The University Of Sheffield.

Investigation of Immunity Related Genes in a Disease Host Using Applied Bioinformatics

By:

Ashraf Mohamed Ahmed

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Under the supervision of Professor David Hornby

The University of Sheffield
Faculty of Science
Department of Molecular Biology and Biotechnology

June 2017

Acknowledgement

I would like to express my sincere gratitude to my supervisor Prof. David Hornby for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D. study.

I am grateful to Dr. Qaiser Sheikh for enlightening me the first glance of research and for the support, advice and guidance throughout my entire course of study especially in the first year.

My sincere thanks also goes to Prof. Richard Pleass for the provision of the badger's raw data, Liverpool life science UTC for supporting mealworm genome sequencing and Dr. Roy Chaudhuri, who provided me an opportunity to access the bioinformatics research facilities. Without their precious support it would not be possible to conduct this research.

I thank my fellow lab colleagues in for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last four years.

Last but not the least, I would like to thank my family: my wife and to my children for supporting me spiritually throughout writing this thesis and my life in general.

Summary

The economic burden and the health risks of bovine tuberculosis have led to an ongoing political and scientific debate on the control of the disease in badgers, perceived to be the main carrier, responsible for spreading the infection among cattle. Although culling of badgers has already been introduced in some parts of Britain, its efficacy remains unclear. Moreover, the implementation of alternative strategies, such as vaccination illustrate the need for a deeper understanding of the badger's immune system. In addition, there is also a need to develop additional models and systems for studying the complexity of the immunological response in host organisms: simple organisms, including many flies and beetles are becoming increasingly popular in this respect.

The first aim of this thesis is to obtain the nucleotide sequence of the badger transcriptome from peripheral blood cells, and to profile the immunity related genes, through critical evaluation of bioinformatics data extracted from public domain databases. In the second part of the thesis, the introduction of the yellow mealworm beetle, *Tenebrio molitor*, is developed through initiation of a genome sequencing project. It is hoped that this simple organism will provide insight into immune challenge and support the annotation of immune-related genes from more complex organisms, including the badger as well as providing an accessible model organism that is easy to manipulate in simpler laboratory environment such as schools.

The sequencing of both the badger transcriptome from peripheral blood cells and the *T. molitor* genome generated large data sets. The transcriptome analysis resulted in the identification of 15967 transcripts related to 698 known immunity genes in different mammals. 1825 transcripts were found to match genes involved in tuberculosis pathogenesis.

It is believed that, these findings will improve our understanding of future attempts to both prevent and treat bovine tuberculosis.

The determination of the *T. molitor* genome will facilitate and improve its use as a model organism to study infections. The genome data have been deposited and assembly of an annotated genome, although incomplete, is currently best described as “work in progress”.

Abbreviations

AHVLA	Animal Health and Veterinary Laboratories Agency
APCs	Antigen presenting cells
BCG	Bacillus Calmette–Guérin vaccine
BGI	Beijing Genomics Institute
BLAST	NCBI Basic Local Alignment Search Tool
Blastn	BLAST nucleotide sequences
Blastp	BLAST protein sequences
bp	Base-pair
bTB	Bovine Tuberculosis
cDNA	Complementary DNA
CMI	Cell-mediated immunity
COG	Clusters of Orthologous Groups database
DCs	Dendritic cells
DEFRA	Department for Environment, Food & Rural Affairs
dNTPs	Deoxynucleotides
ESTs	Expressed sequence tags
GB	Gigabyte
GO	Gene Ontology
IFN-γ	interferon gamma
IKB	Immunome Knowledge Base
IL	interleukin
KEGG	Kyoto Encyclopaedia of Genes and Genomes
LPS	Lipid polysaccharide
LTT	Lymphocyte transformation test
MHC	Major histocompatibility complex
mRNA	Messenger RNA
N50	A statistical measure of average length of a set of sequences
NCBI	National Centre for Biotechnology Information
ncRNA	Non-coding RNA
NGS	Next generation sequencing technology
NK	Natural killer cells
NR (nr)	Non-redundant database
nt	Nucleotide
NT	NCBI nucleotide database
RBCT	Randomised badger culling trail
RNA	Ribonucleic acid
RNA-Seq	RNA sequencing
rpm	Revolutions per minute
snRNA	small nuclear RNA
SRA	Sequence Read Archive of NCBI
TNF	Tumour necrosis factor
tRNA	transfer RNA

Table of Contents

Figures and Tables

Figures

Tables

Chapter I

1 Introduction

The British Badger culling programme started in 2013 in several areas of southwest of England and Wales, in order to control the spread of bovine tuberculosis (Ares, 2014). However, the controversy surrounding the programme itself from an animal rights perspective, and the actual benefits of its implementation, became a debate long before that. Alternative solutions were suggested, including vaccination of both badgers and cattle. However, culling still continues to this day with no clear outcome agreed by both sides of the debate.

Bovine tuberculosis (bTB) infection is endemic in the badger population, which is recognized as a reservoir of infection for cattle and domestic animals in the UK and the Republic of Ireland (Corner et al., 2011). Establishing the identity of the reservoir of the pathogen (*Mycobacterium bovis)* is essential in order for any plan of eradication to be successful (Gormley and Collins, 2000).

1.1 The economic burden of bTB

Although bTB is a zoonotic disease with a significant health impact in developing countries in particular, the problem of zoonosis seems to be rare in the developed world; even in countries with large infected cattle populations such as the UK (Michel et al., 2010). Bovine tuberculosis is known for its economic impact on the cattle industry both in the UK and globally. According to the UK Department for Environment, Food and Rural affairs report on bTB impact assessment (2011), the disease is one of the major challenges facing the farming industry today: the cost of such a burden was estimated at £90 million, and nearly 25,000 cattle were slaughtered in 2010. Control of the disease in cattle can be

particularly challenging when wildlife becomes an integral part of the epidemiological system (DEFRA, 2011).

Internationally, the most significant wildlife reservoirs of bTB for cattle are considered to be the white-tailed deer (*Odocoileus virginianus*) in northern America, the Cape buffalo (*Syncerus caffer*) in South Africa and the brush-tail possum (*Trichosurus vulpecula*) in New Zealand (de Lisle et al., 2002). However, there are many other potential mammalian hosts, some of which may be capable of onward transmission (Delahay et al., 2002).

1.2 Badgers as bTB carriers

1.2.1 Badger Biology

Standard biological classification places the badger (*Meles meles*) under the family "*Mustelidae"*, which also includes weasels, otters and ferrets (Corner et al., 2011). Badgers are social animals that live in groups in underground burrows called setts; and in highly populated areas a single sett can host a family of 8-20 animals of both sexes and different ages (Delahay et al., 2000). Each sett comprises nesting chambers that are inter-connected by tunnels that spread over many metres with multiple entrances. These setts protect the animals from extreme weather and predators beside their main use for resting and breeding (Rogers et al., 2003). Badgers are nocturnal animals that are also territorial and territory boundaries are maintained by bodily secretions like urine, faeces and secretions from inter-digital glands (Corner et al., 2011). Although the *Mustelidae* family is carnivorous, badgers have diverse dietary preferences with seasonal variation depending on food availability (Cleary et al., 2009). Their diet includes invertebrates,

insects, amphibians, small mammals and decaying carcasses, as well as fruits, cereals and vegetation (Cleary et al., 2009). However, they can show highly specific feeding behaviours where diet is composed mainly of one type, such as earthworms in south-west England (Kruuk et al., 1979)

1.2.2 Tuberculosis in Badgers

Tuberculosis infected badgers in England and Ireland were first reported in 1971 (Murhead and Burns, 1974) and since then infected badgers have been found widely. Badgers have, since then, been considered to be the main wildlife reservoir that spreads *M. bovis* infection among cattle in the UK (Woodroffe et al., 2005) and Ireland (Griffin et al., 2005).

Badgers are not the only wildlife species that can acquire bTB infection as surveys have shown in south-west England (Delahay et al., 2001). Other wild mammals can carry the infection as well; however, the prevalence is lower in such cases, with milder lesions (Neal and Cheeseman, 1996).

The Eurasian badger (*Meles meles*) is one of the main sources of infection. The progression of the disease in cattle can cause reduced productivity and premature death, which makes disease control particularly challenging in the cattle population. In recent years, there has been growing interest in the prospective utilisation of farm husbandry and biosecurity measures to reduce the risk of bTB transmission (Krebs et al., 1997).

1.2.3 Routes of bTB transmission

Despite the fact that the exact processes by which cattle become infected with the bacteria has yet to be completely understood and characterised; it is generally assumed that inhalation is the main route of infection. Intra-tracheal inoculation with a single colony forming unit of *M. bovis* is sufficient to cause infection (Dean et al., 2005).

As a consequence, combined with the difficulties associated with laboratory culture *of M. bovis*, positive results from diagnostic methods provide a risk alert, but do not guarantee certainty in respect of transmission.

There is strong evidence that airborne infection is likely to be an important route for transmission of bTB amongst badgers. However, clinical samples from live badgers have shown that *M. bovis* bacilli can be isolated from sputum, faeces, urine, bite wounds and draining abscesses. This suggests that the increased opportunities for transmission of bTB to cattle occur either via direct contact with badgers (aerosols) or through material contaminated with badger secretions and excretory products such as urine, saliva and blood (Delahay et al., 2005).

The levels of increased risk of transmission of bTB through various routes, in accordance with the stage of infection are shown in **Figure 1.1** (Corner et al., 2011).

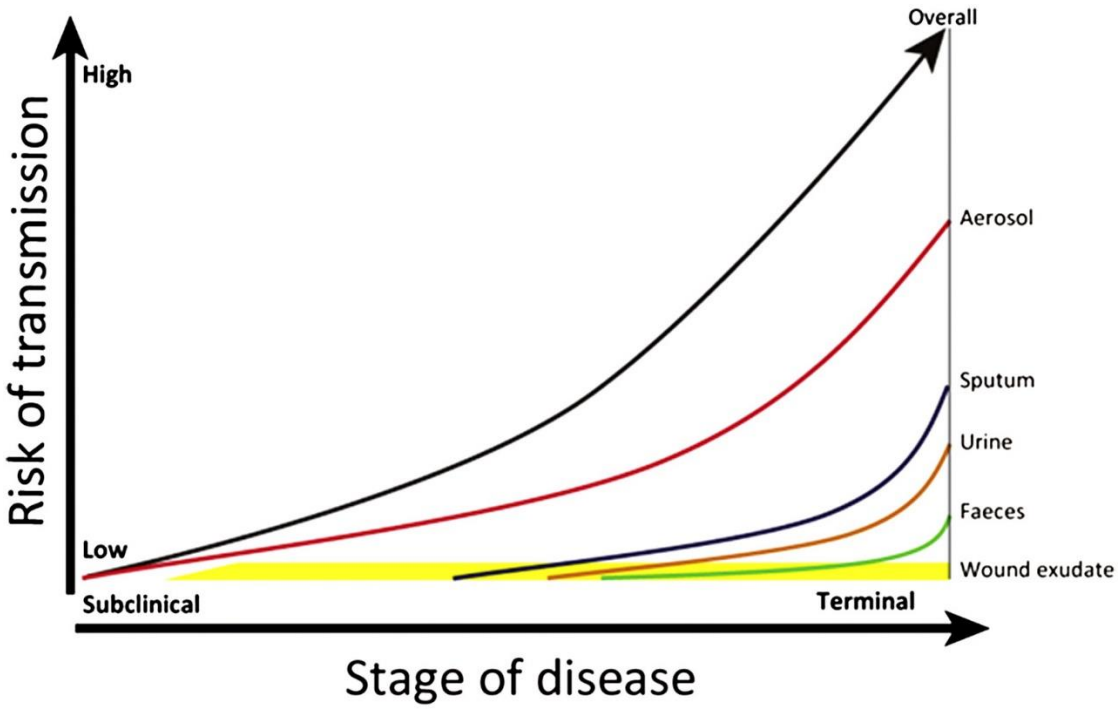


Figure 1.1: Routes of excretion and the risk of transmission of Mycobacterium bovis infection from tuberculosis infected badgers (Corner et al., 2011)

1.3 An overview of the mammalian immune system:

1.3.1 Pathogenesis: Cellular progression of infection

Pathogenesis is a term that defines a series of interactions between the microbe and the host, from the initiation of infection to the development of disease and the appearance of signs and symptoms. bTB is a respiratory and an immuno-pathological disease in which *M. bovis* bacilli enter the host through uptake by alveolar macrophages, and the formation of lesions is a result of cell-mediated immunity to the presence of *M. bovis* bacilli (Corner et al., 2011). Once the bacteria are systemic within the badger, *M. bovis* bacilli are engulfed by macrophages through phagocytosis. The bacilli then acquire

protection from extracellular bactericidal factors by the phagosome membrane. Mycobacteria have the ability to inhibit the fusion of lysosomes with phagosomes and thus protect themselves from intracellular destruction by acidification. The immunological response is then initiated through the complex interaction of numerous cell types, leading to the development of a cell-mediated immune (CMI) response that in turn, leads to the formation of granuloma around infected macrophages. The granuloma is a protective response which serves to localise the infection and to prevent further dissemination by the interactions of immune cells and cytokines (Robinson et al., 2012).

The body is constantly exposed to pathogens and in order to defend itself and the immune system has evolved to protect it against infection. The first line of defence against pathogens is the anatomical and physical barriers, which include the skin, the cilia of the lungs and lysozyme in tears, and these prevent the entry of some pathogenic organisms into the body. Viruses, bacteria, parasites and fungi must penetrate these shields to cause an infection or an infestation (Bradford, 2012). In the human body, the first barrier against infection is the skin even though it only stretches up to two square meters only, whereas the area covered by other epithelial and mucous membranes that line the digestive, respiratory and reproductive systems is two hundred times that of the skin (Sompayrac, 2012). Many pathogens are able to evade these anatomical and physiological barriers, and this is where the adaptive immune system is of a paramount importance. The immune system comprises two major components: (1) innate immunity and (2) adaptive immunity **(Figure 1.2).**

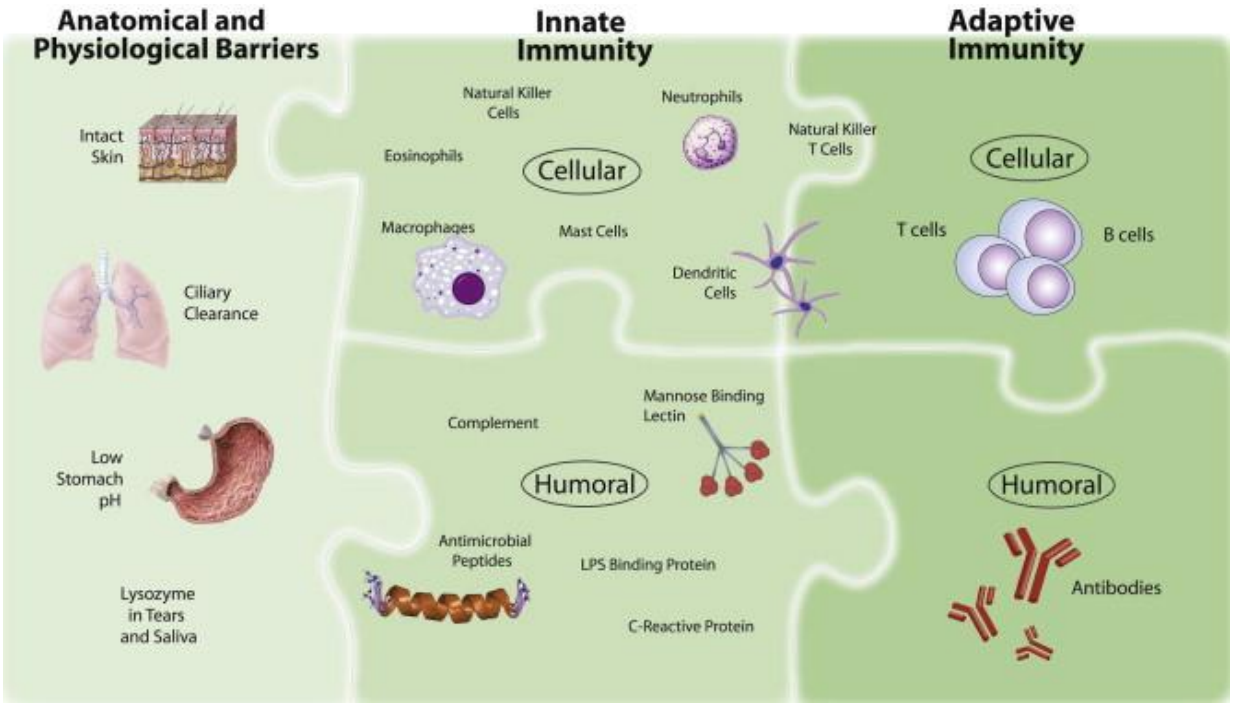


Figure 1.2: Levels of immunological response in mammals. Adapted from (Turvey and Broide, 2010b)

1.3.2 Innate immunity

The innate immune response provides organisms with a rapid means of combating pathogens: within minutes of exposure, a protective inflammatory response is generated. The innate immune system relies on a limited number of receptors referred to as pathogen recognition receptors (PRRs) to detect invading pathogens, but compensates for this limited number by targeting conserved microbial components that are shared by large numbers of pathogens (Turvey and Broide, 2010a). The innate immune system also plays a vital part in activating the adaptive immune response and is composed of both cellular and humoral elements. A number of both hematopoietic cells and non-hematopoietic cells are involved in the innate immune response, including the blood

macrophages, dendritic cells, mast cells, neutrophils, natural killer (NK) cell and NK T cells. Non-hematopoietic cells also play a role in innate immunity, including the skin and epithelial cells lining the respiratory, gastrointestinal, and genitourinary tracts. Innate immunity also has a humoral component that consists of well-characterised components, such as complement proteins, lipid polysaccharide (LPS) binding protein, C-reactive protein and antimicrobial peptides working to enhance the effect of these cellular defenses (Turvey and Broide, 2010a).

The major phagocytic cells of the innate immune system are neutrophils, macrophages and monocytes. Neutrophils destroy pathogens by engulfing them in intracellular vacuoles, where they are exposed to toxic molecules such as nitric oxide and degradative enzymes (Kennedy and DeLeo, 2009). Macrophages and monocytes are also highly phagocytic for microbes and particles that have been marked for clearance either through opsonisation by antibodies and/or complement (Bradford, 2012). Monocytes and macrophages also play vital roles in the adaptive immune response by ingesting microbial antigens, processing them into peptide fragments, and presenting them to T cells (Chaplin, 2010).

Dendritic cells are potent antigen-capture and -presenting cells that play a key role in the initiation and regulation of the adaptive immune response and they are the most potent antigen presenting cells (APCs). Dendritic cells (DCs) are present in most tissues of the body and concentrated in the secondary lymphoid organs (Lambrecht and Hammad, 2009). APCs express class I and II major histocompatibility complex (MHC) molecules that are required for recognition of processed antigen by the T cell receptor and accordingly

DCs are important in initiating the adaptive immune response by migrating from the site of infection to regional lymph nodes where they present pathogen-derived antigen to CD4+ T cells. Activated DCs express co-stimulatory molecules essential to T cell activation and can instruct the differentiation of naive CD4+ T cells into T helper cells, which can lead to clonal expansion and migration of T cells to B cell areas to assist with antibody production (Gallucci et al., 1999).

1.3.3 Adaptive immunity

In order to mount an effective immune response and unlike the innate immune response, the adaptive immune system is based on antigen exposure by both B cells and T cells (Turvey and Broide, 2010a). In contrast to the limited number of pathogen receptors used by the innate immune system, the adaptive immune system has an extremely diverse, randomly generated repertoire of receptors that are encoded by genetic elements that somatically rearrange to assemble antigen-binding molecules with great specificity for individual, unique foreign structures (Chaplin, 2010). Although the diversity in receptors enables adaptive immunity cells to recognise virtually any foreign antigen, this part of the immune system requires a relatively long period of antigen exposure before mounting a response due to clonal expansion of antigen-specific T and B cells. However, the benefit of this key feature is that it produces long-lived cells that persist as memory cells until a second exposure with their specific antigen, where they rapidly re-express effector functions and therefore it manifests an immune memory, which contributes to a more effective host response against specific pathogens following a second encounter (Bradford, 2012).

The major role of the T-cell in the adaptive immune response is to identify and destroy infected cells, and is responsible for cell-mediated immunity. T cells can recognise peptide fragments of antigens that have been taken up by APCs through phagocytosis. The immune system has evolved to enable T cells to recognize foreign antigens through recognizing self-components and foreign peptide molecules on the surface of APCs. MHC molecules are cell surface glycoproteins that bind peptide fragments of proteins that have either been synthesized within the cell i.e. endogenous antigens (class I MHC molecule) or ingested by the cell and proteolytically processed i.e. exogenous antigens (class II MHC molecules) (Chaplin, 2010).

1.3.4 Major histocompatibility complex class I and II

MHC molecules are cell surface proteins that are involved in adaptive immunity and histocompatibility. Class I MHC molecules are cell-surface heterodimers consisting of a highly polymorphic transmembrane polypeptide chains (Bradford, 2012). The main functions of class I MHC molecules are the presentation of peptide antigens to CD8+ cells and serving as inhibitory ligands for NK cell receptors (Li and Jevnikar, 2015). Class II MHC molecules are expressed on a number of immune cells, including B cells, DCs, monocytes and macrophages. They can bind a large repertoire of antigenic peptides which makes them most effective in presenting antigenic proteins of extracellular pathogens, such as most bacteria, parasites, and virus particles that have been released from infected cells to CD4+ T helper cells (Bradford, 2012).

Although MHC class I and class II are generally considered to have separate functions, these molecules most likely have a common evolutionary history. It is difficult to predict

how specificity is achieved in immune recognition because on one hand it is believed that only specific peptides from pathogens are recognized by T cells and on the other hand a single MHC molecule may in principle bind more than a million different peptides (Trowsdale and Knight, 2013).

1.3.5 Humoral Immune Response

Many of infectious bacteria multiply in the extracellular spaces of the host body, and most intracellular pathogens spread the infection by moving from cell to cell through the extracellular fluids. The extracellular spaces are protected by the humoral immune response. The humoral response involves the production of antibodies by B cells any thereby the destruction of extracellular microorganisms and prevent the spread of intracellular infections through extracellular fluids. The activation and differentiation of B cells into antibody-secreting plasma cells is initiated by the presence of antigens and facilitated by helper T cells. Antibodies can neutralize pathogens by preventing them from entering host cells through binding to specific surface molecules. Neutralization by antibodies is important in preventing bacterial toxins from entering cells (Janeway, 2005). Antibodies also protect against bacteria that multiply in the extracellular fluids mainly by facilitating engulfment of the pathogen by macrophages that are dedicated to destroying ingested bacteria. The process is called opsonisation and it involves the recognition of the antibodies that are bound to the pathogen surface antigens by phagocytic cells. Moreover, antibodies binding the surface of a pathogen can activate complement proteins that recruit phagocytic cells to the site of infection, and the terminal components

of this complement system can lyse certain microorganisms directly by forming pores in their membranes (Janeway, 2005).

1.3.6 Cellular immune response

Cellular immune response or cellular-mediated immunity (CMI) is an immune response to an antigen that involve phagocytosis, cytotoxic T-lymphocytes and the release of cytokines. CMI is mainly triggered by pathogens that survive in phagocytes and pathogens that infect non-phagocytic cells. It plays a major role in removing virus-infected cells, and contributes in the immune response against fungi, protozoans, and intracellular bacteria. It also plays a major in fighting cancer and transplant rejection (Janeway, 2005).

Mechanism of CMI involves the activation of antigen-specific cytotoxic T-lymphocytes that are capable of inducing apoptosis in cells displaying fragments of foreign antigen on their surface. Cells that are infected with intracellular bacteria or viruses and tumor display such foreign epitopes on their surfaces. Moreover, CMI involves the activation of macrophages and natural killer cells, enabling them to destroy intracellular pathogens as well as the stimulation of cytokines secretion, which influence the function of other cells involved in immune responses.

Generally, CMI occurs in three common phases: (a) Binding of the cytotoxic cells to the target cells, (b) Release of cytokines and (c) Lysis of the target cell. T-lymphocytes bind to their target cells that display class I MHC or through specific antibodies and receptors. Binding of the cytotoxic T-lymphocytes to the target “abnormal” cells is crucial for the induction of cell apoptosis. The union of two types of cells induces the T-lymphocyte to secrete cytokines, perforin and other cytolytic enzymes into the target cell. Perforin can

induce the formation of pores in the cell membrane which allow the movement of remain of cytotoxic components causing cell death and disintegration into small membrane bound vesicles (Janeway, 2005).

1.4 Immune response to TB and bTB

Although CMI plays a major role in the response to mycobacterial infection, other immune responses are also of importance and will be also discussed in terms of the type of response according to route of infection and the disease progression.

1.4.1 Intradermal route

The main cells involved in protective CMI are the macrophages and the T-lymphocytes as shown by early studies of the response of badgers infected experimentally with *M. bovis* by the intradermal route using the lymphocyte transformation test (LTT) with the *M. bovis* strain bacillus Calmette–Guérin (BCG) as the stimulating antigen, and also using the intradermal tuberculin test together with an enzyme-linked immunosorbent assay (ELISA) (Corner et al., 2011). There was a spectrum of immunological responses observed through different stages of the disease. The main observed pattern in the early stage there was a well-developed CMI response with increased LTT responses and a positive skin test, while in the late stage had an elevated levels if antibody production and CMI responses were depressed (Corner et al., 2011).

1.4.2 Endobronchial route

According to Lesellier *et al*., (2008) a number of immunological studies of badgers infected experimentally with *M. bovis* by the endobronchial route have demonstrated strong

responses starting 3 weeks post-infection to bovine purified protein derivative (PPD-B) *M. bovis*-specific single antigen (CFP-10). However, no CMI or antibody response was detected against ESAT-6, a target specific antigen in most other species (Corner et al., 2011). Moreover, a group of well-known immuno-stimulatory antigens in infected cattle (e.g. MPB70, Rv3019c, Rv3873, Rv3878 and Rv3879) were not recognized by experimentally-infected badgers (Lesellier et al., 2008) which suggests that infected badgers recognize a limited or different repertoire of antigens compared with other species (Corner et al., 2011).

1.4.3 Response to vaccination

Intradermal BCG vaccine can elicit lymphocyte proliferative responses *in vitro*. However, such responses were only obtained after repeated injections of BCG (Southey et al., 2001). Low CMI responses were also reported in other species including white-tailed deer and ferrets vaccinated with subcutaneous BCG injections, which suggests that badgers response to vaccination is mainly innate to lower the bacterial load before initiating a T lymphocyte mediated response (Corner et al., 2011). An immunological tolerance to environmental mycobacterial antigen, which occurs due to repeated exposure in badgers, can cause a delay in detecting adaptive immune responses until the bacterial load (vaccine doses) reach a minimum threshold level, i.e. a BCG dose dependant CMI responses in badgers (Lesellier et al., 2006).

1.4.4 Mycobacterium and Macrophages

Macrophages are the primary targets for mycobacterium infection and the host resistance is highly dependent on the macrophages' capacity to produce reactive nitrogen

and oxygen compounds, TNF and interleukins (Cassidy and Martineau, 2014). After being ingested by macrophages, mycobacteria have the ability to downregulate the expression of cathepsin G, a macrophage lysosomal protease with known antimicrobial properties, during early stages of mycobacterial infection (Rivera-Marrero et al., 2004)). This process is accompanied by an increase the expression of more acidic cathepsins, such as cathepsin B and D, which are weak proteases that can benefit the mycobacteria by contributing in tissue liquefaction and cavity formation in process called cathepsin switch (Cassidy and Martineau, 2014). Apoptosis plays an important role in the macrophages innate response against mycobacteria, and evasion of alveolar macrophage apoptosis is a virulence-associated phenotype in mycobacteria. Mycobacteria can delay the onset of apoptosis in infected macrophages through controlling intracellular prostaglandin E2 synthesis and thereby allowing more time for intracellular mycobacterial replication which infect more cells once the macrophages death occurs (Behar et al., 2010).

1.4.5 Cell-mediated immune response in mycobacterial infection

Recent studies have demonstrated that the adaptive cell-mediated response of cattle is quite similar to that of the human especially at the genomic level where genes encoding cytokines are known to play a role in regulating the immune responses in humans are present in cattle, including cytokines not found in mice (e.g., IL-26) (Waters et al., 2011). Cell-mediated immune responses are predominant in mycobacterial infections. Studies have shown that $CD4^+$ T cells produce Th1 cytokines, such as gamma interferon (IFN-γ), in response to mycobacterial antigens and that the cytolytic activity of $CD8^+$ cells toward infected macrophages is important (Kennedy et al., 2002). The Protective cell mediated

immunity against tuberculosis is dependent upon the interaction of T cells with infected macrophages. $CD4^+$ T-cell sub-population respond to infection principally through the production of cytokines such as gamma interferon (IFN-γ) which are considered to be involved in the activation of macrophages (Kennedy et al., 2002).

The release of IFN-γ is an important function of the CMI response to mycobacterial infection in human and cattle and it is widely used for tuberculosis diagnosis (Vordermeier et al., 2002). However, unlike rodents human and bovine immunity to the disease is less dependent on antigen specific IFN-γ activation of macrophages and more dependent on cytotoxic immune cells. The role of IL-21 and other key regulatory factors for maintenance and induction of IFN-γ (IL-12, IL-18, IL-23, and IL-27) and NK cell function in protective immunity to TB is an important avenue of investigation in the efforts to develop a vaccine for humans and cattle (Waters et al., 2011).

The sub-population of $CD8^+$ T cells is also involved in the production of IFN-γ in a low level during the course of infection (Serbina and Flynn, 1999). The important role of $CD8^+$ cells in early stages of infection is their ability to act as cytotoxic T lymphocytes which may not only be involved in the lysis of specific target cells, but may also release molecules, such as granulysin, which have been shown to directly kill mycobacteria (Stenger et al., 1998).

1.5 Current control strategies of bTB in badgers

1.5.1 Badger culling

The risk of bTB transmission from badgers to cattle can be reduced locally by culling, which has been implemented recently in the UK. Despite the evidence of a beneficial

reduction in bTB incidence rate in the neighbouring Republic of Ireland, after a similar culling programme was implemented (Olea-Popelka et al., 2009), there is still an ongoing debate that culling would not make a significant contribution to bTB control in the UK. This is based on a cost benefit analysis carried out by the Independent Scientific Group on Cattle (ISG) after evaluating data from the randomised badger culling trail (RBCT) (Robinson et al., 2012). Although the repeated culling across accessible land (Proactive badger culling) has reduced bTB incidence in cattle by 23% inside the targeted areas, it has also increased the incidence rate by 25% in the surrounding areas due to the disturbance in badger colonies and migration. Similarly, Reactive Culling, which involves culling in the areas close to farms where recent cases of bTB are reported, has also increased the incidence rate by 20% (ISG, 2007). Although badger culling appears to be an effective interim approach in reducing transmission of bTB to cattle in the cull area, it is unsustainable in the long term from both ethical and economic perspectives (Jenkins et al., 2010) as the public and scientific opposition to culling continues to grow.

1.5.2 Badger and cattle vaccination

Current research has provided evidence that vaccination can be used as an additional tool in any bTB control programme. However, there are still some areas of uncertainty about implementing a vaccination campaign, as more technical information on choosing a strategy of whether tackling the disease in high endemic areas should be started before low prevalence ones or whether or parenteral versus oral vaccination. There is also a limited understanding of the duration of immunity, and whether vaccination reduces *M. bovis* elimination from badgers that are already infected, and the excretion patterns of

vaccinated badgers that subsequently become infected. The most important point in the search for an efficient vaccination strategy, is determining the effects of badger vaccination on the incidence of bTB in cattle (Robinson et al., 2012).

The development of successful strategies to eradicate bovine tuberculosis in cattle and badgers requires a comprehensive knowledge of all of the epidemiological factors controlling the persistence of the infection in both host species, as well as the mode of transmission and a thorough understanding of the obstacles hindering success. The implementation of a local control strategy based on culling is limited by the number of badgers removed versus the total number of infected, and the proportion of the targeted area compared to the whole area affected. Research has yet to be completed in the field of vaccine efficiency and delivery. However, it is hoped that this will contribute to a more effective eradication of bovine tuberculosis on a national level (Gormley and Corner, 2013). Current bTB control strategies focus on vaccination of both badgers and cattle to achieve control and subsequently eradication of the disease. However, there are obstacles in the way of the large-scale use of injectable vaccine in badgers and the delivery mechanisms of oral vaccine in both badgers and cattle (Chambers et al., 2014). Although computer models have shown that maintaining a regular badger vaccination programme can theoretically reduce bTB incidence in cattle (Smith et al., 2012), there is no direct experimental evidence of reducing bTB transmission between badgers and cattle and vice versa (Chambers et al., 2014)

The optimisation of an oral bTB vaccine requires more research on the delivery options and exact mechanisms of its interaction with the badger's immune system. Without this knowledge the timescale and the production of effective and affordable vaccine remains uncertain especially when vaccination of cattle is currently prohibited under European legislations as the ultimate endpoint of using BCG vaccine in cattle without trade restrictions may not be achieved before 2023 (Chambers et al., 2014).

1.6 Contemporary experimental and computational approaches to understanding the transmission of TB

In order to understand disease mechanisms in a comprehensive manner, the combination of systems biology and bioinformatics is currently proving to be a very powerful research approach. Bovine tuberculosis infection can spread from cattle to humans and some domestic animals, which, in addition to the economic burden of the disease, has provided the impetus for the development of a new management strategy, as more traditional approaches seem to be making relatively slow progress. Understanding the fundamental disease mechanism together with the badger's ability to tolerate the infection on a sub-cellular and genomic level, might prove accessible to contemporary bioinformatics research tools.

In the last few years, bioinformatics has become fully embedded in the methodology of research in Biotechnology, through its focus on biological information management, data interpretation and future, including predictive methods. One of the recently developed

bioinformatics approaches is transcriptomics, which deals with the study of mRNA and non-coding RNAs generated by a cell or population of cells that share specific physiological functions. In contemporary, multidisciplinary research projects, global transcriptome analysis and profiling is commonly the first technology to be applied (McGettigan, 2013). Transcriptomics can generate information about which genes are expressed, at what level and can also provide information about different transcript isoforms

1.6.1 Gene transcription

"Gene expression, the process necessary to transmit information from genetically encoded information into functional processes within a living organism, is where it all starts. Basic principles are followed, which have proved so successful that they are used by prokaryotes and eukaryotes alike." (Persson and Mueller, 2015).

In the central dogma of molecular biology (DNA makes RNA makes protein) transcription is the step that precedes protein synthesis (translation), where DNA serves as template for the synthesis of RNA, catalysed by the enzyme RNA polymerase. In eukaryotes, the newly synthesised mRNA is subsequently released into the cytoplasm to be translated into protein by ribosomes (Latchman, 2005).

Generally, transcription takes place in three steps initiation, elongation and termination. Initiation occurs when the enzyme RNA polymerase binds to a promoter region of a gene. This binding signals the DNA to unwind allowing the enzyme to read the bases of a single DNA strand and create a complementary mRNA strand. Elongation refers to the addition

of nucleotides to the mRNA strand directed by the complementary bases with addition of Uracil (U) instead of Thymine (T) to the growing RNA strand. Termination of transcription occurs when RNA polymerase reaches a termination signal in the gene sequence. The mRNA molecule is then detaches from the complex and undergoes further maturation before release into the cytoplasm (Latchman, 2005). The stages are shown diagrammatically in Figure 1.3.

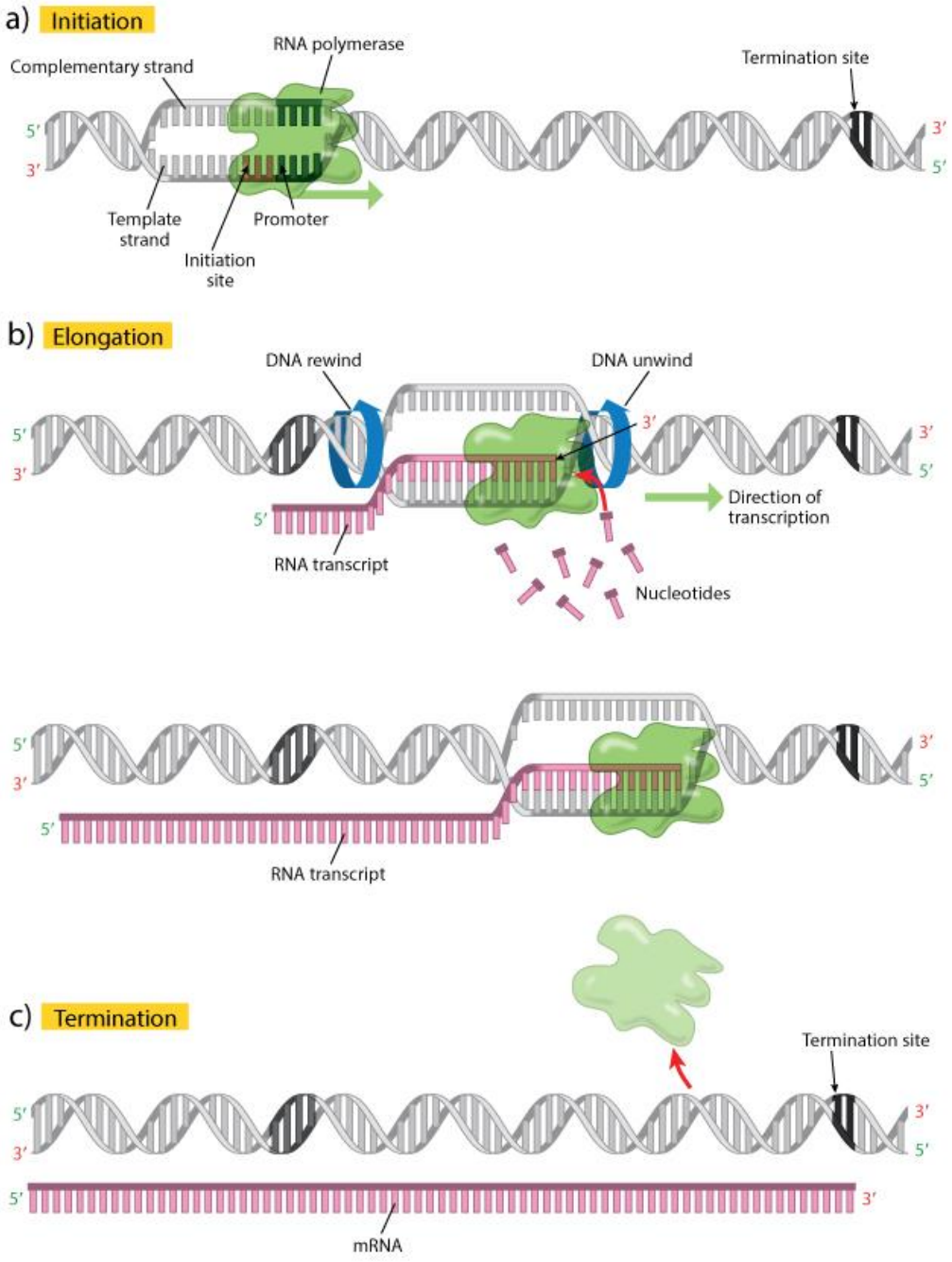


Figure 1.3: (A) “The transcription process is initiated when the enzyme RNA polymerase binds to a DNA template at a promoter sequence. (B) During the elongation process, the DNA double helix unwinds. RNA polymerase reads the template DNA strand and adds nucleotides to the three-prime (3’) end of a growing RNA transcript. (C) When RNA polymerase reaches a termination sequence on the DNA template strand, transcription is terminated and the mRNA transcript and RNA polymerase are released from the complex.” Adapted from Clancy, (2008).

In order for a eukaryotic mRNA molecule to be translated into an amino acid sequence it must first undergo additional processing, usually before it can be released into the cytoplasm. The processing may include mRNA editing, splicing and polyadenylation (Latchman, 2005). These processes allow not only mRNA maturation but also for a single gene to be used in the production of more than one protein. mRNA editing allows the change of some nucleotides in the sequence which as a consequence lead to the production of different forms of the translated protein. An example of this phenomenon is given the human APOB protein which has two different forms as a result of a premature stop signal during mRNA nucleotide editing (Severi and Conticello, 2015). Polyadenylation in a process by which a tail of adenine bases is added to the 3" end of mRNA sequence. Polyadenylation signals the end of mRNA, involved in mRNA export from the nucleus and protects mRNA from hydrolytic enzymes that might break it down in the cytoplasm. Splicing involves the removal of introns which are non-coding short regions of the RNA sequence that separate the coding exons from each other (Latchman, 2005). As a result of splicing the mature mRNA arises by re-joining the remaining exons (the coding regions) to form an mRNA transcript for subsequent translation into a protein.

