Biotechnology Master Program

Joint Biotechnology Master Program

Palestine Polytechnic University
Deanship of Graduate Studies and
Scientific Research

Bethlehem University
Faculty of Science

Identification of cross-protective potential antigens against pathogenic Brucella species through combining pan-genome analysis with reverse vaccinology

By

Yasmin Hisham Al-Dweik

In Partial Fulfillment of the Requirements for the Degree

Master of Science

August 2018

The undersigned hereby certify that they have read and recommend to the Faculty of Scientific Research and Higher Studies at the Palestine Polytechnic University and the Faculty of Science at Bethlehem University for acceptance a thesis entitled:

Identification of cross-protective potential antigens against pathogenic Brucella species through combining pan-genome analysis with reverse vaccinology

by : Yasmin Hisham Al-Dweik

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in biotechnology

Graduate Advisory Committee:

Committee Member (Student's Supervisor) Date

Dr. Yaqoub Ashhab, Palestine Polytechnic University

Committee Member (Internal Examiner) Date

Dr. Rami Arafeh, Palestine Polytechnic University

Committee Member (External Examiner) Date

Dr. Monueef Ayyash, Al-Quds University, Abu Dis

Approved for the Faculties

Dean of Graduate Studies and
Scientific Research

Palestine Polytechnic University

Date

Dean of Faculty of Science

Bethlehem University

Date

Abstract

Brucellosis is an infectious zoonotic disease caused by the *Brucella* bacteria. The three most pathogenic *Brucella* species are *Brucella abortus, Brucella melitensis, Brucella* suis. The most efficient approach to control the disease is to vaccinate domestic animals to prevent the spread to human. Several live attenuated strains are used as vaccines to control animal Brucellosis. However, as well as their pathogenicity to humans, these vaccines have a number of limitations, such as abortion, orchitis, and infertility in animals. Therefore, safer and universal vaccine is needed to control *Brucella* infection.

Reverse vaccinology (RV) has been far and wide applied for identifying potential antigens of pathogens of interest. In this study, a new RV strategy has provided, that combines pan-genome analysis with homemade meta-SCL pipeline, followed by determination of adhesion, antigenicity, and immunogenicity. A total of 90 complete *Brucella* genomes corresponding to the three most important zoonotic species, 55 *B. melitensis*, 17 *B. abortus*, and 18 *B.* suis*,* were subjected to the developed RV antigen selection steps (mentioned above). Using this approach enables the identification of 34 potential cross protective candidates. These candidates found to be associated to few biological functions that are highly referred to virulence such as cell adhesion and ion transport. The resulted antigens represent the most comprehensive set of antigens that can be used to develop a universal vaccine against zoonotic *Brucella* species. Further experiments are needed to clarify the mechanisms, attachment, immunogenicity and the protection level of these proteins.

الملخص

الحُمى المالطية هي أحد الأمراض المُعدية المشتركة بين الإنسان والحيوان والميكروب المسبب لهذا المرض بكتيريا تُسمى "البروسيلا". يتسبب المرض بخسائر اقتصادية و صحية فادحة في قطاع المواشي وتُصنف أنواع بكتيريا الـ بروسيلا حسب الحاضن الحيواني إلى أكثر من عشر أنواع أهمها نوع الـ *Brucella melitensis* الذي يصيب الماعز والأغنام ونوع *Brucella abortus* الذي يصيب الأبقار ونوع *Brucella* suis الذي يصيب الخنازير. وهذه الأنواع الثلاث تتميز بقدرتها العالية على الانتقال من الحيوان إلى الإنسان وتُعد الأكثر قدرة على الإمراض. تكون السيطرة على هذا المرض في الحيوانات من خلال اللقاحات المناعية التي تُعد الطريقة الأنجع للقضاء عليه حيث تتوافر بعض اللُّقاحات المرخصة لِتطعيم الحيوانات. مع ذلك، تعاني هذه اللُّقاحات من بعض المشاكل الجوهرية حيث انها تؤثر على السلامة في الانسان وتسبب عدد من الأعراض الجانبية في الحيوان مثل الإجهاض و العقم. لهذا، تولدت حاجة لتطوير لُقاح جديد أكثر أماناً وفعالية. مع توافر العديد من الخرائط الوراثية لعترات من أنواع البروسيلا المذكورة أعلاه و حيث ان بكتيريا البروسيلا تندرج ضمن ممرضات الدرجة الثالثة, يُعتبر تطبيق تصميم اللقاح العكسي (RV) في هذه الحالة استراتيجية مناسبة للوصول إلى لُقاح واعد. وعليه كان الهدف في هذه الدراسة هو تطوير استراتيجية مبنية على خطوات الُّلقاح العكسي للوصول الى المستفزات المناعية (المستضدات) المرشحة لكي تُستهدف من خلال تطوير لُقاح جديد ضد البروسيلا. آدخل في هذه الدراسة 90 جينوم كامل من انواع الـ بروسيلا الثلاثة الأكثر إمراضاً ولقد تم تحديد مجموعة من المستفزات المناعية الأكثر ترشيحاً حيث انها تحتوي على 34 محفز مناعي ليكونوا ضمن لُقاح شامل ضد الثلاثة أنواع المذكورة أعلاه.

DECLARATION

I declare that the master thesis entitled “Identification of cross protective potential antigens against pathogenic Brucella species through combining pan genome analysis with reverse vaccinology” is my own original work, and thereby certify that unless stated, all work contained within this thesis is my own independent research and has not been submitted for the award of any other degree at any institution, except where due acknowledgment is made in the text.

Name and signature: Yasmin Hisham Al-Dweik

Date: ________________________________

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for the joint master's degree in biotechnology at Palestine Polytechnic University and Bethlehem University, I agree that the library shall make it available to borrowers under rules of the library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made.

Permission for extensive quotation from, reproduction, or publication of this thesis may be granted by my main supervisor, or in his absence, by the Dean of Higher Studies when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature: Yasmin Hisham Al-Dweik

Date: ________________________

DEDICATION

I would like to dedicate this work to my biggest supporters, my precious father, my gorgeous mother, my beloved sisters, and my brothers. I could never thank you enough for being there for me even though we are far in the distance.

Acknowledgment

Foremost, I would like to express my special gratitude to my supervisor Dr.Yaqoub Al-Ashhab who gave me the golden opportunity to do this wonderful project, who also helped me in doing a number of small projects and I came to know about so many new things I am really thankful to him. he was as my father and my best teacher, he never stops encouraging and supporting.

Secondly, I would also like to extend my great appreciation to the BRC staff, Dr.Rami Arafeh, Dr.Robin Abu Ghazaleh, Asma Tamimi, Ayman Al Eadieh, Hassan Taradeh, and Zaid Taradeh, who were family to me.

Besides, I would like to thank my thesis committee: Dr, Rami Arafeh, and Dr. Muneef Ayyash.

I would acknowledge peoples who helped to finalize this project:
Asma Tamimi, Baraa Tamimi, Sakenah Ashhab, Suzan Hassouneh, Yusuf Ahmed, and Hind Hisham.

As well, I should acknowledge all my friends for support, especially AsmaTamimi, Kinda Qabaja, and Tasnim Ahmed.

My gratitude and deep appreciation to all ☺.

List of Abbreviations

%	Percentage
Spp.	Species
B. melitensis	*Brucella melitensis*
B. abortus	*Brucella abortus*
B. suis	*B.* suis
Mb	Mega bases
CDC	Centers for Disease Control and Prevention
BSL	Biological Safety Level
WHO	World Health Organization
RV	Reverse Vaccinology
Ag	Antigen
SCL	Sub-Cellular Localization
OM	Outer Membrane
IM	Inner Membrane
Ex	Extracellular
P	Periplasmic
CYT	Cytoplasmic
MHC	Major Histocompatibility Complex
CD	Conserved Domain
KDa	Kilo Dalton
aa	Amino Acid
ND	Not Determined

List of Figures

List of Tables

Table of Contents

CHAPTER 1: Introduction

1.1 *Brucellae*

Brucella is a genus of Gram-negative facultative intracellular bacteria that belongs to the *Brucella*ceae family of the order Rhizobiales in the class alpha-proteobacteria. This bacteria causes sever and persistent inflammatory disease in mammals that is known as Brucellosis. The genus *Brucella* was named after David Bruce in the late 19th century, a Scottish pathologist and microbiologist, who investigated a Mediterranean Fever that affected British soldiers during their military service in Malta island (Bruce, 1889). Currently, the genus consists of 10 species that are classified based on their host preferences (Ficht, 2010, Olsen and Palmer, 2014). These 10 species are *Brucella abortus, Brucella melitensis, Brucella* suis*, Brucella canis, Brucella ceti, Brucella inopinata, Brucella microti, Brucella neotomae, Brucella ovis,* and *Brucella pinnipedialis* (Figure 1) (Moreno, 2014, Ficht, 2010).

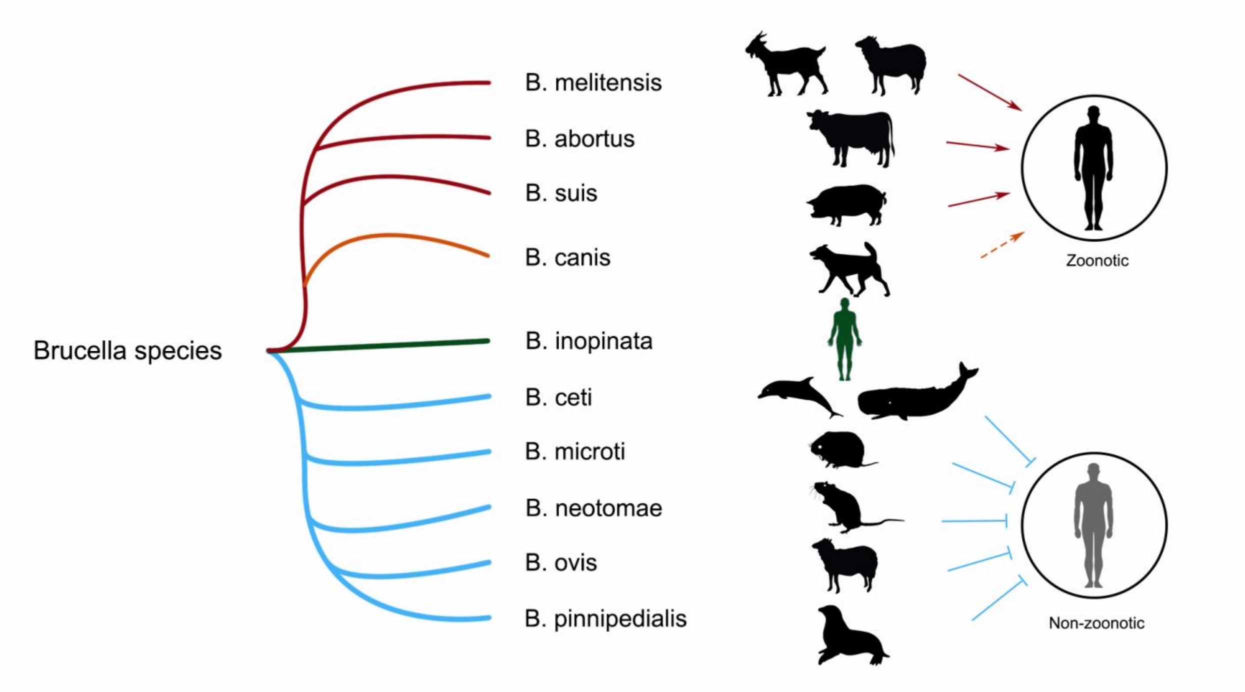


Figure 1 *Brucella* species with their preferred hosts, and zoonotic potential (Moreno, 2014).

Although several *Brucella* species are potentially zoonotic agents, *Brucella melitensis* (*B. melitensis*), *Brucella abortus* (*B. abortus*), and *Brucella* suis (*B.* suis), that primarily infect goats and sheep, cattle and swine respectively, are considered the most pathogenic *Brucella* species that have a serious impact on public health and the livestock industry (Smits, 2013, Seleem et al., 2010).

1.2 *Brucella* Genome

The genome of most *Brucella* species consists of two circular chromosomes with an average total size ~3.29 Mb. Comparative genomics studies revealed a high homology between the different species of *Brucella* and this is reflected in similar CG contents, genome sizes, and gene synteny (Ficht, 2010, Olsen and Palmer, 2014). On the other hand, the genetic phylogeny of *Brucella* genus does not reveal the *Brucella* preferred hosts phylogeny. The average total number of genes per genome is 3460 with 3180 protein coding genes. The number of pseudogenes within *Brucella* genomes is considered high compared to other bacterial genomes (Liu et al., 2004). It was observed that chromosome II has more pseudogenes than chromosome I, suggesting that chromosome II is more susceptible for gene transfer and rearrangements (Wattam et al., 2009). However, analysis of *Brucella* genomes reveals that pseudogene versus functional genes and horizontal gene transfer was proposed to be associated with the variation of the virulence and adaptation to the intracellular lifestyle in *Brucella* spp. (Wattam et al., 2009, Chain et al., 2005). Among the genes that were suggested to be associated with the intracellular adaptation of *Brucella*, and is important for its biogenesis, are the genes involved in VirB Type IV secretion system (T4SS) (Smith et al., 2016). Other genes include LPS biosynthesis, polysaccharide synthesis, ABC transporter, and outer membrane encoding genes (Jenner et al., 2010, O'callaghan and Whatmore, 2011).

1.3 Brucellosis

Brucellosis is a global zoonotic infectious disease that is caused by *Brucella* genus. Brucellosis is known by many other names such as Malta fever, undulant fever, Mediterranean fever, and

Bang disease (Smith and Bhimji, 2017). The incidence rate of this disease is continuously changing worldwide with an average rate of more than 500,000 human cases reported annually (Robinson, 2014). Although it is a worldwide disease, its predominantly found in developing countries of Central Asia, Africa, South America and the Mediterranean region which are currently listed as high risk areas by the Centers for Disease Control and Prevention (CDC) (Musallam et al., 2016, Dean et al., 2012). Several factors can affect the distribution of brucellosis such as socioeconomic, demographic, and occupational factors. However, the reported epidemiological data are likely underestimating the true burden of brucellosis (Dean et al., 2012). This infectious disease was mapped globally in 2006 (Pappas et al., 2006) and no further global mapping were exists. As it is shown in the map (Figure 2), the mediterranean region shows several countries with highest annual incidence and considered as an endemic area.

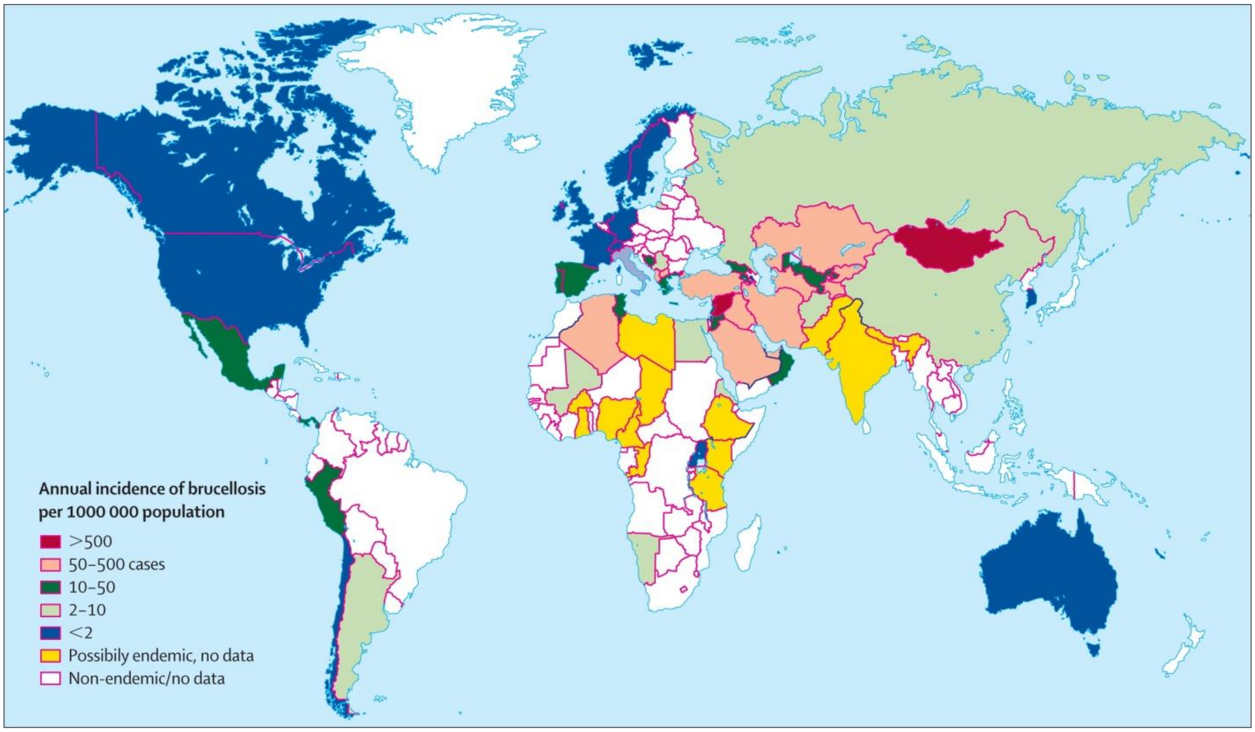


Figure 2 The Annual incidence of human brucellosis worldwide (Pappas et al., 2006).

In small ruminants, brucellosis have several clinical signs and symptoms including reproductive problems such as abortions, stillbirth, and infertility causing a significant economic loses and public health impact. As the reproductive system is the primary target of *Brucella*, the bacteria could be shed in the milk. Thus, infection commonly transmitted to humans by contact with the contaminated animal and/or animal products such as ingestion of unpasteurized milk or dairy products. Aerosol exposure to infected animal fluids is an additional way of transmission so that people who work with animals are at higher risk of exposure to *Brucella* such as veterinarians, livestock producers, laboratory personnel, and slaughterhouse workers (Corbel, 2006).

In humans, brucellosis can cause a severe febrile disease with various clinical complications ranging from mild to severe symptoms including undulant fever, flu-like signs, joint pain arthritis, endocarditis, meningitis and neurological complications (Pal et al., 2017, Gortázar et al., 2007, Franco et al., 2007). The incubation period vary from few days to several weeks, and the symptoms can last for several months or for years as in chronic cases. Treatment of human brucellosis includes a combination of several antibiotics for a minimum of 6–8 weeks. These antibiotics include Doxycycline with either streptomycin, rifampin, gentamicin, or sulfamethoxazole/trimethoprim (Ioannou et al., 2011). Although brucellosis can be treated with antibiotics, prolonged treatment is inconvenient for patients beside the relatively high relapse rate.

Prevention is most effective strategy to fight human brucellosis. Prevention actions include vaccination of domestic animals to reduce brucellosis in human. Increasing society awareness to the hygiene, pasteurize the dairy products, adequate cooking of meat products is an important prevention measure. The strategy employed by the WHO to control brucellosis depends mainly on the massive vaccination of domestic animals so as to prevent disease spread to healthy animals and to humans (Heymann, 2005). Typically, after achieving a very low prevalence rate in domestic animals (below 1%), a strict surveillance strategy can be applied to get rid of infected animals (Blasco, 2010, Pérez-Sancho et al., 2015, Olsen and Stoffregen, 2005).

1.4 Brucellosis vaccines

Currently, all the used vaccines to control Brucellosis in animals were obtained during the mid of 20th century through blind screening and several in vitro passages of the bacteria to obtain a spontaneously attenuated strain that can serve as a good vaccine. The following are the live attenuated vaccines that are used to control brucellosis in animals: *B. abortus* strains S19 and RB51, *B. melitensis* strains Rev.1 and M5, and *B.* suis strain S2 (Avila-Calderón et al., 2013).

In 1923, Dr. John Buck has isolated *Brucella abortus* from the milk of American cattle and discovered the S19 vaccine strain which is a spontaneously attenuated smooth strain (Buck, 1930). At the beginning of the 1930s, S19 has been used worldwide to prevent brucellosis in cattle until the discovery of strain RB51 (McDiarmid, 1957, Schurig et al., 1991). However, its protection efficacy is not stable and still it has several drawbacks (Confer et al., 1985). A major drawback of S19 is that it has intact extracellular O-antigen lipopolysaccharide that gives it the smooth phenotype. This phenotype hampers the serological discrimination between naturally infected cattle from those immunized (Koh and Morley, 1981). Other drawbacks of S19 are that it can cause abortion when given to the animal during pregnancy, and it is virulent for human (Beckett and MacDiarmid, 1985, Nicoletti, 1981, Corner and Alton, 1981).

In contrast, B. *abortus* RB51 discovered in 1982 is a rough mutant strain derived from smooth wild-type B. *abortus* 2308 (Schurig et al., 1991). Hence, RB51 strain is safer than the previous vaccine strain and it does not confuse the diagnosis and it enables distinguishing between infected and vaccinated cattle. Subsequently, it has replaced S19 strain for cattle vaccination. However, S19 showed higher efficiency (Cheville et al., 1996). Furthermore, RB51 suffers from several drawbacks: the major drawback is its resistance to rifampicin, an antibiotic used for treating brucellosis, and as S19 strain, RB51 is virulent for human (Ashford et al., 2004, Schurig et al., 1991).

Rev-1 is a live attenuated vaccine strain derived from several passages of *B.melitensis* 6056 discovered in 1957 (Elberg and Faunce Jr, 1957). This strain is used for goats immunization and produce strong immunity lasting for more than two years (Entessar et al., 1967, Alton, 1966). However, this strong immunity produces strong anti-O antigen antibody titer going back to the problem of distinguishing between vaccinated animals and naturally infected ones (Marín et al., 1999). Another drawback of Rev-1 is that it is more virulent than S19 for human, and it is not safe enough for pregnant goats as it can be excreted in milk (Blasco and Diaz, 1993, Banai, 2002). In addition, Rev-1 vaccine strain is subjected to morphological and immunological changes affecting its impact as a stable vaccine (Bosseray, 1991, Issa and Ashhab, 2016).

*B.*suis strain 2 (S2) in an attenuated strain used animal mainly pigs vaccination orally via adding it to the animal's drinking water. It was discovered in 1953 after revising it from serial passages. S2 has a smooth phenotype but with a lesser virulent of the wild-type strain (Xin, 1986). Nevertheless, the route of immunization affects the dose reproducibility of the S2 vaccine. In addition, this strain is not effective against *B. melitensis* infection (Bosseray and Plommet, 1990, Verger et al., 1995).

Despite their extensive global use, these live attenuated vaccines suffer from various drawbacks, such as their pathogenicity to humans and residual virulence in animals, which can cause abortion, orchitis, and infertility (Adone et al., 2005, Goodwin and Pascual, 2016). Moreover, these vaccines do not enable differentiation of infected animals from vaccinated animals by serological detection methods. These drawbacks have prompted several research groups to attempt the development of safer subunit vaccines.

1.5 Sub-unit vaccine

The subunit vaccine uses part/s of the pathogen instead of the whole microorganism. The used parts can be protein/s, polysaccharide/s, or both to provoke the immune system. Unlike the live-attenuated vaccines, subunit vaccines have a superior safety as there is no risk for multiplication or harmful changes to revert the attenuated microbe to its virulent nature.

Two conditions are essential to design a good subunit vaccine; first, the selection of appropriate protective antigens, and second, the selection of a safe and efficient vehicle to deliver these antigens to the immune system in a way that can evoke a protective immune response.

During the last two decades, a number of *Brucella* antigens have been identified, such as Omp16, Omp19, Omp25, Omp31, SurA, Dnak; trigger factor (TF), ribosomal protein L7L12, bacterioferritin (BFR) P39, and lumazine synthase BLS (Oliveira and Splitter, 1996, Velikovsky et al., 2003, Al-Mariri et al., 2001, Ghasemi et al., 2015, Goel and Bhatnagar, 2012, Delpino et al., 2007, Cassataro et al., 2005, Pasquevich et al., 2009). Despite their ability to induce immune response they did not produce a significant level of protection. These antigens were selected based on empirical screening approaches that are typically laborious, expensive, and require strict safety precautions and particular lab facilities, because *Brucella* is classified as biosafety level (BSL) 3 microorganism. This insufficiency of the empirical methods represents a great need for a rational and comprehensive approach to discover potential antigen candidates that can be used to develop a safe and effective anti-*Brucella* vaccine.

1.6 Reverse vaccinology

In contrast to the conventional vaccine development approaches that require cultivation and extensive empirical screening, the reverse vaccinology (RV) approach is an interesting *in silico* approach to identify protective antigens of a given pathogen starting with its genomic data. The method was first developed by Rino Rappuoli and his co-workers to discover protective antigens of Serogroup B meningococcus (Rappuoli, 2000, Pizza et al., 2000). Since then, RV has been implemneted to identify protective antigens for a large number of pathogens (Seib et al., 2012, Delany et al., 2013). General RV steps for vaccine development are summarized in Figure 3.

