017).

While the aforementioned studies (summarised in Table 2-1) suggest *A. phagocytophilum* detection across the breadth of Africa, caution is urged, as all the fragment sizes of the sequences generated in these studies were short (below 490 nucleotides), so it cannot be categorically proven if indeed *A. phagocytophilum* was detected or other closely related *Anaplasma* species. Furthermore, with the advent of high throughput detection methodologies and 16S rRNA gene survey studies becoming popular, there have been increasing numbers of distinct *Anaplasma*-like 16S rRNA gene sequences being deposited in the public sequence databases. The relationship of these newly detected agents to known pathogens, and their ability to serve as a source of cross-reaction in molecular testing, have not been well assessed.
**Table 2-1:** Studies that have reported the detection of *A. phagocytophilum* in Africa using nucleic acid-based detection methods.

<table>
<thead>
<tr>
<th>Host animal or ticks</th>
<th>Country</th>
<th>Gene used for detection</th>
<th>Fragment size of sequence generated</th>
<th>% identity on BLAST</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dogs</td>
<td>Algeria</td>
<td>msp2</td>
<td>Not sequenced</td>
<td>Not sequenced</td>
<td>(Azzag et al., 2015)</td>
</tr>
<tr>
<td>Horses and Hyalomma marginatum</td>
<td>Tunisia</td>
<td>16S rRNA</td>
<td>431 bp</td>
<td>99% identical to <em>A. phagocytophilum</em> (MH122889)</td>
<td>(M’Ghirbi et al., 2012)</td>
</tr>
<tr>
<td>Baboons and vervet monkeys</td>
<td>Zambia</td>
<td>16S rRNA</td>
<td>305 bp</td>
<td>100% identical to <em>A. phagocytophilum</em> (MH122891)</td>
<td>(Nakayima et al., 2014)</td>
</tr>
<tr>
<td><em>Rhipicephalus sanguineus</em> collected from dogs</td>
<td>Egypt</td>
<td>16S rRNA</td>
<td>219 bp</td>
<td>99% identical to <em>Anaplasma</em> sp. clone (MK138362)</td>
<td>(Ghafar and Amer, 2012)</td>
</tr>
<tr>
<td>Cattle</td>
<td>Algeria</td>
<td>23S rRNA</td>
<td>485 bp</td>
<td>100% identical to <em>A. phagocytophilum</em> (CP006618)</td>
<td>(Dahmani et al., 2015)</td>
</tr>
<tr>
<td>Cattle</td>
<td>Algeria</td>
<td>16S rRNA</td>
<td>313 bp</td>
<td>100% identical to <em>A. phagocytophilum</em> (MH122891)</td>
<td>(Dahmani et al., 2015)</td>
</tr>
<tr>
<td><em>Amblyomma cohaerens-</em> <em>Rhipicephalus pulchellus</em> <em>Rhipicephalus maculatus</em></td>
<td>Ethiopia</td>
<td>msp2</td>
<td>Not sequenced</td>
<td>Not sequenced</td>
<td>(Hornok et al., 2016)</td>
</tr>
<tr>
<td></td>
<td>Ethiopia</td>
<td>16S rRNA</td>
<td>345 bp PCR</td>
<td>Not sequenced</td>
<td>(Teshale et al., 2016)</td>
</tr>
<tr>
<td></td>
<td>Kenya</td>
<td>16S rRNA</td>
<td>319 bp</td>
<td>100% identical to <em>A. phagocytophilum</em> (MG668811) and Candidatus Anaplasma boleense (MH244922)</td>
<td>(Mwamuye et al., 2017)</td>
</tr>
<tr>
<td>Lions, serval and south African wild cat</td>
<td>Zimbabwe</td>
<td>16S RNA</td>
<td>316 bp</td>
<td>99% identical to uncultured <em>Anaplasma</em> sp. (MG869522)</td>
<td>(Kelly et al., 2014)</td>
</tr>
<tr>
<td>Dog</td>
<td>South Africa</td>
<td>16S rRNA</td>
<td>258 bp</td>
<td>100% identical to uncultured <em>Anaplasma</em> sp. (KP823596)</td>
<td>(Kolo et al., 2016)</td>
</tr>
<tr>
<td><em>Rhipicephalus evertsi evertsi, Rh. decoloratus</em> and <em>A. hebraeum</em></td>
<td>South Africa</td>
<td>16S rRNA</td>
<td>250 bp</td>
<td>98% identical to <em>A. phagocytophilum</em> (DQ648489)</td>
<td>(Mtshali et al., 2015)</td>
</tr>
<tr>
<td><em>Rhipicephalus sanguineus</em> and <em>Haemaphysalis elliptica</em></td>
<td>South Africa</td>
<td>16S rRNA</td>
<td>250 bp</td>
<td>93% identical to uncultured <em>Anaplasma</em> sp. (KX417195)</td>
<td>(Mtshali et al., 2017)</td>
</tr>
</tbody>
</table>

2.2.1.3 Lifecycle of *Anaplasma phagocytophilum*

The life cycle of *A. phagocytophilum* involves both mammalian and tick stages. *Anaplasma phagocytophilum* is not effectively transmitted transovarially in ticks so typically larvae are not infected (Rikihisa, 2011). Ticks of different stages (larvae, nymphs and adults) acquire *A. phagocytophilum* when feeding on infected animals. Once infected, *A. phagocytophilum* is maintained in the tick and transstadial transmission to susceptible animals occurs during blood feeding (Rikihisa, 2011).
animals like white footed mice (*Peromyscus leucopus*) and white-tailed deer (*Odocoileus virginianus*) serve as the primary reservoirs of variants of *A. phagocytophilum* in North America, while domestic animals like dogs serve as secondary reservoirs for human infection (Telford et al., 1996). Humans are accidental hosts from the bite of infected ticks and do not form part of the normal lifecycle of the organism (Rikihisa, 2010). Nosocomial transmission of *A. phagocytophilum* in humans has been reported in China though it was not conclusively confirmed by isolation in culture or visualization of the organism in blood smears (Zhang et al., 2008). Upon transmission to a mammalian host, *A. phagocytophilum* multiplies in granulocytes and in monocytes or macrophages by an impairment of the host’s innate defence mechanisms (Rikihisa, 2006).

2.2.1.4 Clinical signs

(a) Human granulocytic anaplasmosis

Clinical signs of infection with *A. phagocytophilum* in humans are reported to include flu-like symptoms of fever, chills, headache and myalgia (Aguero-Rosenfeld et al., 1996, Bakken and Dumler, 2008, Dumler et al., 2005, Dumler et al., 2007). In the United States, HGA can be serious, requiring hospitalization in 36% of patients (Dahlgren et al., 2011). Haematological irregularities like leucopenia, thrombocytopenia, anaemia and an increase in the serum aspartate and alanine aminotransferase liver enzyme activity, which is suggestive of liver damage, have also been observed in 3% of hospitalised patients (Aguero-Rosenfeld et al., 1996, Bakken and Dumler, 2008, Dumler et al., 2005, Dumler et al., 2007).

(b) Granulocytic anaplasmosis in domestic animals

Clinical signs of tick-borne fever in ruminants include a high fever and anorexia accompanied by inclusion bodies (morulae) that are seen in the neutrophils and thrombocytopenia (Engvall and Egenvall, 2002, Stuen, 2007). Lameness and losses in milk production have also been observed in domestic ruminants (Stuen et al., 2002, Tuomi, 1967), as well as abortions in ewes and decreased fertility in rams and bulls (Stuen, 2007). *Anaplasma phagocytophilum* infection is not usually fatal in ruminants unless there are complications with other opportunistic infections (Stuen, 2007).

Clinical signs of CGA include fever, lethargy and anorexia (Little, 2010). Lameness as a result of a neutrophilic polyarthritis, vomiting, diarrhoea and bleeding anomalies like epistaxis have also been observed (Engvall et al., 1997, Engvall and Egenvall, 2002, Greig et al., 1996, Kohn et al., 2008, Poitout et al., 2005). In cats infected with *A. phagocytophilum*, clinical signs that have been seen include fever, inappetance, lethargy, hyperaesthesia, lameness, conjunctivitis and a lack of body coordination (Heikkila et al., 2010).

The gravity of the clinical signs seen in EGA is usually dependent on the age of the horse, the length of the illness, the variant of the organism causing the illness and the immune condition of the host (Madigan and Gribble, 1987, Stuen et al., 1998). Adult horses more than four years of age show signs
of fever, depression, anorexia, distal limb oedema, petechial bleeding of mucosa, jaundice, ataxia and a disinclination to move (Pusterla and Madigan, 2013). Horses younger than four years show less pronounced signs of a moderate fever, depression, a mild limb oedema and ataxia (Pusterla and Madigan, 2013). Pathologic findings in horses infected with *A. phagocytophilum* include leukopenia, thrombocytopenia, anaemia, jaundice and presence of morulae in the neutrophils and eosinophils (Pusterla and Madigan, 2013).

2.2.1.5 Diagnosis, treatment and control of *A. phagocytophilum*

The diagnosis of *A. phagocytophilum* can be facilitated by the identification of the organism in blood smears using a light microscope, serological detection of antibodies to the organism in blood, cultivation of the organism in cell cultures and molecular detection of the organism using polymerase chain reaction (PCR) methods (Stuen et al., 2013). The drug of choice in the treatment of HGA is tetracycline (Brodie et al., 1986, Woldehiwet, 2010). In patients allergic to tetracyclines, rifampin can be used as an alternative drug for treatment (Bakken and Dumler, 2006, Jin et al., 2012). Doxycycline hyclate has been used successfully in the treatment of HGA at a dosage of 100 mg for adults and 2.2 mg/kg for children every 12 hours (Stuen et al., 2013).

Prevention of the disease in domestic animals is based on managing tick infestation by the use of acaricides by dipping or the use of pour-on applications of synthetic pyrethroids (Stuen, 2007). Use of long-acting antibiotics like tetracycline as prophylaxis before movement of animals from tick-free environments to tick-infested grazing land has also been practiced (Brodie et al., 1986). Biological tick control has been proposed as an alternative method to tick management (Samish et al., 2004) but there are limitations on maintaining the biological control of ticks under natural field conditions (Stuen et al., 2013).

There is no vaccine available against *A. phagocytophilum* with attempts to develop one so far unsuccessful (Ge and Rikihisa, 2006), as a vaccine requires an antigen that is broadly conserved among isolates (Stuen et al., 2013). Vaccines against ticks have been developed, and cement antigen vaccines such as TICKGUARDPLUS and Gavac are examples that have been tested and are available commercially. They work by targeting the midgut cells of ticks resulting in their rupture and eventual tick mortality (De la Fuente et al., 2007, Labuda et al., 2006). Vaccines that suppress the expression of subolesin, a tick protective antigen that plays a role in the processes of reproduction, development and digestion of blood meals in the tick, are also being developed (De la Fuente et al., 2006). There is no developed vaccine yet against suppressing the activities of tick salivary glands. Tick vaccines have been advocated as a cost-effective and environmentally friendly method of tick control in comparison to chemical methods (De la Fuente and Kocan, 2006).
2.2.2 Rickettsiae

Rickettsioses are a group of infectious diseases caused by bacteria of the genus *Rickettsia* (Parola et al., 2005, Raoult and Roux, 1997). They are transmitted by arthropod vectors which include ticks, fleas, mites and lice (Kelly et al., 2002). Rickettsial organisms include members of the spotted fever group of rickettsiae, typhus group of rickettsiae and *Orientia tsutsugamushi* (formerly known as *Rickettsia tsutsugamushi*), the cause of scrub typhus (Parola and Raoult, 2006). Spotted fever rickettsiosis has been recognized in South Africa since the beginning of the 20th century with the earliest report of human disease as far back as 1911 (McNaught, 1911). *Rickettsia conorii*, the agent of Mediterranean spotted fever (MSF), has long been associated with human disease in South Africa (Pretorius and Birtles, 2002, Troup and Pijper, 1931). *Rickettsia aeshlimannii*, which causes a Mediterranean spotted fever-like illness, has also been identified in South Africa (Beati et al., 1997). *Rickettsia africae*, the agent of African tick-bite fever (ATBF), caused by *Rickettsia africae*, is also prevalent in South Africa (Jensenius et al., 2003a). *Rickettsia sibirica mongolotimonae*, which causes a lymphangitis-associated rickettsiosis (LAR), has been reported in France from a recent traveller from Egypt (Socolovschi et al., 2010a) and in a construction worker in the Northern Province of South Africa (Pretorius and Birtles, 2004).

2.2.2.1 *Rickettsia africae*

*Rickettsia africae* is the cause of ATBF (Parola et al., 2005). It is transmitted by ticks of the genus *Amblyomma* (Kelly et al., 1996), with *A. hebraeum* and *A. variegatum* being regarded as the reservoir hosts and vectors of the organism (Jensenius et al., 2003a). In South Africa, *R. africae* is transmitted mainly by *A. hebraeum* (Kleinerman et al., 2013). In Africa, *R. africane* has also been detected in a variety of other ticks including *A. lepadum*, *A. compressum*, *Rhipicephalus annulatus*, *R. evertsi evertsi*, *R. decolaratus*, *R. sanguineus*, *R. geigy*, and *Hyalomma dromedarii* in Djibouti, Guinea, Senegal, Nigeria, Liberia, the Democratic Republic of Congo, Botswana, Algeria and Egypt (Abdel-Shafy et al., 2012, Kernif et al., 2012, Mediannikov et al., 2012a, Mediannikov et al., 2010a, Mediannikov et al., 2012b, Mediannikov et al., 2010b, Ogo et al., 2012, Portillo et al., 2007, Reye et al., 2012, Socolovschi et al., 2007). In recent years, it has been detected in *A. hebraeum* (Halajian et al., 2016) and also in *Haemaphysalis elliptica* in South Africa (Kolo et al., 2016).

Signs and symptoms of ATBF include multiple inoculation eschars, headaches, nuchal myalgia, painful lymphadenopathy, and occasionally, a sparse vesicular rash (Jensenius et al., 2003a). *Rickettsia africane* has been detected in ticks and/or humans in 22 countries in sub-Saharan Africa (Parola et al., 2013). Most cases of African tick-bite fever are seen in international travellers returning from sub-Saharan Africa (Althaus et al., 2010, Fournier et al., 1998, Jensenius et al., 2003b). In Cameroon, *R. africane* DNA was detected in 6% of the indigenous patients that presented with an undifferentiated fever that was not caused by malaria or typhoid fever (Ndip et al., 2004). In a recent serological survey conducted in the Mnisi community area, prior exposure to SFG *Rickettsia* was detected in 92.2% of patients that presented with AFI in community clinics and 64.4% of farmers, herdsman and veterinary personnel who were screened (Simpson et al., 2018).
2.2.2.2 *Rickettsia conorii*

A previous study using multi locus sequence typing (MLST) proposed four sub-species of *R. conorii* as *R. conorii* subsp. *conorii*, *R. conorii* subsp. *indica*, *R. conorii* subsp. *caspia* and *R. conorii* subsp. *israelensis* (Zhu et al., 2005). *Rickettsia conorii* subsp. *conorii* is the cause of MSF. It is mainly transmitted by the brown dog tick *Rhipicephalus sanguineus* (Parola et al., 2005). To date, *R. conorii* subsp. *conorii* has been reported in the Central African Republic, Zimbabwe, Senegal, South Africa, Uganda, Swaziland, Kenya, Libya, Egypt, Algeria, Tunisia and Morocco (Mediannikov et al., 2010a, Parola et al., 2013, Socolovschi et al., 2007). It has been detected by polymerase chain reaction (PCR) assays in *R. sanguineus* ticks in Algeria, Tunisia and Morocco (Bitam et al., 2006, Boudebouch et al., Sfar et al., 2009). *Rickettsia conorii* subsp. *conorii* has also been detected in *Rhipicephalus mushumae* ticks from cattle in the Central African Republic, in *Rhipicephalus simus* and *Haemaphysalis leachi* ticks from dogs in Zimbabwe (Parola et al., 2013), in an *R. evertsi evertsi* tick from a horse in Senegal and in a *H. punctaleachi* tick from a dog in Uganda (Mediannikov et al., 2010a, Socolovschi et al., 2007). Clinical signs of MSF include high fever with an acute onset (above 39°C), headache, malaise, arthromyalgias and an eschar at the location of the tick bite (Roverey and Raoult, 2008). There have been increasing numbers of reports of human cases of MSF in Algeria, Tunisia and Morocco (Kernif et al., 2012). In Algeria, 49% of patients hospitalized for MSF had very serious forms of the disease (Kaabia et al., 2009, Mouffok et al., 2009). *Rickettsia conorii* subsp. *conorii* was also detected in the blood of a young girl with fever in Senegal (Mediannikov et al., 2010a) and in a South African man on a trip to Brazil who later died of the infection (De Almeida et al., 2010). Clinical signs seen in this fatal case included fever, chills, vomiting, hematuria, sore throat and a generalized maculopapular rash (De Almeida et al., 2010). *Rickettsia conorii* was also recently detected in a survey of small mammals across seven provinces in South Africa and parts of Namibia (Essbauer et al., 2018).

Other subspecies of *R. conorii* identified in Africa include *R. conorii* subsp. *israelensis* which causes of Israeli spotted fever (ISF). The organism has been detected by PCR in skin biopsy samples from patients in southern Tunisia (Znazen et al., 2011). *Rickettsia conorii* subsp. *caspia*, the cause of Astrakhan fever, was detected in a French traveller who was returning from Chad, and presented with a fever and a maculopapulous rash (Fournier et al., 2003).

2.2.2.3 *Rickettsia aeschlimannii*

*Rickettsia aeschlimannii* causes a Mediterranean spotted fever-like illness. The organism was first isolated from *Hyalomma marginatum marginatum* ticks in Morocco (Beati et al., 1997). *Rickettsia aeschlimannii* has been detected in different species of *Hyalomma* ticks collected from cattle and camels in Egypt, Sudan, Algeria, and Tunisia (Kernif et al., 2012), in *H. marginatum rufipes* and *H. truncatatum* in Senegal (Mediannikov et al., 2010a), in *R. evertsi evertsi* in Nigeria and Senegal (Mediannikov et al., 2010a, Reye et al., 2012) and in *R. annulatus* and *A. variegatum* in Nigeria (Reye et al., 2012). Human infections have been reported in a French traveller after a trip to Morocco (Raoult et al., 2003).
et al., 2002), in two febrile patients in Algeria (Mokrani et al., 2008) and in a South African patient infected after being bitten by a *Rhipicephalus appendiculatus* tick (Pretorius and Birtles, 2002).

2.2.2.4 *Rickettsia sibirica mongolotimonae*

*Rickettsia sibirica mongolotimonae* is the cause of lymphangitis associated rickettsioses (LAR) (Socolovschi et al., 2010a). It has been detected in *H. truncatum* ticks collected from cattle in Niger (Parola et al., 2001) and *H. truncatum* collected from horses, donkeys, cattle, sheep and goats in Senegal (Mediannikov et al., 2010a). Human infection with this organism has been reported in South Africa in a patient that presented with lymphangitis, fever and an eschar on his toe (Pretorius and Birtles, 2004). Two travellers who had recently returned to Southern France from Algeria (Fournier et al., 2005, Socolovschi et al., 2010a) and Egypt (Fournier et al., 2005, Socolovschi et al., 2010a) presented with fever, asthenia, headaches and arthromyalgia and were determined to have *R. sibirica* infection.

2.2.2.5 Other rickettsiae detected in South Africa

A number of other rickettsial pathogens have been reported in South Africa. *Rickettsia massilliae* was detected in ten *Amblyomma sylvaticum* and one *Rhipicephalus simus* collected from a Leopard tortoise in a study conducted in Limpopo and Western Cape Provinces (Halajian et al., 2016). In another study, *R. massilliae*, *R. felis*, *R. helvetica*, and three new Candidatus species, *Candidatus Rickettsia africaustralis*, *Ca R. rhabdomydis* and *Ca R. muridii*, were detected from small mammals collected from seven provinces in the country (Essbauer et al., 2018). A *Rickettsia felis*-like organism, *Ca R. asemboensis*, was also detected in *Ctenocephalides felis strongylus* fleas collected from domestic dogs in Mpumalanga Province (Kolo et al., 2016).

2.2.2.6 Diagnosis, treatment and control of rickettsial infections

Rickettsial diseases are difficult to diagnose clinically because of the similarity of clinical signs with many other infectious diseases. Therefore, laboratory investigations help in ascertaining the cause of rickettsial diseases (Parola and Raoult, 2006). The cultivation of rickettsiae in reference laboratories remains the most definitive way of isolating rickettsiae from clinical samples but it requires that the research is conducted in biosafety level 3 laboratories (Parola and Raoult, 2006). Embryonated chicken egg yolk sacs, laboratory animals, and cell cultures using Vero or L929 cells have been successfully used in the isolation of rickettsiae (La Scola and Raoult, 1997). Serological tests like the Weil-Felix test, indirect fluorescent antibody (IFA) test, enzyme-linked immunosorbent assay (ELISA), the cross-absorption (CA) technique and the western immunoblot assay are also useful tools in the diagnosis of rickettsioses (Ghorbani et al., 1997, La Scola and Raoult, 1997, Osuga et al., 1991, Thiga et al., 2015). Molecular techniques including PCR and sequencing are fast, sensitive and practical tools in the diagnosis of rickettsial pathogens from blood and skin biopsy samples, however these tests suffer from sensitivity issues and work best during acute infection (Parola et al., 2005). Several assays targeting different genes have been developed for the effective diagnosis of rickettsioses (Fournier and Raoult, 2004, Kidd et al., 2008, Renvoise et al., 2012, Stenos et al., 2005). Genomic advances have also
broadened the knowledge base on rickettsiae, with large amounts of genomic data readily available (Fournier et al., 2007, Merhej and Raoult, 2011).

Doxycycline remains the drug of choice for the treatment of rickettsioses at a dosage of 100 mg twice a day for adults while other newer macrolides like clarithromycin and azithromycin show promise as alternatives in the treatment of some rickettsioses (Botelho-Nevers et al., 2012, Dumler et al., 2007, Ghorbani et al., 1997). There were two types of vaccines previously in use against rickettsial diseases in humans whole killed bacteria and live attenuated bacteria (Walker, 2009). However these vaccines have sometimes caused severe illness with low protection (Gear, 1969). There is no vaccine currently available for the prevention of rickettsial infections (Richards, 2004).

2.2.3 Bartonella

*Bartonella* species are gram-negative intracellular bacteria that infect erythrocytes, endothelial cells and macrophages of mammals and are transmitted by arthropod vectors which include ticks, fleas, bat flies, sand flies, ked flies and lice (Billeter et al., 2012, Billeter et al., 2008). Infections with *Bartonella* species have a worldwide distribution, and since the discovery of the genus, more than 30 species of *Bartonella* have been identified from a variety of mammals (Kosoy et al., 2012). A number of *Bartonella* species have been implicated in causing disease in humans. *Bartonella henselae* is the cause of cat-scratch disease (CSD), endocarditis, bacillary peliosis and bacillary angiomatosis (Groves and Harrington, 1994). In South Africa, *B. henselae* was detected using nucleic acid-based techniques in 10% of HIV immunocompromised patients (Frean et al., 2002), and in cats using the IFA test (Kelly, 1996). *Bartonella bacilliformis* is the aetiological agent of Carrion’s disease, a biphasic condition causing Oroya fever in the acute phase and verruga peruana in the chronic stages (Clemente et al., 2012). It is a disease that occurs in Peru and parts of Ecuador and Colombia in South America (Minnick et al., 2014). Oroya fever is a serious disease that can be fatal if left untreated in the acute stages (Clemente et al., 2012). *Bartonella bacilliformis* is transmitted by phlebotomine sand flies (Clemente et al., 2012). *Bartonella quintana* causes trench fever transmitted by the human body louse *Pediculus humanus* variety *corporis* (Ohl and Spach, 2000). It is estimated that over a million troops had trench fever during World War I (Vinson et al., 1969). In immunocompromised individuals, *B. quintana* causes bacillary angiomatosis, endocarditis, and lymphadenopathy (Spach et al., 1993). In western countries of Europe and North America, *B. quintana* infection is linked to homeless people, alcoholism and poverty (Ohl and Spach, 2000).

Rodent-borne *Bartonella* species including *B. elizabethae*, *B. grahamii*, *B. tribocorum*, *B. washoensis*, *B. alsatica* and *B. vinsonii* subsp. *arupensis* have also been implicated in causing human infection especially in immunocompromised individuals (Anderson and Neuman, 1997, Breitschwerdt et al., 2007). In Africa, a serological study detected IgG antibodies against *B. henselae*, *B. quintana* and *B. claridgeiae* in 4.5% of patients that presented with fever in the Democratic Republic of Congo (DRC)
using an IFA test (Laudisoit et al., 2011). Bartonella spp. were also detected in 1% of AFI patients in Tanzania using real-time PCR (Hercik et al., 2017).

In South Africa, Bartonella spp. have been detected in 25% of rodents, 9.5% of cats, 22.5% of HIV patients, 9.5% of apparently healthy volunteers and 9% of dogs screened in Gauteng Province (Trataris et al., 2012). Recently, B. vinsonii subspecies berkoffii, B. henselae, B. quintana, B. thailandensis and Bartonella spp. were detected in 9.5% of AFI patients screened in the Mnisi community area (Simpson et al., 2018). In other studies in the country, Bartonella spp. were detected in small rodents from the Free State Province (Pretoirius et al., 2004) and B. elizabethae in 44% of rock mice (Brettschneider et al., 2012). A recent study reported an infection rate of 15% and the presence of three phylogenetic lineages of Bartonella species circulating in 80 Rhabdomys pumilio rodents captured from the Western Cape Province (Hatyoka et al., 2019). Bartonella elizabethae has been implicated as a cause of bacteraemia and endocarditis in humans (Daly et al., 1993). A wide range of Bartonella species have been detected from small mammals and ectoparasites in other African countries. Bartonella elizabethae and B. tribocorum were detected in rodents and hedgehogs sampled from Algeria, DRC and Tanzania (Bitam et al., 2009, Gundi et al., 2012). In Algeria, B. vinsonii subsp. berkoffii, B. clarridgeiae, and B. elizabethae were detected in 6.2% of domestic dogs (Kernif et al., 2010) and recently, B. tamiae was detected from 63.2% of Ixodes vespertilionis ticks and different genotypes of a Bartonella sp. were detected in soft ticks (Lafri et al., 2017, Leulmi et al., 2016). A diversity of Bartonella species has also been detected in small mammals and fleas in Ethiopia (Kumsa et al., 2014, Meheretu et al., 2013), Uganda (Billeter et al., 2014), Kenya (Halliday et al., 2015, Young et al., 2014) and in fleas from Morocco, DRC and Egypt (Boudebouch et al., 2011, Loftis et al., 2006, Sackal et al., 2008). Bartonella species have also been detected from rodents, fleas and ticks in Nigeria (Kamani et al., 2013).

Clinical presentations seen in humans with Bartonella infection include fever, malaise, joint and muscle pain. Endocarditis and neurological abnormalities have been seen in more severe cases which may be life threatening, especially in individuals with weakened immune systems (Anderson and Neuman, 1997, Breitschwerdt et al., 2007).

2.2.3.1 Diagnosis, treatment and control of Bartonella spp.

Isolation by culture, serology and PCR are effective methods used in the diagnosis of bartonellosis (Anderson and Neuman, 1997). Culture of Bartonella is usually laborious requiring 2 to 6 weeks for the organism to be isolated (Anderson and Neuman, 1997). Serological tests remain the most practical method of detecting Bartonella infections, with commercial kits readily available for the detection of anti-Bartonella antibodies in humans (Regnery et al., 1992). Erythromycin at a dosage of 500 mg four times daily or doxycycline at a dosage of 100 mg twice daily are effective in the treatment of Bartonella spp. infections (Rolain et al., 2004). Control of vector-borne Bartonella species like B. henselae, B. quintana and B. bacilliformis or rodent-borne Bartonella spp. is usually achieved by reducing the population of
the vectors which include fleas, lice and sand flies or the reservoir in the case of rodents (Trataris et al., 2012).

### 2.2.4 *Coxiella burnetii*

*Coxiella burnetii* causes Q fever, an infection that presents as an acute to chronic fever and pneumonia in humans, as well as causing abortion in livestock (Vanderburg et al., 2014). It is transmitted by various tick species, and by direct contact in humans via inhalation of aerosols, eating infected animal products (Maurin and Raoult, 1999) or by direct contact with infected aborted foetuses or other birth products from animals (Vanderburg et al., 2014). Studies in Tanzania have implicated *C. burnetii* as the leading cause of non-malarial AFI in patients from resource-constrained communities (Crump et al., 2013, Prabhu et al., 2011). The organism was also detected in 9% of AFI patients in Tunisia (Kaabia et al., 2006), and 5% of AFI patients in Burkina Faso (Kizerbo et al., 2000). In northern Togo, a serological survey found 26.9% of humans, 16.1% of cattle, 16.2% of sheep and 8.8% of goats positive for *C. burnetii* (Dean et al., 2013). In a study carried out in the Gambia, a seroprevalence of 9.7% in humans and 24.7% in small ruminants was reported (Bok et al., 2017).

In South Africa, antibodies to *C. burnetii* have been detected in cattle (Gummow et al., 1987), and in wild dogs in the Kruger National Park (Van Heerden et al., 1995). The organism was also detected in 44% of *R. sanguineus* and 4% of *H. elliptica* collected from dogs, cats and ruminants across four provinces in the country (Mtshali et al., 2015, Mtshali et al., 2017). A recent seroprevalence study carried out on veterinary staff, cattle farmers and herders in a rural community in Mpumalanga Province also found 60.9% of individuals positive for anti-*Coxiella* antibodies (Simpson et al., 2018). The indirect immunofluorescent antibody test (IFA) is the most common serological test used in the diagnosis of acute or chronic Q fever (Herremans et al., 2013). In acute illness before the administration of antibiotics, PCR can be an effective tool used to detect *C. burnetii* infection from whole blood or serum (Fournier and Raoult, 2003). Doxycycline at a dosage of 100 mg twice daily for adults and 2.2 mg/kg body weight twice daily for children remains the drug of choice in the treatment of acute cases of Q fever (Anderson et al., 2013).
2.3 References


mimicking a hemorrhagic viral fever in a South African traveler in Brazil. *Ticks and Tick-Borne Diseases*, 1, 149-150.


STUEN, S. 2007. Anaplasma phagocytophilum - the most widespread tick-borne infection in animals in Europe. Veterinary Research Communications, 31 Suppl 1, 79-84.


TARELLO, W. 2005. Microscopic and clinical evidence for Anaplasma (Ehrlichia) phagocytophilum infection in Italian cats. The Veterinary Record, 156, 772-774.


Chapter 3
Bacterial blood microbiome of rodents captured from a human/livestock/wildlife interface in the Mnisi community, South Africa

3.1 Abstract
The Mnisi community is located in Bushbuckridge Municipality, Mpumalanga, South Africa and is nestled in the heart of a human/livestock/wildlife interface. Recent research in the area found rodents to be abundant with 76% of sampled households reporting rodents in and around their homes. Although it is well known that wild rodents serve as reservoir hosts for many human pathogens and that they play a key role in the natural circulation of viral, bacterial and parasitic infections, the role that rodents in the Mnisi community play in the transmission of zoonotic pathogens to humans is unknown. The aim of this project was to provide a comprehensive insight into bacterial pathogens in the blood of wild rodents captured from different habitat areas in Mnisi using next-generation sequencing approaches. The 16S rRNA gene was amplified from genomic DNA extracted from the blood of 25 wild Mastomys individuals using barcoded primers. Purified PCR amplicons were submitted for circular consensus sequencing on the Pacific Biosciences platform. A total of 65,060 bacterial sequences were obtained, with the average number of reads per sample being 2,602 which was sufficient to satisfy rarefaction curves indicating that all operational taxonomic units (OTUs) were captured. Notable organisms of zoonotic interest included members of the genus Bartonella, B. grahamii and Bartonella sp. RF255YX, which were the dominant organisms in the rodent blood microbiome. Overall, rodents from Hlalakahle (urban/periurban) and Tlhavekisa (communal rangeland) had higher proportions of Bartonella species (~85%), while those from Gottenburg (urban/periurban) and Manyeleti Game Reserve (protected area) (~45%) had lower Bartonella loads. Other organisms of zoonotic and veterinary significance detected included Ehrlichia sp. (~0.03%), Anaplasma centrale and A. marginale (~0.01%), Brucella spp. (~1%), Anaplasma phagocytophilum (~0.008%) and Coxiella burnetii (~0.02%). Detection of Brucella spp. was validated using conventional PCR techniques. This study serves as the first report on the detection of zoonotic agents: B. grahamii, Ehrlichia sp., A. phagocytophilum, C. burnetii and Brucella spp. in rodents in the Mnisi community and highlights the possible risk these organisms pose to human health within the community.

3.2 Introduction
Zoonotic pathogens make up an important and increasing number of emerging and re-emerging infectious diseases of humans worldwide (Woolhouse and Gowtage-Sequeria, 2005). Rodents are documented to serve as hosts and reservoirs of over 60 zoonotic pathogens that pose significant challenges to human health (Luis et al., 2013). The order Rodentia is the largest and most prolific order of all mammalian species and makes up to 43% of the total number of mammals worldwide (Huchon et al., 2002). The Mnisi community area in Bushbuckridge Municipality, Mpumalanga, South Africa is cradled in the heart of a human/livestock/wildlife interface. In this community, humans, domestic
animals and wildlife have perennial direct and indirect contact. Research in the area has found rodents to be common and abundant (Berrian et al., 2016) with 76% of households having reported seeing rodents around their homes. A recent study conducted in the area also suggested that rodent-borne zoonoses may be implicated as causes of non-malarial acute febrile illness (AFI) in humans (Simpson et al., 2018). In that study, 9.5% of AFI patients tested positive for the rodent-borne zoonotic pathogen *Bartonella* spp. on PCR, while 6.8% of patients showed prior exposure to *Coxiella burnetii*, the cause of Q fever and 2.3% to *Leptospira* spp. using serological tests (Simpson et al., 2018). The surveillance of zoonotic pathogens in rodents in this community is thus of utmost importance as the role they play in the transmission of zoonotic pathogens to humans is unknown.

The advent of next-generation sequencing (NGS) tools have helped to reveal the diversity of bacterial communities that exist in different hosts (Greay et al., 2018). Few NGS studies have been carried out to explore microbial communities in rodents (Cohen et al., 2015, Ge et al., 2018, Razzauti et al., 2015, Rynkiewicz et al., 2015). The studies conducted in China (Ge et al., 2018) and in France (Razzauti et al., 2015), focused on characterizing the bacterial communities in rodent spleens, while studies conducted in Israel and the US (Cohen et al., 2015, Rynkiewicz et al., 2015) explored the bacterial communities in rodent blood and their flea and tick vectors. The aforementioned studies on rodents concluded by stating that characterizing the bacterial communities in hosts and their arthropod vectors is crucial in understanding the abundance of vector-borne pathogens as well as the roles and interactions that exist within these bacterial communities (Cohen et al., 2015, Rynkiewicz et al., 2015). Rodents are important hosts of tick-borne pathogens and have been found to be competent reservoirs for these pathogens (Telfer et al., 2007). Therefore, the aim of this project was to provide an insight into bacterial pathogens in the blood of wild rodents captured from different habitat areas in Mnisi using near full-length 16S rRNA gene circular consensus sequencing on the Pacific Biosciences platform.

### 3.3 Materials and Methods

#### 3.3.1 Ethics approval

The animal ethics committee (AEC) of the Faculty of Veterinary Science, University of Pretoria approved the study with reference number (V105-15). Permission was obtained to conduct the research, in terms of Section 20 of the Animal Diseases Act of 1984, from the Department of Agriculture, Forestry and Fisheries (DAFF), South Africa, with reference numbers 12/11/1/1 and 12/11/1/1/6.

#### 3.3.2 Study area and sample collection

Blood samples were collected from 282 rodents from three habitat areas in the Mnisi community, Mpumalanga Province, South Africa: urban/periurban (Gottenburg and Hlalakahle), communal rangelands (Tlhavekisa) and a protected area (Manyeleti Game Reserve), during three research visits from 2014 to 2015. Wild rodents were trapped using baited Sherman live traps, identified using a field guide (Stuart and Stuart, 2001), and humanely euthanized using isoflurane, an inhalant anaesthetic. Blood was collected immediately after the last heartbeat in EDTA tubes and on FTA cards.
samples were moved to a BSL3 facility at the Faculty of Veterinary Science, University of Pretoria. Subsequently, 25 blood samples from the three habitat areas (six samples per site, except for Gottenburg which had seven samples) all obtained from *Mastomys* spp., identified using both morphological and molecular methods, were then selected for use in this study. Molecular typing of the rodents using the mitochondrial cytochrome b gene (Russo et al., 2006) was done at the Mammal Research Institute Laboratory, Department of Zoology & Entomology, University of Pretoria. Figure 3-1 shows the map of the study area and the locations of rodent capture.