1.6.1.1 Alternative mRNA splicing and its impact on protein diversity

The discovery of the discontinuity of eukaryotic genes with protein coding and non-coding segments was one of the most unanticipated findings in molecular biology, and later it has become clearer with advances in genome sequencing that splicing often parallels the complexity of an organism (Jacquier, 2009).

As an example of the complexity that results from the alternative splicing is the similarity of human and mouse genomes with almost the same number of genes, however, alternative pre-mRNA splicing occurs in more than 95% of human genes, compared with 63% of mouse genes **table 1.1**. This diversity significantly expands the form and function of the human proteome which can serve many regulatory functions, from sex determination and diversity of neuronal wiring in the fruit fly to determination of the physiological function of membrane-bound receptors in the mammalian nervous system (Lee and Rio, 2015).

	Human	Mouse
Genome size	3,300 MB	3,300 MB
Protein-coding genes	22,180	22,740
Multiexonic genes (percentage with 2+ isoforms)	21,144 (88%)	19,654 (63%)
Isoforms (average number per gene)	215,170 (3.4)	94,929 (2.4)
Average number of unique exons per gene(median)	33 (26)	22 (15)
Average number of unique introns per multiexonic gene (median)	28 (21)	19 (12)
Genes (all)	63,677	39,179
Isoforms (all) (average number per gene)	215,170 (3.4)	94,929 (2.4)

Table 1.1: Comparative genomics of splicing levels in human and mouse adapted from (Lee and Rio, 2015).

RNA splicing takes place in a large ribonucleoprotein structure known as the spliceosome which is composed of five small nuclear ribonucleoproteins that recognise and assemble on each intron to ultimately form a catalytically active spliceosome (Will and Luhrmann, 2011).

The human gene contains approximately eight exons and seven introns, producing an average of three or more alternatively spliced mRNA isoforms. Recent high-throughput sequencing studies indicate that the majority of human genes produce at least two alternative mRNA isoforms (Lee and Rio, 2015) **(Figure 1.4)**. Alternative splicing can arise as a result of several different mechanisms including RNA–protein interactions of splicing factors with regulatory sites termed silencers or enhancers, RNA–RNA base-pairing interactions, or chromatin-based effects that can change or determine splicing patterns. Errors in splicing and mutations in splice sites and splicing factors however rare may still can be linked to a number of diseases including cancer (Severi and Conticello, 2015).

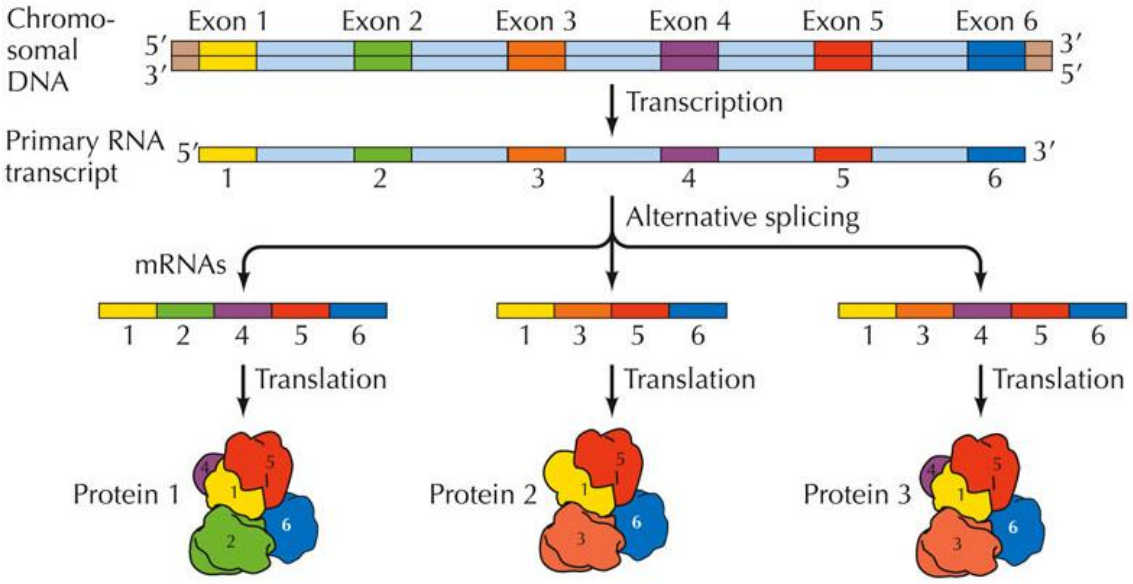


Figure 1.4 : Alternative splicing event and the resultant proteins (Cooper and Hausman, 2009).

1.6.1.2 Transcriptome complexity

Until recently, the description of a transcriptome was fundamentally limited to the characterization of the transcription products of known annotated genes i.e. mRNA, and stable non-coding RNAs such as tRNA and snRNA. However, the sequencing of entire eukaryotic genomes had paved the way to the development of more techniques to determine and catalogue their transcribed sequences and to study the regulation of transcription on a large-scale functional approaches (Jacquier, 2009).

These technologies have revealed that the transcription landscape in higher eukaryotes is more complex than previously had been predicted **(table 1.1)**, with a high proportion of transcripts originating from intergenic regions that were previously thought to be silent or in antisense to genes. The unanticipated level of complexity has led to the fact that the transcripts are not restricted to well-defined functional genes (Jacquier, 2009).

Unlike the genome, the transcriptome is more dynamic and variable depending on cell type and function. It also changes in accordance to physiological states such as growth or pathological conditions such as infections. Transcriptome profiling is an indicator of gene capability of generating different mRNA and protein isoforms through mRNA maturation and splicing and it can also indicate the magnitude of gene expression in terms of number of mRNA copies under different conditions.

1.6.2 Transcriptome analysis

The transcriptome is the entire set of transcripts present in a living cell or group of cells together with the relative abundance of each transcript at a defined developmental stage

or under a specific set of physiological conditions (Wang et al., 2009). Understanding the transcriptome is crucial for interpreting some of the functional elements of the genome and uncovering the molecular components of cells and tissues, and for understanding physiological and pathological development processes (Wang et al., 2009).

Amongst the more common aims of applied transcriptomics are: to classify all different groups of transcripts, such as mRNAs and non-coding RNAs in a given cell or tissue; to establish the transcriptional structure of genes, in terms of their start sites, 5' and 3' ends, splicing patterns and other post-transcriptional modifications and to measure the change in expression levels of each gene during physiological development and under different stress conditions such as disease (Wang et al., 2009).

Transcription is the first key regulatory step of gene expression that can fill the gap between genome expression and cell function. Furthermore, transcriptome analysis mirrors genome expression dynamics, as the transcription patterns are highly specific for each type of cells despite the fact that all cells of a given organism, share the same set of genes (Dong and Chen, 2013). Transcriptomics studies have widened our view field in understanding the structure and function of non-protein-coding RNA (ncRNA) and their rule in gene regulation (Mattick, 2005). As an example for the magnitude of ncRNA, over 93% of the human genome is transcribed into RNA (Carninci et al., 2005) and only 2% is from protein coding region (Green and Chakravarti, 2001). The development of next generation sequencing technology (NGS) has enhanced our understanding of RNA biology, and through that enhancement, the application of transcriptomics has been

expanded (Dong and Chen, 2013). Methodology improvement, particularly of the NGS technology, has led to a higher throughput and resolution level of transcriptome analysis studies, and has produced large amounts of data and correspondingly greater levels of biological information (Wang et al., 2009).

1.6.3 Advantages of transcriptome analysis

A variety of technologies have been developed to study and measure the transcriptome, including hybridization-based and sequence-based techniques. Hybridization-based approaches normally involve hybridizing fluorescently labelled cDNA with customised microarrays or commercially synthesized high-density oligo microarrays (Clark et al., 2002). However, the hybridization approaches have some limitations, which include: dependence upon existing knowledge about genomic sequence and a limited dynamic range of detection due to both background and saturation of signals. Moreover, comparing expression levels across different experiments is often difficult and can require complicated normalization methods (Okoniewski and Miller, 2006).

In contrast to microarray technologies, sequence-based methods can directly determine the cDNA sequence. However, most sequencing techniques are based on Sanger sequencing technology which is rather expensive, and has a number of limitations with short sequences where a significant proportion of the short tags cannot be distinctively mapped to the reference genome. Furthermore, only a portion of the transcript is analysed and isoforms are usually impossible to differentiate from each other. This

disadvantage limits the use of traditional sequencing technology in annotating the structure of transcriptomes (Wang et al., 2009).

In recent years, the development of high-efficiency DNA sequencing technology (NGS) has provided novel methods for both transcriptome mapping and quantification. This method, known as RNA-Seq (RNA sequencing), has significant advantages over existing approaches and is predicted to revolutionize the approach in which eukaryotic transcriptomes are analysed, for example, RNA-Seq has generated consistently high value data for mammalian transcriptome analysis (Mortazavi et al., 2008).

Transcriptome sequencing can provide an inexpensive, rapid approach to access gene sequences, gene expression patterns and provides a quantitative measure of gene expression in different species, regardless of the availability of a reference genome. These advantages can be attributed to the smaller size and the reduced complexity of the transcriptome compared to the whole genome. The successful application of RNA sequencing in combination with *de novo* transcriptome assembly, has facilitated the classification of new genes in a wide range of biochemical pathways. Despite the development in sequencing technologies, however, considerable challenges remain in the processing and analysis of transcriptome sequence data (Gongora-Castillo and Buell, 2013). Some of these challenges are discussed in the limitations of transcriptome analysis section.

1.6.4 Limitations of transcriptome analysis

The achievements in transcriptomics are largely attributed to the high volume of genome research and the desire to obtain functional information to add value to genome data. This has been matched by phenomenal levels of innovation in omics technologies, especially the improvements in NGS technology and its simultaneous reduction in cost. However, like other technology, NGS needs improvements to reduce the bias introduced by RNA amplification and library construction and also to reduce the cost for low input RNA-seq. Further, the optimisation of experimental design and bioinformatics analysis are both required for more efficient and accurate transcriptome characterisation (Dong and Chen, 2013).

Despite these limitations NGS techniques have emerged to be the most dominant genomics technology due to their cost and uses compared to Sanger sequencing (Morozova and Marra, 2008). The applications of RNA-seq techniques in the field of genomics have included genome annotation, gene expression profiling and ncRNA profiling (Morozova and Marra, 2008). NGS approaches have also been used in determining DNA sequences associated with epigenetic modifications of DNA and histones to profile DNA methylations, posttranslational modifications of histones, and nucleosome positions on a genome-wide scale (Callinan and Feinberg, 2006).

1.6.5 The principles of RNA-Seq technology

RNA-Seq technology is based on deep-sequencing methods, which generally involve the conversion of a population of RNA sequences to a library of cDNA fragments, with adaptors attached to one or both ends (Wang et al., 2009).

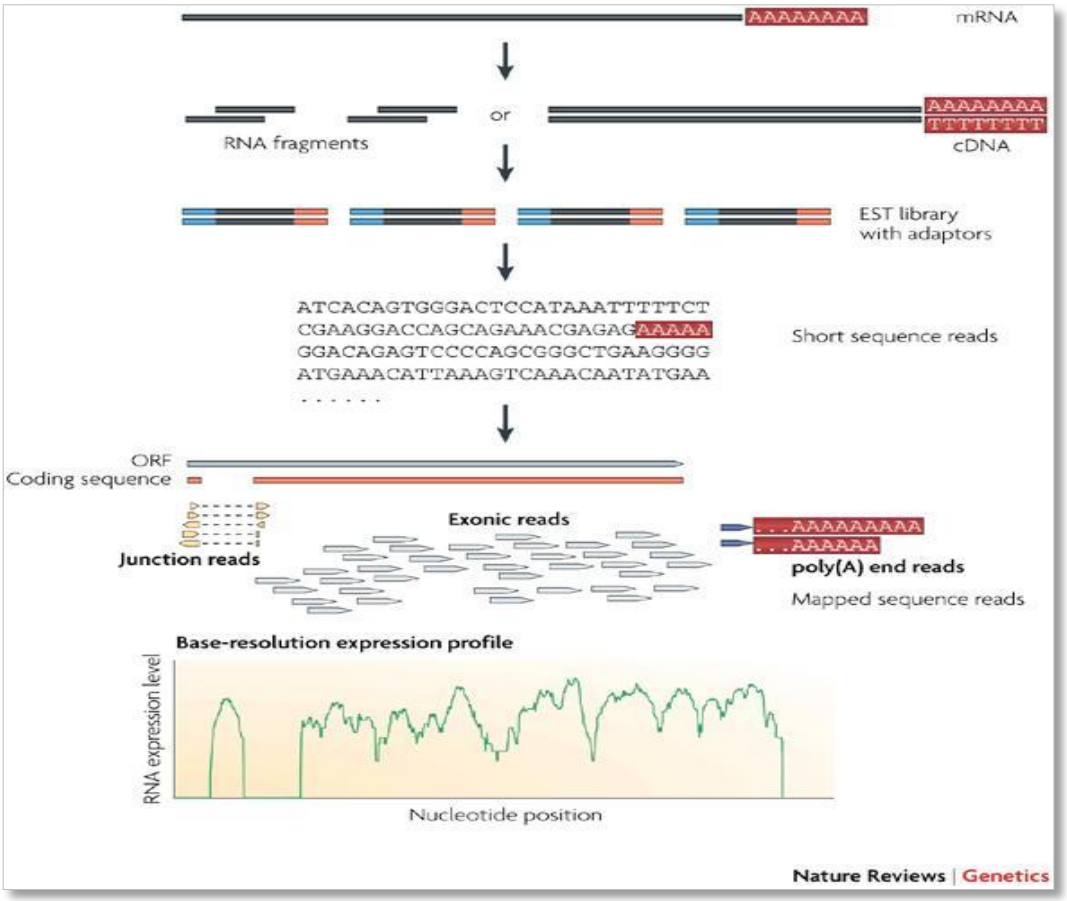


Figure 1.5: RNA-Sequencing, RNAs are converted into a library of cDNA fragments through either RNA fragmentation or DNA fragmentation. Sequencing adaptors (blue) are subsequently added to each cDNA fragment and a short sequence is obtained from each cDNA using high-throughput sequencing technology. The resulting sequence reads are aligned with the reference genome or transcriptome (Wang et al., 2009).

cDNA molecules are then sequenced (either with or without amplification) in a high-throughput approach to acquire short sequences from one end (single-end sequencing),

or both ends (pair-end sequencing). Depending on the DNA-sequencing technology used, the read size is typically in the range of 30-400 bp. After sequencing, the resulting cDNA fragments are either aligned to a reference genome or reference transcripts which enables characterisation of expression profiles **(Figure 1.5)**. *De novo* assembly is performed when the reference genomic sequence is not known to produce a genome-scale transcription map which can illustrate both the transcriptional structure and level of expression for each gene in the sequence (Wang et al., 2009).

1.6.6 Advantages of RNA-Seq technology

Even though RNA-Seq is an emerging technology that seems to be constantly undergoing efficiency and yield related improvement; it already presents a number of advantages over contemporary technologies. RNA-Seq technology is particularly suitable to apply in non-model organisms that do not have reference genomic sequence. The reason for that advantage is that RNA-Seq is not limited to aligning the transcripts to a known DNA sequence (Vera et al., 2008).

RNA-Seq technology can detect variations in the transcribed regions of the genome (Morin et al., 2008). The technology also has a wider dynamic range of expression levels over which transcripts can be detected, i.e. RNA-Seq has no upper quantification limit when compared to DNA microarrays which has a limited sensitivity to gene expressed at low or very high levels (Mortazavi et al., 2008). RNA-Seq has also shown a high accuracy in quantifying expression levels as measured using quantitative PCR (Nagalakshmi et al., 2008).

Further advantages of RNA-Seq include: high technical and biological reproducibility levels, very low background signal interference and the technique does not include cloning steps and therefore less RNA samples are required (Cloonan et al., 2008) and when compared to Sanger EST sequencing or DNA microarrays, RNA-Seq offers better single-base resolution and gene expression levels at much lower cost (Wang et al., 2009).

1.7 Transcriptome assembly and annotation

In bioinformatics, genome annotation is the term that describes two distinctive processes. Structural gene annotation is the process of identifying genes intron–exon structures. Whereas, functional genome annotation is the process of attaching meta-data such as gene ontology terms to structural annotation.

1.7.1 Assembly

A successful annotation is achieved only after comprehensive transcriptome assembly has been completed. There are several statistical summaries that can be used to benchmark the contiguity and completeness of an RNA-Seq experiment. The most important statistic is Scaffold and Contig N50 but other assembly statistics are useful as well (e.g. the average gap size of a scaffold and the average number of gaps per scaffold (Yandell and Ence, 2012). Generally, the existing genomic data are standard, draft assemblies that meet minimum requirements for submission to public databases. However, a high-quality draft assembly (which is 90% complete) is still a better aim for annotation (Chain et al., 2009). A transcriptome assembly with a N50 scaffold length that is gene-size (half of the assembled readings are complete genes) is an accepted annotation target. These 50% of the assembled genes together with the remaining fragments will give a decent resource

for subsequent analysis (Ye et al., 2011). It is recommended to perform more shotgun sequencing when the assembly is incomplete or if the N50 scaffold length is too short as achieving high-quality assembly reads became more efficient in recent genome projects (Husemann and Stoye, 2010).

1.7.2 Annotation

The “Annotation pipeline” is a general term that is used to refer to the different tools and programs that assemble and compute data, and use it to create the primary genome annotation. The process is intrinsically complex and it mainly focuses on annotating protein-coding genes. Annotation pipelines vary in their working details, magnitude and accuracy but still have the common core features. The pipeline is commonly divided into two phases. In the first computational phase the expressed sequence tags (ESTs) and protein-coding genes are aligned to the reference genome or evidence-based gene predictions are generated. In the second annotation phase the collected data are organised and combined into gene annotations (Yandell and Ence, 2012).

1.8 The Mealworm *(Tenebrio molitor)* as a prospective model organism for schools

1.8.1 Definition of model organisms

“Model organisms are usually defined as non-human species that are extensively studied in order to understand a range of biological phenomena that might not be easily researched in advanced organisms, with the hope that data, models and theories generated will be applicable to other organisms, particularly those that are in some way

more complex than the original" (Leonelli and Ankeny, 2013). This definition focuses on the use of model organisms for the primary purpose of research and development, and they play a key role in both drug discovery and testing, in the Pharmaceutical Industry.

One of the earliest systematic experiments, using a model organism, is the use of *Pisum sativum*, by Gregor Mendel during his pioneering work on the "rules" of inherited characteristics (Smýkal et al., 2016). Later, *Drosophila* species were pivotal in the investigation of the harmful effects of radiation at the cellular and genetic level (Lamb and Smith, 1969). Famously, the Guinea Pig Latin name, is cited as the proxy for the test organism in drug trials in particular. However, guinea pigs are used much less frequently today compared with mice and zebra fish, for example. A particular model organism is chosen for experimental work when it closely matches the system under investigation, often in man. However, in Mendel's case, pea plants provided him with a phenotype (plant height) that would unequivocally "report" on the relationship between genotype and phenotype, which is a relationship that he believed would apply across all eukaryotes. Today, plant biologists generally use *Arabidopsis thaliana* as the model organism of choice for exploring the fundamentals of plant physiology.

With the widespread introduction of genome sequencing during the 1990s, traditional model organisms became some of the earliest targets for genome sequencing. Yeast (*Saccharomyces cerevisiae*) *and E.coli* were sequenced alongside the release of the first draft of the human genome sequence, with *Drosophila* and *Arabidopsis* following soon after. It is true to say that the genome sequence will be available for almost all model organism employed as proxies for a more complex organisms.

1.8.2 Characteristics of model organisms

The selection of specific living organisms as experimental models is usually determined by the remit of the experimental investigation. With this in mind, the "simpler" model organisms, such as *E.coli*, yeast and fruit flies have been mainly utilised for the investigation of fundamental biological phenomena, such as metabolism, eukaryotic cell division and embryonic developmental respectively (see for example, Rosenblueth and Wiener, 1945). These organisms grow and divide rapidly, and large populations can be obtained both rapidly and economically, assuming the relevant laboratory facilities are to hand. As the questions become more complex or perhaps more "human", especially in the evaluation of drugs, it is not uncommon to utilise primates for the most critical comparative studies. Of course, over the last 20 years in particular, efforts have been made to keep the use of animals in research to a minimum on ethical grounds.

1.8.3 The use of model organisms in the teaching of Science

Laboratory classes were popularised by John Dewey at the turn of the 19th century and have become embedded in the curriculum of high schools worldwide who teach subjects like Chemistry, Physics and Biology. It is primarily in Biology that living organisms have been employed for traditional dissection classes, simple microscopy and for demonstrations of evolution. However, as part of a project designed to bring a more contemporary flavour to school Science: one which captures the existing curriculum and introduces genome biology, this Chapter is aimed at developing the darkling beetle, *Tenebrio molitor* for such a purpose. In considering which organism to choose, some of the above criteria were incorporated. However, additional criteria were considered and

these are discussed below prior to the experimental section, in which the meal worm (which will be used interchangeably with T. molitor) is developed as a focus for educational experimental biology. One final important point is that this insect should be seen not as a fully understood organism, but one for which many questions still remain, some of which can be asked in suitably designed, schools-led research projects.

Schools, unlike professional research laboratories, are generally “closed for business” for between 10 and 20 weeks per year, imposing major logistical constraints on the choice of a model organism. Plants need regular watering and a regulated light source, bacteria require specialised growth, sterilisation and disposal facilities, mice are relatively expensive, relatively slow to breed, fish are relatively high maintenance and experimental use of both of the latter organisms requires ethical approval. In contrast, insects like drosophila and locusts offer many advantages for school science, except that they are prone to flying around: making their management more “hands-on”. Flightless insects, such as beetles are amongst some of the most diverse species on Earth, and moreover, in most Northern Hemisphere countries fishermen use insect larvae as bait. Hence, inexpensive supplies of larvae such as the meal worm are readily available from pet shops or angling suppliers (angling is the most popular individual sport in the UK) across the UK (and the USA, for example). Importantly, meal worm larvae require minimal “life support”. A small tray of meal worm will complete the life cycle in around 4 weeks, provided simply with bedding and a few pieces of cut fruit or vegetables. Since they are flightless, meal worm larvae and adults are extremely low maintenance, and can be kept at a wide range of room temperatures with little impact on viability. These features are

combined with the more typical, favourable characteristics associated with a model organism including small adult size, a relatively short life span, rapid development, availability of supply and tractability (Bolker, 1995).

1.8.4 The need for another model organism

Since the completion of the human genome project in 2000, the significantly reduced cost and growth in the number of genome sequencing facilities worldwide, has led to a significant increase in the number of model organisms, rather than an experimental consolidation of existing organisms. The argument made here for developing the meal worm (or indeed beetles in general), is primarily made in order to provide access to a living organism that straddles school science and frontline research in a manageable, low-cost way.

Genetic screening using model organisms has proved to be a powerful and valuable approach for over 100 years (in fact the “sanctity” of man possibly influenced the early pioneers of experimental anatomy, including Aristotle). Many remarkable studies using model organism have helped to understand the fundamental principles of vertebrate evolution, which includes resistance to infection (Abnave et al., 2014). Moreover, the main recommendation for model organism work, as stated by Jenner and Wills (2007) is to broaden phylogenetic sampling, to minimize bias in the sample of characteristics that are represented by the chosen models. Our general understanding of phenotypic evolution is therefore better served by deliberately choosing new models with traits that enable them to provide independent illumination of evolutionary developmental biology conceptual themes (Jenner and Wills, 2007).

1.8.5 The yellow mealworm beetle, *(Tenebrio molitor),* as a model organism

In considering the development of the mealworm beetle (*Tenebrio molitor*) as a new model organism, it is important to provide some wider contextual information. *T. molitor* is one of the largest known beetles that feeds on plant products and stored food (typically in a warehouse setting), causing considerable damage to such produce, in terms of total mass, quality and nutritional value. The beetle eats and then contaminates food produce with bodily waste and dead larvae (Siemianowska et al., 2013). Each female *T. molitor* lays around 300-500 eggs: each egg hatches in 4- 17 days, growing to a 3mm long larva **(Figure 1.6)** that reaches an average full size of (25-35mm long and 200mg in weight). The full grown larvae are used in both animal feed and increasingly as a dietary component of the human food chain in some parts of the world (Aguilar-Miranda et al., 2002). Mealworm larvae have a relatively long life span, compared with other beetles: under optimum moisture and temperature, they can survive for up to around six months, followed by 5-6 days of a dormant pupal stage (Siemianowska et al., 2013). More than 20,000 species of Tenebrionidae have been identified and described in different parts of the world (Liu and Wang, 2014). The dried, live larvae of the mealworm beetle are most commonly used as food for birds and other domestic pets (Barker et al., 1998). However, there is considerable interest in developing the meal worm for wider human consumption.

Figure 1.6: *Tenebrio molitor* larvae (Kupferschmidt, 2015).

1.8.6 Research applications of *T. molitor*

A growing number of researchers have used *T. molitor* as an experimental system for studies in biology, biochemistry, evolution, immunology and physiology, owing to its relatively large size, ease of handling, and orthodox genetics (Lee et al., 2015). However, despite this work, a coherent physiological and genetic "profile" of this beetle is lacking (Liu and Wang, 2014), and such information could form the basis of a school-wide programme of directed research.

T. molitor has a particular advantage as a candidate model system for the study of pathogens, since the larvae can be maintained at body temperatures between 25 °C and 37 °C (de Souza et al., 2015). On the other hand, other model organism such as *Drosophila melanogaster* (fruit fly) and *Caenorhabditis elegans* are unable to tolerate this temperature range (Desalermos et al., 2012). Direct inoculation by injecting the pathogen

directly into the *T. molitor* larvae (facilitated by the limited movement of the larvae) make this insect much more favourable than *C. elegans* (Merkx-Jacques et al., 2013). Similar arguments can be made for beetles such as the wax worm. It seems clear that beetles such as *T. molitor* offer some distinct experimental advantages in contemporary biological research: the area of infectious disease research and antimicrobial screening is one clear example.

1.8.6.1 *T. molitor* in immunological research

This thesis provides an introduction to the analysis of those genes involved in the transmission of bTB in cattle (Chapters 1-4), Despite the lack of comparative information regarding the infected versus uninfected badger transcriptomes, an analysis of the immune related genes, reveals a level of complexity that would benefit from investigation in a model system. It is the long term aim of this work, that the meal worm may provide insights into immune responses to infectious disease that may help us understand events in more complex organisms including cattle and man.

T. molitor larvae are known to exhibit a population density-dependant immune response, where lower mortality rates are observed at higher larval density compared to isolated larvae (Barnes and Siva-Jothy, 2000). Studies were also conducted to test the innate immunity responses and its "energetic cost" to *T. molitor* (Moret and Siva-Jothy, 2003) and also to examine the adaptive immune response to repeated exposure to pathogens. In this work, Moret and Siva-Jothy (2003) showed that larvae were found to produce a sustained, antimicrobial response despite that fact that invertebrates do not have acquired immunity. The limited availability of genetic data on *T. molitor*, make it

extremely difficult to provide a molecular framework for these and similar observations, even by drawing on comparative data from *Drosophila melanogaster*, again arguing in favour of the use of a related species such as *T. molitor* larvae, from which significant amounts of haemolymph can be extracted, to help elucidate the mechanisms underlying pathogenic microbe recognition (Park et al., 2010). To date this area of research has revealed the presence of a number of pattern recognition proteins, serine proteases, serpins and antimicrobial peptides and examined how these molecules affect innate immunity (Park et al., 2011).

The difference in gene expression in response to bacterial intoxication by *Bacillus thuringiensis* Cry3Aa Protoxin led to sequencing of *T. molitor* transcriptome, which represents the largest genetic sequence dataset of the organism to date (Oppert et al., 2012a). Other experiments tested the survival of antimicrobial peptides-resistant *Staphylococcus aureus* in response to *T. molitor* as an insect model which led to the suggestion that increased survival of antimicrobial peptides-resistant bacteria almost certainly poses problems to immune-compromised hosts (Dobson et al., 2014).

1.9 T. molitor genomic DNA sequencing

The development of efficient, large-scale and relatively inexpensive DNA sequencing technology has begun to impact significantly on Biological research in the last decade in particular, making DNA sequencing routine in many fields, including forensics, agriculture different aspects of medical and non-medical research (Rosenstein, 2014).

Obtaining the complete sequence of the *T. molitor* genome will greatly enhance our ability to associate genes and mutations with traits and diseases similarly to the process of developing the rat as a model organism and to take full advantage of the wealth of physiological variation among strains and mutants using a map of the genetic variation (Lindblad-Toh, 2004).

The recently published *T. molitor* transcriptome (Oppert et al., 2012b) and mitochondrial genome (Liu and Wang, 2014) offer an opportunity for developing the mealworm as a model organism. *T. molitor* is a valuable model host but its full potential and tolerance has yet to be studied, and the sequencing of *T. molitor* genome could allow the production of different mutants, and contribute to studies on host response to infection (de Souza et al., 2015) just as it has done wlewhere in Biological research.

1.9.1 *T. Molitor* genome sequencing via Illumina platform

“Illumina" sequencing is dependent on solid phase amplification of random DNA fragments followed by sequencing-by-synthesis, by adding fluorescent dNTPs (Fox et al., 2009) to facilitate detection. All “Illumina” sequencing workflow routines are composed of four basic steps that begin with DNA sample preparation, cluster generation, sequencing and data analysis.

1.9.1.1 Sample preparation

There are number of different procedures available for the preparation of genome samples by fragmentation (e.g. nebulization, hydro-shear or sonication). All preparation methods incorporate the addition of adaptors to the ends of the DNA or cDNA fragments,

through reduced cycle amplification **(Figure 1.7)**, additional motifs are introduced, such as the sequence binding site, indices and regions complementary to the flow cell oligonucleotide adapters (Mardis and McCombie, 2017).

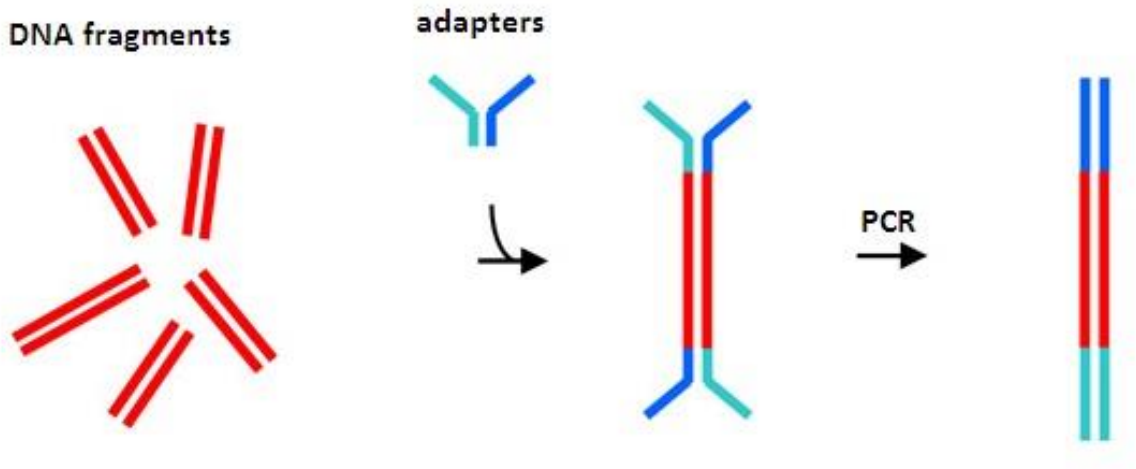


Figure 1.7: DNA fragments are generated by shearing and joined to a pair of oligonucleotides in a forked adapter configuration. The ligated products are amplified using two oligonucleotide primers, resulting in double-stranded blunt-ended fragments with a different adapter sequence on either end (adapted from Bentley et al., 2008)).

1.9.1.2 Clustering

Clustering or cluster generation is a process where each fragment molecule is isothermally amplified. The technology involves attachment of a short DNA fragment to a solid surface called a flow cell. The flow cell is a glass slide with lanes. Each lane is a channel coated with a lawn composed of two types of oligo-nucleotides. Hybridisation is enabled by the first of the two types of oligo-nucleotides on the surface where oligo-nucleotides are complementary to the adapter region on the fragment strands. The attached DNA fragments are PCR amplified to create clusters at a very high density (>10 million DNA clusters per lane) on the surface of the transparent sequencing flow cell. (Fox

et al., 2009). A polymerase creates complements of the hybridised fragments then the double stranded molecules are denatured and the original templates are washed away. The strands are clonally amplified through bridge amplification **(Figure 1.8)**.

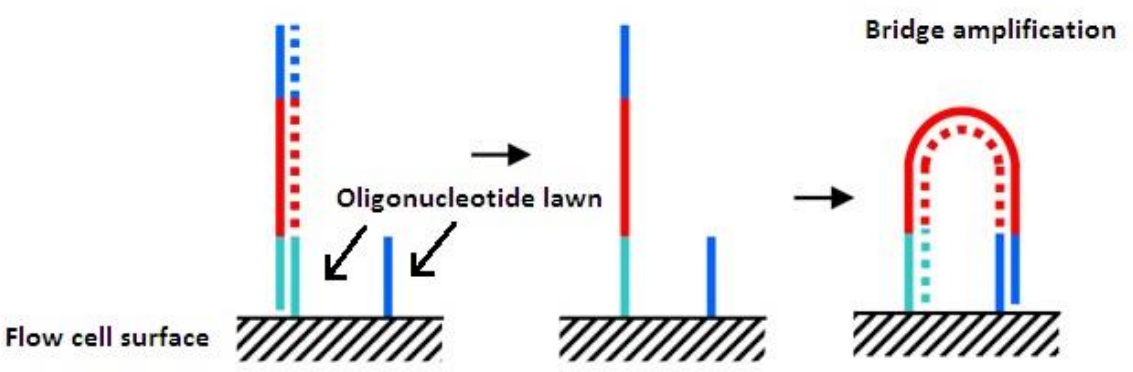


Figure 1.8: Formation of clonal single molecule array. DNA fragments are denatured and single strands are annealed to complementary oligonucleotides on the flow cell surface. A new strand (dotted) is copied from the original strand and the original strand is then removed by denaturation. The adapter sequence is annealed to a new surface bound complementary oligonucleotide, forming a bridge and generating a new site for synthesis of a second strand (shown dotted). (Adapted from Bentley et al., 2008)).

In this process the strands fold over and the adapter regions hybridise to the second type of oligonucleotides on the flow cells. DNA Polymerases generate the complementary strands generating double stranded bridges. Each bridge is then denatured resulting in two single stranded copies of the molecule that are tethered to the flow cell. The process is repeated over and over and occurs simultaneously for millions of clusters resulting in clonal amplification of all the fragments. After bridge amplification, the reverse strands are cleaved and washed off leaving only the forward strands. The 3" prime ends are blocked to prevent unwanted priming. "Solid-phase amplification can produce 100–200

million spatially separated template clusters (Illumina/Solexa), providing free ends to which a universal sequencing primer can be hybridized to initiate the sequencing reaction (Metzker, 2010).

1.9.1.3 Sequencing

Sequencing begins with the extension of the first sequencing primer to produce the first read. With each cycle fluorescently tagged nucleotides compete for the addition to the growing chain (Toh et al., 2017). However, only one nucleotide is incorporated based on the sequence of the template. After the addition of each nucleotide, the clusters are excited by a light source and a characteristic fluorescent signal is emitted. This proprietary process is called “sequencing by synthesis” and the number of cycles determines the length of the reads. The emission wave length along with the signal intensity determine the base call. For a giving cluster all identical strands are read simultaneously and hundreds of millions of clusters are sequenced in a massively parallel process. The sequencing process is repeated for multiple cycles. Amplified fragments representing a cluster are then sequenced and imaged with each reaction step. The system uses dNTPs containing fluorescently labelled 3” reversible terminators, each emitting a different fluorescence signal (Fox et al., 2009).

1.9.1.4 Data analysis

The sequencing process generates millions of short reads representing all the original DNA fragments. Sequences from pooled sample libraries are separated according to the unique indices introduced during sample preparation stage. For each sample reads with similar stretches of base calls are locally clustered and forward and reversed reads are

paired creating contiguous sequences. The contiguous sequences are then either aligned back to a reference genome for variant identification or independently assembled into a genome.

1.10 Aims and Objectives

Understanding the badger's immunological response to bTB might hold the answer to a new treatment or immunisation strategy for cattle and other affected domestic animals as well as zoonosis. It might also help in exploring the human disease and help in developing more potent vaccine or genetic therapy.

In that prospect, this research is more of an exploratory nature to delve deeper into the immunological components that appear to give the badger a unique tolerance to bTB.

In that sense, the main aim of the study is to find out the immunological components of the badger's transcriptome and their phylogenetic relations to other mammals, which may hold some answers of the evolution of the badger's immune system.

The extensive research in immunological responses of mealworm to pathological stress, the flexibility of its use and handling in the laboratory and low cost of maintenance, as well as the known transcriptome and mitochondrial DNA, have made the mealworm an appropriate candidate for genomic DNA sequencing to facilitate its future applications as a model organism for schools other educational institutions as well as a candidate for studying the basic immunological responses for bacterial infections including bTB.

In that prospect, sequencing *T. molitor* genome will set a platform for expanding the current spectrum of model organisms for better understanding of disease and infection as well as to reduce the margin of limitations and error.

The main objectives of this research are:

- To analyse the badger's transcriptome and identify its components.
- To draw a phylogenetic relationship between the badger and other mammals using sequence alignment and phylogenetic tree construction tools.
- To identify the immunity related transcripts involved in bTB pathogenesis using KEGG pathway for a tuberculosis as a reference
- To extract high quality genomic DNA from *T. molitor*, and from this obtain the nucleotide sequence for its genome.

Chapter II

2 Materials and Methods

This chapter is divided into two main sections: the first one describes the bioinformatics approaches for badger transcriptome analysis and the second is dedicated to molecular biology methods relating to the mealworm genomic DNA extraction.

2.1 Materials and methods for transcriptome analysis

The transcriptome data were obtained following isolation of peripheral blood cells at the Animal Health and Veterinary Laboratories Agency (AHVLA) in London. The total RNA was extracted from healthy bTB free animals. All subsequent RNA-seq work was contracted out to the Beijing Genomics Institute (BGI), who also provided a preliminary annotation document. Further analysis was performed using the following databases and software:

2.1.1 NCBI Basic Local Alignment Search Tool (BLAST+)

BLAST (Altschul et al., 1997) is one of the more popular software choices for searching and aligning biological sequence data. BLAST takes a nucleotide or protein sequence as input and compares it with a database of nucleotide or protein sequences respectively. BLAST can translate nucleotide sequences as needed; therefore, BLAST can search a nucleotide query against a protein database or a protein query against a nucleotide database. BLAST uses heuristics to accelerate searches. BLAST also provides statistics that estimate the likelihood of a match occurring by chance (Boratyn et al., 2013). BLAST search was used in conjunction with other tools and platforms (Clustal omega and Galaxy) to draw alignments of several immunity transcripts derived from the badger, and to construct phylogenetic trees for both badger and mealworm.

2.1.2 Galaxy platform

Galaxy platform is a web-based environment in which users can perform genome related computational analyses and importantly, all search details and parameters are automatically tracked for later inspection (Cock et al., 2015). Galaxy was used in data analysis in order to convert the transcriptome sequence data (FASTA) files into a searchable database (makeblastdb), to facilitate the subsequent BLAST+ searches. Galaxy utilises BLAST+ command-line applications (Camacho et al., 2009) through a user-friendly graphical interface.