The RV criteria for selecting antigens is the most challenging step and it varies according to the pathogen and its properties such as the lifestyle, entry, invasion, attachment, survival, replication, and pathogenesis (Seib et al., 2012, Ong et al., 2017, Donati and Rappuoli, 2013).It has been shown that subcellular localization of the antigen plays a critical role in providing a protective antigen. In case of Gram negative bacteria, the surface exposed proteins are the most potential to be good vaccine candidate; while, in Gram positive the extracellular proteins are the most potential antigens (Ong et al., 2017). Furthermore, the proteins that enable pathogens to attach and invade host cells are considered very critical for a successful infection

(Ribet and Cossart, 2015). In addition, the infected host/s and the nature of the immune response involved in the eradication of the pathogen are other criteria that must be taken in consideration in antigen selection.

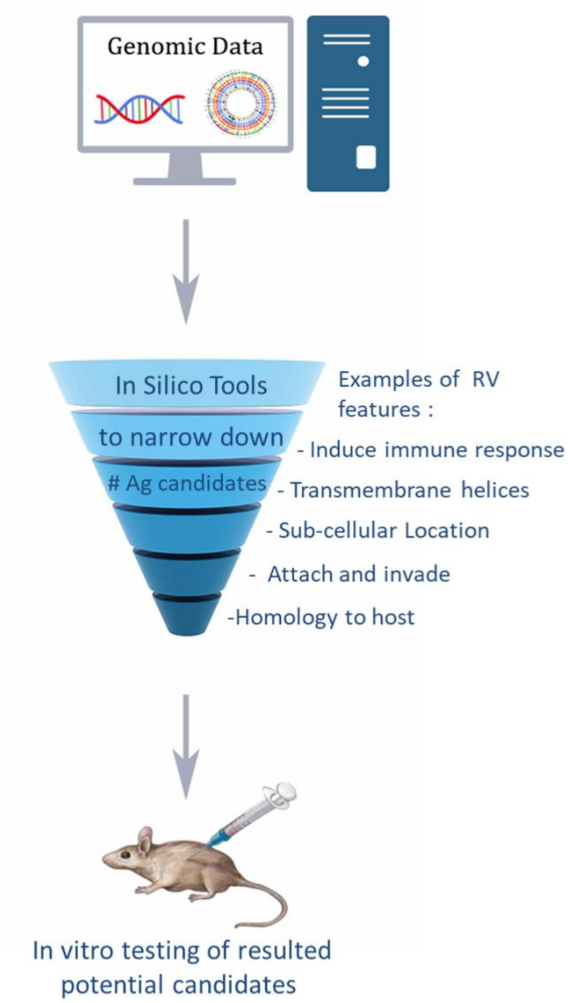


Figure 3 Overview of RV steps for vaccine development.

Limited number of studies have applied RV to identify *Brucella* antigens (Vishnu et al., 2015, Gomez et al., 2013). A major limitation of these studies is that they rely only on one strain, namely *B. melitensis* 16M to perform RV analysis. Moreover, they employed inadequate set of antigen selection criteria.

The concept of pan and core genome was first proposed in 2005 by Tettelin et al, dividing the pan genome of a collection of genomes of a given species into three sets of genes: core, dispensable, and special. Genes that are common in all selected genomes are called core genome; genes that are absent in one or more genomes are called dispensable; and genes that are unique to each genome are special (Tettelin et al., 2005). This analysis has been used for several purposes like studying the genetic diversity and relatedness across species, genera and even at bacterial family level. Due to the inter-strain gene content diversity, it has become crucial to perform a pan-genome analysis of several strains of a given bacterial species or genus to identify the core genome that contains the desired universal protective antigens (Donati and Rappuoli, 2013). Pan genome analysis has been widely used to identify the common genes between several genomes (core genome) that are attractive for the development of cross protective vaccine (Moriel et al., 2016, Zeng et al., 2017).

CHAPTER 2: Problem Statement and Objectives

Brucellosis is a major public health threat causing significant economic losses. Despite the fact that animal vaccination is the main strategy for reducing the disease and controlling it in animals and humans, current licensed vaccines suffer from several limitations. These limitations include abortion, orchitis, pathogenic to human and cannot differentiate between the infected and vaccinated animals. Therefore, there is a great need for developing a safer, more effective, and cross-protective vaccine.

Reverse vaccinology offers a powerful approach for discovering vaccine candidates especially for the pathogens that require special lab facilities. The availability of both free complete genomes of the pathogens and the advanced *in silico* tools, have made the RV approach more efficient. Moreover, as *Brucella* classified as BSL3, applying RV approach can reduce the time needed to develop a new vaccine against *Brucella*. Especially with the publishing of fully sequenced for many *Brucella* genomes.

In this study, the aim was to discover potential antigen candidates that are conserved among *B. melitensis*, *B. abortus*, and *B.* suis, which are the *Brucella* species frequently associated with human and domestic animal disease. The provided reverse vaccinology approach is an improved version based on determining the core genes of an extensive number of genomes from the three aforementioned *Brucella* species, followed by a rational antigen selection strategy. To our knowledge, this is the first study to combine pan-genome and reverse vaccinology approaches to identify potential protective antigen that can be used to develop a universal vaccine against the three most pathogenic *Brucella* spp.

CHAPTER 3: Materials and Methods

The *in silico* antigen prediction protocol is depicted in Figure 4. In the first phase, the retrieved proteomes were analyzed to extract the core proteome (the set of homologous proteins that are present in all analyzed strains of the three *Brucella* species). The identified core proteome is subsequently analyzed using a subcellular localization prediction pipeline to identify outer membrane and periplasmic proteins. In the last stage, various rigorous filters were employed to prioritize proteins based on features that are strongly associated with protective antigens, including adhesion, overall protein antigenicity, and density of B cell and T cell epitopes. Unless otherwise specified, the default parameters were used for all prediction tools.

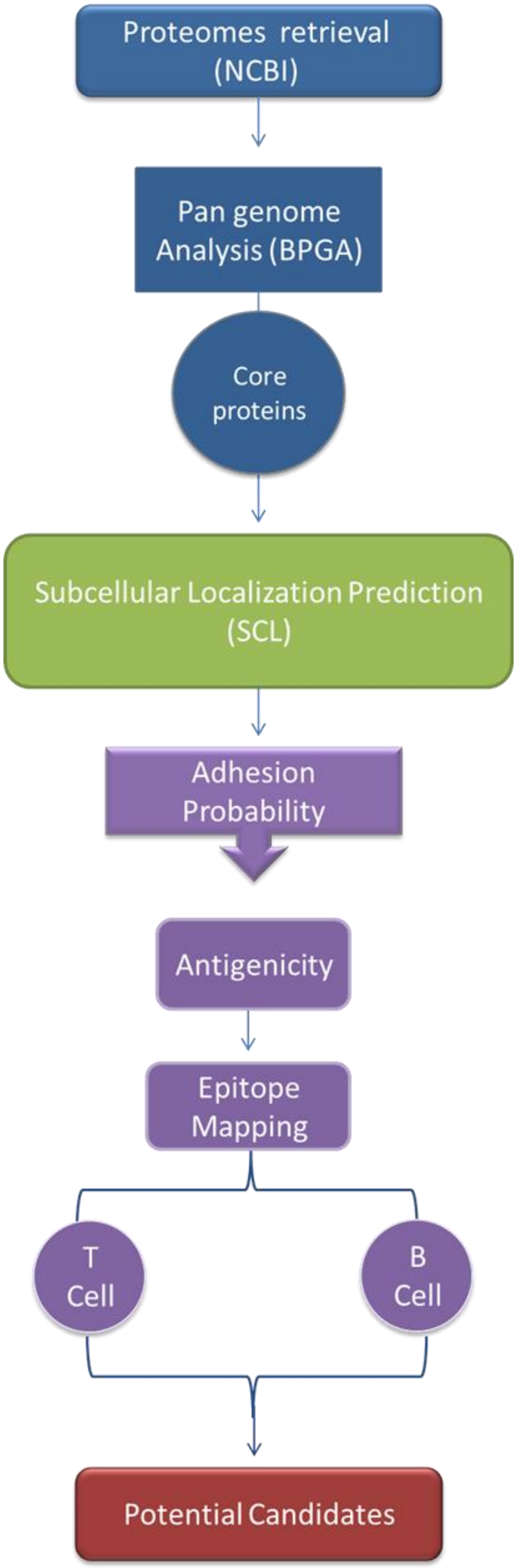


Figure 4 Summary of the reverse vaccinology protocol applied to the selected Brucella species for the selection of vaccine candidates.

3.1 Data retrieval

The full multi-Fasta format protein sequences of 55 *B. melitensis*, 17 *B. abortus*, and 18 *B.* suis genomes were downloaded from the Microbial Genomes Resources - NCBI (as of March 2018). Strain names, and number of proteins are shown in Tables 1, 2, 3 (for accession numbers see Appendix 1).

Table 1 *Brucella melitensis* proteomes selected for the study, their strain name and number of proteins.

	Organism/Name	Strain	Proteins
1.	***Brucella melitensis* bv. 1 str. 16M**	16M	3099
2.	***Brucella melitensis* ATCC 23457**	ATCC 23457	3152
3.	***Brucella melitensis* M28**	M28	3144
4.	***Brucella melitensis* M5-90**	M5-90	3124
5.	***Brucella melitensis* NI**	NI	3115
6.	***Brucella melitensis* bv. 3 str. Ether**	ether	3130
7.	***Brucella melitensis* bv. 1 str. 16M**	16M	3134
8.	***Brucella melitensis***	C-573	3099
9.	***Brucella melitensis***	20236	3140
10.	***Brucella melitensis***	2008724259	3125
11.	***Brucella melitensis***	BwIM_AFG_63	3148
12.	***Brucella melitensis***	BwIM_IRN_37	3149
13.	***Brucella melitensis***	BwIM_IRQ_32	3145
14.	***Brucella melitensis***	BwIM_ITA_45	3131
15.	***Brucella melitensis***	BwIM_ITA_55	3126
16.	***Brucella melitensis***	BwIM_SYR_04	3149
17.	***Brucella melitensis***	BwIM_TKM_56	3148
18.	***Brucella melitensis***	BwIM_TUR_03	3145
19.	***Brucella melitensis***	BwIM_TUR_17	3146
20.	***Brucella melitensis***	BwIM_TUR_19	3149
21.	***Brucella melitensis***	BwIM_TUR_59	3147
22.	***Brucella melitensis***	BwIM_SAU_09	3146
23.	***Brucella melitensis***	BwIM_SYR_26	3145
24.	***Brucella melitensis***	BwIM_TUR_39	3144
25.	***Brucella melitensis***	QY1	3058
26.	***Brucella melitensis***	BY38	3141
27.	***Brucella melitensis***	BL	3146
28.	***Brucella melitensis***	QH61	3131
29.	***Brucella melitensis***	CIIMS-BH-2	3130
30.	***Brucella melitensis***	B3	3117
31.	***Brucella melitensis***	Rev.1 (passage 101)	3111
32.	***Brucella melitensis***	1	3120
33.	***Brucella melitensis***	1	3112
34.	***Brucella melitensis***	1	3120
35.	***Brucella melitensis***	1	3117
36.	***Brucella melitensis***	1	3104
37.	***Brucella melitensis***	1	3124
38.	***Brucella melitensis***	1	3119
39.	***Brucella melitensis***	1	3123
40.	***Brucella melitensis***	1	3122
41.	***Brucella melitensis***	1	3103
42.	***Brucella melitensis***	1	3123

43. *Brucella melitensis*	1	3121
44. *Brucella melitensis*	1	3125
45. *Brucella melitensis*	1	3116
46. *Brucella melitensis*	1	3106
47. *Brucella melitensis*	1	3108
48. *Brucella melitensis*	1	3115
49. *Brucella melitensis*	1	3120
50. *Brucella melitensis*	1	3103
51. *Brucella melitensis*	1	3126
52. *Brucella melitensis*	1	3117
53. *Brucella melitensis*	1	3121
54. *Brucella melitensis*	1	3109
55. *Brucella melitensis*	1	3100

Table 2 *Brucella abortus* proteomes selected for the study, their strain name and number of proteins.

Organism/Name	Strain	Proteins
1. *Brucella abortus* bv. 1 str. 9-941	9-941	3153
2. *Brucella abortus* S19	S19	3151
3. *Brucella abortus* 2308	2308	3153
4. *Brucella abortus* A13334	A13334	3168
5. *Brucella abortus*	BDW	3171
6. *Brucella abortus*	BER	3130
7. *Brucella abortus*	NCTC 10505	3140
8. *Brucella abortus* bv. 9 str. C68	C68	3150
9. *Brucella abortus* bv. 6 str. 870	870	3142
10. *Brucella abortus*	63 75	3147
11. *Brucella abortus*	BFY	3127
12. *Brucella abortus* bv. 2 str. 86/8/59	86/8/59	3159
13. *Brucella abortus*	BAB8416	3121
14. *Brucella abortus* 104M	104M	3152
15. *Brucella abortus*	BD	3115
16. *Brucella abortus*	MC	3109
17. *Brucella abortus*	B4	3123

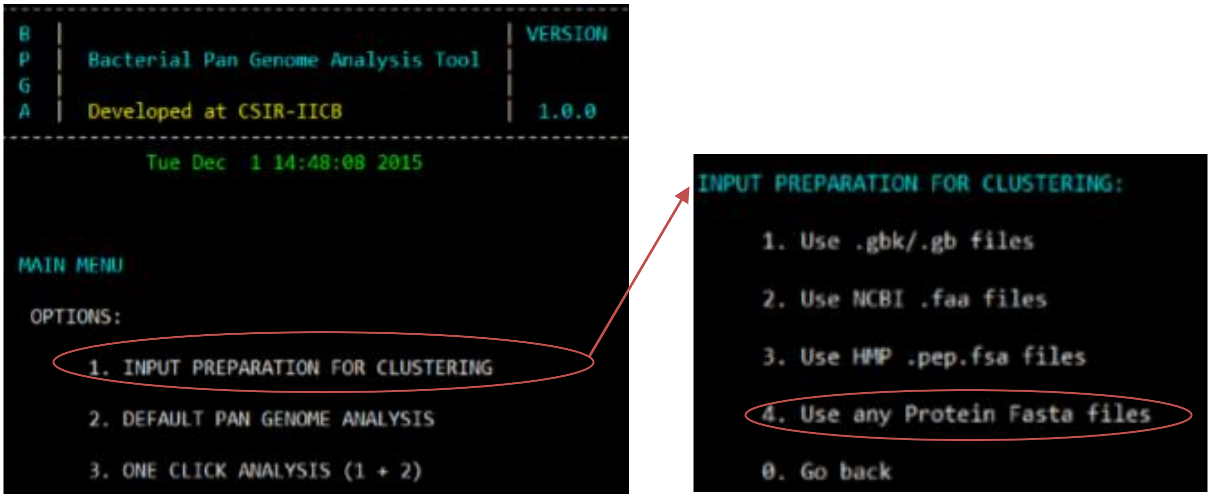
Table 3 *Brucella* suis proteomes selected for the study, their strain name and number of proteins.

Organism/Name	Strain	Proteins
1. *Brucella* suis 1330	1330	3161
2. *Brucella* suis ATCC 23445	ATCC 23445	3158
3. *Brucella* suis 1330	1330	3164
4. *Brucella* suis VBI22	VBI22	3167
5. *Brucella* suis bv. 1 str. S2	S2	3167

6. ***Brucella* suis bv. 2**	PT09143	3162
7. ***Brucella* suis bv. 2**	PT09172	3162
8. ***Brucella* suis bv. 2**	Bs364CITA	3163
9. ***Brucella* suis bv. 2**	Bs396CITA	3164
10. ***Brucella* suis bv. 2**	Bs143CITA	3162
11. ***Brucella* suis**	513UK	3165
12. ***Brucella* suis bv. 3 str. 686**	686	3121
13. ***Brucella* suis**	BSP	3153
14. ***Brucella* suis**	Human/AR/US/1981	3164
15. ***Brucella* suis**	2004000577	3154
16. ***Brucella* suis**	QH05	3151
17. ***Brucella* suis**	ZW046	3124
18. ***Brucella* suis**	ZW043	3090

3.2 Pan genome analysis

In order to identify the core proteins, the 90 proteomes were analyzed by the Bacterial Pan Genome Analysis (BPGA) tool using the default parameters (Chaudhari et al., 2016). In the input preparation for clustering step, option number 4 (use any protein Fasta files) was chosen. To ensure fast and accurate clustering, BPGA uses USEARCH as a default protein clustering tool with an identity cut off = 50%.



```
B |                                      | VERSION
P |  Bacterial Pan Genome Analysis Tool  |
G |                                      |
A |  Developed at CSIR-IICB              |  1.0.0

         Tue Dec  1 14:48:08 2015


MAIN MENU

 OPTIONS:

    1. INPUT PREPARATION FOR CLUSTERING

    2. DEFAULT PAN GENOME ANALYSIS

    3. ONE CLICK ANALYSIS (1 + 2)
```

```
INPUT PREPARATION FOR CLUSTERING:

     1. Use .gbk/.gb files

     2. Use NCBI .faa files

     3. Use HMP .pep.fsa files

     4. Use any Protein Fasta files

     0. Go back
```

Figure 5: Screen shot of the BPGA selected steps.

3.3 Subcellular localization (SCL)

Next, the core proteome was analyzed to predict outer membrane and periplasmic proteins. In this step, a previously developed homemade pipeline for SCL prediction was performed (Y. Ashhab, unpublished data). The pipeline employs different SCL prediction tools in three phases of positive and negative selections (Figure 6). Positive selection was performed for outer membrane (OM) and/or periplasmic (P) proteins. Negative selection was performed for inner membrane (IM), cytoplasmic (CYT), and extracellular (EX) proteins.

The tools used in the first phase were: 1. PSORTb v3.0.2, 2. CELLOv.2.5, and 3. SosuiGramN (Yu et al., 2010, Yu et al., 2006, Imai et al., 2008). In this stage, the positive selection was implemented for proteins that were predicted as OM or P by at least two of the three tools and were therefore included. A negative selection was implemented for proteins that were predicted as IM, EX, or CYT by at least two of the three tools and these proteins were therefore excluded. Proteins that were predicted with "unknown" subcellular location by at least one of the three tools, and OM and/or P by one of the three tools were considered uncertain proteins and were subjected to the second phase of selection.

The tools used in the second phase of selection were: 1. ClubSub-P, and 2. ngLoc (Paramasivam and Linke, 2011, King and Guda, 2007). Again, resulting proteins were divided into three categories. Positive selection was implemented for proteins that were predicted as OM or P by at least one of the two tools and were therefore included. Negative selection was implemented for proteins that were predicted as IM, EX, or CYT by at least one of the two tools and these proteins were therefore excluded. Proteins predicted with "unknown" subcellular location by one of the two tools were defined as uncertain. These uncertain proteins were subjected to a third phase of selection with the meta prediction tool; MetaLoc (Magnus et al., 2012). Proteins in this final step were divided into two categories; included for OM and P or excluded for the other sites. Included proteins from the three phases were collected for further analysis.

3.4 Adhesion probability

Adhesion probability of the surface associated proteins that summed up from the SCL prediction was predicted by Vaxign tool (He et al., 2010). Proteins with adhesion score higher than 0.5 were selected for further analysis.

3.5 Protein Antigenicity

Antigenicity of surface associated proteins was predicted using two tools: AntigenPro which computed antigenicity based on amino acid sequence features (Cheng et al., 2005), and VaxiJen which computed antigenicity based on physicochemical properties of amino acid sequence (Doytchinova and Flower, 2007).

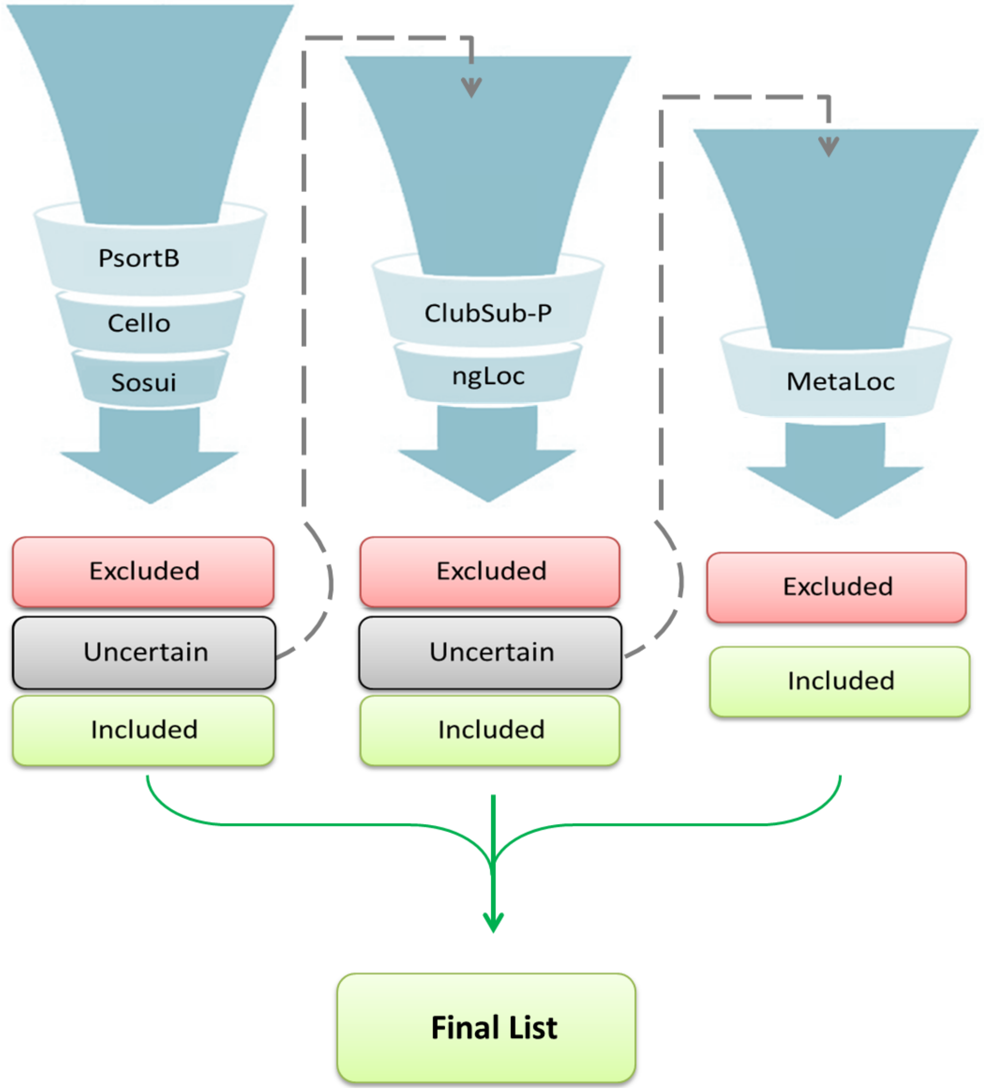


Figure 6: General workflow of the subcellular localization-voting strategy. A total 6 tools were applied to the core proteins (1939 proteins) that resulted from pan genome analyses to select surface- association proteins as novel vaccine candidates. The process starts with first group of tools consist of Psortb, Cello, and Sosui. The resulting proteins were classified to excluded, uncertain, and included. Proteins that predicted as cytosolic, inner membrane, and extracellular, by at least 2 of the 3 tools, were excluded. Proteins that predicted as periplasmic and outer membrane, by at least 2 of the 3 tools, were included. Proteins that predicted as unknown by at least 2 of the 3 tools or, not matched by any of the 3 tools, both types were considered as uncertain. The suspicious proteins were subjected to the second group of prediction tools that consist of ClubSub-P and ngLoc. The resulting proteins were also classified to excluded, uncertain, and included. Proteins that predicted as cytosolic, inner membrane, and extracellular, by at least 1 of the 2 tools, were excluded. Proteins that predicted as periplasmic and outer membrane by at least 1 of the 2 tools were included. Proteins that predicted as unknown by at least 1 of the 2 tools were considered as uncertain. The uncertain proteins were subjected to the final prediction tool; MetaLoc. Resulting proteins were classified in to 2 classes; included or excluded. Proteins predicted as cytosolic, inner membrane and extracellular were excluded. Proteins that predicted as periplasmic and outer membrane were included.

3.6 T-cell Epitope prediction

Surface associated proteins were also subjected to sequential epitope mapping in order to indicate their ability to bind to immune cells. T-cell epitopes were predicted for Major Histocompatibility Complex (MHC) class I and class II, and the number of potential binding alleles for each protein was determined. Propred1, and Propred were used for MHC class I and MHC class II epitopes, respectively (Singh and Raghava, 2003, Singh and Raghava, 2001). The epitope density in a given protein was calculated for each class of MHC by dividing the number of predicted epitopes over the length of that given protein. In addition, epitope coverage was calculated by dividing the number of alleles with positive predictions over the total number of analyzed alleles.

3.7 B-cell Epitope prediction

BCPred and AAPred were used for B-cell epitope prediction (Chen et al., 2007, EL-Manzalawy et al., 2008). Using the default parameters, epitopes with a score of 0.8 and greater were accepted. The density of the B-cell epitope for a given protein was calculated by dividing the number of predicted B-cell epitopes over the protein length.