![Figure 3-1](image_url)

**Figure 3-1**: Map of the Mnisi community study area in Bushbuckridge Municipality Mpumalanga Province, South Africa. Red stars show sites of rodent capture. Gottenburg and Hlalakahle are urban or periurban areas, while Thavekisa is located in communal rangelands. Manyeleti is a protected wildlife reserve. Dark green represents protected areas where wildlife roam freely (map produced by Estelle Mayhew).

### 3.3.3 PCR amplification and sequencing

DNA was extracted from blood spotted on FTA cards using the QIAamp® DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. PCR grade water was not spotted on FTA cards to evaluate potential contaminants on cards. Extracted DNA was kept frozen at -20°C until further analysis. The 16S rRNA gene (V1-V8 variable regions) was amplified from 25 rodent DNA samples using barcoded universal 16S rRNA gene primers, 27F (5’-AGA GTT TGA TCM TGG CTC AGA ACG-3’) and 1435R (5’-CGA TTA CTA GCG ATT CCR RCT TCA-3’) (Lane, 1991, Turner et al., 1999) (Table 3-1) as previously described (Gall et al., 2016). Sample-specific combinations of barcoded primers (Table 3-2) were used in a
final reaction volume of 25 µl containing 1 X Phusion Flash® High Fidelity PCR Master Mix (composed of Phusion Flash II DNA Polymerase, reaction buffer, dNTPs, and MgCl₂ (ThermoFisher Scientific, South Africa), 0.15 µM of each primer, and five µl of template DNA (approximately 100 ng of DNA). For each sample, three technical replicates were performed using the same sample-specific barcoded primer set (Gall et al., 2016). Anaplasma centrale vaccine strain (Onderstepoort Biological Products) was used as the positive control while PCR grade water was used as a no template negative control.

The thermal cycling parameters used were 98°C for 30 seconds, followed by 35 cycles of 98°C for 10 seconds, 60°C for 30 seconds and 72°C for 30 seconds, and a final extension at 72°C for 10 minutes. PCR products were visualized by electrophoresis on a 1.5% agarose gel (1 X TAE buffer, pH 8.0) stained with ethidium bromide and viewed under UV light. PCR products were purified using the QIAquick® PCR purification kit (Qiagen) according to the manufacturer’s instructions. Purified PCR products were couriered to the Genomics Sequencing Core of the Washington State University, Pullman, USA for circular consensus sequencing (CCS) on a Pacific Biosciences (PacBio, Menlo Park, CA) platform. Table 3-2 shows the sample information of the library for circular consensus sequencing.

**Table 3-1:** Barcoded primers used for the amplification of the 16S rRNA gene.

<table>
<thead>
<tr>
<th>Name</th>
<th>Code</th>
<th>*Multiplex identifier (MID) sequence</th>
<th>²Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>MID1-27F</td>
<td>F1</td>
<td>CGT ATC GCC TCC CTC GCG CCA TCAG ACG AGT GCGT</td>
<td>27F</td>
</tr>
<tr>
<td>MID2-27F</td>
<td>F2</td>
<td>CGT ATC GCC TCC CTC GCG CCA TCAG ACG CTC GACA</td>
<td>27F</td>
</tr>
<tr>
<td>MID3-27F</td>
<td>F3</td>
<td>CGT ATC GCC TCC CTC GCG CCA TCAG AGA CGC ACTC</td>
<td>27F</td>
</tr>
<tr>
<td>MID4-27F</td>
<td>F4</td>
<td>CGT ATC GCC TCC CTC GCG CCA TCAG AGC ACT GTAG</td>
<td>27F</td>
</tr>
<tr>
<td>MID5-27F</td>
<td>F5</td>
<td>CGT ATC GCC TCC CTC GCG CCA TCAG TCA GAC ACGA</td>
<td>27F</td>
</tr>
<tr>
<td>MID1-1435R</td>
<td>R1</td>
<td>CTA TGC GCC TTG CCA GCC CGC TCAG ACG AGT GCGT</td>
<td>1435R</td>
</tr>
<tr>
<td>MID2-1435R</td>
<td>R2</td>
<td>CTA TGC GCC TTG CCA GCC CGC TCAG ACG CTC GACA</td>
<td>1435R</td>
</tr>
<tr>
<td>MID3-1435R</td>
<td>R3</td>
<td>CTA TGC GCC TTG CCA GCC CGC TCAG AGA CGC ACTC</td>
<td>1435R</td>
</tr>
<tr>
<td>MID4-1435R</td>
<td>R4</td>
<td>CTA TGC GCC TTG CCA GCC CGC TCAG AGC ACT GTAG</td>
<td>1435R</td>
</tr>
<tr>
<td>MID5-1435R</td>
<td>R5</td>
<td>CTA TGC GCC TTG CCA GCC CGC TCAG ATC AGA CACG</td>
<td>1435R</td>
</tr>
</tbody>
</table>

*red font = Adaptor sequence, blue font= Key sequence, green font= MID sequence.

²27F = 5’-AGA GTT TGA TCM TGG CTC AGA ACG; 1435R= 5’-CGA TTA CTA GCG ATT CCR RCT TCA
Table 3-2: Sample information and primer combinations used to construct the library for circular consensus sequencing (CCS).

<table>
<thead>
<tr>
<th>Origin of rodent</th>
<th>Rodent speciesa</th>
<th>DNA ID</th>
<th>Barcoded primer IDb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tlhavekisa</td>
<td>M. coucha</td>
<td>R2</td>
<td>F1R1</td>
</tr>
<tr>
<td>Gottenburg</td>
<td>M. natalensis</td>
<td>R171</td>
<td>F1R2</td>
</tr>
<tr>
<td>Gottenburg</td>
<td>M. natalensis</td>
<td>R172</td>
<td>F1R3</td>
</tr>
<tr>
<td>Tlhavekisa</td>
<td>M. coucha</td>
<td>R6</td>
<td>F1R4</td>
</tr>
<tr>
<td>Hlalakahle</td>
<td>M. natalensis</td>
<td>R98</td>
<td>F1R5</td>
</tr>
<tr>
<td>Tlhavekisa</td>
<td>M. natalensis</td>
<td>R84</td>
<td>F2R1</td>
</tr>
<tr>
<td>Tlhavekisa</td>
<td>M. coucha</td>
<td>R5</td>
<td>F2R2</td>
</tr>
<tr>
<td>Tlhavekisa</td>
<td>M. coucha</td>
<td>R11</td>
<td>F2R3</td>
</tr>
<tr>
<td>Tlhavekisa</td>
<td>M. coucha</td>
<td>R12</td>
<td>F2R4</td>
</tr>
<tr>
<td>Gottenburg</td>
<td>M. natalensis</td>
<td>R19</td>
<td>F2R5</td>
</tr>
<tr>
<td>Gottenburg</td>
<td>M. natalensis</td>
<td>R29</td>
<td>F3R1</td>
</tr>
<tr>
<td>Gottenburg</td>
<td>M. natalensis</td>
<td>R30</td>
<td>F3R2</td>
</tr>
<tr>
<td>Gottenburg</td>
<td>M. natalensis</td>
<td>R31</td>
<td>F3R3</td>
</tr>
<tr>
<td>Gottenburg</td>
<td>M. natalensis</td>
<td>R20</td>
<td>F3R4</td>
</tr>
<tr>
<td>Manyeleti</td>
<td>M. coucha</td>
<td>R21</td>
<td>F3R5</td>
</tr>
<tr>
<td>Manyeleti</td>
<td>M. coucha</td>
<td>R179</td>
<td>F4R1</td>
</tr>
<tr>
<td>Manyeleti</td>
<td>M. coucha</td>
<td>R53</td>
<td>F4R2</td>
</tr>
<tr>
<td>Manyeleti</td>
<td>M. coucha</td>
<td>R61</td>
<td>F4R3</td>
</tr>
<tr>
<td>Hlalakahle</td>
<td>M. natalensis</td>
<td>R95</td>
<td>F4R4</td>
</tr>
<tr>
<td>Hlalakahle</td>
<td>M. natalensis</td>
<td>R74</td>
<td>F4R5</td>
</tr>
<tr>
<td>Hlalakahle</td>
<td>M. natalensis</td>
<td>R75</td>
<td>F5R1</td>
</tr>
<tr>
<td>Hlalakahle</td>
<td>M. natalensis</td>
<td>R78</td>
<td>F5R2</td>
</tr>
<tr>
<td>Hlalakahle</td>
<td>M. natalensis</td>
<td>R99</td>
<td>F5R3</td>
</tr>
<tr>
<td>Manyeleti</td>
<td>M. coucha</td>
<td>R159</td>
<td>F5R4</td>
</tr>
<tr>
<td>Manyeleti</td>
<td>M. coucha</td>
<td>R177</td>
<td>F5R5</td>
</tr>
</tbody>
</table>

a Species identification was confirmed by molecular typing of the mitochondrial cytochrome b gene (Russo et al., 2006)
b Primer combination used to amplify each sample; See Table 3-1.

3.3.4 Sequence analysis

Binning, trimming and filtering of sequence data obtained was conducted using the Pacific Biosciences software according to the set sequence size range and 99% precision. Reads were then analysed using the Ribosomal Database Project (RDP) 16S classifier (Cole et al., 2009) to classify sequence reads to the genus level with a 95% confidence interval. Filtered data were then analysed against the NCBI BLASTn 16S microbial database using the command line application to ascertain the identity of sequences. Results from BLASTn were filtered to a minimum length of 1275 bp and 98% identity in Microsoft Excel (Gall et al., 2016). Sequence reads that fell below 98% identity were reported at the genus level, while reads with an identity of 98% and above were reported at the species level (Bonnet et al., 2014, Budachetri et al., 2014, Jones et al., 2010). Operational taxonomic units (OTUs) that were less than 1% of the total number of sequences were grouped as ‘rare’ (Gall et al., 2016). Raw microbiome sequence data will be deposited at the sequence read archive at the National Center for Biotechnology information.

Consensus sequences of organisms of interest were extracted in CLC genomics workbench 9.5.1. (Qiagen). An alignment of sequences from Bartonella species along with homologous sequences from GenBank was created from within the workbench. Phylogenetic analyses of the 16S rRNA gene sequences was conducted using the Maximum likelihood method in MEGA 7 (Kumar et al., 2016). An
alignment of representative *Brucella* and *Ochrobactrum* sequences was created within the workbench to determine sequence differences between the two genera.

### 3.3.5 Statistical analysis

The microbial compositions in rodent blood were analysed using the community ecology package vegan version 2.5-2 (Oksanen et al., 2016) in R studio version 1.1 (R Core Team, 2013). Alpha diversity rarefaction curves were plotted to calculate the mean species diversity of bacteria in all rodents from the different habitat areas. Principal Component Analysis (PCA) was done to quantify the compositional similarity or dissimilarity of the bacterial population in the blood of the rodents across the habitat areas where dimensions or principal components were constructed from the linear combinations of the variables. This was done using the package FactoMineR (Lê et al., 2008) for explanatory data analysis in R studio. The individuals that clustered together indicated they shared similar blood bacterial profiles while dispersion indicated variability. The close proximity of the variables on one plane of the projections indicated that the variables are positively correlated while positions of variables on the opposing quadrant of the projection indicated negative correlation. The correlation coefficient ($r$) between variables and the dimensions was extracted from the PCA and was considered significant if the $p$ value was lower than 0.05. Statistical tests of analysis of variance (ANOVA) and Kruskal-Wallis rank sum test were conducted to test if there were statistically significant differences in the bacterial composition between rodents from different habitats. The differences were considered to be statistically significant with a $p$ value $\leq 0.05$.

### 3.3.6 *Brucella* spp. PCR assay

To confirm the detection of *Brucella* spp. identified by microbiome data analysis, a PCR assay amplifying a 214 bp fragment of the 16S-23S rRNA internal transcribed spacer (ITS) of *Brucella* species using primers ITS66 (5'-ACA TAG ATC GCA GGC CAG TCA-3') and ITS279 (5'-AGA TAC CGA CGC AAA CGC TAC-3') was conducted as previously described (Keid et al., 2007). The reaction was performed in a final volume of 15 µl, containing 0.2 µM of each primer, 1 X DreamTaq Green PCR Master Mix (consisting of DreamTaq DNA polymerase, DreamTaq Green buffer, 0.2 mM each of dATP, dCTP, dGTP and dTTP and 2 mM MgCl$_2$) (ThermoFisher Scientific, South Africa), and 2 µl of DNA template. *Brucella abortus* field strain 544 (obtained from *B. abortus* OIE reference Lab, Teramo, Italy) was used as the positive control and PCR grade water was used as the no template negative control. Cycling conditions were 95°C for 3 minutes, 35 cycles of 95°C for 1 minute, 60°C for 2 minutes and 72°C for 2 minutes, followed by a final extension of 72°C for 5 minutes. PCR products were visualized by electrophoresis on a 2% agarose gel (1 X TAE buffer, pH 8.0) stained with ethidium bromide and viewed under UV light.
3.4 Results

3.4.1 Barcoded 16S rRNA gene amplification

PCR amplicons of the near full length 16S rRNA gene were obtained from 25 rodents collected from the three habitat areas. Figure 3-2 shows a representative gel picture of the 16S rRNA gene PCR products amplified using barcoded primers.

![Image of gel picture](image)

Figure 3-2: Example of 16S rRNA gene amplicons obtained from rodent samples. PCR amplicons of the 16S rRNA gene were amplified using barcoded Primers 27F and 1435R and separated on a 1.5% agarose gel (stained with ethidium bromide). Lanes c-q contain rodent PCR samples. Lane b contains the positive control *Anaplasma centrale* and lane r contains the negative water only control. Lane a contains the 1kb DNA molecular ladder (Invitrogen, ThermoFisher® Scientific, South Africa) with sizes in bp indicated on the left.

3.4.2 Sequence and statistical analysis

Sequence analysis generated 65,060 bacterial sequences. After filtering, the mean number of reads per sample was 2,602 sequences, which was sufficient to satisfy a rarefaction curve indicating that all OTUs were captured. The mean species diversity of bacterial populations detected in the rodents plotted using a rarefaction curve plateaued early in the sampling indicating that the majority of bacterial communities were well represented. The rarefaction curve is shown in Figure 3-3.

![Image of rarefaction curve](image)

Figure 3-3: Rarefaction curves of rodents from the Mnisi community. The mean species diversity of bacterial populations in rodent blood were plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied.
Excluding OTUs that grouped into the ‘rare’ and unclassified category, 17 OTUs that grouped into ten species within seven genera were obtained from the rodent samples. *Bartonella grahamii* comprised 29% of the total bacterial sequences in the rodents’ blood, while *Bartonella* sp. RF255YX comprised 23%, and several *Bartonella* spp. that fell below the cut-off point of 98% identity in BLASTn made up 12% of the sequences, respectively. Other organisms detected were *Pseudomonas* spp. (17.9%), *Ochrobactrum* spp. (7.2%), *Brucella* spp. (1%), *Anaplasma* spp. (0.5%), *B. henselae* (0.1%), *Ehrlichia* sp. (0.03%), and *Coxiella burnetii* (0.02%). The ‘rare’ group and unclassified OTUs made up 4.7% and 4.3% of the bacterial sequences in the blood of the rodents, respectively. Ten sequences of *A. centrale*, five sequences of *A. phagocytophilum* and two sequences of *A. marginale* were also detected from the rodents. Figure 3-4 shows the predominant bacterial taxa in the blood of *Mastomys* spp. in the Mnisi community.

![Figure 3-4](image)

*Figure 3-4:* Relative abundance of major taxa of bacteria in the blood of 25 *Mastomys* spp. rodents. Six to seven rodents from each habitat area were sampled. The village name is indicated under the rodent number.

Rodents from Hlalakahle and Tlhavekisa had the highest proportions of *Bartonella* spp. (~85%); while Gottenburg and Manyeleti (~ 45%) had much lower *Bartonella* loads. *Brucella* spp. and *Ochrobactrum* spp. were detected in rodents captured from Gottenburg and Manyeleti. On the other hand, *Ehrlichia* sp. was detected in rodent 75 from Hlalakahle and rodent 172 from Gottenburg while *Coxiella* sp. was only detected in rodent 12 from Tlhavekisa. *Anaplasma phagocytophilum* was detected in rodent 98 from Hlalakahle, while *A. centrale* and *A. marginale* were detected from rodent 20 from Gottenburg. Rodents from Manyeleti had the highest infection burden of *Pseudomonas* spp. (~34%) in their blood compared to rodents captured from other areas, however this is primarily due to two rodents with high
*Pseudomonas* burdens. Figure 3-5 shows the microbial composition of the blood from *Mastomys* spp. based on the sites of rodent capture.

**Figure 3-5:** Relative abundance of major taxa of bacteria in the blood of *Mastomys* spp. from different habitat areas: Hlalakahle and Gottenburg (urban/periurban area), Thhavekisa (communal rangeland), and Manyeleti Game Reserve (protected wild reserve).

Consensus sequences of organisms of interest in the blood microbiome of rodents revealed that the majority of *Bartonella* sequences based on the near full length 16S rRNA gene were 99% identical to *B. grahamii* strain as4aup (CP001562) while the remainder were 99% identical to *Bartonella* sp. strain RF255YX (AY993936). For *Ehrlichia* sp., Rodent 75 (Hlalakahle) and Rodent 172 (Gottenburg) had eight sequences with 98% identity to *Ehrlichia* sp. EH727 (AY309970) and *Ehrlichia* sp. Eh1669 (AY309969) from *Haemaphysalis* ticks in Japan. The sequences also had a 98% identity to *Ehrlichia* sp. Tibet (AF414399) from *Rhipicephalus* (Boophilus) *microplus* in China and *E. chaffeensis* (NR_074500). For *Brucella* spp., detected in rodents from Gottenburg, Rodent 30 had 566 sequences that had a 99% identity to *B. melitensis* (CP025822) and two sequences with 99% identity to *B. abortus* (CP007705) while Rodent 31 had 14 sequences with 99% identity to both *B. melitensis* (CP025822) and *B. abortus* (CP007705). For *Anaplasma* spp., Rodent 98 (Hlalakahle) had five sequences with 99% identity to *A. phagocytophilum* dog strain Dog2 (CP006618), rodent strain JM (CP006617) and human strains HZ (CP000235), HZ2 (CP006616) and Webster (NR044762). For *Ochrobactrum* spp. a total of 4,680 sequences were detected in rodents from Gottenburg and Manyeleti, Rodent 29 had 26 sequences, Rodent 30 had 1,793 sequences, Rodent 31 had 1,649 and Rodent 53 had 1,212 sequences with 99% identity to various species: *O. intermedium* (JN613288), *O. intermedium* (KT696500) and *O. pseintermedium* (DQ365922). For the whole data set, *Pseudomonas* spp. had 11,907 sequences with 98% identity to *P. extremaustralis* (NR114911) *Pseudomonas* sp. BFXJ-8.
(EU013945) and *P. fluorescens* (CP015638) and was detected in all the rodents except Rodent 29. Finally, for *Coxiella* sp., Rodent 12 (Tlhavekisa) had 15 sequences with a 99% identity to *C. burnetii* strain Schperling (CP014563) and *C. burnetii* strain Namibia (CP007555).

Principal component analysis (PCA) of the individuals factor map revealed rodents caught from Gottenburg (R30 and R31) and Manyeleti (R53) were distinct and clearly separated from the rest of the rodents. Similarities were observed in R19 from Gottenburg and R21 and R159 from Manyeleti which grouped in the bottom right quarter and mid-line of the Dim 1 X Dim 2 figure. The remaining rodents irrespective of the habitat area of capture formed one main cluster at the bottom left quarter to the mid-line of the figure. Figure 3-6 shows the PCA plot from individual rodents in the Mnisi communal area and the nearby wildlife area of Manyeleti.

![Figure 3-6: PCA plot of the individuals factor map of rodents from the Mnisi communal area and the nearby wildlife area of Manyeleti. Sample ids in black represent samples from Gottenburg, samples from Hlalakahle are in red, samples from Manyeleti are in green and samples from Tlhavekisa are in blue. Outliers R21, R53 and R159 are from Manyeleti Game Reserve (protected reserve), while R19, R30 and R31 are from Gottenburg (periurban area). Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 32.86% for the first axis and 25.49% for the second axis.](image)

The PCA plot of the variables factor map (Figure 3-7) showed that *Ochrobactrum* spp. expressed by the correlation coefficient (R=0.8) and unclassified Brucellaceae (R=0.9) were significantly and positively correlated to Dimension 1 alongside *Brucella* spp. (R=0.8) and *B. henselae* (R=0.8). On the other hand, *Bartonella* sp. RF255YX (R=0.5), *Bartonella* spp. (R=0.4) and unclassified Rhizobiales (R=0.5) were significantly and positively correlated to Dimension 2. *Bartonella* spp. (R= -0.7) and *Bartonella* sp. RF255YX (R= -0.4) were significantly and negatively correlated to Dimension 1. *Pseudomonas* spp. and unclassified Pseudomonadaceae (R= -0.9) were significantly and negatively correlated to Dimension 2. When linking the PCA plot of the individuals factor map to the variables...
factor map, we see that R30 and R31 (Gottenburg) and R53 (Manyeleti) were linked to the variables *Ochrobactrum* spp., unclassified Brucellaceae, *Brucella* spp. and *B. henselae* that are main positive contributors to Dimension 1. On the other hand, rodents that grouped into the main cluster were linked to the variables *Bartonella* sp. RF255YX, *Bartonella* spp., *B. grahamii* and unclassified Rhizobiales that were positive contributors to Dimension 2. Rodents R19 (Gottenburg), R21 and R159 (Manyeleti) were linked with the variables *Pseudomonas* spp. and unclassified Pseudomonadaceae.

![PCA plot of the variables factor i.e. bacterial populations from the blood of rodents in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 32.86% for the first axis and 25.49% for the second axis.](image)

There was no statistically significant difference between bacterial populations in rodents from Tlhavekisa and Hlalakahle using the Kruskal-Wallis rank sum test (P = 0.1). Statistically significant differences were observed using ANOVA in the bacterial populations in rodents captured from Gottenburg and Tlhavekisa (P < 0.007), Manyeleti and Hlalakahle (P = 0.05) and Gottenburg and Manyeleti (P < 0.005).

### 3.4.3 Phylogenetic analysis of *Bartonella* spp.

Phylogenetic analysis of eight *Bartonella* 16S rRNA gene sequences from representative rodents from the different habitat areas revealed that the rodents had co-infections with different species of *Bartonella*. Figure 3-8 shows the Maximum likelihood phylogenetic tree of *Bartonella* 16S rRNA gene sequences.
Figure 3-8: Maximum likelihood tree based on the Jukes-Cantor model showing the phylogenetic relationship between 16S rRNA gene sequences of previously identified *Bartonella* spp. and *Bartonella* spp. identified in the blood of rodents from different habitat areas in the Mnisi community. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 22 nucleotide sequences. There were 1285 positions in the final dataset.

### 3.4.4 *Brucella* spp. PCR assays

*Brucella* 16S-23S rRNA ITS PCR products were amplified from the DNA samples of Rodents 30 and 31 from Gottenburg, for which *Brucella* spp. were identified in the microbiome data. Figure 3-9 shows the gel picture of the ITS region PCR.
Figure 3-9: *Brucella*-specific amplification from rodents. Photograph of a 2% agarose gel (stained with ethidium bromide) of PCR products of the ITS region amplified using primers ITS66 and ITS279. Lanes d and e show amplification products from rodent DNA (R30 and R31). Lane h contained the positive control *B. abortus* strain 544 while lane i contained the negative control. Lane a and j contain the 100 bp DNA molecular ladder with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa). The other lanes contain samples from another study.

An alignment of consensus sequences of *B. melitensis*, *B. abortus* and *Ochrobactrum* spp. obtained from rodents R30 and R31 from Gottenburg and reference sequences of *B. melitensis*, *B. abortus*, *O. intermedium* and *O. pseuintermedium* from GenBank revealed 28 nucleotide positions where *Brucella* species differ from *Ochrobactrum* species. Table 3-3 shows nucleotide differences observed between the two related genera.
Table 3-3: Nucleotide positions showing differences discriminating between Brucella and Ochrobactrum species in an alignment with reference sequences from GenBank. Accession numbers of reference sequences are enclosed in brackets.

<table>
<thead>
<tr>
<th>*Position</th>
<th>*B.m (CP025822)</th>
<th>R30 B. m</th>
<th>R31 B. m</th>
<th>*B.a (CP007705)</th>
<th>R30 B.a</th>
<th>R31 B. a</th>
<th>*O.i (JN613288)</th>
<th>R30 Och</th>
<th>R31 Och</th>
<th>*O.p (DQ269522)</th>
</tr>
</thead>
<tbody>
<tr>
<td>311</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>319</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>330</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>351</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>506</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>513</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>518</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>609</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>690</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>705-706</td>
<td>TC</td>
<td>TC</td>
<td>TC</td>
<td>TC</td>
<td>TC</td>
<td>TC</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
</tr>
<tr>
<td>759-760</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>AA</td>
<td>TT</td>
<td>TT</td>
<td>TT</td>
<td>TT</td>
</tr>
<tr>
<td>761</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>763</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>1058-1059</td>
<td>GA</td>
<td>GA</td>
<td>GA</td>
<td>GA</td>
<td>GA</td>
<td>GA</td>
<td>AG</td>
<td>AG</td>
<td>AG</td>
<td>AG</td>
</tr>
<tr>
<td>1096-1097</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>TC</td>
<td>TC</td>
<td>TC</td>
<td>TC</td>
</tr>
<tr>
<td>1101-1102</td>
<td>AT</td>
<td>AT</td>
<td>AT</td>
<td>AT</td>
<td>AT</td>
<td>AT</td>
<td>CA</td>
<td>CA</td>
<td>CA</td>
<td>CA</td>
</tr>
<tr>
<td>1104-1105</td>
<td>TG</td>
<td>TG</td>
<td>TG</td>
<td>TG</td>
<td>TG</td>
<td>TG</td>
<td>CA</td>
<td>CA</td>
<td>CA</td>
<td>CA</td>
</tr>
<tr>
<td>1120</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>1123-1126</td>
<td>CGAA</td>
<td>CGAA</td>
<td>CGAA</td>
<td>CGAA</td>
<td>CGAA</td>
<td>CGAA</td>
<td>TCTC</td>
<td>TCTC</td>
<td>TCTC</td>
<td>TCTC</td>
</tr>
<tr>
<td>1129</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>1131-1132</td>
<td>AC</td>
<td>AC</td>
<td>AC</td>
<td>AC</td>
<td>AC</td>
<td>AC</td>
<td>GT</td>
<td>GT</td>
<td>GT</td>
<td>GT</td>
</tr>
<tr>
<td>1136</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>1168</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>1172-1173</td>
<td>CA</td>
<td>CA</td>
<td>CA</td>
<td>CA</td>
<td>CA</td>
<td>CA</td>
<td>TG</td>
<td>TG</td>
<td>TG</td>
<td>TG</td>
</tr>
<tr>
<td>1175-1176</td>
<td>AT</td>
<td>AT</td>
<td>AT</td>
<td>AT</td>
<td>AT</td>
<td>AT</td>
<td>TG</td>
<td>TG</td>
<td>TG</td>
<td>TG</td>
</tr>
<tr>
<td>1181-1182</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>CG</td>
<td>GA</td>
<td>GA</td>
<td>GA</td>
<td>GA</td>
</tr>
<tr>
<td>1214-1216</td>
<td>CCC</td>
<td>CCC</td>
<td>CCC</td>
<td>CCC</td>
<td>CCC</td>
<td>CCC</td>
<td>GTA</td>
<td>GTA</td>
<td>GTA</td>
<td>GTA</td>
</tr>
<tr>
<td>1221-1223</td>
<td>GGG</td>
<td>GGG</td>
<td>GGG</td>
<td>GGG</td>
<td>GGG</td>
<td>GGG</td>
<td>TAC</td>
<td>TAC</td>
<td>TAC</td>
<td>TAC</td>
</tr>
</tbody>
</table>

Numbers in bold in the first column indicate nucleotide positions where differences were observed.

*B.m = Brucella melitensis, B.a = Brucella abortus, O.i = Ochrobactrum intermedium, O.p = Ochrobactrum pseintermedium, Och = Ochrobactrum species.

3.5 Discussion

The multimammate mice, *M. natalensis* and *M. coucha*, are two morphologically similar mice that are well represented in Southern Africa (Skinner and Chimimba, 2005). Both species are commonly found in the veld and are regarded as synanthropic with humans, frequently being found in scrub fences around households and stores in African kraals (De Graaff, 1981). *Mastomys coucha* is the main reservoir host of flea-borne *Yersinia pestis*, the cause of bubonic plague (Davis, 1964), while *M.*
natalensis has been found to carry Lassa fever virus (Skinner and Chimimba, 2005). The close association that both species have with humans means that they are likely to serve as carriers of infection to man (Skinner and Chimimba, 2005). Results of the sequence analysis of bacterial populations in the blood of Mastomys spp. in the Mnisi community revealed that 64% of the bacterial sequences corresponded to Bartonella spp. and all rodents examined were infected with the organism. This is in agreement with previous studies that found Bartonella to be the dominant phylotype in rodent blood and fleas collected from rodents (Cohen et al., 2015) and from rodents in southern Indiana in the USA (Rynkiewicz et al., 2015), though the latter study found much higher numbers of Bartonella spp. which made up 97.8% of the total sequences in that study (Rynkiewicz et al., 2015). In this study, B. grahamii comprised 16.5% of the total bacterial sequences in M. natalensis and 12.4% of the bacterial sequences in M. coucha. It has previously been detected in Tatera leucogaster from the Sand Veld in South Africa (Pretorius et al., 2004) and in Rattus norvegicus from Nigeria (Kamani et al., 2013). Bartonella grahamii has been implicated as causing retinal occlusions in the eye of a man in France (Serratrice et al., 2003), in causing cat scratch disease in an immunocompromised patient in Finland (Oksi et al., 2013) and was also found to be the cause of bilateral neuroretinitis in a female patient in the Netherlands (Kerkhoff et al., 1999). The detection of the zoonotic pathogen B. grahamii from rodents in our study area is notable because of the zoonotic risks posed to human health in the community.

Bartonella sp. RF255YX made up 12.6% of the total bacterial sequences in the blood of M. coucha and 10.6% of sequences in M. natalensis. This organism has previously been detected in Rattus tanezumi flavipectus in Yunnan, China (Li et al., 2004).

Seventy-three sequences of Bartonella henselae were detected in rodent 30 (M. natalensis) from Gottenburg in this study. Antibodies to B. henselae have previously been detected in cats in South Africa (Kelly, 1996). It has also been detected by PCR in impounded cats in Gauteng Province (Trataris et al., 2012) and in HIV infected patients in Johannesburg (Frean et al., 2002). Bartonella henselae has been recognized for its zoonotic potential in causing endocarditis in humans (Hadfield et al., 1993), and has been implicated in causing occult infection in immunocompromised patients (Breitschwerdt et al., 2007). The pathogen was very recently detected in two AFI patients in Mnisi community (Simpson et al., 2018); thus, the detection of this organism in Mastomys spp. from the same area is notable.

Phylogenetic analysis confirmed the microbiome results: Mastomys spp. from all habitat areas had co-infections with different Bartonella spp. This is in agreement with previous studies that found a diverse assembly of Bartonella spp. in rodents from Southern China (Ying et al., 2002) and also from rodents in South-western Spain (Márquez et al., 2008). The phylogenetic findings are also in agreement with a recent study that reported the presence of three phylogenetic lineages of Bartonella species circulating in 80 Rhabdomys pumilio rodents captured from the Western Cape in South Africa (Hatyoka et al., 2019).
Seventeen percent of the bacterial sequences detected in the blood of *Mastomys* spp. in the Mnisi community corresponded to *Pseudomonas* species. *Pseudomonas* spp. are generally known as opportunistic pathogens that cause clinical infection only in rodents that are immunocompromised (Baker, 1998). It may be possible that detection of *Pseudomonas* spp. in rodent blood may be a result of contamination because the genus has been listed as a possible contaminant of high throughput 16S rRNA gene sequencing samples that have been amplified from negative (blank) controls (Salter et al., 2014). However, no detectable PCR band was observed in the negative control of the barcoded 16S rRNA gene PCR conducted in this study so the detection of *Pseudomonas* spp. from rodent blood may well be valid.

*Ochrobactrum* spp. made up 7.2 % of the sequences obtained from rodent blood and were detected in rodents 29, 30 and 31 (*M. natalensis*) from Gottenburg and rodent 53 (*M. coucha*) from Manyeleti. The *Ochrobactrum* genus comprises nine species with *O. intermedium*, *O. anthropi*, and *O. pseudintermedium* having previously been isolated from human clinical samples in France (Teyssier et al., 2005, Teyssier et al., 2007). *Ochrobactrum* spp. are generally regarded as opportunistic and nosocomial pathogens (Hagiya et al., 2013). The clinical manifestations of human infection with *Ochrobactrum* spp. have been found to mimic infection with *Brucella* spp. (Scholz et al., 2008a). In fact, *Ochrobactrum* spp. are phylogenetically closely related to *Brucella* spp. and 16S rRNA gene sequences of the two genera are 98.6% identical (Scholz et al., 2008a). An alignment of *Ochrobactrum* and *Brucella* species in this study indicated 28 nucleotide positions that discriminate between this closely related genera but it was interesting that *Ochrobactrum* spp. and *Brucella* spp. were detected in the same rodents in this study. The 16S rRNA gene alignment could not differentiate between *O. intermedium* (JN613288) and *O. pseuintermedium* (DQ365922) as the two species were identical.

*Brucella* spp. made up 1% of the total bacterial sequences obtained from *Mastomys* spp. and were detected in Rodents 30 and 31 (*M. natalensis*) from Gottenburg and Rodent 53 (*M. coucha*) from Manyeleti. *Brucella* species are gram negative facultative intracellular bacteria causing brucellosis, a major zoonosis in humans, livestock, wildlife and marine mammals (Corbel, 1997). The detection of *Brucella* DNA in Rodents 30 and 31 was validated using conventional PCR amplification of the ITS region. New *Brucella* strains phylogenetically distinct from previously described *Brucella* species have been reported from wild native rodents in Australia (Tiller et al., 2010), and very recently from rodents and shrews in Germany (Hammerl et al., 2017). Rodent-borne *B. neotomae* and *B. microti* have been detected in desert rats and in the common vole *Microtus arvalis* in the Czech Republic (Corbel, 1997, Scholz et al., 2008b). In South Africa, antibodies to *Brucella* spp. have been detected in wild ruminants (De Vos et al., 1969) and *B. abortus* has been isolated from the African buffalo (Gradwell et al., 1977). *Brucella abortus* has also been detected in cattle (Coetzer et al., 1994, Hesterberg et al., 2008) and very recently *B. melitensis*, a species usually associated with goats was isolated from cattle in South Africa (Kolo et al., 2018). The 16S rRNA gene sequence of *Brucella* species is identical (Gee et al., 2004, Vizcaíno et al., 2000); this gene is therefore not a useful tool to discriminate between *Brucella* species. This study serves the first report of the detection of *Brucella* spp. in rodents from an
An urban/periurban area closely situated to human settlements in South Africa. It elucidates that rodents are reservoirs of *Brucella* spp. in the Mnisi community and highlights the potential risks to human health.