2.1.3 Kyoto Encyclopaedia of Genes and Genomes (KEGG)

KEGG is a knowledge base for the systematic analysis of gene function, linking genomic information with higher order functional information. The genomic information is stored in the GENES database, which is a collection of gene catalogues for all completely sequenced genomes, together with some partial genomes with up-to-date annotation of gene functions. The higher order functional information is stored in the PATHWAY database, which contains graphical representations of cellular processes, such as metabolism, membrane transport, signal transduction and cell cycle. The PATHWAY database is supplemented by a set of ortholog group tables for the information about conserved sub-pathways (pathway motifs), which are often encoded by positionally coupled genes on the chromosome and which are especially useful in predicting gene

functions (Kanehisa and Goto, 2000). KEGG was use to extract the KEGG pathway for bTB and the list of genes involved in tuberculosis pathogenesis.

2.1.4 Gene Ontology (GO)

The goal of the Gene Ontology Consortium is to produce a dynamic, controlled vocabulary that can be applied to all eukaryotes even as knowledge of gene and protein roles in cells is accumulating and changing. (Ashburner et al., 2000).

2.1.5 Clusters of Orthologous Groups database (COG)

COGs are groups of three or more orthologue genes, which means that they are direct evolutionary counter-parts and are considered to be part of an 'ancient conserved domain'. A COG is defined as three or more proteins from the genomes of distant species that are more similar to each other than to any other protein within the individual genome. COGs can be used to predict the function of homologous proteins in poorly studied species and can also be used to track the evolutionary divergence from a common ancestor, hence providing a powerful tool for functional annotation of uncharacterized proteins (Tatusov et al., 2000).

Although they were not directly used in the analysis, GO and COG helped in grouping and classification of transcripts involved in immunity and tuberculosis pathway according to their GO terms and COG functional groups.

2.1.6 Immunome Knowledge Base (IKB)

IKB is a dedicated resource for immunological information. IKB contains information for human immunome genes and proteins, phylogenetic trees and evolutionary information

for immunome orthologs, ortholog groups for metazoan immunome, and variation data on genomic, transcriptomic and proteomic level. IKB integrates three previous databases, Immunome, ImmTree and ImmunomeBase with additional data (Ortutay and Vihinen, 2009).

IKB was used as a reference database to extract all the immunity related transcripts form the badger transcriptome annotation data files based on the 983 genes classified as immunity-related in the IKB database.

2.1.7 Clustal Omega

"Clustal Omega is a multiple sequence alignment tool, which can align (in a virtual sense) a large number of protein sequences quickly while delivering accurate and robust alignments. The accuracy of the package on smaller test cases is similar to that of the high-quality alignment software. Clustal Omega also has powerful features for adding sequences to and exploiting information in existing alignments, making use of the vast amount of precomputed information in public databases" (Sievers et al., 2011). Clustal Omega was accessed via EMBL-EBI web service interface (McWilliam et al., 2013). Along with BLAST+ Clustal Omega was used to align sequences and construct phylogenetic trees. Figure 2.1 shows the parameters used in BLAST+ and Clustal Omega searches via the EMBL-EBI website.

In addition to the above databases and platforms, others were used more infrequently to

Blast+:	Clustal Omega:
Database: UniProtKB mammals/ Program: blastp/ Matrix: BLOSUM62/ Expectation value threshold: 1e-5/ Dropoff: 0/ Gap open: -1/ Gap extend: -1/ Filter: F/ Sequence range: START-END/ Gapalign: true/ Composition-based statistics: F/ Align views: 0/ Translation table: -1/ Sequence type:protein	Program: clustalo/ Version: 1.2.4/ Output guide tree: false/ Output distance matrix: false/ Dealign input sequences: false/ mBed-like clustering guide tree: true/ mBed-like clustering iteration: true/ Number of iterations: 0/ Maximum guide tree iterations: -1/ Maximum HMM iterations: -1/ Output alignment format: Clustal/ Output order: aligned/ Sequence Type: protein

Figure 2.1: Search and alignment parameters used in BLAST+ and Clustal Omega

address specific features and characteristics (e.g. full sequence) or functions of some transcripts e.g. UniProt (2017)

2.2 Materials and methods for mealworm genomic DNA extraction

This section describes chemical, enzymes and other materials used in conducting the experiments. A description of the protocols and standard molecular biology methods used is also included.

2.2.1 Chemicals

All chemicals used were of molecular biology grade

Material	**Provider**
5x DNA Loading Dye	QIAGEN
Agarose	Bioline
DNA Hyperladder I	Bioline
ISOLATE Genomic DNA Mini Kit	Bioline

Table 2.1: Chemicals

2.2.2 Solutions

Buffer	Composition
TAE (Tris-Acetate-EDTA) buffer 50X stock)	750 ml of 2.67 M Tris base 57.1 ml of 17.4 M glacial acetic acid 100 ml of 0.5 M EDTA (pH 8.0) Adjust the solution to a final volume of 1 L
Ethidium Bromide Stock solution	25.4 mM (10 mg/ml of deionized distilled Water)

Table 2.2: Solutions

2.3 Basic Molecular Biology protocols

2.3.1 Genomic DNA extraction

Mealworms were obtained from a pet shop and stored in a -80˚C freezer for at least 24 hours before DNA extraction in order to obtain a more consistent powder when grinded. Approximately 50-60 mg of the powdered worm used for the extraction using ISOLATE Genomic DNA Mini Kit according to the following steps:

1- The powdered sample was placed in a 1.5 ml tube and 400µl of lysis buffer and 25 µl of proteinase were added and mixed immediately by vortexing then incubated at 50˚C in an incubator fitted with a rocking platform for continuous mixing for about 3.5 to 4 hours until the sample was completely dispersed.

2- The mixture was centrifuged at 10000x g (12000 rpm) for 1 minute and the supernatant was transferred to another 1.5 ml tube. 200 µl of binding buffer were added and mixed immediately by vortexing for 15 seconds.

3- The sample was then transferred to a spin column with a 2 ml collection tube and centrifuged at 10000x g (12000 rpm) for 2 minutes. The filtrate was discarded and then column was washed twice with 700 µl Wash buffer with a centrifugation at 10000x g (12000 rpm) for 1 minutes each time.

4- To remove all traces of ethanol the sample was centrifuged for 2 minutes at a maximum speed and the collection tube was discarded. The column was placed in a 1.5 elution tube and 200 µl of Elution buffer were added directly to the spin column membrane and incubated at room temperature for 1 minute. The eluted DNA was collected by centrifugation at 6000x g (8000 rpm) for 1 minute.

2.3.2 Confirmation of the genomic DNA extraction by agarose gel electrophoresis

1 g agarose was dissolved in 100 ml TAE buffer by carefully boiling in a microwave oven When the solution had cooled down to about 60˚C , 5 µl of ethidium bromide stock (10 mg/ml) was added to make a final concentration of 0.5 µg/ml. The solution was stirred to disperse the ethidium bromide, and then poured into the gel template. The comb was placed at one side of the gel (about 5-10 mm from the end of the gel) and the gel left until it solidified then it was placed in the tank and TAE buffer was added to just cover the agarose.

After loading the DNA samples and DNA marker electrophoresis was performed at 100 volts for one hour. Gel-separated DNA was stained with ethidium bromide and visualised under ultraviolet light.

Chapter III

3 Transcriptome assembly and annotation

This chapter explores the transcriptome derived from peripheral blood cells taken from live badgers as a first step towards identifying its molecular composition and its coding potential. Transcriptome sequencing, assembly and annotation was performed by BGI (China) and a description of sequencing platform and assembly pipeline and annotation process will be described in addition to a quality control assessment and an analysis of the male/female difference in expression and it significance in the light of bTB infection . This is the first step in building a platform for future genome sequencing. Here is a description of the results of the whole transcriptome assembly and its annotation using different databases.

3.1 Experimental pipeline and Sequencing platform

3.1.1 Experimental (mRNA isolation and sequencing) pipeline as described by BGI:

The steps for the experimental pipeline are shown schematically in **Figure 3.1**. After the total RNA extraction and DNase I treatment, magnetic beads coupled to oligo-dT were used to enrich for polyA-mRNA from the total RNA. mRNA was then mixed with the fragmentation buffer and broken down into short fragments. cDNA was synthesized using the mRNA fragments as templates. Short fragments were purified and resolved with EB buffer for end reparation and single nucleotide A (adenine) addition. After this, the short fragments were connected with adapters. The suitable fragments were selected for PCR amplification as templates. During the quality control (QC) steps, Agilent 2100 Bioanaylzer and ABI StepOnePlus Real-Time PCR System were used in quantification and qualification

of the sample library. Finally, the library was submitted to sequencing via Illumina HiSeq™ 2000. The steps involved in sequence determination using the Illumina platform will be described in more detail in chapter I.

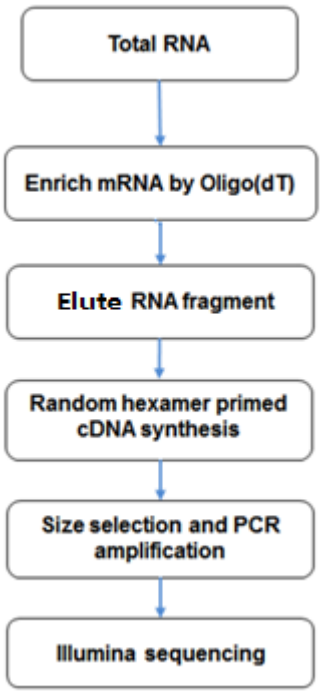


Figure 3.1: Experimental pipeline which mainly involves: mRNA isolation, cDNA synthesis and Illumina sequencing (BGI).

3.1.2 Assembly pipeline RNA sequencing and *de novo* assembly

The fluorescent image data output from Illumina sequencing instruments is transformed by base calling into sequence data, which are referred to as raw data or raw reads and stored in a text (FASTQ) format. The filtering of raw reads was required, as sequences produced from sequencing instruments of this type contain non-clean reads, which contain adapters, and sequences of unknown origin or low quality. If left, these data will negatively impact on downstream bioinformatics analysis.

Therefore, non-clean raw reads were discarded:

- If they are still attached to adaptors.
- If the reads have unknown nucleotides in the sequence larger than 5% of the total length.
- If the percentage of low quality bases in the sequence is more than 20% of the total length.

3.1.2.1 Assembly:

De novo assembly of the Transcriptome was carried out with Trinity, a short reads assembling program (Grabherr et al., 2011). Trinity combines three independent software modules: Inchworm, Chrysalis, and Butterfly, applied sequentially to process large volumes of RNA-seq reads. Trinity partitions the sequence data into many individual de Bruijn graphs, each representing the transcriptional complexity at a given gene or locus, and then processes each graph independently to extract full-length splicing isoforms and to tease apart transcripts derived from paralogous genes. Briefly, the process works as follows:

Inchworm Assembles the RNA-seq data into the unique sequences of transcripts, often generating full-length transcripts for a dominant isoform, but then reports only the unique portions of alternatively spliced transcripts.

Chrysalis Clusters the Inchworm Contigs together and constructs complete de Bruijn graphs for each cluster. Each cluster represents the full transcriptional complexity for a

given gene (or sets of genes that share sequences in common). Chrysalis then partitions the full read set among these disjoint graphs.

Butterfly then processes the individual graphs in parallel, tracing the paths that reads and pairs of reads take within the graph, ultimately reporting full-length transcripts for alternatively spliced isoforms, teasing apart transcripts that correspond to paralogous genes.

3.1.3 Output statistics and bioinformatics

Originally, two samples of blood were taken from one **male (identifier Q828)** and one **female (identifier Q381)** badger. The sequencing process generated over 118 million raw reads from each sample. After removing the adaptors and non-clean reads, a total of 108,193,588 and 105,901,706 high quality, "clean" reads were obtained for the two samples Q828, Q381 respectively **(Table 3.1)**

Samples	Total raw reads	Total clean reads	Total clean nucleotides	GC %
Q828	121,818,354	108,193,588	9,737,422,920	54.09%
Q381	118,836,262	105,901,706	9,531,153,540	51.88%

Table 3.1: Total raw and clean reads generated by sequencing

After assembly, the total number of sequences produced for annotation was 238,295 transcripts, with a total nucleotide length of 305,341,024 bp and an average length of 1281 nucleotides: more than 50% of transcripts are over 2720 nucleotide long **(Table 3.2)**. In the table, a **contig** is defined as any sequence produced by two or more overlapping

reads, and a **unigene** as a hypothetical gene represented by a cluster of similar transcripts that are thought to be isoforms in the *de novo* transcriptome assembly.

	Sample	Total Number	Total Length(nt)	Mean Length(nt)	N50
Contig	Q828	320,860	90,046,338	281	382
	Q381	433,526	115,082,292	265	334
Unigene	Q828	215,117	247,701,860	1151	2636
	Q381	274,052	289,728,809	1057	2625
	All	238,295	305,341,024	1281	2720

Table 3.2: Number of aligned sequences for each sample including contigs, unigenes and the mean length of each group

3.2 Distribution of sequence lengths

The length of transcripts **(Figure 3.2)** was distributed between a minimum of 300bp and 3000bp. Over 50% of the transcripts were in the length range of 300bp to 600bp while 13% were ≥ 3000 and 36 were in the range between > 600bp and < 3000bp.

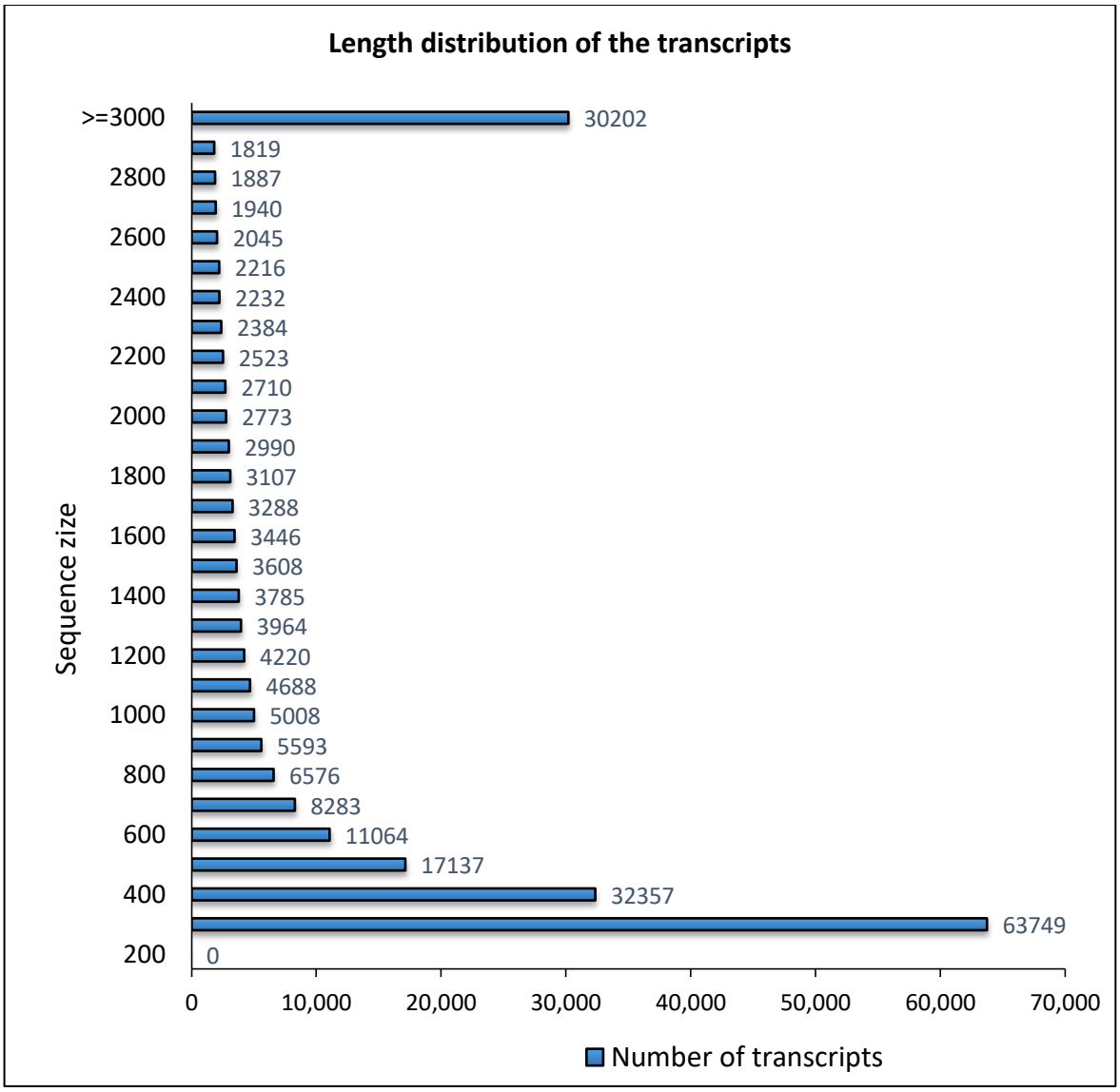


Figure 3.2: The distribution of reads per sequence length.

3.3 Transcriptome Similarity, Functional and Pathway annotation

3.3.1 Transcriptome annotation

Sequence similarity, functional annotation and pathway similarity search for the assembled transcripts resulted in annotation of (39%) 95,245 transcripts with significant

similarity with their correspondent genes in published-online databases. 38% (90,719) of the total transcripts were annotated by sequence similarity using a BLAST search of a Non-redundant (nr) database, 27% (65,384) in Swiss-Prot, 29% (69,924) were matched with 259 KEGG pathways, 23% (57,098) to COG function and 13% (32074) to GO terms **(Table 3.3).**

Database	NR	NT	Swiss-Prot	KEGG	COG	GO
Number of genes annotated	90,719	16,330	65,384	69,924	57,098	32,074
Percentage of genes annotated	38.07%	17.15%	27.44%	29.34%	23.96%	13.45%

Table 3.3: Number and percentage of transcripts annotated in each database

3.3.2 NCBI annotation

In total, 95,245 unigenes were annotated in all databases and 95% (90,719) of all annotated unigenes were identified by sequence similarity alignment in the NCBI non-redundant database. These transcripts were statistically grouped in terms of e-value distribution, the similarity of the sequences to their closest match from the database and also the top matching hits in terms of organismal similarity.

3.3.3 E-value distribution:

The E-value threshold was < 1E-5 and according to E-value distribution the alignments were divided into: Higher end homology from 0 to <1.0E-100 (39%), moderate 1.0E-100

to 1.0E-30 (30%) and lower end homology ranged from 1.0E-30 to 1.0E-5 above which any alignment is insignificant. A detailed distribution of the transcripts according to E-value is shown in the histogram **(Figure. 3.3)**

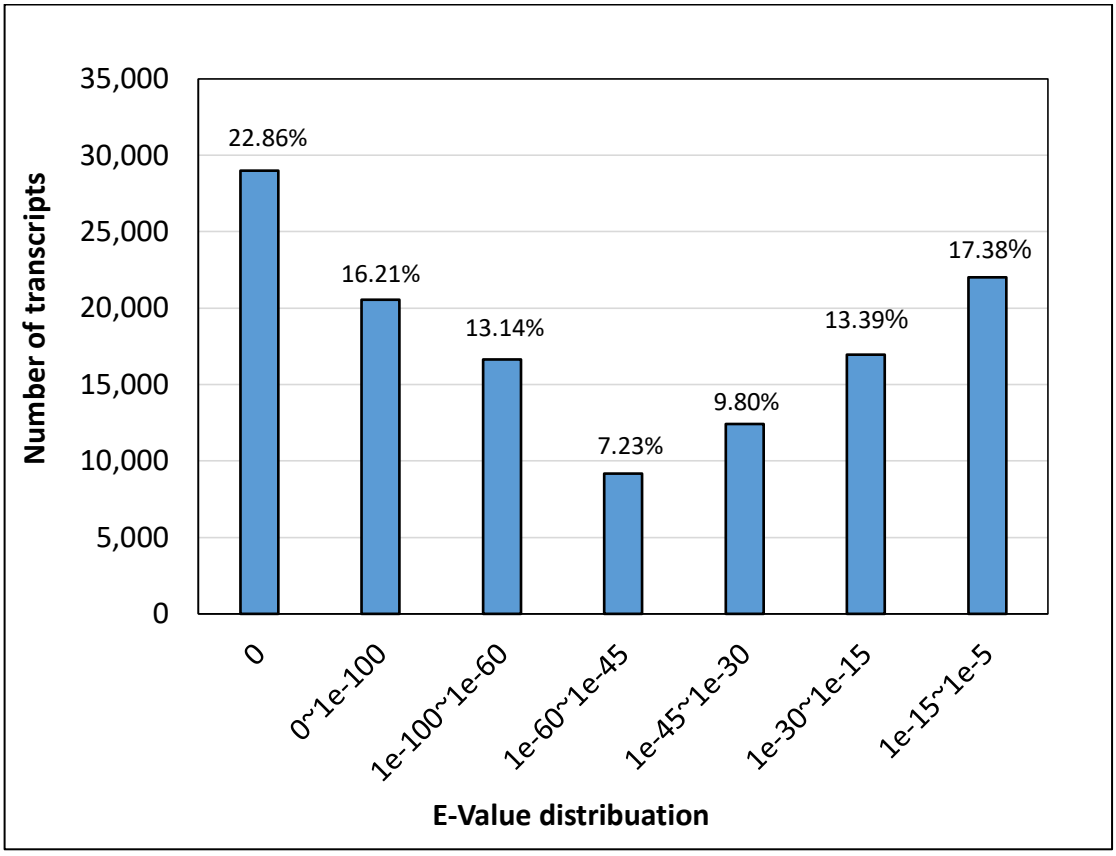


Figure 3.3: E-value distribution of the annotated transcripts in NCBI non-redundant database

3.3.4 Species similarity distribution

Over 70% of the transcripts matching hits belong to five mammals including *Ailuropoda melanoleuca* (the giant panda) with 25%, *Mustela putorius furo* (the ferret) 20%, *Canis lupus familiaris* (the dog) 16%, *Homo sapiens* (human) 4.4%, *Sus scrofa* (the wild boar) 2.3% and *Bos taurus* (the cow) 2.3%. Other species combined represented 29% of the total annotation **(Figure 3.4).**

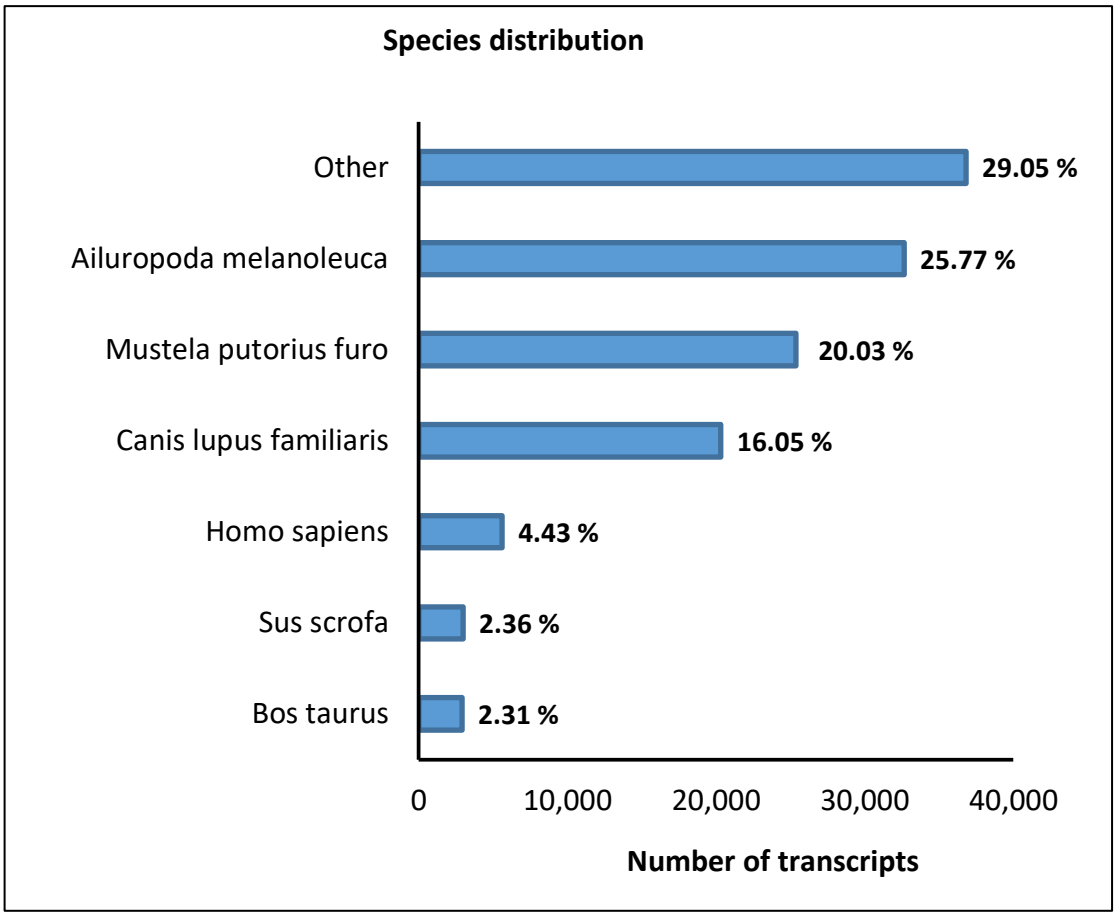


Figure 3.4: Species similarity distribution in NCBI non-redundant database

3.3.5 Sequence similarity distribution

Over 70% of the alignment achieved similarity greater than 80% between the badger blood transcripts and the corresponding genes in the database. Whereas less than 30% achieved a similarity range between 17% and 80% as shown in **Figure 3.5.**

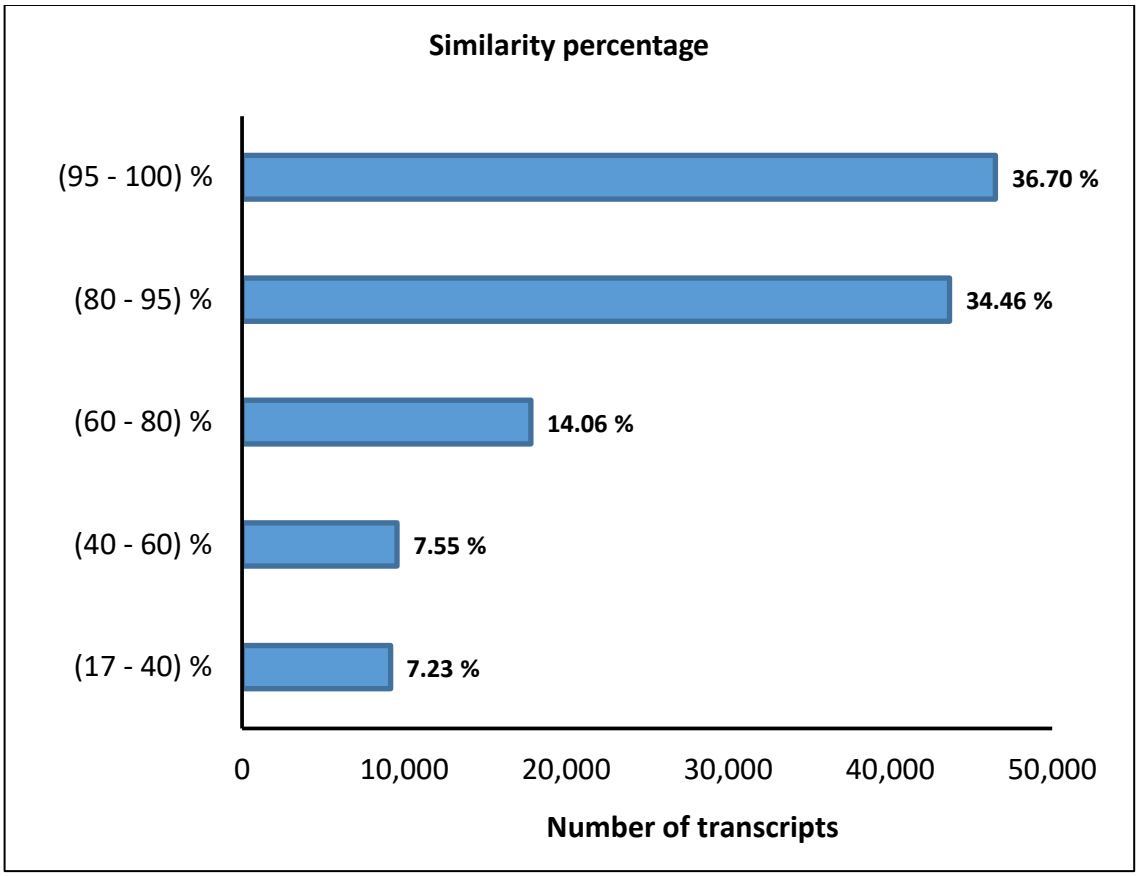


Figure 3.5: Sequence similarity distribution in NCBI non-redundant database

3.3.6 COG Function

Searching the COG database for domain based alignments led to annotation of 57,089 reads and revealed their cellular function classification as shown in **(Fig. 3.6)**. After setting the E-value threshold at 1.0E-5, the homologically significant matches in the COG database were clustered into functional classes, and the cluster (Translation, ribosomal structure and biogenesis) has the highest representation with 13.7% of the total transcripts and (General function) cluster has 13.3% of the transcripts representation. The lowest clusters represented are (Nuclear structure, Extracellular structures and RNA processing and modification) with less than 0.6% combined **(Figure 3.6)**.

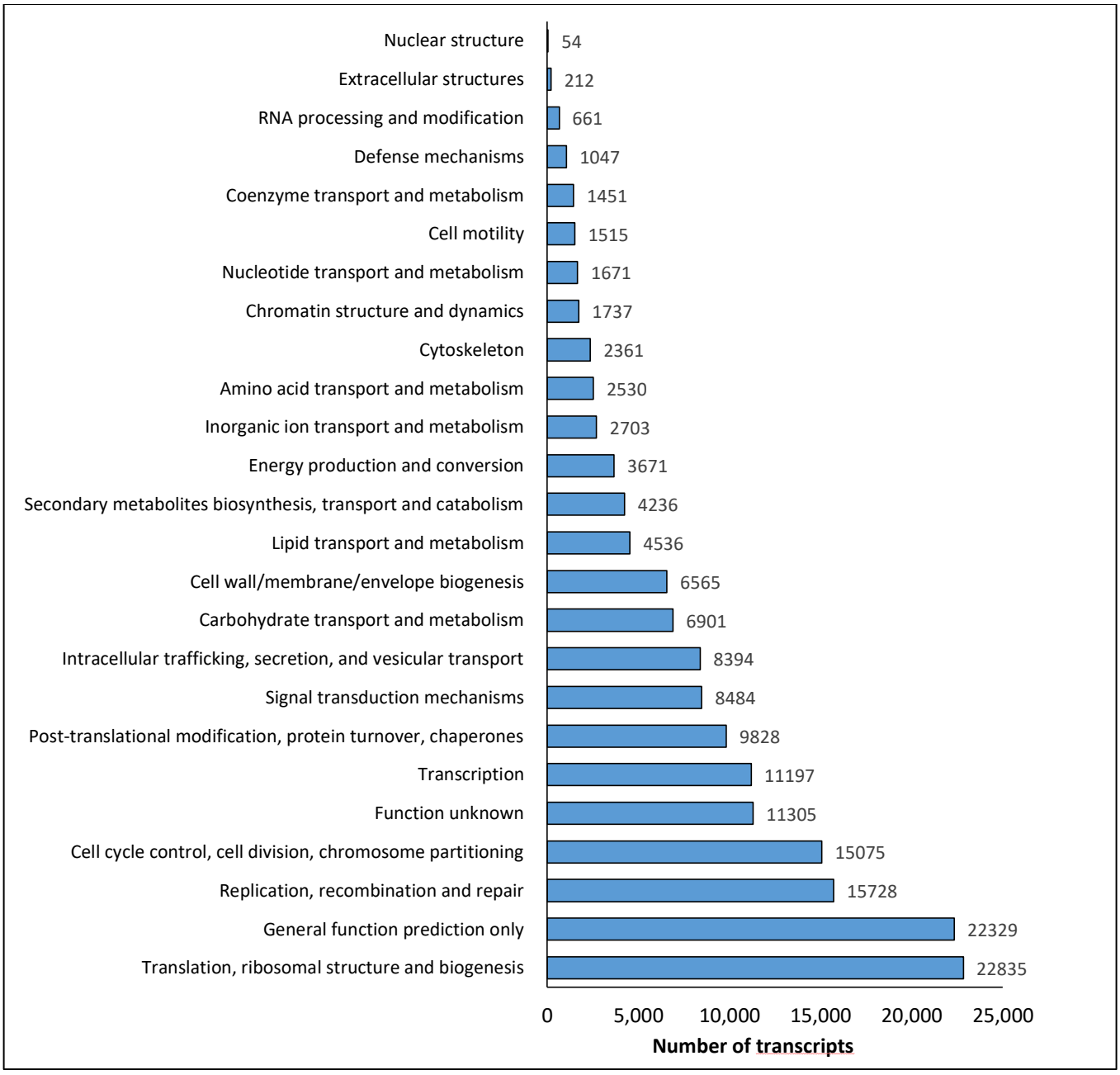


Figure 3.6: COG functional annotation

3.3.7 Gene Ontology

According to the sequence alignment and homology, the total number of GO terms that correspond to all unigenes is 1,028,340. In general, most of the GO term annotations represent biological processes with 499,918 (about 48.6%) GO terms, whereas cellular

components represent 37.2% with 382,212 GO terms, and lastly molecular function with 14.2% with 146,210 GO terms **(Figure 3.7)**. For all GO combined: cell, cell part, binding and cellular process have the highest representation with over 70,000 GO term each while virion, virion part, protein tag, chemorepellent activity, carbon utilization and nutrient reservoir activity have the lowest number of GO terms with less than 10.

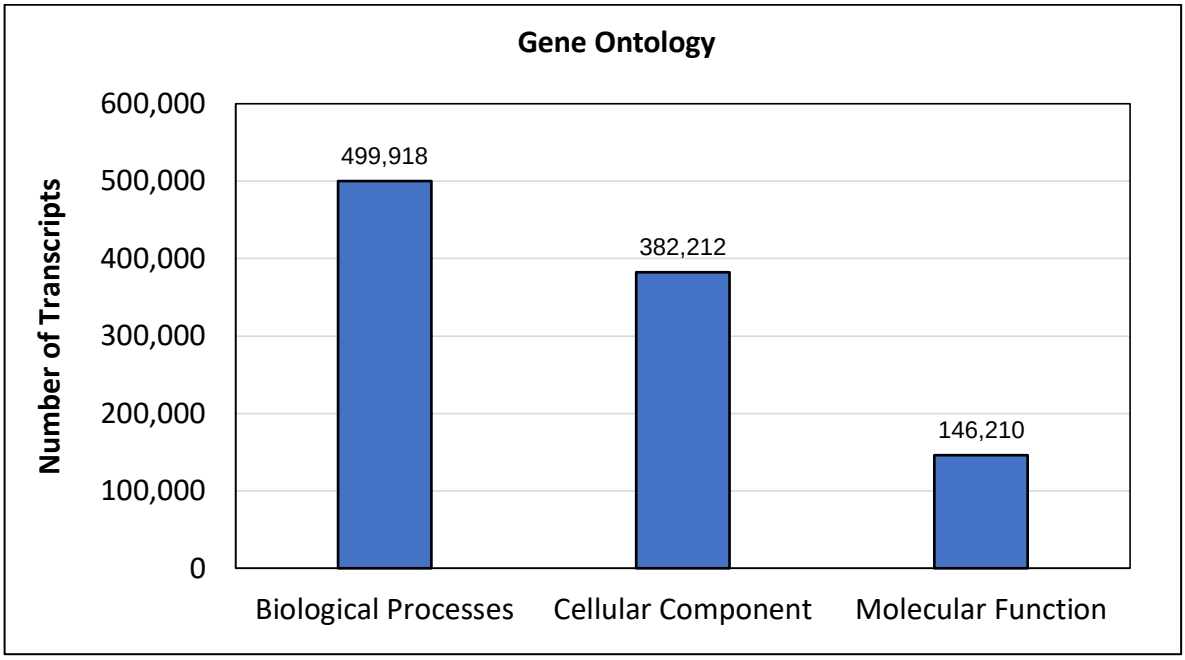


Figure 3.7: Number of transcripts representing each GO function

The distribution of transcripts according to the three groups of GO terms is shown in details in the **figure 3.8.**

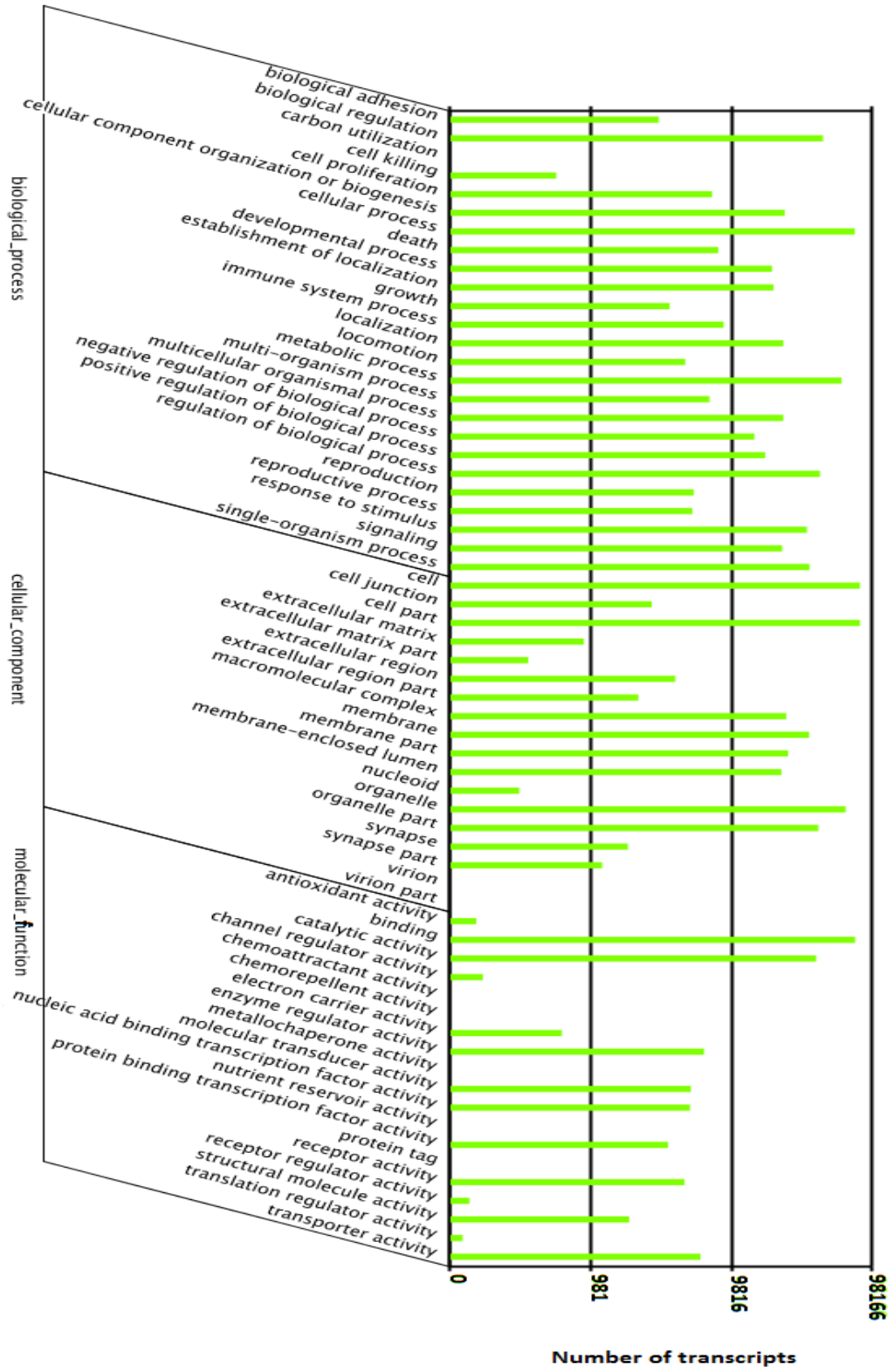


Figure 3.8: Distribution of transcripts on different GO terms.

3.3.8 Pathway annotation KEGG

The mapping of assembled transcripts with pathway annotation to KEGG database resulted in assigning those reads to 259 pathways. Among the mostly represented pathways Metabolic Pathways have 9.97% (9930), Focal adhesion 7.61% (7582), Amoebiasis 7.17% (7136), ECM-receptor interaction 5.94% (5918), Protein digestion and absorption 5.65% (5623), RNA transport 4.17% (4158), Regulation of actin cytoskeleton 4.02% (4001), Pathways in cancer 3.88% (3862) and Herpes simplex infection 3.22% (3209).