3.8 Prioritization of protective antigens

In this step, a cumulative score for the proteins with adhesion score ≥ 0.5 was calculated using the prediction scores of protein antigenicity, MHC-I and MHC-II epitope densities, allele coverage for both classes of MHC, B-cell epitope density. The score for each feature was normalized to "1" as the highest possible value and "0" as the lowest possible value. The protein antigenicity score was the average of the two tools: VaxiJen score and AntigenPro score. The B-cell epitope density score was the average density of the two tools: AAP and BCPred.

3.9 Exclusion of dubious proteins

Proteins that show significant homology to host proteins, or proteins that have low molecular weight were excluded from the final list. To remove proteins with significant homology to host protein sequences, the selected antigens were subjected to homology search against proteomes using protein Blast (BLASTp tool at NCBI) with the following parameters: Database: reference proteins (refseq_protein); Organisms: human, sheep, goat, cattle, and pig; E-value cutoff: 0.001. Antigens that show ≥ 35% identity to any host protein were excluded. Molecular weight of small proteins was estimated using Expasy tool (Gasteiger et al., 2005). Proteins having a molecular weight of < 10 KDa were excluded.

3.10 Protein annotation and domain search

In addition to the "one line annotation description" provided by NCBI, a thorough manual annotation was performed to determine the most likely biological function assigned to the selected antigens. For this purpose, the following protein annotation servers were used: Blannotator, Panzzer, and EggNog (Kankainen et al., 2012, Koskinen et al., 2015, Huerta-Cepas et al., 2015). Furthermore, the conserved domain search was predicted using Blast CD-search tool. BOCTOPUS 2 was used to predict the topology of transmembrane beta-barrel proteins (Hayat et al., 2016).

CHAPTER 4: Results

Results of the established reverse vaccinology analysis to identify potential antigen candidates that can be used to develop a universal vaccine against *Brucella* are summarized in Figure 7. Starting with 282051 proteins, by the pan genome analysis the resulted core proteins are 1939 proteins. Next, they become 177 proteins after SCL-step. Then, number of proteins becomes 87 proteins that were evaluated for their antigenicity and immunogenicity properties. Top 34 proteins are submitted for further analysis.

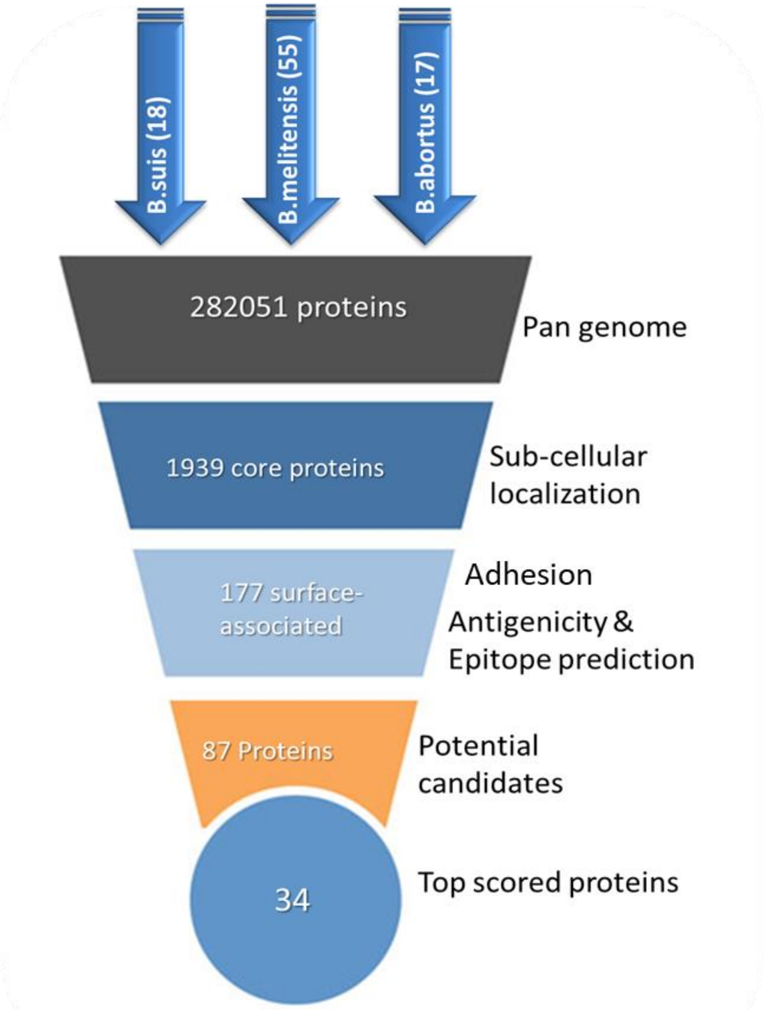


Figure 7: Summary of the resulted proteins in each step of the selection vaccine candidates' criteria.

4.1 Pan-genome analysis

Core proteins were identified first for each species alone, then for the three species together. The number of core proteins for 17 strains of *B. abortus* was 2840 proteins. The number of core proteins for 55 strains of *B.melitensis* was 2578 proteins. As for *B.*suis*,* for 18 strains there were 2484 proteins. The number of core proteins for all 90 proteomes of the three species was 1939 proteins. Figure 8 contains a Venn diagram of core proteins for the three species.

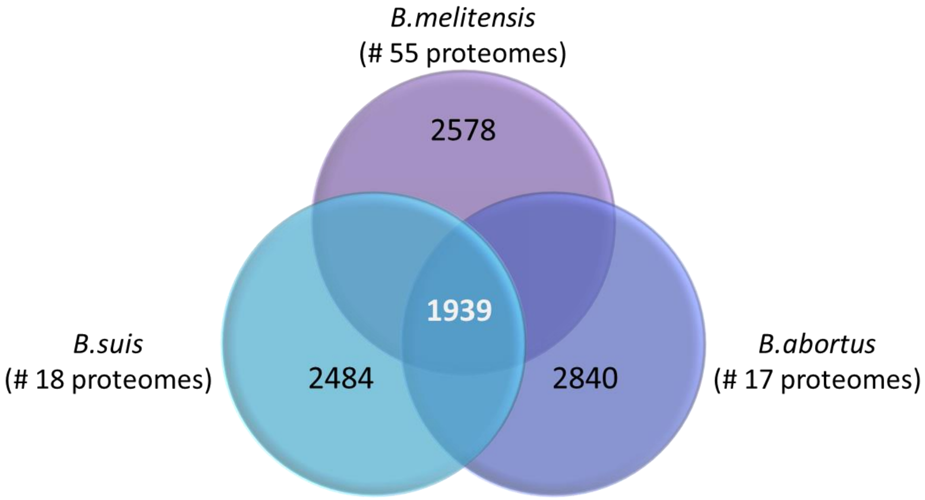


Figure 8: Venn diagram shows the core proteins of the three selected *Brucella* spp. (inter and intra). Core proteins within the 17 proteomes of *B. abortus* are 2840 proteins. While, core proteins within the 18 proteomes of *B.* suis are 2484 proteins. And, core proteins within the 55 proteomes of *B. melitensis* are 2578 proteins. Overall core proteins across three spp. are 1939 proteins.

The curves of gene accumulation for pan and core genes are shown in Figure 9. The number of gene families in both core (red curve) and pan (blue) genomes is plotted alongside the number of included genomes. The higher number of genomes shows the lower number of gene families within the core genome.

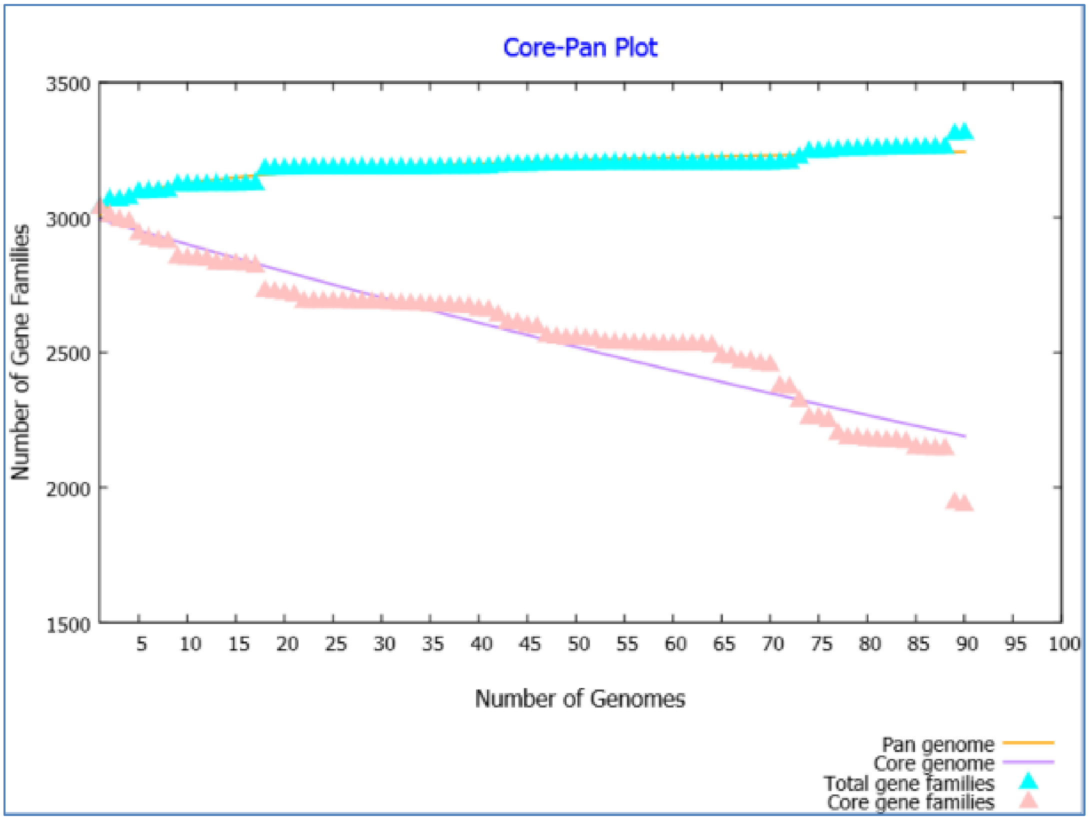


Figure 9: Core - Pan Plot present the distribution of the core and pan proteins.

4.2 Subcellular localization (SCL)

From the 1939 core proteins, the surface associated proteins were selected by this pipeline as shown in Figure 2. In the first phase, 151 proteins were included, 1639 were excluded, and 149 were labeled uncertain. These 149 proteins were subjected to the second phase of analysis in the pipeline, which excluded 104 proteins and included 16 proteins. The rest 29 uncertain proteins were subjected to the final phase of analysis. Of these 29 proteins, 19 were excluded, and 10 were included. Thus, the total number of proteins included from the three phases was 177 proteins, making up the final list of surface associated proteins (see Appendix 2).

4.3 Prioritization of protective antigens

As adhesion probability was shown to be a key feature common to many experimentally verified protective antigens (Ong et al., 2017), it was decide to use the adhesion scores, produced by Vaxign, to scale the 177 surface associated proteins in a descending order. The proteins with adhesion score ≥ 0.5 (38 proteins) were considered antigens with "high potential"; while those with adhesion score between 0.4 and 0.5 are considered antigens with "intermediate potential" (see Appendix 3). The 38 proteins with high potential were ranked based on a cumulative score that was derived from protein antigenicity, density of MHC-I and MHC-II epitopes, MHC allele coverage, and B-cell epitope density scores (Table 4) (For the detailed score calculation see Appendix 4). Of these 38 high potential proteins, Cytochrome c was excluded to avoid autoimmune response because of its homology to host proteins. In addition, 3 proteins with low molecular weight (6.7 KDa, 7.9 KDa, and 9.4 KDa) have been excluded as it is generally known that proteins with a molecular weight < 10 KDa are poorly immunogenic (Mohanty et al., 2013).

Among the 34 proteins classified as antigens with "high potential", 15 were annotated as hypothetical or unknown function. To gain more insight into the biological function of these proteins, the 34 proteins were manually annotated. The number of proteins with unknown function decreased to 4 (Table 4). The domain analysis showed that LomR is a frequently found domain among the antigens with high potential. This domain is a classical domain associated with many outer membrane proteins with transmembrane β-barrel scaffold that belong to Gram-negative porin superfamily.

The results of protein annotation were analyzed to identify any biological pattern that may be associated to predicted antigens (Table 4). Although there are little resources to investigate gene ontology of *Brucella* proteins, the 34 high potential antigens tend to be associated to just a few biological processes; including transmembrane transport (especially ions, iron, and small organic nutrients), membrane assembly, cell adhesion, and pathogenesis.

Table 4: High potential protein list - with their adhesion score, cumulative results, and consensus annotation resulting from Blannotator, Pannzer, and EggNOG tools.

Protein ID (NCBI)	Length (aa)	Single line annotation (NCBI)	Adhesion score	Score of 5	Annotation Note (by Blannotator, Pannzer, and EggNOG)	Domains (CD-search)	No. of transmembrane domains	Biological functions	Ref
WP_004684144.1	274	porin family protein	0.59	4.56	Porin opacity type (Pannzer), heat resistant agglutinin 1 (EggNOG).	**LomR**	8	small solutes transport, colonization, and adhesion	(Mancini et al., 2011, McClean, 2012)
WP_002964666.1	227	OmpW family protein	0.55	4.52	OmpW family outer membrane protein (Pannzer, EggNOG), Uncharacterized outer-membrane protein y4mB (Blannotator).	OmpW	8	Stress response, small solutes transport, bacterial colonization	(Wu et al., 2013, Hong et al., 2006, McClean, 2012)
WP_002969562.1	155	hypothetical protein	0.54	4.5	Not determined.	No Domain hits		ND	
WP_002966849.1	280	DUF1849 domain-containing protein	0.6	4.5	ATP/GTP-binding site domain-containing protein A (Pannzer), DUF1849 domain-containing protein (EggNOG).	DUF1849		Uptake of organic nutrient	
WP_002964611.1	351	DUF1176 domain-containing protein	0.54	4.44	DUF1176 domain-containing protein (EggNOG).	DUF1176		ND	
WP_004690357.1	284	porin family protein	0.55	4.42	Heat resistant agglutinin 1 (Pannzer, EggNOG), Uncharacterized protein BRA0921/BS1330_II0913 (Blannotator).	**LomR**	8	small solutes transport, colonization, and adhesion	(Mancini et al., 2011, McClean, 2012)
WP_004681227.1	238	type IV secretion system protein virB1	0.66	4.38	Type IV secretion system protein VirB 1 (Pannzer, Blannotator), Conjugal transfer protein (EggNOG).	lysozyme_like superfamily		adaptation to intracellular environment	(Comerci et al., 2001, Den Hartigh et al., 2004)
WP_002966226.1	182	hypothetical protein	0.54	4.38	UPF0423 protein BAB2_0840 (Blannotator), Pathogen-specific membrane antigen (Pannzer),	Tpd iron transport		iron acquisition, virulence	(Elhassanny et al., 2013,

					Periplasmic Protein (EggNOG).				Roset et al., 2017)
WP_002963597.1	121	hypothetical protein	0.5	4.37	Membrane-bound lysozyme inhibitor of C-type lysozyme **(by Blannotator, Pannzer, and EggNOG).**	MliC		Immune evasion, colonization/ virulence factor	(Callewaert et al., 2008)
WP_004688070.1	192	hypothetical protein	0.59	4.32	Not determined.	No Domain hit		ND	
WP_002966502.1	126	hypothetical protein	0.62	4.32	Outer membrane lipoprotein omp10 **(by Blannotator, Pannzer, and EggNOG)**	No Domain hit		virulence	(Tibor et al., 2002)
WP_002964322.1	329	hypothetical protein	0.59	4.23	31KDa transporter (Blannotator), Alkanesulfonate transporter substrate-binding subunit (Pannzer), trap transporter solute receptor taxi family (EggNOG).	TRAP_TAXI		nutrient transport, pathogenicity and colonization	(Rosa et al., 2018)
WP_002971090.1	267	hypothetical protein	0.54	4.22	Outer membrane beta-barrel domain protein (Pannzer).	OM_channels superfamily	10	Adhesion	(Fairman et al., 2011, McClean , 2012)
WP_004691650.1	620	TonB-dependent receptor	0.53	4.21	Iron compound TonB-dependent receptor (Pannzer), Involved in the active translocation of vitamin B12 (cyanocobalamin) across the outer membrane to the periplasmic space. It derives its energy for transport by interacting with the trans-periplasmic membrane protein TonB (By similarity) (EggNOG).	BtuB		iron acquisition and vitamin B12 transport	(Elhassanny et al., 2013, Roset et al., 2017, Liao et al., 2015)
WP_002971481.1	168	outer membrane protein assembly factor BamE	0.69	4.21	Outer membrane protein assembly factor BamE (Pannzer), smpa omla domain-containing protein (EggNOG).	BamE		cell envelop biogenesis, OMP assembly	(Sikora et al., 2018)
WP_004683739.1	236	porin family protein	0.56	4.21	Autotransporter outer membrane beta-barrel domain-containing protein-Fragment (Pannzer), hemin binding protein (EggNOG).	**LomR**	8	iron acquisition	(Mancini et al., 2011, McClean , 2012)

WP_004691134.1	403	hypothetical protein	0.62	4.15	Putative L,D-transpeptidase YafK (Blannotator), Pollen allergen Poa pIX/Phl pVI (Pannzer), ErfK ybiS ycfS ynhG family protein (EggNOG).	Yafk		Envelop biogenesis, stress response	(Sanders and Pavelka, 2013)
WP_002965482.1	439	sugar ABC transporter substrate-binding protein	0.55	4.11	ABC-type sugar transport system periplasmic component (Pannzer), extracellular solute-binding protein family 1 (EggNOG).	PBP2_TMBP_like		Uptake of organic nutrient, invasion/virulence	(Rosinha et al., 2002)
WP_002964333.1	220	OmpA family protein	0.56	4.09	Probable lipoprotein YiaD (Blannotator), Cell envelope biogenesis protein OmpA (Pannzer), Ompa motb domain protein (EggNOG).	OmpA		cell envelop biogenesis, adhesion, invasion/ intracellular survival, and evasion of host defense	(Confer and Ayalew, 2013)
WP_023080793.1	661	heme transporter BhuA	0.51	4.05	Heme transporter BhuA (Blannotator, Pannzer), receptor (EggNOG).	CirA superfamily		iron acquisition, virulence, association for bacterial persistence	(Elhassanny et al., 2013, Roset et al., 2017, Paulley et al., 2007)
WP_002964719.1	261	porin family protein	0.57	4.04	31 kDa outer-membrane immunogenic protein (Omp31) **(by Blannotator, Pannzer, and EggNOG)**.	**LomR**	8	haemin-binding proteins, virulence	(Delpino et al., 2006, Carroll et al., 2000, McClean , 2012)
WP_002966352.1	156	DUF2271 domain-containing protein	0.7	4.02	Tat pathway signal protein (Pannzer), Predicted periplasmic protein (DUF2271) (EggNOG).	DUF2271		ND	
WP_004690579.1	429	cell wall hydrolase	0.51	4	Cell wall hydrolase (Pannzer, EggNOG).	CwlJ		cell envelop biogenesis	(Nair et al., 2015)
WP_004683944.1	212	porin family protein	0.62	3.98	Omp 25 (Pannzer), Membrane (EggNOG).	**LomR**	**8**	Virulence, adhesion	(Cloeckaert et al., 2002, Fairman et al., 2011, McClean , 2012)

WP_011068938.1	792	LPS-assembly protein LptD	0.5	3.93	LPS-assembly protein LptD (Pannzer), involved in the assembly of LPS in the outer leaflet of the outer membrane. Determines N-hexane tolerance and is involved in outer membrane permeability. Essential for envelope biogenesis (By similarity) (EggNOG).	LptD		cell envelop biogenesis	(Chng et al., 2010)
WP_002967296.1	166	BA14K family protein	0.6	3.92	Immunoreactive BA14K (Pannzer, EggNOG).	BA14K		lectin-like activity, virulence	(Vemulapalli et al., 2006, Chirhart-Gilleland et al., 1998)
WP_002964622.1	170	BA14K family protein	0.58	3.92	Glutelin (Pannzer), BA14K (EggNOG).	BA14K		lectin-like activity, virulence	(Vemulapalli et al., 2006, Chirhart-Gilleland et al., 1998)
WP_004683466.1	213	membrane protein	0.55	3.9	25 kDa outer-membrane immunogenic protein Omp 25 (Blannotator, Pannzer), membrane (EggNOG).	**LomR**	8	Virulence, adhesion	(Cloeckaert et al., 2002, Fairman et al., 2011, McClean, 2012)
WP_002963776.1	115	DUF2147	0.6	3.89	sn-glycerol-3-phosphate ABC transporter ATP-binding protein (Pannzer), Uncharacterized protein conserved in bacteria (DUF2147) (EggNOG).	COG4731		nutrient transport, invasion/virulence	(Rosinha et al., 2002)
WP_004681306.1	367	iron ABC transporter substrate-binding protein	0.51	3.88	Periplasmic binding ABC transporter (Pannzer), solute-binding protein (EggNOG).	AfuA		iron acquisition, invasion/virulence	(Rosinha et al., 2002)
WP_002964998.1	177	hypothetical protein	0.67	3.72	Outer membrane lipoprotein omp19 **(by Blannotator, Pannzer, and EggNOG)**	Inh		protease inhibitor, alters the outer membrane properties	(Tibor et al., 2002, Coria et al., 2016)

WP_002964530.1	287	outer membrane protein assembly factor BamD	0.52	3.69	Outer membrane protein assembly factor BamD (Blannotator, Pannzer), Part of the outer membrane protein assembly complex, which is involved in assembly and insertion of beta-barrel proteins into the outer membrane (EggNOG).	BamD	cell envelop biogenesis, OMPs assembly, required for bacterial viability	(Sikora et al., 2018)
WP_006278325.1	261	hypothetical protein	0.5	3.69	Proline-rich region:Proline-rich extensin (Pannzer).	DNA_pol3_ gamma3 superfamily	ND	
WP_002963780.1	216	hypothetical protein	0.7	3.58	Not determined	No Domain hit	ND	

CHAPTER 5: Discussion

Brucellosis is a global zoonotic infection with a devastating economic impact on livestock sector and public health in many developing countries (McDermott et al., 2013). There is an unmet need to develop safe and efficient vaccine to fight brucellosis. This need was addressed in 2017 by launching a global prize competition of $30 million US for developing a safe and efficient vaccine against brucellosis (2018). The first step in developing such a vaccine would be to determine the protective antigens of these bacteria. Therefore, the aim of this study was to determine a set of universal and protective antigens that can be used to develop a vaccine against the three most pathogenic species of *Brucella* (*B. melitensis*, *B. abortus*, and *B.* suis) that are responsible for most cases of brucellosis among domestic animals and humans. A combined pan genome analysis with rational selection steps of reverse vaccinology to determine a manageable shortlist of *Brucella* antigens was performed. As a result 34 potential cross-protective antigens were identified from 90 complete proteomes covering the three species.

Although two recent studies have published their pan genome analysis results of *Brucella* spp. (Sankarasubramanian et al., 2015, Yang et al., 2016), it was decided to perform a new pan genome analysis because these two studies were performed with a relatively limited number of genomes to study the variation and relatedness among almost all species of *Brucella*, while the objective was to identify the core genome for *B. melitensis*, *B. abortus*, and *B.* suis.

A critical factor in applying a successful RV approach is to have a good understanding of the natural immune response to the pathogen of interest. In the case of *Brucella* infection, immunity is achieved by triggering both cellular and humoral mechanisms. Cell mediated immunity plays a critical role in protection against this intracellular bacteria and it is mainly mediated by Th1 response (Vitry et al., 2014). On the other hand, passive immunization of animals with antibodies from immunized animals provides protection against *Brucella* infection (Araya and Winter, 1990, Adone et al., 2012, Jain et al., 2017). Several studies have shown that surface associated antigens of Gram negative bacteria are not only essential to confer protective humoral immunity but also cell-mediated immunity against intracellular bacteria

(Barat et al., 2012, Bumann, 2014, Bras-Gonçalves et al., 2014). Therefore, the first RV filter was to identify outer membrane and periplasmic proteins of *Brucella*. Instead of using a single tool to identify these surface-associated proteins, a home-made pipeline was used which outperforms the currently available SCL prediction tools (Y. Ashhab, unpublished data). The established pipeline minimizes the possibility of excluding proteins that are assigned with unknown SCL, a scenario common in all SCL prediction tools.

In addition to surface-associated localization, it was suggested to use a feature that is strongly associated to protective immune response. Ong et al. investigated a large group of protective bacterial antigens to reveal the most prominent biological features shared among these proteins. They found that the two most important features among protective antigens of Gram-negative bacteria are adhesion and association with cell surface (Ong et al., 2017). Consequently, after predicting the list of surface-associated proteins (177 proteins), adhesion capability was predicted and used to rank these proteins.

It has been proven that proteins with high epitope density have significantly greater immunogenicity (Liu and Chen, 2005, Sette et al., 1994). Accordingly, proteins with high density of predicted epitopes are more potential vaccine candidates. Despite the growing numbers of immunobioinformatic tools that can predict MHC class I and Class II binding peptides; these tools are almost exclusive to human and mouse MHC alleles. Unfortunately, domestic animals, such as sheep, goats, and cows have limited MHC epitope data and prediction tools. However, a good agreement was noticed between the epitope prediction results of human and cow MHC alleles using the Propred server (See Appendix 3). This similar binding behavior would support the validity of the MHC scoring and its contribution to enhance the selection of universal antigens.