*Anaplasma* spp. comprised 0.5% of the sequences obtained from the blood of *Mastomys* spp. in this study and was detected in *M. natalensis*. Ten sequences of *A. centrale* and two sequences of *A. marginale* were detected from rodent 20 from Gottenburg and five sequences of *A. phagocytophilum* were detected from rodent 98 from Hlalakahle. *Anaplasma marginale* is generally known to infect cattle as well as other wild ruminant species including American bison, antelope, deer, water buffalo and giraffes (Aubry and Geale, 2011). *Anaplasma centrale*, as well, has been recognized as a parasite of cattle but has also been detected in sika deer (Wu et al., 2015), sheep (Hosseini-Vasoukolaei et al., 2014), buffalo and wildebeest (Khumalo et al., 2016). The detection of *A. marginale* and *A. centrale* in wild rodents was curious as these species of *Anaplasma* are known to infect ruminants. Because the numbers of sequences obtained were quite low, we speculate on one possible scenario as follows: If an infected tick “injected” these pathogens into a rodent, where they failed to establish infection, their presence in the blood could still be detected due to the depth of the sequence coverage that the Pac Bio platform provides (Rhoads and Au, 2015). However, it has been reported that *M. natalensis* is a good host for *Haemaphysalis elliptica* and *Rhipicephalus simus*, with a high burden of infestation of larvae having been noted on these rodents (Horak et al., 2005). *Rhipicephalus simus* has been proven to be experimentally competent of transmitting *A. marginale* and *A. centrale* (Potgieter, 1987). Zoonotic *A. phagocytophilum* is known to have a much wider host range which includes humans, dogs, cattle, horses, reptiles, wild ruminants, insectivores and birds (Dugat et al., 2015). Rodents are generally considered as reservoir hosts of *A. phagocytophilum* with a high prevalence rate of infection detected in dusky footed wood rats and brush mice in the USA (Drazenovich et al., 2006, Foley et al., 2008). In Europe, *A. phagocytophilum* has been detected in more than nine rodent species (Stuen et al., 2013), from yellow necked mice in the Czech republic (Hulinska et al., 2004) to field voles in the UK (Bown et al., 2003). *Anaplasma phagocytophilum* sequences detected were closely related to *A. phagocytophilum* dog strain Dog2, rodent strain JM and human strains HZ, HZ2 and Webster. This study serves the first report of the detection of *A. centrale*, *A. marginale* and *A. phagocytophilum* from rodents in South Africa.

*Ehrlichia* sp. comprised 0.03% of the total sequences obtained from the blood of *Mastomys* spp. in Mnisi and was detected in *M. natalensis* from Rodent 75 from Hlalakahle and Rodent 172 from Gottenburg. The sequences had 99% identity to *Ehrlichia* sp. Ehf669 (AY309969) detected from *Haemaphysalis* ticks collected from dogs in Japan (Inokuma et al., 2004) and were closely related to *E. chaffeensis*, with 98% similarity to *E. chaffeensis* (NR_074500). *Ehrlichia chaffeensis* causes human monocytic ehrlichiosis (HME), an important tick-borne zoonosis in the USA transmitted by *Amblyomma americanum* ticks (Paddock and Childs, 2003). In Africa, serological evidence of human infection with *E. chaffeensis* has been reported from patients in Burkina Faso and Mozambique (Brouqui et al., 1994) and in an American traveller on return from travel to Mali, West Africa (Uhaa et al., 1992). Clinical infection with *E. chaffeensis* has also been detected and reported from undifferentiated AFI patients in
Cameroun (Ndip et al., 2009). In South Africa, a serological survey has detected antibodies to *E. chaffeensis* in dogs in the Free State province (Pretorius and Kelly, 1998). This study thus serves the first report of the detection of an *Ehrlichia* sp., similar to *E. chaffeensis*, in rodents in South Africa.

*Coxiella burnetii* was detected in only Rodent 12 (*M. coucha*) from Tlhavekisa. *Coxiella burnetii* causes Q fever, an infection that presents as an acute to chronic fever and pneumonia in humans, and also causes abortion in livestock (Vanderburg et al., 2014). It is transmitted by various tick species, and by direct contact in humans via inhalation of aerosols or eating infected animal products (Maurin and Raoult, 1999). Previous studies have detected *C. burnetii* from rodents in Spain and Central Italy (Barandika et al., 2007, Pascucci et al., 2015). Studies in Tanzania have implicated *C. burnetii* as the leading cause of non-malarial AFI in patients from resource-constrained communities (Crump et al., 2013, Prabhu et al., 2011). In South Africa, antibodies to *C. burnetii* have been detected in cattle (Gummow et al., 1987), and in wild dogs in the Kruger National Park (Van Heerden et al., 1995). The organism was also detected from 44% of *R. sanguineus* and 4% of *H. elliptica* ticks collected from dogs, cats and ruminants across four provinces in the country (Mtshali et al., 2015, Mtshali et al., 2017). *Haemaphysalis elliptica* is frequently found on *M. natalensis* in South Africa (Horak et al., 2005) and Kenya (Guerra et al., 2016). The probability of transmitting *C. burnetii* to humans via tick bites is, however, very rare (Eklund et al., 1947). This study serves the first detection of zoonotic *C. burnetii* from a wild rodent in a communal grazing area in the Mnisi community.

We did not detect any *Rickettsia* spp. in the blood of the *Mastomys* spp. in this study, this is contrary to a recent study reported by (Essbauer et al., 2018) in which the pathogenic rickettsial organisms, *R. conorii*, *R. massiliae*, *R. felis* and *R. helvetica*, were detected in wild rodents sampled across seven provinces in South Africa and Namibia. In a recent study in China (Ge et al., 2018), *Rickettsia* spp. were detected and comprised 60% of the bacterial communities in the spleen of wild mice and shrews. However, the authors of both of these studies extracted DNA from ear tissue samples and spleen, while we used DNA extracted from blood (Essbauer et al., 2018, Ge et al., 2018). Previous studies have stated that rickettsial pathogens are usually enclosed in the dermis, vascular endothelium and spleen (Bayliss et al., 2009, Hawley et al., 2007) so it is more likely to detect rickettsiae in such tissues rather than in blood.

The number of sequences of *Anaplasma*, *Ehrlichia* and *Coxiella* species detected in *Mastomys* spp. in this study was low and may not be statistically significant, however because these were genera of zoonotic and veterinary importance, it necessitated the mention of their detection. No significant similarities were observed when comparing bacterial communities based on the rodent species sampled: either *M. coucha* or *M. natalensis*. Statistical analysis of bacterial communities in the blood of *Mastomys* spp. revealed that the blood microbiome of rodents was diverse across the different habitat areas. Rodents from the communal grazing area, Tlhavekisa, on the most part shared similar bacterial profiles with rodents from Hlalakahle, an urban/periurban area. The physical distance from Tlhavekisa to Hlalakahle is about 8.1 km, further than the distance from Tlhavekisa to Gottenburg (about 7.7 km).
so it is not certain why rodents from these two areas shared similar blood bacterial profiles. On the other hand, rodents captured from Gottenburg, also an urban/periurban area, and rodents from Manyeleti Game Reserve, a protected wildlife reserve, had diverse blood bacterial profiles. This is in agreement with a previous study (Gall et al., 2017) where the authors speculated that geographic location plays a role in the ecological variation that is seen in bacterial populations in hosts and vectors. PCA revealed an association of positive correlations observed between B. grahamii, Bartonella sp. RF255YX, Bartonella spp. and unclassified Rhizobiales as well as positive correlations between Ochrobactrum spp., unclassified Brucellaceae, Brucella spp. and B. henselae from the blood bacterial profile of rodents R30 and R31 from Gottenburg and R53 from Manyeleti. Overall, the sample numbers in our study were low, such that individual variation had a large impact on the statistical outcome.

Due to regulations by the Department of Agriculture, Forestry and Fisheries (DAFF) South Africa, we were not allowed to collect the ticks that were on the body of the rodents when they were captured. All ectoparasites were frozen with the rodent carcasses at -80°C. In ideal circumstances, we would have collected the ticks, allowed them to digest their blood meals, performed dissections and characterized the bacterial microbiome of the tick salivary gland and midgut. Prior studies conducted on rodents and their ectoparasites have underpinned how characterizing the bacterial populations in hosts and their arthropod vectors is important in understanding the abundance of vector-borne pathogens as well as the roles and interactions that exist within these bacterial populations (Cohen et al., 2015, Rynkiewicz et al., 2015). Unfortunately, we were unable to do so. A limitation of the study was that PCR grade water was not spotted on FTA cards to evaluate potential contaminants on cards; this would have been useful as a DNA extraction control.

In conclusion, this study provided a synoptic insight into the bacterial diversity of zoonotic pathogens in the blood of wild rodents in the Mnisi community. It revealed Mastomys spp. as important reservoirs of bartonellae in the Mnisi community, hosting a variety of species. Finally, it reports the detection of other important zoonotic pathogens: Brucella spp., Ehrlichia sp., A. phagocytophilum and C. burnetii from rodents in South Africa.
3.6 References


2014. *Anaplasma* infection in ticks, livestock and human in Ghaemshahr, Mazandaran Province, Iran. *Journal of Arthropod-Borne Diseases*, 8, 204-211.


INOKUMA, H., BEPPU, T., OKUDA, M., SHIMADA, Y. & SAKATA, Y. 2004. Detection of *Anaplasma phagocytophilum* in *Haemaphysalis* ticks recovered from dogs in Japan that is closely related to a novel *Ehrlichia* sp. found in cattle ticks from Tibet, Thailand, and Africa. *Journal of Clinical Microbiology*, 42, 1353-1355.


de la rétine révélant une infection à *Bartonella grahamii*. *La Revue de Médecine Interne*, 24, 629-630.


Chapter 4
Composition of the bacterial blood microbiome of dogs and cattle in the Mnisi community

4.1 Abstract
All over the world, dogs are nurtured and kept as domestic pets. In most parts of Sub-Saharan Africa, the majority of the dog population are owned by households yet wander about freely. Cattle on the other hand are prized stock known to generally improve the standard and quality of life of communal farmers. In the Mnisi community, livestock farming is by far the most significant agricultural activity and cattle form the vast majority of domestic livestock reared. Recent research in the area has detected tick-borne pathogens of zoonotic and economic importance from domestic dogs and cattle. Because of the close contact dogs and cattle share with humans, the aim of this project was surveillance of the bacterial populations in the blood of dogs and cattle in order to understand the role these domestic animals play as reservoirs of zoonotic and other important tick-borne pathogens in the Mnisi community. The 16S rRNA gene was amplified from genomic DNA extracted from the blood of 10 domestic dogs and nine cattle using barcoded primers. Purified PCR amplicons were submitted for circular consensus sequencing on the Pacific Biosciences platform. Sequencing of 16S rRNA gene amplicons from the dogs resulted in 30,340 bacterial sequences. Approximately 24% of the sequences from the blood of the dogs corresponded to *Ehrlichia canis*, 19.3% to *Anaplasma platys*, 14.8% to *Anaplasma* sp. ZAM dog, and 21.4% to *Achromobacter xylosoxidans*. Several species were represented by relatively small numbers of sequences as follows: 5% of sequences corresponded to *Mycoplasma haemocanis*, 0.3% to *A. phagocytophilum*, while 1.6% of the total sequences obtained from canine blood corresponded to other *Anaplasma* spp. A similar number of bacterial sequences, 34,559 were obtained from cattle. Of the total number of sequences obtained from cattle blood, 54% corresponded to *A. marginale*, 22.2% to *Anaplasma* sp. Mymensingh, 10.5% to *Anaplasma* spp., and 5.4% to *Anaplasma* sp. Dedessa. Species with low representation included: *Anaplasma* sp. Hadesa at 2.7%, *A. centrale* at 1.4%, *Bartonella* spp. at 0.5%, *A. platys* at 0.2%, *Anaplasma* sp. Saso at 0.2% and *A. phagocytophilum* at 0.01% of the total sequences obtained from cattle blood. Sequences of *Borrelia* sp., *Brucella* sp., *Bartonella bovis* and the novel pathogen *Ehrlichia minasensis* were also detected in cattle. The average number of sequence reads obtained per sample was 3,034 sequences for dogs and 3,839 sequences for cattle, sufficient to satisfy rarefaction curves that all OTUs were captured. In conclusion, this study reveals that domestic dogs and cattle in Mnisi community serve as reservoirs of zoonotic and other economically important *Anaplasma* spp. This study also serves as the first report of the detection of recently described *Anaplasma* species and *E. minasensis* in cattle and the pathogen *M. haemocanis* in dogs in South Africa.
4.2 Introduction

Dogs play a dominant role in the lives of humans, as the popular saying goes: they are man’s best friend. Because of the close relationship they share with humans, they can serve as sentinels of infection between humans, livestock and wildlife (Damborg et al., 2016, Macpherson, 2005). The majority of dogs in the Mnisi community are owned by households and are also free roaming (Conan et al., 2015). Recent research conducted in the area reported the molecular detection of *Anaplasma* spp. from domestic dogs (Kolo et al., 2016). In the Mnisi community cattle rearing is the main agricultural activity, and cattle are regarded as the most valuable livestock to communal farmers (Choopa, 2016). A study reported that 65% of cattle screened for tick-borne haemoparasites were positive for *Ehrlichia, Anaplasma, Theileria* and *Babesia* species (Choopa, 2016). People in this community have close interactions with both dogs and cattle, and these animals could serve as conduits of zoonotic pathogens.

The advent of next-generation sequencing technologies coupled with the advances made in bioinformatics analyses have led to the discovery of pathogens and their genes within and on the body of humans and companion animals (Barko et al., 2018, Deng and Swanson, 2015). The majority of microbiome studies that have been conducted on domestic animals have focused purely on the microbiome of the gastrointestinal tract in mice (Gu et al., 2013), dogs and cats (Hand et al., 2013, Handl et al., 2011, Ritchie et al., 2008, Suchodolski et al., 2008, Tun et al., 2012), and cattle (Dowd et al., 2008, Myer et al., 2016). Other studies conducted in dogs have characterized the microbiota of the upper and lower respiratory tracts and that of the skin (Ericsson et al., 2016, Hoffmann et al., 2014). Studies in cattle have characterized the microbiome of the respiratory tract (Gaeta et al., 2017, Holman et al., 2015, Lima et al., 2016, Nicola et al., 2017). The microbiome of the bovine mammary gland has also been described (Falentin et al., 2016, Le Loir, 2017, Oikonomou et al., 2014), while other studies have characterized the vaginal and uterine microbiota of cows (Rodrigues et al., 2015, Santos and Bicalho, 2012). However, there is little or no information about using next generation sequence technology to describe the bacterial composition of the circulating blood in domestic dogs and cattle. The aim of this study was to perform a survey of the bacterial populations in the blood of dogs and cattle using circular consensus sequencing on the Pacific Biosciences platform in order to understand the role these domestic animals play as reservoirs of zoonotic and other important pathogens in the Mnisi community.

4.3 Materials and Methods

4.3.1 Ethics approval

The study was conducted with approval number (V105-15) from the Animal Ethics committee (AEC) of the Faculty of Veterinary Science, University of Pretoria. Permission was obtained to carry out the research in terms of Section 20 of the Animal Diseases Act 1984, from the Department of Agriculture, Forestry and Fisheries (DAFF) South Africa with reference numbers: 12/11/1/1; 12/11/1/1/6.
4.3.2 Samples used for the study

A set of 10 DNA samples extracted from blood collected from domestic dogs and nine DNA samples extracted from cattle blood was used in the study. Blood samples were collected during a Health and Demographic Surveillance System (HDSS) programme that ran in the Mnisi community (Conan et al., 2015) and was stored in the Hans Hoheisen Wildlife Research Station Biobank, University of Pretoria, at the Orpen Gate, Kruger National Park. Blood samples from domestic dogs were collected from households in Hluvukani village between April to May 2012. Blood samples from cattle were collected at dip tanks in Seville A, Seville B, Hlalakahle, Thhavekisa and Gottenburg between April to September 2013. Figure 4-1 shows the map of the study area and the locations of blood sampling.

![Figure 4-1: Map of study area with red stars showing locations of blood sampling from dogs and cattle. Dark green represents protected areas where wildlife roam freely (map produced by Estelle Mayhew).](image)

4.3.2 PCR amplification and sequencing

DNA used in this study was extracted from 200 µl EDTA blood using the QIAamp DNA mini kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions and stored at -20°C. The V1-V8 variable regions (i.e. near full length) of the 16S rRNA gene was amplified from samples using barcoded sample specific primers as described in Chapter 3 of this thesis using five µl of DNA (approximately 100 ng of DNA) to make a final PCR reaction volume of 25 µl. Purified PCR products were couriered to the Genomics Sequencing Core at Washington State University, Pullman, WA, USA for Pacific Biosciences (PacBio, Menlo Park, CA) circular consensus sequencing (CCS). Table 4-1 shows the sample information of the library prepared for CCS.
Table 4-1: Sample and primer information on the library prepared for CCS.

<table>
<thead>
<tr>
<th>Origin of samples</th>
<th>DNA ID</th>
<th>Host animal</th>
<th>Barcoded primer ID*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hluvukani</td>
<td>D2</td>
<td>Dog</td>
<td>F1R1</td>
</tr>
<tr>
<td>Hluvukani</td>
<td>D25</td>
<td>Dog</td>
<td>F1R2</td>
</tr>
<tr>
<td>Hluvukani</td>
<td>D6</td>
<td>Dog</td>
<td>F1R3</td>
</tr>
<tr>
<td>Hluvukani</td>
<td>D26</td>
<td>Dog</td>
<td>F1R4</td>
</tr>
<tr>
<td>Hluvukani</td>
<td>D13</td>
<td>Dog</td>
<td>F1R5</td>
</tr>
<tr>
<td>Hluvukani</td>
<td>D14</td>
<td>Dog</td>
<td>F2R1</td>
</tr>
<tr>
<td>Hluvukani</td>
<td>D21</td>
<td>Dog</td>
<td>F2R2</td>
</tr>
<tr>
<td>Hluvukani</td>
<td>D24</td>
<td>Dog</td>
<td>F2R3</td>
</tr>
<tr>
<td>Hluvukani</td>
<td>D36</td>
<td>Dog</td>
<td>F2R4</td>
</tr>
<tr>
<td>Hluvukani</td>
<td>D28</td>
<td>Dog</td>
<td>F2R5</td>
</tr>
<tr>
<td>Hlalakahle</td>
<td>C5</td>
<td>Cattle</td>
<td>F4R2</td>
</tr>
<tr>
<td>Tlhavekisa</td>
<td>C10</td>
<td>Cattle</td>
<td>F4R3</td>
</tr>
<tr>
<td>Seville A</td>
<td>C13</td>
<td>Cattle</td>
<td>F4R4</td>
</tr>
<tr>
<td>Hlalakahle</td>
<td>C17</td>
<td>Cattle</td>
<td>F4R5</td>
</tr>
<tr>
<td>Tlhavekisa</td>
<td>C38</td>
<td>Cattle</td>
<td>F5R1</td>
</tr>
<tr>
<td>Seville B</td>
<td>C42</td>
<td>Cattle</td>
<td>F5R2</td>
</tr>
<tr>
<td>Hlalakahle</td>
<td>C47</td>
<td>Cattle</td>
<td>F5R3</td>
</tr>
<tr>
<td>Seville A</td>
<td>C91</td>
<td>Cattle</td>
<td>F5R4</td>
</tr>
<tr>
<td>Gottenburg</td>
<td>C96</td>
<td>Cattle</td>
<td>F5R5</td>
</tr>
</tbody>
</table>

* See Table 3-1 for primer information

4.3.3 Sequence analysis

Binning, trimming and filtering of sequence data obtained was conducted using the Pacific Biosciences software according to the set sequence size range and 99% precision. Reads were then analysed using the Ribosomal Database RDP 16S classifier (Cole et al., 2009) to classify sequence reads to the genus level using a 95% confidence interval. Sequence data was then blasted against the NCBI BLASTn 16S rRNA gene microbial database using the command line application to ascertain the identity of sequences. Sequence data was subsequently blasted against a local database created from Anaplasma spp. sequences downloaded from GenBank to ascertain precise assignment of Anaplasma spp. sequences within the microbiome data. Results from BLAST were filtered to a minimum length of 1275 bp and 98% identity in Microsoft Excel (Gall et al., 2016). Sequence reads that fell below 98% identity were reported at the genus level, while reads that were 98% and above in identity, where appropriate, were reported at the species level as described by (Bonnet et al., 2014, Budachettri et al., 2014, Jones et al., 2010). OTUs that were less than 1% of the total number of sequences per sample were grouped into a ‘rare’ category as described by (Gall et al., 2016). Sequence analysis was carried out using the CLC genomics workbench 9.5.1 (Qiagen). Raw microbiome sequence data will be deposited at the sequence read archive at the National Center for Biotechnology Information.

4.3.4 Statistical analysis

The bacterial compositions in dog and cattle blood were analysed using the community ecology package vegan version 2.5-2 (Oksanen et al., 2016) in R studio version 1.1 (R Core Team, 2013). Alpha diversity rarefaction curves were plotted to calculate the mean species diversity of bacteria in domestic dogs and cattle from the different cattle dip tanks. Principal Component Analysis (PCA) was done to quantify the compositional similarity or dissimilarity of the bacterial population in the blood of the dogs.
and cattle across the different dip tanks where dimensions or principal components were constructed from the linear combinations of the variables. This was done using the package FactoMineR (Lê et al., 2008) for explanatory data analysis in R studio. The individuals that clustered together indicated they shared similar blood bacterial profiles while dispersion indicated variability. The close proximity of the variables on one plane of the projections indicated that the variables were positively correlated while positions of variables on the opposing quadrant of the projection indicated negative correlation. The correlation coefficient (r) between variables and the dimensions was extracted from the PCA and was considered significant if the p value was lower than 0.05. Statistical tests of analysis of variance (ANOVA) and Kruskal-Wallis rank sum test were conducted to test for statistical significant differences in the bacterial composition in cattle blood from the different dip tanks. The differences were statistically significant with a p value < 0.05.

4.4 Results

4.4.1 Barcoded 16S rRNA gene amplification

PCR products of the near full length 16S rRNA gene were obtained from the 10 DNA samples from dogs and nine DNA samples from cattle using the barcoded primers. Figure 4-2 shows the representative gel picture of 16S rRNA gene PCR.

Figure 4-2: Example of 16S rRNA gene amplicons from dog and cattle samples. Photograph of a 1.5% agarose gel (stained with ethidium bromide) showing 16S rRNA gene PCR amplicons. Lanes c-u contain dog and cattle PCR samples. Lane b contains the positive control (A. centrale) and lane v contains the negative water only control. Lane a and w contain the 1kb DNA molecular ladder, with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

4.4.2 Sequence and statistical analysis

4.4.2.1 Dogs

PacBio CCS sequencing of 10 sets of 16S rRNA gene amplicons from dog samples yielded a total of 30,340 bacterial sequences. After filtering, the mean number of reads per sample was 3,034 sequences. The mean species diversity of bacterial populations detected in the dogs plotted using rarefaction curves plateaued early in the sampling for most of the samples indicating that majority of bacterial communities were well represented. Figure 4-3 shows rarefaction curves for the dog samples.
Sequences from dog 28 (D28) did not meet the rarefaction standard because its plot showed an increasing trend beyond the vertical line in Figure 4-3 indicating a greater sequencing depth was needed for this sample to capture appropriate representation of the OTUs.

**Figure 4-3**: Rarefaction curves of dogs from the Mnisi community. The mean species diversity of bacterial populations in canine blood were plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied.

Sixteen OTUs were obtained from dogs that grouped into 13 species within nine genera excluding OTUs grouped into the ‘rare’ category. A look at all the bacterial sequences detected in canine blood revealed that 23.8% of the sequences were classified as *E. canis*, 19.3% as *Anaplasma platys*, 14.8% as *Anaplasma* sp. ZAM dog, while 0.3% of the sequences from the blood corresponded to *A. phagocytophilum*. Other organisms detected were *Achromobacter xylosoxidans* (21.4% of the sequences), *Anaplasma* spp. (1.6%), ‘rare’ category (10.6%), *Mycoplasma haemocanis* (4.9%), *Achromobacter* sp. (2.1%) and *Ehrlichia* sp. comprised (1.1%) of the total sequences from canine blood. One sequence each corresponding to *A. centrale* was detected in the blood microbiomes of dog 21 and 28, while four sequences corresponding to *A. marginale* were detected from dog 36. Consensus sequences extracted from CLC genomics workbench revealed that *E. canis* sequences had 99% identity to *E. canis* strain YZ-1 (CP025749). Other *Anaplasma* species detected included 4,479 sequences with 99% identity to *Anaplasma* sp. ZAM dog strain (LC269823) in dog 36 and 5,846 sequences with 99% to *A. platys* ZAM dog (LC269822) that were distributed across the microbiome data from most of the dogs, but highly prevalent in dog 25. There were 41 sequences with 99% identity to *A. phagocytophilum* strain Webster (U02521) detected in dog 24, and 10 and 31 sequences with 99% identity to *A. phagocytophilum* strain Norway variant2 (CP015376) detected in dogs 26 and 28, respectively. *Achromobacter xylosoxidans* sequences (6,494) had 98% identity to *A. xylosoxidans* strain FDAQGOS_147 (CP014060), and was detected in all but two of the dogs. *Mycoplasma haemocanis* sequences (1,511) were 99% identical to *M. haemocanis* strain Illinois (CP003319), and were detected in dogs 2 and 13.
Dogs 21 and 2 had a proportion of *E. canis*, with 89% and 78.5% of the sequences, respectively, from their blood corresponding to this pathogen. Dog 25 had the proportion of *A. platys* sequences making up to (94.7%) of its total sequences while dog 36 had the proportion of *Anaplasma* sp. ZAM dog comprising (93.3%) of the sequences obtained from this animal. Dog 24 had the proportion of 2.1% of the total sequences obtained corresponding to *A. phagocytophilum* while in dog 28, 76% of the total sequences obtained corresponded to ‘rare’ organisms. Dog 14 had the highest burden of *A. xylosoxidans* comprising (76.5%) of the total sequences obtained from its blood followed by dog 24 with (66.5%) of sequences corresponding to *A. xylosoxidans*. Finally in dog 13, 43.7% of the total sequences obtained from this animal corresponded to *M. haemocanis*. Figure 4-4 shows the bacterial composition of the blood microbiome from dogs in the Mnisi community.

![Relative abundance of major taxa of bacteria in the blood of 10 domestic dogs in the Mnisi community.](image)

**Figure 4-4**: Relative abundance of major taxa of bacteria in the blood of 10 domestic dogs in the Mnisi community.

PCA of the individuals factor map revealed that Dogs 6, 14 and 24 shared some similarities due to their proximity on the bottom right quarter on the Dim 1 X Dim 2 figure. Dog 13 was located in the top right quarter of the graph and was closer to the trio of D6, D14 and D24. Dogs 2 and 21 also shared some similarities and grouped closely to each other on the top left quarter of the graph. Dogs 25 and 36 shared some similarities in the bottom left quarter of the graph. Dogs 26 and 28 were distinct from the other dogs. Figure 4-5 shows figure of PCA plot of dogs in the Mnisi community.
Figure 4-5: PCA plot of individuals factor map of dogs in the Mnisi community. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 29.90% for the first axis and 25.19% for the second axis.

The PCA plot of the variables factor map (Figure 4-6) showed that A. xylosoxidans and Achromobacter spp. expressed by the correlation coefficient (R= 0.8) were significantly and positively correlated to Dimension 1. On the other hand, E. canis and Ehrlichia spp. were significantly and positively correlated to Dimension 2 (R= 0.8). Anaplasma spp. was significantly and negatively correlated to both Dimension 1 (R= -0.6) and Dimension 2 (R= -0.7). When linking the PCA plot of the individuals factor map to the variables factor map, we can see that dogs D2, D21 and D26 were linked to the variables E. canis and Ehrlichia spp. that are main positive contributors to Dimension 2. On the other hand, D25 and D36 were linked to the variables Anaplasma spp., Anaplasma sp. ZAM dog and A. platys. Dogs D6, D13, D14 and D24 to the right of the graph, were linked to the variables of A. xylosoxidans and Achromobacter spp. that are positive contributors to Dimension 1. Dog D28 was linked to the variable ‘rare’ group.

Figure 4-6: PCA plot of the variables factor i.e. bacterial populations from the blood of dogs in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 29.90% for the first axis and 25.19% for the second axis.
4.4.2.2 Cattle

Sequencing of the 16S rRNA gene amplicons obtained from cattle resulted in 34,559 bacterial sequences. After filtering, the mean number of reads per sample were 3,839 sequences. The mean species diversity of bacterial populations in cattle blood calculated using rarefaction curves plateaued for the majority of the samples indicating that the majority of the bacterial communities were well represented. Figure 4-7 shows the rarefaction curves for species diversity in cattle.

![Rarefaction curves of species diversity in Mnisi cattle. The mean species diversity of bacterial populations in cattle blood was plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied.](image)

Twenty-two OTUs were obtained from cattle that grouped into 17 species and within seven genera excluding OTUs grouped into the ‘rare’ category. *Anaplasma marginale* made up 54% of the total bacterial sequences obtained from cattle blood followed by *Anaplasma* sp. Mymensingh with 22.2%, 10.5% of the sequences corresponded to *Anaplasma* spp. and 5.4% of sequences to *Anaplasma* sp. Dedessa. *Anaplasma* sp. Hadesa sequences made up 2.7%, *A. centrale* 1.4%, *A. platys* 0.2%, *Anaplasma* sp. Saso 0.2% and *A. phagocytophilum* sequences made up 0.01% of the total sequences obtained. Three percent of the total sequences corresponded to the rare category. Other organisms of interest detected in far lower numbers include *Bartonella bovis* (~0.4%), *Bartonella* sp. (~0.03%) and *Ehrlichia minasensis* (~0.02%). Sequence analysis also revealed the detection of one sequence each of *Brucella* sp., *Rickettsia* sp. and *Borrelia* sp. *Bartonella* sp. was detected in cattle 38 and 42, and *Rickettsia* sp. in bull 42, *Borrelia* sp. was detected in cow 5, *Ehrlichia* sp. was only detected in cow 13 while *Brucella* sp. was only detected in cow 17.

Consensus sequences revealed that 18,742 *A. marginale* sequences had >99% identity to *A. marginale* strain Veld from South Africa (AF414873). For the other *Anaplasma* spp., there were 7,675 sequences with >99% identity to *Anaplasma* sp. Mymensingh (MF576175), and 1,851 sequences had a >99% identity to *Anaplasma* sp. Dedessa (KY924886). There were 922 sequences with >99% identity to
Anaplasma sp. Hadesa (KY924884) while 65 sequences had a >99% identity to Anaplasma sp. Saso (KY924885), finally 82 sequences had a >99% identity to A. platys (LC269882). Three thousand six hundred and nineteen sequences corresponded to Anaplasma spp. with less than 98% identity and were assigned to the genus level. There were five sequences with 99% identity to A. phagocytophilum str. Norway variant2 (CP015376).

In cow 38, three Bartonella sequences had a 99% identity to Bartonella sp. BVS12 (KU859917) while bull 42 had 141 sequences with 99% identity to B. bovis (KM371095). In cow 5, the Borrelia sp. had 99% identity to Borrelia sp. (AB897891). In cow 13, eight sequences of Ehrlichia spp. had a 99% identity to E. minasensis strain UFMG-EV (JX629805) and three sequences had 99% identity to Ehrlichia sp. (KX57742). Finally, in cow 17 the Brucella sp. sequence was 99% identical to B. melitensis strain CIT31 (CP025822).

Cow 47 had the most number of A. marginale sequences making up 89% of the total sequences obtained from the blood followed by cow 13 with 75.7% and cow 5 with 73.7% of the overall sequences. Cow 10 had the highest number of sequences of Anaplasma spp. (20.5%) in its blood followed by cow 96 with 18.9% of the total sequences. Cow 10 also had the highest number of Anaplasma sp. Mymensingh sequences in its blood (61.5%) followed by cow 96 with 39.4% and bull 42 with 28.2% of the overall sequences respectively. Cow 17 had the highest number of A. centrale sequences making 20.5% of the total sequences obtained in this animal. Anaplasma phagocytophilum was detected in cattle 10 and 96 only. Figure 4-8 shows the composition of major taxa of bacteria in the blood of Mnisi cattle.
Principal component analysis of the individuals factor map revealed the trio of C13 from Seville A and C5 and C47 from Hlalakahle shared some similarities due to their proximity to each other on the Dim 1 X Dim 2 figure. C38 from Tlhavekisa and C91 from Hlalakahle also shared slight similarities and were placed to the right of the graph. All other cattle were distinct from each other with C17 from Hlalakahle clearly separated from the other animals. Figure 4-9 shows figure of PCA plot in cattle.

Figure 4-9: PCA plot of the individuals factor map of cattle in the Mnisi community. Sample id in black font represents a sample from Gottenburg, samples from Hlalakahle are in red, samples from Seville A are in green, a sample from Seville B is in blue and samples from Tlhavekisa are in turquoise. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 42.93% for the first axis and 24.20% for the second axis.

The PCA plot of the variables factor map (Figure 4-10) showed that *Anaplasma* sp. Mymensingh expressed by the correlation coefficient (R= 0.8) was significantly and positively correlated to Dimension 1 alongside *Anaplasma* spp. (R= 0.7). *Anaplasma* sp. Dedessa (R= - 0.8), *A. centrale* (R= - 0.8) and the ‘rare’ group (R= - 0.7) were significantly and negatively correlated to Dimension 1. *Anaplasma marginale* (R= - 0.9) was significantly and negatively correlated to Dimension 2. When linking the PCA plot of the individuals factor map to the variables factor map we see that C17 was linked with the variables *Anaplasma* sp. Dedessa, the ‘rare’ group and *A. centrale* that are positive contributors to Dimension 2. On the other hand, C5, C13 and C47 were linked to the variable *A. marginale* that was negatively correlated with Dimension 2. Cattle C10, C38 and C96 to the right of the graph were linked to the variables *Anaplasma* sp. Mymensingh and *Anaplasma* spp. that are positive contributors to Dimension 1. Individual C96 was also linked to the variable *Anaplasma* sp. Hadesa.
Figure 4-10: PCA plot of the variables factor i.e. bacterial populations from the blood of cattle in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 42.93% for the first axis and 24.20% for the second axis.

Statistically significant differences were observed in the bacterial populations in cattle blood from the different dip tanks. Microbiome data from C17 was excluded from the analysis to reduce bias since its blood bacterial profile was so clearly distinct from the other cattle. Statistically significant differences was observed between bacterial populations in cattle blood using ANOVA from Seville A and Hlalakahle (P < 0.002), Thhavekisa and Hlalakahle (P = 0.03), Seville A and Thhavekisa (P = 0.01), and Seville B and Gottenburg represented by (P = 0.01). Kruskal Wallis rank sum test did not reveal any statistically significant differences in bacterial populations in cattle blood from any of the dip tanks (P = 0.4).

4.5 Discussion

This study serves as the first report of the comprehensive description of the blood microbiome of dogs and cattle using NGS techniques based on the 16S rRNA gene in a rural community in South Africa. Anaplasma species made up approximately 36% of the detected sequences from the blood of dogs in the Mnisi community area and were the dominant genus from the microbiome analysis. The organisms were detected from 8/10 (80%) of the dogs tested. OTU clustering classified A. platys as comprising 19.3% of the sequences from canine blood, Anaplasma sp. ZAM dog made up 14.8% of the sequences and Anaplasma spp. that fell below the 98% identity cut-off value in BLAST made up 1.6% of the sequences. Anaplasma phagocytophilum corresponded to 0.3% of the total sequences obtained from canine blood. These findings indicate that the dogs in the Mnisi community had co-infections of these closely related Anaplasma species. Anaplasma platys causes infectious canine thrombocytopaenia (Harvey et al., 1978) and the brown dog tick R. sanguineus is thought to be responsible for its transmission (Sanogo et al., 2003). Anaplasma phagocytophilum causes canine granulocytic anaplasmosis (CGA) a disease in dogs characterized by fever, anorexia and depression (Stuen et al.,
In Europe, America and Asia, *A. phagocytophilum* is transmitted by ticks of the *Ixodes persulcatus*-complex (Stuen et al., 2013). In the Mnisi community, a study found that the most common ticks found infesting domestic dogs were *R. sanguineus*, *A. hebraeum* and *H. elliptica* (Mulder, 2014). Since *R. sanguineus* is known to occasionally feed on humans (Dantas-Torres, 2008), it is possible it could transmit *A. platys* and *A. phagocytophilum* to humans in Mnisi. *Anaplasma platys* has been recognized as an emerging zoonotic pathogen (Arraga-Alvarado et al., 2014, Breitschwerdt et al., 2014, Maggi et al., 2013b) and *A. phagocytophilum* is a known zoonotic agent (Dumler et al., 2005), therefore the detection of these pathogens in domestic dogs from Mnisi is notable. Consensus sequences of *A. platys*, *A. phagocytophilum* and *Anaplasma* sp. ZAM dog were also extracted to be used in the phylogenetic analysis of *Anaplasma* spp. to be discussed in Chapter 7 of this thesis.

*Anaplasma centrale* was also detected from two dogs and *A. marginale* from a single dog. Tick species mainly found on cattle *Rhipicephalus simus*, *R. microplus*, *R. decoloratus*, *R. evertsi evertsi* and *Hyalomma marginatum rufipes* are known as vectors that transmit *A. marginale* and/or *A. centrale* (*R. simus*) in South Africa (Potgieter, 1987). However, since dogs in the Mnisi community are free roaming and are often used to escort cattle to the dip tanks, it is possible these ticks may have fed on the cattle. Dogs can also be infested with *R. simus* (Nyangiwe et al., 2006) and *R. sanguineus* commonly found infesting dogs in Mnisi have also been stated to be capable of transmitting *A. marginale* (Kocan et al., 2004), so the detection of *A. centrale* and *A. marginale* in dogs in this study is likely an incidental finding due to the deep sequencing and the fact that the ticks feeding on these dogs carry these pathogens rather than these pathogens truly infecting the dogs.