In addition to the major pathways represented by the transcripts in the **figure 3.9** there are other 239 pathways shown in the **Appendix** including some infectious diseases pathways like influenza A, measles, toxoplasmosis and tuberculosis.

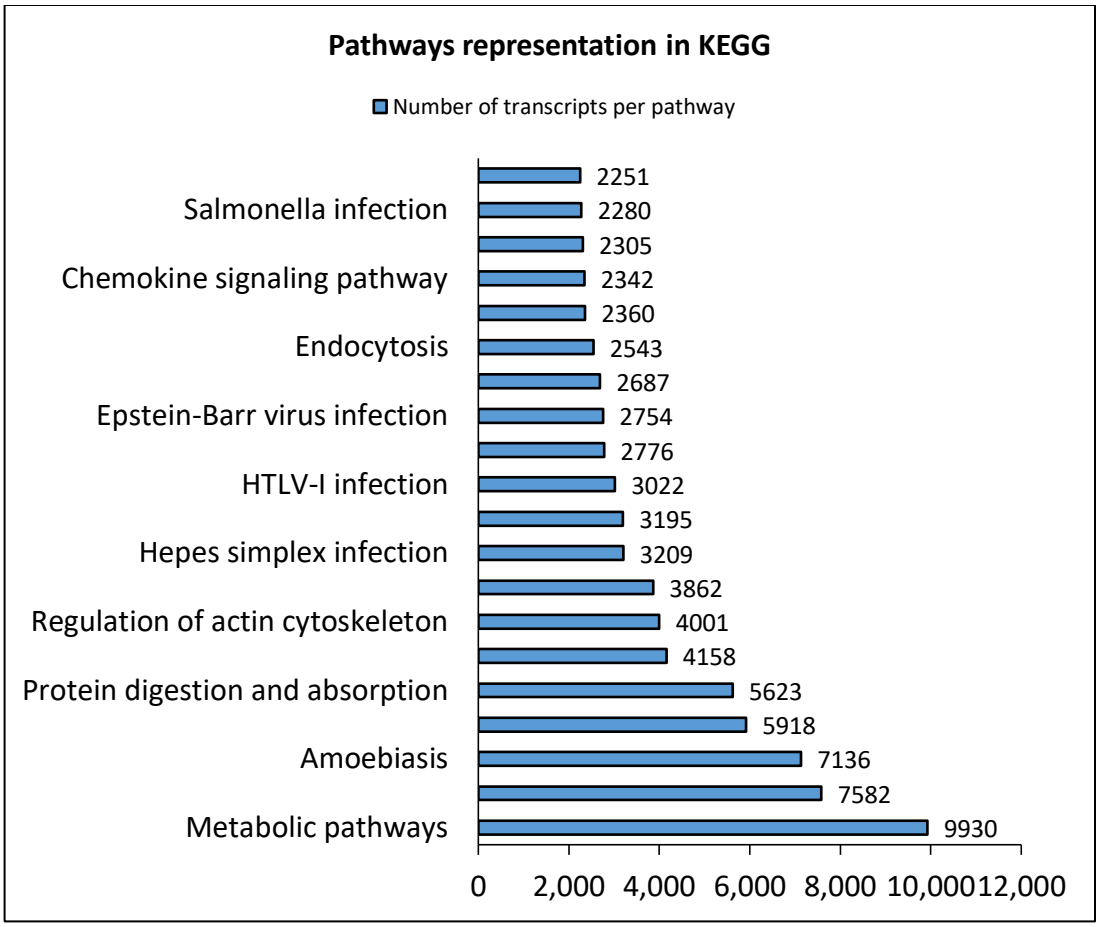


Figure 3.9: Number of transcripts in the main KEGG pathways

In general, there are more transcripts annotated by all four databases combined (9806) than transcripts annotated in one or two or three database as shown in Venn diagram **(Figure 3.10)**. NR has the majority of transcripts annotated in an independent database (8800) whereas COG has the lowest (34).

The pattern in this is approximately consistent with other Venn diagrams in other transcriptome analysis studies and will be further discussed in the discussion section

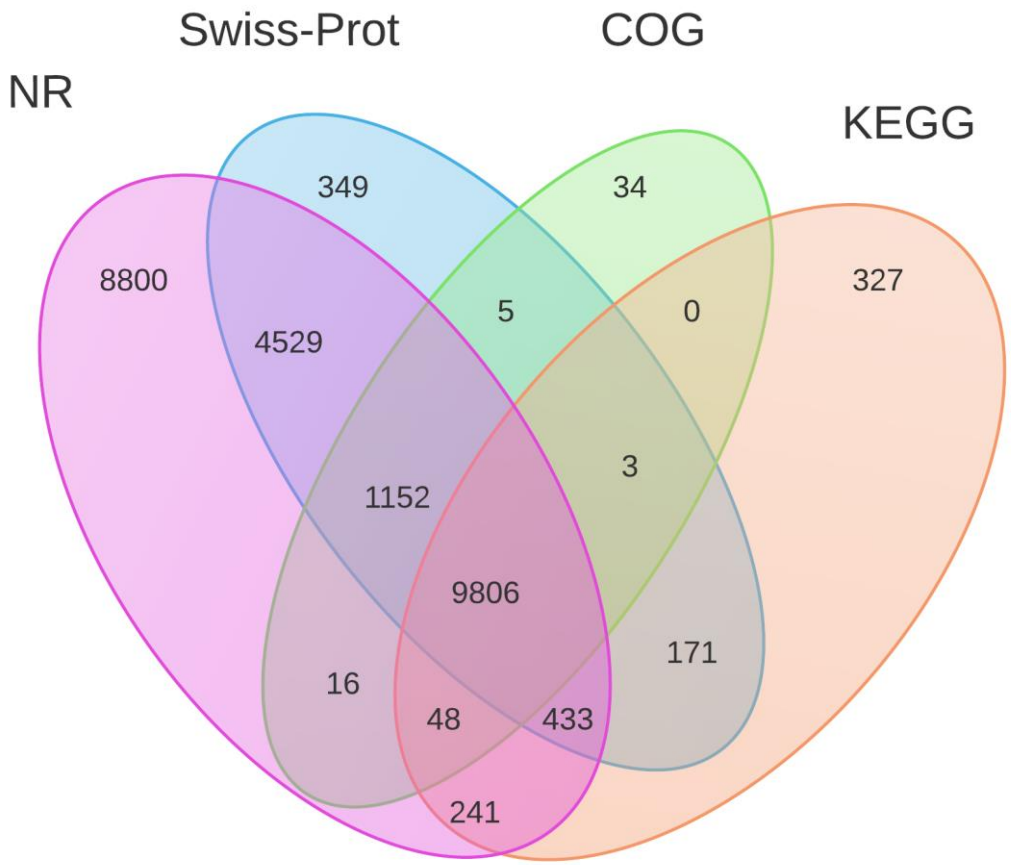


Figure 3.10: Venn diagram illustrating shared and unique transcripts annotated in databases of Nr, Swiss-Prot, COG and KEGG.

3.1 Quality Control using BLAST sequence alignment tool

In order to check the quality of the acquired transcriptome a BLAST search was performed against some of the badgers mRNA sequences available in the NCBI database. To date, there are 20-deposited partial sequences of badger's mRNA in the NCBI database that represent three genes; MCH class I antigen, MCH class II antigen and Interferon gamma.

The first comparison was between a transcript which is a partial sequence (399bp) of MHC class I antigen and a partial sequence of mRNA of the same gene (975bp) published by (Sin et al., 2012). The comparison show that the transcript in hand is 89% identical to that of the same organism currently available in public databases **(Figure 3.11)**. The variability in the number of class I sequences in *M. meles* is intermediate compared to other carnivores (Sin et al., 2012).

Description	Max score	Total score	Query cover	E value	Ident	Accession
Meles meles MHC class I antigen (Meme-MHCI) mRNA, Meme-MHCI*04 allele, partial cds	515	515	100%	3e-142	89%	JQ425430.1

Meles meles MHC class I antigen (Meme-MHCI) mRNA, Meme-MHCI*04 allele, partial cds

Sequence ID: gb|JQ425430.1| **Length:** 975 **Number of Matches:** 1

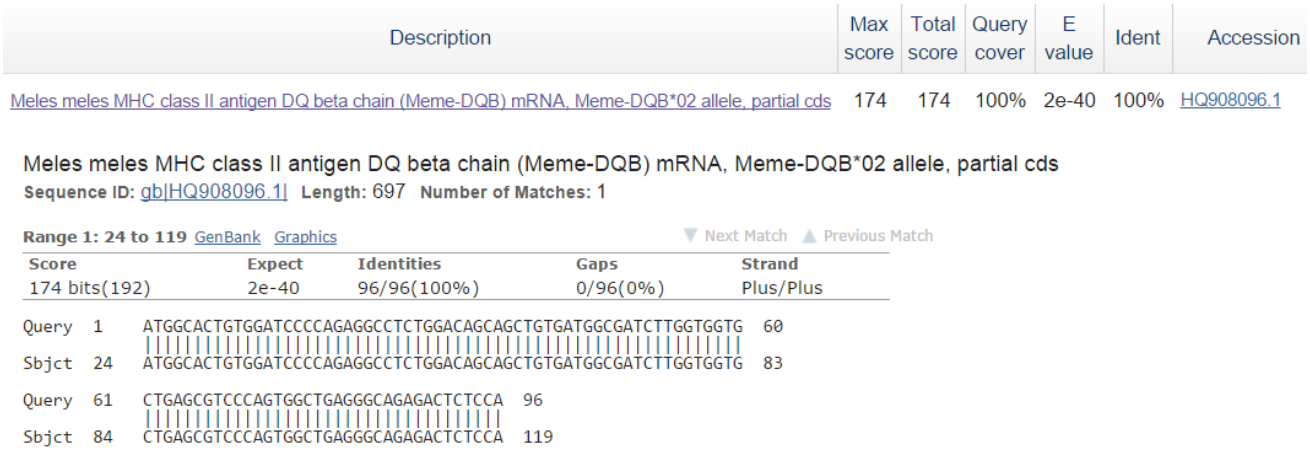
Range 1: 553 to 948 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
515 bits(570)	3e-142	354/399(89%)	3/399(0%)	Plus/Plus

```
Query  1    CGCTCAGAACTGCCCAACACACACGTGACCCACCACCCCATCTCTGACCATGCTAACACC  60
            ||| |||||   |||||| ||||||||||||| |||||||||||||||||| || |  ||||
Sbjct  553  CGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGACCGTGATGTCACC  612

Query  61   CTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACCTGGCAGAGGGAT  120
            |||||||| ||||||||||||||||||||| |||||||||||||||||| || | |||
Sbjct  613  CTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGAAGCGAGAT  672

Query  121  GGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGAGATGGAACC  180
            | |||||||| |||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  673  GAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCAGGAGATGGAACC  732

Query  181  TTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATACACATGCCAT  240
            |||||||||||||||||||||| ||||| |||||| ||||||||||||||||||| ||
Sbjct  733  TTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAGATACACATGCTAT  792

Query  241  GTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCTCCTCCCACC  300
            |||||||| |||||||| |||| |||||||||  ||| ||| |||||   |||| ||||
Sbjct  793  GTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCCA---CCTCACACC  849

Query  301  ATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTGGTGGTTGGA  360
            ||||||||||  |||||||||||||| |||| ||||||||  |||| ||  ||| |||||
Sbjct  850  ATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGGTCATTGCAGTGATTGGA  909

Query  361  GCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGA  399
            |||||||||||| ||||||||||||||||||| |||||||
Sbjct  910  GCTGTGATCTGGTGGAAGAAGCGCTCAGGAGAAAAAGGA  948
```

Figure 3.11: Two sequence alignment using online NCBI blastn application: MHC class I antigen partial sequence

The second comparison **(Figure 3.12)** was between a partial sequence (96bp) of MHC class II antigen DQ beta chain from the transcriptome and a partial sequence (697bp) of the same mRNA from the NCBI public database by (Sin et al., 2011). The two sequences were 100% identical.



Description	Max score	Total score	Query cover	E value	Ident	Accession
Meles meles MHC class II antigen DQ beta chain (Meme-DQB) mRNA, Meme-DQB*02 allele, partial cds	174	174	100%	2e-40	100%	HQ908096.1

Meles meles MHC class II antigen DQ beta chain (Meme-DQB) mRNA, Meme-DQB*02 allele, partial cds

Sequence ID: gb|HQ908096.1| **Length:** 697 **Number of Matches:** 1

Range 1: 24 to 119 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
174 bits(192)	2e-40	96/96(100%)	0/96(0%)	Plus/Plus

```
Query  1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  24   ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  83

Query  61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCA  96
            ||||||||||||||||||||||||||||||||||||
Sbjct  84   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCA  119
```

Figure 3.12: Two sequence alignment using online NCBI blastn application: MHC class II antigen partial sequence

The last comparison **(Figure 3.13)** was between a partial sequence (204bp) of interferon gamma from the transcriptome and a partial sequence of the same mRNA in NCBI database (501bp) submitted by (Zhou et al., 2014)

Description	Max score	Total score	Query cover	E value	Ident	Accession
M.meles mRNA for interferon gamma, partial	369	369	100%	1e-98	100%	Y11647.2

M.meles mRNA for interferon gamma, partial

Sequence ID: emb|Y11647.2| **Length:** 501 **Number of Matches:** 1

Range 1: 184 to 387 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
369 bits(408)	1e-98	204/204(100%)	0/204(0%)	Plus/Plus

```
Query  1    GAGAGTGACAAAACAATCATTCAAAGCCAAATTGTCTCCTTCTACTTGAAACTGTTTGAA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  184  GAGAGTGACAAAACAATCATTCAAAGCCAAATTGTCTCCTTCTACTTGAAACTGTTTGAA  243

Query  61   AACTTTAAAGATAACCAGATCATTCAAAGGAGCATGGATACCATCAAGGAAGACATGCTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  244  AACTTTAAAGATAACCAGATCATTCAAAGGAGCATGGATACCATCAAGGAAGACATGCTT  303

Query  121  GTCAGGTTCTTCAATAGCAGCAGCAGTAAGCGGGAGGACTTTCTTAAGCTGATTCGAATT  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  304  GTCAGGTTCTTCAATAGCAGCAGCAGTAAGCGGGAGGACTTTCTTAAGCTGATTCGAATT  363

Query  181  CCCGTGAATGATCTGCAGGTCCAG  204
            |||||||||||||||||||||||||
Sbjct  364  CCCGTGAATGATCTGCAGGTCCAG  387
```

Figure 3.13: Two sequence alignment using online NCBI blastn application: Interferon gamma

3.2 Using the transcriptome as a searchable database for MHC genes sequences

MHC are one of the major regulators of cell-mediated adaptive immune response. On average, there is a 10% difference between any two unrelated Individuals in the diversity of MHC combinations, which provide a protective function against pathogens. For an organism, the repertoire of MHC is polygenic, co-dominantly expressed from both sets of inherited alleles and highly polymorphic (Janeway, 2005). Moreover, the evidence of antigenic peptide splicing which can combine peptides from different proteins resulting in increased MHC antigen diversity (Vigneron et al., 2004)and low MHC diversity may pose a threat to the organism's pathogen resistance and on the long term the species survival (Zhu et al., 2007).

Here is an attempt to extract the MHC transcripts from the badger's transcriptome that correspond to 20 RNA-seq sequences in NCBI database and comments on their abundance and significance comparing to that of the giant panda and the effect of that on the whole immune response to infection.

The web-based platform Galaxy (Cock et al., 2015) was used in data analysis. Galaxy utilises BLAST+ command-line applications (Camacho et al., 2009a) through a user-friendly graphical interface. The first task performed was the conversion of the badger's transcriptome sequences data (FASTA) files into a searchable database (makeblastdb) followed by extraction of the badger's MCH class I and II RNA-seq genes sequences

(FASTA) from NCBI. Both nucleotide and protein sequences were combined into two query sequences lists (**Figures 3.14 and 3.15)**

JQ425440.1 Meles meles MHC class I antigen (Meme-MHCI) pseudogene mRNA, Meme-MHCI*PS03 allele, partial sequence
JQ425439.1 Meles meles MHC class I antigen (Meme-MHCI) pseudogene mRNA, Meme-MHCI*PS02 allele, partial sequence
JQ425438.1 Meles meles MHC class I antigen (Meme-MHCI) pseudogene mRNA, Meme-MHCI*PS01 allele, partial sequence
JQ425432.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA, Meme-MHCI*06 allele, partial cds
JQ425433.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA, Meme-MHCI*07 allele, partial cds
JQ425431.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA, Meme-MHCI*05 allele, partial cds
JQ425430.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA, Meme-MHCI*04 allele, partial cds
JQ425429.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA, Meme-MHCI*03 allele, partial cds
JQ425428.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA, Meme-MHCI*02 allele, partial cds
JQ425427.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA, Meme-MHCI*01 allele, partial cds
HQ908107.1 Meles meles nonfunctional MHC class II antigen (Meme-DQB) pseudogene mRNA, Meme-DQB*PS01 allele, partial sequence
HQ908099.1 Meles meles MHC class II antigen DR alpha chain (Meme-DRA) mRNA, Meme-DRA*02 allele, partial cds
HQ908098.1 Meles meles MHC class II antigen DQ alpha chain (Meme-DQA) mRNA, Meme-DQA*02 allele, partial cds
HQ908097.1 Meles meles MHC class II antigen DQ alpha chain (Meme-DQA) mRNA, Meme-DQA*01 allele, partial cds
HQ908096.1 Meles meles MHC class II antigen DQ beta chain (Meme-DQB) mRNA, Meme-DQB*02 allele, partial cds
HQ908095.1 Meles meles MHC class II antigen DR beta chain (Meme-DRB) mRNA, Meme-DRB*04 allele, partial cds
HQ908094.1 Meles meles MHC class II antigen DR beta chain (Meme-DRB) mRNA, Meme-DRB*03 allele, partial cds
HQ908093.1 Meles meles MHC class II antigen DR beta chain (Meme-DRB) mRNA, Meme-DRB*02 allele, partial cds
HQ908092.1 Meles meles MHC class II antigen DR beta chain (Meme-DRB) mRNA, Meme-DRB*01 allele, partial cds
Y11647.2 M.meles mRNA for interferon gamma, partial

Figure 3.14: Query nucleotide sequences extracted from NCBI database

AFR54067.1 MHC class I antigen, partial [Meles meles]
AFR54066.1 MHC class I antigen, partial [Meles meles]
AFR54065.1 MHC class I antigen, partial [Meles meles]
AFR54064.1 MHC class I antigen, partial [Meles meles]
AFR54063.1 MHC class I antigen, partial [Meles meles]
AFR54062.1 MHC class I antigen, partial [Meles meles]
AFR54061.1 MHC class I antigen, partial [Meles meles]
AFR54060.1 MHC class I antigen, partial [Meles meles]
AFR54059.1 MHC class I antigen, partial [Meles meles]
AFR54058.1 MHC class I antigen, partial [Meles meles]
AFR54057.1 MHC class I antigen, partial [Meles meles]
AFR54056.1 MHC class I antigen, partial [Meles meles]
AFR54055.1 MHC class I antigen, partial [Meles meles]
AFR54054.1 MHC class I antigen, partial [Meles meles]
AET36883.1 MHC class II antigen, partial [Meles meles]
AET36881.1 MHC class II antigen, partial [Meles meles]
AET36880.1 MHC class II antigen, partial [Meles meles]
AET36875.1 MHC class II antigen DR alpha chain, partial [Meles meles]
AET36874.1 MHC class II antigen DQ alpha chain, partial [Meles meles]
AET36873.1 MHC class II antigen DQ alpha chain, partial [Meles meles]
AET36872.1 MHC class II antigen DQ beta chain, partial [Meles meles]
AET36871.1 MHC class II antigen DR beta chain, partial [Meles meles]
AET36870.1 MHC class II antigen DR beta chain, partial [Meles meles]
AET36869.1 MHC class II antigen DR beta chain, partial [Meles meles]
AET36868.1 MHC class II antigen DR beta chain, partial [Meles meles]

Figure 3.15: Query nucleotide sequences extracted from NCBI database

Traditional megablast was used to perform BLAST search to find highly similar sequences (Megablast is used for intra-species or closely related species) with e-value cut-off point

of 1e-05. Both forward and reverse strands of the query sequences were searched against the database with no restriction in the maximum number of hits.

3.2.1 Nucleotide BLAST search (tabular format)

The blast results were first tabulated to obtain a general visualisation of the data in hand. The initial observations of this alignment are:

The total number of achieved alignments between the query sequences and the transcriptome database is 195 **(Appendix)** with an identity percentage range between 76% and 100%. The e-value range of the alignments was between 0.0 and 4.00e-14.

From the total alignments 6 query NCBI sequence achieved a 100% identity with 19 transcripts **(Figure 3.16)**. 49 aligned transcripts with 18 query sequences came up with an E-value of 0 **(Figure 3.17)**.

Query Seq-id	Subject seq-id	Identity %	Alignment length	E-vlue	Query seq. length	Subject seq. length	Sequence classification
JQ425440.1	Unigene117550_All	100	83	2.00E-36	848	171	MHC I
JQ425440.1	Unigene12618_All	100	83	2.00E-36	848	171	MHC I
JQ425440.1	Unigene26995_All	100	73	8.00E-31	848	126	MHC I
JQ425439.1	Unigene117550_All	100	84	7.00E-37	932	171	MHC I
JQ425439.1	Unigene12618_All	100	84	7.00E-37	932	171	MHC I
JQ425439.1	Unigene26995_All	100	73	9.00E-31	932	126	MHC I
JQ425432.1	CL1150.Contig22_All	100	110	3.00E-51	975	360	MHC I
HQ908107.1	Unigene4377_All	100	96	1.00E-43	680	96	MHC II
HQ908107.1	CL4065.Contig4_All	100	96	1.00E-43	680	96	MHC II
HQ908107.1	CL4065.Contig2_All	100	96	1.00E-43	680	96	MHC II
HQ908107.1	CL4065.Contig1_All	100	96	1.00E-43	680	96	MHC II
HQ908098.1	CL10050.Contig2_All	100	582	0	615	582	MHC II
HQ908098.1	CL10050.Contig1_All	100	582	0	615	582	MHC II
HQ908096.1	Unigene4377_All	100	96	1.00E-43	697	96	MHC II
HQ908096.1	CL4065.Contig4_All	100	96	1.00E-43	697	96	MHC II
HQ908096.1	CL4065.Contig2_All	100	96	1.00E-43	697	96	MHC II
HQ908096.1	CL4065.Contig1_All	100	96	1.00E-43	697	96	MHC II
HQ908092.1	CL13896.Contig1_All	100	96	1.00E-43	822	96	MHC II
Y11647.2	Unigene54563_All	100	204	7.00E-104	501	204	Interferon Gamma

Figure 3.16: Nucleotide alignments with a 100% identity score

Query Seq-id	Subject seq-id	Identity %	Alignment length	E-vlue	Query seq. length	Subject seq. length	Sequence classification
JQ425440.1	CL1150.Contig16_All	90.12	506	0	848	972	MHC I
JQ425439.1	CL1150.Contig16_All	90.04	562	0	932	972	MHC I
JQ425439.1	CL1150.Contig4_All	89.09	550	0	932	702	MHC I
JQ425438.1	CL1150.Contig16_All	90.8	848	0	900	972	MHC I
JQ425438.1	CL1150.Contig4_All	91.31	587	0	900	702	MHC I
JQ425432.1	CL1150.Contig16_All	92.08	972	0	975	972	MHC I
JQ425432.1	CL1150.Contig4_All	93.28	699	0	975	702	MHC I
JQ425433.1	CL1150.Contig16_All	94.44	972	0	975	972	MHC I
JQ425433.1	CL1150.Contig4_All	98.71	699	0	975	702	MHC I
JQ425431.1	CL1150.Contig16_All	91.98	972	0	975	972	MHC I
JQ425431.1	CL1150.Contig4_All	92.7	699	0	975	702	MHC I
JQ425430.1	CL1150.Contig16_All	93.44	975	0	975	972	MHC I
JQ425430.1	CL1150.Contig4_All	93.87	702	0	975	702	MHC I
JQ425429.1	CL1150.Contig16_All	95.16	972	0	975	972	MHC I
JQ425429.1	CL1150.Contig4_All	98	699	0	975	702	MHC I
JQ425428.1	CL1150.Contig16_All	94.32	546	0	543	972	MHC I
JQ425428.1	CL1150.Contig4_All	92.49	546	0	543	702	MHC I
JQ425427.1	CL1150.Contig4_All	93.59	546	0	543	702	MHC I
JQ425427.1	CL1150.Contig16_All	93.04	546	0	543	972	MHC I
HQ908107.1	Unigene57188_All	97.18	674	0	680	795	MHC II
HQ908107.1	Unigene67843_All	97.23	650	0	680	651	MHC II
HQ908107.1	Unigene75449_All	97.28	551	0	680	774	MHC II
HQ908099.1	CL2981.Contig3_All	99.27	688	0	691	762	MHC II
HQ908099.1	CL2981.Contig1_All	99.27	688	0	691	762	MHC II
HQ908099.1	Unigene42913_All	99.67	603	0	691	603	MHC II
HQ908098.1	CL10050.Contig3_All	99.67	615	0	615	765	MHC II
HQ908098.1	CL10050.Contig2_All	100	582	0	615	582	MHC II
HQ908098.1	CL10050.Contig1_All	100	582	0	615	582	MHC II
HQ908097.1	CL10050.Contig3_All	91.73	617	0	615	765	MHC II
HQ908097.1	CL10050.Contig2_All	91.61	584	0	615	582	MHC II
HQ908097.1	CL10050.Contig1_All	91.61	584	0	615	582	MHC II
HQ908096.1	Unigene57188_All	99.7	674	0	697	795	MHC II
HQ908096.1	Unigene67843_All	99.85	650	0	697	651	MHC II
HQ908096.1	Unigene75449_All	97.28	551	0	697	774	MHC II
HQ908096.1	Unigene75450_All	85.32	654	0	697	774	MHC II
HQ908095.1	Unigene27967_All	94.83	774	0	822	774	MHC II
HQ908095.1	Unigene75450_All	94.6	741	0	822	774	MHC II
HQ908095.1	CL4065.Contig3_All	96.64	684	0	822	693	MHC II
HQ908094.1	Unigene27967_All	99.22	774	0	822	774	MHC II
HQ908094.1	Unigene75450_All	97.17	672	0	822	774	MHC II
HQ908094.1	CL4065.Contig3_All	97.17	672	0	822	693	MHC II
HQ908094.1	Unigene75449_All	84.09	773	0	822	774	MHC II
HQ908093.1	Unigene27967_All	99.48	774	0	822	774	MHC II
HQ908093.1	Unigene75450_All	97.62	672	0	822	774	MHC II
HQ908093.1	CL4065.Contig3_All	97.62	672	0	822	693	MHC II
HQ908093.1	Unigene75449_All	83.44	773	0	822	774	MHC II
HQ908092.1	Unigene27967_All	96.51	774	0	822	774	MHC II
HQ908092.1	Unigene75450_All	92.31	741	0	822	774	MHC II
HQ908092.1	CL4065.Contig3_All	94.79	672	0	822	693	MHC II

Figure 3.17: Nucleotide alignments with E-value = 0

The percentage of identity for the alignments with E-value=0 is in the range between 83% and 100%. The alignment lengths were between 506 and 975 bases.

Only one query sequence (ID: HQ908098.1 MHC class II antigen DQ alpha chain) achieved both an alignment with 100% identity and E-value=0 with two transcripts (ID: CL10050.Contig1_All, CL10050.Contig2_All).

3.2.2 Nucleotide alignment (Query-anchored)

The tabulated alignment was also converted into an alignment (query-anchored format). Due to the large data file only one example of the alignment is mentioned below **(Figure 3.18)**. The reset of the alignments can be viewed in the **(Appendix)**.

Example:

Query= JQ425440.1 *Meles meles* MHC class I antigen (Meme-MHCI) pseudogene mRNA, Meme-MHCI*PS03 allele, partial sequence. Length=848bp

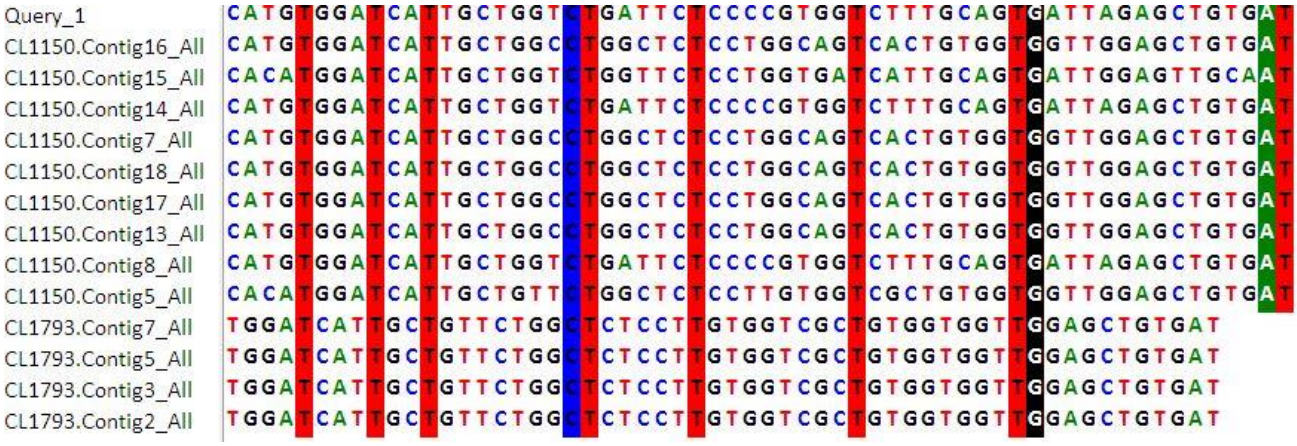


Figure 3.18: A section of the nucleotide alignment between a MHC class one (query) and the badger transcripts

On average, 10-17 alignments were achieved between each query sequence and transcripts from the badger transcriptome database. Most of the first hits in the alignments for each query were with transcripts that have been previously annotated as the badger's own in NCBI blast search (during transcriptome assembly).

The remainder alignments were with closely related mammals like the ferret (*Mustela putorius furo*) or extensively studied (for economical or conservational reasons) mammals such as panda (*Ailuropoda melanoleuca*), cow (*Bos taurus*), horse (*Equus caballus*), Californian sea lion (*Zalophus californianus*), olive baboon (*Papio anubis*) and marmoset (*Callithrix jacchus)* as shown in the **(Figure 3.19).**

Subject seq-id	NR-annotation
CL10050.Contig1_All	MHC class II antigen DQ alpha chain, partial [Meles meles]
CL10050.Contig2_All	MHC class II antigen DQ alpha chain, partial [Meles meles]
CL10050.Contig3_All	PREDICTED: SLA class II histocompatibility antigen, DQ haplotype D alpha chain-like [Ailuropoda melanoleuca]
CL1150.Contig1_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig13_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig14_All	hypothetical protein PANDA_022308 [Ailuropoda melanoleuca]
CL1150.Contig15_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig16_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig17_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig18_All	MHC class I antigen [Ailuropoda melanoleuca]
CL1150.Contig2_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig22_All	PREDICTED: LOW QUALITY PROTEIN: popy Class I histocompatibility antigen, A-1 alpha chain-like [Papio anubis]
CL1150.Contig23_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig24_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig3_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig4_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig5_All	MHC class I antigen [Bos taurus]
CL1150.Contig7_All	MHC class I antigen [Ailuropoda melanoleuca] >gi|163636633|gb|ABY27208.1| MHC class I antigen [Ailuropoda melanoleuca]
CL1150.Contig8_All	hypothetical protein PANDA_022308 [Ailuropoda melanoleuca]
CL13896.Contig1_All	MHC class II antigen DR beta chain, partial [Meles meles]
CL1793.Contig2_All	PREDICTED: MHC class I polypeptide-related sequence B-like [Equus caballus]
CL1793.Contig3_All	PREDICTED: MHC class I polypeptide-related sequence B-like [Equus caballus]
CL1793.Contig5_All	PREDICTED: MHC class I polypeptide-related sequence B-like [Equus caballus]
CL1793.Contig7_All	PREDICTED: MHC class I polypeptide-related sequence B-like [Equus caballus]
CL2981.Contig1_All	PREDICTED: HLA class II histocompatibility antigen, DR alpha chain-like [Ailuropoda melanoleuca]
CL2981.Contig3_All	PREDICTED: HLA class II histocompatibility antigen, DR alpha chain-like [Ailuropoda melanoleuca]
CL4065.Contig1_All	MHC class II antigen DQ beta chain, partial [Meles meles]
CL4065.Contig2_All	MHC class II antigen [Zalophus californianus] >gi|22023813|gb|AAM89234.1| MHC class II antigen [Zalophus californianus]
CL4065.Contig3_All	MHC class II antigen DR beta chain, partial [Meles meles]
CL4065.Contig4_All	MHC class II antigen [Zalophus californianus] >gi|22023813|gb|AAM89234.1| MHC class II antigen [Zalophus californianus]
CL7631.Contig3_All	MHC class I antigen, partial [Meles meles]
Unigene117550_All	PREDICTED: patr class I histocompatibility antigen, A-126 alpha chain-like, partial [Callithrix jacchus]
Unigene12618_All	PREDICTED: patr class I histocompatibility antigen, A-126 alpha chain-like, partial [Callithrix jacchus]
Unigene26995_All	hypothetical protein PANDA_022308 [Ailuropoda melanoleuca]
Unigene27967_All	hypothetical protein PANDA_022308 [Ailuropoda melanoleuca]
Unigene42913_All	MHC class II antigen DR alpha chain, partial [Meles meles]
Unigene4377_All	MHC class II antigen DQ beta chain, partial [Meles meles]
Unigene54563_All	interferon gamma [Mustela putorius furo]
Unigene57188_All	MHC class II antigen [Zalophus californianus] >gi|22023813|gb|AAM89234.1| MHC class II antigen [Zalophus californianus]
Unigene67843_All	MHC class II antigen DQ beta chain, partial [Meles meles]
Unigene75449_All	MHC class II antigen [Zalophus californianus] >gi|22023813|gb|AAM89234.1| MHC class II antigen [Zalophus californianus]
Unigene75450_All	MHC class II antigen DR beta chain, partial [Meles meles]
Unigene81593_All	MHC class II antigen, partial [Meles meles]

Figure 3.19: Aligned transcripts NR annotation during transcriptome assembly and annotation (yellow cells indicate sequences previously annotated as MHC genes from the badger)

3.2.3 Protein sequences BLAST search (tabular format)

Search criteria for protein sequence BLAST included the use of traditional BLASTP to compare a protein query to a protein database with e-value cut-off point of 1e-05 and a scoring matrix of BLOSUM90 (which is used to compare highly related and less divergent sequences). BLOSUM only accepts a mutation in the protein primary structure if it is commonly found in conservative substitutions in nature.

The total number of achieved alignments between query sequences and transcripts is 248. The Identity percentage observed fell in a range of between 37% and 100%. The e-value range was between 0.0 and 8.00e-13.

Five alignments between two query NCBI sequences and the transcriptome database showed a 100% match **(Figure 3.20)** and only seven alignments with seven query sequences produced an e-value of 0 **(Figure 3.21)** i.e. the probability of the alignments arising by chance is zero.

Query Seq-id	Subject seq-id	Identity %	Alignment length	E-vlue	Bit score	Query seq. length	Subject seq. length	Sequence classification
AET36874.1	CL10050.Contig2_All	100	194	3.00E-146	423	205	194	MHC II
AET36874.1	CL10050.Contig1_All	100	194	3.00E-146	423	205	194	MHC II
AET36883.1	Unigene42913_All	100	81	8.00E-54	177	81	201	MHC II
AET36883.1	CL2981.Contig3_All	100	81	3.00E-53	178	81	254	MHC II
AET36883.1	CL2981.Contig1_All	100	81	3.00E-53	178	81	254	MHC II

Figure 3.20: Protein alignments with 100% identity score

Query Seq-id	Subject seq-id	Identity %	Alignment length	E-vlue	Bit score	Query seq. length	Subject seq. length	Sequence classification
AFR54060.1	CL1150.Contig16_All	88.58	324	0	618	325	324	MHC I
AFR54059.1	CL1150.Contig16_All	85.19	324	0	596	325	324	MHC I
AFR54058.1	CL1150.Contig16_All	84.88	324	0	593	325	324	MHC I
AFR54057.1	CL1150.Contig16_All	86.73	324	0	604	325	324	MHC I
AFR54056.1	CL1150.Contig16_All	89.81	324	0	625	325	324	MHC I
AET36870.1	Unigene27967_All	98.84	258	0	554	258	258	MHC II
AET36869.1	Unigene27967_All	99.22	258	0	557	258	258	MHC II

Figure 3.21: Protein alignments with an E-value = 0

3.2.4 Protein alignment (Query-anchored)

The tabulated protein alignment was also converted into a query-anchored alignment format as in the example **(Figure 3.22).**

Database: BLAST Database, 127,401 sequences; 28,354,467 total letters

Query = AFR54067.1 MHC class I antigen, partial [*Meles meles*], Length=180

```
Query_1              1    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  60
CL1150.Contig4_All   3    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  62
CL1150.Contig16_All  6    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  65
CL154.Contig12_All   2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  59
CL154.Contig11_All   2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  59
CL154.Contig24_All   54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  111
CL154.Contig19_All   54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  111
CL154.Contig27_All   41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig23_All   41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig2_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig1_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98

Query_1              61   QQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYSQHSYDGA  120
CL1150.Contig4_All   63   QQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQFAYDGA  122
CL1150.Contig16_All  66   RQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQFAYDGA  125
CL154.Contig12_All   60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ  117
CL154.Contig11_All   60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ  117
CL154.Contig24_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ  169
CL154.Contig19_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ  169
CL154.Contig27_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ  156
CL154.Contig23_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ  156
CL154.Contig2_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ  156
CL154.Contig1_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ  156

Query_1              121  DYIALNEDLRSWTAADTAAQITQRKWE-DAGEAERWRNYVEGTCVEWLGRYLENGKESLL  179
CL1150.Contig4_All   123  DYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESLL  181
CL1150.Contig16_All  126  DYIALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESLL  184
CL154.Contig12_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT  177
CL154.Contig11_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT  177
CL154.Contig24_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT  229
CL154.Contig19_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT  229
CL154.Contig27_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT  216
CL154.Contig23_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT  216
CL154.Contig2_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT  216
CL154.Contig1_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT  216
```

Figure 3.22: An example of an alignment of a protein query sequence (MHC class I antigen) and 10 transcripts from the transcriptome database.