To examined the virulence and pathogenicity of the resulted proteins, Virulentpred, a virulence prediction tool (Garg and Gupta, 2008), and MP3, a meta pathogenicity prediction tool (Gupta et al., 2014) were used, respectively (data not shown). However, the results of these two tools were not informative or useful in ranking the antigens; the majority of the 177 surface associated proteins gave a positive prediction. Therefore, it was decided to exclude these two tools.

In this study, a rational reverse vaccinology approach was provided against the three most clinically important *Brucella* species. Two previous studies have employed reverse vaccinology to identify antigens of *B. melitensis* 16M (Vishnu et al., 2015, Gomez et al., 2013). However, these studies suffered from a number of limitations. The major limitation is that they were restricted to one genome and therefore their results cannot be extrapolated either to different strains of *B. melitensis* or to the different pathogenic species of *Brucella*. Although the two studies were performed on the same strain of *B. melitensis*, they have no overlapping in the final list of selected antigens.

In this study 34 proteins were identified as potential protective antigens that can serve to develop a novel universal vaccine against brucellosis. As 15 of these proteins have been deposited in GenBank without assigned function (11 hypothetical proteins and 4 proteins containing domains of unknown function “DUF”), it was decided to perform a thorough *in silico* analysis to gain more insight on the function of all the 34 proteins.

As shown in Table 2, the potential antigens tend to fall into a few categories of biological functions. An interesting protein family under these categories is the outer membrane proteins (OMPs) that possess 8-10 stands of β sheet. Of the 34 proteins, 8 belong to this subfamily of OMPs. Despite their involvement in transport of small solutes, it was found that small size OMPs (8-10 β sheet strands) tend to have a key role in adhesion, invasion and evasion to contribute to the tissue damage and bacterial spread across tissue barriers (McClean, 2012). Indeed, most of the shortlisted OMPs such as Omp10, Omp19, Omp25, Omp31, OmpA, and OmpW are associated with *Brucella* virulence and some of them showed a significant level of immune response when used as subunit vaccines (Tibor et al., 2002, Tadepalli et al., 2016, Confer and Ayalew, 2013, Simborio et al., 2016, Wu et al., 2013, Hur et al., 2011, Martín-Martín et al., 2009).

A second interesting group are proteins related to iron acquisition including the hypothetical protein "WP_002966226.1", TonB-dependent receptor "WP_004691650.1", heme transporter

BhuA "WP_023080793.1", and the iron ABC transporter substrate-binding protein "WP_004681306.1". The importance of iron for survival and virulence of *Brucella* is well documented and targeting proteins essential for iron acquisition is a promising strategy to develop effective bacterial vaccines (Cassat and Skaar, 2013).

A third group of proteins is the ABC transporters. This family of transporters is essential to secure uptake of various vital nutrients that cannot be produced by *Brucella*. It is believed that the ABC transporter proteins play a role in *Brucella* survival within the host cell during its infectious life cycle (Rosinha et al., 2002). Furthermore, it has been reported that the ABC proteins are able to induce immunity, making them potential vaccine targets (Riquelme-Neira et al., 2013, Issa and Ashhab, 2016).

An interesting identified candidate is VirB1, which is a component of the type IV secretion system (T4SS) of *Brucella* spp. This secretion system in *Brucella* is a well-known virulence factor, which is responsible for the survival, intracellular trafficking, and replication of *Brucella* inside the infected host cells (Comerci et al., 2001, Boschiroli et al., 2002, Dozot et al., 2006).

In the resulted list, some potential antigens were identified, that are periplasmic proteins with critical roles in outer membrane biogenesis and integrity. Among these proteins are BamD, and BamE, which are critical components of the β-barrel assembly machinery (BAM) (Konovalova et al., 2017).

Another interesting protein is the LPS-assembly protein LptD that is an essential component of the lipopolysaccharide transport (Lpt) machinery (Sperandeo et al., 2017). It is plausible that targeting one of these essential outer membrane biogenesis machineries would have a severe effect on bacterial survival.

Among the list of potential antigens, two proteins belong to a known immunoreactive protein family, BA14K. It has been mentioned that this family can strongly induce both cellular and humoral immune responses (Chirhart-Gilleland et al., 1998, Vemulapalli et al., 2006). Further investigation is needed to understand the functions of these two factors and their potential as protective antigens.

As the aim was to identify universal antigens that are conserved among the three pathogenic species (*B. melitensis*, *B. abortus*, and *B.* suis), it is possible that this approach could miss some interesting species specific antigens.

Although, the 177 surface associated proteins were ranked using adhesion as it is a crucial biological property strongly associated with a significant number of experimentally verified protective antigens, the possibility that some potential antigens are missed from the "high potential" 34 antigens cannot be exclude. In fact, a few interesting candidates were ranked in the "intermediate potential" antigens (see Appendix 3).

Among these interesting candidates are 26 kDa periplasmic immunogenic protein (Bp26) and Superoxide dismutase (SOD). Bp26, or immunoreactive Omp28, an antigen protein that is widely described as a potential vaccine candidate (Gomez et al., 2013, Cannella et al., 2012, Liang et al., 2010, Hur et al., 2011). In addition, it has been found to be immunogenic in both goats and humans and it provides a significant protection rate in BALB/c mice (Liang et al., 2010, Yang et al., 2005). SOD proteins have been reported in *Brucella abortus* and found to be responsible for host macrophage bursts. Thus, it is considered a promising antigen (Gee et al., 2005). This antigen has also been found in *Brucella melitensis* as an immunodominant protein (Yang et al., 2011). Moreover, SOD is considered a potential antigen with promising protective properties (Hur et al., 2011, He, 2012, Sáez et al., 2008). Here, two of superoxide dismutases, SOD_Cu-Zn, and SOD_Mn, were identified and found within the long list of universal vaccine targets.

It is worth to mention that the extended list of antigens, either with high and/or intermediate potential, does not contain various proteins that were previously suggested as possible antigens. The major reason for this is that most of these suggested antigens were not selected on rational basis that can differentiate simple antigens from antigens that can confer protective immunity. For example, the cytoplasmic proteins trigger factor (TF), bacterioferritin (BFR), and *Brucella* lumazine synthase (BLS), did not confer protective immunity when tested in mice (Velikovsky et al., 2003, Ghasemi et al., 2015, Al-Mariri et al., 2001). It is important to recall that these antigens are not surface-associated proteins.

CHAPTER 6: Conclusion

Bioinformatics is a very strong approach for vaccine candidate discovery as it offers a faster, cheaper and safer method to identify potential vaccine targets when compared with traditional laboratory identification methods, particularly when dealing with risk group 3 microorganisms such as *Brucella*. Here, a RV strategy that combines pan-genome analysis with a meta-SCL pipeline, followed by a rational based selection was provided that can rank surface-associated antigens according to their potential protective immunogenicity.

Using the established approach enables the identification of several potential cross protective candidates. The majority of the top ranked antigens are strongly associated to bacterial virulence and, therefore, it is plausible to assume that some of these antigens can form a solid base to design an efficient and safe vaccine against animal and human brucellosis. Further experiments are needed to test immunogenicity and protection level of these proteins.

References

2018. *APPLICATIONS NOW OPEN FOR THE BRUCELLOSIS VACCINE PRIZE!* [Online]. Available: https://brucellosisvaccine.org/ [Accessed].

ADONE, R., CIUCHINI, F., MARIANELLI, C., TARANTINO, M., PISTOIA, C., MARCON, G., PETRUCCI, P., FRANCIA, M., RICCARDI, G. & PASQUALI, P. 2005. Protective properties of rifampin-resistant rough mutants of *Brucella melitensis*. *Infection and immunity,* 73**,** 4198-4204.

ADONE, R., FRANCIA, M., PISTOIA, C., PETRUCCI, P., PESCIAROLI, M. & PASQUALI, P. 2012. Protective role of antibodies induced by *Brucella melitensis* B115 against B. *melitensis* and *Brucella abortus* infections in mice. *Vaccine,* 30**,** 3992-3995.

AL-MARIRI, A., TIBOR, A., MERTENS, P., DE BOLLE, X., MICHEL, P., GODFROID, J., WALRAVENS, K. & LETESSON, J.-J. 2001. Induction of Immune Response in BALB/c Mice with a DNA Vaccine Encoding Bacterioferritin or P39 of*Brucella* spp. *Infection and immunity,* 69**,** 6264-6270.

ALTON, G. 1966. Duration of the immunity produced in goats by the rev 1 *Brucella melitensis* vaccine. *J Comp Pathol*.

ARAYA, L. & WINTER, A. 1990. Comparative protection of mice against virulent and attenuated strains of *Brucella abortus* by passive transfer of immune T cells or serum. *Infection and immunity,* 58**,** 254-256.

ASHFORD, D. A., DI PIETRA, J., LINGAPPA, J., WOODS, C., NOLL, H., NEVILLE, B., WEYANT, R., BRAGG, S. L., SPIEGEL, R. A. & TAPPERO, J. 2004. Adverse events in humans associated with accidental exposure to the livestock brucellosis vaccine RB51. *Vaccine,* 22**,** 3435-3439.

AVILA-CALDERÓN, E. D., LOPEZ-MERINO, A., SRIRANGANATHAN, N., BOYLE, S. M. & CONTRERAS-RODRÍGUEZ, A. 2013. A history of the development of *Brucella* vaccines. *BioMed research international,* 2013.

BANAI, M. 2002. Control of small ruminant brucellosis by use of *Brucella melitensis* Rev. 1 vaccine: laboratory aspects and field observations. *Veterinary microbiology,* 90**,** 497-519.

BARAT, S., WILLER, Y., RIZOS, K., CLAUDI, B., MAZÉ, A., SCHEMMER, A. K., KIRCHHOFF, D., SCHMIDT, A., BURTON, N. & BUMANN, D. 2012. Immunity to intracellular Salmonella depends on surface-associated antigens. *PLoS pathogens,* 8**,** e1002966.

BECKETT, F. & MACDIARMID, S. 1985. The effect of reduced-dose *Brucella abortus* strain 19 vaccination in accredited dairy herds. *British Veterinary Journal,* 141**,** 507-514.

BLASCO, J. 2010. Control and eradication strategies for *Brucella melitensis* infection in sheep and goats. *Prilozi,* 31**,** 145-165.

BLASCO, J. & DIAZ, R. 1993. *Brucella melitensis* Rev-1 vaccine as a cause of human brucellosis. *The Lancet,* 342**,** 805.

BOSCHIROLI, M. L., OUAHRANI-BETTACHE, S., FOULONGNE, V., MICHAUX-CHARACHON, S., BOURG, G., ALLARDET-SERVENT, A., CAZEVIEILLE, C., LIAUTARD, J. P., RAMUZ, M. & O'CALLAGHAN, D. 2002. The *Brucella* suis virB operon is induced intracellularly in macrophages. *Proceedings of the National Academy of Sciences,* 99**,** 1544-1549.

BOSSERAY, N. 1991. *Brucella melitensis* Rev. 1 living attenuated vaccine: stability of markers, residual virulence and immunogenicity in mice. *Biologicals,* 19**,** 355-363.

BOSSERAY, N. & PLOMMET, M. 1990. *Brucella* suis S2, *Brucella melitensis* Rev. 1 and *Brucella abortus* S19 living vaccines: residual virulence and immunity induced against three *Brucella* species challenge strains in mice. *Vaccine,* 8**,** 462-468.

BRAS-GONÇALVES, R., PETITDIDIER, E., PAGNIEZ, J., VEYRIER, R., CIBRELUS, P., CAVALEYRA, M., MAQUAIRE, S., MOREAUX, J. & LEMESRE, J.-L. 2014. Identification and characterization of new Leishmania promastigote surface antigens, LaPSA-38S and LiPSA-50S, as major immunodominant excreted/secreted components of L. amazonensis and L. infantum. *Infection, Genetics and Evolution,* 24**,** 1-14.

BRUCE, D. 1889. Observations on Malta fever. *British medical journal,* 1**,** 1101.

BUCK, J. 1930. Studies of vaccination during calfhood to prevent bovine infectious abortion. *J. Agric. Res,* 41**,** 667-689.

BUMANN, D. 2014. Identification of protective antigens for vaccination against systemic salmonellosis. *Frontiers in immunology,* 5**,** 381.

CALLEWAERT, L., AERTSEN, A., DECKERS, D., VANOIRBEEK, K. G., VANDERKELEN, L., VAN HERREWEGHE, J. M., MASSCHALCK, B., NAKIMBUGWE, D., ROBBEN, J. & MICHIELS, C. W. 2008. A new family of lysozyme inhibitors contributing to lysozyme tolerance in gram-negative bacteria. *PLoS pathogens,* 4**,** e1000019.

CANNELLA, A. P., TSOLIS, R. M., LIANG, L., FELGNER, P. L., SAITO, M., SETTE, A., GOTUZZO, E. & VINETZ, J. M. 2012. Antigen-specific acquired immunity in human brucellosis: implications for diagnosis, prognosis, and vaccine development. *Frontiers in cellular and infection microbiology,* 2.

CARROLL, J. A., COLEMAN, S. A., SMITHERMAN, L. S. & MINNICK, M. F. 2000. Hemin-binding surface protein fromBartonella quintana. *Infection and immunity,* 68**,** 6750-6757.

CASSAT, J. E. & SKAAR, E. P. 2013. Iron in infection and immunity. *Cell host & microbe,* 13**,** 509-519.

CASSATARO, J., VELIKOVSKY, C. A., DE LA BARRERA, S., ESTEIN, S. M., BRUNO, L., BOWDEN, R., PASQUEVICH, K. A., FOSSATI, C. A. & GIAMBARTOLOMEI, G. H. 2005. A DNA vaccine coding for the *Brucella* outer membrane protein 31 confers protection against B. *melitensis* and B. ovis infection by eliciting a specific cytotoxic response. *Infection and immunity,* 73**,** 6537-6546.

CHAIN, P. S., COMERCI, D. J., TOLMASKY, M. E., LARIMER, F. W., MALFATTI, S. A., VERGEZ, L. M., AGUERO, F., LAND, M. L., UGALDE, R. A. & GARCIA, E. 2005. Whole-genome analyses of speciation events in pathogenic *Brucella*e. *Infection and immunity,* 73**,** 8353-8361.

CHAUDHARI, N. M., GUPTA, V. K. & DUTTA, C. 2016. BPGA-an ultra-fast pan-genome analysis pipeline. *Scientific reports,* 6.

CHEN, J., LIU, H., YANG, J. & CHOU, K.-C. 2007. Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. *Amino acids,* 33**,** 423-428.

CHENG, J., RANDALL, A. Z., SWEREDOSKI, M. J. & BALDI, P. 2005. SCRATCH: a protein structure and structural feature prediction server. *Nucleic acids research,* 33**,** W72-W76.

CHEVILLE, N. F., OLSEN, S., JENSEN, A., STEVENS, M., PALMER, M. & FLORANCE, A. 1996. Effects of age at vaccination on efficacy of *Brucella abortus* strain RB51 to protect cattle against brucellosis. *American journal of veterinary research,* 57**,** 1153-1156.

CHIRHART-GILLELAND, R. L., KOVACH, M. E., ELZER, P. H., JENNINGS, S. R. & ROOP, R. M. 1998. Identification and characterization of a 14-kilodalton *Brucella abortus* protein reactive with antibodies from naturally and experimentally infected hosts and T lymphocytes from experimentally infected BALB/c mice. *Infection and immunity,* 66**,** 4000-4003.

CHNG, S.-S., RUIZ, N., CHIMALAKONDA, G., SILHAVY, T. J. & KAHNE, D. 2010. Characterization of the two-protein complex in Escherichia coli responsible for lipopolysaccharide assembly at the outer membrane. *Proceedings of the National Academy of Sciences,* 107**,** 5363-5368.

CLOECKAERT, A., VIZCAíNO, N., PAQUET, J.-Y., BOWDEN, R. A. & ELZER, P. H. 2002. Major outer membrane proteins of *Brucella* spp.: past, present and future. *Veterinary microbiology,* 90**,** 229-247.

COMERCI, D. J., MARTÍNEZ-LORENZO, M. J., SIEIRA, R., GORVEL, J. P. & UGALDE, R. A. 2001. Essential role of the VirB machinery in the maturation of the *Brucella abortus*-containing vacuole. *Cellular microbiology,* 3**,** 159-168.

CONFER, A., HALL, S., FAULKNER, C., ESPE, B., DEYOE, B., MORTON, R. & SMITH, R. 1985. Effects of challenge dose on the clinical and immune responses of cattle vaccinated with reduced doses of *Brucella abortus* strain 19. *Veterinary microbiology,* 10**,** 561-575.

CONFER, A. W. & AYALEW, S. 2013. The OmpA family of proteins: roles in bacterial pathogenesis and immunity. *Veterinary microbiology,* 163**,** 207-222.

CORBEL, M. J. 2006. *Brucellosis in humans and animals*, World Health Organization.

CORIA, L. M., IBAÑEZ, A. E., TKACH, M., SABBIONE, F., BRUNO, L., CARABAJAL, M. V., BERGUER, P. M., BARRIONUEVO, P., SCHILLACI, R. & TREVANI, A. S. 2016. A *Brucella* spp. protease inhibitor limits antigen lysosomal proteolysis, increases cross-presentation, and enhances CD8+ T cell responses. *The Journal of Immunology,* 196**,** 4014-4029.

CORNER, L. & ALTON, G. 1981. Persistence of *Brucella abortus* strain 19 infection in adult cattle vaccinated with reduced doses. *Research in veterinary science,* 31**,** 342-344.

DEAN, A. S., CRUMP, L., GRETER, H., SCHELLING, E. & ZINSSTAG, J. 2012. Global burden of human brucellosis: a systematic review of disease frequency. *PLoS neglected tropical diseases,* 6**,** e1865.

DELANY, I., RAPPUOLI, R. & SEIB, K. L. 2013. Vaccines, reverse vaccinology, and bacterial pathogenesis. *Cold Spring Harbor perspectives in medicine,* 3**,** a012476.

DELPINO, M. V., CASSATARO, J., FOSSATI, C. A., GOLDBAUM, F. A. & BALDI, P. C. 2006. *Brucella* outer membrane protein Omp31 is a haemin-binding protein. *Microbes and infection,* 8**,** 1203-1208.

DELPINO, M. V., ESTEIN, S. M., FOSSATI, C. A., BALDI, P. C. & CASSATARO, J. 2007. Vaccination with *Brucella* recombinant DnaK and SurA proteins induces protection against *Brucella abortus* infection in BALB/c mice. *Vaccine,* 25**,** 6721-6729.

DEN HARTIGH, A. B., SUN, Y.-H., SONDERVAN, D., HEUVELMANS, N., REINDERS, M. O., FICHT, T. A. & TSOLIS, R. M. 2004. Differential requirements for VirB1 and VirB2 during *Brucella abortus* infection. *Infection and immunity,* 72**,** 5143-5149.

DONATI, C. & RAPPUOLI, R. 2013. Reverse vaccinology in the 21st century: improvements over the original design. *Annals of the New York Academy of Sciences,* 1285**,** 115-132.

DOYTCHINOVA, I. A. & FLOWER, D. R. 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC bioinformatics,* 8**,** 4.

DOZOT, M., BOIGEGRAIN, R. A., DELRUE, R. M., HALLEZ, R., OUAHRANI-BETTACHE, S., DANESE, I., LETESSON, J. J., DE BOLLE, X. & KÖHLER, S. 2006. The stringent response mediator Rsh is required for *Brucella melitensis* and *Brucella* suis virulence, and for expression of the type IV secretion system virB. *Cellular microbiology,* 8**,** 1791-1802.

EL-MANZALAWY, Y., DOBBS, D. & HONAVAR, V. 2008. Predicting linear B-cell epitopes using string kernels. *Journal of molecular recognition,* 21**,** 243-255.

ELBERG, S. S. & FAUNCE JR, K. 1957. IMMUNIZATION AGAINST *BRUCELLA* INFECTION VI. *Brucella melitensis*: Immunity Conferred on Goats by a Nondependent Mutant from a Streptomycin-Dependent Mutant Strain of1. *Journal of bacteriology,* 73**,** 211.

ELHASSANNY, A. E., ANDERSON, E. S., MENSCHER, E. A. & ROOP, R. M. 2013. The ferrous iron transporter FtrABCD is required for the virulence of *Brucella abortus* 2308 in mice. *Molecular microbiology,* 88**,** 1070-1082.

ENTESSAR, F., ARDALAN, A., EBADI, A. & JONES, L. 1967. Effect of living Rev. 1 vaccine in producing long-term immunity against *Brucella melitensis* infection in sheep in Iran. *Journal of comparative pathology,* 77**,** 367-376.

FAIRMAN, J. W., NOINAJ, N. & BUCHANAN, S. K. 2011. The structural biology of β-barrel membrane proteins: a summary of recent reports. *Current opinion in structural biology,* 21**,** 523-531.

FICHT, T. 2010. *Brucella* taxonomy and evolution. *Future microbiology,* 5**,** 859-866.

FRANCO, M. P., MULDER, M., GILMAN, R. H. & SMITS, H. L. 2007. Human brucellosis. *The Lancet infectious diseases,* 7**,** 775-786.

GARG, A. & GUPTA, D. 2008. VirulentPred: a SVM based prediction method for virulent proteins in bacterial pathogens. *BMC bioinformatics,* 9**,** 62.

GASTEIGER, E., HOOGLAND, C., GATTIKER, A., WILKINS, M. R., APPEL, R. D. & BAIROCH, A. 2005. Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook.* Springer.

GEE, J. M., VALDERAS, M. W., KOVACH, M. E., GRIPPE, V. K., ROBERTSON, G. T., NG, W.-L., RICHARDSON, J. M., WINKLER, M. E. & ROOP, R. M. 2005. The *Brucella abortus* Cu, Zn superoxide dismutase is required for optimal resistance to oxidative killing by murine macrophages and wild-type virulence in experimentally infected mice. *Infection and immunity,* 73**,** 2873-2880.

GHASEMI, A., JEDDI-TEHRANI, M., MAUTNER, J., SALARI, M. H. & ZARNANI, A.-H. 2015. Simultaneous immunization of mice with Omp31 and TF provides protection against *Brucella melitensis* infection. *Vaccine,* 33**,** 5532-5538.

GOEL, D. & BHATNAGAR, R. 2012. Intradermal immunization with outer membrane protein 25 protects Balb/c mice from virulent B. *abortus* 544. *Molecular immunology,* 51**,** 159-168.

GOMEZ, G., PEI, J., MWANGI, W., ADAMS, L. G., RICE-FICHT, A. & FICHT, T. A. 2013. Immunogenic and invasive properties of *Brucella melitensis* 16M outer membrane protein vaccine candidates identified via a reverse vaccinology approach. *PloS one,* 8**,** e59751.

GOODWIN, Z. I. & PASCUAL, D. W. 2016. Brucellosis vaccines for livestock. *Veterinary immunology and immunopathology,* 181**,** 51-58.

GORTÁZAR, C., FERROGLIO, E., HÖFLE, U., FRÖLICH, K. & VICENTE, J. 2007. Diseases shared between wildlife and livestock: a European perspective. *European Journal of Wildlife Research,* 53**,** 241.

GUPTA, A., KAPIL, R., DHAKAN, D. B. & SHARMA, V. K. 2014. MP3: a software tool for the prediction of pathogenic proteins in genomic and metagenomic data. *PloS one,* 9**,** e93907.

HAYAT, S., PETERS, C., SHU, N., TSIRIGOS, K. D. & ELOFSSON, A. 2016. Inclusion of dyad-repeat pattern improves topology prediction of transmembrane β-barrel proteins. *Bioinformatics,* 32**,** 1571-1573.

HE, Y. 2012. Analyses of *Brucella* pathogenesis, host immunity, and vaccine targets using systems biology and bioinformatics. *Frontiers in cellular and infection microbiology,* 2.

HE, Y., XIANG, Z. & MOBLEY, H. L. 2010. Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *BioMed Research International,* 2010.

HEYMANN, E. 2005. BRUCELLOSIS IN HUMANS AND ANIMALS.: WHO GUIDANCE. GENEVA. *World Health Organization*.

HONG, H., PATEL, D. R., TAMM, L. K. & VAN DEN BERG, B. 2006. The outer membrane protein OmpW forms an eight-stranded β-barrel with a hydrophobic channel. *Journal of biological chemistry,* 281**,** 7568-7577.

HUERTA-CEPAS, J., SZKLARCZYK, D., FORSLUND, K., COOK, H., HELLER, D., WALTER, M. C., RATTEI, T., MENDE, D. R., SUNAGAWA, S. & KUHN, M. 2015. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic acids research,* 44**,** D286-D293.

HUR, J., XIANG, Z., FELDMAN, E. L. & HE, Y. 2011. Ontology-based *Brucella* vaccine literature indexing and systematic analysis of gene-vaccine association network. *BMC immunology,* 12**,** 49-49.