*Ehrlichia canis* made up 24% of the detected sequences in the blood of Mnisi dogs and was detected in 6/10 (60%) of the dogs tested. In dogs 21, 2 and 26 it made up approximately 89%, 79% and 56% of the overall sequences obtained from their blood. *Ehrlichia canis* is the cause of canine monocytic ehrlichiosis (CME) a major disease of dogs with a worldwide distribution except in Australia and New Zealand (Harrus et al., 1997). In Southern Africa, the organism has been detected using molecular techniques in 3% of domestic dogs in South Africa (Matjila et al., 2008) in 19% of ticks collected from dogs and cats in the country (Mtschali et al., 2017) and in 13.5% of dogs screened in the Western Cape (Allan, 2017). In Namibia, using serological techniques, the organism was detected in 53.8% of dogs screened (Manyarara et al., 2015). Since the dogs sampled in this study were apparently healthy dogs that were not showing any clinical signs of the infection, it is presumed that the dogs were sub-clinically infected having overcome the acute phase of the infection (Waner et al., 1997). It is interesting to note that a negative correlation of infection was observed between *Anaplasma* spp. and *E. canis* infection in the dogs. This observation was supported by the PCA. Dogs that had high burdens of *E. canis* infection in their blood had lower burdens of *Anaplasma* spp. and vice versa. This phenomenon of bacterial interference between rickettsial species has been previously described in *Dermacentor* ticks (Gall et al., 2016, Macaluso et al., 2002). PCA also revealed a positive correlation association between *E. canis* and *Ehrlichia* spp. infection in dogs in the Mnisi community.
Achromobacter xylosoxidans made up 21.4% of the total bacterial sequences obtained from the blood of the domestic dogs in this study. The organism was detected in 8/10 (80%) of the dogs screened. *Achromobacter xylosoxidans* is a gram negative bacilli described as an emerging pathogen because of the opportunistic role it plays in immunosuppressed patients battling cystic fibrosis (Dupont et al., 2018). Bacteremia caused by *A. xylosoxidans* has only been reported from humans (Aisenberg et al., 2004, Dai et al., 2015, Duggan et al., 1996, Raghuraman et al., 2015). PCA revealed a positive correlation association between *A. xylosoxidans* and *Achromobacter* spp. infection in dogs. Thus, this study marks the first report of the detection of this emerging pathogen in domestic dogs in South Africa.

This study serves the first report of the hemotropic *Mycoplasma haemocanis* in dogs in South Africa. The organism made up approximately 5% of the total sequences obtained from the blood of dogs in this study. *Mycoplasma haemocanis* was detected in 2/10 (20%) of the dogs screened. In dogs 2 and 13, it made up 6.5% and 43.7% of the total sequences detected in their blood. Infection with the pathogen is usually associated with haemolytic anaemia (Compton et al., 2012). Recently around the world, hemotropic *Mycoplasma* spp. have been detected from 77.1% of dogs in Argentina (Mascarelli et al., 2016), in 1.9% of dogs tested in China (Zheng et al., 2017), in 4.5% of dogs in Italy (Ravagnan et al., 2017) and in 2.5% of dogs in Thailand (Kaewmongkol et al., 2017). It has been suggested that *R. sanguineus* may be the vector responsible for the transmission of *M. haemocanis* though this has not yet been proven (Novacco et al., 2010). *Mycoplasma* species such as *Mycoplasma ovis*, *M. haemofelis*, *M. suis*, *Candidatus* Mycoplasma haemohominis, and *Ca. M. haematoparvum* have been described as emerging zoonotic pathogens posing health risks in animals and humans (Maggi et al., 2013a). There is currently no information on the zoonotic potential of *M. haemocanis* with published reports of its detection only in dogs (Kaewmongkol et al., 2017, Mascarelli et al., 2016, Ravagnan et al., 2017, Zheng et al., 2017).

In a five month census study of the dog population that was carried out across four villages in the Mnisi community in 2013, (Conan et al., 2017) reported an owned dog population of 942 dogs from 2969 households. The same authors in a previous study also reported a population of 7.9 dogs per 100 people in the area (Conan et al., 2015). There are approximately 8500 households in Mnisi (Statistics, 2012) with 31% of them owning dogs (Berrian et al., 2016). A prior study has also stated that domestic dogs play a role as sentinels of infection to humans (Cleaveland et al., 2006). Thus, the detection of zoonotic pathogens in dogs in Mnisi in this study highlights the health risks posed to humans because of the close relationship shared between dogs and humans.

*Anaplasma* species were by far the dominant pathogens of the blood microbiome of cattle comprising 96.8% of the total bacterial sequences obtained. *Anaplasma marginale* alone made up 54% of the total bacterial sequences in the blood of cattle and was detected in eight out of the nine cattle. *Anaplasma marginale* infection in cattle is common in South Africa. The organism has been detected in eight out of the nine provinces in the country (Hove et al., 2018, Potgieter, 1979). *Anaplasma centrale* made up 1.4% of the total sequences in the blood of cattle and was detected in 3/9 (33%) of cattle tested.
Anaplasma centrale was recently corroborated as a separate species from A. marginale (Khumalo et al., 2018), and is the cause of a less virulent form of bovine anaplasmosis (Carelli et al., 2008). It was recently shown to be freely circulating in cattle and wildlife species in South Africa (Khumalo et al., 2016).

Other Anaplasma species made up 40.9% of the total sequences detected from the blood of cattle. Anaplasma sp. Mymensingh comprised 22.2% of the total sequences obtained in cattle. A novel Anaplasma species, it was recently detected from 13% of cattle in Bangladesh and is phylogenetically closely related to A. platys (Roy et al., 2018). Anaplasma sp. Dedessa, Anaplasma sp. Hadesa and Anaplasma sp. Saso made up 5.4%, 2.7% and 0.2% of the total sequences obtained from cattle in Mnisi community. All novel Anaplasma species, Anaplasma sp. Hadesa was recently detected in 12.5%, Anaplasma sp. Saso in 14.3% and Anaplasma sp. Dedessa in 5.6% of cattle in Ethiopia (Hailemariam et al., 2017). Zoonotic pathogens, A. platys and A. phagocytophilum occupied a small fraction of the bacterial population in the blood of cattle with 82 sequences of A. platys and five sequences of A. phagocytophilum detected. Due to the advancement of new sequencing technologies, many novel pathogens are being discovered and new species sequences have been deposited in Genbank. It is interesting to note that in this study, new species of Anaplasmata very recently described were detected. A negative correlation of infection was also observed between A. marginale, Anaplasma sp. Mymensingh and Anaplasma sp. Dedessa infection in cattle. This observation was supported by the PCA. Cattle that had a higher proportion of A. marginale infection in their blood had a lower proportion of Anaplasma sp. Mymensingh or Anaplasma sp. Dedessa and vice versa. This finding is in agreement with the bacterial interference phenomenon described among rickettsial species in the aforementioned studies (Gall et al., 2016, Macaluso et al., 2002). PCA also revealed positive correlations observed between Anaplasma sp. Mymensingh, Anaplasma sp. Hadesa and Anaplasma spp. and positive correlations associated with Anaplasma sp. Dedessa, A. centrale and the ‘rare’ group.

Other important bacterial pathogens made up a very small fraction of the sequences detected from the blood of Mnisi cattle. Bartonella bovis was detected in cattle 38 and 42. Bartonella bovis is known to cause bovine endocarditis (Cherry et al., 2009). The organism has been detected in cattle in France (Bermond et al., 2002, Maillard et al., 2007) and in the United States (Breitschwerdt et al., 2001). Bartonella bovis has also been isolated from cattle in Ivory Coast (Raoult et al., 2005), in Italy (Martini et al., 2008) and from cattle herds in Israel (Rudoler et al., 2014). The organism has also been detected in 6.8% of cattle tested in Poland (Welc-Falęciak and Grono, 2013). This study serves the first report of the detection of B. bovis in cattle in South Africa.

An Ehrlichia sp. was detected in only cow 13 and consensus sequences had a 99% to Ehrlichia minasensis. Ehrlichia minasensis is a novel ehrlichial species isolated from R. microplus (Cabezas-Cruz et al., 2016). The organism has been detected from R. microplus and cattle in Brazil (Aguiar et al., 2014, Carvalho et al., 2016) and is similar to the genotype of Ehrlichia species previously detected in
naturally infected cattle (Gajadhar et al., 2010) and deer in Canada (Lobanov et al., 2012). This is the first report of the detection of this novel pathogen in cattle in South Africa.

A *Brucella* sp. was detected in cow 17. The consensus sequence had 99% identity to *B. melitensis* strain CIT31. *Brucella melitensis* is a highly contagious emerging zoonotic pathogen that is usually associated with infections in sheep and goats (Corbel, 1997). In South Africa, *B. melitensis* have been detected from a human patient (Wojno et al., 2016) and very recently the organism was isolated from slaughter cattle in Gauteng Province (Kolo et al., 2018). This finding is in agreement with a recent study in Mnisi who reported a low seroprevalence of 1.4% in cattle, and a non-existence of *Brucella* seropositivity in goats and dogs in the study area (Simpson et al., 2018a).

A *Borrelia* sp. sequence was detected in cow 5 in this study. *Borrelia* species are blood-borne spirochetes that cause tick-borne relapsing fever (TBRF) a zoonotic vector-borne disease worldwide (Cutler, 2015). *Borrelia burgdorferi* sensu lato causes Lyme disease an infection described by the WHO as the fastest growing vector-borne zoonosis worldwide (Donohoe et al., 2015). It has been reported in over 60 countries and is endemic in North America, Europe and Asia (Jongejan and Uilenberg, 2004, Piesman and Eisen, 2008). In South Africa, *Borrelia* spp. have been isolated from *Ornithodorus* ticks (Zumpt and Organ, 1961) and more recently a novel *Borrelia* sp. was detected from African penguins in the Eastern and Western Cape Provinces (Yabsley et al., 2012). To the best of our knowledge, this is the first report of the detection of *Borrelia* sp. in cattle in South Africa. A prior study in the Mnisi community reported a cattle population of 12,000 with 15 cattle dip tanks (Musoke et al., 2015). The authors of that study also speculated on interactions between cattle and wildlife in the area resulting in spill over of zoonotic infections to livestock and humans (Musoke et al., 2015). Household dogs usually accompany their owners to the cattle dip tanks, thus there is a risk of transmission of ectoparasites and infections between dogs and cattle in Mnisi. In the survey for zoonotic diseases carried out in Mnisi by (Simpson et al., 2018b) veterinary staff, livestock farmers and herders, which the authors termed “dip-tanksters” 92.2% of individuals tested positive for zoonotic SFG *Rickettsia* spp. and *C. burnetti*. Thus, it shows there is a high risk of zoonotic disease transmission between livestock, dogs and humans in the area.

In conclusion, this study demonstrates a highly diverse blood microbiome of domestic dogs and cattle in the Mnisi community. It reports on the detection of recently described novel species in dogs and cattle. It confirms that dogs and cattle are reservoirs of *Anaplasma* species and other infectious pathogens in the community and elucidates the zoonotic risks posed to human health.
4.6 References


minasensis sp. nov., isolated from the tick Rhipicephalus microplus. *International Journal of Systematic and Evolutionary Microbiology*, 66, 1426-1430.


LE LOIR, Y. Microbiota of bovine udder and susceptibility to mastitis in dairy cows. Presented at Scientific MEM days: Journées Scientifiques MEM (Métaomiques et Ecosystèmes Microbiens), 2017 Paris, France.


WELC-FALECIAK, R. & GRONO, K. 2013. The first cases of *Bartonella bovis* infection in cattle from Central Europe. *Veterinary Microbiology*, 162, 954-956.


Chapter 5

Bacterial microbiome of *Amblyomma hebraeum* ticks collected from cattle in the Mnisi community

5.1 Abstract

Ixodid ticks are vectors of pathogenic organisms of humans and animals worldwide. *Amblyomma hebraeum*, an ixodid tick native to Southern Africa and known for its particular aggressiveness and readiness to bite humans, transmits *Rickettsia africae* the cause of African tick bite fever an emerging tick-borne zoonosis in humans. *Amblyomma hebraeum* has been found to be the most abundant tick in the Mnisi community, and in Southern Africa, cattle are its most preferred host. The aim of this study was to characterize the bacterial microbiome of *A. hebraeum* in order to gain an understanding of the microbial communities that reside in this important tick vector. This was achieved by screening salivary glands and midgut tissue of 348 adult *A. hebraeum* ticks collected from cattle across eight different dip tanks in the Mnisi community. The 16S rRNA gene was amplified from genomic DNA from salivary gland and midgut pools of *A. hebraeum* using barcoded primers. Purified PCR products were submitted for circular consensus sequencing on the Pacific Biosciences platform. Binning, trimming and analysis of sequence data resulted in 84,405 sequences from the salivary glands and 86,244 sequences from the midgut. The average number of reads was 3,836 from each salivary gland pool and 3,920 from each midgut pool satisfying rarefaction curves that all operational taxonomic units were captured. *Rickettsia africae*, *Rickettsia* spp., *Coxiella* spp. and *Ehrlichia ruminantium* were among organisms of interest detected from sequence analysis of *A. hebraeum* data. Overall, 83% of the total sequences obtained from the salivary glands and 81.7% of the sequences from the midgut pools corresponded to *R. africae*. Results of the sequence analysis were validated using a *Rickettsia* genus-specific qPCR assay that amplified the citrate synthase gene. In conclusion, this study showed *R. africae* as the dominant symbiont in the bacterial microbiome of *A. hebraeum* in the Mnisi community. The insight from this study provides a baseline data for future microbiome studies on *A. hebraeum* in Southern Africa.

5.2 Introduction

Ticks transmit a larger number of pathogenic organisms, including protozoa, bacteria and viruses, to domestic animals and wildlife worldwide than any other blood sucking arthropod vector (Colwell et al., 2011, Jongejan and Uilenberg, 2004) and are second only to the mosquitoes in the number of pathogens transmitted to humans (Jongejan and Uilenberg, 2004). *Amblyomma hebraeum*, an ixodid tick of ungulates native to Southern Africa is the reservoir host and vector of *R. africae* the cause of African tick bite fever (ATBF), an infection mostly detected in travellers to South Africa (Jensenius et al., 2003a). *Amblyomma hebraeum* is also the vector of *Ehrlichia ruminantium* the cause of heartwater an important tick-borne disease affecting ruminants in Africa (Uilenberg, 1983). A particularly aggressive tick, *A. hebraeum* is known to inflict numerous bites on exposed individuals (Jensenius et
al., 2003a). Most cases of ATBF are seen in international travellers returning from sub-Saharan Africa (Althaus et al., 2010, Fournier et al., 1998, Jensenius et al., 2003b). As stated in earlier chapters, the Mnisi community is located in a human/livestock/wild-life interface. Previous research in the area detected sequences for the zoonotic pathogens \textit{R. africae} and an \textit{R. felis}-like agent from ticks and fleas collected from domestic dogs (Kolo et al., 2016). A recent serological study conducted in the area reported that 63.4\% of patients with non-malarial acute febrile illness (AFI) showed a past exposure and 24.3\% of patients showed a recent exposure to spotted fever group (SFG) \textit{Rickettsia} (Simpson et al., 2018). The same study also screened veterinary staff, cattle farmers and herders that attend the cattle dip tanks in the area with 92.2\% of those individuals showing prior exposure to SFG \textit{Rickettsia} (Simpson et al., 2018).

The advent of next-generation sequencing (NGS) has led to the vast exploration of microbial communities in a fast, massive and economical manner (Greay et al., 2018). The first study of the tick microbiome using NGS was described by (Andreotti et al., 2011) and since then a large number of studies in this area have been published. Studies conducted on the Illumina platform have described bacterial communities in \textit{Ixodes ricinus} ticks in Europe (Bonnet et al., 2014, Carpi et al., 2011, Vayssier-Taussat et al., 2013) and \textit{I. pacificus} ticks in the USA (Swei and Kwan, 2016). Studies on the Ion Torrent platform have described bacterial communities in \textit{I. holocyclus} and \textit{I. ricinus} ticks in Australia (Gofton et al., 2015) and in \textit{Haemaphysalis} ticks in Malaysia (Khoo et al., 2016). Another study focused on the description of microbiome changes observed following blood feeding in \textit{Amblyomma maculatum} ticks in the USA (Menchaca et al., 2013). Studies using the Pacific Biosciences platform have examined the wild type and disrupted the bacterial microbiome of the salivary glands and midgut of \textit{Dermacentor andersoni} ticks in the USA (Gall et al., 2016, Gall et al., 2017).

There is very little information about the bacterial communities that exists within ixodid ticks in Africa using NGS approaches. There has been one published study describing the bacterial microbiota of \textit{R. sanguineus} collected from France, USA and Senegal (René-Martellet et al., 2017). The aim of this study was to use NGS on the Pacific Biosciences platform to characterise the bacterial microbiome of the salivary glands and midgut tissue of \textit{A. hebraeum} collected from domestic cattle. This was done to reveal the range of endosymbionts and pathogens found in this important tick vector in the Mnisi community.

5.3 Materials and Methods

5.3.1 Ethics approval

The animal ethics committee (AEC) of the Faculty of Veterinary Science, University of Pretoria approved the study with approval number: (V105-15). Section 20 approval in terms of the Animal diseases act of 1984 was obtained from the Department of Agriculture, Forestry and Fisheries (DAFF) South Africa with reference numbers: 12/11/1/1: 12/11/1/1/6.
5.3.2 Study site and sample collection

The Mnisi community, described in Chapter 1 is located in Mpumalanga Province, South Africa; and is at the livestock/wildlife/human interface of the western boundary of the Kruger National Park (KNP). The community shares 75% of its boundaries with wildlife reserves which include the KNP, Andover Nature reserves, Manyeleti Game Reserve, Timbavati Game Reserve and Sabi Sands Game Reserve.

Three hundred and forty-eight adult *A. hebraeum* ticks were collected manually from domestic cattle at eight communal dip tanks including Clare A, Welverdiend A, Welverdiend B, Seville B, Utha, Shorty, Gottenburg and Eglington, during the summer months of November and December 2017. The underside of the animals were targeted for tick collection around the axillae and sternum as these are preferred attachment sites for *A. hebraeum* (Petney et al., 1987). Six ticks were collected per animal and collection was biased towards male ticks because it has been stated male ticks tend to move more frequently between hosts in the quest for mating possibilities (Lysyk, 2013). Ticks that were least engorged with blood were preferentially selected. Figure 5-1 shows a map of the study site.

![Map of study site](image)

*Figure 5-1: Map of study site with red stars indicating communal dip tanks where *A. hebraeum* was collected from cattle. Dark green indicates game reserves where wildlife roam freely (map produced by Estelle Mayhew).*

All ticks were identified to the species level using the relevant taxonomic tools (Walker et al., 2003). Ticks were pooled in groups of 12, according to the location of collection and kept for 72 hours to ensure digestion of the blood meal. Ticks were then surface sterilized by vigorous shaking in 5% bleach solution, then 5% ethanol solution followed by a final rinse with double distilled water as described by
(Scoles et al., 2005). Salivary glands and midgut tissue were carefully removed and placed into separate tubes containing 200 µl of phosphate buffered saline (PBS) and kept frozen at -20°C. Tick dissections were performed at the Hans Hoheisen Wildlife Research Station Laboratory of the University of Pretoria located at the Orpen gate of the Kruger National Park, South Africa. All dissection tools were sterilized between each dissection.

5.3.3 DNA extraction and 16S rRNA gene amplification

Genomic DNA was extracted from pooled *A. hebraeum* salivary gland and midgut tissue (29 pools each of salivary glands and midgut tissue) using the QiAamp DNA mini kit® (Qiagen) according to the manufacturer’s instructions and stored at -20°C for further analysis. Amplification of the 16S rRNA gene using barcoded universal primers 27F (5’-AGA GTT TGA TCM TGG CTC AGA ACG-3’) and 1435R (5’-CGA TTA CTA GCG ATT CCR RCT TCA-3’) (Lane, 1991, Turner et al., 1999) used genomic DNA from salivary glands and midgut tissue as described in Chapter 3 of this thesis using 1 µl of template DNA (approximately 100 ng of DNA). These primers generate a near full length 16S rRNA gene amplicon. Purified PCR products were couriered to the Washington State University Pullman USA Genomics Sequencing Core for Pacific Biosciences (PacBio, Menlo Park, CA, USA) circular consensus sequencing (CCS). Tables 5-1 and 5-2 shows the sample and primer information of PCR amplicons for the libraries constructed for CCS.

Table 5-1: Sample and primer information of PCR amplicons for library 1.

<table>
<thead>
<tr>
<th>Sample ID*</th>
<th>Dip tank of origin</th>
<th>Primer ID†</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Clare A</td>
<td>F1R1</td>
</tr>
<tr>
<td>M1</td>
<td>Clare A</td>
<td>F1R2</td>
</tr>
<tr>
<td>M2</td>
<td>Clare A</td>
<td>F1R4</td>
</tr>
<tr>
<td>S3</td>
<td>Clare A</td>
<td>F1R5</td>
</tr>
<tr>
<td>M3</td>
<td>Clare A</td>
<td>F2R1</td>
</tr>
<tr>
<td>S4</td>
<td>Welverdiend A</td>
<td>F2R2</td>
</tr>
<tr>
<td>M4</td>
<td>Welverdiend A</td>
<td>F2R3</td>
</tr>
<tr>
<td>S5</td>
<td>Utha</td>
<td>F2R4</td>
</tr>
<tr>
<td>M5</td>
<td>Utha</td>
<td>F2R5</td>
</tr>
<tr>
<td>M7</td>
<td>Shorty</td>
<td>F3R4</td>
</tr>
<tr>
<td>M9</td>
<td>Welverdiend B</td>
<td>F4R3</td>
</tr>
<tr>
<td>S10</td>
<td>Welverdiend B</td>
<td>F4R4</td>
</tr>
<tr>
<td>M11</td>
<td>Welverdiend B</td>
<td>F5R2</td>
</tr>
<tr>
<td>S12</td>
<td>Eglington</td>
<td>F1R3</td>
</tr>
<tr>
<td>M12</td>
<td>Eglington</td>
<td>F3R1</td>
</tr>
<tr>
<td>S13</td>
<td>Eglington</td>
<td>F3R2</td>
</tr>
<tr>
<td>M13</td>
<td>Eglington</td>
<td>F3R3</td>
</tr>
<tr>
<td>S14</td>
<td>Eglington</td>
<td>F3R5</td>
</tr>
<tr>
<td>M14</td>
<td>Eglington</td>
<td>F4R1</td>
</tr>
<tr>
<td>S15</td>
<td>Eglington</td>
<td>F4R2</td>
</tr>
<tr>
<td>S18</td>
<td>Gottenburg</td>
<td>F5R5</td>
</tr>
<tr>
<td>S8</td>
<td>Shorty</td>
<td>F5R1</td>
</tr>
<tr>
<td>M8</td>
<td>Shorty</td>
<td>F5R3</td>
</tr>
<tr>
<td>M10</td>
<td>Welverdiend B</td>
<td>F5R4</td>
</tr>
</tbody>
</table>

*S=Salivary gland pool, M= Midgut pool
†Refer to primer codes in Table 3-1
Table 5-2: Sample and primer information for PCR amplicons for library 2.

<table>
<thead>
<tr>
<th>Sample ID*</th>
<th>Dip tank of origin</th>
<th>Primer ID†</th>
</tr>
</thead>
<tbody>
<tr>
<td>M18</td>
<td>Gottenburg</td>
<td>F1R1</td>
</tr>
<tr>
<td>M19</td>
<td>Gottenburg</td>
<td>F1R3</td>
</tr>
<tr>
<td>S20</td>
<td>Utha</td>
<td>F1R2</td>
</tr>
<tr>
<td>M20</td>
<td>Utha</td>
<td>F1R4</td>
</tr>
<tr>
<td>S21</td>
<td>Utha</td>
<td>F1R5</td>
</tr>
<tr>
<td>M21</td>
<td>Utha</td>
<td>F2R1</td>
</tr>
<tr>
<td>S22</td>
<td>Utha</td>
<td>F2R2</td>
</tr>
<tr>
<td>M22</td>
<td>Utha</td>
<td>F2R3</td>
</tr>
<tr>
<td>S23</td>
<td>Utha</td>
<td>F2R4</td>
</tr>
<tr>
<td>M23</td>
<td>Utha</td>
<td>F2R5</td>
</tr>
<tr>
<td>S24</td>
<td>Utha</td>
<td>F3R1</td>
</tr>
<tr>
<td>M24</td>
<td>Utha</td>
<td>F3R2</td>
</tr>
<tr>
<td>S25</td>
<td>Utha</td>
<td>F3R3</td>
</tr>
<tr>
<td>M25</td>
<td>Utha</td>
<td>F3R4</td>
</tr>
<tr>
<td>S28</td>
<td>Seville B</td>
<td>F4R4</td>
</tr>
<tr>
<td>M28</td>
<td>Seville B</td>
<td>F4R5</td>
</tr>
<tr>
<td>S29</td>
<td>Seville B</td>
<td>F5R1</td>
</tr>
<tr>
<td>M29</td>
<td>Seville B</td>
<td>F5R2</td>
</tr>
<tr>
<td>S11</td>
<td>Welverdiend B</td>
<td>F3R5</td>
</tr>
<tr>
<td>S17</td>
<td>Eglinton</td>
<td>F4R2</td>
</tr>
<tr>
<td>M17</td>
<td>Eglinton</td>
<td>F4R3</td>
</tr>
<tr>
<td>S19</td>
<td>Gottenburg</td>
<td>F5R3</td>
</tr>
<tr>
<td>S26</td>
<td>Seville B</td>
<td>F5R4</td>
</tr>
<tr>
<td>S2</td>
<td>Clare A</td>
<td>F5R5</td>
</tr>
</tbody>
</table>

*S=Salivary gland pool, M= Midgut pool
†Refer to primer codes in Table 3-1

5.3.4 Sequence analysis of NGS data

Binning, trimming and filtering of sequence data obtained was conducted using the Pacific Biosciences software according to the set sequence size range and 99% precision. Reads were then analysed using the Ribosomal Database RDP 16S classifier (Cole et al., 2009) to classify sequence reads to the genus level using a 95% confidence interval. Filtered data were blasted against the National Center for Biotechnology information (NCBI) BLASTn 16S microbial database using the command line application to ascertain the sequence identities. Results from BLAST were filtered to a minimum length of 1250 bp and 98% identity in Microsoft Excel as described by (Gall et al., 2016). Sequence reads that fell below 98% identity were reported at the genus level, while reads that were 98% and above in identity were reported at the species level when appropriate (Bonnet et al., 2014, Budachetri et al., 2014, Jones et al., 2010). OTUs that were less than 1% of the total number of sequences obtained were grouped as ‘rare’ as described by (Gall et al., 2016). Raw microbiome sequence data will be deposited at the sequence read archive at the NCBI.

Consensus sequences for *R. africae* were extracted in CLC genomics workbench 9.5.1 (Qiagen). An alignment with homologous sequences from GenBank of *Rickettsia* species was created within the workbench. Phylogenetic analyses of the 16S rRNA gene sequences was conducted using the Maximum Likelihood method in MEGA 7 (Kumar et al., 2016).
5.3.5 Statistical analysis

The microbial compositions of the A. hebraeum salivary glands and midgut pools were analysed using the community ecology package vegan version 2.5-2 (Oksanen et al., 2016) in R studio version 1.1. (R Core Team, 2013). Alpha diversity rarefaction curves were plotted to calculate the mean species diversity in all tick tissue from the different dip tanks. Principal Component Analysis (PCA) was done to quantify the compositional similarity or dissimilarity of the bacterial population in the salivary glands and midgut tissue of A. hebraeum collected from cattle across the dip tanks where dimensions or principal components were constructed from the linear combinations of the variables. This was done using the package FactoMineR (Lê et al., 2008) for explanatory data analysis in R studio. The tissue pools that clustered together indicated they shared similar bacterial microbiomes while dispersion indicated variability. The close proximity of the variables on one plane of the projections indicated that the variables are positively correlated while positions of variables on the opposing quadrant of the projection indicated negative correlation. The correlation coefficient (r) between variables and the dimensions was extracted from the PCA and was considered significant if the p value was lower than 0.05. Statistical tests of analysis of variance (ANOVA) and Kruskal-Wallis rank sum test were conducted to test for statistically significant differences in the bacterial composition of the salivary glands and midgut tissue of A. hebraeum collected from the different dip tanks. The differences were considered statistically significant with a p value≤0.05.

5.3.6 Rickettsia detection from samples

The sequence data were verified with a Rickettsia genus-specific qPCR assay that amplified a 74 base pair fragment of the citrate synthase gene of all Rickettsia species as described by (Stenos et al., 2005). Primers CS-F (5´-TCG CAA ATG TTC ACG GTA CTT T-3´) and CS-R (5´-TCG TGC ATT TCT TTC CAT TGT G-3´) and the probe CS-P (5´-6-FAM-TGC AAT AGC AAG AAC CGT AGG CTG GAT G-TAMRA-1-3´) were used. Reactions were performed in a final volume of 20 µl comprising 10 µl of 1X Taqman® Universal PCR Master Mix (composed of AmpliTaq Gold® DNA Polymerase, Uracil-N-glycosylase, dNTPs with dUTP, ROX™ passive reference and optimized buffer components) (Thermofisher Scientific, South Africa), 200 nM of each primer and probe, 5 mM Magnesium Chloride, 7.24 µl of PCR grade water and 1.0 µl of template DNA. The reactions were run on a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA), with cycling conditions at 50°C for 2 minutes, 95°C for 10 minutes followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Positive and negative controls were included with each run. An R. africae sample from H. elliptica confirmed by sequence of the 16S rRNA gene (Kolo et al., 2016) was used as the positive control while PCR grade water was used as the no template negative control. The results of the assay were analysed using the StepOne Plus software version 2.2.
5.4 Results

5.4.1 Results of barcoded 16S rRNA gene amplification

Using the barcoded universal primers, PCR amplicons of the 16S rRNA gene were amplified from 24 and 23 out of 29 pools of salivary gland and midgut tissue, respectively. Figure 5-2 shows a representative gel picture of 16S rRNA gene amplicons obtained.

![Image of 16S rRNA gene amplicons](image)

**Figure 5-2**: Example of 16S rRNA gene amplicons from *A. hebraeum* salivary gland and midgut pools. Photograph of a 1.5% agarose gel (stained with ethidium bromide) of amplicons. Lanes b-i contain *A. hebraeum* PCR samples. Lane m contains the positive control (*A. centrale*) and lane n contains the negative water only control. Lane a and o contain the 1kb DNA molecular ladder with sizes in bp indicated on the left (Invitrogen, ThermoFisher® Scientific, South Africa).

5.4.2 Sequence and statistical analysis

5.4.2.1 Salivary gland pools

Sequencing of *A. hebraeum* salivary gland pools generated 84,405 bacterial sequences. After filtering, the mean number of reads from each pool was 3,836 bacterial sequences. Salivary gland pools 12 and 25 (S12 and S25) had less than five bacterial sequences and were excluded from the analysis. The mean species diversity of the bacterial populations detected in the salivary glands was calculated and plotted using a rarefaction curve which plateaued for all the samples indicating that the majority of the OTUs were well represented. The rarefaction curve from *A. hebraeum* salivary gland pools is shown in Figure 5-3.
Ten OTUs were obtained from salivary gland pools that grouped into seven species and within seven genera excluding OTUs grouped into the ‘rare’ category. The sequences were mostly *Rickettsia* spp. with 82.9% corresponding to *R. africae* and 6.8% to *Rickettsia* spp. The remaining sequences were classified as follows: 6% corresponded to *Ehrlichia ruminantium*, 1.8% to the ‘rare’ category and 1.4% to *Coxiella* spp. Other pathogens that were detected at levels of ≤ 1% of the sequences obtained in a given salivary gland pool included *Ehrlichia* sp., *Paenibacillus alvei*, *Serratia nematodiphila*, and *Escherichia fergusonii*. Seven sequences of *A. marginale* and a sequence of *E. minasensis* were also detected from the salivary gland pools. Figure 5-4 shows the bacterial composition of the microbiome of salivary glands of *A. hebraeum* ticks in the Mnisi community.
When comparing bacterial sequences detected in the salivary glands of *A. hebraeum* according to the dip tank of origin, the salivary gland pools from ticks collected at Welverdiend A had 93% of the sequences corresponding to *R. africae* while salivary gland pools of *A. hebraeum* collected from Utha had 73.2% of the sequences obtained corresponding to *R. africae*. Salivary gland pools from Utha also had the proportion of *E. ruminantium* sequences making up 18.1% of the total sequences obtained from ticks collected at the dip tank followed by salivary gland pools from Seville B which had 10.7% of the sequences obtained corresponding to *E. ruminantium*. Salivary gland pools from Gottenburg had the proportion of *Coxiella* spp. making up 5.1% of the total sequences obtained from the dip tank. *Paenibacillus alvei*, *S. nematodiphila* and *E. fergusonii* were only detected in salivary gland pools from Welverdiend B. Figure 5-5 shows the bacterial composition of *A. hebraeum* salivary gland pools according to the dip tank of origin.
Principal component analysis of the individuals factor map revealed that salivary gland pools S5 (Utha), S26 (Seville B), S10 and S11 (Welverdiend B) were distinct from the other salivary gland pools. Pools S21 and S24 from Utha shared some similarities and were placed in the top left quarter of the graph. The other salivary gland pools irrespective of dip tank of origin showed similarities and grouped into one main cluster at the bottom quadrant of the graph. Figure 5-6 shows PCA plot of salivary gland pools of *A. hebraeum* collected from dip tanks in the Mnisi community.

**Figure 5-5:** Relative abundance of important taxa of bacteria in the salivary glands of *A. hebraeum* from the dip tanks of origin in the Mnisi community.

**Figure 5-6:** PCA plot of the individuals factor map of the salivary gland pools of *A. hebraeum* collected from cattle in the Mnisi community. Sample IDs in black represent samples from Clare A, samples from Eglington are in red, green samples are from Gottenburg, blue samples from Seville B, samples in turquoise are from Shorty, samples in purple are from Utha, samples in grey are from Welverdiend A and samples in brown are from Welverdiend B. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 41.07% for the first axis and 33.51% for the second axis.
The PCA plot of the variables factor map (Figure 5-7) showed that *P. alvei*, *E. fergusonii*, *S. nematodiphila* expressed by the correlation coefficient (R= 0.8) and the ‘rare’ group (R= 0.9) were significantly and positively correlated to Dimension 1. On the other hand, *E. ruminantium* (R= 0.8) and *Ehrlichia* spp. (R= 0.7) were significantly and positively correlated to Dimension 2. *Rickettsia africae* (R= -0.8) and *Rickettsia* spp. (R= -0.7) were significantly and negatively correlated to Dimension 2. When linking the PCA plot of the individuals factor map to the variables factor map, we see that the main cluster of salivary gland pools on the bottom of the plot were linked to the variables *R. africae* and *Rickettsia* spp. that are negatively correlated to Dimension 2. On the other hand, S5, S21, S24 (Utha) and S26 (Seville B) were linked with the variables *E. ruminantium* and *Ehrlichia* spp. that are positive contributors to Dimension 2. Pools S10 and S11 (Welverdiend B) were linked to the variables *P. alvei*, *E. fergusonii*, *S. nematodiphila* and the ‘rare’ group that are positive contributors to Dimension 1.

![PCA plot of the variables factor map](image)

**Figure 5-7:** PCA plot of the variables factor i.e. bacterial populations from the salivary gland pools of *A. hebraeum* in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 41.07% for the first axis and 33.51% for the second axis.

Statistically significant differences were observed in some of the bacterial populations in salivary glands of *A. hebraeum* collected from the different dip tanks. Significant differences using ANOVA were observed between bacterial populations in the salivary glands between Utha and Welverdiend B (P < 0.003) and Welverdiend B and Seville B (P < 0.001). There were no statistically significant differences between the bacterial population of the salivary gland pools from Utha and Seville B using the Kruskal Wallis rank sum test (P = 0.2), Clare A and Welverdiend A (P = 0.1), Shorty and Eglington (P = 0.2) and Gottenburg and Shorty (P = 0.2).
5.4.2.2 Midgut pools

Sequence analysis generated 86,244 sequences from the midgut pools of *A. hebraeum*. After filtering, the mean number of reads obtained from the pools was 3,920 bacterial sequences. Midgut pool 7(M7) had less than five bacterial sequences and was excluded from the analysis. The mean species diversity of bacterial populations in the midgut pools plotted using a rarefaction curve plateaued in all the samples indicating that bacterial communities were well represented. The rarefaction curve from *A. hebraeum* midgut pools is shown in Figure 5-8.