The 25 protein query sequences (NCBI) achieved 248 alignments with 42 transcripts from the badger transcriptome database. 10 of the alignments achieved were with transcripts previously annotated as badger MHC sequences genes. The remainder alignments were with closely related mammals like the ferret (*Mustela putorius furo*) or extensively studied (for economical or conservational reasons) mammals such as panda (*Ailuropoda*

melanoleuca), dog (*Canis lupus familiaris*), and Californian sea lion (*Zalophus californianus*) as shown in the **(Figure 3.23).**

Subject seq-id	NR-annotation
CL10050.Contig1_All	MHC class II antigen DQ alpha chain, partial [Meles meles]
CL10050.Contig2_All	MHC class II antigen DQ alpha chain, partial [Meles meles]
CL10050.Contig3_All	PREDICTED: SLA class II histocompatibility antigen, DQ haplotype D alpha chain-like [Ailuropoda melanoleuca]
CL1150.Contig13_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig15_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig16_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig17_All	MHC class I antigen, partial [Meles meles]
CL1150.Contig18_All	MHC class I antigen [Ailuropoda melanoleuca]
CL1150.Contig4_All	MHC class I antigen, partial [Meles meles]
CL12484.Contig1_All	HLA class II histocompatibility antigen, DO alpha chain precursor [Mustela putorius furo]
CL12484.Contig2_All	PREDICTED: HLA class II histocompatibility antigen, DO alpha chain-like [Ailuropoda melanoleuca]
CL12484.Contig3_All	MHC class II antigen DO alpha [Canis lupus familiaris]
CL154.Contig1_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig11_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig12_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig13_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig14_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig15_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig16_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig17_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig19_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig2_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig20_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig21_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig23_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig24_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig27_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL154.Contig33_All	class I histocompatibility antigen, GOGO-C0203 alpha chain-like protein [Mustela putorius furo]
CL2981.Contig1_All	PREDICTED: HLA class II histocompatibility antigen, DR alpha chain-like [Ailuropoda melanoleuca]
CL2981.Contig3_All	PREDICTED: HLA class II histocompatibility antigen, DR alpha chain-like [Ailuropoda melanoleuca]
CL5174.Contig1_All	hypothetical protein PANDA_002284 [Ailuropoda melanoleuca]
CL6815.Contig2_All	PREDICTED: HLA class II histocompatibility antigen, DO beta chain-like [Ailuropoda melanoleuca]
CL6815.Contig3_All	PREDICTED: HLA class II histocompatibility antigen, DO beta chain-like [Ailuropoda melanoleuca]
CL6815.Contig4_All	major histocompatibility complex, class II, DO beta [Mustela putorius furo]
CL6815.Contig5_All	PREDICTED: HLA class II histocompatibility antigen, DO beta chain-like [Ailuropoda melanoleuca]
CL6815.Contig6_All	major histocompatibility complex, class II, DO beta [Mustela putorius furo]
Unigene27967_All	hypothetical protein PANDA_022308 [Ailuropoda melanoleuca]
Unigene42913_All	MHC class II antigen DR alpha chain, partial [Meles meles]
Unigene57188_All	MHC class II antigen [Zalophus californianus] >gi|22023813|gb|AAM89234.1| MHC class II antigen [Zalophus californianus]
Unigene67843_All	MHC class II antigen DQ beta chain, partial [Meles meles]
Unigene75449_All	MHC class II antigen [Zalophus californianus] >gi|22023813|gb|AAM89234.1| MHC class II antigen [Zalophus californianus]
Unigene75450_All	MHC class II antigen DR beta chain, partial [Meles meles]

Figure 3.23: List of the transcripts that achieved and alignment with all 25 query protein sequences

An example of an alignment using Clustal Omega (Sievers et al., 2011) between the first query sequence (AFR54067.1, a partial sequence of MHC class I antigen) and two transcripts is shown in the **(figure 3.24)** to visualise the differences in the sequences and to show different types of mutations found.

```
CLUSTAL O(1.2.4) multiple sequence alignment


AFR54067.1              -----SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG
CL1150.Contig4_All      ---AGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG
CL1150.Contig16_All     ETWAGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG
                             ************************************************************

AFR54067.1              PEYWDQQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYSQH
CL1150.Contig4_All      PEYWDQQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQF
CL1150.Contig16_All     PEYWDRQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQF
                        *****:**:  ******** *** **************:**:************** *.

AFR54067.1              SYDGADYIALNEDLRSWTAADTAAQITQRKWEDAGEAERWRNYVEGTCVEWLGRYLENGK
CL1150.Contig4_All      AYDGADYLALNEDLRSWTAADTAAQISRRKWEDAGEAERYRNYVEGTCVEWLGRYLENGK
CL1150.Contig16_All     AYDGADYIALNEDLRSWTAADTAAQISRRKWEDAGEAERYRNYVEGTCVEWLGRYLENGK
                        :******:******************::**********:********************

AFR54067.1              ESLLR-----------------------------------------------------
CL1150.Contig4_All      ESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWKRDEEDLTQDTELVET---
CL1150.Contig16_All     ESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRPA
                        *****

AFR54067.1              -----------------------------------------------------
CL1150.Contig4_All      -----------------------------------------------------
CL1150.Contig16_All     GDGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLALLAVT
```

Figure 3.24: An alignment of one query sequence (AFR54067.1) and two transcripts where (*) represents conserved sequence, (:) conservative mutations, (.) semi-conservative mutations, and the space ( ) for non-conservative mutations.

3.2.5 A comparison between the badger and the giant panda

A comparison of the most abundant MHC transcripts in the badger and panda transcriptomes (Du et al., 2015a) in the **(Table 3.4).** The comparison is based on the abundance of the transcripts despite the annotation of the sequence in terms of the closest matching in the public databases, providing it is in the MHC gene family. The data are expressed as the total number of fragments (reads) per kilo-base of gene length per million reads of the transcriptome (FPKM). The data show that the total abundance of badger MHC transcripts is 4936.818 FPKM, which is about 2.5 times more than that of the giant panda (1995.43 FPKM).

Badger Transcript ID	**Transcript Length**	**FPKM**	**Giant Panda Transcript ID**	**Transcript Length**	**FPKM**
CL1150.Contig13_All	1363	1467.8917	asmbl_17923	1012	22.47
CL1150.Contig15_All	1385	1106.5719	asmbl_17924	2228	26.63
CL1150.Contig16_All	1435	671.4784	asmbl_17925	2257	55.28
CL1150.Contig4_All	1187	582.748	asmbl_42999	336	14.59
CL1150.Contig18_All	2627	277.4338	asmbl_43000	1024	16.42
Unigene27967_All	1466	250.1088	asmbl_43001	1408	4.31
CL10050.Contig1_All	1750	185.9091	asmbl_43002	1029	15.11
CL7631.Contig3_All	559	67.2444	asmbl_43007	193	3.16
Unigene62040_All	269	60.309	asmbl_49924	2383	12.4
Unigene42913_All	1020	53.2618	asmbl_49925	4142	15.09
CL4065.Contig2_All	3331	50.9938	asmbl_49929	455	1.96
CL1150.Contig6_All	251	44.5692	asmbl_55085	320	1744.82
CL10050.Contig2_All	954	33.539	asmbl_75564	2816	27.39
CL13896.Contig1_All	508	25.0588	asmbl_77939	2947	1.41
CL7631.Contig1_All	556	22.8955	asmbl_77940	1870	2.14
CL4065.Contig1_All	3796	20.1605	asmbl_77941	508	3.55
CL12484.Contig3_All	627	10.459	asmbl_92419	125	27.35
CL11221.Contig2_All	492	6.1853	asmbl_92575	104	1.35
Total	23576	4936.818	Total	25157	1995.43

Table 3.4: The most abundant MHC transcripts in the badger transcriptome and in the giant panda transcriptome

3.3 Using the most abundant transcripts to build a phylogenetic tree

In this section, transcripts with the highest raw reads, highest FPKM value and the longest amino acid sequences were used to build three phylogenetic trees using BLAST+ (Camacho et al., 2009b) and Clustal Omega (Sievers et al., 2011) from European Bioinformatics Institute (EMBL-EBI) web services (McWilliam et al., 2013).

The nature of the transcripts will be also discussed in terms of the conservative regions nature and whether they can be classified as highly conserved sequences.

The transcript with most abundant raw reads (CL4057.Contig1_All), transcript with highest FPKM value (Unigene65050_All) and transcript with the longest amino acid sequence (CL3144.Contig26_All) were chosen to build a tree.

3.3.1 The construction of a phylogenetic tree using the transcript with most abundant raw reads

The query sequence (ID: CL4057.Contig1_All, sequence length: 514 aa) was aligned with highest identical blast search results and a phylogenetic tree was drawn using Clustal Omega as in **(Figure 3.26)**

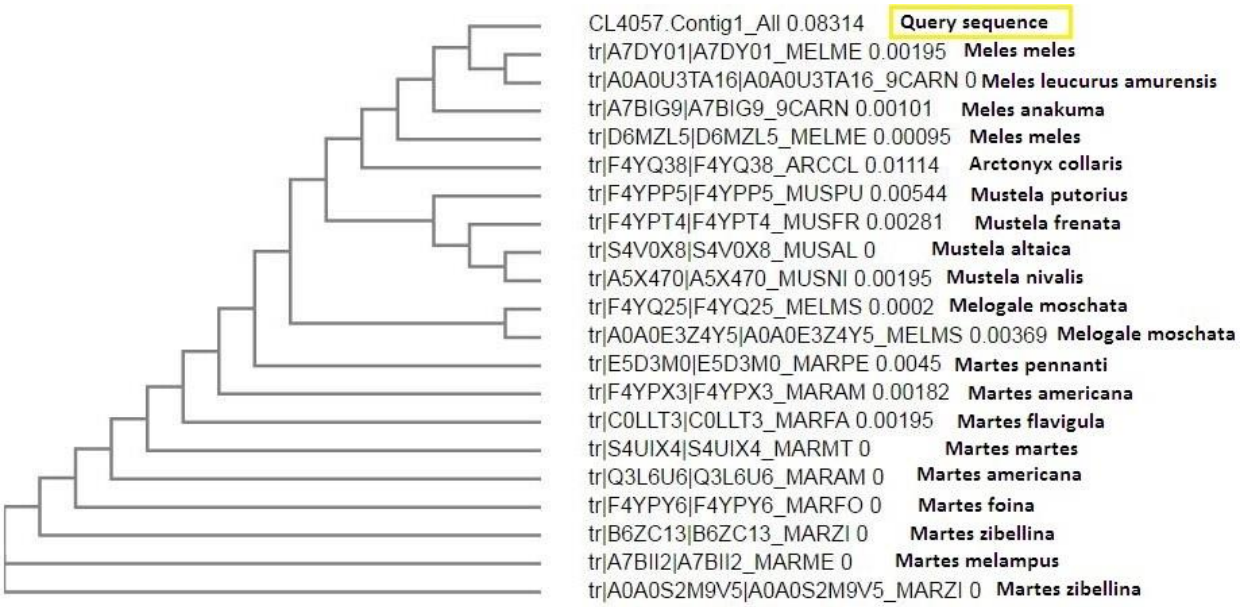


Figure 3.25: A phylogenetic tree model using the most abundant raw read (ID: CL4057.Contig1_All) in the badger transcriptome

Sequence (CL4057.Contig1_All) annotation in the BLAST search is the cytochrome c oxidase subunit I (COI), a key component in the mitochondrial electron transport chain. The gene is used as a "DNA barcode" to identify species. Although the sequence tends to be conserved among members of the same species, it has a fast enough mutation rate that enables distinction between closely related species where more than 2% sequence divergence cab be detected (Hebert et al., 2003).

The graph shows a degree of distinction within the *Mustelidae* family among the subfamilies such as *Mustelinae* (ferrets) and *Melinae* (badgers) and also on the species level within the subfamily which tend to be less diverse. This model agrees with multigene

phylogeny of the *Mustelidae* by (Koepfli et al., 2008a). However, it does not reflect the chronical evolution and divergence of the species.

3.3.2 Construction of a phylogenetic tree using the transcript with the highest FPKM value

Transcript (Unigene65050_All) has the highest FPKM value (38044.2458) in the transcriptome annotation data. A phylogenetic tree using this transcript is shown in the **(Figure 3.27).**

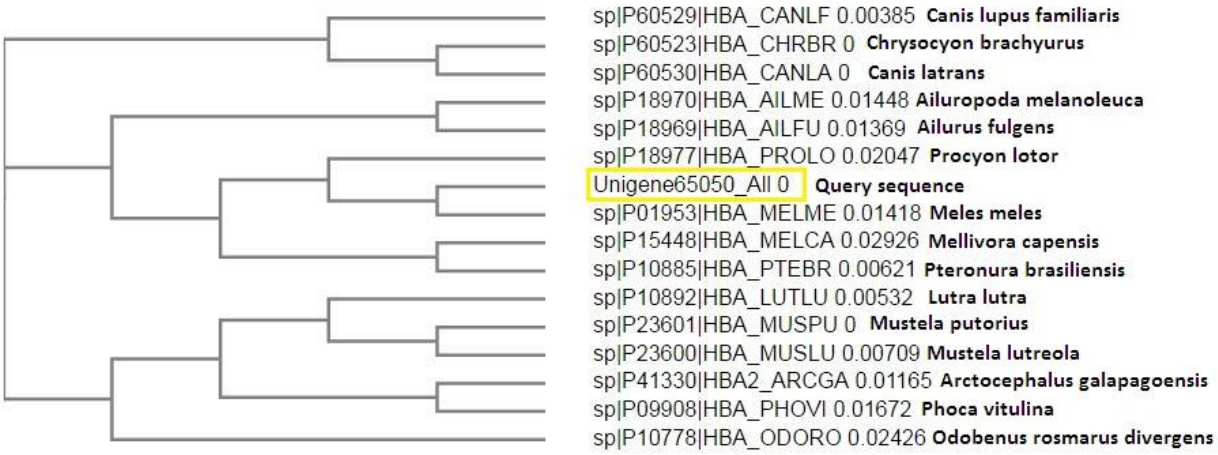


Figure 3.26: A phylogenetic tree model using the transcript with the highest FPKM value (Unigene65050_All) in the badger transcriptome

The query sequence (Unigene65050_All) has a length of (141 aa) and is annotated as Haemoglobin subunit alpha (HBA). The length of HBA gene in mammals is about 142 aa including human, panda, cow and badgers.

In this example, using a single gene to draw a phylogenetic (gene) tree does not necessarily represent the actual evolutionary pathway of a species (whole genome) tree.

3.3.3 Building a phylogenetic tree using the longest transcript sequence

In the transcriptome data files (FASTA), the query sequence (CL3144.Contig26_All) has the longest amino acid sequence with 4719 aa. A BLAST search followed by a Clustal Omega alignment and tree formation is in the **(figure 3.28).**

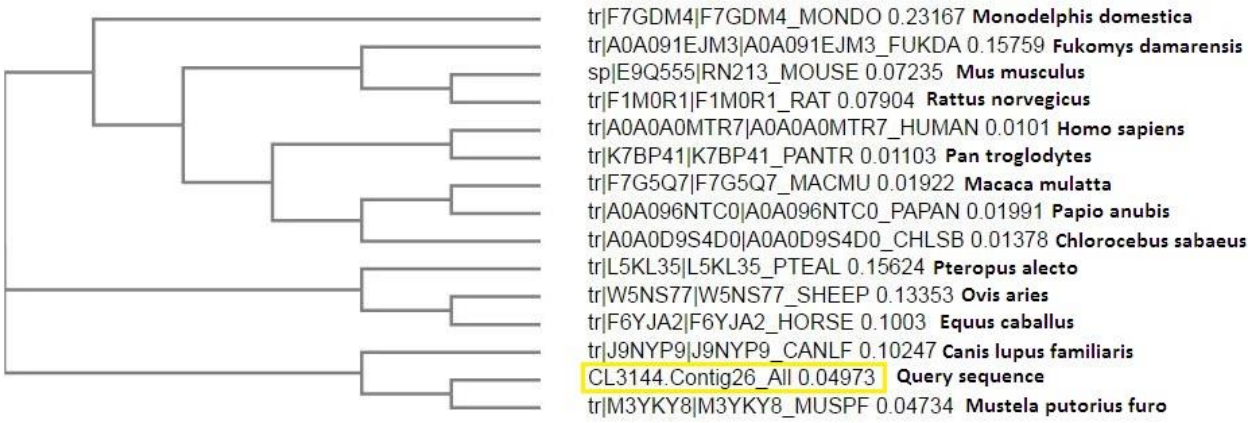


Figure 3.27: A phylogenetic tree model using the transcript with the longest amino acid sequence (CL3144.Contig26_All) in the badger transcriptome

The query sequence annotation E3 ubiquitin-protein ligase RNF213 suggests a protein involved in angiogenesis and vascular development. The function of the sequence is predicted in the transcriptome annotation as well as similar sequences from other mammals based on the well-known sequence from human genome.

3.4 Discussion

3.4.1 General description

To my knowledge, this is the first attempt to analyse the composition of any transcriptome from the European badger. The data appear to be of high quality compared

with other published blood transcriptome studies e.g. giant panda blood transcriptome (Du et al., 2015b). A large volume of data (about 3.52GB) has been generated using next generation (RNA-seq) technology. The quality criteria are discussed below, together with the interpretation of the sequencing results in an overall Biological context.

Although sequence-quality checks are performed at the sequencing service provider (BGI) some of the transcripts were aligned to similar single sequences of RNA available on the NCBI database in order to provide an independent assessment. From 20 RNA sequences of MHC class I and II antigens and interferon gamma, it was found that two of the transcripts scored 100% similarity and one achieved 89%. The similarity score is limited by the number sequences available on public databases and might not be statistically significant. However, still two out of the three aligned transcripts are identical despite possible differences in techniques and sequencing bias.

The annotation generated from different databases provide different perspectives on the sequencing data in terms of similarity to known genes, functionality and the predicted roles the sequenced transcripts play in physiological and pathological pathways.

After assembly, 95,245 transcripts were found to have significant similarities in six databases. In the NCBI non-redundant database, all the significant similarities between each two sequences were under the E-value of E≤1.0E-5 and above which the alignments considered insignificant. Over 70% of the alignments showed a similarity of ≥ 80% between the transcripts and the aligned sequences. More than 70% of the aligned sequences match those of genes from six mammals: *Ailuropoda melanoleuca* (giant panda), *Mustela putorius furo* (ferret), *Canis lupus familiaris* (Dog), *Homo sapiens*

(human), *Sus scrofa* (wild boar) and *Bos taurus* (cow). Other species combined represented 29.05% of the total annotation.

Recent studies suggest that increasing taxon sampling can enhance phylogenetic accuracy and resolution (Flynn et al., 2005) and more informative characters are required to confidently resolve close species relationships when studying phylogenetic relationships (Fulton and Strobeck, 2006). In that sense, even though the transcriptome data places the badger under the class “Mammalia” the spectrum of species that emerge as closest matches, indicate the need for more similar transcriptome and genome sequencing. However, when considering the availability of data at the time of transcriptome annotation, the European badger (M. meles) still holds its position within the superfamily “Musteloidea” and the family “Mustelidae” which can be attributed to the sequenced ferret (Mustela putorius furo) transcriptome (Bruder et al., 2010).

Similarly, over 45% of the annotated transcripts places the badger under infraorder “Arctoidae” of the order “Carnifora” which can be explained by the extensive research on the endangered giant panda (Ailuropoda melanoleuca) whose mitogenome (Peng et al., 2007), genome (Angelia et al., 2010) and transcriptome (Du et al., 2015b) have been sequenced. Moreover, the family “Ursidae” has been characterized by rapid radiation events, making phylogenetic inference of the species relationships difficult, and thus, often contentious (Fulton and Strobeck, 2006) and this may also affect the taxonomic relationship between the two families “Mustelidae and Ursidae”.

A search in Cluster of Orthologous Groups (COG) database was performed In order to uncover the homologous relationships of the transcriptome to the well-known conserved

domains. Most of the transcripts were matched to the major groups of (Cellular Processes and Signalling), (Information Storage and Processing) and (Metabolism). However, there were also some transcripts classified as “Poorly Characterized” including 22,329 transcripts under “General function prediction only” and 11,305 transcripts under “Function unknown” categories.

Transcripts were also mapped against all known GO terms to obtain a consistent description of all their corresponding genes and gene products, which may facilitate the future update of their characteristics and functions. In total, the 62 GO terms have 1,028,340 corresponding transcripts and the majority of transcripts (499,918) are assigned to “Biological Processes” GO terms, 382,212 transcripts to “Cellular components” and 146,210 transcripts to “Molecular functions”. KEGG pathway annotation led to the association of 69,924 transcripts to 259 biological pathways including those of biomolecules metabolism, drug metabolism and pathological pathways such as cancers and infection.

3.4.2 Venn diagrams

The distribution of annotated transcripts among databases and the number of shared annotated transcripts appears to be database-dependent rather than a function of the nature of the annotated transcripts in a non-target gene study. In another word, a database that basically depend on annotation based on simple sequence similarity like NR would normally produce more annotations than one that is manually curated (Swiss-Prot), pathway specific (KEGG) or COG which compare orthologs and paralogs of genes. A similar distribution pattern was observed in other transcriptome analysis studies of

animals like the panda (Du et al., 2015b) or even trees like *Hevea brasiliensis* (Para Rubber Tree) (Fang et al., 2016).

3.4.3 Phylogenetic trees

Three transcripts were used to construct three different phylogenetic trees as shown in the results. These transcripts were the most abundant in terms of raw reads, assembled transcript and the longest assembled protein sequence in the transcriptome analysis.

A phylogenetic (evolutionary) tree is a representation of the evolutionary relationships among a set of organisms known as taxa. The tips of the tree represent groups of descendent taxa (species) and the nodes on the tree represent the common ancestors of those descendants. Two descendants that split from the same node are called sister groups.

All three examples showed that the closest relative sequence of the query transcript sequence is either from the same animal (M. meles) or a member from the same family (*Mustelidae*).

Using sequences with highly conserved regions such as haemoglobin alpha subunit and cytochrome c oxidase might be useful in comparing deep phylogenies among close species. However, caution is required when interpreting the chronological order of speciation and divergence. Such an approach might provide satisfactory results in distantly related taxa, but suffers from a number of issues when dealing with evolutionary relationships at shallow time depths (Nater et al., 2015).

It is important to recognise that if the phylogenetic tree is computed from data coming from a single gene (gene tree) is sometimes different from a whole-genome tree (species tree). One of the important factors that cause this difference is genetic polymorphism in the ancestral species (Pamilo and Nei, 1988). Even under standardised alignment parameters different genes provide different levels of speciation leading to formation of different trees. This could be useful in studying the evolution of a single gene and its variants but also could be distracting when studying a whole organism despite the fact that the alignments presented here consistently showbadger or ferret is the closest because this is in part true but is also a consequence of the sequences deposited and searchable in the database.

It has been shown that a combination fossil record, observation, ecological and biological studies in addition to multi-genic phylogeny can describe the classification of a family such as Mustelidae and the position of a species such as the European badger in that family (Koepfli et al., 2008b). Nevertheless, these data still show a relative degree of similarity in accord with the multigene phylogeny of the Mustelidae study (Koepfli et al., 2008b) in respect of species divergence.

3.4.4 MHC transcripts

On investigating the abundance of MHC genes in the badger transcriptome, the sequences data files for both nucleotide and protein were converted into a searchable database. 20 nucleotide sequence and 25 protein sequence achieved respectively 195 and 248 alignments with nucleotide and protein sequences from the transcriptome database. The aligned transcripts were previously annotated (during transcriptome assembly by

BGI) as MHC class I, class II and interferon gamma from several mammals including badger *(Meles meles*) and other closely related or more studied mammals.

MHC genes are highly diverse both within species and among species populations of mammals and the confirmation that alignment of the MHC query sequence with transcripts that have been previously annotated as MHC transcripts, is a possible indicator for both similarity and diversity of badgers MHC transcripts. However, one of the limitations of searching a local database (transcriptome) is that it gives an overly significant e-value as it is calculated within the local database, in another word, it calculates the chances of the alignment occurring by chance with other transcripts.

The identity range of the first hits in the protein alignment is between 74% and 100% which, despite the length of both query and subject sequences, showed regions of highly conserved sequences which showed the intra-species similarities and few regions of less conserved regions witch may reflect the diversity of MHC genes. However, this would require corroboration by sequencing the full length gene of interest.

Comparing the abundance of MHC transcripts between badger and giant panda blood transcriptomes can provide a crude measure of the differences, as the two experiments are different to some extent in terms of different living conditions of both animals (laboratory controlled vs captivity) which might affect the level of exposure to environmental pathogens. Moreover the experimental design (number of samples, pipeline, statistical analysis, etc.) may also affect the outcome. Despite this, the use of FPKM may overcome some the obstacles in the experimental differences as it not just

calculate the crude number of transcripts for a given gene but also gives more statistical sense by dividing that number by the actual length (in kilo-base) of that gene then by million fragments mapped.

In this respect, comparing the highly abundant MHC transcripts from both mammals showed that the badger has higher expression levels of both classes of MHC genes. Although, this result may require further experimental investigation it may reflect the findings of some researches where MHC genes diversity is much lower in this endangered species (Zhang et al., 2015) and diversity of MHC genes in the giant panda was relatively lower than other vertebrates (Zhu et al., 2007).

Chapter IV

4 A comparison between male and female transcriptome

Biological differences between males and females can affect susceptibility to mycobacterial infection, and these differences may provide valuable insight into the components that constitute an effective immune response to this pathogen (Nhamoyebonde and Leslie, 2014). Infectious diseases rarely affect males and females equally. As a general rule, females present a more-robust immune responses to antigenic challenges, such as infection and vaccination, than males (Griesbeck et al., 2015). Sex hormones have diverse effects on many immune cell types, including B cells, T cells, neutrophils, dendritic cells, macrophages, and natural killer cells (Fish, 2008). In badgers, differences in disease prevalence between males and females may affect differences in the physiological impact of infection and may also have significant effects on pathogen persistence within the host population (Tomlinson et al., 2013). It has been suggested that the resilience of female badgers to the negative effects of bTB infection may favour its persistence. Females not only exhibited enhanced survival, but also continued to reproduce successfully, whilst excreting bacteria. This suggests that infected females may play a pivotal role in the maintenance of infection in social groups, particularly through transmission to their cubs (Tomlinson et al., 2013).

In this chapter, the over-arching differences in transcription in terms of a whole transcriptome and also for a limited subset of immune/bTB related transcripts, are explored. The most significantly affected transcripts are discussed, and a small number of

examples of immunologically important pathways with up/down regulated genes were also included.

Fragments per kilo-base per million or FPKM is a normalized count for the sequencing depth (million): sequencing runs with greater depth will have more reads mapped to each gene. FPKM is also normalized for the length of the original gene (expressed in kilo-bases), as longer genes will have more mapped transcripts. The female transcriptome was used as a reference for the comparison, i.e. up regulated transcripts means that they are expressed at higher levels in male and down regulated transcripts means that they are expressed at lower levels in the female badger. The fold change (FC), or difference in expression was calculated as a log2, i.e. positive FC for up regulated and negative FC for down regulated transcripts. The false discovery rate (FDR) which calculates the expected proportion of difference in a comparison (between male and female transcripts counts in this case) that is false, thereby ensuring that the differences in expression are statistically insignificant.

4.1 Results

As mentioned above, male and female FPKM were compared. 2690 transcripts were found to be upregulated in the male badger and 18691 transcripts were downregulated (**Figure 4.1**). A lower FC threshold of ± 1.5 was applied to highlight a significant expression level difference (Hausen et al., 2015). The FDR (Benjamini and Hochberg, 1995) cutoff value was set to ≤ 0.001 as used by most studies of this kind. In (**Figure 4.2**) the majority

of differentially expressed transcripts were in the region of 5 to -5 FC which represents log2 of the male FPKM/Female FPKM.

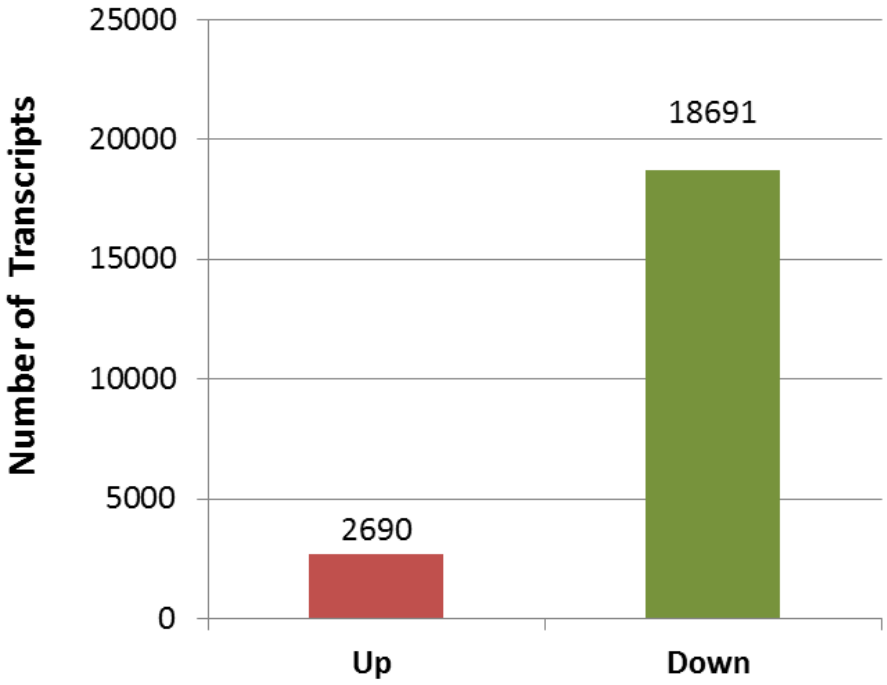


Figure 4.1: Numbers of differentially expressed transcripts (FDR ≤ 0.001) in male badger's transcriptome compared to the female transcriptome with fold change cut-off of ± 1.5.

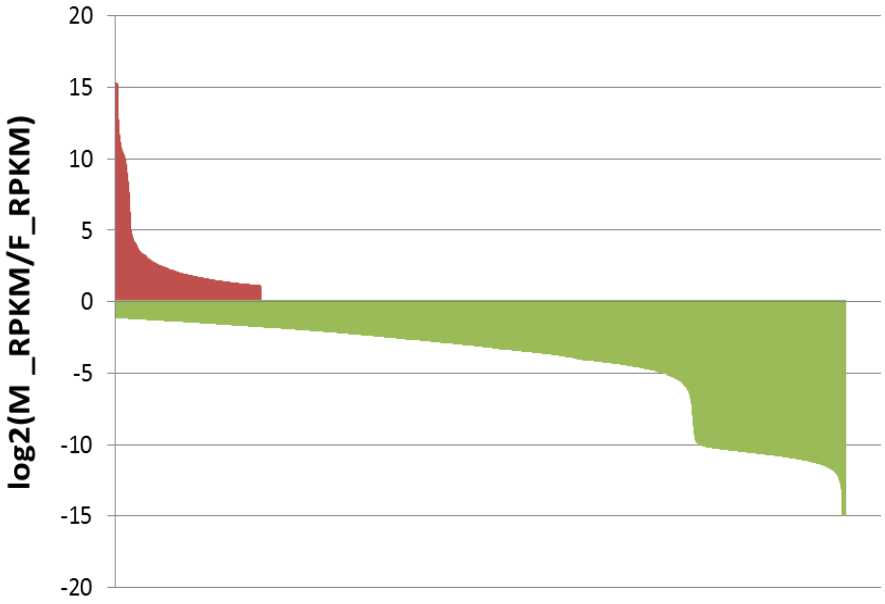


Figure 4.2: the expression levels of differentially expressed transcripts in male transcriptome compared to the female transcriptome. The expression level was presented as Log2 of male FPKM / female FPKM ratio.

4.1.1 The major differentially regulated transcripts

Tables 4.1 and 4.2 show top 10 upregulated and down regulated transcripts and their annotation in male badger transcriptome when compared to the female transcriptome.

	Transcript Nr-annotation	FC	FDR
1	Eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked, isoform CRA_c [*Mus musculus*]	15.18	2.5E-205
2	Lysine (K)-specific demethylase 5C [*Bos taurus*]	14.54	1.3E-97
3	Probable ubiquitin carboxyl-terminal hydrolase FAF-Y [*Bos taurus*]	13.97	2.4E-157
4	PREDICTED: eukaryotic translation initiation factor 1A, Y-chromosomal-like [*Papio anubis*]	13.32	8.9E-29
5	Lysine -specific demethylase 6A [*Mustela putorius furo*]	13.28	8.7E-41
6	Probable ubiquitin carboxyl-terminal hydrolase FAF-Y [*Bos taurus*]	12.95	2.6E-45
7	PREDICTED: ATP-dependent RNA helicase DDX3X isoform 3 [*Saimiri boliviensis boliviensis*]	12.93	4.7E-256
8	Environmental lipopolysaccharide-responding gene protein [*Macaca fascicularis*]	12.25	7.1E-17
9	PREDICTED: synaptonemal complex protein 3-like [*Sus scrofa*]	12.21	2.0E-20
10	Tetratricopeptide repeat protein [*Canis lupus familiaris*]	12.21	1.4E-30

Table 4.1: Table of 10 most upregulated transcripts and their annotation in male badger transcriptome compared to the female transcriptome.

	Transcript Nr-annotation	FC	FDR
1	PREDICTED: uncharacterized protein LOC101053426 [*Saimiri boliviensis boliviensis*]	-14.88	2.5E-99
2	Immunoglobulin heavy chain variable region subgroup 1 [*Felis catus*]	-14.61	6.0E-162
3	Immunoglobulin lambda light chain variable region [*Homo sapiens*]	-13.78	1.8E-40
4	Immunoglobulin G heavy chain variable region [*Homo sapiens*]	-13.58	3.1E-37
5	Immunoglobulin heavy chain [*Homo sapiens*]	-13.49	3.1E-37
6	Immunoglobulin heavy chain variable region [*Homo sapiens*]	-13.43	3.4E-30
7	Immunoglobulin kappa light chain V-J region [*Equus caballus*]	-13.15	1.3E-29
8	Hypothetical protein [*Trypanosoma cruzi strain CL Brener*]	-13.02	5.3E-34
9	Nucleolar RNA-binding protein [*Trypanosoma cruzi strain CL Brener*]	-12.87	3.8E-28
10	Hypothetical protein PANDA_022421 [Ailuropoda melanoleuca]	-12.85	9.3E-24

Table 4.2: Table of most 10 downregulated transcripts and their annotation in male badger transcriptome compared to the female transcriptome.

4.1.2 Gene ontology for differentially expressed transcripts

Gene Ontology (GO) is an international standardized gene functional classification system which offers a dynamic, updated and controlled vocabulary, and a strictly defined rule-set to comprehensively describe the properties of genes and their products in any organism. GO has three ontologies: biological process, molecular function and cellular component. The basic unit of GO is the GO-term. Every GO-term belongs to a type of ontology. GO functional analysis provides a functional classification and annotation for differentially expressed genes as well as GO functional enrichment analysis for differentially expressed genes. GO functional classification annotation provides a gene list and gene numbers for each GO term.

GO functional enrichment analysis **(Figure 4.3)** provides GO terms that are significantly enriched in differentially expressed genes, showing which differentially expressed genes are connected to their respective biological functions. The analysis firstly maps all differentially expressed genes to GO terms in the database (http://www.geneontology.org/), calculates gene numbers for every term then finds significantly enriched GO terms in differentially expressed genes comparing to the database background.

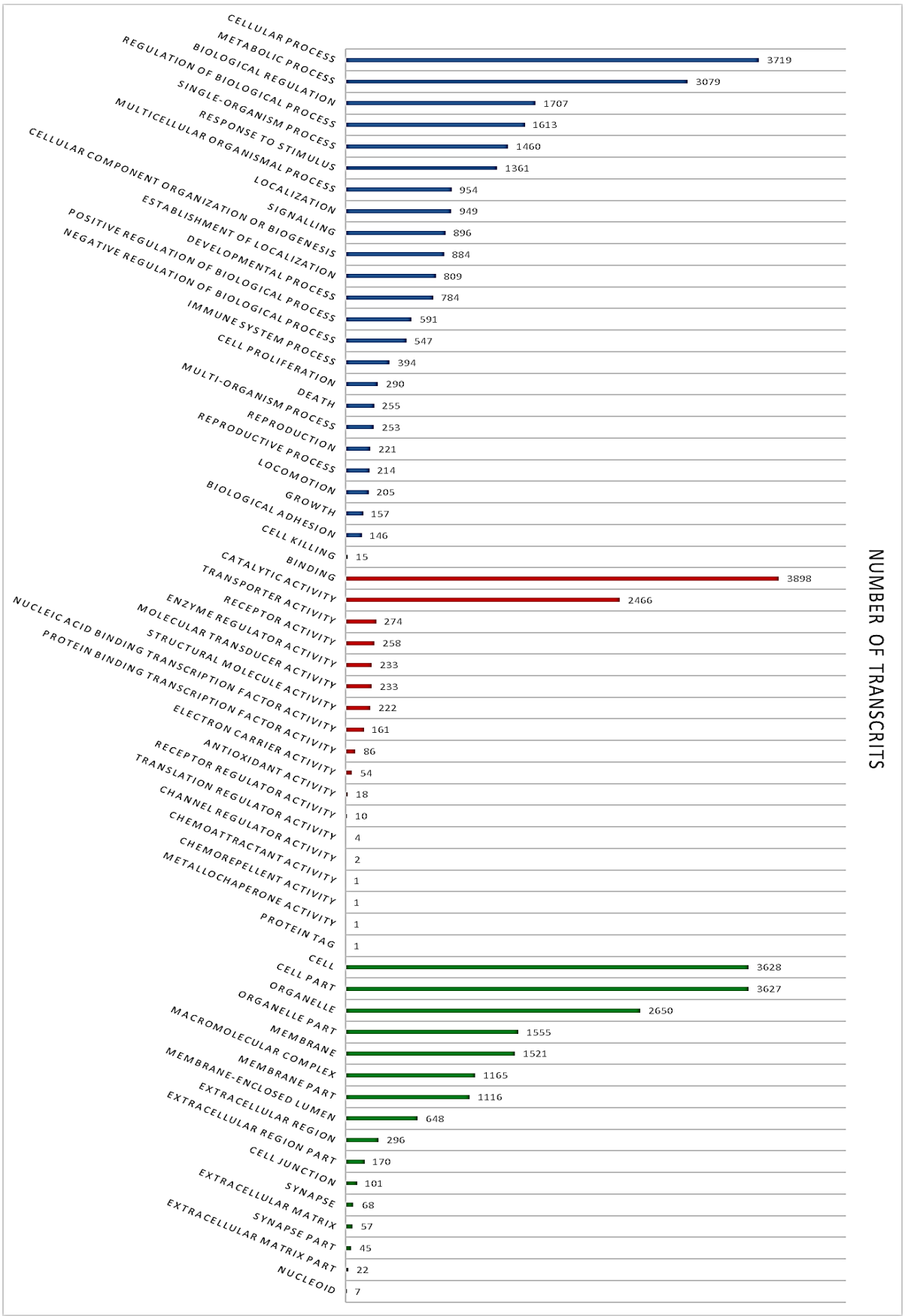

CELLULAR PROCESS 3719
METABOLIC PROCESS 3079
BIOLOGICAL REGULATION 1707
REGULATION OF BIOLOGICAL PROCESS 1613
SINGLE-ORGANISM PROCESS 1460
RESPONSE TO STIMULUS 1361
MULTICELLULAR ORGANISMAL PROCESS 954
LOCALIZATION 949
CELLULAR COMPONENT ORGANIZATION OR BIOGENESIS 896
SIGNALLING 884
ESTABLISHMENT OF LOCALIZATION 809
DEVELOPMENTAL PROCESS 784
POSITIVE REGULATION OF BIOLOGICAL PROCESS 591
NEGATIVE REGULATION OF BIOLOGICAL PROCESS 547
IMMUNE SYSTEM PROCESS 394
CELL PROLIFERATION 290
MULTI-ORGANISM PROCESS 255
DEATH 253
REPRODUCTION 221
REPRODUCTIVE PROCESS 214
LOCOMOTION 205
GROWTH 157
BIOLOGICAL ADHESION 146
CELL KILLING 15
BINDING 3898
CATALYTIC ACTIVITY 2466
TRANSPORTER ACTIVITY 274
RECEPTOR ACTIVITY 258
ENZYME REGULATOR ACTIVITY 233
MOLECULAR TRANSDUCER ACTIVITY 233
STRUCTURAL MOLECULE ACTIVITY 222
NUCLEIC ACID BINDING TRANSCRIPTION FACTOR ACTIVITY 161
PROTEIN BINDING TRANSCRIPTION FACTOR ACTIVITY 86
ELECTRON CARRIER ACTIVITY 54
ANTIOXIDANT ACTIVITY 18
RECEPTOR REGULATOR ACTIVITY 10
TRANSLATION REGULATOR ACTIVITY 4
CHANNEL REGULATOR ACTIVITY 2
CHEMOATTRACTANT ACTIVITY 1
CHEMOREPELLENT ACTIVITY 1
METALLOCHAPERONE ACTIVITY 1
PROTEIN TAG 1
CELL 3628
CELL PART 3627
ORGANELLE 2650
ORGANELLE PART 1555
MEMBRANE 1521
MACROMOLECULAR COMPLEX 1165
MEMBRANE PART 1116
MEMBRANE-ENCLOSED LUMEN 648
EXTRACELLULAR REGION 296
EXTRACELLULAR REGION PART 170
CELL JUNCTION 101
SYNAPSE 68
EXTRACELLULAR MATRIX 57
SYNAPSE PART 45
EXTRACELLULAR MATRIX PART 22
NUCLEOID 7
NUMBER OF TRANSCRITS

Figure 4.3: Gene Ontology annotation of differentially expressed transcripts. GO biological process (blue), GO molecular function (red) and GO cellular components (green).