IMAI, K., ASAKAWA, N., TSUJI, T., AKAZAWA, F., INO, A., SONOYAMA, M. & MITAKU, S. 2008. SOSUI-GramN: high performance prediction for sub-cellular localization of proteins in Gram-negative bacteria. *Bioinformation,* 2**,** 417.

IOANNOU, S., KARADIMA, D., PNEUMATICOS, S., ATHANASIOU, H., PONTIKIS, J., ZORMPALA, A. & SIPSAS, N. 2011. Efficacy of prolonged antimicrobial chemotherapy for *Brucella*r spondylodiscitis. *Clinical Microbiology and Infection,* 17**,** 756-762.

ISSA, M. N. & ASHHAB, Y. 2016. Identification of *Brucella melitensis* Rev. 1 vaccine-strain genetic markers: Towards understanding the molecular mechanism behind virulence attenuation. *Vaccine,* 34**,** 4884-4891.

JAIN, L., RAWAT, M., RAMAKRISHNAN, S. & KUMAR, B. 2017. Active immunization with *Brucella abortus* S19 phage lysate elicits serum IgG that protects guinea pigs against virulent B. *abortus* and protects mice by passive immunization. *Biologicals,* 45**,** 27-32.

JENNER, D. C., DASSA, E., WHATMORE, A. M. & ATKINS, H. S. 2010. ATP-binding cassette systems of *Brucella*. *Comparative and Functional Genomics,* 2009.

KANKAINEN, M., OJALA, T. & HOLM, L. 2012. BLANNOTATOR: enhanced homology-based function prediction of bacterial proteins. *BMC bioinformatics,* 13**,** 33.

KING, B. R. & GUDA, C. 2007. ngLOC: an n-gram-based Bayesian method for estimating the subcellular proteomes of eukaryotes. *Genome biology,* 8**,** R68.

KOH, S. & MORLEY, F. 1981. The effect of calfhood vaccination with strain 19 on the serological diagnosis and eradication of bovine brucellosis. *Australian veterinary journal,* 57**,** 551-553.

KONOVALOVA, A., KAHNE, D. E. & SILHAVY, T. J. 2017. Outer membrane biogenesis. *Annual review of microbiology,* 71**,** 539-556.

KOSKINEN, P., TÖRÖNEN, P., NOKSO-KOIVISTO, J. & HOLM, L. 2015. PANNZER: high-throughput functional annotation of uncharacterized proteins in an error-prone environment. *Bioinformatics,* 31**,** 1544-1552.

LIANG, L., LENG, D., BURK, C., NAKAJIMA-SASAKI, R., KAYALA, M. A., ATLURI, V. L., PABLO, J., UNAL, B., FICHT, T. A. & GOTUZZO, E. 2010. Large scale immune profiling of infected humans and goats reveals differential recognition of *Brucella melitensis* antigens. *PLoS neglected tropical diseases,* 4**,** e673.

LIAO, H., LIU, M. & CHENG, A.-C. 2015. Structural features and functional mechanism of TonB in some Gram-negative bacteria-a review. *Wei sheng wu xue bao= Acta microbiologica Sinica,* 55**,** 529-536.

LIU, W. & CHEN, Y. H. 2005. High epitope density in a single protein molecule significantly enhances antigenicity as well as immunogenicity: a novel strategy for modern vaccine development and a preliminary investigation about B cell discrimination of monomeric proteins. *European journal of immunology,* 35**,** 505-514.

LIU, Y., HARRISON, P. M., KUNIN, V. & GERSTEIN, M. 2004. Comprehensive analysis of pseudogenes in prokaryotes: widespread gene decay and failure of putative horizontally transferred genes. *Genome biology,* 5**,** R64.

MAGNUS, M., PAWLOWSKI, M. & BUJNICKI, J. M. 2012. MetaLocGramN: A meta-predictor of protein subcellular localization for Gram-negative bacteria. *Biochimica ET Biophysica Acta (BBA)-Proteins and Proteomics,* 1824**,** 1425-1433.

MANCINI, J., WECKSELBLATT, B., CHUNG, Y. K., DURANTE, J. C., ANDELMAN, S., GLAUBMAN, J., DORFF, J. D., BHARGAVA, S., LIJEK, R. S. & UNGER, K. P. 2011. The heat-resistant agglutinin family includes a novel adhesin from enteroaggregative Escherichia coli strain 60A. *Journal of bacteriology,* 193**,** 4813-4820.

MARÍN, C., MORENO, E., MORIYON, I., DIAZ, R. & BLASCO, J. 1999. Performance of competitive and indirect enzyme-linked immunosorbent assays, gel immunoprecipitation with native hapten polysaccharide, and standard serological tests in diagnosis of sheep brucellosis. *Clinical and diagnostic laboratory immunology,* 6**,** 269-272.

MARTÍN-MARTÍN, A. I., CARO-HERNÁNDEZ, P., SANCHO, P., TEJEDOR, C., CLOECKAERT, A., FERNÁNDEZ-LAGO, L. & VIZCAÍNO, N. 2009. Analysis of the occurrence and distribution of the Omp25/Omp31 family of surface proteins in the six classical *Brucella* species. *Veterinary microbiology,* 137**,** 74-82.

MCCLEAN, S. 2012. Eight stranded β-barrel and related outer membrane proteins: role in bacterial pathogenesis. *Protein and peptide letters,* 19**,** 1013-1025.

MCDERMOTT, J., GRACE, D. & ZINSSTAG, J. 2013. Economics of brucellosis impact and control in low-income countries. *Rev Sci Tech,* 32**,** 249-61.

MCDIARMID, A. 1957. degree and duration of immunity in cattle resulting from vaccination with s. 19 *abortus* vaccine and its implication in the future control and eventual eradication of brucellosis. *Veterinary record*.

MOHANTY, S. K., MOHANTY, S. K. & LEELA, K. S. 2013. *Textbook of immunology*, JP Medical Ltd.

MORENO, E. 2014. Retrospective and prospective perspectives on zoonotic brucellosis. *Frontiers in microbiology,* 5**,** 213.

MORIEL, D. G., TAN, L., GOH, K. G., PHAN, M.-D., IPE, D. S., LO, A. W., PETERS, K. M., ULETT, G. C., BEATSON, S. A. & SCHEMBRI, M. A. 2016. A novel protective vaccine antigen from the core Escherichia coli genome. *mSphere,* 1**,** e00326-16.

MUSALLAM, I., ABO-SHEHADA, M., HEGAZY, Y., HOLT, H. & GUITIAN, F. 2016. Systematic review of brucellosis in the Middle East: disease frequency in ruminants and humans and risk factors for human infection. *Epidemiology & Infection,* 144**,** 671-685.

NAIR, N., VINOD, V., SURESH, M. K., VIJAYRAJRATNAM, S., BISWAS, L., PEETHAMBARAN, R., VASUDEVAN, A. K. & BISWAS, R. 2015. Amidase, a cell wall hydrolase, elicits protective immunity against Staphylococcus aureus and S. epidermidis. *International journal of biological macromolecules,* 77**,** 314-321.

NICOLETTI, P. 1981. Prevalence and persistence of *Brucella abortus* strain 19 infections and prevalence of other biotypes in vaccinated adult dairy cattle. *Journal of the American Veterinary Medical Association,* 178**,** 143-145.

O'CALLAGHAN, D. & WHATMORE, A. M. 2011. *Brucella* genomics as we enter the multi-genome era. *Briefings in functional genomics,* 10**,** 334-341.

OLIVEIRA, S. C. & SPLITTER, G. A. 1996. Immunization of mice with recombinant L7L12 ribosomal protein confers protection against *Brucella abortus* infection. *Vaccine,* 14**,** 959-962.

OLSEN, S. & PALMER, M. 2014. Advancement of knowledge of *Brucella* over the past 50 years. *Veterinary pathology,* 51**,** 1076-1089.

OLSEN, S. C. & STOFFREGEN, W. 2005. Essential role of vaccines in brucellosis control and eradication programs for livestock. *Expert review of vaccines,* 4**,** 915-928.

ONG, E., WONG, M. U. & HE, Y. 2017. identification of new Features from Known Bacterial Protective Vaccine antigens enhances rational Vaccine Design. *Frontiers in immunology,* 8.

PAL, M., GIZAW, F., FEKADU, G., ALEMAYEHU, G. & KANDI, V. 2017. Public Health and Economic Importance of Bovine Brucellosis: An Overview. *American Journal of Epidemiology,* 5**,** 27-34.

PAPPAS, G., PAPADIMITRIOU, P., AKRITIDIS, N., CHRISTOU, L. & TSIANOS, E. V. 2006. The new global map of human brucellosis. *The Lancet infectious diseases,* 6**,** 91-99.

PARAMASIVAM, N. & LINKE, D. 2011. ClubSub-P: cluster-based subcellular localization prediction for Gram-negative bacteria and archaea. *Frontiers in microbiology,* 2.

PASQUEVICH, K. A., ESTEIN, S. M., SAMARTINO, C. G., ZWERDLING, A., CORIA, L. M., BARRIONUEVO, P., FOSSATI, C. A., GIAMBARTOLOMEI, G. H. & CASSATARO, J. 2009. Immunization with recombinant *Brucella* species outer membrane protein Omp16 or Omp19 in adjuvant induces specific CD4+ and CD8+ T cells as well as systemic and oral protection against *Brucella abortus* infection. *Infection and immunity,* 77**,** 436-445.

PAULLEY, J. T., ANDERSON, E. S. & ROOP, R. M. 2007. *Brucella abortus* requires the heme transporter BhuA for maintenance of chronic infection in BALB/c mice. *Infection and immunity,* 75**,** 5248-5254.

PÉREZ-SANCHO, M., GARCÍA-SECO, T., DOMÍNGUEZ, L. & ÁLVAREZ, J. 2015. Control of Animal Brucellosis—The Most Effective Tool to Prevent Human Brucellosis. *Updates on Brucellosis.* InTech.

PIZZA, M., SCARLATO, V., MASIGNANI, V., GIULIANI, M. M., ARICÒ, B., COMANDUCCI, M., JENNINGS, G. T., BALDI, L., BARTOLINI, E. & CAPECCHI, B. 2000. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science,* 287**,** 1816-1820.

RAPPUOLI, R. 2000. Reverse vaccinology. *Current opinion in microbiology,* 3**,** 445-450.

RIBET, D. & COSSART, P. 2015. How bacterial pathogens colonize their hosts and invade deeper tissues. *Microbes and Infection,* 17**,** 173-183.

RIQUELME-NEIRA, R., RETAMAL-DÍAZ, A., ACUÑA, F., RIQUELME, P., RIVERA, A., SÁEZ, D. & OÑATE, A. 2013. Protective effect of a DNA vaccine containing an open reading frame with homology to an ABC-type transporter present in the genomic island 3 of *Brucella abortus* in BALB/c mice. *Vaccine,* 31**,** 3663-3667.

ROBINSON, R. K. 2014. *Encyclopedia of food microbiology*, Academic press.

ROSA, L. T., BIANCONI, M. E., THOMAS, G. H. & KELLY, D. J. 2018. Tripartite ATP-independent periplasmic (TRAP) transporters and tripartite tricarboxylate transporters (TTT): from uptake to pathogenicity. *Frontiers in cellular and infection microbiology,* 8**,** 33.

ROSET, M., ALEFANTIS, T., DELVECCHIO, V. & BRIONES, G. 2017. Iron-dependent reconfiguration of the proteome underlies the intracellular lifestyle of *Brucella abortus*. *Scientific reports,* 7**,** 10637.

ROSINHA, G., FREITAS, D. A., MIYOSHI, A., AZEVEDO, V., CAMPOS, E., CRAVERO, S. L., ROSSETTI, O., SPLITTER, G. & OLIVEIRA, S. 2002. Identification and characterization of a *Brucella abortus* ATP-binding cassette transporter homolog to Rhizobium meliloti ExsA and its role in virulence and protection in mice. *Infection and immunity,* 70**,** 5036-5044.

SÁEZ, D., GUZMÁN, I., ANDREWS, E., CABRERA, A. & ONATE, A. 2008. Evaluation of *Brucella abortus* DNA and RNA vaccines expressing Cu–Zn superoxide dismutase (SOD) gene in cattle. *Veterinary microbiology,* 129**,** 396-403.

SANDERS, A. N. & PAVELKA, M. S. 2013. Phenotypic analysis of Eschericia coli mutants lacking L, D-transpeptidases. *Microbiology,* 159**,** 1842-1852.

SANKARASUBRAMANIAN, J., VISHNU, U. S., SRIDHAR, J., GUNASEKARAN, P. & RAJENDHRAN, J. 2015. Pan-Genome of *Brucella* Species. *Indian journal of microbiology,* 55**,** 88-101.

SCHURIG, G. G., ROOP II, R. M., BAGCHI, T., BOYLE, S., BUHRMAN, D. & SRIRANGANATHAN, N. 1991. Biological properties of RB51; a stable rough strain of *Brucella abortus*. *Veterinary microbiology,* 28**,** 171-188.

SEIB, K. L., ZHAO, X. & RAPPUOLI, R. 2012. Developing vaccines in the era of genomics: a decade of reverse vaccinology. *Clinical Microbiology and Infection,* 18**,** 109-116.

SELEEM, M. N., BOYLE, S. M. & SRIRANGANATHAN, N. 2010. Brucellosis: a re-emerging zoonosis. *Veterinary microbiology,* 140**,** 392-398.

SETTE, A., VITIELLO, A., REHERMAN, B., FOWLER, P., NAYERSINA, R., KAST, W. M., MELIEF, C., OSEROFF, C., YUAN, L. & RUPPERT, J. 1994. The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes. *The Journal of Immunology,* 153**,** 5586-5592.

SIKORA, A. E., WIERZBICKI, I. H., ZIELKE, R. A., RYNER, R. F., KOROTKOV, K. V., BUCHANAN, S. K. & NOINAJ, N. 2018. Structural and functional insights into the role of BamD and BamE within the β-barrel assembly machinery in Neisseria gonorrhoeae. *Journal of Biological Chemistry,* 293**,** 1106-1119.

SIMBORIO, H., REYES, A., HOP, H. T., ARAYAN, L. T., MIN, W., LEE, H. J., LEE, J. J., CHANG, H. H. & KIM, S. 2016. Immune Modulation of Recombinant OmpA against *Brucella abortus* 544 Infection in Mice. *Journal of microbiology and biotechnology,* 26**,** 603-609.

SINGH, H. & RAGHAVA, G. 2001. ProPred: prediction of HLA-DR binding sites. *Bioinformatics,* 17**,** 1236-1237.

SINGH, H. & RAGHAVA, G. 2003. ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics,* 19**,** 1009-1014.

SMITH, E. P., MILLER, C. N., CHILD, R., CUNDIFF, J. A. & CELLI, J. 2016. Postreplication roles of the *Brucella* VirB type IV secretion system uncovered via conditional expression of the VirB11 ATPase. *MBio,* 7**,** e01730-16.

SMITH, M. E. & BHIMJI, S. S. 2017. Brucellosis.

SMITS, H. 2013. Brucellosis in pastoral and confined livestock: prevention and vaccination. *Rev Sci Tech,* 32**,** 219-228.

SPERANDEO, P., MARTORANA, A. M. & POLISSI, A. 2017. The lipopolysaccharide transport (Lpt) machinery: A nonconventional transporter for lipopolysaccharide assembly at the outer membrane of Gram-negative bacteria. *Journal of Biological Chemistry,* 292**,** 17981-17990.

TADEPALLI, G., SINGH, A. K., BALAKRISHNA, K., MURALI, H. S. & BATRA, H. V. 2016. Immunogenicity and protective efficacy of *Brucella abortus* recombinant protein cocktail (rOmp19+ rP39) against B. *abortus* 544 and B. *melitensis* 16M infection in murine model. *Molecular immunology,* 71**,** 34-41.

TETTELIN, H., MASIGNANI, V., CIESLEWICZ, M. J., DONATI, C., MEDINI, D., WARD, N. L., ANGIUOLI, S. V., CRABTREE, J., JONES, A. L. & DURKIN, A. S. 2005. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial “pan-genome”. *Proceedings of the National Academy of Sciences,* 102**,** 13950-13955.

TIBOR, A., WANSARD, V., BIELARTZ, V., DELRUE, R.-M., DANESE, I., MICHEL, P., WALRAVENS, K., GODFROID, J. & LETESSON, J.-J. 2002. Effect of omp10 or omp19 deletion on *Brucella abortus* outer membrane properties and virulence in mice. *Infection and immunity,* 70**,** 5540-5546.

VELIKOVSKY, C. A., GOLDBAUM, F. A., CASSATARO, J., ESTEIN, S., BOWDEN, R. A., BRUNO, L., FOSSATI, C. A. & GIAMBARTOLOMEI, G. H. 2003. *Brucella* lumazine synthase elicits a mixed Th1-Th2 immune response and reduces infection in mice challenged with *Brucella abortus* 544 independently of the adjuvant formulation used. *Infection and immunity,* 71**,** 5750-5755.

VEMULAPALLI, T. H., VEMULAPALLI, R., SCHURIG, G. G., BOYLE, S. M. & SRIRANGANATHAN, N. 2006. Role in virulence of a *Brucella abortus* protein exhibiting lectin-like activity. *Infection and immunity,* 74**,** 183-191.

VERGER, J.-M., GRAYON, M., ZUNDEL, E., LECHOPIER, P. & OLIVIER-BERNARDIN, V. 1995. Comparison of the efficacy of *Brucella* suis strain 2 and *Brucella melitensis* Rev. 1 live vaccines against a *Brucella melitensis* experimental infection in pregnant ewes. *Vaccine,* 13**,** 191-196.

VISHNU, U. S., SANKARASUBRAMANIAN, J., GUNASEKARAN, P. & RAJENDHRAN, J. 2015. Novel vaccine candidates against *Brucella melitensis* identified through reverse vaccinology approach. *Omics: a journal of integrative biology,* 19**,** 722-729.

VITRY, M.-A., MAMBRES, D. H., DE TREZ, C., AKIRA, S., RYFFEL, B., LETESSON, J.-J. & MURAILLE, E. 2014. Humoral immunity and CD4+ Th1 cells are both necessary for a fully protective immune response upon secondary infection with *Brucella melitensis*. *The Journal of Immunology,* 192**,** 3740-3752.

WATTAM, A. R., WILLIAMS, K. P., SNYDER, E. E., ALMEIDA, N. F., SHUKLA, M., DICKERMAN, A., CRASTA, O., KENYON, R., LU, J. & SHALLOM, J. 2009. Analysis of ten *Brucella* genomes reveals evidence for horizontal gene transfer despite a preferred intracellular lifestyle. *Journal of bacteriology,* 191**,** 3569-3579.

WU, X.-B., TIAN, L.-H., ZOU, H.-J., WANG, C.-Y., YU, Z.-Q., TANG, C.-H., ZHAO, F.-K. & PAN, J.-Y. 2013. Outer membrane protein OmpW of Escherichia coli is required for resistance to phagocytosis. *Research in microbiology,* 164**,** 848-855.

XIN, X. 1986. Orally administrable brucellosis vaccine: *Brucella* suis strain 2 vaccine. *Vaccine,* 4**,** 212-216.

YANG, X., HUDSON, M., WALTERS, N., BARGATZE, R. F. & PASCUAL, D. W. 2005. Selection of protective epitopes for *Brucella melitensis* by DNA vaccination. *Infection and immunity,* 73**,** 7297-7303.

YANG, X., LI, Y., ZANG, J., LI, Y., BIE, P., LU, Y. & WU, Q. 2016. Analysis of pan-genome to identify the core genes and essential genes of *Brucella* spp. *Molecular genetics and genomics,* 291**,** 905-912.

YANG, Y., WANG, L., YIN, J., WANG, X., CHENG, S., LANG, X., WANG, X., QU, H., SUN, C. & WANG, J. 2011. Immunoproteomic analysis of *Brucella melitensis* and identification of a new immunogenic candidate protein for the development of brucellosis subunit vaccine. *Molecular immunology,* 49**,** 175-184.

YU, C. S., CHEN, Y. C., LU, C. H. & HWANG, J. K. 2006. Prediction of protein subcellular localization. *Proteins: Structure, Function, and Bioinformatics,* 64**,** 643-651.

YU, N. Y., WAGNER, J. R., LAIRD, M. R., MELLI, G., REY, S., LO, R., DAO, P., SAHINALP, S. C., ESTER, M. & FOSTER, L. J. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics,* 26**,** 1608-1615.

ZENG, L., WANG, D., HU, N., ZHU, Q., CHEN, K., DONG, K., ZHANG, Y., YAO, Y., GUO, X. & CHANG, Y.-F. 2017. A novel pan-genome reverse vaccinology approach employing a negative-selection strategy for screening surface-exposed antigens against leptospirosis. *Frontiers in microbiology,* 8**,** 396.

Appendix

Appendix1: (Table of Proteome file selected for the study, strain name, their accession number, and number of proteins).

Organism/Name	Strain	Replicons	Proteins
***Brucella melitensis* bv. 1 str. 16M**	16M	chromosome I:NC_003317.1/AE008917.1; chromosome II:NC_003318.1/AE008918.1	3099
***Brucella melitensis* ATCC 23457**	ATCC 23457	chromosome I:NC_012441.1/CP001488.1; chromosome II:NC_012442.1/CP001489.1	3152
***Brucella melitensis* M28**	M28	chromosome 1:NC_017244.1/CP002459.1; chromosome 2:NC_017245.1/CP002460.1	3144
***Brucella melitensis* M5-90**	M5-90	chromosome I:NC_017246.1/CP001851.1; chromosome II:NC_017247.1/CP001852.1	3124
***Brucella melitensis* NI**	NI	chromosome I:NC_017248.1/CP002931.1; chromosome II:NC_017283.1/CP002932.1	3115
***Brucella melitensis* bv. 3 str. Ether**	ether	chromosome 1:NZ_CP007760.1/CP007760.1; chromosome 2:NZ_CP007761.1/CP007761.1	3130
***Brucella melitensis* bv. 1 str. 16M**	16M	chromosome 1:NZ_CP007763.1/CP007763.1; chromosome 2:NZ_CP007762.1/CP007762.1	3134
Brucella melitensis	C-573	chromosome I:NZ_CP019679.1/CP019679.1; chromosome II:NZ_CP019680.1/CP019680.1	3099
Brucella melitensis	20236	chromosome 1:NZ_CP008750.1/CP008750.1; chromosome 2:NZ_CP008751.1/CP008751.1	3140
Brucella melitensis	2008724259	chromosome 1:NZ_CP016983.1/CP016983.1; chromosome 2:NZ_CP016984.1/CP016984.1	3125
Brucella melitensis	BwIM_AFG_63	chromosome 1:NZ_CP018478.1/CP018478.1; chromosome 2:NZ_CP018479.1/CP018479.1	3148
Brucella melitensis	BwIM_IRN_37	chromosome 1:NZ_CP018486.1/CP018486.1; chromosome 2:NZ_CP018487.1/CP018487.1	3149
Brucella melitensis	BwIM_IRQ_32	chromosome 1:NZ_CP018490.1/CP018490.1; chromosome 2:NZ_CP018491.1/CP018491.1	3145
Brucella melitensis	BwIM_ITA_45	chromosome 1:NZ_CP018494.1/CP018494.1; chromosome 2:NZ_CP018495.1/CP018495.1	3131
Brucella melitensis	BwIM_ITA_55	chromosome 1:NZ_CP018496.1/CP018496.1; chromosome 2:NZ_CP018497.1/CP018497.1	3126
Brucella melitensis	BwIM_SYR_04	chromosome 1:NZ_CP018512.1/CP018512.1; chromosome 2:NZ_CP018513.1/CP018513.1	3149
Brucella melitensis	BwIM_TKM_56	chromosome 1:NZ_CP018536.1/CP018536.1; chromosome 2:NZ_CP018537.1/CP018537.1	3148
Brucella melitensis	BwIM_TUR_03	chromosome 1:NZ_CP018540.1/CP018540.1; chromosome 2:NZ_CP018541.1/CP018541.1	3145
Brucella melitensis	BwIM_TUR_17	chromosome 1:NZ_CP018544.1/CP018544.1; chromosome 2:NZ_CP018545.1/CP018545.1	3146
Brucella melitensis	BwIM_TUR_19	chromosome 1:NZ_CP018546.1/CP018546.1; chromosome 2:NZ_CP018547.1/CP018547.1	3149
Brucella melitensis	BwIM_TUR_59	chromosome 1:NZ_CP018560.1/CP018560.1; chromosome 2:NZ_CP018561.1/CP018561.1	3147
Brucella melitensis	BwIM_SAU_09	chromosome 1:NZ_CP018504.1/CP018504.1; chromosome 2:NZ_CP018505.1/CP018505.1	3146
Brucella melitensis	BwIM_SYR_26	chromosome 1:NZ_CP018526.1/CP018526.1; chromosome 2:NZ_CP018527.1/CP018527.1	3145
Brucella melitensis	BwIM_TUR_39	chromosome 1:NZ_CP018554.1/CP018554.1; chromosome 2:NZ_CP018555.1/CP018555.1	3144
Brucella melitensis	QY1	chromosome I:NZ_CP022204.1/CP022204.1; chromosome II:NZ_CP022205.1/CP022205.1	3058