![Rarefaction curve](image)

**Figure 5-8**: Rarefaction curve of *A. hebraeum* midgut pools from the Mnisi community. The mean species diversity of bacterial populations in midgut pools was plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied.

Seventeen OTUs were obtained from the midgut pools that grouped into 12 species that were within 10 genera excluding sequences grouped as ‘rare’. Looking at the midgut pools as a whole, 81.7% of the sequences corresponded to *R. africae*, 6.9% to *Rickettsia* species, 3.5% to the ‘rare’ group, 3.4% to *E. ruminantium*, 1.9% to *P. alvei*, and 1.2% to *Coxiella* species. Pathogens that were each represented by less than 1% of the total sequences (i.e. they are included in the rare category) in a midgut pool included *Ehrlichia* spp., *Shigella flexneri*, *S. nematodiphila*, *E. fergusonii*, *Pseudomonas plecoglossicida*, *Paenibacillus* spp., *Pseudomonas monteilii* and *Serratia marcescens*. Eight sequences of *A. marginale*, five sequences of *Anaplasma* sp. Mymensingh and one sequence of *Bartonella bovis* were also detected from the midgut pools. Figure 5-9 shows the bacterial microbiome composition of the midgut pools.
When comparing bacterial populations of the *A. hebraeum* midgut pools grouped according to the dip tank of origin, the pooled midgut sequences showed that Welverdiend A had 92.3% of sequences corresponding to *R. africae* followed by midgut pools from Shorty and Gottenburg with *R. africae* making up 91.9% and 91.4% of the total sequences, respectively. Midgut pools from Welverdiend B, on the other hand, had the proportion of *R. africae* sequences corresponding to 56.6% of the total sequences. Midgut pools from Seville B had the proportion of *E. ruminantium* sequences making up 11.4% of the total sequences followed by midgut pools from Utha with 7.9% of the sequences corresponding to *E. ruminantium*. Midgut pools from Eglington had the proportion of *Coxiella* spp. comprising 4.2% of the total sequences followed by midgut pools from Seville B with 2.7%. *Pseudomonas plecoglossicida, P. alvei, Pseudomonas monteilii, and S. marcescens* were only detected from midgut pools from Welverdiend B. Figure 5-10 shows the combined bacterial microbiome composition of midgut pools of *A. hebraeum* based on the dip tank of origin.
Figure 5-10: Relative abundance of important taxa of bacteria in the midgut pools of *A. hebraeum* from the dip tank of origin in the Mnisi community.

Principal component analysis of the individuals factor map showed that midgut pools from Welverdiend B (M10) and Eglington (M12 and M14) were distinct as well as midgut pools from Seville B (M29) and Utha (M5 and M24). All the other midgut pools irrespective of the dip tank of origin showed similarities by clustering together. Figure 5-11 shows PCA plot of midgut pools.
Figure 5-11: PCA plot of the individuals factor map of the midgut pools of *A. hebraeum* collected from cattle in the Mnisi community. Sample IDs in black represent samples from Clare A, samples from Eglington are in red, green samples are from Gottenburg, blue samples from Seville B, samples in turquoise are from Shorty, samples in purple are from Utha, samples in grey are from Welverdiend A and samples in brown are from Welverdiend B. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 46.68% for the first axis and 21.31% for the second axis.

The PCA plot of the variables factor map (Figure 5-12) showed that *P. alvei* expressed by the correlation coefficient (R= 0.5) and the ‘rare’ group (R= 0.8) were significantly and positively correlated to Dimension 1. *Rickettsia africæ* (R= -0.9) and *Rickettsia* spp. (R= -0.8) were significantly and negatively correlated to Dimension 1. *Ehrlichia ruminantium* (R= -0.7) on the other hand, was significantly and negatively correlated to Dimension 2. When linking the PCA plot of the individuals factor map to the variables factor map we see that the main cluster of midgut pools were linked to the variables *R. africæ* and *Rickettsia* spp. that were negatively correlated to Dimension 1. On the other hand midgut pools, M5 and M24 (Utha) and M29 (Seville B) were linked to the variable *E. ruminantium* which was negatively correlated to Dimension 2. Midgut pools M10 (Welverdiend B) and M14 (Eglington) were linked with the variable ‘rare’ group with pool M10 also linked with the variable *P. alvei*, while pool M29 (Seville B) was linked to the variable *Coxiella* spp.
Figure 5-12: PCA plot of the variables factor map i.e. bacterial populations from the midgut pools of *A. hebraeum* in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 46.48% for the first axis and 21.31% for the second axis.

Statistically significant differences was observed in the bacterial populations in midgut pools of *A. hebraeum* collected from cattle from four of the dip tanks. Significant differences were observed between bacterial populations in the midgut pools between Welverdiend B and Eglington using ANOVA (P <0.0007), Utha and Seville B (P < 0.0007), Seville B and Welverdiend B (P < 0.0004) and Eglington and Utha (P < 0.001). There was no statistically significant differences between bacterial populations in the midgut pools from the different dip tanks using the Kruskal-Wallis rank sum test (P= 0.4).

Based on the near full-length 16S rRNA gene (1251-1310 nt), we identified two 16S rRNA gene sequence variants of *R. africae*. Sequences from (M4, M8, M10, M28, S1, S17 and S23) were identical while sequence S18 differed from the others by 1 nt at position 1126. *Rickettsia africae* sequences in this study differed from the *R. africae strain* ESF-5 (CP001612) by 2 nt. Phylogenetic analysis of 16S rRNA gene sequences from *Rickettsia* showed that the *R. africae* sequences obtained from the *A. hebraeum* midgut and salivary gland pools in this study grouped with *R. africae* strain ESF-5. Figure 5-13 shows the phylogenetic tree of 16S rRNA gene *Rickettsia* sequences with representative sequences from the different dip tanks of origin.
Figure 5-13: Maximum likelihood tree of 16S rRNA gene sequences showing the phylogenetic relationship between *Rickettsia* species. *Ehrlichia chaffeensis* was used as the outgroup sequence. Representative *R. africae* sequences generated in this study are in bold. The percentage of the bootstrap (1000 replicates) are shown next to the branches. All positions containing gaps and missing data were eliminated. Evolutionary distances based on the Jukes Cantor model were computed in MEGA 7. *M*=Midgut pool; *S*= Salivary gland pool.

### 5.4.3 *Rickettsia* genus-specific qPCR assay

*Rickettsia* DNA was successfully amplified from 22/23 salivary gland pools and 17/22 midgut pools of *A. hebraeum* using the *Rickettsia*-genus specific qPCR assay that amplifies a 74 bp fragment of the citrate synthase gene for all *Rickettsia* species. This was done to validate results of the microbiome analysis. *Rickettsia* DNA could not be detected in salivary gland pool 29 and midgut pools (2, 13, 18, 19 and 24). *Rickettsia* DNA was detected from samples at a CT mean of ~21 cycles. Figure 5-14 shows the amplification plot of *Rickettsia*-genus specific qPCR assay.
5.5 Discussion

This study presents a comprehensive insight into the bacterial populations that exist in *A. hebraeum* ticks in the Mnisi community, South Africa obtained using NGS on the Pacific Biosciences platform. Three hundred and forty-eight ticks were sampled from domestic cattle across eight different cattle dip tanks that are run by the Mnisi community programme. Previous studies in the area have recognized *A. hebraeum* as the most abundant tick in the community either in its free living form on vegetation (Horak, 1995, Horak et al., 2011, Spickett et al., 1991), or on domestic cattle (Zumpt, 2009). In general, cattle have been reported to be the most important hosts of *A. hebraeum* in Southern Africa (Norval, 1979).

This study reports that the bacterial population of *A. hebraeum* in the Mnisi community based on the different tissue types sampled, salivary glands and midgut, was significantly dominated by the pathogen *R. africae*. *Rickettsia africae* is the cause of ATBF an acute AFI in humans characterized by headaches, inoculation eschars, neck pain and inflammation of the lymph nodes (Jensenius et al., 2003a). Approximately 82% of the sequences obtained from the salivary gland and midgut pools corresponded to *R. africae* and every pool was infected with the pathogen. These results are in agreement with the statement by (Ross et al., 2018) that *Rickettsia* spp. play a significant role as an important driver in the bacterial abundance that is present in ixodid ticks. Previous studies have found *R. africae* infection rates of between 50% to 100% in *Amblyoma* spp. (Kelly and Mason, 1991, Maina et al., 2014, Parola et al., 2001), it is therefore not surprising that that all tissue pools in this study were positive for *R. africae*. There is a range of Rickettssias that have been detected from previous tick microbiome studies: *Rickettsia bellii* making up to 82% of the bacterial microbiome in the salivary glands of *Dermacentor andersoni* (Gall et al., 2016) and *Rickettsia* spp. making up 89% and 99% of the bacterial sequences in
Ixodes scapularis (Rynkiewicz et al., 2015, Zolnik et al., 2016). *Rickettsia* also made up 35% of sequences obtained from *A. americanum* (Ponnusamy et al., 2014), and comprised 46% and 7% of sequences from *A. maculatum* midgut and salivary glands (Budachetri et al., 2014). Finally, *Rickettsia* made up to 10% of the bacterial sequences detected in *Rhipicephalus turanicus* (Lalzar et al., 2012). The findings from these aforementioned studies support the statement that specific rickettsiae have evolved mechanisms that are adapted for or are associated with particular tick species (Socolovschi et al., 2009).

In this study, the abundance of *R. africae* was mostly similar in both the salivary gland and midgut pools of *A. hebraeum* collected from five dip tanks of Clare A, Welverdiend A, Gottenburg, Eglington and Shorty. Significant differences in bacterial profiles were observed in some salivary gland and midgut pools from Utha and Seville B and in some midgut pools from Welverdiend B and Eglington. Maximum likelihood analysis of *R. africae* sequences obtained from the salivary gland and midgut pools of *A. hebraeum* in this study showed sequences grouping with *R. africae* reference strain ESF-5 isolated from *A. variegatum* collected from cattle in Ethiopia (Fournier et al., 2009), supported with a fairly low bootstrap value. The high prevalence of *R. africae* in the salivary gland and midgut pools of *A. hebraeum* in this study correlates with the finding of previous exposure that 63.4% of febrile patients in the Mnisi community and 92.2% of cattle farmers, veterinary staff and herders tested in the community had antibodies of SFG *Rickettsia* (Simpson et al., 2018). This shows that there is a high risk of transmission of *R. africae* to the human population in the Mnisi community given the highly aggressive nature of the *A. hebraeum* vector which is known to infli-ct numerous bites on exposed individuals (Jensenius et al., 2003a). PCA revealed an association of positive correlations observed between *R. africae* and *Rickettsia* spp. infection in the salivary glands and midgut pools of *A. hebraeum* in this study.

A recent study in South Africa reported the detection of pathogenic rickettsial organisms *R. conorii*, *R. massiliae*, *R. felis* and *R. helvetica* from wild rodents across seven provinces in the country and neighbouring Namibia (Essbauer et al., 2018). Thus indicating that there is a diversity of *Rickettsia* species in the country.

Six percent of the total sequences from the salivary gland pools and 3.4% of the sequences from the midgut pools corresponded to *E. ruminantium*. It was detected in 12/22 (54.5%) salivary gland pools and 11/22 (50%) midgut pools. The proportion of *E. ruminantium* sequences obtained was more in salivary gland and midgut pools from Utha and Seville B. *Ehrlichia ruminantium* is the cause of heartwater, a notifiable disease by the World Organisation of Animal Health that affects domestic and wild ruminants (Allsopp, 2015). Recent studies on tick-borne haemoparasites in cattle in the Mnisi community detected *E. ruminantium* in 5.8% of animals tested on the reverse line blot (RLB) hybridization assay (Choopa, 2016). Differences in the relative abundance of *E. ruminantium* were observed in salivary glands and midgut pools of *A. hebraeum* collected from the different cattle dip tanks with salivary gland pools from Utha having 18.1% of the total bacterial sequences corresponding to *E. ruminantium*. Midgut pools from Seville B had the proportion of *E. ruminantium* making up 11.4%
of the total sequences. These observations are in agreement with what was stated that a geographic location plays a determining factor in dissimilarities observed in the makeup of the microbiome (Gall et al., 2017). A negative correlation of infection was observed between E. ruminantium and R. africae infection in the salivary glands and midgut pools of A. hebraeum, this was supported by the PCA. Pools that had a higher proportion of R. africae infection had a lower proportion of E. ruminantium and vice versa. This finding is thus in agreement with the bacterial interference phenomenon described among rickettsial species in previous studies (Gall et al., 2016, Macaluso et al., 2002). PCA also revealed an association of positive correlations observed between E. ruminantium, Ehrlichia spp., Coxiella spp., P. alvei and the ‘rare’ group in the salivary glands and midgut pools of A. hebraeum.

In this study, 1.4% of the total sequences in the salivary gland and 1.2% of the sequences obtained from the midgut pools of A. hebraeum corresponded to Coxiella spp. It was detected in 19/22 (86%) of both the salivary gland and midgut pools. Coxiella-like endosymbionts (Coxiella-LEs) have been found to have a mutualistic association with ticks of the genera Rhipicephalus, Amblyomma and Ornithodorus (Bonnet et al., 2017, Duron et al., 2017, Klyachko et al., 2007, Lalzar et al., 2014, Machado-Ferreira et al., 2011). Abundance of Coxiella-LEs have been rare or not detected at all in other ticks like I. ricinus (Carpi et al., 2011), Dermacentor, Hyalomma and Antricola (Duron et al., 2017). Coxiella-LEs have evolved to be maternally transmitted endosymbionts and have been found to be more associated with female ticks, primarily colonising the ovaries and the Malpighian tubules (Duron et al., 2015, Lalzar et al., 2014, Lalzar et al., 2012). There have been reports of Coxiella-LEs detection from the salivary glands of A. americanum (Klyachko et al., 2007), A. cajennense (Machado-Ferreira et al., 2011), Ixodes ovatus, I. persulcatus and Haemaphysalis flav (Qiu et al., 2014). Since tick collection was biased toward male ticks in this study, it could explain the proportion of Coxiella sequences that was obtained. An aforementioned study has suggested that the abundance of Coxiella-LE in the salivary glands of A. americanum is likely to inhibit the horizontal transmission of disease causing pathogens (Klyachko et al., 2007).

Sixteen Anaplasma marginale, five Anaplasma sp. Mymensingh, and an E. minasensis and Bartonella bovis sequence were detected from pools of A. hebraeum. It is interesting that these organisms were also detected from cattle blood discussed in the previous chapter of this thesis. However, the fact that these organisms were detected in such low numbers from tissue pools could be linked to small remnants of the blood meal from cattle still present within the tissues of the ticks. Ticks that were also least engorged with blood were preferentially collected in this study. It has also been proven that R. simus, R. microplus, R. decolaratus, R. evertsi evertsi and Hyalomma marginatus rufipes are competent vectors and are more associated with the transmission of A. marginale (Potgieter, 1979).

Paenibacillus alvei, Pseudomonas plecoglossicida, P. monteilii, Serratia marcescens, and S. nematodiphila were detected from A. hebraeum pools from Welverdiend B. Shigella flexneri and Escherichia fergusonii were also detected from pools from Welverdiend B and Eglington. Overall, these pathogens occupied a minute proportion of the microbiome of A. hebraeum pools. Paenibacillus alvei
is a spore forming bacilli found in soil and commonly detected in honey bee colonies (Djordjevic et al., 2000). *Pseudomonas* spp. and *S. marcescens* are opportunistic pathogens that are environmental contaminants (Bogaerts et al., 2011, KhAnNA et al., 2013) while *S. nematodiphila* is an enteric pathogen associated with the intestine of the nematode *Heterorhabditoides chongmingensis* (Zhang et al., 2009). *Shigella flexneri* and *E. fergusonii*, both members of the family Enterobacteriaceae, are usually associated with human infection causing bacillary dysentery for *S. flexneri* (Philpott et al., 2000) and urinary tract infections and diarrhoea for *E. fergusonii* (Gaastra et al., 2014). Even though *A. hebraeum* ticks were surface sterilized before dissections, the possibility exists that these organisms were environmental contaminants. As previously mentioned, ticks in this study were surface sterilized with both bleach and ethanol. We also observed that the bacterial microbiome of the salivary glands and midgut tissue of *A. hebraeum* in this study exhibited a low bacterial diversity. This finding is in agreement with a recent study that found ticks surface sterilized with bleach for 30 seconds exhibited a low bacterial diversity in their internal microbiome as opposed to ticks surface sterilized with ethanol indicating that bleach is a more efficient agent in denaturing external bacterial DNA on ticks than ethanol (Binetruy et al., 2019).

In conclusion, this study provides an insight into the bacterial microbiome of *A. hebraeum* ticks. It also reveals the overwhelming dominance of the zoonotic pathogen *R. africae* as the principal symbiont of the bacterial microbiome of *A. hebraeum* in the Mnisi community.
5.6 References


HORAK, I. G. The relationships between ticks, hosts and the environment in the Kruger National Park, South Africa. International Conference on Tick-borne Pathogens, Kruger National Park (South Africa), 1995.

HORAK, I. G., GALLIVAN, G. J. & SPICKETT, A. M. 2011. The dynamics of questing ticks collected for 164 consecutive months off the vegetation of two landscape zones in the Kruger National Park


Chapter 6
Characterization of the blood microbiome of acute febrile illness patients

6.1 Abstract

Acute febrile illness (AFI) is defined as a sudden fever with no identified cause. Studies carried out in Sub-Saharan Africa have reported the over diagnosis of malaria and the evident underestimation of rates of AFI occurrence in hospital settings. In the Mnisi community, undifferentiated non-malarial AFI is among the most common presenting sign in patients that seek healthcare at the community clinics. Recent serological analysis of non-malarial AFI patients in the area found 63.4% had prior exposure to spotted fever group (SFG) rickettsiae and 38.3% to Coxiella burnetii, the cause of Q fever. Another test showed that 9.5% of AFI patients were PCR positive for zoonotic rodent-borne Bartonella species. The aim of this study was to provide a comprehensive description of the circulating blood microbiome of a set of AFI patients using NGS techniques. The 16S rRNA gene was amplified from genomic DNA from nine non-malarial AFI patients using barcoded primers. Purified PCR products were submitted for circular consensus sequencing on the Pacific Biosciences platform. A total of 13,725 bacterial sequences were obtained from the AFI patients, with an average of 1,513 sequence reads per sample, satisfying rarefaction curves that all operational taxonomic units were captured. Sequence analysis showed that opportunistic pathogens Herbaspirillum huttiense and Stenotrophomonas maltophilia comprised 27% and 15.1% of the total sequences obtained, respectively. Rickettsia africae, the cause of African tick-bite fever, made up 16% of the total sequences obtained and was detected in three of the nine AFI patients. Also noteworthy was the detection of the zoonotic bacterial pathogen Brucella melitensis from the blood of an AFI patient. In conclusion, this study demonstrated a diverse blood microbiome from AFI patients and reports the detection of R. africae, Rickettsia spp. and B. melitensis. It recommends that healthcare workers in the community consider rickettioses in the differential diagnosis of non-malarial AFI.

6.2 Introduction

The vast majority of pathogenic organisms causing disease in humans are zoonotic (Taylor et al., 2001). A meta-analysis of severe febrile illness in Low- and Middle-Income countries showed that AFI is understudied in Africa, with relatively few published studies examining aetiology (Prasad et al., 2015). From those that have been done, we know that AFI in sub-Saharan Africa has links to zoonotic causes (Simpson et al., 2018). For example, R. africae was detected in 6% of non-malarial AFI patients in Cameroon (Ndip et al., 2004), while 7% of AFI patients in western Kenya were found to be positive for Rickettsia felis (Maina et al., 2012). Bartonella species have been detected in AFI patients in the Congo and Tanzania (Hercik et al., 2017, Laudisoit et al., 2011) and brucellosis and leptospirosis was documented in AFI patients in eastern Kenya and Tanzania (Biggs et al., 2011, Njeru et al., 2016). In the Mnisi community, recent research demonstrated that 24.1% of AFI patients were positive for Rickettsia spp., 2.3% to C. burnetii, and 6.8% to Leptospira spp. on serology, while 9.5% of patients
were positive for *Bartonella* species based on PCR testing (Simpson et al., 2018). A recent review of febrile illnesses in Africa listed brucellosis, rickettsioses, leptospirosis, Q fever and relapsing fever as important bacterial zoonotic diseases on the continent (Maze et al., 2018). The close relationship shared between humans, livestock and wildlife in rural villages and urban centres on the African continent has also been postulated as being a significant driver for the high occurrence of zoonotic diseases (Maze et al., 2018).

Next-generation sequencing (NGS) techniques offer a solid alternative to bacterial culture in studying the diversity of bacterial communities in different organs and systems within the human body (Whittle et al., 2019). Surprisingly, a recent study described a diverse array of bacteria in the blood of apparently healthy blood donors in France (Paisse et al., 2016). Another study identified infectious agents in the blood of patients diagnosed with sepsis (Grumaz et al., 2016), while other studies compared the diversity of pathogens resident in the blood of septic pancreatitis patients with blood collected from healthy volunteers (Gosiewski et al., 2017, Li et al., 2018). NGS studies have also been reported that compared the bacterial profile of the circulating blood from patients with cardiovascular disease with healthy individuals (Dinakaran et al., 2014, Rajendhran et al., 2013). These NGS approaches have not been brought to bear on the blood of AFI patients in South Africa to examine bacterial diversity which, in turn could shed light on potential zoonotic and/or emerging pathogens. The aim of this study was to use NGS techniques to characterise the bacterial microbiome of the circulating blood of a set of AFI patients in the Mnisi community.

### 6.3 Materials and Methods

#### 6.3.1 Ethics approval

The study was approved by the Human ethics committee of the Faculty of Health Sciences, University of Pretoria with approval number 152/2016.

#### 6.3.2 Study area and sample collection

The Mnisi community, described in Chapter 1, is located in Mpumalanga Province, South Africa. Blood samples from 74 patients with clinical signs of non-malarial AFI fever (>37.5°C) presenting at the community clinic in Hluvukani were made available to the project by the National Institute of Communicable Diseases (NICD), Johannesburg, South Africa. Patient blood samples were collected from adults that presented with a documented axillary temperature, or a history of fever within the last 72 hours. Patients with AFI were assessed by the clinic staff, a routine malaria smear was done and, if negative, the patients were referred to the study nurse for enrolment into the study. If the patient agreed to participate, a questionnaire related to his/her contact with animals, presence at dip tanks and tick bites were then completed; and two blood tubes for acute serology and molecular tests were taken. At the NICD, blood samples were tested for *Brucella* (serology), leptospirosis (serology), Q fever (serology and PCR), rickettsioses (serology), bartonellosis (PCR) and arboviruses (West Nile, Rift Valley fever, chikungunya and Sindbis) using serology as part of an ongoing zoonosis surveillance (Simpson et al.,...
Out of 74 blood samples, a set of nine samples (from four male and five female patients) were selected to be utilized in this study.

### 6.3.3 PCR amplification and sequencing

DNA used in this study was extracted from blood using the QIAamp DNA mini kit \(^\circledast\) (Qiagen, Hilden, Germany) according to the manufacturer’s instructions and stored at -20°C. The V1-V8 variable regions of the 16S rRNA gene was amplified from samples using barcoded sample specific primers as described in Chapter 3 of this thesis using 10 µl of human DNA (approximately 200 ng of DNA). Table 6-1 shows the sample and primer information of PCR amplicons for the library constructed for CCS.

#### Table 6-1: Sample and primer information for PCR amplicons from humans for library.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Gender</th>
<th>Clinic of origin</th>
<th>Primer ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>male</td>
<td>Utha</td>
<td>F1R1</td>
</tr>
<tr>
<td>H4</td>
<td>male</td>
<td>Welverdiend</td>
<td>F2R4</td>
</tr>
<tr>
<td>H8</td>
<td>female</td>
<td>Gottenburg</td>
<td>F2R3</td>
</tr>
<tr>
<td>H18</td>
<td>male</td>
<td>Welverdiend</td>
<td>F3R1</td>
</tr>
<tr>
<td>H27</td>
<td>female</td>
<td>Welverdiend</td>
<td>F3R2</td>
</tr>
<tr>
<td>H47</td>
<td>male</td>
<td>Welverdiend</td>
<td>F3R3</td>
</tr>
<tr>
<td>H53</td>
<td>female</td>
<td>Utha</td>
<td>F3R4</td>
</tr>
<tr>
<td>H59</td>
<td>female</td>
<td>Gottenburg</td>
<td>F3R5</td>
</tr>
<tr>
<td>H69</td>
<td>female</td>
<td>Welverdiend</td>
<td>F4R1</td>
</tr>
</tbody>
</table>

*H=AFI patient  
†Refer to primer codes in Table 3-1  
*Human blood samples were collected from AFI patients that reported to the Gottenburg, Utha and Welverdiend clinics, but patients could have come from neighbouring villages in the Mnisi community.

### 6.3.4 Sequence analysis

Binning, trimming and filtering of sequence data obtained was conducted using the Pacific Biosciences software according to the set sequence size range and 99% precision. Reads were then analysed using the Ribosomal Database RDP 16S classifier (Cole et al., 2009) to classify sequence reads to the genus level using a 95% confidence interval. Sequence data was then blasted against the NCBI BLASTn 16S rRNA gene microbial database using the command line application to ascertain the identity of sequences. Results from BLAST were filtered to a minimum length of 1275 bp and 98% identity in Microsoft Excel (Gall et al., 2016). Sequence reads that fell below 98% identity were reported at the genus level, while reads that were 98% and above in identity were reported at the species level as described, when appropriate (Bonnet et al., 2014, Budachetri et al., 2014, Jones et al., 2010). OTUs that were less than 1% of the total number of sequences were grouped into the ‘rare’ category as described by (Gall et al., 2016). Raw microbiome sequence data will be deposited at the sequence read archive at the National Center for Biotechnology information. Consensus sequences for *R. africae* were extracted in CLC genomics workbench 9.5.1 (Qiagen). An alignment with homologous sequences from GenBank of *Rickettsia* species was created within the workbench. Phylogenetic analyses of the 16S
rRNA gene sequences was conducted using the Maximum Likelihood method in MEGA 7 (Kumar et al., 2016).

6.3.5 Statistical analysis

The bacterial make up of human blood was analysed using the community ecology package vegan version 2.5-2 (Oksanen et al., 2016) in R studio version 1.1 (R Core Team, 2013). Alpha diversity rarefaction curves were plotted to calculate the mean species diversity of bacteria in AFI patients. Principal Component Analysis (PCA) was done to quantify the compositional similarity or dissimilarity of the bacterial population in the blood of the AFI patients where dimensions or principal components were constructed from the linear composition of the variables. This was done using the package FactoMineR (Lê et al., 2008) for explanatory data analysis in R studio. The individuals that clustered together indicated they shared similar blood bacterial profiles while dispersion indicated variability. The close proximity of the variables on one plane of the projections indicated that the variables are positively correlated while positions of variables on the opposing quadrant of the projection indicated negative correlation. The correlation coefficient (r) between variables and the dimensions was extracted from the PCA and was considered significant if the p value was lower than 0.05. Statistical tests of analysis of variance (ANOVA) and Kruskal-Wallis rank sum test were conducted to test if there were statistically significant differences in the bacterial composition between the genders of AFI patients. The differences were considered statistically significant with a p-value ≤0.05.

6.4 Results

6.4.1 Barcoded 16S rRNA gene PCR

PCR products of the near full length 16S rRNA gene were obtained from the nine DNA samples from AFI patients using the barcoded primers. Figure 6-1 shows a representative gel picture of 16S rRNA gene PCR.
Figure 6-1: Example of 16S rRNA gene amplicons from AFI patients. Photograph of a 1.5% agarose gel (stained with ethidium bromide) of amplicons. Lanes b-m contain human samples. Lane n contains the positive control *Anaplasma centrale* and lane o contains the negative water only control. Lane a and p contain the 1kb DNA molecular ladder, with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

### 6.4.2 Sequence and statistical analysis

PacBio CCS sequencing of nine 16S rRNA gene amplicons from AFI patients yielded 13,725 bacterial sequences. After filtering, the mean number of reads per sample was 1,513 sequences. The mean species diversity of bacterial populations detected in the AFI patients plotted using rarefaction curves plateaued early in the sampling for all of the samples indicating that majority of bacterial OTUs were well represented. Figure 6-2 shows the rarefaction diversity in AFI patients.

Figure 6-2: Rarefaction curves of AFI patients from the Mnisi community. The mean species diversity of bacterial populations in human blood were plotted as a function of read depth. The vertical line in the plot indicates the value where rarefaction criterion was satisfied.
Examining the overall sequence set, 32 OTUs were obtained from AFI patients that grouped into 23 species and within 22 genera excluding OTUs grouped into the ‘rare’ category and unclassified. Approximately 27% of the total sequences from the blood of AFI patients corresponded to \textit{Herbaspirillum huttiense}, 16.1% to \textit{R. africae}, 15.1% to \textit{Stenotrophomonas maltophilia}, 11.3% to \textit{Stenotrophomonas} spp. while 11.2% of the sequences were categorized in the ‘rare’ group. Other organisms detected in lower numbers include \textit{Achromobacter xylosoxidans}, corresponding to 3.6% of the sequences, \textit{Delftia lacustris} (3%), \textit{Sphingomonas paucimobilis} (2.8%), \textit{Beijerinckia fluminensis} (1.9%), \textit{Pseudomonas putida} (1.9%), \textit{Rhizobium} spp. (1.9%), \textit{Sphingobium yanoikuyae} (1.2%) while \textit{Rickettsia} spp. and \textit{Herbaspirillum} spp. each comprised 1% of the total sequences. Two sequences corresponding to \textit{Brucella melitensis} were detected in AFI patient H47.

Consensus sequences for the 16S rRNA gene from organisms of interest were extracted from GenBank and compared to sequences obtained in this study, revealing that the \textit{H. huttiense} sequences (3,698 in total) that were detected from all of the AFI patients had >99% identity to \textit{H. huttiense} strain NBRC (NR114139). \textit{Rickettsia africae} sequences from AFI patients H18, H27 and H59 (2,195 in total) had >99% identity to \textit{R. africae} strain ESF-5 (CP001612). \textit{Stenotrophomonas maltophilia} sequences that were detected in eight patients (1,918 in total) had 99% identity to \textit{S. maltophilia} strain AB550 (CP028899). \textit{Achromobacter xylosoxidans} sequences (491 in total) detected in seven of the nine AFI patients had 99% identity to \textit{A. xylosoxidans} strain FDAARGOS_147 (CP014060). \textit{Brucella melitensis} sequences from AFI patient H47 were 100% identical to \textit{B. melitensis} strain CIT 31 (CP025822). AFI patient H27 had the highest proportion of sequences corresponding to \textit{R. africae}, with 90.6% of the sequences obtained from this patient, while AFI patient H47 had the highest proportion of sequences corresponding to \textit{H. huttiense}, equalling 58.2% of the sequences obtained from this patient. AFI patient H4 had the highest proportion of sequences corresponding to \textit{S. maltophilia}, with 25.1%. Figure 6-3 shows the bacterial microbiome composition of the blood of AFI patients in the Mnisi community.
**Figure 6-3**: Relative abundance of major bacterial taxa in the blood of AFI patients from the Mnisi community.*H=AFI patient.

PCA of the individuals factor map of the nine AFI patients in the Mnisi community confirmed that AFI patient H27 (female) was very distinct from the other patients. AFI patients H1 and H4 (males) and H8 (female) shared some similarities due to their close proximity on the top right quarter of the figure with the duo of H4 and H8 being more similar to each other. The other patients H18 and H47 (males) and H53, H59 and H69 (females) also shared some similarities due to their location on the bottom left quarter of the figure. The PCA plot also showed that there was no association with AFI patients of the same gender grouping together. Figure 6-4 shows PCA plot of AFI patients in the Mnisi community.
Figure 6-4: PCA plot of the individuals factor map of AFI patients in the Mnisi community. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. Patients from Gottenburg clinic are represented in black, patients from Utha clinic are in red and patients from Welverdiend clinic are in green. The percentage of variability for each dimension is given 59.81% for the first axis and 16.64% for the second axis. H = AFI patient.

The PCA plot of the variables factor map (Figure 6-5) showed that *D. lacustris*, *S. maltophilia*, *Stenotrophomonas* spp., *S. yanoikuyae*, *P. putida*, *S. paucimobilis* expressed by the correlation coefficient (R = 0.9) were significantly and positively correlated to Dimension 1 alongside *H. huttiense* (R = 0.8) and the ‘rare’ group (R = 0.7). *Rickettsia africae* and *Rickettsia* spp. were significantly and positively correlated to Dimension 2 (R = 0.8). On the other hand, *Herbaspirillum* spp. (R = -0.7) was significantly and negatively correlated to Dimension 2. When relating the PCA plot of the individuals factor map to the variables factor map, we can see that H27 was linked with the variables *R. africae* and *Rickettsia* spp. that are main positive contributors to the Dimension 2. On the other hand, H18, H47, H53, H59 and H69 were linked to the variable *Herbaspirillum* spp. that was negatively correlated to Dimension 2. AFI patients H1, H4 and H8 were linked with the variables *P. putida*, *H. huttiense*, *S. maltophilia*, *Stenotrophomonas* spp., *S. yanoikuyae*, *D. lacustris*, *S. paucimobilis* and the ‘rare’ group that were main positive contributors to Dimension 1.
Figure 6-5: PCA plot of the variables factor i.e. bacterial populations from the blood of AFI patients in the Mnisi community area. Dim 1 and 2 are dimensions or principal components produced from the linear combinations of the variables. The percentage of variability for each dimension is given 59.81% for the first axis and 16.64% for the second axis.

There were statistically significant differences in the bacterial populations between male and female AFI patients using the Kruskal-Wallis rank sum test (P= 0.007) and ANOVA (P<0.001). AFI patient H27 which had a clearly distinct blood bacterial profile was excluded from the analysis to reduce bias.

Phylogenetic analysis of *Rickettsia* sequences showed *R. africae* sequences from AFI patients 18 (male), 27 and 59 (female) grouping with *R. africae* strain ESF-5 (CP001612). Figure 6-6 shows the phylogenetic tree of the 16S rRNA gene from *Rickettsia* sequences.
**Figure 6-6:** Maximum likelihood tree of 16S rRNA gene sequences showing the phylogenetic relationship between *Rickettsia* species. *Ehrlichia chaffeensis* was used as the outgroup sequence. Representative *R. africae* sequences from AFI patients generated in this study are in bold. The percentage of the bootstrap (1000 replicates) are shown next to the branches. All positions containing gaps and missing data were eliminated. Evolutionary distances based on the Jukes Cantor model were computed in MEGA 7. H = AFI patient.

**6.5 Discussion**

This study provides an understanding into the bacterial blood microbiome of a subset of AFI patients in the Mnisi community. It has been stated that undifferentiated AFI is the most prevalent sign reported by patients from resource-constrained communities seeking healthcare in sub-Saharan Africa (Prasad et al., 2015). Previous studies carried out on febrile illness in East Africa have reported that AFI is often misclassified as malaria and there is a need for improved diagnosis (Amexo et al., 2004, Reyburn et al., 2004).

*Herbaspirillum huttiense* made up 27% of the total sequences detected from the blood of the AFI patients in this study and was detected in all of the patients tested. It is a gram-negative bacteria that normally lives in soil and water (Regunath et al., 2015). It has previously been implicated as the cause of pneumonia and bacteremia in an immunocompromised patient in the United States (Regunath et al.,...
2015) and is generally regarded as an opportunistic organism in patients battling cancer and other immunosuppressive conditions (Chemaly et al., 2014, Chen et al., 2011, Spilker et al., 2008, Ziga et al., 2010). The detection of this organism in the blood of AFI patients in the Mnisi community suggests that they were all immunocompromised individuals. In the study conducted on febrile patients in the community (Simpson et al., 2018), the authors stated that AFI patients screened were not tested for HIV, but their median age of 34 years correlated with a study on the prevalence of HIV in South Africa that reported a prevalence rate of 14.1% infection within that age demography in Mpumalanga Province (Shisana et al., 2014).