1.1.1 KEGG pathways for differentially expressed transcripts

Individual genes often cooperate with each other to fully express their biological functions. Pathway-based analysis helps to further understand the biological functions associated with specific genes. KEGG is the major public pathway-related database (Kanehisa et al., 2008) operating in this analysis space. Pathway enrichment analysis identifies significantly enriched metabolic pathways or signal transduction pathways in differentially expressed genes comparing with the whole genome background. A pathway with an FDR score of ≤ 0.05 is called significantly enriched in a group of differentially expressed genes (Nhamoyebonde and Leslie, 2014). There were 83 significantly enriched KEGG pathways where FDR ≤ 0.05 and an example is shown in **(Figure 4.4).**

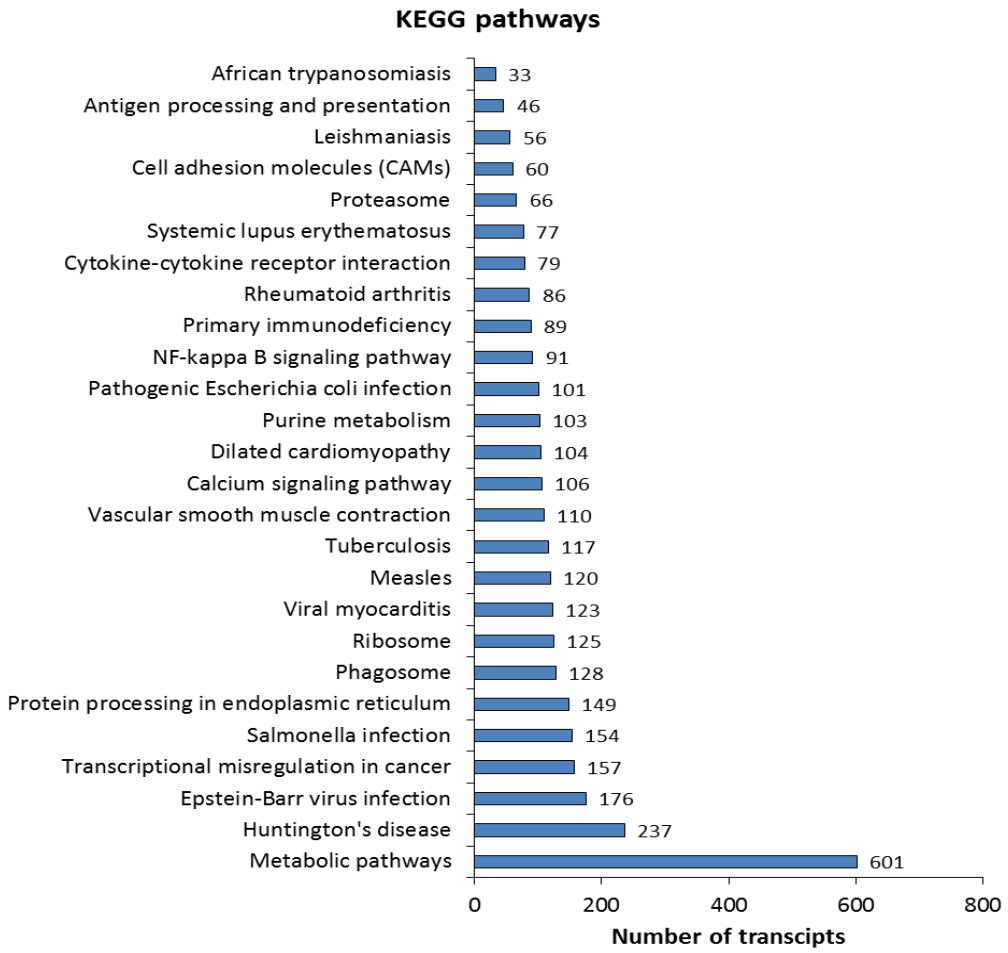


Figure 4.4: KEGG pathway for differentially expressed transcripts.

4.1.3 Difference in expression of bTB related transcripts

There were 117 transcripts that encode candidate genes involved in the bTB pathway, with a significant difference in expression levels between male and female badgers. 77 transcripts were downregulated in the male badger. Immunoglobulin heavy chain variable regions annotated as genes from human and cats were the main down regulated transcripts in the male badger **(Table 4.3)**

Nr-annotation	log2 (Male RPKM / Female RPKM)
immunoglobulin heavy chain variable region subgroup 1 [*Felis catus*]	-14.6079
immunoglobulin G heavy chain variable region [*Homo sapiens*]	-13.576
immunoglobulin heavy chain [*Homo sapiens*]	-13.4881
immunoglobulin heavy chain variable region [*Homo sapiens*]	-13.4288
immunoglobulin heavy chain variable region subgroup 3 [*Felis catus*]	-12.589
immunoglobulin heavy chain variable region [*Homo sapiens*]	-12.4574
immunoglobulin heavy chain variable region [*Homo sapiens*]	-12.1266
immunoglobulin heavy chain variable region subgroup 3 [*Felis catus*]	-11.9554
immunoglobulin heavy chain variable region [*Homo sapiens*]	-11.7744
immunoglobulin heavy chain variable region subgroup 1 [*Felis catus*]	-11.4248

Table 4.3: Most downregulated bTB related transcripts in the male badger

40 transcripts were upregulated in the male badger. The most upregulated transcripts (**Table 4.4)** correspond to TB pathway related genes from different species.

Nr-annotation	log2 (Male RPKM / Female RPKM)
PREDICTED: nuclear transcription factor Y subunit gamma isoform 3 [*Equus caballus*]	2.1209
PREDICTED: interleukin-23 subunit alpha [*Canis lupus familiaris*]	2.3726
PREDICTED: galectin-3-like [*Oreochromis niloticus*]	2.3844
PREDICTED: HLA class II histocompatibility antigen, DO alpha chain-like [*Ailuropoda melanoleuca*]	2.3909
PREDICTED: bcl2 antagonist of cell death-like [*Ailuropoda melanol*euca]	2.6656
PREDICTED: toll-like receptor 2-like [*Ailuropoda melanoleuca*]	2.8286
immunoglobulin mu heavy chain [*Pteropus alecto*]	3.0128
PREDICTED: LOW QUALITY PROTEIN: complement receptor type 1-like [*Equus caballus*]	3.7766
IgM heavy chain VH1 region precursor [*Homo sapiens*]	8.7529
PREDICTED: hypothetical protein LOC467582 [*Pan troglodytes*]	11.5732

Table 4.4: Most upregulated bTB related transcripts in the male badge

Differential expression of bTB related transcripts was also illustrated in the KEGG pathway of tuberculosis (**Figure 4.5**), providing a clearer picture of the involvement of differentaily expressed transcripts at different stages of disease. The pathway shows transcription upregulation in genes involved in the JAK-STAT signalling cascade in the male badger. The JAK-STAT cascade is involved in nuclear transcription of genes involved in inflammatory response to an infection and antigen processing and presentation. The **Figure 4.5** also shows a down regulation of transcripts involved in MAPK signalling cascade which affects cell proliferation and proinflammatory cytokine production. There was also a downregulation of transcripts involved in (NOD)-like receptor signalling cascade which is important in the regulation of the host innate immune response (Franchi et al., 2009). There was also a dowonregulation of several transcripts involved in phagcytosis and phagosome-lysosome fusion.

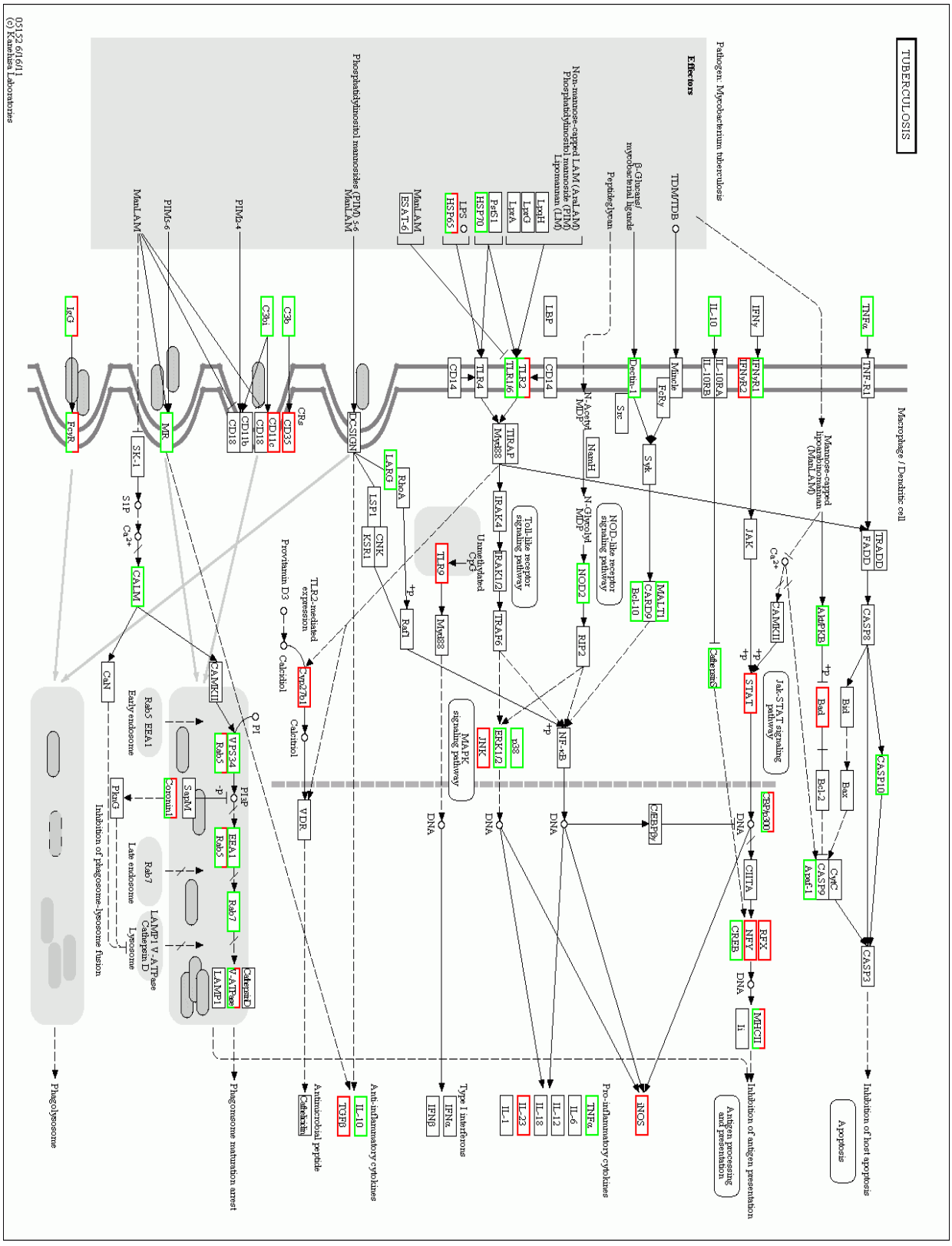


Figure 4.5: Differential expression of bTB related transcripts for male compared to female badger illustrated on KEGG pathway of tuberculosis. Upregulated (red) and downregulated (green)

4.2 Discussion

Gene-expression variation may play a significant role in gender-specific health disparities, probably through upregulating or downregulating genes within physiological pathways. Males and females could have gender-specific transcriptional or translational regulation, leading to differential mRNAs or protein products for some genes (Zhang et al., 2009). Sex-biased gene expression has been observed in a number of mammals including humans, in different tissues that are thought to be phenotypically similar, including the brain were it involves about 2.5% of all expressed genes (Trabzuni et al., 2013). Similar observations were also recorded in cell lines were gender-specific differential expression in lymphoblastoid cell lines in male and female human were studied (Zhang et al., 2007). In animal models, genes in the pathway of bone development were shown to be differentially expressed between the male and female of mice (Huang et al., 2014). Sexual dimorphism is a term used to describe the difference in characteristics - other than sexual organs – such as body size, shape and colour between male and female of the same species. Transcriptome analysis has showed that sexual dimorphism in gene expression was much greater than previously recognized as thousands of genes showed sexual dimorphism in liver, adipose, and muscle where the sexually dimorphic genes were also found to be highly tissue-specific (Yang et al., 2006). In this chapter, 12.6% of transcripts were found upregulated in the male badger and 87.4% transcripts were downregulated when compared to the female badger. The highest 10 upregulated transcripts were found to be involved in cellular functions such as protein synthesis and folding, vesicular trafficking, cytoskeletal assembly and regulation of cell growth and apoptosis which may

be attributed to many physiological and environmental factors that can cause this sex-biased expression such as levels of sex hormones and growth hormones as shown in other studies that showed sex-differences in gene expression can be dependent on the hormonal status (Jansen et al., 2014). Whereas, the 10 most downregulated transcripts in the male badger are of unknown specific function (No. 8 and 9 in table 4.2) or of critical regulators of immunity, stress responses, apoptosis differentiation and cell signaling. In general, topographical classification of the highest 10 differentially expressed transcripts in both tables has showed that upregulated transcripts encode more often, intracellular functions, in particular in the cytoplasm; whereas down regulated transcripts generally encode integral membrane proteins, which may reflect the source of the transcriptome data: human peripheral blood cells (Jansen et al., 2014).

However as stated by (Schurch et al., 2016): “It is worth recalling that identifying a gene as significantly differentially expressed does not necessarily equate to identifying it as biologically significant and that it is important to consider both the magnitude of the measured fold change and existing biological knowledge alongside the statistical significance when inferring a biological significance for the results of gene expression difference experiments” (Schurch et al., 2016).

Gender is one variable that influences innate and adaptive immune responses to foreign and self-antigens in infections, malignancies and as a response to vaccination. It shows distinctions in innate and adaptive immune responses where specific immunological sex-biased responses are present throughout life and are sex-hormone independent, whereas others appear only after the age of puberty which may suggest that both genes and

hormones are involved (Klein and Flanagan, 2016). In infectious diseases, human males tend to be more susceptible to some infections such as Ebola, hepatitis B and tuberculosis whereas females are more susceptible to other infections like malaria, influenza and toxoplasmosis (Klein and Flanagan, 2016).

In humans, mycobacterial infection tends to be gender biased and males exhibit a higher prevalence of tuberculosis infection even if confounding factors, such as social and economic factors, awareness, exposure and are taken into account in epidemiological studies (Neyrolles and Quintana-Murci, 2009). In other mammals such as brush-tail possums, bTB prevalence is often substantially higher in males than in females (Nugent et al., 2015). In cattle, males tended to show higher prevalence rates for bTB (Moiane et al., 2014). Female badgers are more resilient to established bTB infection than male badgers, with longer survival times following the detection of bacterial excretion (Tomlinson et al., 2013).

The majority of immune cells express specific receptors for sex hormones and are responsive to changes in hormone levels. Females pose a more-robust set of immune responses to antigenic challenges, such as infection and vaccination, than males. Sex hormones have diverse effects on many immune cell types, including dendritic cells, macrophages, B cells, T cells, neutrophils and natural killer cells **(Figure 4.6)** (Nhamoyebonde and Leslie, 2014).

Although it cannot be anticipated whether the observed differential expression between male and female badgers is a direct result of the hormonal effect, and despite the fact

that these data were from uninfected animals, some of the expression levels difference were consistent with the biological differences between the sexes and their susceptibility to tuberculosis model by Nhamoyebonde and Leslie (2014) in the **figure 4.6.**

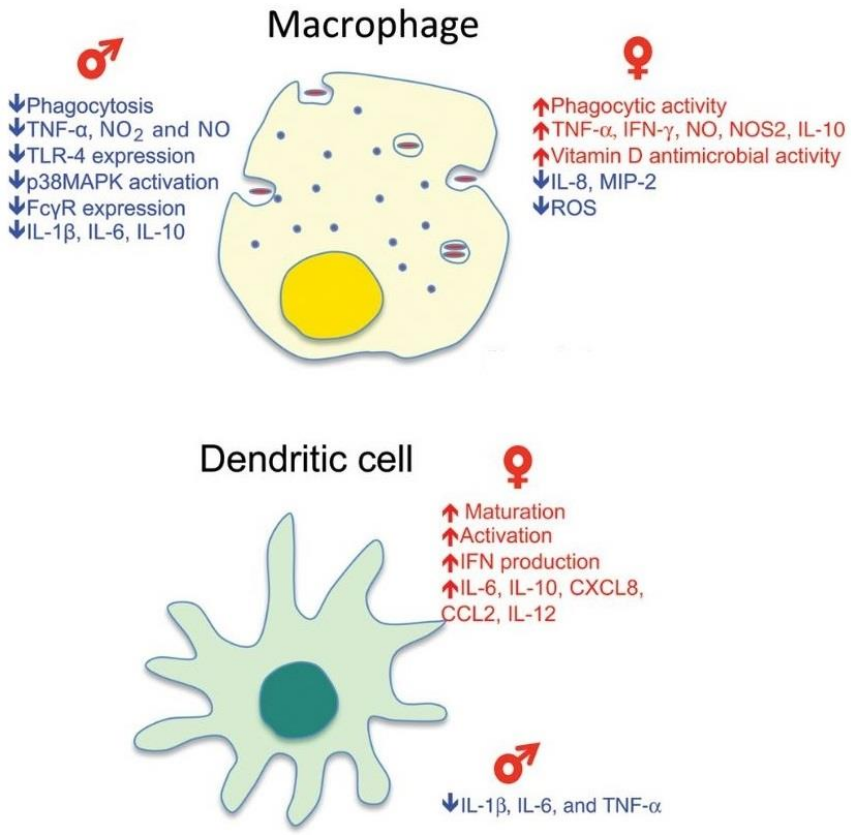


Figure 4.6: The effects of sex hormones on the modulation of immune responses of macrophages and dendritic cells to tuberculosis. Adapted from (Nhamoyebonde and Leslie, 2014)

As suggested by the KEGG pathway for tuberculosis infection in macrophages and dendritic cells**,** there was an observed downregulation of the tumour necrosis factor superfamily, member 2 (TNFA), Toll-like receptor 4 (TLR4), mitogen-activated protein

kinases (MAPK). However, there was a mixed set of expression levels amongst the FC receptors, Interleukins and an upregulation of nitric oxide synthase activity (NOS), unlike the findings in this **figure 4.6**. This differences in expression may lead to differences in the capability of detection and elimination of mycobacterial cells from the host body through phagocytosis, bactericidal action and granuloma formation.

There was also upregulation of interferon gamma receptor 2 (IFNGR2) and downregulation of interferon gamma receptor 1 (IFNGR1) and interleukin 8 (CXCL8) which might affect the migration of macrophages and neutrophils to the site of infection.

In conclusion, the results of the transcriptome analysis and expression difference show some consistencies with some aspects of the observed patterns of the body responses to tuberculosis, despite the fact that the studied animals are bTB. It is also consistent with the epidemiological and observational findings in which males are found to be more prone and less resistant to TB and bTB infection. Taking gender and the influence of sex hormones into account may help in improving our understanding of the immune response required to prevent and control tuberculosis.

Chapter V

5 Immunity related and bTB specific transcripts

This chapter describes the immunity related transcripts in terms of numbers and classes of the immunity genes they represent. It also illustrates some of the key genes that play a potential role in the immune response to infection and the evolutionary relatedness of the badger to other mammals as revealed by a bioinformatic analysis. Finally, it also describes the key genes in the bTB pathway (as per KEGG) and the matching transcripts from the analysis of transcriptome data from the badger.

5.1 Immunity-related transcripts

Several search criteria were applied to extract the immuno-components of the badger's transcriptome. After transcriptome assembly and annotation, a basic (crude) search was performed using the prefix (immun-) as a search term in the four databases and the numbers are shown in the **Table 5.1** below.

Database	NR	Swiss-Prot	COG	KEGG
Number of annotations	358	194	0	266

Table 5.1: Number of transcripts with (immun-) term in each database annotation

The distribution of the annotations among four different databases is shown in the **Figure 5.1**

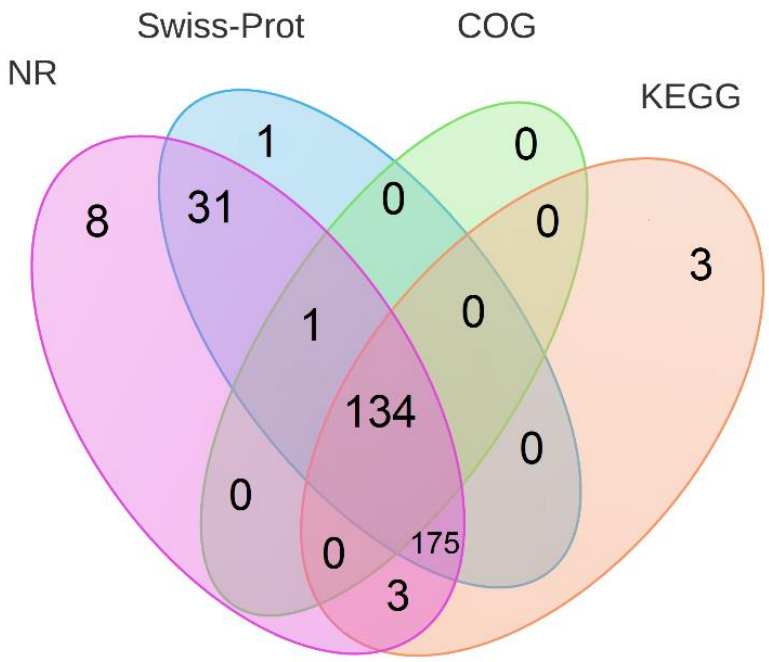


Figure 5.1: Venn diagram of distribution of the annotations among four databases

From the **figure 5.1** Eight transcripts were found annotated only in the NR database (suppressive immunomodulating factor, putative [*Trypanosoma cruzi marinkellei*], immunodominant antigen [*Trypanosoma cruzi strain CL Brener*], PREDICTED: immunity-related GTPase family M protein 1-like [*Canis lupus familiaris*], immunoglobulin superfamily, DCC subclass, member 4-like [*Bos taurus*], and leukocyte immunoglobulin-like receptor B2 [*Halichoerus grypus*]). A single transcript was found to be annotated only in Swiss-Prot database (Immunoglobulin lambda-like polypeptide 5 OS=Homo sapiens GN=IGLL5 PE=2 SV=2). Three transcripts were found to be annotated only in KEGG (as human immunodeficiency virus type I enhancer-binding protein and autoimmune regulator) and no transcripts were found annotated in COG database alone.

A more thorough search was performed using the immunity genes list from the IKB database (Ortutay and Vihinen, 2009) as search queries. The search has shown the number of transcripts in all databases align with each gene of **893** immunity genes. Of the 11724 annotations found the 20 most abundant are shown in **table 5.2**, with a comprehensive list of annotations for each gene given in the **appendix.**

Gene	Full name	Number of annotations
WAS	Wiskott-Aldrich syndrome protein.	421
MME	membrane metallo-endopeptidase.	295
CD22	CD22 molecule.	295
CD2	CD2 molecule.	273
TMC8	EVIN2.	187
IL12RB1	interleukin 12 receptor, beta 1 isoform 1 precursor.	171
TRAF3	TNF receptor-associated factor 3 isoform 1.	151
LYST	lysosomal trafficking regulator.	128
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog isoform 1 precursor.	126
CD19	CD19 antigen.	112
CD44	CD44 antigen isoform 5 precursor.	109
CD46	CD46 antigen, complement regulatory protein isoform 12 precursor.	107
ADAM17	ADAM metallopeptidase domain 17 preproprotein.	106
C4B	complement component 4B preproprotein.	97
RFX1	regulatory factor X1.	97
ITGA5	integrin alpha 5 precursor.	88
DKC1	dyskerin.	88
CD5	CD5 molecule.	87
BLNK	B-cell linker.	86
STAT3	signal transducer and activator of transcription 3 isoform 2.	84

Table 5.2: Immunity genes with the most abundant number of annotation found using IKB

5.2 Immunoglobulins

Members of the immunoglobulin family, an important component of the immune system in identifying and tagging foreign antigens for subsequent neutralization, were identified

in each database. The number and annotation of immunoglobulins in the badger transcriptome are shown in **the table 5.3** for three data bases.

Database	NR	NT	Swiss-Prot
Number of immunoglobulin related transcripts	289	460	170

Table 5.3: Number of annotations for immunoglobulin in each database

The most abundant annotations of immunoglobulins and immunoglobulin receptors were also extracted and shown in the **table 5.4** below.

Gene	Full name	Number of annotations
IGSF8	immunoglobulin superfamily, member 8.	48
LAIR1	leukocyte-associated immunoglobulin-like receptor 1 isoform a precursor.	11
LILRB2	leukocyte immunoglobulin-like receptor, subfamily B, member 2 isoform 1.	10
LILRB3	leukocyte immunoglobulin-like receptor, subfamily B, member 3 isoform 2.	10
VPREB1	immunoglobulin iota chain preproprotein.	7
LILRA5	leukocyte immunoglobulin-like receptor subfamily A member 5 isoform 1.	5
FCGR3B	low affinity immunoglobulin gamma Fc region receptor III-B precursor.	4
IGJ	immunoglobulin J chain.	4
LILRA6	leukocyte immunoglobulin-like receptor, subfamily A, member 6.	4
IGSF2	immunoglobulin superfamily, member 2.	3
LILRA2	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 2.	2
CADM2	immunoglobulin superfamily, member 4D.	1
PILRA	paired immunoglobulin-like type 2 receptor alpha isoform 1 precursor.	1
KIR2DL4	killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 4 isoform a.	1
KIR3DL2	killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 2 precursor.	1
IGHG2	immunoglobulin gamma-2 heavy chain.	1

Table 5.4: most abundant immunoglobulins in the transcriptome

1.1.2 NCBI Non-redundant database annotation of immunoglobulins

The figure 5.2 below shows the number of immunoglobulin sequences from different mammals that aligned with transcripts from the badger's RNA-seq data. Sequences from

the family Carnivora are the highest in terms of number of alignments and sequence similarity.

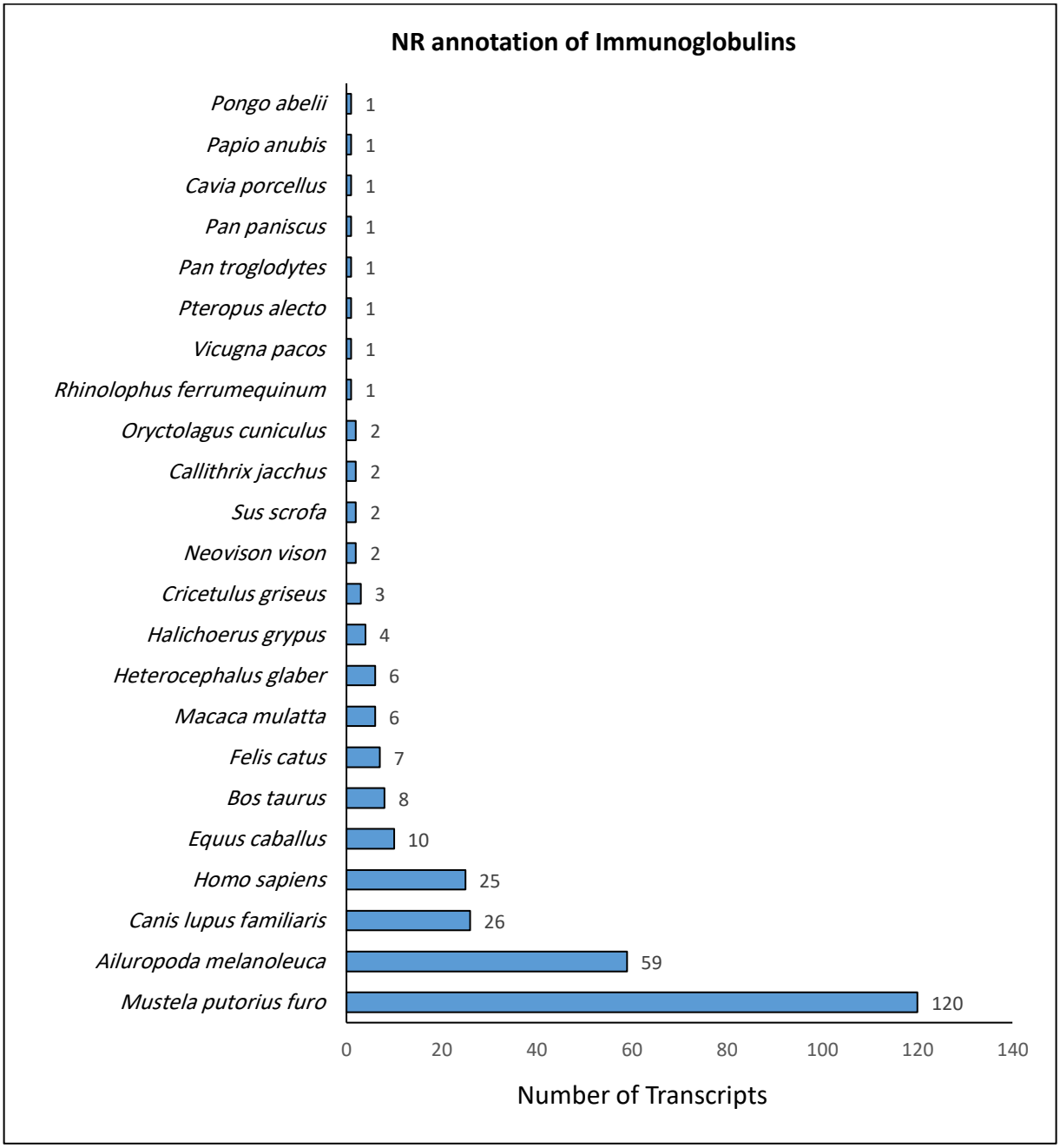


Figure 5.2: NCBI Non-redundant database annotation of immunoglobulins

5.2.1 Swiss-Prot database annotation of immunoglobulins

Using SWISS-PROT (Bairoch and Apweiler, 2000) which is a manually curated protein sequence database can provide a high level of annotation such as the description of the function of a protein with a minimal level of redundancy. In that sense it is naturally that most stored sequences are the most studied ones (human and model animals) and the similarity shown in **the figure 5.3** is mostly driven by function rather than the animal source of that sequence

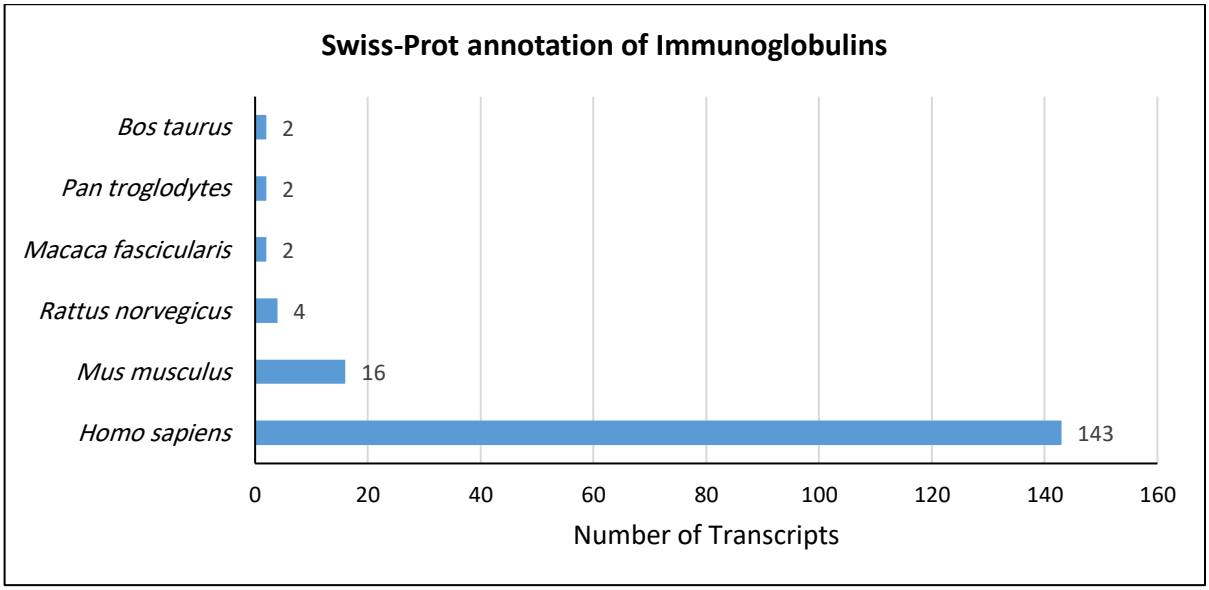


Figure 5.3: Swiss-Prot database annotation of immunoglobulins

5.3 Immunoglobulin variable regions

Because of the importance of the immunoglobulin variable regions for recognition and binding of the antigens, the corresponding sequences of the variable genes from the transcriptome data files were extracted, yielding 75 reads that were annotated in different databases as follows:

5.3.1 NCBI Non-redundant database annotation of variable genes' transcripts

Most of the variable genes transcripts of the badger are similar to those of human and primates' **figure 5.4** and carnivorous animals with four transcripts unaligned with any known variable gene sequence.

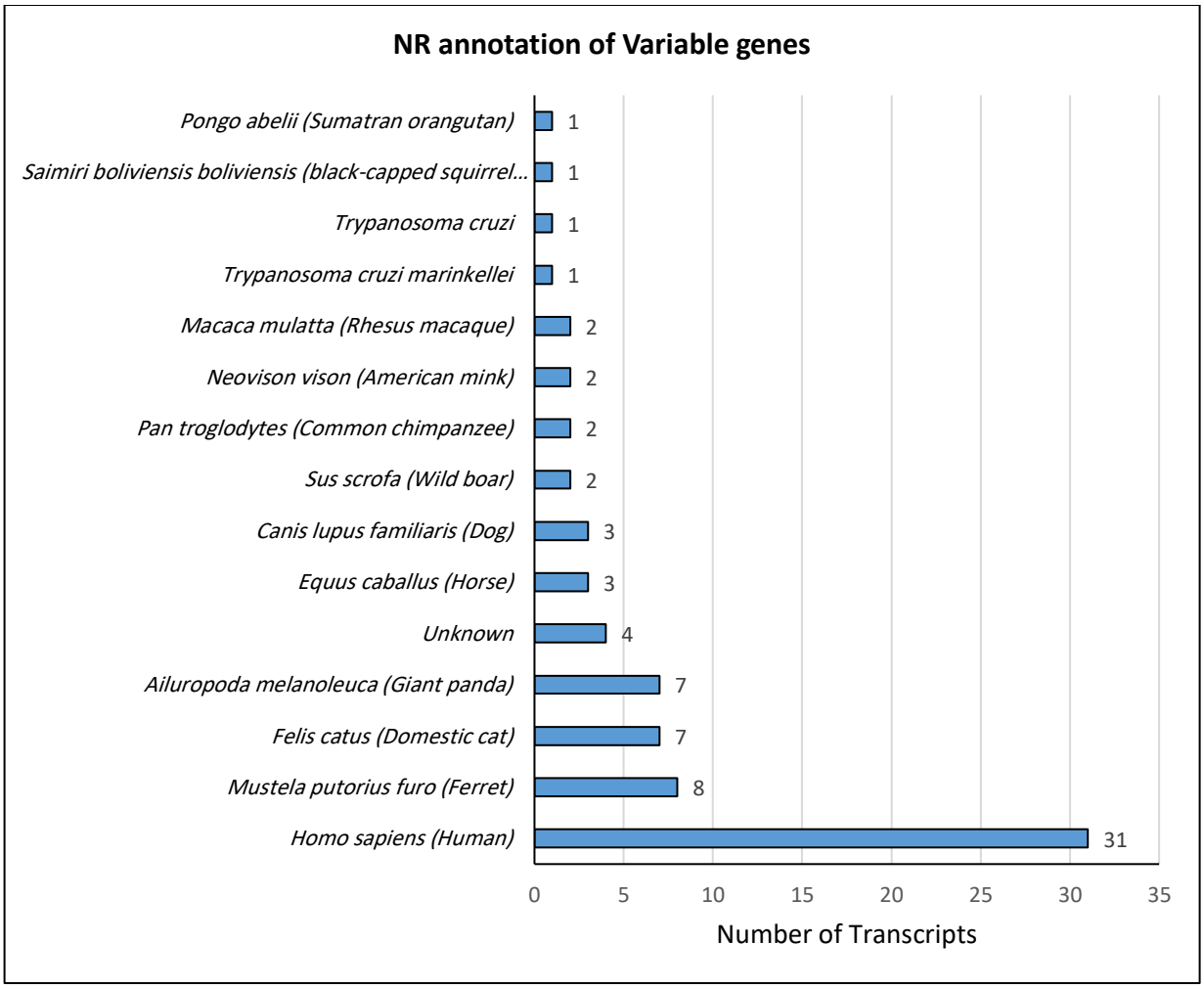


Figure 5.4: NCBI Non-redundant database annotation of variable genes

5.3.2 KEGG annotation of Variable genes

Most of the badger's variable genes transcripts were found to be highly similar to sequences from dog, rhesus monkey, cow and other mammals with exception of 14 sequences that could not be mapped to any KEGG pathway **Figure 5.5**

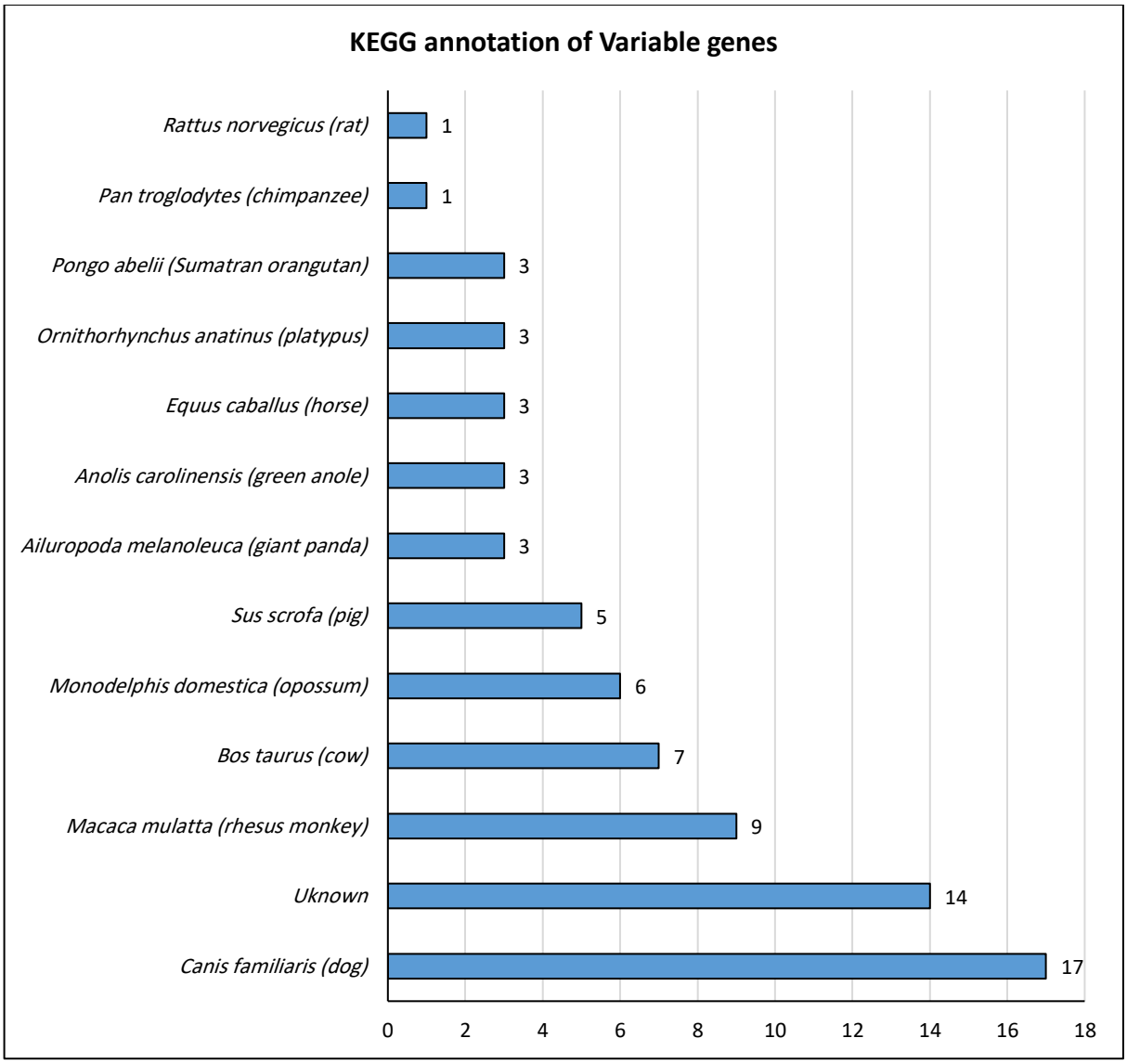


Figure 5.5: KEGG annotation of Variable genes.