Brucella melitensis	BY38	chromosome 1:NZ_CP022827.1/CP022827.1; chromosome 2:NZ_CP022828.1/CP022828.1	3141
Brucella melitensis	BL	chromosome 1:NZ_CP022875.1/CP022875.1; chromosome 2:NZ_CP022876.1/CP022876.1	3146
Brucella melitensis	QH61	chromosome 1:NZ_CP024653.1/CP024653.1; chromosome 2:NZ_CP024654.1/CP024654.1	3131
Brucella melitensis	CIIMS-BH-2	chromosome I:NZ_CP025680.1/CP025680.1; chromosome II:NZ_CP025681.1/CP025681.1	3130
Brucella melitensis	B3	chromosome 1:NZ_CP026005.1/CP026005.1; chromosome 2:NZ_CP026006.1/CP026006.1	3117
Brucella melitensis	Rev.1 (passage 101)	chromosome I:NZ_CP024715.1/CP024715.1; chromosome II:NZ_CP024716.1/CP024716.1	3111
Brucella melitensis	1	chromosome 1:NZ_LT962928.1/LT962928.1; chromosome 2:NZ_LT962929.1/LT962929.1	3120
Brucella melitensis	1	chromosome 1:NZ_LT963350.1/LT963350.1; chromosome 2:NZ_LT963351.1/LT963351.1	3112
Brucella melitensis	1	chromosome 1:NZ_LT962945.1/LT962945.1; chromosome 2:NZ_LT962946.1/LT962946.1	3120
Brucella melitensis	1	chromosome 1:NZ_LT962940.1/LT962940.1; chromosome 2:NZ_LT962941.1/LT962941.1	3117
Brucella melitensis	1	chromosome 1:NZ_LT962947.1/LT962947.1; chromosome 2:NZ_LT962948.1/LT962948.1	3104
Brucella melitensis	1	chromosome 1:NZ_LT962910.1/LT962910.1; chromosome 2:NZ_LT962911.1/LT962911.1	3124
Brucella melitensis	1	chromosome 1:NZ_LT962943.1/LT962943.1; chromosome 2:NZ_LT962944.1/LT962944.1	3119
Brucella melitensis	1	chromosome 1:NZ_LT962914.1/LT962914.1; chromosome 2:NZ_LT962915.1/LT962915.1	3123
Brucella melitensis	1	chromosome 1:NZ_LT962926.1/LT962926.1; chromosome 2:NZ_LT962927.1/LT962927.1	3122
Brucella melitensis	1	chromosome 1:NZ_LT962922.1/LT962922.1; chromosome 2:NZ_LT962923.1/LT962923.1	3103
Brucella melitensis	1	chromosome 1:NZ_LT962916.1/LT962916.1; chromosome 2:NZ_LT962917.1/LT962917.1	3123
Brucella melitensis	1	chromosome 1:NZ_LT962924.1/LT962924.1; chromosome 2:NZ_LT962925.1/LT962925.1	3121
Brucella melitensis	1	chromosome 1:NZ_LT962912.1/LT962912.1; chromosome 2:NZ_LT962913.1/LT962913.1	3125
Brucella melitensis	1	chromosome 1:NZ_LT962920.1/LT962920.1; chromosome 2:NZ_LT962921.1/LT962921.1	3116
Brucella melitensis	1	chromosome 1:NZ_LT962918.1/LT962918.1; chromosome 2:NZ_LT962919.1/LT962919.1	3106
Brucella melitensis	1	chromosome 1:NZ_LT963348.1/LT963348.1; chromosome 2:NZ_LT963349.1/LT963349.1	3108
Brucella melitensis	1	chromosome 1:NZ_LT962930.1/LT962930.1; chromosome 2:NZ_LT962931.1/LT962931.1	3115
Brucella melitensis	1	chromosome 1:NZ_LT962932.1/LT962932.1; chromosome 2:NZ_LT962933.1/LT962933.1	3120
Brucella melitensis	1	chromosome 1:NZ_LT962938.1/LT962938.1; chromosome 2:NZ_LT962939.1/LT962939.1	3103
Brucella melitensis	1	chromosome 1:NZ_LT962934.1/LT962934.1; chromosome 2:NZ_LT962935.1/LT962935.1	3126
Brucella melitensis	1	chromosome 1:NZ_LT962936.1/LT962936.1; chromosome 2:NZ_LT962937.1/LT962937.1	3117
Brucella melitensis	1	chromosome 1:NZ_LT962953.1/LT962953.1; chromosome 2:NZ_LT962954.1/LT962954.1	3121
Brucella melitensis	1	chromosome 1:NZ_LT962951.1/LT962951.1; chromosome 2:NZ_LT962952.1/LT962952.1	3109

Brucella melitensis	1	chromosome 1:NZ_LT962949.1/LT962949.1; chromosome 2:NZ_LT962950.1/LT962950.1	3100
***Brucella abortus* bv. 1 str. 9-941**	9-941	chromosome I:NC_006932.1/AE017223.1; chromosome II:NC_006933.1/AE017224.1	3153
***Brucella abortus* S19**	S19	chromosome 1:NC_010742.1/CP000887.1; chromosome 2:NC_010740.1/CP000888.1	3151
***Brucella abortus* 2308**	2308	chromosome I:NC_007618.1/AM040264.1; chromosome II:NC_007624.1/AM040265.1	3153
***Brucella abortus* A13334**	A13334	chromosome 1:NC_016795.1/CP003176.1; chromosome 2:NC_016777.1/CP003177.1	3168
Brucella abortus	BDW	chromosome 1:NZ_CP007681.1/CP007681.1; chromosome 2:NZ_CP007680.1/CP007680.1	3171
Brucella abortus	BER	chromosome 1:NZ_CP007682.1/CP007682.1; chromosome 2:NZ_CP007683.1/CP007683.1	3130
Brucella abortus	NCTC 10505	chromosome 1:NZ_CP007700.1/CP007700.1; chromosome 2:NZ_CP007701.1/CP007701.1	3140
***Brucella abortus* bv. 9 str. C68**	C68	chromosome 1:NZ_CP007705.1/CP007705.1; chromosome 2:NZ_CP007706.1/CP007706.1	3150
***Brucella abortus* bv. 6 str. 870**	870	chromosome 1:NZ_CP007709.1/CP007709.1; chromosome 2:NZ_CP007710.1/CP007710.1	3142
Brucella abortus	63 75	chromosome 1:NZ_CP007663.1/CP007663.1; chromosome 2:NZ_CP007662.1/CP007662.1	3147
Brucella abortus	BFY	chromosome 1:NZ_CP007738.1/CP007738.1; chromosome 2:NZ_CP007737.1/CP007737.1	3127
***Brucella abortus* bv. 2 str. 86/8/59**	86/8/59	chromosome 1:NZ_CP007765.1/CP007765.1; chromosome 2:NZ_CP007764.1/CP007764.1	3159
Brucella abortus	BAB8416	chromosome 1:NZ_CP008774.1/CP008774.1; chromosome 2:NZ_CP008775.1/CP008775.1	3121
***Brucella abortus* 104M**	104M	chromosome 1:NZ_CP009625.1/CP009625.1; chromosome 2:NZ_CP009626.1/CP009626.1	3152
Brucella abortus	BD	chromosome 1:NZ_CP022877.1/CP022877.1; chromosome 2:NZ_CP022878.1/CP022878.1	3115
Brucella abortus	MC	chromosome 1:NZ_CP022879.1/CP022879.1; chromosome 2:NZ_CP022880.1/CP022880.1	3109
Brucella abortus	B4	chromosome I:NZ_CP025743.1/CP025743.1; chromosome II:NZ_CP025744.1/CP025744.1	3123
***Brucella* suis 1330**	1330	chromosome I:NC_004310.3/AE014291.4; chromosome II:NC_004311.2/AE014292.2	3161
***Brucella* suis ATCC 23445**	ATCC 23445	chromosome I:NC_010169.1/CP000911.1; chromosome II:NC_010167.1/CP000912.1	3158
***Brucella* suis 1330**	1330	chromosome I:NC_017251.1/CP002997.1; chromosome II:NC_017250.1/CP002998.1	3164
***Brucella* suis VBI22**	VBI22	chromosome I:NC_016797.1/CP003128.1; chromosome II:NC_016775.1/CP003129.1	3167
***Brucella* suis bv. 1 str. S2**	S2	chromosome I:NZ_CP006961.1/CP006961.1; chromosome II:NZ_CP006962.1/CP006962.1	3167
***Brucella* suis bv. 2**	PT09143	chromosome I:NZ_CP007691.1/CP007691.1; chromosome II:NZ_CP007692.1/CP007692.1	3162
***Brucella* suis bv. 2**	PT09172	chromosome I:NZ_CP007693.1/CP007693.1; chromosome II:NZ_CP007694.1/CP007694.1	3162
***Brucella* suis bv. 2**	Bs364CITA	chromosome I:NZ_CP007697.1/CP007697.1; chromosome II:NZ_CP007698.1/CP007698.1	3163
***Brucella* suis bv. 2**	Bs396CITA	chromosome I:NZ_CP007720.1/CP007720.1; chromosome II:NZ_CP007721.1/CP007721.1	3164
***Brucella* suis bv. 2**	Bs143CITA	chromosome I:NZ_CP007695.1/CP007695.1; chromosome II:NZ_CP007696.1/CP007696.1	3162
***Brucella* suis**	513UK	chromosome 1:NZ_CP007717.1/CP007717.1; chromosome 2:NZ_CP007716.1/CP007716.1	3165

***Brucella* suis bv. 3 str. 686**	686	chromosome 1:NZ_CP007719.1/CP007719.1; chromosome 2:NZ_CP007718.1/CP007718.1	3121
***Brucella* suis**	BSP	chromosome 1:NZ_CP008757.1/CP008757.1; chromosome 2:NZ_CP008756.1/CP008756.1	3153
***Brucella* suis**	Human/AR/US/1981	chromosome I:NZ_CP010850.1/CP010850.1; chromosome II:NZ_CP010851.1/CP010851.1	3164
***Brucella* suis**	2004000577	chromosome 1:NZ_CP016981.1/CP016981.1; chromosome 2:NZ_CP016982.1/CP016982.1	3154
***Brucella* suis**	QH05	chromosome 1:NZ_CP024420.1/CP024420.1; chromosome 2:NZ_CP024421.1/CP024421.1	3151
***Brucella* suis**	ZW046	chromosome 1:CP009096.1; chromosome 2:CP009097.1	3124
***Brucella* suis**	ZW043	chromosome 1:CP009094.1; chromosome 2:CP009095.1	3090

Appendix 2: (Table contains the 177 proteins resulted from SCL).

Protein Info			**SCL-Pipeline**					
NCBI ID	**Length**	**Description**	**Psortb**	**SosuiGramN**	**Cello**	**ngLoc**	**ClubSub-P**	**MetaLoc**
WP_002966516.1	57	type IV secretion system lipoprotein virB7	unKnown	unKnown	**P**	Cyt/IM	unKnown	**OM** 42.92%.
WP_002971893.1	60	hypothetical protein	unKnown	**OM**	Cyt	Cyt/IM	**OM**	
WP_002970216.1	62	hypothetical protein	unKnown	**P**	**P**			
WP_006135036.1	71	DNA gyrase inhibitor YacG	unKnown	**P**	**P**			
WP_006189877.1	78	hypothetical protein	unKnown	**P**	**P**			
WP_002963878.1	82	hypothetical protein	unKnown	**P**	**P**			
WP_002964124.1	90	hypothetical protein	unKnown	**P**	**P**			
WP_002968043.1	93	hypothetical protein	IM	**P**	**P**			
WP_004686827.1	94	pentapeptide MXKDX repeat protein	unKnown	**P**	**P**			
WP_002971526.1	98	hypothetical protein	unKnown	**P**	**P**		-	-
WP_002964730.1	105	hypothetical protein	unKnown	**P**	**P**			
WP_002964559.1	107	DUF2293 domain-containing protein	unKnown	**P**	**P**			
WP_004681873.1	108	hypothetical protein	unKnown	**OM**	**P**	-	-	-
WP_002963717.1	111	hypothetical protein	unKnown	**P**	**P**		-	-
WP_002964582.1	114	hypothetical protein	unKnown	**P**	**P**			
WP_002966248.1	114	acid stress chaperone HdeA	unKnown	**P**	**P**			
WP_002963776.1	115	DUF2147	unKnown	**P**	**P**			
WP_002964396.1	115	hypothetical protein	**P**	**P**	**P**			
WP_002966091.1	115	hypothetical protein	unKnown	**P**	**P**			
WP_002965199.1	118	septal ring lytic transglycosylase RlpA family protein	unKnown	**P**	**P**	-	-	-

WP_002963597.1	121	hypothetical protein	OM	P	P			
WP_006168377.1	123	30S ribosomal protein S12	Cyto	P	P			
WP_002966502.1	126	hypothetical protein	OM	OM	Ex			
WP_004686521.1	127	TIGR02301 family protein	unKnown	P	P		-	-
WP_002965104.1	131	DUF4354 domain-containing protein	unKnown	P	P		-	-
WP_002966353.1	131	hypothetical protein	unKnown	P	P			
WP_002968479.1	131	hypothetical protein	IM	P	P			
WP_002966407.1	131	cytochrome c family protein	P	P	P			
WP_002963970.1	133	hypothetical protein	unKnown	P	P			
WP_004686362.1	134	hypothetical protein	unKnown	P	P		-	-
WP_004690392.1	135	hypothetical protein	unKnown	P	P			
WP_002963960.1	137	hypothetical protein	unKnown	P	P			
WP_004681387.1	138	DUF2195 domain-containing protein	unKnown	P	P			
WP_002966303.1	146	pseudoazurin [*Brucella*]	P	unKnown	P			
WP_002966682.1	151	hypothetical protein	unKnown	P	P	-	-	-
WP_004683280.1	154	hypothetical protein	unKnown	P	P			
WP_002969562.1	155	hypothetical protein	IM	P	P			
WP_002966352.1	156	DUF2271 domain-containing protein	unKnown	OM	P			
WP_002963473.1	160	CAP domain-containing protein	unKnown	P	P			
WP_002972060.1	165	DUF995 domain-containing protein	unKnown	P	P			
WP_002967296.1	166	BA14K family protein	unKnown	P	P			
WP_002966257.1	166	hypothetical protein	unKnown	P	P			
WP_002971481.1	168	outer membrane protein assembly factor BamE	unKnown	P	P			
WP_002966947.1	168	hypothetical protein	OM	OM	OM			
WP_002964622.1	170	BA14K family protein	unKnown	P	P			
WP_004689960.1	172	type IV secretion system protein [*Brucella*] VirB12	unKnown	P	P	-	-	-
WP_002972093.1	173	superoxide dismutase [Cu-Zn]	P	P	P			
WP_002964998.1	177	hypothetical protein	OM	P	Ex			
WP_002971445.1	177	hypothetical protein	unKnown	P	P		-	-
WP_002966306.1	179	hypothetical protein	unKnown	P	P			
WP_002964988.1	180	hypothetical protein	unKnown	P	P			
WP_004689355.1	181	hypothetical protein	unKnown	P	P			
WP_002966226.1	182	hypothetical protein	unKnown	P	P			
WP_006154070.1	184	hypothetical protein	unKnown	OM	P	Cyto	P	
WP_002964648.1	187	ubiquinol-cytochrome c reductase iron-sulfur subunit	IM	P	P			
WP_004685190.1	190	lytic transglycosylase domain-containing protein	unKnown	P	P			

WP_004688070.1	192	hypothetical protein	unKnown	Ex	OM	IM	OM	
WP_002967573.1	193	NADH-quinone oxidoreductase subunit B	IM	P	P			
WP_002964883.1	193	hypothetical protein	unKnown	P	OM	Cyto	OM	
WP_002963919.1	195	polysaccharide export protein	unKnown	P	OM	-	-	-
WP_002964222.1	196	peptidyl-prolyl cis-trans isomerase	P	P	P			
WP_004684011.1	196	SCO family protein	unKnown	P	P			
WP_004690660.1	199	superoxide dismutase	P	Cyto	P			
WP_004684614.1	202	cytochrome c family protein	IM	P	P			
WP_002963732.1	203	DUF1214 domain-containing protein	unKnown	P	P			
WP_002971565.1	203	pyridoxine/pyridoxamine 5'-phosphate oxidase	P	unKnown	P			
WP_002967079.1	203	DsbE family thiol:disulfide interchange protein	P	P	P			
WP_004686596.1	204	DsbA family protein	unKnown	P	P			
WP_002971243.1	205	invasion associated locus B family protein	unKnown	P	P			
WP_004686655.1	206	L,D-transpeptidase	unKnown	P	P			
WP_004683722.1	207	hypothetical protein	P	P	P			
WP_002968302.1	207	lipase	unKnown	P	P			
WP_002963624.1	208	invasion associated locus B family protein	unKnown	P	P			
WP_002964140.1	209	ABC transporter	unKnown	P	OM		-	-
WP_004683944.1	212	porin family protein	OM	OM	Ex			
WP_004683466.1	213	membrane protein	OM	OM	OM	-	-	-
WP_002963780.1	216	hypothetical protein	unKnown	P	P			
WP_004691352.1	216	organic solvent tolerance protein OstA	unKnown	OM	P	Cyto	P	
WP_002971551.1	217	DsbA family protein	unKnown	P	P			
WP_004682501.1	218	peptide-methionine (S)-S-oxide reductase	unKnown	P	P			
WP_002964333.1	220	OmpA family protein	IM	OM	OM			
WP_002971287.1	223	hypothetical protein	unKnown	P	P			
WP_002964666.1	227	OmpW family protein	OM	OM	P			
WP_002967534.1	229	ribonuclease T(2)	P	P	P			
WP_002966482.1	234	DUF1007 domain-containing protein	unKnown	Cyto	P	IM	P	
WP_004683739.1	236	porin family protein	OM	OM	OM			
WP_004681227.1	238	type IV secretion system protein virB1	unKnown	P	P			
WP_002966514.1	238	type IV secretion system protein virB5	unKnown	OM	P	-	-	-
WP_002966517.1	239	type IV secretion system protein virB8	unKnown	OM	OM			
WP_002965297.1	241	L,D-transpeptidase	unKnown	P	P			
WP_002965333.1	241	arylesterase	P	P	P			
WP_002971753.1	244	DUF1223 domain-containing protein	unKnown	Cyto	P	OM	P	

WP_002964581.1	250	SIMPL domain-containing protein	P	OM	P			
WP_004689974.1	251	molybdate ABC transporter substrate-binding protein	P	P	P			
WP_002964476.1	252	DUF4198 domain-containing protein	unKnown	P	P			
WP_002963812.1	254	hypothetical protein	unKnown	OM	P			
WP_004681957.1	257	amino acid ABC transporter substrate-binding protein	P	P	P			
WP_002964719.1	261	porin family protein	OM	OM	OM			
WP_006278325.1	261	hypothetical protein	unKnown	P	P			
WP_004690620.1	266	hypothetical protein	unKnown	P	P			
WP_002971090.1	267	hypothetical protein	unKnown	OM	OM			
WP_004684144.1	274	porin family protein	unKnown	OM	OM			
WP_004682863.1	280	amino acid ABC transporter	P	P	P			
WP_002966849.1	280	DUF1849 domain-containing protein	unKnown	P	P			
WP_004690357.1	284	hypothetical protein	OM	OM	OM			
WP_002964174.1	284	3-mercaptopyruvate sulfurtransferase	P	Cyto	P			
WP_002964530.1	287	outer membrane protein assembly factor BamD	OM	P	Cyto	Cyto	OM	OM 54.64 %.
WP_004686020.1	288	glycine/betaine ABC transporter substrate-binding protein	unKnown	P	P			
WP_004686339.1	295	cytochrome c1 family protein	P	P	P			
WP_004685781.1	303	extensin	unKnown	Cyt	P	unKnown	OM	OM 39.37%.
WP_004681548.1	306	formylglycine-generating enzyme family protein	P	P	P			
WP_002971682.1	309	ABC transporter substrate-binding protein	IM	P	P			
WP_002966164.1	314	hemolysin secretion protein D	IM	OM	P	Cyto	OM	
WP_002969381.1	314	penicillin-insensitive murein endopeptidase	P	P	P			
WP_002965965.1	315	iron ABC transporter substrate-binding protein	P	unKnown	P			
WP_004691921.1	318	molecular chaperone SurA	unKnown	P	P			
WP_002965661.1	319	mononuclear molybdenum enzyme YedY	unKnown	P	P			
WP_002964322.1	329	hypothetical protein	unKnown	P	P			
WP_002967978.1	331	peptidylprolyl isomerase	IM	P	P			
WP_004684204.1	334	thiamine-binding periplasmic protein	P	P	P			
WP_002965301.1	334	DUF1775 domain-containing protein	unKnown	P	P			
WP_006191074.1	340	cell wall hydrolase	unKnown	P	P		-	-
WP_002963881.1	343	amino acid ABC transporter substrate-binding protein	P	Ex	OM			
WP_002965202.1	348	phosphonate ABC transporter substrate-binding protein	IM	P	P			
WP_004684153.1	350	sel1 repeat family protein	unKnown	P	P			
WP_002964611.1	351	DUF1176 domain-containing protein	unKnown	OM	P		-	-
WP_004686075.1	355	BMP family ABC transporter substrate-binding protein	unKnown	P	P			
WP_006132773.1	355	glycoside hydrolase family 43	IM	OM	OM			

WP_004689211.1	356	hypothetical protein	**P**	**P**	**P**			
WP_004691678.1	363	ABC transporter substrate-binding protein	**P**	**P**	Cyto			
WP_004681306.1	367	iron ABC transporter substrate-binding protein	unKnown	**P**	**P**			
WP_002968345.1	368	peptidase S58 family protein	**P**	unKnown	**P**			
WP_002965732.1	368	efflux RND transporter periplasmic adaptor subunit	unKnown	**P**	Cyto	Cyto	**OM**	**OM** 39.53%
WP_004685038.1	372	cystathionine gamma-synthase	unKnown	**P**	**P**			
WP_002966326.1	376	nitrite reductase, copper-containing	**P**	**P**	**P**			
WP_002964109.1	396	D-alanyl-D-alanine carboxypeptidase	IM	**P**	**P**			
WP_004690733.1	398	hypothetical protein	unKnown	**P**	**P**	-	-	-
WP_004691134.1	403	hypothetical protein	unKnown	**P**	**P**			
WP_004688921.1	403	branched-chain amino acid ABC transporter substrate-binding protein	**P**	**P**	**OM**			
WP_004682068.1	411	sugar ABC transporter substrate-binding protein	**P**	**P**	**P**			
WP_004690137.1	415	sugar ABC transporter substrate-binding protein	unKnown	unKnown	**OM**	Cyt	**P**	**OM** 39.06 %
WP_004681076.1	421	ABC transporter substrate-binding protein	**P**	**P**	**P**			
WP_002968069.1	427	lytic murein transglycosylase	IM	**P**	**P**			
WP_004690579.1	429	cell wall hydrolase	**P**	unKnown	**P**			
WP_004686494.1	433	amidase	unKnown	**P**	**P**			
WP_002965482.1	439	sugar ABC transporter substrate-binding protein	unKnown	**P**	**P**			
WP_002964915.1	442	S41 family peptidase	IM	**P**	**P**			
WP_002964785.1	443	protein TolB	**P**	**OM**	**OM**			
WP_002966309.1	464	nitrous oxide reductase family maturation protein NosD	**P**	**P**	**P**	-	-	-
WP_002969408.1	471	M48 family peptidase	unKnown	**P**	**P**			
WP_002964782.1	488	tol-pal system protein YbgF	**OM**	**P**	**P**			
WP_004684782.1	492	ABC transporter substrate-binding protein	**P**	**P**	**P**			
WP_004686294.1	493	metalloprotease	unKnown	**P**	**OM**			
WP_004686222.1	494	GMC family oxidoreductase	IM	**P**	**P**			
WP_002966234.1	499	catalase	**P**	**P**	**P**			
WP_002968232.1	512	hypothetical protein	**OM**	Cyt	**OM**	-	-	-
WP_002963760.1	513	DegP-like serine endoprotease	**P**	**P**	**P**			
WP_004685925.1	518	ABC transporter substrate-binding protein	**P**	unKnown	Cyto	Cyto	**P**	**P** 41.94%
WP_004688535.1	524	DegQ family serine endoprotease	**P**	unKnown	Cyto	**P**	**P**	
WP_002966094.1	525	peptide ABC transporter substrate-binding protein	**P**	Ex	**P**			
WP_004681740.1	527	peptide ABC transporter substrate-binding protein	**P**	**P**	**P**			
WP_002965856.1	537	ABC transporter substrate-binding protein	**P**	Cyto	**P**			
WP_006133510.1	539	ABC transporter substrate-binding protein	**P**	unKnown	**P**			
WP_002963770.1	563	electron transfer flavoprotein-ubiquinone oxidoreductase	unKnown	**P**	**P**			

WP_002963823.1	577	M23 family peptidase	unKnown	**P**	**P**			
WP_004686326.1	586	hypothetical protein	**OM**	**OM**	**OM**			
WP_002965256.1	615	ABC transporter substrate-binding protein	**P**	unKnown	**P**			
WP_004691650.1	620	TonB-dependent receptor	**OM**	**OM**	**OM**			
WP_004684597.1	622	ABC transporter substrate-binding protein	**P**	**OM**	**P**			
WP_002971772.1	639	outer membrane protein assembly factor	**OM**	**OM**	**OM**			
WP_002964936.1	651	M23 family peptidase	Cyto	**OM**	**OM**			
WP_005974736.1	653	DUF882 domain-containing protein	unKnown	**P**	**P**		-	-
WP_023080793.1	661	heme transporter BhuA	**OM**	**OM**	**OM**			
WP_002971180.1	761	penicillin amidase	**P**	**P**	Cyto			
WP_011068938.1	792	LPS-assembly protein LptD	**OM**	**OM**	**OM**			
WP_006133056.1	1579	translocation/assembly module TamB	unKnown	IM	OM	**Cyto**	**unKnown**	**OM** 52.36 %
WP_006139398.1	1582	kinesin	unKnown	**OM**	**OM**			

Appendix 3: (containing the 87 proteins with their adhesion, antigenicity, and immunogenicity results).