*Rickettsia africae* corresponded to 16% of the total sequences obtained from the blood of the patients but was only detected in three (33.3%) of the patients screened. *Rickettsia africae* causes African tick bite fever (ATBF) an infection in humans typified by headaches, inoculation eschars, neck pain and inflammation of the lymph nodes (Jensenius et al., 2003). In AFI patient 18 who was male, the organism made up 16.5% of sequences obtained while in AFI patient 27 who was female, approximately 90% of the total sequences detected corresponded to the pathogen. In AFI patient 59, who was also female, only one sequence of the organism was detected. Reports of the detection of the organism in indigenous populations is usually scant, with most cases of ATBF being reported in recent travellers to regions where *A. hebraeum* is endemic (Althaus et al., 2010, Binder and Gupta, 2015, Hauser et al., 2016). Previous studies in areas endemic for *A. hebraeum* have reported seropositivity in individuals as high as 70% (Jensenius et al., 2003).

Of recent in the Mnisi community, a serological survey revealed that 63.4% of AFI patients tested had past exposure with detection of Immunoglobulin G (IgG) anti-*Rickettsia* antibodies and in 24.5% of patients a detection of Immunoglobulin M (IgM) anti-*Rickettsia* antibodies indicating a recent exposure to SFG rickettsial pathogens (Simpson et al., 2018). The study by (Simpson et al., 2018) also screened cattle farmers, herders and veterinary staff and found a much higher proportion (92.2%) showing past exposure to *Rickettsia* species. This is understandable because cattle are kept more than any other livestock species in the Mnisi area and *Amblyomma hebraeum* the main vector of *R. africae* is also the most abundant tick vector on cattle in the community (Zumpt, 2009). Chapter 5 of this thesis on the microbiome of *A. hebraeum*, and a previous study on tick and flea pools collected from domestic dogs in the study area (Kolo et al., 2016) have shown the high prevalence and dominance of *R. africae* as the principal symbiont of *A. hebraeum* ticks in the Mnisi community. It is interesting that AFI patient 27 had such a high proportion of *R. africae* in her blood microbiome. Although the rickettsemia in this individual was not quantitated, the profile suggests that she may have been experiencing an active infection. This finding from our study correlates with a recent study that found that females in the Mnisi community have an increased risk to SFG rickettsiosis (Berrian et al., 2019). Phylogenetic analysis of *R. africae* sequences from AFI patients 18, 27 and 59 showed sequences grouping with *R. africae* reference strain ESF-5 (Fournier et al., 2009) supported with a fair bootstrap value. This study thus serves as the first report of the molecular detection of zoonotic *R. africae* from the blood of rural indigenous AFI patients in the Mnisi community, South Africa.
Stenotrophomonas maltophilia and Stenotrophomonas spp. made up 15.1% and 11.3% of the total sequences obtained from the blood of the AFI patients and were detected in all the patients. *Stenotrophomonas maltophilia* is an emerging global multidrug resistant (MDR) pathogen (Brooke, 2012) and has previously been detected in the blood microbiome of apparently healthy blood donors in France (Paisse et al., 2016). It is also regarded as an opportunistic pathogen causing pneumonia and bacteremia in immunocompromised patients (Garazi et al., 2012).

*Achromobacter xylosoxidans* made up 3.6% of the total sequences obtained from the blood of the AFI patients and was detected in eight out of nine of the patients tested. An uncommon opportunistic pathogen, it is also usually associated with immunocompromised patients (Dai et al., 2015). It is interesting that *A. xylosoxidans* occupied 21.4% of the sequences obtained from the blood of domestic dogs from the same area - see Chapter 4 of this thesis. This is in agreement with a previous study where the authors reported that owners of dogs share a similar bacterial microbiome with their pet dogs compared with other dogs, although the finding from the study described the skin microbiota (Song et al., 2013).

*Pseudomonas putida* made up 1.9% of the sequences obtained from the blood of the AFI patients and was detected in seven out of nine of the patients tested. This is an organism that is usually associated with injury and with immunocompromised individuals (Thomas et al., 2013). In general, *Pseudomonas* spp. are regarded as opportunistic nosocomial organisms causing bacteremia in patients with neutropenia (Bodey et al., 1985, Denham et al., 2017). A few studies have been published that have reported the emergence of multidrug resistant *P. putida* in immunocompromised patients in Argentina and the United States (Almuzara et al., 2007, Bennett et al., 2009).

Two *Brucella* spp. sequences were detected in AFI patient 47 who was an elderly male patient. Sequence analysis revealed a 100% identity to *B. melitensis* strain CIT31 identified from a human in China (CP025822). *Brucella melitensis* is an emerging zoonotic Alphaproteobacterium that is associated with infection in sheep and goats but it has also been incriminated as the dominant cause of human brucellosis (Corbel, 1997). In South Africa, *B. melitensis* has been isolated from a human patient in the Western Cape (Wojno et al., 2016) and from slaughtered cattle in Gauteng Province (Kolo et al., 2018). This correlates with a previous report that detected antibodies to *Brucella* spp. in 1.4% of AFI patients screened (Simpson et al., 2018). Although very minimal, the detection of this very important zoonotic pathogen in the blood of an AFI patient in this study is notable.

We observed in this study that the blood microbiome of AFI patients in the Mnisi community was dominated by members of the phylum Proteobacteria. This is consistent with previous blood microbiome studies that observed a similar occurrence in the blood of patients with sepsis and patients battling cardiovascular disease (Dinakaran et al., 2014, Gosiewski et al., 2017, Li et al., 2018, Rajendhran et al., 2013). This study detected in low proportions the organisms *Delftia lacustris* comprising (3%),
Sphingomonas paucimobilis (2.8%), Beijerinckia fluminensis (1.9%), Rhizobium spp. (1.9%) and Sphingobium yanoikuyae (1.2%) of the total sequences obtained from the blood of the AFI patients. However, it is important to state that these organisms have previously been listed as possible contaminant genera from DNA extraction kits, PCR reagents and molecular grade water and have been sequenced from negative controls in high throughput 16S rRNA gene profiling studies (Salter et al., 2014). No amplification was observed from the negative controls of the PCRs conducted in this study however it cannot be categorically ruled out that these organisms were contaminants.

In this study, the average age of AFI patients from which zoonotic bacterial pathogens were detected was 48 years, which supports the statement that advancing age may play a role as a risk factor for exposure to zoonotic pathogens (Maze et al., 2018). Statistically significant differences in the bacterial populations between genders of AFI patients was also observed. PCA revealed that there was no association with AFI patients of the same gender having similar profiles of bacterial populations in their blood. PCA also revealed positive correlations observed between R. africae and Rickettsia spp., and an association between H. huttiiense, D. lacustris, S. maltophilia, Stenotrophomonas spp., S. yanoikuyae, P. putida, S. paucimobilis and the ‘rare’ group which were also positively correlated with each other in the blood of the AFI patients. A limitation of this study was not comparing the bacterial diversity that exists in the blood of AFI patients versus healthy non-AFI patients to determine if significant differences exist between both sets as was conducted in other blood microbiome studies by (Gosiewski et al., 2017, Li et al., 2018, Rajendhran et al., 2013).

In conclusion, this study reports a diverse blood microbiome from AFI patients in the Mnisi community. It reports the detection of the zoonotic pathogen R. africae from the blood of three AFI patients in the community. It also reports the detection of the globally important zoonotic pathogen B. melitensis from an AFI patient in the community. This study recommends that tick-borne rickettsioses should be considered in the differential diagnosis of non-malarial AFI in the Mnisi community, South Africa.
6.6 References


HERCIK, C., COSMAS, L., MOGENI, O. D., WAMOLA, N., KOHI, W., OMBALLA, V., OCHIENG, M., LIDECHI, S., BONVENTURE, J., OCHIENG, C., ONYANGO, C., FIELDS, B. S., MFINANGA,


Chapter 7

*Anaplasma phagocytophilum* and other *Anaplasma* spp. in various hosts in the Mnisi community, Mpumalanga Province, South Africa

7.1 Abstract

DNA samples from 74 patients with acute febrile illness, 282 rodents, 100 cattle, 56 dogs and 160 *Rhipicephalus sanguineus* ticks were screened for *Anaplasma phagocytophilum* using a quantitative PCR assay targeting the *msp2* gene. The test was found to detect both *A. phagocytophilum* and *Anaplasma* sp. SA/ZAM dog DNA. Sequencing of the 16S rRNA and *gltA* genes confirmed the presence of *A. phagocytophilum* DNA in humans, dogs and rodents, highlighting it as a possible contributing cause of AFI in rural South Africa. A number of recently described *Anaplasma* species and *A. platys* were also detected in the study. However, the cross-reaction of the diagnostic test with *Anaplasma* sp. SA/ZAM dog indicates that this very closely related species may contribute to erroneous findings of *A. phagocytophilum* unless careful controls are implemented. Until this species is formally described, it should be referred to as *Anaplasma* sp. SA dog.

7.2 Introduction

*Anaplasma phagocytophilum* is a zoonotic tick-borne intracellular pathogen that causes granulocytic anaplasmosis in humans, dogs and horses, and tick-borne fever in ruminants (Stuen et al., 2013). Clinical signs of the disease in humans can range from a mild febrile illness to a life-threatening condition (Bakken and Dumler, 2008, Rikihisa, 2011). The main vectors are hard ticks of the *Ixodes* genus, namely *Ixodes ricinus* in Europe, *I. scapularis*, and *I. pacificus* in the eastern and western parts of the United States respectively, and *I. persulcatus* in Asia (Woldehiwet, 2010). In Europe, *A. phagocytophilum* has been detected in the yellow-necked mouse (Hulinska et al., 2004) and field voles (Bown et al., 2009) while in the eastern United States, the white footed mouse is considered the main reservoir host (Massung et al., 2003).

There have been reports of the detection of *A. phagocytophilum* in Africa, mainly identified using nucleic acid-based detection methods. *Anaplasma phagocytophilum* DNA has been detected in horses, ticks and cattle from Tunisia (M’Ghirbi et al., 2016, M’Ghirbi et al., 2012), vervet monkeys and baboons in Zambia (Nakayima et al., 2014), lions, African wild cats and serval in Zimbabwe (Kelly et al., 2014), and dogs and cattle in Algeria (Azzag et al., 2015, Dahmani et al., 2015). *Anaplasma phagocytophilum* DNA has also been detected in a number of tick species including *Rhipicephalus sanguineus* (Egypt) (Ghafar and Amer, 2012), *Amblyomma cohaerens* and *Rhipicephalus pulchellus* (Ethiopia) (Hornok et al., 2016, Teshale et al., 2016) and *Rhipicephalus maculatum* (Kenya) (Mwamuye et al., 2017).

*This chapter has been submitted to the Emerging Infectious Diseases journal for publication.*
Anaplasma phagocytophilum was also reported in ticks collected from cattle, sheep and goats in South Africa (Mtshali et al., 2015); however, this finding should be verified since close inspection of the PCR primers used in that study indicates that they could amplify any Anaplasma species, and the sequence data presented does not conclusively prove the presence of A. phagocytophilum.

A closely related organism, Anaplasma sp. SA dog strain, has been detected in dogs from South Africa (Inokuma et al., 2005, Kolo et al., 2016), and a similar sequence, designated Anaplasma sp. ZAM dog, was detected in Zambia (Vlahakis et al., 2018). In a recent study, A. phagocytophilum-like 16S rDNA sequences were reported from ticks collected from dogs and cats (Mtshali et al., 2017).

While the preceding studies suggest A. phagocytophilum presence throughout Africa, caution is urged, as we show that organisms, such as Anaplasma sp. ZAM dog and Anaplasma sp. SA dog, are very closely related to A. phagocytophilum and PCR primers could cross-react in some of the assays that have been employed. With the advent of high throughput detection methodologies and the consequent spate of 16S rRNA gene survey studies, there have been increasing numbers of distinct Anaplasma-like 16S rRNA gene sequences deposited in the public databases. The relationship of these newly-detected agents to known pathogens, and their ability to serve as a source of cross-reaction in molecular testing, has not been well assessed. Therefore, we urge caution when assigning species designation for organisms in the Anaplasma genus, as the 16S rRNA genes are highly similar, and a similarity score of 97%, as used in many studies, will lead to misclassification.

In the present study, we have confirmed detection of both A. phagocytophilum and Anaplasma sp. ZAM dog in South Africa. As the importance of A. phagocytophilum and its potential role in febrile illness in South Africa is not known, it is necessary to develop methods that are rapid, sensitive and specific for the detection of this pathogen. The aim of this study was to explore the occurrence and genetic diversity of A. phagocytophilum in various hosts in Mnisi, an agro-pastoral community at the wildlife-livestock-human interface in South Africa, and to understand its circulation within the community.

7.3 Methods

7.3.1 Ethics approval

The study was approved by the Animal Ethics Committee of the Faculty of Veterinary Science (V105-15), and the Human Ethics Committees of the Faculty of Health Sciences of the University of Pretoria (152/2016) and the University of the Witwatersrand (M120667). Rodent trapping, tick collection and use of cattle and dog biobanked samples were approved by the Department of Agriculture, Forestry and Fisheries under section 20 of the Animal Diseases Act of 1984 (reference numbers 12/11/1/7/5, 12/11/1/1, 12/11/1/6).

7.3.2 Study area

Samples were collected in the Mnisi community described in Chapter 1 located in the north-eastern corner of the Bushbuckridge municipality in Mpumalanga Province, South Africa (Figure 7-1). The Mnisi community shares 75% of its boundary with adjacent wildlife areas. Livestock farming is the main
agricultural activity and more cattle are kept than any other livestock species. Most cattle owners also own dogs, which accompany herders in the field and to the cattle dip tanks. Rodents are widespread and abundant in the community (Berrian et al., 2016).

![Figure 7-1: Map of the study area with red stars indicating locations of sample collection. Dark green indicates protected areas where wildlife roam freely (map produced by Estelle Mayhew).](image)

### 7.3.3 Collection of blood samples

Blood samples were collected from 282 wild rodents from three habitat areas: urban/periurban (Gottenburg and Hlalakahle), communal rangelands (Tlhavekisa) and a protected area (Manyeleti) – in three research visits from 2014 to 2015. A description of wild rodent trapping was provided in Chapter 3 of this thesis. Blood samples were moved under a DAFF veterinary Red Cross permit to a BSL3 facility at the Faculty of Veterinary Science, University of Pretoria.

Blood samples from 53 domestic dogs and 100 cattle, collected previously (Conan et al., 2015) and stored in the biobank at the Hans Hoheisen Wildlife Research Station, were utilized. Blood samples were from dogs from households in Hluvukani (Figure 7-1). Blood samples from cattle were from dip tanks at Seville A, Seville B, Hlalakahle, Tlhavekisa, Gottenburg and Utha (Figure 7-1).

Human blood samples from 74 patients with non-malarial acute febrile illness (AFI) from Gottenburg, Utha and Welverdiend community clinics were used and were made available to the project by the National Institute of Communicable Diseases (NICD), Johannesburg, South Africa (Table 7-1). These
also formed part of a recent study (Simpson et al., 2018) that assessed the prevalence of selected zoonotic pathogens in these patient samples.

7.3.4 Collection of ticks

Adult male *R. sanguineus* ticks (160) were manually collected from domestic dogs from households in Athol and Hluvukani (Table 7-1) (Figure 7-1). Ticks were identified to species level using the relevant taxonomic keys (Walker et al., 2003), pooled in groups of 8 (1 pool = 8 adult male ticks), and whole ticks manually disrupted with a Tissue Lyzer® (Qiagen).

Table 7-1: Origin and sample sizes of the specimens used in the study.

<table>
<thead>
<tr>
<th>Origin</th>
<th>AFI patients</th>
<th>Rodents</th>
<th>Dogs</th>
<th>Cattle</th>
<th>Ticks*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protected Area:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manyeleti</td>
<td>76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Communal rangeland:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tlhavekisa</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Urban/periurban:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Athol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10 pools</td>
</tr>
<tr>
<td>Gottenburg</td>
<td>22</td>
<td>103</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hi hlakahle</td>
<td>63</td>
<td></td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hluvukani</td>
<td>5</td>
<td>56</td>
<td></td>
<td></td>
<td>10 pools</td>
</tr>
<tr>
<td>Seville A</td>
<td></td>
<td></td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seville B</td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Utha</td>
<td>20</td>
<td></td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Welverdiend</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>74</strong></td>
<td><strong>282</strong></td>
<td><strong>56</strong></td>
<td><strong>100</strong></td>
<td><strong>20 pools</strong></td>
</tr>
</tbody>
</table>

*Ticks were male *R. sanguineus* collected from dogs.
†Human blood samples were collected from AFI patients that reported to the Gottenburg, Utha and Welverdiend clinics, but patients could have come from neighbouring villages in the Mnisi community.

7.3.5 DNA extraction, quantitative real-time PCR and assay specificity

DNA was extracted from all samples using the QIAamp DNA mini kit ® (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. Sample DNA was screened for *A. phagocytophilum* using a previously-reported qPCR assay (Courtney et al., 2004). Primers, ApMSP2 forward (5'-ATG GAA GGT AGT GTT GGT TAT GGT ATT-3') and ApMSP2 reverse (5'-TTG GTC TTG AAGCGC TCG TA-3'), were used to amplify a 77 bp fragment of the *msp2* gene; a TaqMan probe, ApMSP2p (FAM- 5'-TGG TGC CAG GGT TGA GCT TGA GAT TG-3'-TAMRA) was used for detection. Reactions were performed in a final volume of 20 µl comprising 1x Taqman® Universal PCR Master Mix (Thermofisher Scientific, South Africa), 900 nM of each primer, 125 nM of the probe and 2.5 µl of template DNA. The reactions were run on a Step OnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA), using the cycling conditions reported previously (Courtney et al., 2004) with the modification of UNG incubation at 50°C for 2 minutes, prior to cycling. Positive DNA from the L610 [dog] strain (Alberdi et al., 2015) and negative (PCR grade water) controls were included with each run. The results were analyzed using the StepOne Plus software version 2.2. The analytical specificity of the qPCR assay was evaluated using DNA from *A. marginale, A. centrale, Anaplasma* sp. Omatajenne and DNA from two jackal samples (Penzhorn et al., 2018) confirmed to contain *Anaplasma* sp. ZAM dog by 16S rDNA
sequence analysis. Note that the latter were initially described as *Anaplasma* sp. SA dog (Penzhorn et al., 2018), but re-analysis of the sequence data after the *Anaplasma* sp. ZAM dog sequences became available on GenBank, confirmed that these sequences were identical to *Anaplasma* sp. ZAM dog.

7.3.6 Characterization of *A. phagocytophilum* by multi-locus gene sequencing

Four genes (16S rRNA, *gltA*, *msp4* and *ankA*) known to be useful for phylogenetic analysis of *A. phagocytophilum* (Al-Khedery and Barbet, 2014, Inokuma et al., 2001, Vlahakis et al., 2018, Walls et al., 2000, Weisburg et al., 1991) were amplified and sequenced from 32 qPCR-positive samples (cattle=4; dogs=11; AFI patients=4; rodents=8; ticks=5 pools). Amplicons could not be obtained from the remaining qPCR-positive samples. Table 7-2 shows the primers used for amplification. For samples that did not produce amplicons with 16S and *gltA* primer set 1, a nested PCR was performed with primer set 2 as in Table 7-2.

Primers were used at a final concentration of 0.2 µM in a 20 µl reaction containing 10 µl of Phusion Flash® High Fidelity PCR Master Mix (Thermofisher Scientific, South Africa), and 4 µl of template DNA. A second PCR using the same primers was performed using 2.5 µl of the primary PCR product as template. Cycling conditions were as recommended by the manufacturer, with 35 cycles for the primary PCR and 30 cycles for the second PCR. PCR products were purified using the Qiaquick® PCR purification kit (Qiagen, Germany), then cloned using the Clone Jet® PCR Cloning Kit (Thermofisher Scientific, South Africa) according to the manufacturer’s instructions. Clones were screened by colony PCR and at least 10 positive clones per sample were sequenced on an ABI 3500XL Genetic Analyzer using vector primers pJET1.2F and pJET1.2R at Inqaba Biotechnical Industries (Pty) Ltd (Pretoria, South Africa).

Seven near full-length 16S rRNA *Anaplasma* spp. gene sequences obtained from Chapter 3 and 4 of this thesis on the characterization of the bacterial blood microbiome of dogs, cattle and rodents from the study area were also included in the sequence and phylogenetic analysis (C5, C13, C91, D24, D28, D36 and R98).
Table 7-2: Primers used for amplification of the *A. phagocytophilum* DNA.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer set</th>
<th>Primer name</th>
<th>Primer sequence (5’-3’)</th>
<th>Amplicon length (bp/aa)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S rRNA</td>
<td>1</td>
<td>fD1, rP2</td>
<td>AGAGTTTGATCCTGGCTCAG, ACGGCTACCTTGTTACGACCT, AAAAAATCCCCACATTCAAGCA</td>
<td>1470</td>
<td>(Weisburg et al., 1991)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>16SAp1, 16SAp2</td>
<td>ATGGAGGATAATGGCGCA, TGGTGCCAGGGTTGAGCTTG - TAMRA</td>
<td>700</td>
<td>This study</td>
</tr>
<tr>
<td>gltA</td>
<td>1</td>
<td>F1, 1085R</td>
<td>CCGGTTTTATGCTACTGCA, ACTATAACCAGATAAAAATGC</td>
<td>956/318</td>
<td>(Inokuma et al., 2001)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1085F, 1085R</td>
<td>GATCATGARCARAATGCTTC, ACTATAACCAGATAAAAATGC</td>
<td>422/140</td>
<td>(Vlahakis et al., 2018)</td>
</tr>
<tr>
<td>msp4</td>
<td>AB1692F, AB1693R</td>
<td>TAAATGACGCTCTGATATGAGCAC, CACCACCTGTGATGTGTTAAGC</td>
<td>690/230</td>
<td>(Al-Khedery and Barbet, 2014)</td>
<td></td>
</tr>
<tr>
<td>ankA</td>
<td>LA6-F, LA1-R</td>
<td>GAGAGATGCTTTATGGTAAGAC, CGTTCAGCCTCATTTGAGAC</td>
<td>444/148</td>
<td>(Walls et al., 2000)</td>
<td></td>
</tr>
</tbody>
</table>

* Amplicon length is in base pairs (bp), some analyses are converted to amino acids (aa).
† For samples from which no amplicon could be generated using primer set 1 a second PCR, amplifying a shorter fragment of the gene, was attempted using primer set 2.

### 7.3.7 Sequence and phylogenetic analysis

The gene sequences were assembled, edited and aligned using CLC Main Workbench 7.9 (Qiagen). Sequence identities were determined from Genbank using BLASTn (Altschul et al., 1990). Sequences were aligned with appropriate reference sequences from Genbank; sequence variation was inferred using the alignment tool in Workbench. Alignments were edited and trimmed in Bioedit 7 (Hall, 1999). Bayesian inferences were deduced for 16S rRNA and *gltA* gene sequences using Mr Bayes 3.2 (Ronquist et al., 2012). The best nucleotide substitution model predicted for the 16S rRNA gene sequences was GTR + I + G using the Jmodel test 1.3 (Darriba et al., 2012). ProTest 3.0 predicted JTT + G as the best model for *gltA* sequences (Darriba et al., 2011). Phylogenetic trees for 16S rRNA and *gltA* genes were constructed using the Maximum Likelihood (ML) method carried out in PhyML 3.1 (Guindon et al., 2010) and Bayesian inferences in Mr Bayes 3.2 (Ronquist et al., 2012). Trees generated were edited in MEGA 7 (Kumar et al., 2016). Accession numbers of reference sequences are shown in Table 1 of the appendix. Sequences generated in this study have been deposited under accession numbers MK814402-MK814450 and MK804077-MK804111 (Table 2 of the Appendix).

### 7.4 Results

#### 7.4.1 Specificity of the qPCR assay

Amplification of the msp2 gene was not observed with DNA from *A. marginale*, *A. centrale* and *Anaplasma* sp. Omatjenne; however, amplification occurred with *Anaplasma* sp. ZAM dog DNA and *A. phagocytophilum* DNA. No amplification was observed from the negative control. Figure 7-2 shows amplification plot of qPCR assay with noted amplification for *Anaplasma* sp. ZAM dog.
7.4.2 *Anaplasma phagocytophilum* and/or *Anaplasma* sp. ZAM dog occurred in all hosts tested

Based on the qPCR assay 11% (8/74) of AFI patients, 59% (166/282) of wild rodents, 83% (46/56) of domestic dogs, 85% (17/20) of *R. sanguineus* tick pools and 10% (10/100) of cattle samples were positive for DNA from *A. phagocytophilum* and/or *Anaplasma* sp. ZAM dog (Figure 7-3). The percentage from positive samples from rodents in the different habitats was 50% in Manyeleti (protected area), 60% in Tlhavekisa (communal rangeland), 65% in Gottenburg, 75% in Hlalakahle, and 60% in Hluvukani (urban/periurban communities). This was not statistically significantly different (chi-squared test, \( p = 0.24 \)). For the *R. sanguineus* ticks, 90% (9/10) of the pools from Hluvukani and 80% (8/10) of pools from Athol were positive. For cattle, 11% (2/19) of samples from Tlhavekisa, 5% (1/20) from Gottenburg, 15% (3/20) from Hlalakahle, 6% (1/18) from Utha, 13% (2/16) from Seville A and 14% (1/7) from Seville B were positive using this test. Figure 7-4 shows a representative amplification plot of the qPCR assay.
7.4.3 PCR assays for sequencing

PCR products were obtained from the near full length 16S rRNA gene primers (fD1 and rP2) from nine dog DNA samples positive on the qPCR assay and cloned using the Clone Jet® PCR Cloning Kit. Figure 7-5 shows a representative gel picture of colony PCR products from near full-length 16S rRNA gene.

Due to the difficulty of obtaining amplicons from other hosts using the full length 16S rRNA primers, partial 16S rRNA gene nested PCR products of ~717 bp in length were obtained using primers (16S Ap-F and 16S Ap-R) designed in this study from six dogs, three R. sanguineus pools, three cattle, two rodents, and a human DNA sample. Figure 7-6 shows a representative gel picture of positive samples.
PCR products of ~ 950 bp in length were obtained from the *glt*A gene PCR using F4B and HG1085R primers from three dogs and a rodent DNA sample. For samples from which no amplicon could be generated, another PCR, which amplified a shorter fragment of the gene ~ 430 bp was attempted using a second set of primers (F1B and 1085R). Figure 7-7 shows a gel picture of the *glt*A gene PCR with F4B and 1085R.

Nested PCR products of ~ 690 bp of the *msp4* gene were obtained from (two dogs, seven rodents, three humans, five ticks and a bull). Figure 7-8 shows gel picture of the *msp4* gene PCR assay.
Figure 7-8: Gel electrophoresis picture of a 1.5% agarose gel (stained with ethidium bromide) of nested PCR products of the *msp4* gene from positive samples on the qPCR assay. Primers AB1692-F and AB1693-R were used for the PCR. Lanes 2 to 22 contained nested PCR samples. Lane 21 contained the positive control *A. phagocytophilum* strain L610 while lane 2 contained the no template negative control. Lane 1 contained the 100 bp DNA ladder with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

PCR products of ~444 bp in length were obtained from the *ankA* gene PCR from three dogs, three rodents and an *R. sanguineus* pool. Figure 7-9 shows the gel picture of the *ankA* gene PCR.

Figure 7-9: Gel electrophoresis picture of a 1.5% agarose gel (stained with ethidium bromide) of nested PCR products of the *ankA* gene from positive samples on the qPCR assay. Primers LA6-F and LA1-R were used for the PCR. Lanes 2 to 12 contained nested PCR products. Lane marked 11 contained the positive control *A. phagocytophilum* strain L610 while lane marked 12 contained the no template negative control. Lanes 1 and 13 contain the 100 bp DNA ladder with sizes in bp indicated on the right (Invitrogen, ThermoFisher® Scientific, South Africa).

7.4.4 Sequence analysis of the 16S rRNA, *gltA*, *msp4* and *ankA* genes

We analysed the 16S rRNA, *gltA*, *msp4* and *ankA* gene sequences from four AFI patients, eight rodents, four cattle, 11 dogs and five *R. sanguineus* tick pools (Table 7-3). Based on both 16S rRNA and *gltA* gene sequence analyses, none of the mammalian species investigated were found to be simultaneously co-infected with two or more *Anaplasma* species; except for dog D25 which was co-infected with *A. platys* (*Apla1/16S*) and *A. phagocytophilum* (*Aph1/GltA*).
7.4.4.1 16S rRNA
Based on the near full-length 16S rRNA gene (1262-1465 nt), we identified two 16S rRNA sequence variants for *A. phagocytophilum* (Aph1/16S and Aph2/16S), two for *Anaplasma* sp. ZAM dog (Adog1/16S and Adog2/16S), and one each of *A. platys* (Apla1/16S), *Candidatus* Anaplasma boleense (Cab1/16S) and *Anaplasma* sp. Mymensingh (Asm1/16S) (Table 7-3). The Aph1/16S and Aph2/16S sequences differed by 1 and 2 nt, respectively, from the *A. phagocytophilum* type strain Webster (U02521). Aph1/16S was obtained from one dog (D2) (five identical clone sequences), while Aph2/16S was obtained from one rodent (R98; *Mastomys natalensis* trapped in Hlalakahle) and two dogs (D24 and D28). A further three partial 16S rRNA gene sequences (690-693 nt) were obtained from two rodents (R102 and R103; *Rattus tanezumi* and *M. natalensis*, respectively; both trapped in Tlhavekisa) and one patient (H59) sample. These could unfortunately not be assigned to either Aph1/16S or Aph2/16S sequence variants due to the conserved nature of this region of the *A. phagocytophilum* 16S rRNA gene. This result highlights the importance of generating and using (near) full-length 16S rRNA gene sequences when assigning species designations and in constructing phylogenies.

The Adog2/16S sequences obtained in this study were identical to the *Anaplasma* sp. ZAM dog sequences previously described in Zambia (LC269823) (Vlahakis et al., 2018), while the Adog1/16S sequences differed by a single deletion. Adog1/16S and Adog2/16S differed by 4 nt from the *Anaplasma* sp. SA dog sequences previously described in South Africa (AY570538, AY570539, AY570540) (Inokuma et al., 2005). Three near full-length Adog2/16S sequences were obtained from three dogs (D9, D27 and D36), while eight Adog1/16S sequences were obtained from four dogs (D3, D5, D27 and D37). A further eight partial sequences (628-1031 nt) obtained from two dogs (D36 and D37) were assigned to Adog1/16S, while eight partial sequences (687-698 nt) from four dogs (D3, D5, D27 and D37) and three *R. sanguineus* tick pools (RA3, RH3 and RH8) were assigned to Adog2/16S. Both Adog1/16S and Adog2/16S sequence variants were found in five dogs (D3, D5, D27, D36 and D37) suggesting that co-infections of different *Anaplasma* sp. ZAM dog strains may occur in individual hosts.

The nine *A. platys* 16S rRNA gene sequences obtained in this study from two dogs (D25 and D33) (Apla1/16S) were conserved and identical to the *A. platys* 16S rRNA gene sequences described from dogs in Zambia (LC269820, LC269821, LC269822) (Vlahakis et al., 2018). Apla1/16S differed by 5 nt from the *A. platys* type strain 16S rRNA gene sequence (M82801) published in 1992 (Anderson et al., 1992).

The *Candidatus* Anaplasma boleense sequence (Cab1/16S) obtained in this study from one cattle sample (C13; Seville A) differed by 1-2 nt from the *Candidatus* Anaplasma boleense sequences recently described from *Anopheles sinensis* from Wuhan, China (KU586025, KU586041, KU586169) (Guo et al., 2016). It differed by 3 nt from the *Candidatus* Anaplasma boleense sequences originally described from *Hyalomma asiaticum* ticks from the Bole region of Xinjiang, China (KJ410247, KJ410248, KJ410249) (Kang et al., 2014) and by 3 nt from *Anaplasma* sp. Dedessa described from cattle from South-western Ethiopia (KY924886) (Hailermariam et al., 2017).
The *Anaplasma* sp. Mymensingh sequences (Asm1/16S) obtained from two cattle samples (C5 and C91; from Hlalakahle and Seville A, respectively) were conserved and identical to *Anaplasma* sp. Mymensingh originally described from cattle in the Mymensingh district of Bangladesh (MF576175) (Roy et al., 2018).
Table 7-3: *Anaplasma* species 16S rRNA, GltA, Msp4 and AnkA sequence variants circulating in the Mnisi community.

<table>
<thead>
<tr>
<th>Sample nr</th>
<th>Origin</th>
<th>Location</th>
<th>Species</th>
<th>16S rRNA</th>
<th>GltA</th>
<th>Msp4</th>
<th>AnkA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>Hlalakahle</td>
<td>Cattle</td>
<td>Aph1</td>
<td>1††</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C13</td>
<td>Seville A</td>
<td>Cattle</td>
<td>Adog1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>C42</td>
<td>Seville B</td>
<td>Cattle</td>
<td>Apla1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>C91</td>
<td>Seville A</td>
<td>Cattle</td>
<td>Asm1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D2</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apost1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D3</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D5</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Asm1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D9</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Asp1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D24</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apost1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D25</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D27</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Asm1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D28</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Asp1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D33</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apost1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D36</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>D37</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Asm1</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>H27</td>
<td>Welverdiend</td>
<td>Human</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>H47</td>
<td>Welverdiend</td>
<td>Human</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>H53</td>
<td>Utah</td>
<td>Human</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>H69</td>
<td>Utah</td>
<td>Human</td>
<td>UM</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>R98</td>
<td>Hlalakahle</td>
<td>M. natalensis</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>R102</td>
<td>Tlhavekisa</td>
<td>R. tanezumi</td>
<td>U</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>R103</td>
<td>Tlhavekisa</td>
<td>M. natalensis</td>
<td>U</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>R104</td>
<td>Tlhavekisa</td>
<td>M. natalensis</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>R105</td>
<td>Tlhavekisa</td>
<td>M. natalensis</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>R124</td>
<td>Hlalakahle</td>
<td>Saccostomus sp.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>R125</td>
<td>Hlalakahle</td>
<td>G. leucocaster</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>R138</td>
<td>Hlalakahle</td>
<td>G. leucocaster</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>RA1</td>
<td>Athol</td>
<td>R. sanguineus</td>
<td>U</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>RA3</td>
<td>Athol</td>
<td>R. sanguineus</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>RH1</td>
<td>Hluvukani</td>
<td>R. sanguineus</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>RH3</td>
<td>Hluvukani</td>
<td>R. sanguineus</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>RH8</td>
<td>Hluvukani</td>
<td>R. sanguineus</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

*C, cattle; D, dog; H, human; R, rodent; RA, *R. sanguineus* (Athol); RH, *R. sanguineus* (Hluvukani)*

1† Aph, *A. phagocytophilum*
2‡ Adog, *Anaplasma* sp. ZAM dog
3§ Apla, *A. platys*
4¶ Cab, *Candidatus* Anaplasma boleense
5# Asm, *Anaplasma* sp. Mymensingh
6** Asp, *Anaplasma* sp.
†† numbers represent sequence variants
‡‡ M, Msp4 *Anaplasma* sequence
§§ A, AnkA *Anaplasma* sequence
¶¶ U, sequence variant unknown/could not be assigned due to short sequence length.

7.4.4.2 GltA

We identified one GltA sequence variant for each of *A. phagocytophilum* (Aph1/GltA), and *Anaplasma* sp. ZAM dog (Adog1/GltA), based on partial GltA deduced amino acid sequences (138-302 amino acids) (Table 7-3). *Anaplasma phagocytophilum* Aph1/GltA deduced amino acid sequences obtained from three dogs (D24, D25 and D28) and one rodent (R102) were conserved and identical to the sheep isolate *A. phagocytophilum* str. Norway variant2 (CP015376), as well as to *A. phagocytophilum* strain RD1 (SCV65315) and an *A. phagocytophilum* from *Ixodes* ticks (AKZ20811). Aph1/GltA differed by 4 amino acids from the *A. phagocytophilum* type strain Webster (AF304136).
One Anaplasma sp. ZAM dog GltA deduced amino acid sequence variant (Adog1/GltA) was identified in a dog (D36); it was identical to the Anaplasma sp. ZAM dog GltA sequence previously described in dogs in Zambia (LC269827) (Vlahakis et al., 2018), and differed by one amino acid from Anaplasma sp. SA dog GltA sequences (AY570541, AAT74599) (Inokuma et al., 2005).