5.4 Genes involved in the bTB pathway

Using KEGG pathway for tuberculosis as a reference, 1825 transcripts were extracted from the transcriptome sequencing data that correspond to 147 of the 183 genes that are involved in the bTB pathway from early infection until cell death. Some of the transcripts were found to encode several signalling protein receptors, but not the signalling proteins

themselves. The IKB immunological classification of the transcripts involved in bTB pathogenesis is shown in the **table 5.5.**

IKB Immunological function	Number of transcripts
Phagocytosis	5
CD molecules	23
Chemokines and receptors	30
Cellular immunity	12
Humoral Immunity	13
Antigen processing and presenting	20
inflammation	25
Complement system	6
Transcription factors	6

Table 5.5: IKB immunological classification of transcripts involved in bTB KEGG pathway.

However, number of genes that are involved in bTB pathogenesis were found to have no matching transcripts in the badger's RNA-seq. These genes include interleukins, interferons and transforming growth factors **(table 5.6)**.

IL6	interleukin 6 (interferon, beta 2)	IFNA13	interferon, alpha 13
IFNA1	interferon, alpha 1	IFNA14	interferon, alpha 14
IFNA2	interferon, alpha 2	IFNA16	interferon, alpha 16
IFNA4	interferon, alpha 4	IFNA17	interferon, alpha 17
IFNA5	interferon, alpha 5	IFNA13	interferon, alpha 13
IFNA6	interferon, alpha 6	IFNA21	interferon, alpha 21
IFNA7	interferon, alpha 7	IFNB1	interferon, beta 1, fibroblast
IFNA8	interferon, alpha 8	TGFB2	transforming growth factor, beta 2
IFNA10	interferon, alpha 10	TGFB3	transforming growth factor, beta 3

Table 5.6: IKB Genes without matching sequences in the transcriptome.

1.1.3 NR annotation of TB related transcripts

NR annotation of bTB related transcripts shows that the most similar sequences are from another mustelid (ferret) and two other carnivores (giant panda and dog) **Figure 5.6.**

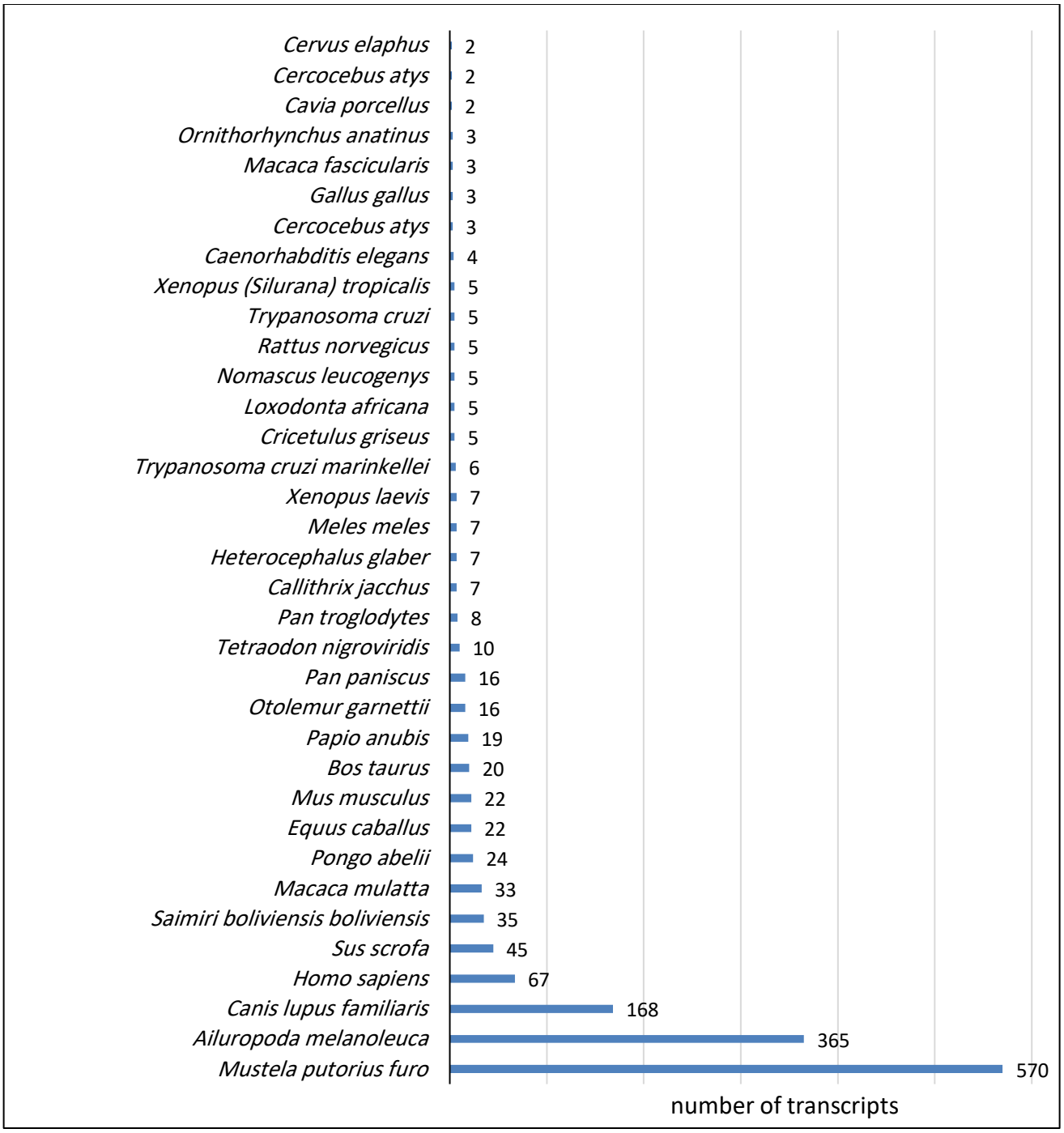


Figure 5.6: NR annotation of TB related transcripts.

5.4.1 KEGG pathway

Host genes involved in KEGG pathway for both *M. tuberculosis* and *M. bovis* infection pathway (Kanehisa et al., 2016) were used to extract transcripts for those genes from the

badger's annotated transcriptome. According to KEGG pathway for tuberculosis pathogenesis there are 183 genes involved in encoding cellular and nuclear receptors, intracellular signalling molecules and signalling cascades. All the genes involved in bTB pathway found matching transcripts with exception of Mincle (Macrophage-inducible C-type lectin receptor) which has not been detected in the transcriptome. For IL-6 (Interleukin 6), INFα (Interferon alpha) and INFβ (Interferon beta) only transcripts encode cellular receptors were found in the badger's transcriptome **Figure 5.7.**

Tuberculosis pathway can be viewed more clearly on http://www.genome.jp/kegg/ reference entry for tuberculosis is hsa05152**.**

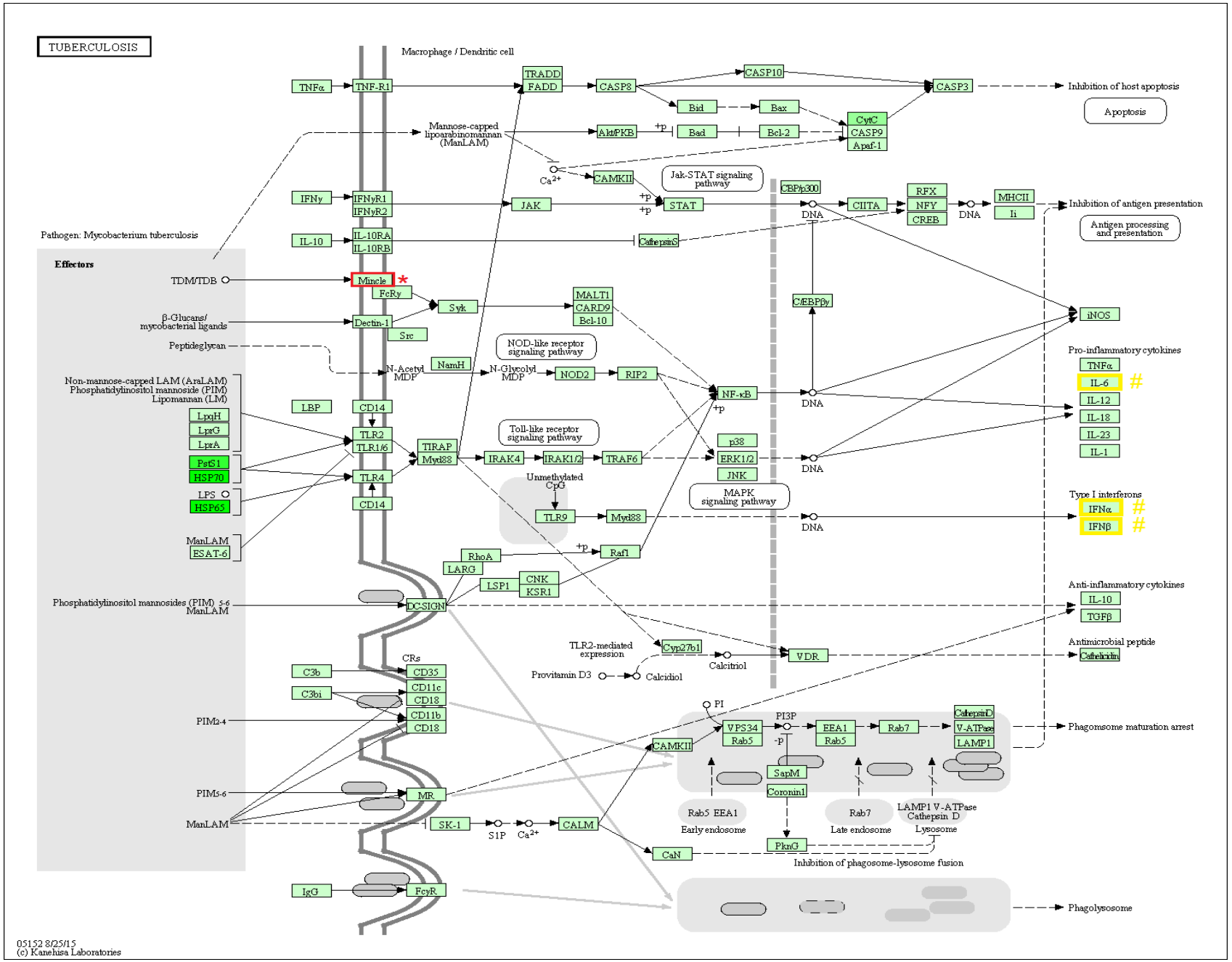

TUBERCULOSIS
Macrophage / Dendritic cell
TNFα
TNF-R1
TRADD
FADD
CASP8
CASP10
CASP3
Inhibition of host apoptosis
Apoptosis
Bid
Bax
CytC
CASP9
Apaf-1
Mannose-capped lipoarabinomannan (ManLAM)
Akt/PKB
Bad
Bcl-2
Ca2+
CAMKII
Jak-STAT signaling pathway
IFNγ
IFNγR1
IFNγR2
JAK
STAT
CBP/p300
CIITA
RFX
NFY
CREB
DNA
MHCII
Ii
Inhibition of antigen presentation
Antigen processing and presentation
Pathogen: Mycobacterium tuberculosis
IL-10
IL-10RA
IL-10RB
CathepsinS
Effectors
TDM/TDB
Mincle
FcRγ
Syk
MALT1
CARD9
Bcl-10
C/EBPβγ
iNOS
β-Glucans/ mycobacterial ligands
Dectin-1
Src
Peptideglycan
NOD-like receptor signaling pathway
N-Acetyl MDP
NamH
N-Glycolyl MDP
NOD2
RIP2
NF-κB
Pro-inflammatory cytokines
TNFα
IL-6
IL-12
IL-18
IL-23
IL-1
Non-mannose-capped LAM (AraLAM)
Phosphatidylinositol mannoside (PIM)
Lipomannan (LM)
LBP
CD14
LpqH
LprG
LprA
TLR2
TLR1/6
TIRAP
Myd88
Toll-like receptor signaling pathway
IRAK4
IRAK1/2
TRAF6
p38
ERK1/2
JNK
MAPK signaling pathway
PstS1
HSP70
TLR4
CD14
LPS
HSP65
Unmethylated CpG
TLR9
Myd88
Type I interferons
IFNα
IFNβ
ManLAM
ESAT-6
RhoA
Raf1
LARG
LSP1
CNK
KSR1
Anti-inflammatory cytokines
IL-10
TGFβ
Phosphatidylinositol mannosides (PIM) 5-6
ManLAM
DC-SIGN
TLR2-mediated expression
Cyp27b1
VDR
Antimicrobial peptide
Cathelicidin
Provitamin D3
Calcidiol
Calcitriol
CRs
C3b
CD35
C3bi
CD11c
CD18
CD11b
CD18
PIM2-4
PI
PI3P
VPS34
Rab5
EEA1
Rab5
Rab7
CathepsinD
V-ATPase
LAMP1
Phagomsome maturation arrest
CAMKII
SapM
Coronin1
PIM5-6
MR
ManLAM
SK-1
S1P
Ca2+
CALM
Rab5 EEA1
Early endosome
Rab7
Late endosome
LAMP1 V-ATPase Cathepsin D
Lysosome
CaN
PknG
Inhibition of phagosome-lysosome fusion
IgG
FcγR
Phagolysosome
05152 8/25/15
(c) Kanehisa Laboratories

Figure 5.7: bTB pathway (KEGG), (*) Genes with no matching transcripts. (#) Genes with the receptors only found in the transcriptome sequences. (Kanehisa et al., 2016).

5.4.2 Constructing a phylogenetic tree using key transcripts in bTB infection

"Genetic factors play a major role in determining differential susceptibility to infection and disease outcome. Genetic variation in an increasing number of genes (e.g., NRAMP1, HLA class II, VDR, DC-SIGN, TLR8) has been found to be associated with complex susceptibility to pulmonary TB" (Neyrolles and Quintana-Murci, 2009).

Natural resistance-associated macrophage protein 1 (NRAMP1) polymorphisms have been found to be associated with susceptibility to tuberculosis in cattle (Liu et al., 2017). However, a systematic review on HLA class II genes can highly vary in their role against tuberculosis infection even within the same species (Oliveira-Cortez et al., 2016). Vitamin D receptor (VDR) polymorphisms and vitamin D deficiency are also associated with tuberculosis susceptibility in humans (Lee et al., 2016). DC-sign protein (encoded by CD209 gene) is a pathogen-recognition receptor expressed on the surface of immature dendritic cells (DCs) and involved in initiation of primary immune response through binding mannose containing antigens bacterial cell surface (Tanne and Neyrolles, 2010) and thereby considered one of the early systematic responses to tuberculosis. TLR8 genetic polymorphisms are associated with susceptibility to mycobacterial infection and

the severity of clinical manifestation by affecting phagocytosis in monocytes (Lai et al., 2016).

Here we use BLAST+ and Clustal Omega (through EMBL-EBI) to draw phylogenetic trees using transcripts of these genes with standard search criteria of the animal orthologs.

Both input criteria for BALSTP and Clustal Omega are shown below.

Blast+	**Clustal Omega**
Database: UniProtKB mammals/ Program: blastp/ Matrix: BLOSUM62/ Expectation value threshold: 1e-5/ Dropoff: 0/ Gap open: -1/ Gap extend: -1/ Filter: F/ Sequence range: START-END/ Gapalign: true/ Composition-based statistics: F/ Align views: 0/ Translation table: -1/ Sequence type:protein	Program: clustalo/ Version: 1.2.4/ Output guide tree: false/ Output distance matrix: false/ Dealign input sequences: false/ mBed-like clustering guide tree: true/ mBed-like clustering iteration: true/ Number of iterations: 0/ Maximum guide tree iterations: -1/ Maximum HMM iterations: -1/ Output alignment format: Clustal/ Output order: aligned/ Sequence Type: protein

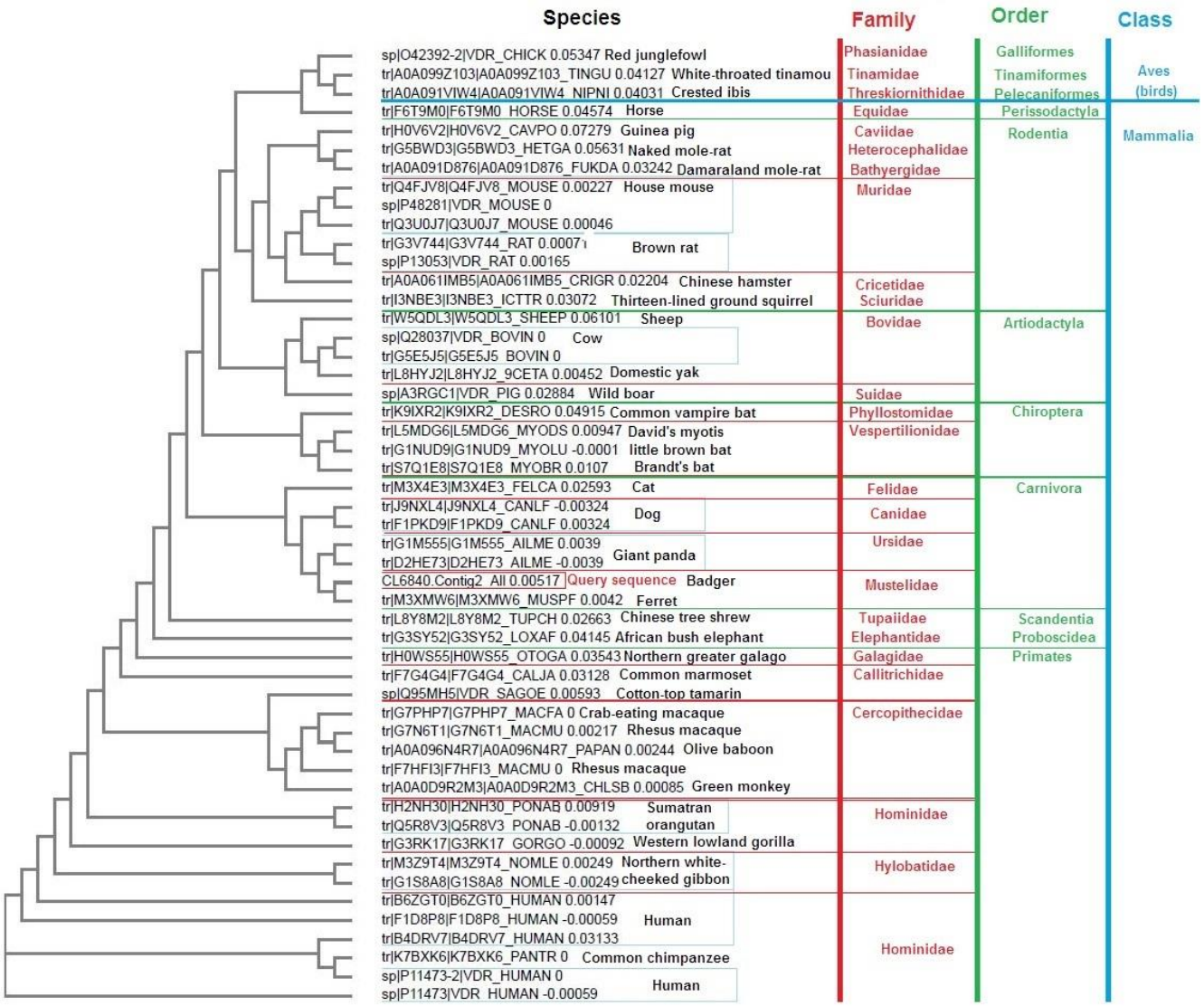


Figure 5.8: A phylogenetic tree of the VDR gene

Figure 5.8 In this example a Blastp search of the orthologs of VDR gene was performed and 50 aligned sequences where then used to build a phylogenetic tree (gene tree) using Clustal Omega. The 50 orthologs were from 3 species of birds and 38 species of mammals including the badger. These species were originated from 25 families which came from 11 orders from both classes (birds and Mammals).

The "tips" of the tree branches represent the taxa in the study. These taxa include taxonomic levels of species, families, orders and classes.

In the other **figure 5.9** DC-sign transcript was used to build a gene phylogeny using 50 orthologs from 33 species divided into 19 families and 8 orders mammals.

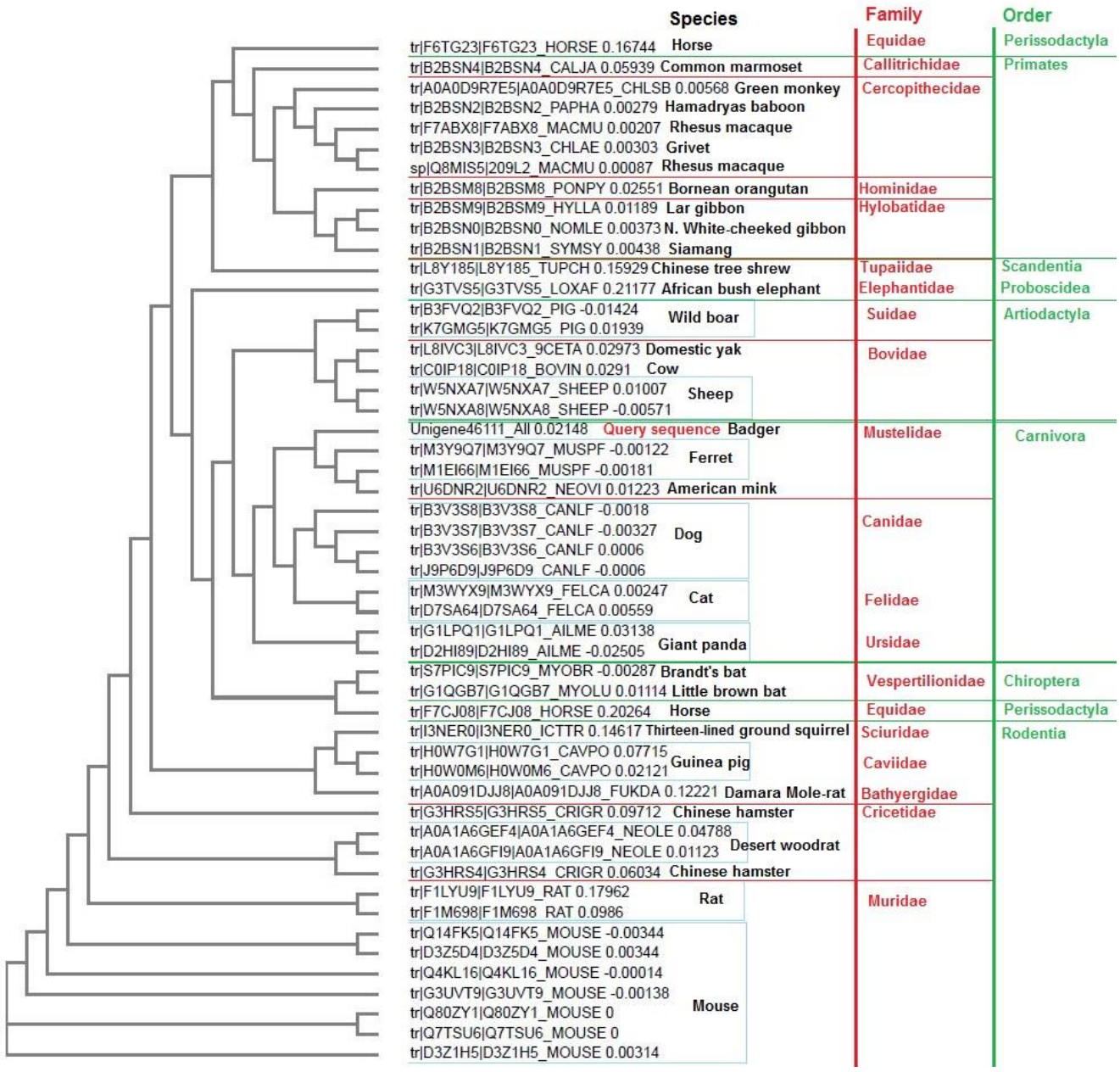


Figure 5.9: A phylogenetic tree of DC-sign gene

Similar results were obtained for HLA class II, NRAMP1 and TLR8 orthologs **Figure 5.10** where species share the same ancestor gene but with different "adaptive radiation" levels which led to speciation (appearance of different species) and different homologs of the gene within the same species.

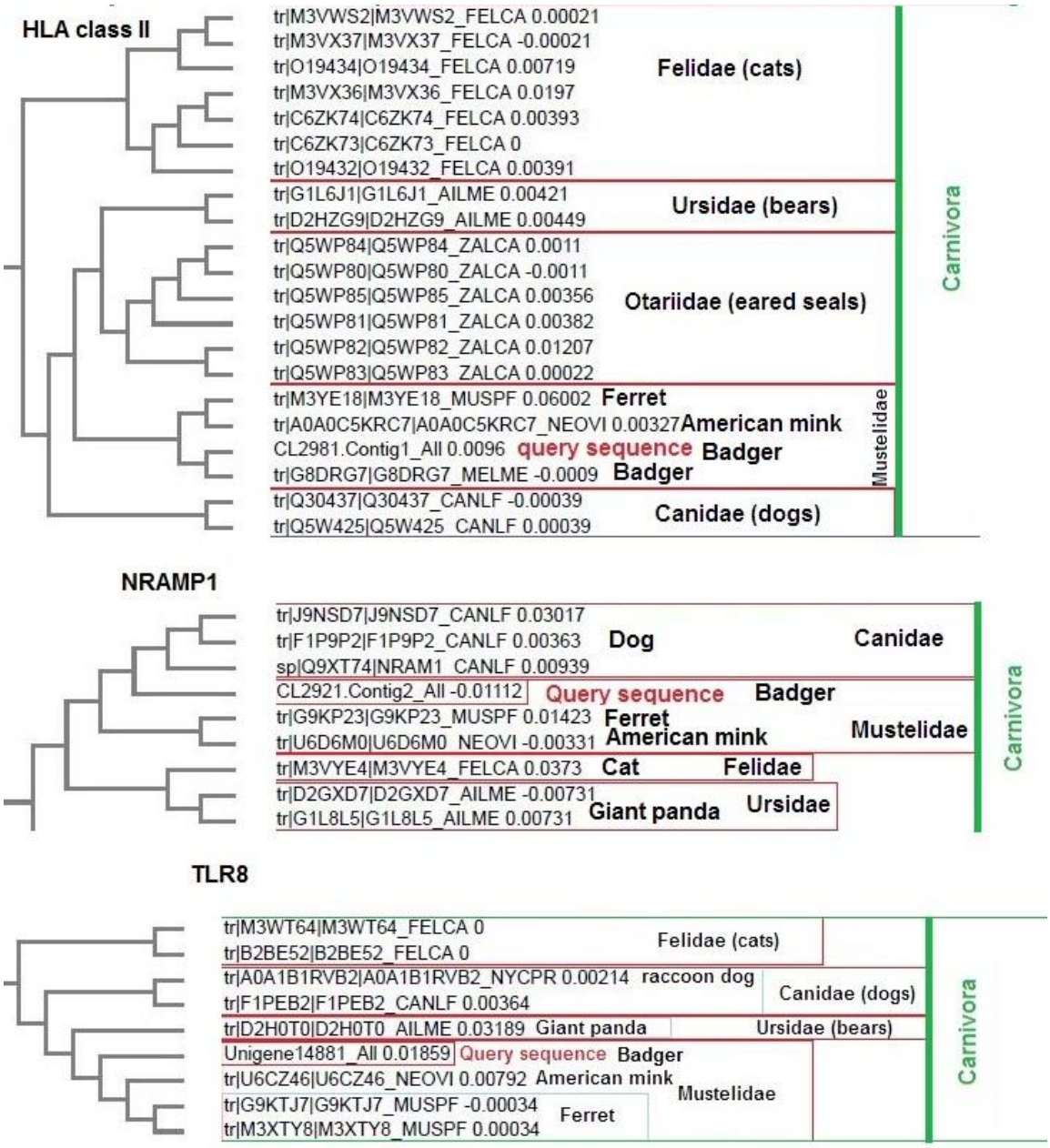


Figure 5.10: the position of the badger in the order Carnivora from three sections of the phylogenetic trees of HLA class II, NRAMP1 and TLR8 genes

5.5 Discussion

This chapter provides an *in silico* characterisation of the coding potential of the badger transcriptome, as a platform for characterising and investigating the immune component of the assembled transcriptome. A pilot search has shown that 15967 annotations in four databases are related to the search term (immun-), however, when a more thorough gene-by-gene search was performed using the IKB immunome database as a reference, and it showed that 698 genes out of 893 immunity genes in the database have corresponding transcripts in the badger transcriptome (11724) annotations. 195 genes from IKB show no match in the available data.

A total of 919 transcripts annotated as immunoglobulins in three databases reveal a similarity to immunoglobulin genes from a number of mammals; mainly *Ailuropoda melanoleuca* (Giant panda), *Mustela putorius furo* (Ferret), *Canis lupus familiaris* (Dog) and *Homo sapiens* (Human). Swiss-Prot annotation of immunoglobulin transcripts showed over 84% similarity of the transcripts to human immunoglobulins, in terms of predicted function.

Venn diagrams are commonly used to visualize the overlap between data sets, including differential gene expression data under various condition and genes or transcripts annotation in different databases. NR database contains sequences from both, non-curated and curated databases including non-reviewed sequences submitted from individual laboratories and large-scale sequencing projects deposited in GenBank and non-reviewed section of Uni-Prot. Thereby NR has the highest number of annotated

sequences both independently (8 annotations) or shared with other databases. Swiss-Prot is a manually annotated, non-redundant protein sequence database with over 500,000 manually annotated, curated and reviewed protein sequence. If we compare the number of deposited sequence in Swiss-Prot to that of Genbank alone (200,877,884 nucleotide sequence on April 2017) we find that the number of annotations in GenBank for any given sequence is higher. This can be useful for preliminary investigation of a sequence source or closest similar organism. However, in order to obtain a higher level of annotation such as description of the function of a protein, its domain structure and post-translational modifications then Swiss-Prot will have the advantage over the NR database. The COG database is a tool for identifying ortholog and paralog proteins and classifying those proteins into identifiable clusters according to protein function i.e. protein grouping rather than direct identification of the sequences and thus show less independence in annotation. The KEGG pathway is a tool to assign sequences to their known or predicted metabolic pathway utilizing genomic and non-genomic data (e.g. drug and disease effects) and thereby it is another useful yet less independent database where primary annotations are performed through NCBI blast searching. Similar Venn diagram distribution patterns were also found in the swamp buffalo transcriptome analysis (Deng et al., 2016) and giant panda blood transcriptome analysis (Du et al., 2015) which might explain that the difference in number of annotations among databases is more dependent on the characteristics of the database itself (such as availability of similar sequences or the level of annotation and classification in each database) rather than the significance of sequence in question.

For the 75 transcripts that match variable genes, NR and NT annotation showed a similarity to *Mustela putorius furo* (Ferret), *Canis lupus familiaris* (Dog), *Files catus* (Cat) and *Homo sapiens* (Human). However, KEGG annotation showed that 14 transcripts could not be mapped to any known pathway which may indicate a different spectrum of antigen binding capabilities of the badger's immunoglobulins to foreign antigens. Some mapped to pathways from *Bos taurus* (Cow) and *Monodelphis domestica* (Opossum). Both mammals can acquire bTB infection.

KEGG pathway annotation for both *M. bovis* and *M. tuberculosis* infection suggests that 183 mammalian genes are involved in the pathogenesis of the disease. 1825 transcripts were extracted that correspond to 147 of those 183 genes that are involved in bTB pathway from early infection until cell death. Some transcripts were found to match genes that encode receptors for signalling proteins but not the proteins themselves for the remaining 36 genes in bTB pathway. Only subunits of IL-6 receptors (IL6R, IL6ST) and INFα and INFβ receptors (IFNAR1, IFNAR2) were found in the annotated transcriptome. Production of interferons occurs mainly in response to infection and the inability to detect their presence in the transcriptome might be explained by the fact that the animals were kept all their lives in a (disease-free) controlled environment. Mincle is essential in detection of microbial glycolipids and it has been found that Mincle-dependent macrophage activation is regulated by IL-4 which has a strong downregulation of the mRNA expression of Mincle (Hupfer et al., 2016). 24 transcripts were annotated as IL-4 and signal transducer and activator of transcription 6 which might explain the undetected presence of Mincle transcripts.

A phylogeny is a "tree", which estimates the "historical" connections between species or genes that they carry. Five different transcripts were used to construct gene phylogenies compared to other fifty animals from different classes and orders. "Gene" trees represent the evolutionary history of the genes included in the study and provide evidence for gene duplication events, as well as speciation events by including different homologs in a gene tree and clustering orthologs to demonstrate the evolutionary history of the gene orthologs.

The phylogeny was built using only molecular data with no regard to morphological, behavioural or any other type of data and only single transcripts were used, which may neglect some important inputs drawn by non-coding sequences or the whole genome. However, the phylogenetic trees generated by those transcripts managed to some extent place the badger in the right position in terms of gene speciation i.e. within the Mustelidae family. Although these trees show the approximate speciation process which led to the emergence of species, families and orders they do not show the exact chronological order for the evolution of these taxa.

Gene trees showed that phylogenetics can add biological meaning to the data. It is important to keep in mind that a single line on the phylogeny is in fact a population, and populations can have genetic variation. These gene variations represented as paralogs which resulted from duplication of homologs within the genome and orthologs which are defined as evolutionary counterparts derived from a single ancestral gene in the last common ancestor of the given two species (Koonin and Galperin, 2003). The genetic variations within the population can affect the outcome of the phylogeny when single

sequence is used to build the tree. Some immunity genes are highly polymorphic and different hosts usually have different genotypes and thereby recognize different spectrums of pathogens. However, even with such variability it is still possible to produce this phylogeny.

Chapter VI

6 Mealworm (*Tenebrio molitor*)

"Model organisms have always played a leading role in genetic research, starting with Mendel's experimentation with pea plants in the 19th century. Today, geneticists continue to rely on these organisms, especially when investigating questions of gene expression, function, and mutation. Interestingly, even simple organisms can reveal much about the molecular basis of disease; in fact, the simple structure, short life span, and easy manipulability of these species make it particularly amenable to ongoing use in the research environment." (Adams, 2008). Thereby, a model organism can be defined as an organism appropriate for studying a specific disease or phenomenon owing to its relatively short lifetime in addition to characterized genome, well known biology and ease of accessibility for laboratory studies (Hedges, 2002).

"The phylogeny and timescale of life are becoming better understood as the analysis of genomic data from model organisms continues to grow. As a result, discoveries are being made about the early history of life and the origin and development of complex multicellular life" (Hedges, 2002). Insects were among the first animals on land, and the diversity and distribution of the current living insects has often been described as "astonishing". Within excess of one million species, insects are the most diverse organisms in the history of life, both in the number of species and the variety of structures and behaviours (Grimaldi and Engel, 2005). In comparison to the number of insect species which is estimated at over 950,000, the diversity of their genome size remains poorly understood when compared to all vertebrates and not just mammals (over 66,000 species), with sparse and sporadic sampling of a few selected insect orders and with many

orders unrepresented or underrepresented in genome databases. The sizes of insect genomes are estimated to range from 91 to 7,752 megabase. (Hanrahan and Johnston, 2011). In terms of genome size, the closest insect to *T. molitor* is *Tribolium castaneum* (the red flour beetle) from the same family "Tenebrionidae" and it has a genome size of 160 megabase (Richards et al., 2008). On the other hand, the size of the fully sequenced genome of *Drosophila melanogaster* (the fruit fly) is estimated at 120 megabase (Adams et al., 2000). Understanding the difference in genome sizes between these two species provides one justification for widening the genomic pool of model organisms to be analysed by sequencing. In the case of the mealworm genome, the availability of the mitogenome and transcriptome make the task easier. As previously mentioned, the other advantages of using the mealworm beetle as a model organism, include its use in immunological research, which will be substantially enhanced with the availability of a sequenced genome. On the negative side of course, simple, model organisms are perhaps easier to understand, but provide less transferable information, owing to their specialised features: and in this respect, the mealworm is not an exception. However, acknowledging that all model organisms have limitations, the opportunity to derive the sequence of the mealworm genome, seems on balance worthwhile, in particular in respect of its use in a wider educational setting, as discussed below.

The ease of managing the mealworm in teaching and research laboratories, (as has been observed with school students at the Liverpool Life Sciences University Technical College (UTC)., has provided invaluable support for extending the knowledge base of this organism, using genomic technology. Moreover, the significant and experimentally

convenient, amounts of protein extraction that can be obtained from a single insect, further support the case for making this insect an addition to the list of model organisms. Advantages and disadvantages of common animal models are frequently under discussion and the two main disadvantages of using models are the evolutionary distance from human and the issue of manipulating a small flying insect or its embryo (Wheeler and Brandli, 2009). Although not essential for a school-based research organism, simple - manipulations, the mealworm is much easier to handle, dissect and generally maintain, compared with other known insect models such as the fruit fly. The fact that the mealworm is evolutionary distant to human more than other mammalian or vertebrate models is not a real hindrance when it comes to teaching fundamental aspects of Biology. On the contrary it might be easier to use an insect model to avoid conflicts in respect of ethics, religion, etc.

The cost of purchase (around £14 per kg of mealworms) and maintenance, and the level of sophistication in mealworm "husbandry" is relatively low. Simple experiments such as total protein and DNA extraction and UV light exposure to observe the morphological changes such as pigmentation have been performed at Liverpool UTC by students aged 14-16 years old which provides a promising opportunity for attracting young people to science, and also helps lay the foundations for their subsequent career or educational opportunities.

6.1 Genomic DNA extraction

Mealworm larvae were obtained from a local supplier (any general pet shop will suffice) and after removal of dead organisms and rinsing in water to remove debris, were dried on paper towels before freezing at -80˚C and storing for at least 24 hours before DNA extraction. This method improved the consistency of extraction compared with grinding the larvae directly. Typical mealworm larvae weigh around 150 mg, and 50 mg of homogenised larvae were used for genomic DNA extraction using an ISOLATE Genomic DNA Mini Kit (see M and M, give section no.).

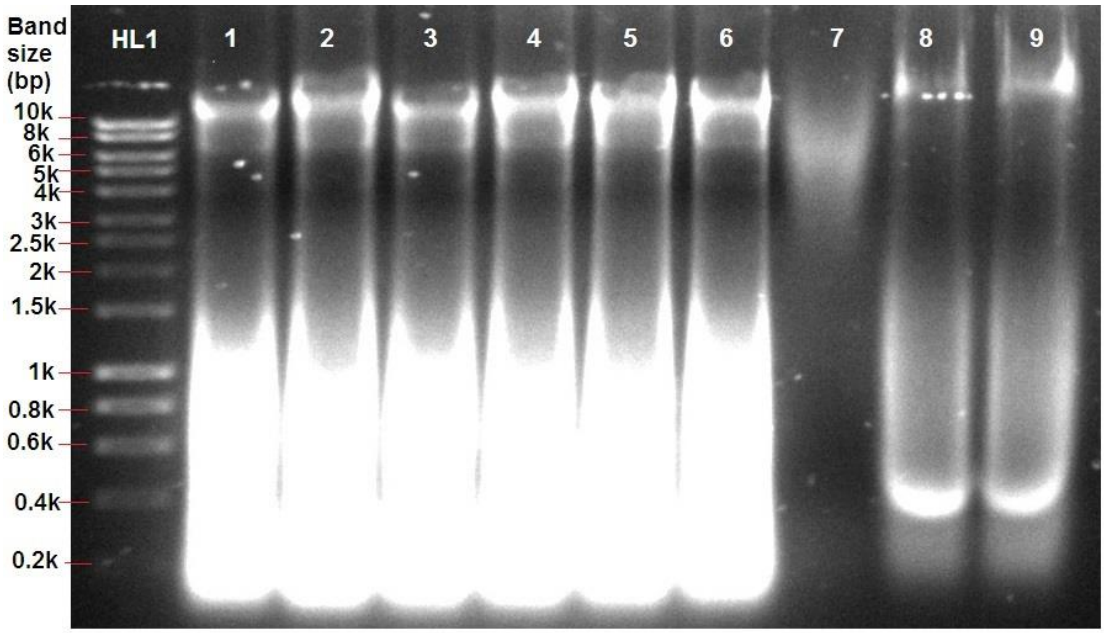


Figure 6.1: The analysis of genomic DNA extracted from T molitor, using the Bioline I Genomic DNA extraction kit followed by 1% agarose gel electrophoresis. Lane 1 (HL1) DNA Hyperladder I (lanes 2-6) X ul Extracted DNA (7) the wash buffer flow-through and lanes 8 and 9 whole lysate.