ANTIGENpro Protein Antigenicity	VaxiJen Prediction Result	Propred I (MHCI)		Propred II (MHCII)		BCPred			AAP		
		No. of Alleles in query	Density (Sum epitopes/47 (Alleles)/length)	No. of Alleles in query	Density (Sum epitopes/51 (Alleles)/length)	N	length aa	Density:N/length	N	length aa	Density :N/length
0.902746	0.6967 (Probable ANTIGEN).	47/47	0.028467297	44/51	0.008442266	5	216	0.023148148	6	216	0.027777778
0.882738	0.8005 (Probable ANTIGEN).	47/47	0.032051282	47/51	0.012317748	4	156	0.025641026	4	156	0.025641026
0.916867	0.7013 (Probable ANTIGEN).	47/47	0.03064843	49/51	0.016339869	4	168	0.023809524	4	168	0.023809524
0.931412	0.6298 (Probable ANTIGEN).	46/47	0.029210242	48/51	0.009859311	5	177	0.028248588	5	177	0.028248588
0.776825	0.4183 (Probable ANTIGEN).	46/47	0.030931522	51/51	0.022326578	7	238	0.029411765	6	238	0.025210084
0.958819	0.4441 (Probable ANTIGEN).	47/47	0.029888551	51/51	0.019763461	2	126	0.015873016	3	126	0.023809524
0.95578	0.5920 (Probable ANTIGEN).	47/47	0.028615173	51/51	0.016299324	10	403	0.024813896	9	403	0.022332506
0.941443	0.7910 (Probable ANTIGEN).	47/47	0.029104777	50/51	0.011931188	5	212	0.023584906	5	212	0.023584906
0.934726	0.5498 (Probable ANTIGEN).	47/47	0.032336212	51/51	0.016598719	4	202	0.01980198	7	202	0.034653465
0.89195	0.6621 (Probable ANTIGEN).	45/47	0.029047179	45/51	0.013810742	2	115	0.017391304	3	115	0.026086957
0.532994	0.4659 (Probable ANTIGEN).	47/47	0.022814663	51/51	0.020198441	4	166	0.024096386	4	166	0.024096386
0.902471	0.4221 (Probable ANTIGEN).	47/47	0.031306991	51/51	0.022689076	4	280	0.014285714	7	280	0.025
0.896893	0.5546 (Probable ANTIGEN).	47/47	0.030141844	51/51	0.019301471	4	192	0.020833333	6	192	0.03125
0.5	0.7078 (Probable ANTIGEN).	47/47	0.033305309	51/51	0.01829668	5	329	0.015197568	8	329	0.024316109
0.695898	0.6353 (Probable ANTIGEN).	47/47	0.032225501	51/51	0.023185917	9	274	0.032846715	7	274	0.025547445
0.814179	0.6525 (Probable ANTIGEN).	47/47	0.024780976	50/51	0.015340254	3	170	0.017647059	4	170	0.023529412
0.643178	0.6844 (Probable ANTIGEN).	47/47	0.030977419	51/51	0.015250545	4	261	0.01532567	3	261	0.011494253
0.909901	0.7150 (Probable ANTIGEN).	47/47	0.033075435	50/51	0.012745098	4	220	0.018181818	5	220	0.022727273
0.758413	0.5458 (Probable ANTIGEN).	47/47	0.031283808	51/51	0.01811233	4	236	0.016949153	6	236	0.025423729
0.964208	0.5768 (Probable ANTIGEN).	47/47	0.03355516	51/51	0.020039734	3	227	0.013215859	3	227	0.013215859
0.800794	0.4671 (Probable ANTIGEN).	47/47	0.031599864	50/51	0.016838626	9	439	0.020501139	10	439	0.022779043
0.928539	0.4565 (Probable ANTIGEN).	46/47	0.028841608	44/51	0.01372549	1	90	0.011111111	2	90	0.022222222
0.669913	0.7673 (Probable ANTIGEN).	47/47	0.031465388	51/51	0.02029826	9	284	0.031690141	5	284	0.017605634
0.860866	0.9522 (Probable ANTIGEN).	43/47	0.028826356	47/51	0.019924099	1	62	0.016129032	0	62	0
0.240098	0.8212 (Probable ANTIGEN).	47/47	0.031665168	51/51	0.014913007	3	213	0.014084507	6	213	0.028169014
0.671545	0.6612 (Probable ANTIGEN).	47/47	0.032592238	51/51	0.017404715	6	267	0.02247191	6	267	0.02247191
0.793744	0.5366 (Probable ANTIGEN).	47/47	0.034369885	51/51	0.020110608	7	351	0.01994302	8	351	0.022792023
0.809882	0.6365 (Probable ANTIGEN).	47/47	0.030885381	51/51	0.021884883	2	155	0.012903226	3	155	0.019354839
0.335187	0.4879 (Probable ANTIGEN).	47/47	0.031914894	51/51	0.024671407	5	182	0.027472527	5	182	0.027472527
0.905643	0.6549 (Probable ANTIGEN).	47/47	0.032292382	51/51	0.015528147	13	620	0.020967742	13	620	0.020967742
0.650735	0.7504 (Probable ANTIGEN).	47/47	0.030988213	51/51	0.008881601	4	287	0.013937282	6	287	0.020905923
0.773868	0.4910 (Probable ANTIGEN).	41/47	0.024822695	43/51	0.015082956	2	78	0.025641026	2	78	0.025641026
0.811642	0.6011 (Probable ANTIGEN).	47/47	0.030128432	51/51	0.015187921	18	661	0.027231467	14	661	0.02118003
0.923882	0.4799 (Probable ANTIGEN).	47/47	0.030958316	51/51	0.012074585	7	367	0.019073569	10	367	0.027247956
0.672117	0.4026 (Probable ANTIGEN).	47/47	0.030005456	51/51	0.017322547	10	429	0.023310023	9	429	0.020979021
0.756992	0.3146 (Probable NON-ANTIGEN).	47/47	0.026703202	51/51	0.017726282	14	792	0.017676768	13	792	0.016414141
0.88846	0.5064 (Probable ANTIGEN).	47/47	0.033937049	50/51	0.01895965	1	121	0.008264463	3	121	0.024793388
0.785869	0.5356 (Probable ANTIGEN).	46/47	0.026493845	50/51	0.012245511	7	261	0.026819923	7	261	0.026819923
NA (too big)	0.4449 (Probable ANTIGEN).	47/47	0.030630075	50/51	0.016576679	5	207	0.024154589	5	207	0.024154589
0.574888	0.6721 (Probable ANTIGEN).	47/47	0.031364637	51/51	0.013860717	4	348	0.011494253	7	348	0.020114943
0.907002	0.5756 (Probable ANTIGEN).	47/47	0.030746526	51/51	0.01700102	5	173	0.028901734	4	173	0.023121387
0.190419	0.5930 (Probable ANTIGEN).	46/47	0.03140289	50/51	0.016042781	4	187	0.021390374	4	187	0.021390374
0.080124	0.4561 (Probable ANTIGEN).	47/47	0.030371934	51/51	0.018859452	3	131	0.022900763	2	131	0.015267176
0.477155	0.8020 (Probable ANTIGEN).	47/47	0.019142051	51/51	0.018211815	4	309	0.012944984	6	309	0.019417476
0.614675	0.4591 (Probable ANTIGEN).	47/47	0.0224339	51/51	0.024366232	6	239	0.025104603	5	239	0.020920502
0.793408	0.7103 (Probable ANTIGEN).	47/47	0.017647059	51/51	0.015282584	4	340	0.011764706	7	340	0.020588235
0.95874	0.5769 (Probable ANTIGEN).	46/47	0.030209518	51/51	0.029187247	1	131	0.007633588	1	131	0.007633588
0.887067	0.9210 (Probable ANTIGEN).	47/47	0.033106383	51/51	0.017568627	4	250	0.016	6	250	0.024
0.704516	0.5615 (Probable ANTIGEN).	47/47	0.034110098	51/51	0.016308746	5	315	0.015873016	7	315	0.022222222
0.907579	0.4899 (Probable ANTIGEN).	47/47	0.033066811	51/51	0.019732734	6	471	0.012738854	6	471	0.012738854
0.929297	0.7312 (Probable ANTIGEN).	46/47	0.02901354	51/51	0.02150921	4	165	0.024242424	2	165	0.012121212
0.821862	0.4732 (Probable ANTIGEN).	47/47	0.033000434	51/51	0.021608643	3	196	0.015306122	3	196	0.015306122
0.304549	0.5542 (Probable ANTIGEN).	47/47	0.03318568	51/51	0.016496018	10	586	0.017064846	8	586	0.013651877
0.892487	0.5397 (Probable ANTIGEN).	47/47	0.031134101	49/51	0.013311747	6	218	0.027522936	5	218	0.02293578
0.950173	0.4525 (Probable ANTIGEN).	47/47	0.028775724	49/50	0.009964642	10	488	0.020491803	12	488	0.024590164
0.949494	0.6481 (Probable ANTIGEN).	47/47	0.030322334	50/51	0.016613831	6	334	0.017964072	9	334	0.026946108
0.890226	0.2502 (Probable NON-ANTIGEN).	47/47	0.028171789	51/51	0.014070443	5	216	0.023148148	5	216	0.023148148
0.80023	0.4254 (Probable ANTIGEN).	47/47	0.031648936	47/51	0.015808824	2	160	0.0125	3	160	0.01875
0.682792	0.3816 (Probable NON-ANTIGEN).	47/47	0.032239452	51/51	0.018079096	6	295	0.020338983	5	295	0.016949153
0.920054	0.5437 (Probable ANTIGEN).	47/47	0.034115225	51/51	0.016259255	10	527	0.018975332	11	527	0.020872865
0.904344	0.7246 (Probable ANTIGEN).	47/47	0.031859773	51/51	0.020217413	5	193	0.025906736	5	193	0.025906736
0.784612	0.9127 (Probable ANTIGEN).	46/47	0.025555042	50/51	0.023337596	3	184	0.016304348	2	184	0.010869565
0.940451	0.4962 (Probable ANTIGEN).	47/47	0.033054711	51/51	0.016316527	6	280	0.021428571	4	280	0.014285714
0.791845	0.6115 (Probable ANTIGEN).	47/47	0.032891611	51/51	0.020726901	7	403	0.017369727	5	403	0.012406948
0.831711	0.5512 (Probable ANTIGEN).	46/47	0.030750116	45/51	0.012594819	3	137	0.02189781	2	137	0.01459854
0.925761	0.5793 (Probable ANTIGEN).	47/47	0.03021148	44/51	0.027249571	7	331	0.021148036	6	331	0.018126888
0.850223	0.6279 (Probable ANTIGEN).	46/47	0.028307123	19/51	0.004092072	3	115	0.026086957	3	115	0.026086957
0.937913	0.6904 (Probable ANTIGEN).	47/47	0.032089292	51/51	0.015750563	7	244	0.028688525	7	244	0.028688525
0.704674	0.4800 (Probable ANTIGEN).	47/47	0.030925981	51/51	0.012537973	7	355	0.01971831	7	355	0.01971831
0.319204	0.5790 (Probable ANTIGEN).	47/47	0.033856189	51/51	0.016697677	8	411	0.01946472	8	411	0.01946472
0.72435	0.5146 (Probable ANTIGEN).	47/47	0.030649244	51/51	0.016297837	12	622	0.019292605	16	622	0.025723473
0.529446	0.3299 (Probable NON-ANTIGEN).	47/47	0.03114228	51/51	0.011392266	4	179	0.022346369	5	179	0.027932961
0.538797	0.5232 (Probable ANTIGEN).	47/47	0.032396023	50/51	0.014090058	4	199	0.020100503	3	199	0.015075377
0.855834	0.6663 (Probable ANTIGEN).	47/47	0.034641292	51/51	0.017390876	7	398	0.01758794	6	398	0.015075377
0.938587	0.4415 (Probable ANTIGEN).	47/47	0.029589968	51/51	0.015842098	3	151	0.01986755	3	151	0.01986755
0.784243	0.5847 (Probable ANTIGEN).	47/47	0.035598973	50/51	0.017547875	3	275	0.010909091	2	275	0.007272727
0.845601	0.6269 (Probable ANTIGEN).	47/47	0.030235162	36/51	0.010663915	3	114	0.026315789	3	114	0.026315789
0.722375	0.6082 (Probable ANTIGEN).	46/47	0.032996754	48/51	0.016616816	0	118	0	1	118	0.008474576
0.812169	0.6132 (Probable ANTIGEN).	47/47	0.035460993	51/51	1.529411765	2	138	0.014492754	2	138	0.014492754
0.636442	0.4531 (Probable ANTIGEN).	47/47	0.031550152	51/51	0.013781513	4	350	0.011428571	3	350	0.008571429
0.831638	0.6871 (Probable ANTIGEN).	47/47	0.031614377	50/51	0.029577933	4	177	0.02259887	5	177	0.028248588
0.739556	0.5079 (Probable ANTIGEN).	47/47	0.030669986	51/51	0.017991239	6	376	0.015957447	9	376	0.02393617
0.873975	0.4828 (Probable ANTIGEN).	47/47	0.032386542	51/51	0.020187385	5	203	0.024630542	5	203	0.024630542
0.868607	0.5571 (Probable ANTIGEN).	47/47	0.031344058	49/51	0.016738403	5	205	0.024390244	4	205	0.019512195
0.384736	0.4855 (Probable ANTIGEN).	47/47	0.034943853	51/51	0.015182462	3	288	0.010416667	6	288	0.020833333
0.903362	0.5020 (Probable ANTIGEN).	47/47	0.03269028	51/51	0.019150517	4	343	0.011661808	7	343	0.020408163
0.820348	0.6526 (Probable ANTIGEN).	45/47	0.026675135	51/51	0.013169447	1	134	0.007462687	2	134	0.014925373

Appendix 4 : (contains the detailed cumulative calculations for the antigenicity, MHC-I density, MHC-II density, allele coverage, and B-cell density. And the results of conserved domain search, and annotation from the three tools, Blannotator, Pannzer, and EggNOG of the top 34 proteins).

Appendix 4.1

NCBI ID	Length	NCBI-Description	Adhesion Probability (vaxign)	Overall Antigenicity	MHC-I Density	MHC-II Density	Allele coverage	Overall B cell Density	Score of 5
WP_004684144.1	274	porin family protein	0.59	0.694460966	0.960374	0.939789	1	0.969894	4.564518
WP_002964666.1	227	OmpW family protein	0.558	0.802877547	1	0.812266	1	0.906133	4.521276
WP_002969562.1	155	hypothetical protein	0.544	0.754198664	0.920436	0.887055	1	0.943527	4.505217
WP_002966849.1	280	DUF1849 domain-containing protein	0.6	0.689630255	0.933001	0.919651	1	0.959825	4.502107
WP_002964611.1	351	DUF1176 domain-containing protein	0.546	0.693372671	1.02428	0.815138	1	0.907569	4.44036
WP_004690357.1	284	porin family protein	0.553	0.750299346	0.937721	0.822744	1	0.911372	4.422137
WP_004681227.1	238	type IV secretion system protein virB1	0.666	0.622479847	0.921811	0.904958	0.989796	0.947377	4.386421
WP_002966226.1	182	hypothetical protein	0.54	0.430010852	0.951117	1	1	1	4.381128
WP_002963597.1	121	hypothetical protein	0.502	0.726630617	1.011381	0.768487	0.989796	0.879141	4.375436
WP_004688070.1	192	hypothetical protein	0.596	0.756313445	0.898277	0.782342	1	0.891171	4.328103
WP_002966502.1	126	hypothetical protein	0.627	0.730402286	0.890729	0.801067	1	0.900534	4.322732
WP_002964322.1	329	hypothetical protein	0.592	0.630945772	0.992554	0.741615	1	0.870807	4.235922
WP_002971090.1	267	hypothetical protein	0.547	0.695432551	0.971303	0.705461	1	0.85273	4.224927
WP_004691650.1	620	TonB-dependent receptor	0.537	0.813518354	0.962367	0.629399	1	0.814699	4.219983
WP_002971481.1	168	outer membrane protein assembly factor BamE	0.693	0.843703304	0.913375	0.6623	0.979592	0.820946	4.219915
WP_004683739.1	236	porin family protein	0.562	0.679882335	0.93231	0.734143	1	0.867071	4.213406
WP_004691134.1	403	hypothetical protein	0.623	0.806488637	0.85278	0.660656	1	0.830328	4.150253
WP_002965482.1	439	sugar ABC transporter substrate-binding protein	0.557	0.660534087	0.941729	0.682516	0.989796	0.836156	4.110731
WP_002964333.1	220	OmpA family protein	0.562	0.84728488	0.985703	0.516594	0.989796	0.753195	4.092573
WP_023080793.1	661	heme transporter BhuA	0.518	0.736522799	0.897878	0.615608	1	0.807804	4.057813
WP_002964719.1	261	porin family protein	0.57	0.692904865	0.923179	0.618147	1	0.809073	4.043304
WP_002966352.1	156	DUF2271 domain-containing protein	0.701	0.878095257	0.955182	0.499272	0.959184	0.729228	4.020961
WP_004690579.1	429	cell wall hydrolase	0.514	0.559938367	0.894213	0.70213	1	0.851065	4.007347
WP_004683944.1	212	porin family protein	0.62	0.903548892	0.867371	0.483604	0.989796	0.7367	3.98102
WP_011068938.1	792	LPS-assembly protein LptD	0.509	0.557742394	0.7958	0.718495	1	0.859247	3.931285
WP_002967296.1	166	BA14K family protein	0.602	0.521033517	0.679915	0.818698	1	0.909349	3.928996
WP_002964622.1	170	BA14K family protein	0.588	0.764828514	0.738515	0.621783	0.989796	0.805789	3.920711
WP_004683466.1	213	membrane protein	0.55	0.555717224	0.943675	0.604465	1	0.802233	3.90609
WP_002963776.1	115	DUF2147	0.609	0.810198426	0.865655	0.559787	0.918367	0.739077	3.893085
WP_004681306.1	367	iron ABC transporter substrate-binding protein	0.516	0.731083916	0.92261	0.489416	1	0.744708	3.887818
WP_002964998.1	177	hypothetical protein	0.679	0.813701131	0.870514	0.399625	0.959184	0.679404	3.722428
WP_002964530.1	287	outer membrane protein assembly factor BamD	0.526	0.73148021	0.923501	0.359996	1	0.679998	3.694974
WP_006278325.1	261	hypothetical protein	0.502	0.688763909	0.789561	0.496344	0.979592	0.737968	3.692229
WP_002963780.1	216	hypothetical protein	0.706	0.833965255	0.848373	0.342188	0.928571	0.63538	3.588478

Appendix 4.2

NCBI ID	Length	NCBI-Description	Adhesion Probability (vaxign)	Protein Antigenicity (ANTIGENpro)	Protein Antigenicity (VaxiJen)	MHC-I (Propred I) No. of Alleles in query	MHC-I Density (Sum epitopes/47 (Alleles)/length)	MHC-II (Propred II) No. of Alleles in query	MHC-II Density (Sum epitopes/47 (Alleles)/length)	B-Cell Density (BCPred) Density:N/length		B-Cell Density (AAP) Density:N/length	
WP_002963780.1	216	hypothetical protein	0.706	0.902746	0.6967	47/47	0.028467297	44/51	0.008442266	5	0.023148148	6	0.02777778
WP_002966352.1	156	DUF2271 domain-containing protein	0.701	0.882738	0.8005	47/47	0.032051282	47/51	0.012317748	4	0.025641026	4	0.02564103
WP_002971481.1	168	outer membrane protein assembly factor BamE	0.693	0.916867	0.7013	47/47	0.03064843	49/51	0.016339869	4	0.023809524	4	0.02380952
WP_002964998.1	177	hypothetical protein	0.679	0.931412	0.6298	46/47	0.029210242	48/51	0.009859311	5	0.028248588	5	0.02824859
WP_004681227.1	238	type IV secretion system protein virB1	0.666	0.776825	0.4183	46/47	0.030931522	51/51	0.022326578	7	0.029411765	6	0.02521008
WP_002966502.1	126	hypothetical protein	0.627	0.958819	0.4441	47/47	0.029888551	51/51	0.019763461	2	0.015873016	3	0.02380952
WP_004691134.1	403	hypothetical protein	0.623	0.95578	0.592	47/47	0.028615173	51/51	0.016299324	10	0.024813896	9	0.02233251
WP_004683944.1	212	porin family protein	0.62	0.941443	0.791	47/47	0.029104777	50/51	0.011931188	5	0.023584906	5	0.02358491
WP_002963776.1	115	DUF2147	0.609	0.89195	0.6621	45/47	0.029047179	45/51	0.013810742	2	0.017391304	3	0.02608696
WP_002967296.1	166	BA14K family protein	0.602	0.532994	0.4659	47/47	0.022814663	51/51	0.020198441	4	0.024096386	4	0.02409639
WP_002966849.1	280	DUF1849 domain-containing protein	0.6	0.902471	0.4221	47/47	0.031306991	51/51	0.022689076	4	0.014285714	7	0.025
WP_004688070.1	192	hypothetical protein	0.596	0.896893	0.5546	47/47	0.030141844	51/51	0.019301471	4	0.020833333	6	0.03125
WP_002964322.1	329	hypothetical protein	0.592	0.5	0.7078	47/47	0.033305309	51/51	0.01829668	5	0.015197568	8	0.02431611
WP_004684144.1	274	porin family protein	0.59	0.695898	0.6353	47/47	0.032225501	51/51	0.023185917	9	0.032846715	7	0.02554745
WP_002964622.1	170	BA14K family protein	0.588	0.814179	0.6525	47/47	0.024780976	50/51	0.015340254	3	0.017647059	4	0.02352941
WP_002964719.1	261	porin family protein	0.57	0.643178	0.6844	47/47	0.030977419	51/51	0.015250545	4	0.01532567	3	0.01149425
WP_002964333.1	220	OmpA family protein	0.562	0.909901	0.715	47/47	0.033075435	50/51	0.012745098	4	0.018181818	5	0.02272727
WP_004683739.1	236	porin family protein	0.562	0.758413	0.5458	47/47	0.031283808	51/51	0.01811233	4	0.016949153	6	0.02542373
WP_002964666.1	227	OmpW family protein	0.558	0.964208	0.5768	47/47	0.03355516	51/51	0.020039734	3	0.013215859	3	0.01321586
WP_002965482.1	439	sugar ABC transporter substrate-binding protein	0.557	0.800794	0.4671	47/47	0.031599864	50/51	0.016838626	9	0.020501139	10	0.02277904
WP_004690357.1	284	porin family protein	0.553	0.669913	0.7673	47/47	0.031465388	51/51	0.02029826	9	0.031690141	5	0.01760563
WP_004683466.1	213	membrane protein	0.55	0.240098	0.8212	47/47	0.031665168	51/51	0.014913007	3	0.014084507	6	0.02816901
WP_002971090.1	267	hypothetical protein	0.547	0.671545	0.6612	47/47	0.032592238	51/51	0.017404715	6	0.02247191	6	0.02247191
WP_002964611.1	351	DUF1176 domain-containing protein	0.546	0.793744	0.5366	47/47	0.034369885	51/51	0.020110608	7	0.01994302	8	0.02279202
WP_002969562.1	155	hypothetical protein	0.544	0.809882	0.6365	47/47	0.030885381	51/51	0.021884883	2	0.012903226	3	0.01935484
WP_002966226.1	182	hypothetical protein	0.54	0.335187	0.4879	47/47	0.031914894	51/51	0.024671407	5	0.027472527	5	0.02747253
WP_004691650.1	620	TonB-dependent receptor	0.537	0.905643	0.6549	47/47	0.032292382	51/51	0.015528147	13	0.020967742	13	0.02096774
WP_002964530.1	287	outer membrane protein assembly factor BamD	0.526	0.650735	0.7504	47/47	0.030988213	51/51	0.008881601	4	0.013937282	6	0.02090592
WP_023080793.1	661	heme transporter BhuA	0.518	0.811642	0.6011	47/47	0.030128432	51/51	0.015187921	18	0.027231467	14	0.02118003
WP_004681306.1	367	iron ABC transporter substrate-binding protein	0.516	0.923882	0.4799	47/47	0.030958316	51/51	0.012074585	7	0.019073569	10	0.02724796
WP_004690579.1	429	cell wall hydrolase	0.514	0.672117	0.4026	47/47	0.030005456	51/51	0.017322547	10	0.023310023	9	0.02097902
WP_011068938.1	792	LPS-assembly protein LptD	0.509	0.756992	0.3146	47/47	0.026703202	51/51	0.017726282	14	0.017676768	13	0.01641414
WP_002963597.1	121	hypothetical protein	0.502	0.88846	0.5064	47/47	0.033937049	50/51	0.01895965	1	0.008264463	3	0.02479339
WP_006278325.1	261	hypothetical protein	0.502	0.785869	0.5356	46/47	0.026493845	50/51	0.012245511	7	0.026819923	7	0.02681992