7.4.4.3 Msp4
For Msp4, based on partial deduced amino acid sequences (197-214 amino acids), only one Anaplasma Msp4 sequence variant was identified (Table 7-3). While the deduced amino acid sequence was identical to several other A. phagocytophilum Msp4 sequences and the nucleotide sequence differed by 11 nt from the A. phagocytophilum type strain Webster (EU857674), we could not designate the Msp4 sequence to a particular species, as limited Anaplasma sp. Msp4 sequence data is available in the public databases, and Msp4 sequences of the more recently described Anaplasma spp., are currently unknown. The sequence obtained was thus designated Anaplasma sp./Msp4 (Asp/Msp4).

A total of 23 partial Msp4 deduced amino acid sequences (197-214 amino acids) were obtained from three patient samples (H27, H47 and H53), two dogs (D3 and D33), seven rodents from Thhavekisa and Hlalakahle (R102, R103, R104 [M. natalensis], R105 [Mastomys natalensis], R124 [Saccostomus sp.], R125 [Gerbillicus leucogaster] and R138 [Gerbillicus leucogaster]) and from one cattle sample (C42; Seville B). These were all conserved and identical to A. phagocytophilum type strain Webster (ACH70064) as well as several other A. phagocytophilum Msp4 sequences described from various hosts (ACH70059, ACH70060, AGH02966, AGH02967, AGH02970, AGH02971, AHG97932, AJP32949, AVH68963, AVH68960).

No Anaplasma sp. ZAM dog, A. platys, A. bovis, Candidatus Anaplasma boleense or Anaplasma sp. Mymensingh msp4 gene sequences or Msp4 deduced amino acid sequences are currently available in the public databases. We also failed to amplify the msp4 gene from samples we found to harbour either Anaplasma sp. ZAM dog (dog D36), Candidatus Anaplasma boleense (cattle C13) or Anaplasma sp. Mymensingh (cattle C5 and C91). We can therefore, not relate the Msp4 sequences we obtained to A. phagocytophilum with absolute certainty. While rodents R102 and R103 were shown to harbour only A. phagocytophilum using 16S rRNA gene and/or GltA deduced amino acid sequence analysis (Table 7-3), several of the other samples shown to contain the Asp/Msp4 sequence variant were also shown to harbour either Anaplasma sp. ZAM dog (D3, RA3, RH3 and RH8) or A. platys (D33) 16S rRNA gene sequence variants; these data suggest that Msp4 deduced amino acid sequences might be the same in these two species.

7.4.4.4 AnkA
Based on partial AnkA deduced amino acid sequences (137-144 amino acids), only one Anaplasma sp. ankA sequence variant (Anaplasma sp./AnkA; (Asp/AnkA)) could be identified. As in the case of Msp4, limited Anaplasma sp. AnkA sequences are available in the public databases apart from A. phagocytophilum. Seven partial AnkA deduced amino acid sequences obtained from three dogs (D24,
D25 and D28), three rodents (R102, R103 and R124) and one R. sanguineus tick pool (RH1) were conserved and identical to the A. phagocytophilum strain Dog2 AnkA deduced amino acid sequence (ADA72255), as well as to several other A. phagocytophilum AnkA sequences described from various hosts (AAS21270, ADV02358, ADV02361, ADV02363, KJV67204). As in the case of Msp4, no Anaplasma sp. ZAM dog, A. platys, A. bovis, Candidatus Anaplasma boleense or Anaplasma sp. Mymensingh ankA gene sequences or AnkA deduced amino acid sequences are currently available in the public databases. We did however, obtain a partial AnkA sequence from rodent R102 that was shown to harbour A. phagocytophilum based on both 16S rRNA gene and GltA deduced amino acid sequence analyses. On the other hand, a partial AnkA sequence was obtained from dog D36 shown to harbour Anaplasma sp. ZAM dog using both 16S rRNA gene and GltA deduced amino acid sequence analyses; and from dog D33 shown to harbour A. platys using 16S rRNA gene sequence analysis. Due to the apparent conserved nature of this region of the ankA gene, it is therefore difficult to relate the sequences obtained in this study to only A. phagocytophilum; more sequence data is needed to clarify this point.

7.4.5 Phylogenetic analyses

Separate phylogenetic trees were generated for the 16S rRNA gene (Figure 7-10) and the GltA deduced amino acid (Figure 7-11) sequence datasets obtained in this study; gene sequences were not concatenated.

The sequence similarities observed in the 16S rRNA sequence variants were confirmed by phylogenetic analyses (based on an alignment truncated to 1195 nt). The phylogenetic tree topologies obtained using three tree algorithms were similar, and the maximum likelihood tree was chosen as a representative tree (Figure 7-10).
Figure 7-10: Maximum-likelihood phylogenetic tree based on 16S rRNA nucleotide sequences. The tree shows the phylogenetic relationship between the 16S rRNA gene sequence variants obtained in the study and related Anaplasma species. The 16S RNA variants were designated as *A. phagocytophilum* Aph1 and Aph2, *Anaplasma* sp. ZAM dog Adog1 and Adog2, *A. platys* Apla1, *Candidatus* Anaplasma boleense Cab1 and *Anaplasma* sp. Mymensingh Asm1. Sequence accession numbers are shown in parentheses. The evolutionary history was inferred by using the Maximum Likelihood method based on the General Time Reversible model (Nei and Kumar, 2000). The tree with the highest log likelihood (-3074.44) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.2734)). The rate variation model allowed for some sites to be evolutionarily invariable (+I, 39.92% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 31 nucleotide sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There was a total of 1191 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016).
Phylogenetic trees based on the GltA deduced amino acid sequences were constructed (Figure 7-11); a phylogenetic tree generated from gltA gene sequences was similar to the GltA deduced amino acid sequence tree (data not shown). Comparison with the 16S rRNA phylogenetic analysis revealed similar groupings, with Anaplasma sp. SA dog and ZAM dog consistently grouping together in a separate clade from A. phagocytophilum.

Figure 7-11: Maximum likelihood phylogenetic tree based on GltA deduced amino acid sequences showing the phylogenetic relationship between the obtained Anaplasma GltA sequence variants and related Anaplasma species. The GltA variants were designated as A. phagocytophilum Aph1 and Anaplasma sp. ZAM dog Adog1/GltA. Sequence accession numbers are shown in parentheses. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood (-657.40) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1.7688)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 22.99% sites). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There was a total of 66 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016).

For both the Msp4 and AnkA deduced amino acid sequence datasets, the apparent observed sequence similarities with A. phagocytophilum were confirmed by phylogenetic analyses (data not shown).

7.5 Discussion
This study provides the first report of the detection of the zoonotic agent A. phagocytophilum in humans, dogs and rodents in Mnisi, a rural community in South Africa. This was confirmed by 16S rRNA (humans, dogs, rodents) and gltA (dogs, rodents) gene sequence analysis. Reports of human granulocytic anaplasmosis occurring in Africa have been few (Elhamiani Khatat et al., 2017, Elhamiani Khatat et al., 2016). In South Africa, there have been no reported A. phagocytophilum infections in humans. The significance of detecting A. phagocytophilum DNA in humans, dogs and rodents, as well
as the potential role of *A. phagocytophilum* as a cause of AFI in South Africa, is not known. Recent research in the Mnisi community (Simpson et al., 2018) assessed the prevalence of selected zoonotic pathogens in patients that presented with non-malarial fever (≥37.5°C) at the community health clinics. Organisms for which there was evidence of recent or past infection/exposure to included *Bartonella* spp., spotted fever group *Rickettsia* spp., *Coxiella burnetii*, and *Leptospira* spp. which could have been the cause of their fevers. Low levels of exposure to West Nile and Sindbis viruses, and not Rift Valley fever virus were found. No screening for *Anaplasma* species was done as part of this study (Simpson et al., 2018).

In our study, we obtained one *A. phagocytophilum* (*Aph*) 16S rRNA sequence from patient H59, whereas from patients H27, H47 and H53 we obtained an *Anaplasma* (*AspM*) Msp4 sequence that we could not, due to the conserved region of the *msp4* gene, with certainty assign to *A. phagocytophilum*. These coincided with the detection of IgG antibodies in these patients to the spotted fever group *Rickettsia* spp. (H27, H47, H59), *C. burnetii* (H47) (Berrian et al., 2019) and West Nile virus (H53); and IgM antibodies to Sindbis virus (H47). *Bartonella* spp. were furthermore detected by PCR in one patient (H27). Although no causal link between the detection of these pathogens and AFI can be made, we infer that the role of all of these pathogens should be considered in the investigation of febrile patients in rural areas of South Africa.

We furthermore detected *A. phagocytophilum* DNA in four dogs and three rodents (two *M. natalensis* and one *R. tanezumi* trapped in the urban/periurban area of Hlalakahle and in the communal rangelands of Tlhavekisa). This was based on 16S rRNA and GltA sequence analysis. It has been stated that domestic dogs play a role as sentinels of infection to humans (Cleaveland et al., 2006). *Mastomys natalensis* and *R. tanezumi* are also known synanthropes of humans; the close association both species have with humans indicates they are likely to serve as carriers of infection to man (Bonwitt et al., 2017, Skinner and Chimimba, 2005).

*Anaplasma phagocytophilum* Msp4 sequences in this study were conserved. All ankA sequences in this study were found to be 100% identical, even though this gene has been used in previous studies to show variability between *A. phagocytophilum* strains (Matei et al., 2015). This may be because the region of the gene we targeted was short, unlike previous studies that amplified the entire open reading frame (Scharf et al., 2011, Von Loewenich et al., 2003). It is important to note that we cannot be certain whether the *msp4* and ankA gene sequences obtained from dogs and *R. sanguineus* were derived from *A. phagocytophilum* or *Anaplasma* sp. SA/ZAM dog because there are currently no *msp4* or ankA sequences available from this species for comparison.

It should be further noted that partial Msp4 deduced amino acid sequences were obtained from all host species investigated in this study. Furthermore, partial Anka deduced amino acid sequences were obtained from dogs, rodents and *R. sanguineus* ticks. Both Msp4 and Anka sequences were found to be identical to *A. phagocytophilum*. Since no *Anaplasma* sp. ZAM dog, *A. platys*, *A. bovis*, *Candidatus Anaplasma boleense* or *Anaplasma* sp. Mymensingh Msp4 or Anka sequences are currently available in the public databases, and due to the apparent conserved nature of this partial region of the *msp4*
and ankA genes, we could not with any certainty relate the sequences obtained in this study to only *A. phagocytophilum*. More sequence data is needed to clarify this point.

In screening our sample set initially with a qPCR assay designed to specifically detect *A. phagocytophilum* DNA (Courtney et al., 2004), we found the test to cross-react with *Anaplasma* sp. ZAM dog DNA; suggesting that the high number of qPCR positives obtained from dogs (82%) and *R. sanguineus* (85%) could probably be attributed to the presence of *Anaplasma* sp. ZAM dog. *Anaplasma platys* control DNA was not available and we therefore cannot discount the possibility that the qPCR assay might also cross-react with *A. platys* DNA. We subsequently identified *Anaplasma* sp. ZAM dog 16S rRNA gene sequence variants (Adog1 and Adog2) in six dogs and three *R. sanguineus* tick pools (Adog2). For one dog (D36), an *Anaplasma* sp. ZAM dog GltA sequence variant (Adog1) was also identified. In addition to this, *A. platys* was identified in a further two dogs based on 16S rRNA gene sequence analysis.

Not much is known about *Anaplasma* sp. ZAM dog and *Anaplasma* sp. SA dog apart from the initial studies reporting these sequences in dogs from South Africa and Zambia (Inokuma et al., 2005, Vlahakis et al., 2018). In South Africa, *Anaplasma* sp. SA dog was first described from three dogs presented at the Veterinary Teaching Hospital of the Medical University of South Africa (Inokuma et al., 2005) and shown to be genetically closely related to *A. phagocytophilum* based on 16S rRNA and gltA gene sequence analysis. Unfortunately, laboratory records and clinical data on these dogs were not available. In the subsequent Zambian study, similar 16S rRNA and gltA gene sequences were obtained from apparently healthy dogs and the organism was designated *Anaplasma* sp. ZAM dog. The authors also indicated dogs as a possible reservoir host in the transmission of this *Anaplasma* species. It should also be noted that these authors (Vlahakis et al., 2018) stated that “the same *Anaplasma* species” was previously reported in sheep in South Africa (Allsopp et al., 1997) and a goat in Mozambique (Bekker et al., 2001); and that transmission of this *Anaplasma* species was suspected to be through the bite of *R. sanguineus* ticks (Inokuma et al., 2005). The species the authors was referring to is in fact *Anaplasma* sp. Omatjenne that groups phylogenetically with *A. platys*.

In sub-Saharan Africa, *A. platys* has been detected from dogs in North Central Nigeria (Kamani et al., 2013), ticks from dogs in the Congo (Sanogo et al., 2003), and ticks and dogs from Cote d’Ivoire and Kenya (Matei et al., 2016). In southern Africa, it has been detected in *R. evertsi evertsi* collected from domestic and wild ruminants in South Africa (Berggoetz et al., 2014), and more recently in domestic dogs in Zambia (Vlahakis et al., 2018). In Africa, *R. sanguineus* is thought to be the reservoir host that plays a role in the transmission of *A. platys* (Sanogo et al., 2003); however, we did not detect *A. platys* in any of the *R. sanguineus* ticks sampled in this study.

The phylogenetic trees inferred from the 16S rDNA (Figure 7-10) and GltA (Figure 7-11) sequence data that obtained in our study, consistently grouped *A. phagocytophilum*, *A. platys* and *Anaplasma* sp. ZAM dog (and *Anaplasma* sp. SA dog) into three distinct clades, indicative of a divergence between these organisms. The bootstrap values were, however, only poorly supportive of these relationships; this was also reported previously (Inokuma et al., 2005).
There is only a four-nucleotide difference between the 16S rRNA gene sequences of *Anaplasma* sp. ZAM dog and *Anaplasma* sp. SA dog; and only one GltA deduced amino acid sequence difference. These data suggest that *Anaplasma* sp. ZAM dog and *Anaplasma* sp. SA dog are variants of the same species. Phylogenetic analyses furthermore grouped *Anaplasma* sp. ZAM dog and *Anaplasma* sp. SA dog into a distinct clade; providing sufficient divergence from other *Anaplasma* species to warrant classification as a separate species. Until appropriate type-material can be deposited and the species can be formally described, we will refer to this novel organism as *Anaplasma* sp. SA dog for *Anaplasma* sp. southern Africa dog. Our findings would further suggest that *R. sanguineus* should be considered as a possible vector for *Anaplasma* sp. SA dog in South Africa. It will be necessary, however, to undertake a tick transmission study to confirm the vectorial capacity of this tick species.

16S rRNA gene sequences closely related to the novel organism *Candidatus Anaplasma boleense* is reported from one cattle sample from the study area; this is also the first description of *Candidatus Anaplasma boleense* in South Africa. This agent was first detected from *Hyalomma asiaticum* collected from livestock in the Bole region of Xinjiang China. Phylogenetic analysis of 16S rRNA, *gltA* and *groEL* sequence data (Kang et al., 2014) revealed a lineage clearly differentiated from other *Anaplasma* species (Kang et al., 2014). The organism was subsequently described from *Anopheles sinensis* mosquitoes in Wuhan, China (Guo et al., 2016). The zoonotic potential and pathogenicity of this agent are unknown.

We furthermore described *Anaplasma* sp. Mymensingh 16S rRNA gene sequences from two cattle samples. *Anaplasma* sp. Mymensingh was originally described from cattle in the Mymensingh district of Bangladesh (Roy et al., 2018). Phylogenetic analysis of combined 16S rRNA and *groEL* data (Roy et al., 2018) revealed that *Anaplasma* sp. (Mymensingh) clustered with *A. platys*. The zoonotic potential, pathogenicity, tick vector and reservoir hosts of this agent are unknown.

In conclusion, this study serves as the first report of the detection of *A. phagocytophilum* in humans, dogs and rodents in South Africa. We recommend that health care practitioners in the Mnisi community also consider *A. phagocytophilum* in the differential diagnosis of non-malarial AFI which will help to guide appropriate treatment. The study furthermore provided evidence that *Anaplasma* sp. ZAM dog and *Anaplasma* sp. SA dog strain were phylogenetically distinct from other *Anaplasma* species and warrant classification as a separate species. We report the first detection of *Candidatus Anaplasma boleense* and *Anaplasma* sp. Mymensingh in cattle in South Africa.
7.6 References


Chapter 8
General Discussion, Recommendations and Conclusions

The Mnisi community is located at the northeastern part of the Bushbuckridge municipality in Mpumalanga Province in South Africa. It is made up of an estimated 29,500 hectares of land and shares 75% of its boundaries with wildlife reserves (Berrian et al., 2016). The total population in the community is approximately 40,000 people in 8500 households (Statistics, 2012). The Mnisi community has been classified as falling under one of South Africa’s 14 rural poverty nodes with an estimated HIV prevalence rate of 14.1% (Shisana et al., 2014). Prior research conducted in the area suggested that undifferentiated non-malarial AFI was the most common presenting sign in patients that sought healthcare at the community clinics (Quan et al., 2014). A survey carried out in the community revealed 72% of all households own an animal, and 68% of households own livestock (Berrian et al., 2016). Out of the number of animal-owning households, 55% of households owned chickens, 31% dogs, cattle were owned in 25% of households, goats 16%, cats 9%, pigs were owned by 5% of households and ducks and pigeons were owned in less than 1% of the households sampled (Berrian et al., 2016). In the aforementioned study, respondents in the community were asked if they believed animals could transmit diseases to humans with 61% of individuals answering in the affirmative, they also believed zoonotic diseases transmitted to humans from animals were from dogs, cattle and cats (Berrian et al., 2016). Cattle owners in the community take cattle to designated dip tanks for weekly dipping as part of control measures against Corridor disease transmitted by Rhipicephalus appendiculatus and for mandated weekly inspections against food and mouth disease (FMD) carried out by the Mpumalanga Veterinary Services (Berrian et al., 2016).

This study characterized the bacterial blood microbiome of 25 Mastomys spp. molecularly profiled as M. natalensis and M. coucha collected from three habitat areas in the vicinity of the Mnisi community. Previous research in the area had found rodents to be common and abundant with 76% of households sampled reporting rodents around their homes (Berrian et al., 2016). Another study conducted in the area suggested that rodent-borne zoonoses may be implicated as causes of non-malarial AFI in humans and reported that 6.5% of AFI patients tested were positive for the rodent-borne zoonotic pathogen Bartonella using PCR (Simpson et al., 2018).

Sequence analysis of bacterial populations in the blood of Mastomys spp. in the Mnisi community revealed that 64% of the total sequences corresponded to Bartonella spp. and all rodents were infected with the organism. This was in agreement with studies that found Bartonella to be the dominant phylotype in rodent blood (Cohen et al., 2015, Rynkiewicz et al., 2015). This study reported the detection of the zoonotic pathogens B. grahamii, B. henselae and Bartonella sp. RF255YX from Mastomys spp. in the area. It is interesting that B. henselae was recently detected in two AFI patients in the Mnisi community (Simpson et al., 2018), thus its detection in Mastomys spp. from the same area was notable. This study also detected Bartonella bovis in cattle and A. hebraeum collected from cattle in Mnisi.
Phylogenetic analysis reaffirmed the microbiome results that *Mastomys* spp. from all habitat areas had co-infections with different *Bartonella* spp. This was in agreement with previous studies that found a diverse assembly of *Bartonella* spp. in rodents from Southern China (Ying et al., 2002), from rodents in South-western Spain (Márquez et al., 2008) and the recent report of the presence of three phylogenetic lineages of *Bartonella* species circulating in 80 *Rhabdomys pumilio* rodents captured from the Western Cape Province in South Africa (Hatyoka et al., 2019). Statistical analysis on bacterial communities in the blood of *Mastomys* spp. revealed that the blood microbiome of rodents was diverse across the different habitat areas. This was in agreement with a previous study (Gall et al., 2017) where the authors showed that geographic location plays a role in the ecological variation that is seen in bacterial populations in tick vectors.

This study found that *Anaplasma* species made up 36% of the total bacterial population in the blood of dogs with the zoonotic *A. platys* comprising 19.3%, *Anaplasma* sp. ZAM dog 14.8%, *Anaplasma* spp. 1.6% and *A. phagocytophilum* 0.3% of the total sequences obtained from canine blood. The findings indicated that dogs in the Mnisi community had co-infections of these closely related *Anaplasma* species. *Anaplasma centrale* was also detected from two dogs and *A. marginale* from a single dog, which was a rare finding. Since dogs in the Mnisi community are free roaming and are often used to escort cattle to the dip tanks, it is presumed tick vectors that transmit both organisms may have fed on them.

*Anaplasma* species were also the dominant pathogens of the blood microbiome of cattle comprising 96.8% of the total bacterial population. *Anaplasma marginale* alone made up 54% of the total sequences obtained from the blood of cattle, followed by the novel species, *Anaplasma* sp. Mymensingh, which comprised 22.2% of the sequences obtained. This is the first detection of *Anaplasma* sp. Mymensingh in cattle in South Africa. *Anaplasma* sp. Mymensingh was recently detected from 13% of cattle tested in Bangladesh (Roy et al., 2018). Sequences of other novel *Anaplasma* species recently detected in cattle in Ethiopia (Hailemariam et al., 2017): *Anaplasma* sp. Dedessa, *Anaplasma* sp. Hadesa and *Anaplasma* sp. Saso were detected in cattle in this study. A negative correlation of infection was observed between *A. marginale*, *Anaplasma* sp. Mymensingh and *Anaplasma* sp. Dedessa infection in cattle. Cattle that had a higher proportion of *A. marginale* infection in their blood had a lower proportion of *Anaplasma* sp. Mymensingh or *Anaplasma* sp. Dedessa and vice versa, in agreement with the bacterial interference phenomenon described among rickettsial species in previous studies (Gall et al., 2016, Macaluso et al., 2002), this observation was supported by PCA. *Anaplasma marginale* and *Anaplasma* sp. Mymensingh were also detected in relatively low numbers from pools of *A. hebraeum* collected from cattle in this study. That these organisms were detected in only a small proportion of *A. hebraeum* pools could be attributed to studies that have shown that ticks of the *Rhipicephalus* genus: *R. simus*, *R. decoloratus*, *R. microplus*, *R. evertsi evertsi* and other ticks like *D. andersoni* are predominantly responsible for the transmission of *A. marginale* and *A. centrale* (Futse et al., 2003, Potgieter, 1987) and not *A. hebraeum*. 
This study provides the first report of the detection of the zoonotic agent *A. phagocytophilum* in humans, dogs and rodents in the Mnisi community in South Africa. This was confirmed by 16S rRNA (humans, dogs, rodents) and *gltA* (dogs, rodents) gene sequence analysis.

Previous reports of *A. phagocytophilum* and the closely related *Anaplasma* sp. SA dog strain in South Africa include the detection of *Anaplasma* sp. SA dog strain in three dogs in Bloemfontein (Inokuma et al., 2005) and domestic dogs in the Mnisi community area (Kolo et al., 2016). *Anaplasma phagocytophilum* was reportedly detected by PCR in 50% of *Rhipicephalus* spp. ticks collected from ruminants across four provinces in South Africa (Mtshali et al., 2015). There are other reports of the detection of *A. phagocytophilum* across Africa (Dahmani et al., 2015, Ghafar and Amer, 2012, Kelly et al., 2014, Kolo et al., 2016, M'Ghirbi et al., 2012, Mwamuye et al., 2017, Nakayima et al., 2014, Teshale et al., 2016). While the preceding studies suggest *A. phagocytophilum* detection across the breadth of Africa, caution is urged, as we have shown that organisms such as *Anaplasma* sp. ZAM dog and *Anaplasma* sp. SA dog are very closely related to *A. phagocytophilum*, and PCR primers could cross-react in some of the assays that have been employed. With the advent of high throughput detection methodologies and the consequent spate of 16S rRNA gene survey studies, there have been increasing numbers of distinct *Anaplasma*-like 16S rRNA gene sequences deposited in the databases. The relationship of these newly-detected agents to known pathogens, and their ability to serve as a source of cross-reaction in molecular testing, has not been well assessed. Therefore, we urge caution when assigning species designation for organisms in the *Anaplasma* genus, as the 16S rRNA genes are highly similar, and a similarity score of 97%, as used in many studies, will lead to misclassification.

The qPCR assay targeting the *msp2* gene used in this study cross-reacted with *Anaplasma* sp. ZAM dog DNA, explaining the very high percentage positives of 82% and 85% obtained from dogs and *R. sanguineus*. The assay therefore detected both *A. phagocytophilum* and *Anaplasma* sp. ZAM dog in dogs and *R. sanguineus*, and is not specific for *A. phagocytophilum* in South Africa.

This study reported the first detection of *Anaplasma* sp. ZAM dog in *R. sanguineus*, thus expanding the known host range of the organism. The study also elucidated the presence of two sequence variants of the organism circulating in dogs and *R. sanguineus* in the Mnisi community. We found a four-nucleotide difference between the 16S rRNA gene sequences of *Anaplasma* sp. ZAM dog and *Anaplasma* sp. SA dog strain; and only one difference in the *gltA* deduced amino acid sequences. These data suggest that *Anaplasma* sp. ZAM dog and *Anaplasma* sp. SA dog strain are variants of the same species. Phylogenetic analyses furthermore grouped *Anaplasma* sp. ZAM dog and *Anaplasma* sp. SA dog into a distinct clade; providing sufficient divergence from other *Anaplasma* species to warrant classification as a separate species. Until appropriate type-material can be deposited and the species can be formally described, we will refer to this novel organism as *Anaplasma* sp. SA dog for *Anaplasma* sp. southern Africa dog. Our findings would further suggest that *R. sanguineus* should be considered as a possible vector for *Anaplasma* sp. SA dog in South Africa. It will be necessary, however, to undertake a tick transmission study to confirm the vectorial capacity of this tick species.
We obtained a single sequence variant of *A. platys* in dogs in the Mnisi community. In Africa, *R. sanguineus* is thought to be the reservoir host that plays a role in the transmission of *A. platys* (Sanogo et al., 2003), however, we did not detect *A. platys* in any of the *R. sanguineus* ticks sampled in this study, suggesting the presence of another vector in the Mnisi community. *Anaplasma platys* was also detected in low proportions in cattle in Mnisi. Since *A. platys* has been recognised for its zoonotic potential (Arraga-Alvarado et al., 2014, Breitschwerdt et al., 2014, Maggi et al., 2013b), its detection in dogs and cattle in this study is notable.

The *Anaplasma* Msp4 sequences in this study were conserved and all *ankA* sequences in this study were found to be 100% identical. It is important to note that we cannot be certain whether the *msp4* and *ankA* gene sequences obtained from dogs and *R. sanguineus* were derived from *A. phagocytophilum* or *Anaplasma* sp. SA dog because there are currently no *msp4* or *ankA* sequences available from *Anaplasma* sp. SA dog for comparison.

It is important to state that it was very difficult to amplify all the different gene products of *A. phagocytophilum* from dog, cattle and human samples using conventional PCR, and in fact all of the amplicons obtained for sequencing were from nested PCRs. This indicates that the level of *A. phagocytophilum* DNA in the array of samples tested was low. The EDTA blood samples used in this study were also relatively old, with samples from dogs collected in 2012 and those from cattle and humans collected in 2013.

This study served as the first report of the detection of *A. centrale* and *A. marginale* from rodents in South Africa. The detection of *A. marginale* and *A. centrale* in wild rodents was interesting as these species of *Anaplasma* are known to infect mostly ruminants. It has been reported previously (Horak et al., 2005) that *M. natalensis* is a good host for *Haemaphysalis elliptica* and *Rhipicephalus simus*, with a high burden of infestation of larvae having been observed on these rodents. *Rhipicephalus simus* has been proven to be experimentally competent for transmitting *A. marginale* and *A. centrale* (Potgieter, 1987).

This study also reported the first description of *Candidatus* Anaplasma boleense in South Africa from one cattle sample. Previous reports of the detection of this organism have been from mosquitoes and ruminants in China (Guo et al., 2016, Kang et al., 2014).

This study characterized the bacterial microbiome of 348 *A. hebraeum* collected from cattle in Mnisi, and reported that the bacterial population of *A. hebraeum* in the Mnisi community was significantly dominated by the zoonotic *R. africae* comprising 82.9% of the total sequences obtained from the salivary gland pools and 81.7% of the sequences from the midgut pools. It was also reported that all tick pools were infected with the pathogen. *Rickettsia africae* also made up 16% of the total sequences obtained from the blood of the AFI patients and was detected in three of the patients screened. In AFI patient 27, approximately 90% of the sequences corresponded to *R. africae*.

*Rickettsia africae* causes ATBF, an AFI in humans, characterized by headaches, inoculation eschars, neck pain and inflammation of the lymph nodes (Jensenius et al., 2003). A prior study on tick and flea pools collected from dogs in the study area (Kolo et al., 2016) and the results of the *A. hebraeum*
microbiome demonstrated the high prevalence and dominance of *R. africae* as the principal symbiont of *A. hebraeum* in the Mnisi community. It was therefore not surprising that one of the AFI patients had such a high burden of *R. africae* bacteraemia. There is usually a paucity of information on the detection of *R. africae* in indigenous populations, and more cases of ATBF have been reported in recent travellers to endemic regions of *A. hebraeum* (Althaus et al., 2010, Binder and Gupta, 2015, Hauser et al., 2016). Our findings correlated with a recent serological survey that revealed that 63.4% of AFI patients and 92.2% of cattle farmers, veterinary staff and herders tested in Mnisi had a past exposure and 24.3% of AFI patients had a recent exposure, having anti-*Rickettsia* antibodies (Simpson et al., 2018).

The findings of the *A. hebraeum* microbiome are in agreement with a previous study (Ross et al., 2018) that suggested *Rickettsia* spp. play a significant role as an important driver in the bacterial abundance that is present in ixodid ticks. Phylogenetic analysis of *R. africae* sequences from pools of *A. hebraeum* in this study and the three AFI patients showed sequences grouping with *R. africae* reference strain ESF-5 isolated from *A. variegatum* collected from cattle in Ethiopia (Fournier et al., 2009) supported with a fair bootstrap value. This study served as the first report of the molecular detection of the zoonotic *R. africae* from the blood of rural indigenous AFI patients in the Mnisi community, South Africa. It also highlighted the risk of transmission of *R. africae* to the human population of Mnisi given the highly aggressive nature of the vector *A. hebraeum* known to inflict numerous bites on exposed individuals (Jensenius et al., 2003).

*Brucella* spp. and *Ochrobactrum* spp., two phylogenetically closely related organisms, were detected from two rodents from Gottenburg, the periurban habitat, and a rodent from the Manyeleti protected area. *Brucella melitensis* was detected from the blood of a febrile patient and a cow in Mnisi. This finding correlates with the detection of antibodies to *Brucella* spp. in one of the AFI patients (Simpson et al., 2018). The detection of this very important zoonotic pathogen in this study is notable. We noted that the 16S rRNA gene was not a useful tool to discriminate between species of *Brucella* as sequences were 100% identical, which was in agreement with findings from previous studies (Gee et al., 2004, Vizcaíno et al., 2000).

The first detection of the zoonotic agent *C. burnetii*, the cause of Q fever, from a wild rodent in the communal grazing area of Tlhavekisa was reported in this study. Antibodies to *C. burnetii* were previously detected in 38.3% of AFI patients and 60.9% of cattle farmers, veterinary staff and herders tested in the Mnisi community (Simpson et al., 2018). The detection of this important pathogen in a rodent is thus notable. This study also detected *Coxiella* spp. symbionts making up 1.4% of the total sequences obtained from the salivary gland pools and 1.2% of the sequences obtained from the midgut pools of *A. hebraeum*. *Coxiella*-like endosymbionts (*Coxiella*-LEs) have been found to have a mutualistic association with ticks of the genera *Rhipicephalus, Amblyomma* and *Ornithodorus* (Bonnet et al., 2017, Duron et al., 2017, Klyachko et al., 2007, Lalzar et al., 2014, Machado-Ferreira et al., 2011), and it has been suggested that the abundance of *Coxiella*-LEs in the salivary glands of *A. americanum* is likely to inhibit the horizontal transmission of disease-causing pathogens (Klyachko et al., 2007). *Coxiella*-LEs have evolved to be maternally transmitted endosymbionts and have been found to be more associated with female ticks, primarily colonising the ovaries and the Malpighian tubules.
Since tick collection was biased toward male ticks in this study, this could explain the low proportion of Coxiella sequences that were obtained. However, it has been suggested that Coxiella may not be an obligate endosymbiont in some tick hosts (Fryxell and DeBruyn, 2016), and Coxiella-LEs have been rarely detected or not detected at all in ticks like I. ricinus (Carpi et al., 2011), Dermacentor, Hyalomma and Antricola (Duron et al., 2017).

Ehrlichia canis, the cause of canine monocytic ehrlichiosis (CME), made up 24% of the total sequences obtained from the blood of dogs in Mnisi and was detected in 60% of the dogs tested. It is interesting that a negative correlation of infection was observed between Anaplasma spp. and E. canis infection. Dogs that had high burdens of E. canis infection in their blood had lower burdens of Anaplasma spp. and vice versa. This phenomenon of bacterial interference between rickettsial species has been previously described in Dermacentor ticks (Gall et al., 2016, Macaluso et al., 2002).

This study also reported the detection of E. ruminantium, the cause of heartwater, comprising 6% of the total sequences obtained from the salivary gland pools and 3.4% of the sequences obtained from the midgut pools of A. hebraeum ticks. A previous study on tick-borne haemoparasites in cattle in the Mnisi community detected E. ruminantium in 5.8% of animals tested on the reverse line blot (RLB) hybridization assay (Choopa, 2016). A negative correlation of infection was also observed between E. ruminantium and R. africae infection in the salivary glands and midgut pools of A. hebraeum. Pools that had a higher proportion of R. africae infection had a lower proportion of E. ruminantium and vice versa. This finding is also in agreement with the bacterial interference phenomenon described among rickettsial species in aforementioned studies (Gall et al., 2016, Macaluso et al., 2002).

This study detected the novel species Ehrlichia minasensis in cattle and A. hebraeum pools and Borrelia spp. from the blood of cattle in Mnisi. It also reported the detection of Ehrlichia sp. in rodents in Mnisi. Previous research conducted in the community speculated that interactions between cattle and wildlife in the area could result in spill over of zoonotic infections to livestock and humans (Musoke et al., 2015). Household dogs usually accompany their owners to the cattle dip tanks, thus there is a risk of transmission of ectoparasites and infections between dogs and cattle in Mnisi. Since prior studies have shown evidence of zoonotic diseases in Mnisi (Simpson et al., 2018), there is a possible risk of zoonotic disease transmission between livestock, dogs and humans in the area.

The emerging pathogen, Achromobacter xylosoxidans, was detected in this study, and comprised 21.4% of the total sequences obtained from dogs in the Mnisi area and 3.6% of the sequences from the AFI patients. This could be an important finding since a previous study speculated that owners of dogs share a more similar bacterial microbiome with their pet dogs than with other dogs, though the finding from that study was described from the skin microbiota (Song et al., 2013).

Mycoplasma haemocanis comprised 5% of the total sequences obtained from the dogs in the Mnisi community. Mycoplasma species such as M. ovis, M. haemofelis, M. suis, Candidatus Mycoplasma haemohominis, and Ca. M. haematoparvum have been described as emerging zoonotic pathogens posing health risks in animals and humans (Maggi et al., 2013a). It has been suggested that Rhipicephalus sanguineus may be the vector responsible for the transmission of M. haemocanis though
this has not yet been proven (Novacco et al., 2010). There is currently no information on the zoonotic potential of *M. haemocanis* with published reports of its detection only in dogs (Kaewmongkol et al., 2017, Mascarelli et al., 2016, Ravagnan et al., 2017, Zheng et al., 2017).

The opportunistic pathogens, *Herbaspirillum huttiense*, *Stenotrophomonas maltophilia*, and *Stenotrophomonas* spp., were detected in the blood of the AFI patients. These pathogens have been implicated in causing infection in immunosuppressed individuals. The detection of these organisms in the blood of AFI patients in the Mnisi community suggests, together with a prior study on the prevalence of HIV in South Africa that reported a 14.1% prevalence rate in rural residents in Mpumalanga Province (Shisana et al., 2014), that they may have been immunocompromised individuals.

All the microbiome analyses carried out in this study detected very low numbers of organisms that have been listed as possible contaminant genera from DNA extraction kits, PCR reagents and molecular grade water and have been sequenced from negative controls in high throughput 16S rRNA gene profiling studies (Salter et al., 2014). Although there was no observation of any amplification products in the negative controls of the 16S rRNA gene PCRs that were conducted, it is probable that these organisms were contaminants.

We expected to detect *Leptospira* spp. in the rodents from this study given that rodents are important reservoir hosts of the pathogen (Cosson et al., 2014), and a prior study conducted in the Mnisi community detected antibodies to *Leptospira* in 6.8% of AFI patients and 21.9% of cattle farmers, veterinary staff and herders (Simpson et al., 2018). Leptospires are usually shed in the urine of infected animals so it is usually more effective to detect the organism in DNA from urine or kidney samples using molecular methods (Merien et al., 1992) rather than in blood, the tissue that was sampled for our study. Leptospires are also usually detectable in the blood only in the acute phase of the infection (Merien et al., 1992).

We did not detect any *Rickettsia* spp. in the blood of the rodents in this study, contrary to what was reported previously (Essbauer et al., 2018). These authors detected pathogenic rickettsial organisms *R. conorii*, *R. massiliae*, *R. felis* and *R. helvetica* from wild rodents sampled across seven provinces in South Africa and Namibia. A recent study in China (Ge et al., 2018) also detected *Rickettsia* spp. which made up 60% of the bacterial sequences from the spleen of wild mice and shrews. However, these authors (Essbauer et al., 2018, Ge et al., 2018) used DNA from ear tissue samples and spleen while we used DNA extracted from blood. It has previously been shown that rickettsial pathogens are usually enclosed in the dermis, vascular endothelium and spleen (Bayliss et al., 2009, Hawley et al., 2007) so it is more likely to detect rickettsiae in such tissues rather than from blood.

A limitation of this study was that our samples were not from a related set of individuals, i.e. dogs, cattle and their associated ticks that all belonged to an AFI patient. A clearer picture of zoonosis shared in Mnisi community could be painted in future if dogs and cattle (and their associated ticks) belonging to AFI patients were sampled, or if rodents from households of AFI patients were caught. This was impossible to achieve in our study, although the best possible set of samples was chosen given the circumstances. Studies conducted on rodents and their ectoparasites have emphasized that the
characterization of the bacterial populations in hosts and their arthropod vectors is important in understanding the abundance of vector-borne pathogens as well as the roles and interactions that exist within these bacterial populations (Cohen et al., 2015, Rynkiewicz et al., 2015). However due to the existing regulations, we were unable to do so. Due to regulations by the Department of Agriculture, Forestry and Fisheries (DAFF) South Africa, we were not allowed to collect the ticks that were on the body of the rodents when they were captured. All ectoparasites were frozen with the rodent carcasses at -80°C. In ideal circumstances we would have collected the ticks, allowed them to digest their blood meals, performed salivary gland and midgut dissections and also gone ahead to characterize the bacterial microbiome of the ticks. Another limitation of this study was the small sample sizes of dogs, cattle and AFI patients tested in the microbiome studies; this was because available research funds played a role in the number of samples that could be screened using NGS techniques on the PacBio platform which is still relatively expensive. Therefore, we acknowledge that the statistical inferences from the study of the blood microbiome of the dogs, cattle and AFI patients may not be fully significant, but as a pilot study it provides a synoptic insight into the blood microbiome of dogs, cattle and AFI patients in the Mnisi community.

Given the number of potential zoonotic agents identified in this study, the human population of the Mnisi community should be educated on effective tick and rodent control, keeping food out of the reach of rodents and observing good sanitary practices to reduce the rodent population around households. These practices would help in the overall reduction of zoonotic disease burden in the community. This study also recommends that the current cattle dipping regimen practised in the community be improved, as there is a persistent infestation of cattle with ixodid ticks.

Recommendations for future studies include the development of a qPCR assay specific for the detection of *A. phagocytophilum* in South Africa and the exploration of the occurrence of *A. phagocytophilum* in other provinces in South Africa. Given the close relationship between *A. phagocytophilum* and *Anaplasma* sp. SA dog, these organisms should be further examined. Isolation of the organisms in culture and whole genome sequencing of both organisms is recommended. Vector competence studies should be attempted using *R. sanguineus* to determine if it can experimentally transmit *A. phagocytophilum* and *Anaplasma* sp. SA dog in the country. The molecular characterization of *Bartonella* spp., *R. africae*, *A. platys*, *Anaplasma* sp. SA dog, *Coxiella* spp., *Brucella* spp. and *M. haemocanis* in the Mnisi community could also be undertaken. A comparison of the bacterial diversity that exists in the blood of AFI patients versus healthy non-AFI patients could be undertaken to determine if significant differences exist between both sets. Finally, the molecular detection of *Rickettsia* spp. and *Leptospira* spp. from different hosts in Mnisi using other sample types rather than blood can be undertaken.

In conclusion, this study detected an array of zoonotic bacterial pathogens: *B. grahamii*, *B. henselae*, *Bartonella* spp., *C. burnetii*, *R. africae*, *Brucella* spp., *A. platys* and *A. phagocytophilum* from wild rodents, domestic dogs, cattle and their ticks and humans and highlights their significance as possible causes of non-malarial febrile illness in the Mnisi community area. Health care practitioners in the community should consider these pathogens in the differential diagnosis of non-malarial AFI.
8.1 References


of Anaplasma phagocytophilum and Anaplasma platys in cattle from Algeria. Comparative Immunology, Microbiology and Infectious Diseases, 39, 39-45.


Appendix 1

Table 1: Genbank accession numbers of sequences used in the phylogenetic analysis of *Anaplasma* species.

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Organism</th>
<th>Strain/Country of origin</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>16S rRNA gene</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U02521</td>
<td>A. phagocytophilum</td>
<td>Webster/US</td>
<td>Human</td>
</tr>
<tr>
<td>NC007797</td>
<td>A. phagocytophilum</td>
<td>HZ/US</td>
<td>Human</td>
</tr>
<tr>
<td>CP006618</td>
<td>A. phagocytophilum</td>
<td>Dog2/US</td>
<td>Dog</td>
</tr>
<tr>
<td>AF093789</td>
<td>HGE agent</td>
<td>US</td>
<td>Human</td>
</tr>
<tr>
<td>M82801</td>
<td>A. platys</td>
<td>US</td>
<td>Dog</td>
</tr>
<tr>
<td>KT982643</td>
<td>A. platys</td>
<td>India</td>
<td>Dog</td>
</tr>
<tr>
<td>LC269822</td>
<td>A. platys</td>
<td>Zambia</td>
<td>Dog</td>
</tr>
<tr>
<td>AY570540</td>
<td><em>Anaplasma</em> sp. South Africa dog 1245</td>
<td>South Africa</td>
<td>Dog</td>
</tr>
<tr>
<td>AY570539</td>
<td><em>Anaplasma</em> sp. South Africa dog 1076</td>
<td>South Africa</td>
<td>Dog</td>
</tr>
<tr>
<td>LC269823</td>
<td><em>Anaplasma</em> sp. ZAM dog</td>
<td>Zambia</td>
<td>Dog</td>
</tr>
<tr>
<td>KUS56025</td>
<td><em>Candidatus Anaplasma</em> boleense WHANSP</td>
<td>China</td>
<td><em>Anopheles sinensis</em></td>
</tr>
<tr>
<td>KUS56169</td>
<td><em>Candidatus Anaplasma</em> boleense WHANSA</td>
<td>China</td>
<td><em>Anopheles sinensis</em></td>
</tr>
<tr>
<td>MF576175</td>
<td><em>Anaplasma</em> sp. Mymensingh</td>
<td>Bangladesh</td>
<td>Cattle</td>
</tr>
<tr>
<td>KY924886</td>
<td>Uncultured <em>Anaplasma</em> sp. Dedessa</td>
<td>Ethiopia</td>
<td>Cattle</td>
</tr>
<tr>
<td>KJ410249</td>
<td><em>Anaplasma</em> sp. BL102-7</td>
<td>China</td>
<td><em>Hyalomma asiaticum</em></td>
</tr>
<tr>
<td>KP006405</td>
<td>Uncultured <em>Anaplasma</em> sp. Dog9</td>
<td>Philippines</td>
<td>Dog</td>
</tr>
<tr>
<td>U037775</td>
<td>A. bovis</td>
<td>South Africa</td>
<td>Cattle</td>
</tr>
<tr>
<td>AF309869</td>
<td>A. centrale</td>
<td>Strain Israel</td>
<td>Cattle</td>
</tr>
<tr>
<td>AY048816</td>
<td>A. marginale</td>
<td>St. Maries</td>
<td>Cattle</td>
</tr>
<tr>
<td>AF414870</td>
<td>A. ovis</td>
<td>South Africa</td>
<td>Sheep</td>
</tr>
<tr>
<td>U54806</td>
<td><em>Anaplasma</em> sp. Omatjenne</td>
<td>Namibia</td>
<td><em>Hyalomma truncatum</em></td>
</tr>
<tr>
<td>KP314237</td>
<td>A. capra</td>
<td>China</td>
<td>Tick</td>
</tr>
<tr>
<td>U11021</td>
<td>Rickettsia rickettsii</td>
<td>Sawtooth/US</td>
<td>Dermacentor andersoni</td>
</tr>
<tr>
<td><strong>GltA gene</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AF304140</td>
<td>A. marginale</td>
<td>Florida/US</td>
<td>Cattle</td>
</tr>
<tr>
<td>KJ410283</td>
<td>A. ovis</td>
<td>China</td>
<td>Tick</td>
</tr>
<tr>
<td>CP001759</td>
<td>A. centrale</td>
<td>Israel</td>
<td>Cattle</td>
</tr>
<tr>
<td>KUS56317</td>
<td>A. bovis</td>
<td>China</td>
<td>Aedes</td>
</tr>
<tr>
<td>AQW38804</td>
<td>A. bovis</td>
<td>China</td>
<td>Aedes albopictus</td>
</tr>
<tr>
<td>AF304137</td>
<td>A. phagocytophilum</td>
<td>California/US</td>
<td>Horse</td>
</tr>
<tr>
<td>CP015376</td>
<td>A. phagocytophilum</td>
<td>Norway variant2/Norway</td>
<td>Sheep</td>
</tr>
<tr>
<td>AGR81592</td>
<td>A. phagocytophilum</td>
<td>Dog2/US</td>
<td>Dog</td>
</tr>
<tr>
<td>SVC65315</td>
<td>A. phagocytophilum</td>
<td>France</td>
<td></td>
</tr>
<tr>
<td>AKZ90811</td>
<td>A. phagocytophilum</td>
<td>Spain</td>
<td>Ixodes</td>
</tr>
<tr>
<td>KX987359</td>
<td><em>Candidatus Anaplasma</em> boleense</td>
<td>WHBMXZ/China</td>
<td>Boophilus</td>
</tr>
<tr>
<td>KR261628</td>
<td>A. capra</td>
<td>China</td>
<td>Haemaphysalis</td>
</tr>
<tr>
<td>KP861636</td>
<td>A. phagocytophilum</td>
<td>NY18/US</td>
<td>Ixodes</td>
</tr>
<tr>
<td>KP861637</td>
<td>A. phagocytophilum</td>
<td>DogL610/Germany</td>
<td>Ixodes</td>
</tr>
<tr>
<td>DQ525687</td>
<td>A. platys</td>
<td>Sicily/Italy</td>
<td>Dog</td>
</tr>
<tr>
<td>LC269826</td>
<td>A. platys</td>
<td>ZAM dog/Zambia</td>
<td>Dog</td>
</tr>
<tr>
<td>AY570541</td>
<td><em>Anaplasma</em> sp. SA dog</td>
<td>South Africa</td>
<td>Dog</td>
</tr>
<tr>
<td>AAT74599</td>
<td><em>Anaplasma</em> sp. SA dog</td>
<td>South Africa</td>
<td>Dog</td>
</tr>
<tr>
<td>LC269827</td>
<td><em>Anaplasma</em> sp. ZAM dog</td>
<td>Zambia</td>
<td>Dog</td>
</tr>
<tr>
<td>U59729</td>
<td>Rickettsia rickettsii</td>
<td>Iowa/US</td>
<td>Vero cells</td>
</tr>
</tbody>
</table>


Table 2: Sample information for the 16S rRNA, _gilA_, _msp4_ and _ankA_ sequences generated in this study.

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Sample ID</th>
<th>Species</th>
<th>Origin</th>
<th>Host</th>
<th>Sequence variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>MK814405</td>
<td>D2-1</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Aphi1/16S</td>
</tr>
<tr>
<td>MK814406</td>
<td>D2-2</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Aphi1/16S</td>
</tr>
<tr>
<td>MK814404</td>
<td>D2-4</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Aphi1/16S</td>
</tr>
<tr>
<td>MK814402</td>
<td>D2-5</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Aphi1/16S</td>
</tr>
<tr>
<td>MK814403</td>
<td>D2-8</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Aphi1/16S</td>
</tr>
<tr>
<td>MK814410</td>
<td>*D24</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Aphi2/16S</td>
</tr>
<tr>
<td>MK814412</td>
<td>*D28</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Aphi2/16S</td>
</tr>
<tr>
<td>MK814407</td>
<td>H59</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Human</td>
<td>Aphi1or2/16S</td>
</tr>
<tr>
<td>MK814411</td>
<td>*R98</td>
<td>A. phagocytophilum</td>
<td>Hialakahile</td>
<td>Mastomys natalensis</td>
<td>Aphi2/16S</td>
</tr>
<tr>
<td>MK814408</td>
<td>R102</td>
<td>A. phagocytophilum</td>
<td>Thhavekisa</td>
<td>Rattus tanezumi</td>
<td>Aphi1or2/16S</td>
</tr>
<tr>
<td>MK814409</td>
<td>R103</td>
<td>A. phagocytophilum</td>
<td>Thhavekisa</td>
<td>Mastomys natalensis</td>
<td>Aphi1or2/16S</td>
</tr>
<tr>
<td>MK814426</td>
<td>D3-1</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814428</td>
<td>D3-5</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814425</td>
<td>D3-6</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814423</td>
<td>D3-7</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814427</td>
<td>D5-1</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814441</td>
<td>D9-6</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog2/16S</td>
</tr>
<tr>
<td>MK814431</td>
<td>D27-2</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814435</td>
<td>D27-9</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog2/16S</td>
</tr>
<tr>
<td>MK814422</td>
<td>D27-1</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814429</td>
<td>D27-13</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814433</td>
<td>D36-3</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814432</td>
<td>D36-9</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814430</td>
<td>D36-10</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814434</td>
<td>D36-11</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814446</td>
<td>D37-1</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814447</td>
<td>D37-4</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814424</td>
<td>D37-8</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814445</td>
<td>D37-9</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/16S</td>
</tr>
<tr>
<td>MK814436</td>
<td>D3</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog2/16S</td>
</tr>
<tr>
<td>MK814438</td>
<td>D5</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog2/16S</td>
</tr>
<tr>
<td>MK814443</td>
<td>D27</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog2/16S</td>
</tr>
<tr>
<td>MK814439</td>
<td>*D36</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog2/16S</td>
</tr>
<tr>
<td>MK814437</td>
<td>D37</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog2/16S</td>
</tr>
<tr>
<td>MK814442</td>
<td>RA3</td>
<td>Anaplasma sp. SA dog</td>
<td>Athol</td>
<td>Rhipicephalus sanguineus</td>
<td>Adog2/16S</td>
</tr>
<tr>
<td>MK814440</td>
<td>RH3</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Rhipicephalus sanguineus</td>
<td>Adog2/16S</td>
</tr>
<tr>
<td>MK814444</td>
<td>RH8</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Rhipicephalus sanguineus</td>
<td>Adog2/16S</td>
</tr>
<tr>
<td>MK814416</td>
<td>D25-2</td>
<td>A. platys</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apla1/16S</td>
</tr>
<tr>
<td>MK814419</td>
<td>D25-5</td>
<td>A. platys</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apla1/16S</td>
</tr>
<tr>
<td>MK814414</td>
<td>D25-6</td>
<td>A. platys</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apla1/16S</td>
</tr>
<tr>
<td>MK814417</td>
<td>D25-8</td>
<td>A. platys</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apla1/16S</td>
</tr>
<tr>
<td>MK814418</td>
<td>D25-9</td>
<td>A. platys</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apla1/16S</td>
</tr>
<tr>
<td>MK814415</td>
<td>D33-3</td>
<td>A. platys</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apla1/16S</td>
</tr>
<tr>
<td>MK814421</td>
<td>D33-4</td>
<td>A. platys</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apla1/16S</td>
</tr>
<tr>
<td>MK814420</td>
<td>D33-5</td>
<td>A. platys</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apla1/16S</td>
</tr>
<tr>
<td>MK814413</td>
<td>D33-7</td>
<td>A. platys</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Apla1/16S</td>
</tr>
<tr>
<td>MK814448</td>
<td>*C5</td>
<td>Anaplasma sp. Mymensingh</td>
<td>Hialakahile</td>
<td>Dog</td>
<td>Asm1/16S</td>
</tr>
<tr>
<td>MK814450</td>
<td>*C13</td>
<td>Candidatus Anaplasma boleense</td>
<td>Seville A</td>
<td>Dog</td>
<td>Cab1/16S</td>
</tr>
<tr>
<td>MK814449</td>
<td>*C91</td>
<td>Anaplasma sp. Mymensingh</td>
<td>Seville A</td>
<td>Dog</td>
<td>Asm1/16S</td>
</tr>
</tbody>
</table>

**Gila**

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Sample ID</th>
<th>Species</th>
<th>Origin</th>
<th>Host</th>
<th>Sequence variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>MK804077</td>
<td>D24</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Aphi1/GltA</td>
</tr>
<tr>
<td>MK804079</td>
<td>D25</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Aphi1/GltA</td>
</tr>
<tr>
<td>MK804078</td>
<td>D28</td>
<td>A. phagocytophilum</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Aphi1/GltA</td>
</tr>
<tr>
<td>MK804080</td>
<td>R102</td>
<td>A. phagocytophilum</td>
<td>Thhavekisa</td>
<td>Rattus tanezumi</td>
<td>Aphi1/GltA</td>
</tr>
<tr>
<td>MK804081</td>
<td>D36</td>
<td>Anaplasma sp. SA dog</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>Adog1/GltA</td>
</tr>
</tbody>
</table>

**Msp4**

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Sample ID</th>
<th>Species</th>
<th>Origin</th>
<th>Host</th>
<th>Sequence variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>MK804088</td>
<td>D3</td>
<td>Anaplasma sp.</td>
<td>Hluvukani</td>
<td>Dog</td>
<td>AspM/Msp4</td>
</tr>
</tbody>
</table>
Table 3: Genbank accession numbers of sequences used for the phylogenetic analysis of *Bartonella* species.

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Organism</th>
<th>Strain/Country</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>NR104902</td>
<td>Bartonella vinsoni subsp. arupensis</td>
<td>OK 94-513/Sweden</td>
<td>Culture</td>
</tr>
<tr>
<td>DQ228135</td>
<td>Bartonella vinsoni subsp. berkhoffii</td>
<td>Q64SHD/China</td>
<td>Dog</td>
</tr>
<tr>
<td>AJ250247</td>
<td>Bartonella quintana</td>
<td>Italy</td>
<td>Pediculus humanus corporis</td>
</tr>
<tr>
<td>NR074335</td>
<td>Bartonella henselae</td>
<td>Houston-1/USA</td>
<td>Cat</td>
</tr>
<tr>
<td>EU111756</td>
<td>Bartonella queenslandensis</td>
<td>AUST/NH8/Australia</td>
<td>Melomys sp.</td>
</tr>
<tr>
<td>NR_074354</td>
<td>Bartonella tribocorum</td>
<td>IBS 566/France</td>
<td>Rattus sp.</td>
</tr>
<tr>
<td>NR115255</td>
<td>Bartonella rattimassiliensis</td>
<td>15908/France</td>
<td>Rattus norvegicus</td>
</tr>
<tr>
<td>CP001562</td>
<td>Bartonella grahamii</td>
<td>as4aup/Sweden</td>
<td>Apodemus flavicollis</td>
</tr>
<tr>
<td>NR116175</td>
<td>Bartonella rattraustraliani</td>
<td>AUST/NH4/Australia</td>
<td>Rattus tunneyi</td>
</tr>
<tr>
<td>NR025889</td>
<td>Bartonella elizabethae</td>
<td>F9251/USA</td>
<td>Human</td>
</tr>
<tr>
<td>NR115254</td>
<td>Bartonella phoceensis</td>
<td>16120/France</td>
<td>Rattus norvegicus</td>
</tr>
<tr>
<td>AB602532</td>
<td>Bartonella pachyuromydis</td>
<td>FN18-1/Netherlands</td>
<td>Pachyuromys duprasi</td>
</tr>
<tr>
<td>AY993936</td>
<td>Bartonella sp.</td>
<td>RF255Y/China</td>
<td>Rattus tanezumi flavipectus</td>
</tr>
<tr>
<td>NR042460</td>
<td>Brucella abortus</td>
<td>544/India</td>
<td>Cattle</td>
</tr>
</tbody>
</table>

* Sequences were extracted from data generated from the characterization of the bacterial blood microbiome of dogs, cattle and rodents in the Mnisi community area, Mpumalanga Province South Africa using next-generation sequencing techniques (Refer to Chapter 3 and 4).
Table 4: Genbank accession numbers of sequences used in the phylogenetic analysis of *Rickettsia* species.

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Organism</th>
<th>Strain/Country</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP001612</td>
<td><em>Rickettsia africarum</em></td>
<td>ESF-5/Ethiopia</td>
<td>Amblyomma variegatum</td>
</tr>
<tr>
<td>NR118678</td>
<td><em>Rickettsia sibirica</em></td>
<td>246/Russia</td>
<td>Human</td>
</tr>
<tr>
<td>CP003341</td>
<td><em>Rickettsia parkeri</em></td>
<td>Portsmouth/USA</td>
<td>Amblyomma maculatum</td>
</tr>
<tr>
<td>NR074480</td>
<td><em>Rickettsia conorii</em></td>
<td>Malish 7/South Africa</td>
<td>Human</td>
</tr>
<tr>
<td>CP002428</td>
<td><em>Rickettsia slovaca</em></td>
<td>13-B/Slovakia</td>
<td>Human</td>
</tr>
<tr>
<td>NR118837</td>
<td><em>Rickettsia peacockii</em></td>
<td>Skalkaho/USA</td>
<td>Dermacentor andersoni</td>
</tr>
<tr>
<td>CP018914</td>
<td><em>Rickettsia rickettsii</em></td>
<td>Iowa/USA</td>
<td>Dermacentor variabilis</td>
</tr>
<tr>
<td>CP003308</td>
<td><em>Rickettsia philipii</em></td>
<td>364D/USA</td>
<td>Human</td>
</tr>
<tr>
<td>NR025967</td>
<td><em>Rickettsia honei</em></td>
<td>RB/Austria</td>
<td>Human</td>
</tr>
<tr>
<td>AP017602</td>
<td><em>Rickettsia japonica</em></td>
<td>YH, Japan</td>
<td>Human</td>
</tr>
<tr>
<td>CP002912</td>
<td><em>Rickettsia heilongiangensis</em></td>
<td>054/China</td>
<td>Dymicoccus sylvanum</td>
</tr>
<tr>
<td>KY474575</td>
<td><em>Rickettsia raoultii</em></td>
<td>IM16/China</td>
<td>Human</td>
</tr>
<tr>
<td>RIRRGDK</td>
<td><em>Rickettsia helvetica</em></td>
<td>C9P9/Switzerland</td>
<td>Ixodes ricinus</td>
</tr>
<tr>
<td>CP013133</td>
<td><em>Rickettsia rhipicephali</em></td>
<td>HJ#5/Brazil</td>
<td>Haemaphysalis juxtakochi</td>
</tr>
<tr>
<td>CP015012</td>
<td><em>Rickettsia amblyommati</em></td>
<td>An13/Argentina</td>
<td>Amblyomma neumanni</td>
</tr>
<tr>
<td>NR074483</td>
<td><em>Rickettsia felis</em></td>
<td>URRWXCal2/USA</td>
<td>Ctenocephalides felis</td>
</tr>
<tr>
<td>RIRRGDF</td>
<td><em>Rickettsia bellii</em></td>
<td>USA</td>
<td>Dermacentor andersoni</td>
</tr>
<tr>
<td>CP003319</td>
<td><em>Rickettsia massiliae</em></td>
<td>AZT80/USA</td>
<td>Rhipicephalus sanguineus</td>
</tr>
<tr>
<td>NR044656</td>
<td><em>Rickettsia prowazekii</em></td>
<td>Brein1/Poland</td>
<td>Human</td>
</tr>
<tr>
<td>NR029155</td>
<td><em>Rickettsia canadensis</em></td>
<td>2678/Canada</td>
<td>Haemaphysalis leporispalustris</td>
</tr>
<tr>
<td>NR029154</td>
<td><em>Rickettsia akari</em></td>
<td>MK (Kaplan)/USA</td>
<td>Human</td>
</tr>
<tr>
<td>RIRRGDU</td>
<td><em>Rickettsia typhi</em></td>
<td>Wilmington/USA</td>
<td>Human</td>
</tr>
<tr>
<td>NR074500</td>
<td><em>Ehrlichia chaffeensis</em></td>
<td>Arkansas/USA</td>
<td>Human</td>
</tr>
</tbody>
</table>
# Appendix 2

## Approval documents

### Animal Ethics Committee

<table>
<thead>
<tr>
<th>PROJECT TITLE</th>
<th>Zoonotic tick-borne bacterial pathogens as a cause of acute febrile illness in South Africa</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROJECT NUMBER</td>
<td>V105-15</td>
</tr>
<tr>
<td>RESEARCHER/PRINCIPAL INVESTIGATOR</td>
<td>A Kolo</td>
</tr>
<tr>
<td>STUDENT NUMBER (where applicable)</td>
<td>UP_11233517</td>
</tr>
<tr>
<td>DISSERTATION/THESIS SUBMITTED FOR</td>
<td>PhD</td>
</tr>
<tr>
<td>ANIMAL SPECIES</td>
<td>Wild rodents x 600</td>
</tr>
<tr>
<td>NUMBER OF ANIMALS</td>
<td>Dogs x 100</td>
</tr>
<tr>
<td></td>
<td>Cattle x 100</td>
</tr>
<tr>
<td></td>
<td>Ticks x</td>
</tr>
<tr>
<td>Approval period to use animals for research/testing purposes</td>
<td>November 2015-November 2016</td>
</tr>
<tr>
<td>SUPERVISOR</td>
<td>Prof. M Oosthuizen</td>
</tr>
</tbody>
</table>

**KINDLY NOTE:**

Should there be a change in the species or number of animal/s required, or the experimental procedure/s - please submit an amendment form to the UP Animal Ethics Committee for approval before commencing with the experiment.

### APPROVED

<table>
<thead>
<tr>
<th>Date</th>
<th>30 November 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAIRMAN: UP Animal Ethics Committee</td>
<td>Signature</td>
</tr>
</tbody>
</table>

54285-J5
# Animal Ethics Committee

## Extension No. 1

<table>
<thead>
<tr>
<th>PROJECT TITLE</th>
<th>Zoonotic tick-borne bacterial pathogens as a cause of acute febrile illness in South Africa</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROJECT NUMBER</td>
<td>V105-15</td>
</tr>
<tr>
<td>RESEARCHER/PRINCIPAL INVESTIGATOR</td>
<td>Dr. A Kolo</td>
</tr>
<tr>
<td>STUDENT NUMBER (where applicable)</td>
<td>UP_11233517</td>
</tr>
<tr>
<td>DISSERTATION/THESIS SUBMITTED FOR</td>
<td>PhD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANIMAL SPECIES</th>
<th>Wild rodents x 600</th>
<th>Cattle x 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUMBER OF ANIMALS</td>
<td>Dogs x 100</td>
<td>Ticks x</td>
</tr>
</tbody>
</table>

Approval period to use animals for research/testing purposes: November 2016-November 2017

<table>
<thead>
<tr>
<th>SUPERVISOR</th>
<th>Prof. M Oosthuizen</th>
</tr>
</thead>
</table>

**KINDLY NOTE:**

Should there be a change in the species or number of animal/s required, or the experimental procedure/s - please submit an amendment form to the UP Animal Ethics Committee for approval before commencing with the experiment.

**APPROVED**

<table>
<thead>
<tr>
<th>Date</th>
<th>22 November 2016</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>CHAIRMAN: UP Animal Ethics Committee</th>
</tr>
</thead>
</table>

**Signature**

S4285-15
Animal Ethics Committee

Extension No. 1

<table>
<thead>
<tr>
<th>PROJECT TITLE</th>
<th>Zoonotic tick-borne bacterial pathogens as a cause of acute febrile illness in South Africa</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROJECT NUMBER</td>
<td>V105-15</td>
</tr>
<tr>
<td>RESEARCHER/PRINCIPAL INVESTIGATOR</td>
<td>Dr. A Kelo</td>
</tr>
<tr>
<td>STUDENT NUMBER (where applicable)</td>
<td>UP_11233517</td>
</tr>
<tr>
<td>DISSERTATION/THESIS SUBMITTED FOR</td>
<td>PhD</td>
</tr>
<tr>
<td>ANIMAL SPECIES</td>
<td>Wild rodents x 600</td>
</tr>
<tr>
<td></td>
<td>Cattle x 100</td>
</tr>
<tr>
<td>NUMBER OF ANIMALS</td>
<td>Dogs x 56</td>
</tr>
<tr>
<td></td>
<td>Ticks</td>
</tr>
<tr>
<td>Approval period to use animals for research/testing purposes</td>
<td>January 2017-January 2018</td>
</tr>
<tr>
<td>SUPERVISOR</td>
<td>Prof. M Oosthuizen</td>
</tr>
</tbody>
</table>

**KINDLY NOTE:**

Should there be a change in the species or number of animal/s required, or the experimental procedure/s - please submit an amendment form to the UP Animal Ethics Committee for approval before commencing with the experiment

<table>
<thead>
<tr>
<th>APPROVED</th>
<th>Date</th>
<th>15 March 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAIRMAN: UP Animal Ethics Committee</td>
<td>Signature</td>
<td></td>
</tr>
</tbody>
</table>

S4285-15
2nd December 2015

RE Request for access to biobanked samples at Hans Hoheisen Wildlife Research Station

Dear Dr Agatha Kolo

We hereby give you permission to access up to 100 canine samples and 200 bovine samples from the Hans Hoheisen Wildlife Research Station biobank to conduct your testing. Please observe the standard operating procedures for utilisation of samples and be aware the samples are to remain at the laboratory and return to the biobank when your testing is complete.

Kind regards

[Signature]

Greg Simpson
The Research Ethics Committee, Faculty of Health Sciences, University of Pretoria complies with ICH-GCP guidelines and has US Federal-wide Assurance.
- IRB 0000 2235 IORG0001762 Approved dd 22/04/2014 and Expires 22/04/2017.

Faculty of Health Sciences Research Ethics Committee

Approval Certificate
New Application

26/05/2016

Ethics Reference No.: 152/2016

Title: Zoonotic tick-borne bacterial pathogens as a cause of acute febrile illness in South Africa

Dear Agatha Kolo

The New Application as supported by documents specified in your cover letter dated 26/04/2016 for your research received on the 26/04/2016, was approved by the Faculty of Health Sciences Research Ethics Committee on its quorate meeting of 25/05/2016.

Please note the following about your ethics approval:
- Ethics Approval is valid for 2 years
- Please remember to use your protocol number (152/2016) on any documents or correspondence with the Research Ethics Committee regarding your research.
- Please note that the Research Ethics Committee may ask further questions, seek additional information, require further modification, or monitor the conduct of your research.

Ethics approval is subject to the following:
- The ethics approval is conditional on the receipt of 6 monthly written Progress Reports, and
- The ethics approval is conditional on the research being conducted as stipulated by the details of all documents submitted to the Committee. In the event that a further need arises to change who the investigators are, the methods or any other aspect, such changes must be submitted as an Amendment for approval by the Committee.

Additional Conditions:
- Conditional approval, pending permission letter from NICD that patient specimens can be used for study

We wish you the best with your research.

Yours sincerely

Kindly collect your original signed approval certificate from our offices, Faculty of Health Sciences, Research Ethics Committee, Tswelopele Building, Level 4-59

Dr R Sommers; MBChB; MMed (Int); MPharmMed, PhD
Deputy Chairperson of the Faculty of Health Sciences Research Ethics Committee, University of Pretoria

The Faculty of Health Sciences Research Ethics Committee complies with the SA National Act 61 of 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 and 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki, the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles Structures and Processes 2004 (Department of Health).

012 356 3085 http://www.up.ac.za/healthethics
Private Bag X323, Arcadia, 0007 - Tswelopele Building Level 4-59, Gezina, Pretoria
agriculture, forestry & fisheries
Department: Agriculture, Forestry and Fisheries
REPUBLIC OF SOUTH AFRICA

Directorate Animal Health, Department of Agriculture, Forestry and Fisheries
Private Bag X138, Pretoria 0001
Enquiries: Mr Henry Gololo • Tel: +27 12 319 7532 • Fax: +27 12 319 7470 • E-mail: HenryG@daff.gov.za
Reference: 12/11/1/1/6; 12/11/1/1

Prof Marinda Oosthuizen
Department of Veterinary Tropical Diseases
Faculty of Veterinary Science
University of Pretoria

Dear Prof Oosthuizen,

RE: Permission to do research in terms of Section 20 of the ANIMAL DISEASES ACT, 1984 (ACT NO. 35 of 1984)

Your fax / memo / letter/ Email dated 19 October 2015, requesting an amendment to permission granted on 23 December 2015 under Section 20 of the Animal Disease Act, 1984 (Act No. 35 of 1984) to perform a research project or study, refers.

I am pleased to inform you that permission is hereby granted to perform the following amended research/study, with the following conditions:

Conditions:
1. All conditions stipulated within the approval dated 2014-12-23 must still be complied with;
2. The only allowable deviation from the research protocol for which Section 20 approval was granted on 2014-12-23 is the inclusion of the following postgraduate students to partake in the study:
   (i) Dr Agatha Kolo
   (ii) Dr Liesl De Boni
   (iii) Ms Samantha Willis

Title of research/study: Discovering emerging tick-borne pathogens that could impact on human health and livestock production in South Africa: Developing real-time PCR assays for the specific and sensitive detection of these pathogens

Researcher (s): Prof Marinda Oosthuizen, Dr Agatha Kolo, Dr Liesl De Boni, Ms Samantha Willis

Institution: Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, UP

Your Ref./ Project Number:
Our Ref Number: 12/11/1/1; 12/11/1/1/6

Kind regards,

[Signature]

DR. MPHLO MAJA
DIRECTOR OF ANIMAL HEALTH

Date: 2015-11-13
VETERINARY IMPORT PERMIT FOR ANIMAL SAMPLES IMPRINTED ONTO FTA PAPER FOR RESEARCH OR DIAGNOSTIC TESTING

[Issued in terms of the Animal Diseases Act, 1984 (Act No. 35 of 1984)]

Authority is hereby granted for you to import the following into Republic of South Africa:

2 X ANAPLASMA PHAGOCYTOPHILUM DNA CANINE ISOLATE (L610) SAMPLES IMPRINTED ONTO FTA PAPER FOR PCR VALIDATION

from: FREIE UNIVERSITAET BERLIN, GERMANY
subject to the following conditions:

1. The consignment must be accompanied by:
   1.1 this original veterinary import permit and
   1.2 a veterinary health certificate issued by a veterinarian authorised there to by the Veterinary Authorities of the exporting country to the effect that:
   i. He/she is not aware of any animal or human health or infectious risks associated with the import of these samples
   ii. These samples do not originate from an area under restrictions for any disease the species is susceptible to.

2. The impression samples/blood spots on the FTA paper must be completely dry, and securely packed in leakproof containers and sealed by an authorised official of the Veterinary Authorities of the exporting country. No fresh samples may be imported using this permit.

3. The consignment to be airfreighted through port of entry OR TAMBO INTERNATIONAL AIRPORT. Samples may only be imported as manifest cargo under an airwaybill number and may not be imported as personal luggage.

4. The consignment must be accompanied by this permit and its arrival reported immediately to the inspecting veterinary official: KEMPTON PARK Tel: 011 973 2827, and may not be released without his/her written permission.

5. Upon arrival the inspecting veterinary official will inspect the consignment and release it to the importer only after he/she is satisfied that all the import conditions have been complied with in full.

6. The specimens must be kept and used at the BSL2+, DEPT. OF VETERINARY TROPICAL DISEASES, FACULTY OF VETERINARY SCIENCE, under the personal supervision of MS ILSE VORSTER

7. On completion of tests/research the specimens, including all contaminated/infectious things or animal products (as defined by the Animal Diseases Act, 1984 [Act No. 35 of 1984]) derived/produced from or that came into contact with the above-mentioned specimens, must be destroyed by incineration. Records of the incinerations must be maintained for a period of 5 years, and made available for auditing to the Veterinary Authority upon request.

8. This permit does not absolve the importer from compliance with the provisions of any other legislation relating to this import