Appendix 4.3

NCBI ID	**Length**	**NCBI-Description**	Adhesion Probability (vaxign)	Protein Antigenicity (ANTIGENpro)	Protein Antigenicity (VaxiJen)	ANTIGENpro normalization	VaxiJen normalization	**overall antigenicity**
WP_002963780.1	216	hypothetical protein	0.706	0.902746	0.6967	0.936256492	0.731674018	0.833965255
WP_002966352.1	156	DUF2271 domain-containing protein	0.701	0.882738	0.8005	0.915505783	0.84068473	0.878095257
WP_002971481.1	168	outer membrane protein assembly factor BamE	0.693	0.916867	0.7013	0.950901673	0.736504936	0.843703304
WP_002964998.1	177	hypothetical protein	0.679	0.931412	0.6298	0.965986592	0.661415669	0.813701131
WP_004681227.1	238	type IV secretion system protein virB1	0.666	0.776825	0.4183	0.805661227	0.439298467	0.622479847
WP_002966502.1	126	hypothetical protein	0.627	0.958819	0.4441	0.994410957	0.466393615	0.730402286
WP_004691134.1	403	hypothetical protein	0.623	0.95578	0.592	0.991259147	0.621718126	0.806488637
WP_004683944.1	212	porin family protein	0.62	0.941443	0.791	0.976389949	0.830707834	0.903548892
WP_002963776.1	115	DUF2147	0.609	0.89195	0.6621	0.925059738	0.695337114	0.810198426
WP_002967296.1	166	BA14K family protein	0.602	0.532994	0.4659	0.552779068	0.489287965	0.521033517
WP_002966849.1	280	DUF1849 domain-containing protein	0.6	0.902471	0.4221	0.935971284	0.443289225	0.689630255
WP_004688070.1	192	hypothetical protein	0.596	0.896893	0.5546	0.930186225	0.582440664	0.756313445
WP_002964322.1	329	hypothetical protein	0.592	0.5	0.7078	0.518560311	0.743331233	0.630945772
WP_004684144.1	274	porin family protein	0.59	0.695898	0.6353	0.721730166	0.667191766	0.694460966
WP_002964622.1	170	BA14K family protein	0.588	0.814179	0.6525	0.84440183	0.685255198	0.764828514
WP_002964719.1	261	porin family protein	0.57	0.643178	0.6844	0.667053167	0.718756564	0.692904865
WP_002964333.1	220	OmpA family protein	0.562	0.909901	0.715	0.94367709	0.75089267	0.84728488
WP_004683739.1	236	porin family protein	0.562	0.758413	0.5458	0.786565762	0.573198908	0.679882335
WP_002964666.1	227	OmpW family protein	0.558	0.964208	0.5768	1	0.605755093	0.802877547
WP_002965482.1	439	sugar ABC transporter substrate-binding protein	0.557	0.800794	0.4671	0.830519971	0.490548204	0.660534087
WP_004690357.1	284	porin family protein	0.553	0.669913	0.7673	0.694780587	0.805818105	0.750299346
WP_004683466.1	213	membrane protein	0.55	0.240098	0.8212	0.249010587	0.862423861	0.555717224
WP_002971090.1	267	hypothetical protein	0.547	0.671545	0.6612	0.696473168	0.694391934	0.695432551
WP_002964611.1	351	DUF1176 domain-containing protein	0.546	0.793744	0.5366	0.82320827	0.563537072	0.693372671
WP_002969562.1	155	hypothetical protein	0.544	0.809882	0.6365	0.839945323	0.668452006	0.754198664
WP_002966226.1	182	hypothetical protein	0.54	0.335187	0.4879	0.34762935	0.512392355	0.430010852
WP_004691650.1	620	TonB-dependent receptor	0.537	0.905643	0.6549	0.939261031	0.687775677	0.813518354
WP_002964530.1	287	outer membrane protein assembly factor BamD	0.526	0.650735	0.7504	0.674890687	0.788069733	0.73148021
WP_023080793.1	661	heme transporter BhuA	0.518	0.811642	0.6011	0.841770655	0.631274942	0.736522799
WP_004681306.1	367	iron ABC transporter substrate-binding protein	0.516	0.923882	0.4799	0.958177074	0.503990758	0.731083916
WP_004690579.1	429	cell wall hydrolase	0.514	0.672117	0.4026	0.697066401	0.422810334	0.559938367
WP_011068938.1	792	LPS-assembly protein LptD	0.509	0.756992	0.3146	0.785092013	0.330392775	0.557742394
WP_002963597.1	121	hypothetical protein	0.502	0.88846	0.5064	0.921440187	0.531821046	0.726630617
WP_006278325.1	261	hypothetical protein	0.502	0.785869	0.5356	0.815040946	0.562486873	0.688763909

Appendix 4.4

NCBI ID	**Length**	**NCBI-Description**	Adhesion Probability (vaxign)	MHC-I (Propred I) No. of Alleles in query	MHC-I Density (Sum epitopes/47 (Alleles)/length)	MHC-II (Propred II) No. of Alleles in query	MHC-II Density (Sum epitopes/51 (Alleles)/length)	MHC-I	MHC-II	Allele counts for both MHC (98)	Allele coverage
WP_002963780.1	216	hypothetical protein	0.706	47	0.028467297	44	0.008442266	0.848373164	0.342188258	91	0.9285714
WP_002966352.1	156	DUF2271 domain-containing protein	0.701	47	0.032051282	47	0.012317748	0.955181922	0.499272198	94	0.9591837
WP_002971481.1	168	outer membrane protein assembly factor BamE	0.693	47	0.03064843	49	0.016339869	0.913374568	0.662299854	96	0.9795918
WP_002964998.1	177	hypothetical protein	0.679	46	0.029210242	48	0.009859311	0.870514156	0.399624997	94	0.9591837
WP_004681227.1	238	type IV secretion system protein virB1	0.666	46	0.030931522	51	**0.022326578**	0.921811183	0.904957616	97	0.9897959
WP_002966502.1	126	hypothetical protein	0.627	47	0.029888551	51	0.019763461	0.890728917	0.801067443	98	1
WP_004691134.1	403	hypothetical protein	0.623	47	0.028615173	51	0.016299324	0.852780127	0.66065643	98	1
WP_004683944.1	212	porin family protein	0.62	47	0.029104777	50	0.011931188	0.867371139	0.483603856	97	0.9897959
WP_002963776.1	115	DUF2147	0.609	45	0.029047179	45	0.013810742	0.865654603	0.559787355	90	0.9183673
WP_002967296.1	166	BA14K family protein	0.602	47	0.022814663	51	0.020198441	0.679915191	0.818698374	98	1
WP_002966849.1	280	DUF1849 domain-containing protein	0.6	47	0.031306991	51	**0.022689076**	0.933000798	0.919650655	98	1
WP_004688070.1	192	hypothetical protein	0.596	47	0.030141844	51	0.019301471	0.898277467	0.782341703	98	1
WP_002964322.1	329	hypothetical protein	0.592	47	**0.033305309**	51	0.01829668	0.992554041	0.741614791	98	1
WP_004684144.1	274	porin family protein	0.59	47	0.032225501	51	**0.023185917**	0.960373935	0.939788991	98	1
WP_002964622.1	170	BA14K family protein	0.588	47	0.024780976	50	0.015340254	0.738514624	0.621782687	97	0.9897959
WP_002964719.1	261	porin family protein	0.57	47	0.030977419	51	0.015250545	0.923179006	0.618146531	98	1
WP_002964333.1	220	OmpA family protein	0.562	47	**0.033075435**	50	0.012745098	0.985703403	0.516593886	97	0.9897959
WP_004683739.1	236	porin family protein	0.562	47	0.031283808	51	0.01811233	0.932309914	0.734142551	98	1
WP_002964666.1	227	OmpW family protein	0.558	47	**0.03355516**	51	0.020039734	1	0.812265548	98	1
WP_002965482.1	439	sugar ABC transporter substrate-binding protein	0.557	47	0.031599864	50	0.016838626	0.941728917	0.682515841	97	0.9897959
WP_004690357.1	284	porin family protein	0.553	47	0.031465388	51	0.02029826	0.9377213	0.822744326	98	1
WP_004683466.1	213	membrane protein	0.55	47	0.031665168	51	0.014913007	0.943675086	0.604465219	98	1
WP_002971090.1	267	hypothetical protein	0.547	47	0.032592238	51	0.017404715	0.971303329	0.705460969	98	1
WP_002964611.1	351	DUF1176 domain-containing protein	0.546	47	**0.034369885**	51	0.020110608	1.024280189	0.815138282	98	1
WP_002969562.1	155	hypothetical protein	0.544	47	0.030885381	51	0.021884883	0.920436115	0.887054515	98	1
WP_002966226.1	182	hypothetical protein	0.54	47	0.031914894	51	**0.024671407**	0.951117318	1	98	1
WP_004691650.1	620	TonB-dependent receptor	0.537	47	0.032292382	51	0.015528147	0.962367093	0.629398507	98	1
WP_002964530.1	287	outer membrane protein assembly factor BamD	0.526	47	0.030988213	51	0.008881601	0.923500672	0.35999574	98	1
WP_023080793.1	661	heme transporter BhuA	0.518	47	0.030128432	51	0.015187921	0.897877771	0.615608216	98	1
WP_004681306.1	367	iron ABC transporter substrate-binding protein	0.516	47	0.030958316	51	0.012074585	0.922609715	0.489416132	98	1
WP_004690579.1	429	cell wall hydrolase	0.514	47	0.030005456	51	0.017322547	0.894212863	0.702130475	98	1
WP_011068938.1	792	LPS-assembly protein LptD	0.509	47	0.026703202	51	0.017726282	0.795800181	0.718494994	98	1
WP_002963597.1	121	hypothetical protein	0.502	47	**0.033937049**	50	0.01895965	1.01138095	0.768486773	97	0.9897959
WP_006278325.1	261	hypothetical protein	0.502	46	0.026493845	50	0.012245511	0.789560992	0.496344259	96	0.9795918

Appendix 4.5

NCBI ID	**Length**	**NCBI-Description**	Adhesion Probability (vaxign)	B-Cell Density (BCPred) Density:N/length		B-Cell Density (AAP) Density:N/length		BCPred normalization	AAP normalization	**overall B cell density**
WP_002963780.1	216	hypothetical protein	0.706	5	0.023148148	6	0.027777778	0.70473251	0.801587302	0.753159906
WP_002966352.1	156	DUF2271 domain-containing protein	0.701	4	0.025641026	4	0.025641026	0.780626781	0.73992674	0.76027676
WP_002971481.1	168	outer membrane protein assembly factor BamE	0.693	4	0.023809524	4	0.023809524	0.724867725	0.68707483	0.705971277
WP_002964998.1	177	hypothetical protein	0.679	5	0.028248588	5	0.028248588	0.860012555	0.815173527	0.837593041
WP_004681227.1	238	type IV secretion system protein virB1	0.666	7	0.029411765	6	0.025210084	0.895424837	0.727490996	0.811457916
WP_002966502.1	126	hypothetical protein	0.627	2	0.015873016	3	0.023809524	0.48324515	0.68707483	0.58515999
WP_004691134.1	403	hypothetical protein	0.623	10	0.024813896	9	0.022332506	0.755445272	0.644452322	0.699948797
WP_004683944.1	212	porin family protein	0.62	5	0.023584906	5	0.023584906	0.71802935	0.680592992	0.699311171
WP_002963776.1	115	DUF2147	0.609	2	0.017391304	3	0.026086957	0.529468599	0.752795031	0.641131815
WP_002967296.1	166	BA14K family protein	0.602	4	0.024096386	4	0.024096386	0.733601071	0.69535284	0.714476955
WP_002966849.1	280	DUF1849 domain-containing protein	0.6	4	0.014285714	7	0.025	0.434920635	0.721428571	0.578174603
WP_004688070.1	192	hypothetical protein	0.596	4	0.020833333	6	0.03125	0.634259259	0.901785714	0.768022487
WP_002964322.1	329	hypothetical protein	0.592	5	0.015197568	8	0.024316109	0.462681527	0.701693443	0.582187485
WP_004684144.1	274	porin family protein	0.59	9	0.032846715	7	0.025547445	1	0.737226277	0.868613139
WP_002964622.1	170	BA14K family protein	0.588	3	0.017647059	4	0.023529412	0.537254902	0.678991597	0.608123249
WP_002964719.1	261	porin family protein	0.57	4	0.01532567	3	0.011494253	0.466581524	0.331691297	0.399136411
WP_002964333.1	220	OmpA family protein	0.562	4	0.018181818	5	0.022727273	0.553535354	0.655844156	0.604689755
WP_004683739.1	236	porin family protein	0.562	4	0.016949153	6	0.025423729	0.516007533	0.733656174	0.624831854
WP_002964666.1	227	OmpW family protein	0.558	3	0.013215859	3	0.013215859	0.402349486	0.381371932	0.391860709
WP_002965482.1	439	sugar ABC transporter substrate-binding protein	0.557	9	0.020501139	10	0.022779043	0.624145786	0.657338106	0.640741946
WP_004690357.1	284	porin family protein	0.553	9	0.031690141	5	0.017605634	0.964788732	0.50804829	0.736418511
WP_004683466.1	213	membrane protein	0.55	3	0.014084507	6	0.028169014	0.428794992	0.812877264	0.620836128
WP_002971090.1	267	hypothetical protein	0.547	6	0.02247191	6	0.02247191	0.684144819	0.64847512	0.66630997
WP_002964611.1	351	DUF1176 domain-containing protein	0.546	7	0.01994302	8	0.022792023	0.607154163	0.657712658	0.63243341
WP_002969562.1	155	hypothetical protein	0.544	2	0.012903226	3	0.019354839	0.392831541	0.558525346	0.475678443
WP_002966226.1	182	hypothetical protein	0.54	5	0.027472527	5	0.027472527	0.836385836	0.79277865	0.814582243
WP_004691650.1	620	TonB-dependent receptor	0.537	13	0.020967742	13	0.020967742	0.638351254	0.605069124	0.621710189
WP_002964530.1	287	outer membrane protein assembly factor BamD	0.526	4	0.013937282	6	0.020905923	0.424312815	0.603285217	0.513799016
WP_023080793.1	661	heme transporter BhuA	0.518	18	0.027231467	14	0.02118003	0.829046899	0.611195159	0.720121029
WP_004681306.1	367	iron ABC transporter substrate-binding protein	0.516	7	0.019073569	10	0.027247956	0.580684226	0.78629817	0.683491198
WP_004690579.1	429	cell wall hydrolase	0.514	10	0.023310023	9	0.020979021	0.70966071	0.605394605	0.657527658
WP_011068938.1	792	LPS-assembly protein LptD	0.509	14	0.017676768	13	0.016414141	0.538159371	0.473665224	0.505912298
WP_002963597.1	121	hypothetical protein	0.502	1	0.008264463	3	0.024793388	0.251606979	0.715466352	0.483536665
WP_006278325.1	261	hypothetical protein	0.502	7	0.026819923	7	0.026819923	0.816517667	0.77394636	0.795232014

Appendix 4.6

Query	Hit type	From	To	E-Value	Bitscore	Accession	Short name	Incomplete	Superfamily
Q#2 - WP_002966352.1	specific	24	140	3.11E-41	136.942	pfam10029	DUF2271	-	cl21544
Q#3 - WP_002971481.1	specific	8	154	1.93E-56	177.212	COG2913	BamE	-	cl19285
Q#4 - WP_002964998.1	specific	86	175	1.33E-22	88.5378	pfam02974	Inh	-	cl03811
Q#5 - WP_004681227.1	superfamily	7	165	4.00E-21	89.1483	cl00222	lysozyme_like superfamily	C	-
Q#7 - WP_004691134.1	specific	2	299	3.56E-135	392.951	COG3034	YafK	-	cl21614
Q#7 - WP_004691134.1	superfamily	238	396	0.005997	38.6772	cl26386	DNA_pol3_gamma3 superfamily	NC	-
Q#8 - WP_004683944.1	specific	40	212	1.14E-25	100.534	COG3637	LomR	-	cl21487
Q#10 - WP_002963776.1	specific	1	115	6.73E-22	86.3671	COG4731	COG4731	N	cl01951
Q#11 - WP_002967296.1	specific	134	163	9.05E-13	60.1561	pfam07886	BA14K	-	cl06774
Q#12 - WP_002966849.1	specific	31	270	7.58E-116	336.919	pfam08904	DUF1849	-	cl07482
Q#14 - WP_002964322.1	specific	28	326	1.13E-129	377.056	TIGR02122	TRAP_TAXI	-	cl26668
Q#15 - WP_004684144.1	specific	11	256	2.68E-31	117.098	COG3637	LomR	-	cl21487
Q#16 - WP_002964622.1	specific	139	168	2.72E-10	53.2225	pfam07886	BA14K	-	cl06774
Q#17 - WP_002964719.1	specific	9	261	1.77E-35	127.884	COG3637	LomR	-	cl21487
Q#18 - WP_002964333.1	specific	77	216	4.36E-59	187.749	COG2885	OmpA	N	cl28145
Q#19 - WP_004683739.1	specific	31	236	6.49E-41	141.366	COG3637	LomR	-	cl21487
Q#20 - WP_002964666.1	specific	43	227	4.49E-83	249.966	COG3047	OmpW	-	cl21487
Q#21 - WP_002965482.1	specific	42	429	4.18E-56	193.005	cd13585	PBP2_TMBP_like	-	cl21456
Q#23 - WP_004690357.1	specific	1	283	9.46E-34	124.032	COG3637	LomR	-	cl21487
Q#25 - WP_004683466.1	specific	5	213	1.74E-41	142.136	COG3637	LomR	-	cl21487
Q#26 - WP_002971090.1	superfamily	189	267	5.82E-05	43.0689	cl21487	OM_channels superfamily	N	-
Q#27 - WP_002964611.1	specific	32	349	6.04E-91	278.874	pfam06674	DUF1176	-	cl05951
Q#29 - WP_002966226.1	specific	1	178	3.78E-104	300.562	COG3470	Tpd	-	cl01377
Q#30 - WP_004691650.1	specific	21	620	0	757.343	COG4206	BtuB	-	cl26861
Q#31 - WP_002964530.1	specific	17	267	3.97E-117	340.83	COG4105	BamD	-	cl26155
Q#33 - WP_023080793.1	superfamily	38	655	7.15E-67	234.209	cl26861	CirA superfamily	-	-
Q#34 - WP_004681306.1	specific	56	365	1.06E-59	197.658	COG1840	AfuA	-	cl25412
Q#35 - WP_004690579.1	specific	180	429	2.06E-109	326.317	COG3773	CwlJ	-	cl26889
Q#36 - WP_011068938.1	specific	7	791	0	725.363	COG1452	LptD	-	cl27152
Q#37 - WP_002963597.1	specific	27	121	3.71E-40	132.625	COG3895	MliC	-	cl01604
Q#38 - WP_006278325.1	superfamily	78	164	4.80E-06	47.0164	cl26386	DNA_pol3_gamma3 superfamily	NC	-

Appendix 4.7

NCBI ID	Length	NCBI Description	Blannotator	PANNZER	EggNoc
WP_002963776.1	115	DUF2147	putative protein without homology	sn-glycerol-3-phosphate ABC transporter ATP-binding protein	Uncharacterized protein conserved in bacteria (DUF2147)
WP_002963597.1	121	hypothetical protein	Membrane-bound lysozyme inhibitor of C-type lysozyme	Membrane-bound lysozyme inhibitor of C-type lysozyme	Membrane-bound lysozyme-inhibitor of c-type lysozyme
WP_002966502.1	126	hypothetical protein	Outer membrane lipoprotein omp10	Outer membrane protein 10	Outer membrane lipoprotein
WP_002969562.1	155	hypothetical protein	putative protein without homology	Uncharacterized protein	_
WP_002966352.1	156	DUF2271 domain-containing protein	putative protein without homology	Tat pathway signal protein	Predicted periplasmic protein (DUF2271)
WP_002967296.1	166	BA14K family protein	putative protein without homology	Immunoreactive protein BA14k	BA14K family
WP_002971481.1	168	outer membrane protein assembly factor BamE	putative protein without homology	Outer membrane protein assembly factor BamE	smpa omla domain-containing protein
WP_002964622.1	170	BA14K family protein	putative protein without homology	Glutelin	BA14K-like protein
WP_002964998.1	177	hypothetical protein	Outer membrane lipoprotein omp19	Outer membrane lipoprotein omp19	Outer membrane lipoprotein
WP_002966226.1	182	hypothetical protein	UPF0423 protein BAB2_0840	Pathogen-specific membrane antigen	Periplasmic Protein
WP_004688070.1	192	hypothetical protein	putative protein without homology	Uncharacterized protein	NA
WP_004683944.1	212	porin family protein	putative protein without homology	Outer-membrane protein Omp25	membrAne
WP_004683466.1	213	membrane protein	25 kDa outer-membrane immunogenic protein	Outer membrane protein 25	membrAne
WP_002963780.1	216	hypothetical protein	putative protein without homology	Uncharacterized protein	NA
WP_002964333.1	220	OmpA family protein	Probable lipoprotein YiaD	Cell envelope biogenesis protein OmpA	Ompa motb domain protein
WP_002964666.1	227	OmpW family protein	Uncharacterized outer-membrane protein y4mB	OmpW family outer membrane protein	outer membrane protein W
WP_004683739.1	236	porin family protein	putative protein without homology	Autotransporter outer membrane beta-barrel domain-containing protein (Fragment)	hemin binding protein
WP_004681227.1	238	type IV secretion system protein virB1	Type IV secretion system protein virB1	Type IV secretion system protein virB1	Conjugal transfer protein
WP_002964719.1	261	porin family protein	31 kDa outer-membrane immunogenic protein	Outer membrane protein Omp31	Outer membrane protein Omp31
WP_006278325.1	261	hypothetical protein	putative protein without homology	Proline-rich region:Proline-rich extensin	
WP_002971090.1	267	hypothetical protein	putative protein without homology	Outer membrane beta-barrel domain protein	NA
WP_004684144.1	274	porin family protein	putative protein without homology	Porin opacity type	heat resistant agglutinin 1
WP_002966849.1	280	DUF1849 domain-containing protein	putative protein without homology	ATP/GTP-binding site domain-containing protein A	Domain of unknown function (DUF1849)
WP_004690357.1	284	porin family protein	Uncharacterized protein BRA0921/BS1330_II0913	Heat resistant agglutinin 1	heat resistant agglutinin 1
WP_002964530.1	287	outer membrane protein assembly factor BamD	Outer membrane protein assembly factor BamD	Outer membrane protein assembly factor BamD	Part of the outer membrane protein assembly complex, which is involved in assembly and insertion of beta-barrel proteins into the outer membrane
WP_002964322.1	329	hypothetical protein	31 kDa immunogenic protein	Alkanesulfonate transporter substrate-binding subunit	trap transporter solute receptor taxi family
WP_002964611.1	351	DUF1176 domain-containing protein	putative protein without homology	Uncharacterized protein	Protein of unknown function (DUF1176)
WP_004681306.1	367	iron ABC transporter substrate-binding protein	putative protein without homology	Periplasmic binding ABC transporter	solute-binding protein
WP_004691134.1	403	hypothetical protein	Putative L,D-transpeptidase YafK	Pollen allergen Poa pIX/Phl pVI	ErfK ybiS ycfS ynhG family protein
WP_004690579.1	429	cell wall hydrolase	putative protein without homology	Cell wall hydrolyses involved in spore germination	Cell wall hydrolase
WP_002965482.1	439	sugar ABC transporter substrate-binding protei	putative protein without homology	ABC-type sugar transport system periplasmic component	extracellular solute-binding protein family 1
WP_004691650.1	620	TonB-dependent receptor	putative protein without homology	Iron compound TonB-dependent receptor	Involved in the active translocation of vitamin B12 (cyanocobalamin) across the outer membrane to the periplasmic space. It derives its energy for transport by interacting with the trans-periplasmic membrane protein TonB (By similarity)
WP_023080793.1	661	heme transporter BhuA	Heme transporter BhuA	Heme transporter bhuA	receptor
WP_011068938.1	792	LPS-assembly protein LptD	putative protein without homology	LPS-assembly protein LptD	involved in the assembly of LPS in the outer leaflet of the outer membrane. Determines N-hexane tolerance and is involved in outer membrane permeability. Essential for envelope biogenesis (By similarity)