6.2 Methodology used to determine the nucleotide sequence of genomic DNA from T molitor

In view of the specialised techniques employed, the description of the sequencing experiments are included in this Chapter and not Chapter 2. Mealworm DNA samples prepared as above, were used to generate three DNA Illumina paired-end libraries with different insert sizes for sequencing and downstream genome assembly. The “Illumina” method is one of the most common methods used to sequence whole genomes. Genomic DNA was sheared using the Covaris S2 sonicator and used as input material for the TruSeq Nano DNA LT Sample Prep Kit. The libraries were purified throughout using AMPure XP bead clean at a ratio of 1:1.6 sample to beads. Following 6 cycles of amplification libraries were purified using AMPure XP beads. Each library was quantified using Qubit. Base-calling and de-multiplexing of indexed reads was performed by CASAVA version 1.8.2 (Illumina) to produce sequence data, in FASTQ format. The raw FASTQ files were trimmed to remove Illumina adapter sequences using Cutadapt version 1.2.1. The reads were further trimmed to remove low quality bases. If both reads from a pair passed this filter, each was included in the R1 (forward reads) or R2 (reverse reads) file. If only one of a read pair passed this filter, it was included in the R0 (unpaired reads) file **Table 6.1.**

Library	Raw reads	Trimmed reads (%)	Paired (R1/ R2) reads	Single (R0) reads (%)
Sample_2_350	153,873,144	153,405,605 (99.70)	76,513,117	379,371 (0.25)
Sample_3_550	136,112,914	135,963,239 (99.89)	67,921,367	120,505 (0.09)
Sample_2_3000	130,794,224	130,180,043 (99.53)	64,895,761	388,521 (0.30)

Table 6.1: summary of the read counts before and after adapter and quality trimming for all the samples. Percentage of the reads remaining after adapter and quality trimming are is shown in brackets.

6.3 Preliminary sequencing data

The primary results received to date are the raw reads of the genome (25GB of data) which represent sequencing data from three genomic DNA samples. The number of sequences produced ranges between 130-153 million short reads each.

6.3.1 Data files description

For paired-end sequence data, there are three sequence file types **(Figure 6.2)**. The files labelled R1 and R2 contain the corresponding paired-end sequences. The singlet files contain sequences whose pair has been removed due to poor sequence quality or adapter contamination. If a sample has been sequenced several times, there will be several sets of sequence files in the sample directory. These will need to be concatenated before downstream analysis.

6.3.2 Sequence size for each sample

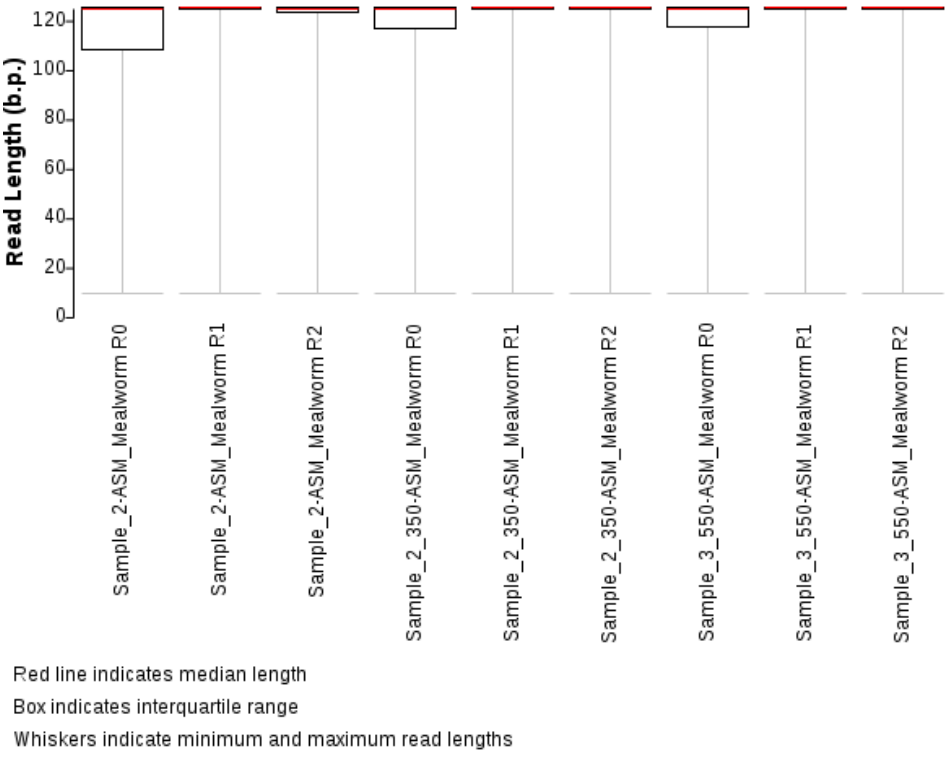


Figure 6.2: A graph showing the distribution of trimmed read lengths for the forward strand (R1), reverse strand (R2) and singlet (R0) reads. Note that it is common for a small number of reads to consist of mostly adapter-derived sequence, so it is expected that the distribution will show a long tail.

6.4 Mealworm (*T. molitor*) extracted genome BLASTN search results

The received data files incorporated three repeat sequencing experiments on the genomic DNA. Each library comprises three raw reads: R1 (Forward reads), R2 (Reverse reads) and R0 for single reads, all in FASTQ format. FASTQ files were converted to FASTA format using Galaxy platform (Afgan et al., 2016) in order to choose an unbiased selection of short reads for a BLASTN search.

A 1000 short reads were randomly selected from each file for the three sequencing samples making the total short reads around 9000 with a nucleotide length range from 45 bp to 125 bp per short read. The BLAST? search was optimised for highly similar sequences only (megablast) within the class Insecta (taxa ID: 50557) to avoid any interference or noise caused by the input sequences, as short input sequences have higher chance to align with sequences from organisms other than insects.

The percentage of short reads that were found to matching sequences are around 10-20%, i.e. for each 1000 short reads only 100-200 were found to have alignments in the public databases. This is probably attributed to the relatively "strict" search parameters employed, and the availability of sequences from similar organisms in the database, and of course the nature of query sequence themselves.

55% of the total observed alignments are with sequences from *T. molitor* with identity range between 90-100%. The sequences of T. molitor were from mitogenome, satellite repeats and other sequences such as antifreeze protein.

If we consider species diversity in the observed alignments, we find that over 80% of the alignments were with sequences from beetles (Coleoptera) order including mealworm, red flour beetle genome (2008) and Asian long-horned beetle genome (McKenna et al., 2016). The rest of the alignments are from different insect orders including flies (Diptera) and bees (Hymenoptera) among other orders **(Table 6.2)**.

Order	Family	Genus	Species
Coleoptera	*Tenebrionidae*	*Tenebrio*	*T. Molitor* (Mealworm)
Coleoptera	*Tenebrionidae*	*Tribolium*	*T. castaneum* (Red flour beetle)
Coleoptera	*Boridae*	*Synercticus*	*Synercticus spp.*
Coleoptera	*Cerambycidae*	*Anoplophora*	*A. glabripennis* (Asian long-horned beetle)
Coleoptera	*Silphidae*	*Nicrophorus*	*N. vespilloides* (Burying beetle)
Coleoptera	*Eucnemidae*	*Anischia*	*Anischia bicolor*
Coleoptera	*Cerylonidae*	*Mychocerus*	*Mychocerus spp.*
Coleoptera	*Scarabaeidae*	*Onthophagus*	*Onthophagus vinctus*
Diptera	*Tephritidae*	*Bactrocera*	*B. latifrons* (Solanum fruit fly)
Hymenoptera	*Formicidae*	*Dinoponera*	*D. quadriceps* (Ants species)
Diptera	*Drosophilidae*	*Drosophila*	*Drosophila busckii* (Fruit fly species)
Lepidoptera	*Saturniidae*	*Samia*	*Philosamia cynthia ricini* (Silkmoth)
Hemiptera	*Machaerotidae*	*Pectinariophyes*	*Pectinariophyes reticulata*
Phthiraptera	*Pediculidae*	*Pediculus*	*Pediculus humanus corporis* (Body louse)
Hymenoptera	*Formicidae*	*Monomorium*	*M. pharaonic* (Pharaoh ant)
Diptera	*Drosophilidae*	*Drosophila*	*Drosophila ficusphila* (Fruit fly species)
Diptera	*Tephritidae*	*Rhagoletis*	*Rhagoletis zephyria*
Hemiptera	*Cicadellidae*	*Ulopa*	*Ulopa reticulate*

Table 6.2: The most common insect species that were found to have sequence alignment with mealworm's query sequences. Species from the same order (Coleoptera) are highlighted in yellow.

6.5 Using short read to draw a phylogenetic tree

A short read (HISEQ: 115:C6KP7ANXX/ 18S Ribosomal RNA) was selected after showing an alignment with multiple insects in blastn to build a gene phylogenetic tree **(Figure 6.3)**.

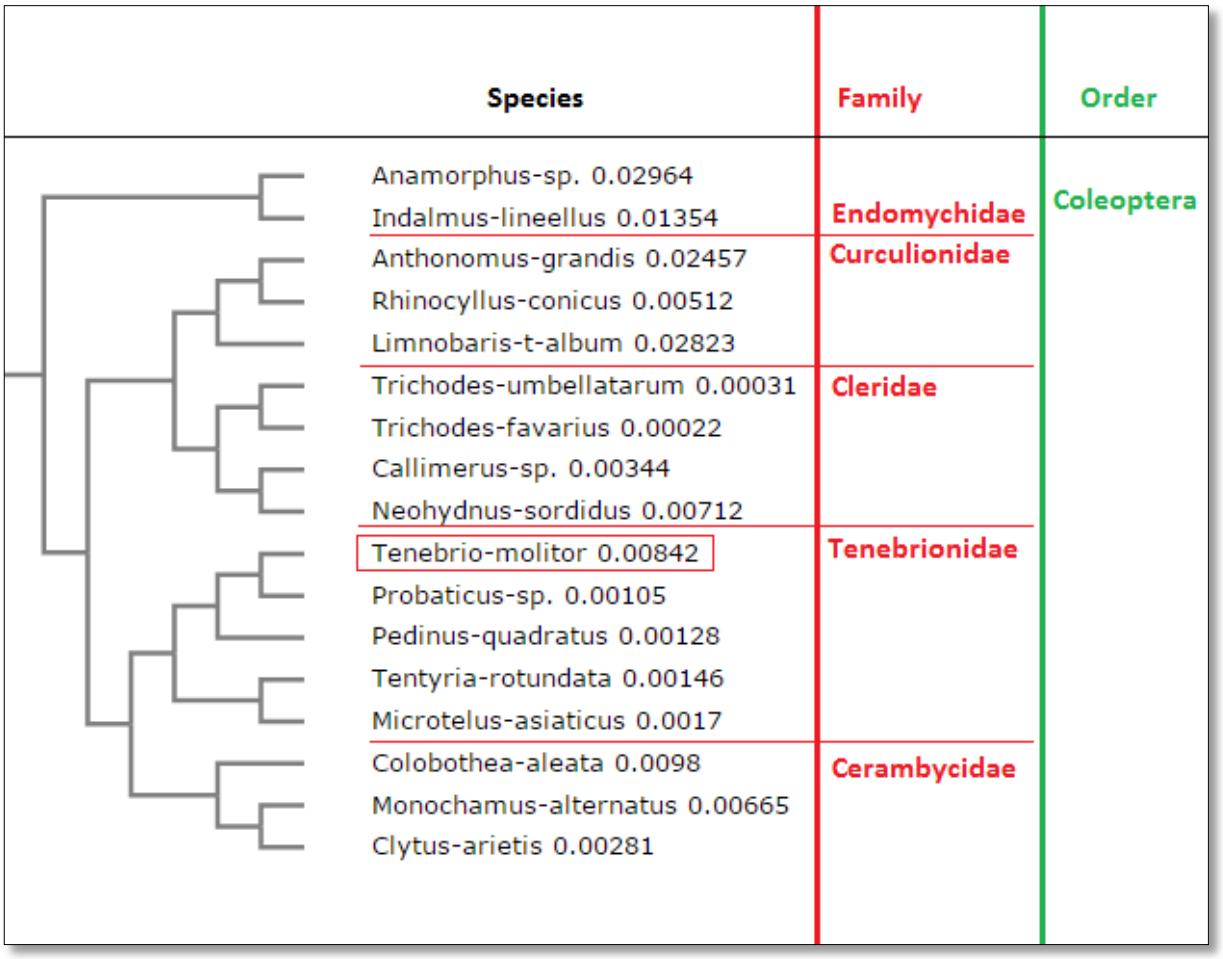


Figure 6.3: a gene phylogenetic tree for 18S ribosomal RNA sequence from different beetles of the order Coleoptera.

Using a partial sequence of 18S ribosomal RNA gene which has a highly conserved function among species can draw a relatively well-ordered species tree which congregate beetle species into their families under the order Coleoptera as shown in the (**figure 6.3).**

6.6 Discussion

6.6.1 Extracted DNA Quality

Genomic DNA samples were submitted for sequencing at the Liverpool University Genome Centre, and were relatively free from interfering contaminants having been and purified as described in **Chapter II** In order to circumvent the inhibition of downstream, enzymatic processing. A gel image of all samples, following analytical agarose gel electrophoresis provides a reasonable assessment of sample purity and integrity, and was used to "quality control" samples before submitting to sequencing. Genomic DNA was visible as a single prominent band >12 Kb (resolution diminishes significantly above fragment sizes in excess of 10kb). Samples with significant levels of degraded DNA were excluded (If more than one band or a smear was observed, the DNA may be degraded or contain components that might impair library preparation.) A low level of RNA usually appeared as a smear on gels, despite the use of RNase, however, it did appeared not to significantly interfere with sequencing. Purified DNA was dissolved in nuclease-free TE buffer for short and long term storage.

6.6.2 Quality of the "raw" sequence data

Although the current sequencing data are still non-assembled, the quality of raw reads was interrogated and tested using three adapter trimming programs (by the service provider): Sickle, EAUtils and Cutadapt which has been extensively used by sequencing providers (Martin, 2011).

6.6.3 Data significance

To my knowledge this is the first attempt to obtain the nucleotide sequence the mealworm genome. Although the data in hand are raw data, theyare suitable for submission to The Sequence Read Archive (SRA) of the NCBI database (Wheeler et al., 2008) to allow public access, free annotation and use. The availability of a *T. molitor* genome sequence, will facilitate the wider use of mealworm beetle in immunological research and go a long way in developing the insect as a new model organism.

6.6.4 Sequence alignment and phylogenetic tree

Around 9000 short reads from all the sequencing data FASTQ files were used in a sequence alignment search. The read lengths was in the range between 90 to 125 nucleotides. A highly constrained search was performed where BLASTN was limited to insects sequences in the database,s and only those sequences with an identity index of 95% or more. From each data file around 23% were found to show an alignment via NCBI blastn. Around 99% of the alignments were to satellite DNA from *T. molitor*. *Tenebrio molitor* contains an unusually abundant and homogeneous fraction of satellite DNA, which constitutes up to 60% of its genome (Davis and Wyatt, 1989). The remaining aligned sequences came from T. molitor mitochondrial DNA and other insects including species of beetles and fruit flies.

The E-value was in the range of 1e-58 to 9e-5 for the observed alignments. Most of the aligned sequences with high E-value were those of conserved genes from other insects (e.g. *Bactrocera latifrons:* basic-leucine zipper transcription factor D gene) or *T. molitor* specific sequences (e.g. *Tenebrio molitor:* antifreeze protein BST1 (BST1) gene). The

identity score in both examples were 94% and 100% respectively, which may indicate that however the E-value was at the high end, it is still identical to well-known conserved genes from other insects or the mealworm's own gene sequences. The identical alignments (i.e. 100%) were from *T. Molitor* mitochondrial genome (Liu and Wang, 2014) and ribosomal RNA genes and DNA satellite repeats (Davis and Wyatt, 1989) which may be helpful in a school based blastn searches for confirming the identity of the DNA under investigation (say by a PCR experiment).

Although there was considerable difficulty experienced in assembling the data in hand, on a genome level a simple NCBI blast search could be used to answer simple questions in a school environment, where students can take "ownership" of the project and learn about the importance of DNA sequencing and the value of sequence alignments in modern biological research. In this case a single or multiple short sequences could be used to answer questions such as "to what organism does this sequences belong?" and "how similar is this sequence to others from other organisms?" and the significance of sequence similarity in understanding evolution by using a partial sequence as shown in the phylogenetic tree of 18S ribosomal RNA sequence.

The demand for animal protein is expected to rise by 70–80% in the next five decades and mealworm is a suggested as a more sustainable source of protein (Adams, 2008). Simple lifecycle, protein extraction, DNA sequencing and bioinformatics experiments can help to draw young people's attention and acceptance for this possible futuristic food.

Chapter VI

7 Discussion

7.1 Analysis of the badger transcriptome

De novo transcriptome assembly can provide large amounts of data from which it is possible to build a preliminary platform for a future, total genome sequencing programme. These large data sets have become much easier to generate as a result of the simplification and cost reductions in NGS technology. Furthermore, the data we have obtained in this study, can be also used in further studies to explore the immune components of the badger's transcriptome, which represents another step in understanding the tolerance exhibited by the badger to infectious diseases such as bovine tuberculosis. However, in such a scenario it would be desirable to perform a parallel programme of sequencing of the blood transcriptome derived from an infected animal. Whilst this was an initial aim of this thesis, insurmountable logistical barriers prevented this work from being undertaken.

In comparison with the giant panda (*Ailuropoda melanoleuca*) transcriptome analysis (Du et al., 2015) the Illumina sequencing experiment produced approximately **3.52 Gb** of "clean" sequence data, which was assembled into **238,295** transcripts using a *de novo* transcriptome assembly without access to a reference genome. This compares with **92,598** for the panda. The average length of the assembled transcripts was **1281bp** with **N50** over **2720bp** long. The mean length of the panda transcripts was **1626 bp** with **N50** of **2842 bp**. The number of annotated transcripts for the panda was **38,726** whereas **95,245** of badger's transcripts were annotated using BLAST searches against public

databases. These annotated transcripts will provide a valued resource for improving the sequencing and annotation of the badger's complete genomic DNA. The rationale behind the using the giant panda to compare with the badger is that the panda was the closest mammal to badger in terms or RNA sequence similarity and sequence function prediction in the transcriptome annotation. Moreover, the comparison was between two peripheral blood transcriptome analyses data which might eliminating the tissue expression level difference bias as it would be difficult to compare two different tissue transcriptomes from two animals. Eckalbar et al., (2013) achieved similar improvement through reannotation of the genomic DNA of the green anole lizard by transcriptome sequencing.

E-value distribution in the NR database showed that 39% of the aligned transcripts exhibit a highly significant homology (<1.0E-100) whereas for the panda's transcriptome, about 44% of the transcripts have similar homology. Over 70% of both badger and panda transcripts achieved a similarity of 80% or more. The most important observation in this comparison is that, despite the panda's genome and mitochondrial DNA being sequenced, the NR alignment of the panda's transcripts give only 51% identity to the panda as the closest match, whereas for the badger 70% of the transcripts where aligned to five mammals with their genome already sequenced.

The GO annotation was performed to explore the assembled transcriptome and the results reflected a functional diversity of the transcripts in the three main categories of GO terms. The GO annotation seems consistent with similar results from a GO comparison of giant panda and polar bear blood transcriptomes by et al., (2015).

The mapping of assembled genes with pathway annotation to the KEGG database resulted in assigning those reads to 259 pathways compared to 324 KEGG pathways identified from the panda's transcriptome. Among the highest representation are metabolic pathways, at 97% (9930), Focal adhesion 7.61% (7582), Amoebiasis 7.17% (7136), ECM-receptor interaction 5.94% (5918), Protein digestion and absorption 5.65% (5623), RNA transport 4.17% (4158), Regulation of actin cytoskeleton 4.02% (4001), Pathways in cancer 3.88% (3862) and Herpes simplex infection 3.22% (3209). In addition to the major pathways represented by the transcripts in **Chapter III** there are other 239 pathways shown in the **appendix** including some infectious diseases pathways like influenza A, measles, toxoplasmosis and tuberculosis.

A search in the Cluster of Orthologous Groups (COG) database was performed, in order to uncover the homologous relationships of the transcriptome to well-known conserved domains. Most of the transcripts were matched to the major groups of (Cellular Processes and Signalling), (Information Storage and Processing) and (Metabolism). However, there was also some transcripts classified as (Poorly Characterized) which includes 22,329 transcripts under (General function prediction only) and 11,305 transcripts under (Function unknown) categories.

A phylogenetic (evolutionary) tree is a representation of the evolutionary relationships among a set of organisms known as taxa. It is important to recognise that if the phylogenetic tree is computed from data derived from a single gene (gene tree) is sometimes different from a whole-genome tree (species tree). One of the important factors that cause this difference is genetic polymorphism in the ancestral species (Pamilo

and Nei, 1988). Even under standardised alignment parameters different genes provide different levels of speciation leading to formation of different trees. This could be useful in studying the evolution of a single gene and its variants but also could be misleading when studying a whole organism despite the fact that our alignments always show badger or ferret is the closest because that's in part true but is also what is available in the database.

MHC genes are highly diverse both within species and among species. The identity range of the first hits in the protein alignment is between 74% and 100% which, despite the length of both query and subject sequences showed regions of highly conserved sequences which revealed the intra-species similarities and few regions of less conserved regions, which may reflect the diversity of MHC genes. However, this can perhaps be more easily observed through sequencing the full length of the gene of interest. In a comparison of the abundance of MHC transcripts it was found that the badger has higher expression levels of both classes of MHC genes. Although, this result may require further experimental investigation, it may reflect the findings of other work where MHC genes diversity is much lower in this endangered species (Zhang et al., 2015) and the diversity of MHC genes in the giant panda was relatively lower than other vertebrates (Zhu et al., 2007).

7.2 Data from the immune component of the badger's transcriptome

The search using the IKB immunome database (Ortutay et al., 2007) as a reference for immunity-related genes, revealed that **698** out of IKB's **893** immunity genes have **11724** annotations and **195** have no match to the badger's transcriptome.

The total **919** annotations of immunoglobulins in three databases do show a similarity to genes from different mammals; mainly: giant panda, ferret, dog and human. Swiss-Prot annotation of immunoglobulin-related transcripts showed over **84%** similarity of the transcripts to human immunoglobulin in terms of predicted function. Immunoglobulin heavy and light chains are composed of constant and variable regions and the amino-terminal variable or V domains of the heavy and light chains together make up the V region of the antibody and confer upon it the ability to bind specific antigen(s) (Janeway, 2001). The **75** transcripts that matches the variable genes by annotation, showed a similarity to *Mustela putorius furo* (Ferret), *Canis lupus familiaris* (Dog), *Files catus* (Cat) and *Homo sapiens* (Human). However, KEGG annotation showed that **14** transcripts could not be mapped to any known pathological pathway and that some mapped to pathways from *Bos taurus* (Cow) and *Monodelphis domestica* (Opossum). Both mammals can acquire bTB infection.

KEGG pathway annotation for both *M. bovis* and *M. tuberculosis* infection inside their hosts showed that **183** mammalian genes are involved in the pathogenesis of the diseases. **1825** transcripts from the badger were found to correspond to **147** of those **183** genes that are involved in bTB pathway from the early infection until cell death.

Interestingly, only sequences encoding the subunits of IL-6, INFα and INFβ receptors were found in the annotated transcriptome. Production of interferons occurs mainly in response to infection, and the inability to detect their presence in the transcriptome might be explained by the fact that the animals had been maintained for their life course in a (disease-free) controlled environment. The Mincle receptor is essential for the detection of microbial glycolipids, and it has been found that Mincle-dependent macrophage activation is regulated by IL-4, which leads to a strong downregulation of expression of Mincle mRNA (Hupfer et al., 2016). 24 transcripts were annotated as IL-4 and signal transducer and activator of transcription 6 which might explain the undetected presence of Mincle transcripts.

A phylogeny is a "tree", which estimates the "historical" connections between species or genes that they carry. Five different transcripts were used to construct gene phylogenies compared to other fifty animals from different classes and orders. "Gene" trees represent the evolutionary history of the genes included in the study and provide evidence for gene duplication events, as well as speciation events by including different homologues in a gene tree and clustering orthologues to demonstrate the evolutionary history of the gene orthologues.

7.3 Gender differences in gene expression

Transcriptome analysis has showed that sexual dimorphism in gene expression was much greater than previously recognized, since thousands of genes showed sexual dimorphism in liver, adipose, and muscle where the sexually dimorphic genes were also found to be

highly tissue-specific (Yang et al., 2006). Over 12% of transcripts were found upregulated in the male badger while 87% of transcripts were downregulated when compared to the female badger.

Although it cannot anticipated whether the observed differential expression between male and female badgers is a direct result of the hormonal effect, and despite the fact that these data were from uninfected animals, some of the expression levels difference were consistent with the biological differences between the sexes and their susceptibility to tuberculosis

7.4 Implications for the evolution of the European Badger

The phylogenetic gene-trees were constructed using molecular data with no regard to morphological, behavioural or any other type of data. In the construction of gene trees only transcripts were used, which may neglect some important inputs drawn by non-coding sequences or the whole genome. However, the phylogenetic trees generated by those transcripts managed to some extent place the badger in the appropriate (by consensus) position in terms of gene speciation i.e. within the Mustelidae family. Although these trees show the approximate speciation process which led to the emergence of species, families and orders they do not show the exact chronological order for the evolution of these taxa.

Gene-trees showed that phylogenetics can add biological meaning to the data. It is important to keep in mind that a single line on the phylogeny is in fact a population, and populations can have genetic variation. These gene variations represented as paralogs

which resulted from duplication of homologues within the genome and orthologues which are defined as evolutionary counterparts derived from a single ancestral gene in the last common ancestor of the given two species (Koonin and Galperin, 2003). The genetic variations within the population can affect the outcome of the phylogenetic conclusion when single sequence is used to build the tree. Some immunity genes are highly polymorphic and different hosts usually have different genotypes and thereby recognize different spectrums of pathogens. However, even with such variability it is still possible to produce this phylogeny.

Moreover, the limited availability of sequenced data from other related mammals can also contribute to the position of badger in the evolutionary timeline. However, it is still possible to make a reasonable, albeit tentative, judgement from such high quality data. The anatomical records place the Eurasian badger within the mustelid subfamily where it is closely related to the hog badger (Marmi et al., 2006), and this could be either sustained, shifted or changed when the whole genome is sequenced. In conclusion, whilst the transcriptome data point to the position of the badger**,** the definitive position must await the publication of the badger genome.

7.5 The data from the mealworm

The characteristics and the ease of handling of the mealworm make it a suitable choice as a model organism in biological and evolutionary research, which in turn made obtaining the full genomic data of the organism an experimental priority (Liu and Wang, 2014).

To my knowledge this is the first attempt to sequence the mealworm genome. Although the data in hand are at this stage "raw" data, they will shortly be submitted to The Sequence Read Archive (SRA) of NCBI database (Wheeler et al., 2008) to allow public access and free annotation and use.

As previously mentioned in the introduction, the mealworm has a number of advantages that support its selection as another model organism. Phenomena ranging from variable temperature tolerance to the ease of induction of infection can be easily investigated and combined with a full omics based approach. The availability of the *T. molitor* genome will facilitate the use of mealworm beetle in immunological research and in developing the insect as a new model organism for as yet, undetermined areas of Biology. As a very attractive by-product of this work, we have been able to succeed in making this model organism a suitable choice as a model organism for introducing school students to Systems Biology. This development well also help widening the spectrum of available model organisms leading to an increase in experimental variability and a decrease in bias.

Although there was some difficulty encountered in assembling the data in hand on a genome level: a simple NCBI blast search could be used to answer simple question in a school environment where students can own the project and learn about the importance of DNA sequencing and sequence alignment significance in modern biological research. In this case single or multiple short sequences could be used to answer questions like "to what organism does this sequences belong?" and "how similar is this sequence to others from other organisms?" and the significance of sequence similarity in understanding

evolution by using a partial sequence as shown in the phylogenetic tree of 18S ribosomal RNA sequence.

In addition, the demand for animal protein is expected to rise by 70–80% in the next five decades and mealworm is a suggested as a more sustainable source of protein (Adams, 2008). Simple lifecycle, protein extraction, DNA sequencing and bioinformatics experiments can help to draw young people's attention and acceptance for this possible futuristic food.

7.6 Limitations and proposed work

Current transcriptome analysis data alone do not provide an explanation for the mechanism, by which badgers develop a tolerance to bTB, or indeed what makes them effectively transmit the disease to other animals within the badger population or among farms. Moreover, the current data do not provide a prevention or treatment strategy for both wild and domestic animals. However, it still can be argued that these data can contribute to long-term future work such as genome sequencing. Furthermore, the available data encourage a further investigation of the badger immunome through a transcriptomic comparison between healthy and bTB infected animals.

The observed similarity between the badger's bTB-related transcripts that aligned to KEGG pathway of both bTB and Human tuberculosis would encourage a prospective study of the difference in the key genes in bTB in badgers, cow and human. The available transcriptome-immunome would facilitate designing further experiments (e.g. isolation

and amplification of specific proteins to study their interaction with the pathogen). Moreover, the identification of some variable regions genes with no mention of their annotation in the public databases could be a platform for further research to find if they play a role that is unique to the badger's immunoglobulins through specific binding to mycobacterial antigens. This can be achieved through a variety of techniques used in antibody research including monoclonal and polyclonal antibody production and phage display research.

Simultaneous RNA sequencing (also known as dual RNA-seq) of both mycobacteria and badger can be another approach to study the changes in gene expression in both organisms along with interactions among the products of those genes during the course of infection. However, dual RNA-seq for bTB should be carefully planned in terms of the sequencing time and whether a repeated sequencing is required for different stages of the disease. Such a procedure could be relatively expensive in an infectious disease that exhibits variable manifestation varying from active to latent tuberculosis.

Single-cell transcriptome analysis which another economical approach that can yield a high-throughput, high-resolution transcriptomic analysis of cell states and dynamics (Liu and Trapnell, 2016). Single-cell transcriptome analysis has the possibility of shedding the light on macrophages behaviour during the course of infection and their possibility to override the phagosome/lysosome fusion inhibition effect induced by mycobacteria and if that is achieved by over expression of certain genes and nuclear receptors.

References

8 References

2008. The genome of the model beetle and pest Tribolium castaneum. *Nature,* 452**,** 949-955.

2017. UniProt: the universal protein knowledgebase. *Nucleic Acids Research,* 45**,** D158-D169.

ABNAVE, P., CONTI, F., TORRE, C. & GHIGO, E. 2014. What RNAi screens in model organisms revealed about microbicidal response in mammals? *Frontiers in Cellular and Infection Microbiology,* 4**,** 184.

ADAMS, J. 2008. C. elegans: Model organism in the discovery of PKD. *Nature Education,* 1**,** 196.

ADAMS, M. D., CELNIKER, S. E., HOLT, R. A., EVANS, C. A., GOCAYNE, J. D., AMANATIDES, P. G., SCHERER, S. E., LI, P. W., HOSKINS, R. A., GALLE, R. F., GEORGE, R. A., LEWIS, S. E., RICHARDS, S., ASHBURNER, M., HENDERSON, S. N., SUTTON, G. G., WORTMAN, J. R., YANDELL, M. D., ZHANG, Q., CHEN, L. X., BRANDON, R. C., ROGERS, Y.-H. C., BLAZEJ, R. G., CHAMPE, M., PFEIFFER, B. D., WAN, K. H., DOYLE, C., BAXTER, E. G., HELT, G., NELSON, C. R., GABOR, G. L., MIKLOS, ABRIL, J. F., AGBAYANI, A., AN, H.-J., ANDREWS-PFANNKOCH, C., BALDWIN, D., BALLEW, R. M., BASU, A., BAXENDALE, J., BAYRAKTAROGLU, L., BEASLEY, E. M., BEESON, K. Y., BENOS, P. V., BERMAN, B. P., BHANDARI, D., BOLSHAKOV, S., BORKOVA, D., BOTCHAN, M. R., BOUCK, J., BROKSTEIN, P., BROTTIER, P., BURTIS, K. C., BUSAM, D. A., BUTLER, H., CADIEU, E., CENTER, A., CHANDRA, I., CHERRY, J. M., CAWLEY, S., DAHLKE, C., DAVENPORT, L. B., DAVIES, P., PABLOS, B. D., DELCHER, A., DENG, Z., MAYS, A. D., DEW, I., DIETZ, S. M., DODSON, K., DOUP, L. E., DOWNES, M., DUGAN-ROCHA, S., DUNKOV, B. C., DUNN, P., DURBIN, K. J., EVANGELISTA, C. C., FERRAZ, C., FERRIERA, S., FLEISCHMANN, W., FOSLER, C., GABRIELIAN, A. E., GARG, N. S., GELBART, W. M., GLASSER, K., GLODEK, A., GONG, F., GORRELL, J. H., GU, Z., GUAN, P., HARRIS, M., HARRIS, N. L., HARVEY, D., HEIMAN, T. J., HERNANDEZ, J. R., HOUCK, J., HOSTIN, D., HOUSTON, K. A., HOWLAND, T. J., WEI, M.-H., et al. 2000. The Genome Sequence of Drosophila melanogaster. *Science,* 287**,** 2185-2195.

AFGAN, E., BAKER, D., VAN DEN BEEK, M., BLANKENBERG, D., BOUVIER, D., ČECH, M., CHILTON, J., CLEMENTS, D., CORAOR, N., EBERHARD, C., GRÜNING, B., GUERLER, A., HILLMAN-JACKSON, J., VON KUSTER, G., RASCHE, E., SORANZO, N., TURAGA, N., TAYLOR, J., NEKRUTENKO, A. & GOECKS, J. 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research,* 44**,** W3-W10.

AGUILAR-MIRANDA, E. D., LÓPEZ, M. G., ESCAMILLA-SANTANA, C. & BARBA DE LA ROSA, A. P. 2002. Characteristics of Maize Flour Tortilla Supplemented with Ground Tenebrio molitor Larvae. *Journal of Agricultural and Food Chemistry,* 50**,** 192-195.

ALTSCHUL, S. F., MADDEN, T. L., SCHÄFFER, A. A., ZHANG, J., ZHANG, Z., MILLER, W. & LIPMAN, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research,* 25**,** 3389-3402.

ANGELIA, M. R. N., GARCIA, R. N., CALDO, K. M. P., PRAK, K., UTSUMI, S. & TECSON-MENDOZA, E. M. 2010. Physicochemical and Functional Characterization of Cocosin, the Coconut 11S Globulin. *Food Science and Technology Research,* 16**,** 225-232.

ARES, E. H., O. 2014. Badger Culling: TB Control Policy. *In:* (SGS), D. E. A. S. O. H. (ed.).

ASHBURNER, M., BALL, C. A., BLAKE, J. A., BOTSTEIN, D., BUTLER, H., CHERRY, J. M., DAVIS, A. P., DOLINSKI, K., DWIGHT, S. S., EPPIG, J. T., HARRIS, M. A., HILL, D. P., ISSEL-TARVER, L., KASARSKIS, A., LEWIS, S., MATESE, J. C., RICHARDSON, J. E., RINGWALD, M., RUBIN, G. M. & SHERLOCK, G. 2000. Gene Ontology: tool for the unification of biology. *Nat Genet,* 25**,** 25-29.

BAIROCH, A. & APWEILER, R. 1996. The SWISS-PROT Protein Sequence Data Bank and Its New Supplement TREMBL. *Nucleic Acids Research,* 24**,** 21-25.

BARKER, D., FITZPATRICK, M. P. & DIERENFELD, E. S. 1998. Nutrient composition of selected whole invertebrates. *Zoo Biology,* 17**,** 123-134.

BARNES, A. I. & SIVA-JOTHY, M. T. 2000. Density-dependent prophylaxis in the mealworm beetle Tenebrio molitor L. (Coleoptera: Tenebrionidae): cuticular melanization is an indicator of investment in immunity. *Proceedings of the Royal Society B: Biological Sciences,* 267**,** 177-182.

BENJAMINI, Y. & HOCHBERG, Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological),* 57**,** 289-300.

BENTLEY, D. R., BALASUBRAMANIAN, S., SWERDLOW, H. P., SMITH, G. P., MILTON, J., BROWN, C. G., HALL, K. P., EVERS, D. J., BARNES, C. L., BIGNELL, H. R., BOUTELL, J. M., BRYANT, J., CARTER, R. J., CHEETHAM, R. K., COX, A. J., ELLIS, D. J., FLATBUSH, M. R., GORMLEY, N. A., HUMPHRAY, S. J., IRVING, L. J., KARBELASHVILI, M. S., KIRK, S. M., LI, H., LIU, X., MAISINGER, K. S., MURRAY, L. J., OBRADOVIC, B., OST, T., PARKINSON, M. L., PRATT, M. R., RASOLONJATOVO, I. M. J., REED, M. T., RIGATTI, R., RODIGHIERO, C., ROSS, M. T., SABOT, A., SANKAR, S. V., SCALLY, A., SCHROTH, G. P., SMITH, M. E., SMITH, V. P., SPIRIDOU, A., TORRANCE, P. E., TZONEV, S. S., VERMAAS, E. H., WALTER, K., WU, X., ZHANG, L., ALAM, M. D., ANASTASI, C., ANIEBO, I. C., BAILEY, D. M. D., BANCARZ, I. R., BANERJEE, S., BARBOUR, S. G., BAYBAYAN, P. A., BENOIT, V. A., BENSON, K. F., BEVIS, C., BLACK, P. J., BOODHUN, A., BRENNAN, J. S., BRIDGHAM, J. A., BROWN, R. C., BROWN, A. A., BUERMANN, D. H., BUNDU, A. A., BURROWS, J. C., CARTER, N. P., CASTILLO, N., CATENAZZI, M. C. E., CHANG, S., COOLEY, R. N., CRAKE, N. R., DADA, O. O., DIAKOUMAKOS, K. D., DOMINGUEZ-FERNANDEZ, B., EARNSHAW, D. J., EGBUJOR, U. C., ELMORE, D. W., ETCHIN, S. S., EWAN, M. R., FEDURCO, M., FRASER, L. J., FAJARDO, K. V. F., FUREY, W. S., GEORGE, D., GIETZEN, K. J., GODDARD, C. P., GOLDA, G. S., GRANIERI, P. A., GREEN, D. E., GUSTAFSON, D. L., HANSEN, N. F., HARNISH, K., HAUDENSCHILD, C. D., HEYER, N. I., HIMS, M. M., HO, J. T.,

HORGAN, A. M., et al. 2008. Accurate Whole Human Genome Sequencing using Reversible Terminator Chemistry. *Nature,* 456**,** 53-59.
BOLKER, J. A. 1995. Model systems in developmental biology. *BioEssays,* 17**,** 451-455.
BORATYN, G. M., CAMACHO, C., COOPER, P. S., COULOURIS, G., FONG, A., MA, N., MADDEN, T. L., MATTEN, W. T., MCGINNIS, S. D., MEREZHUK, Y., RAYTSELIS, Y., SAYERS, E. W., TAO, T., YE, J. & ZARETSKAYA, I. 2013. BLAST: a more efficient report with usability improvements. *Nucleic Acids Research,* 41**,** W29-W33.
BRADFORD, C. M. 2012. *The role of citrullination of central nervous system proteins in multiple sclerosis.* Ph.D, Sheffield Hallam University.
BRONNER, I. F., QUAIL, M. A., TURNER, D. J. & SWERDLOW, H. 2009. Improved Protocols for Illumina Sequencing. *Current protocols in human genetics / editorial board, Jonathan L. Haines ... [et al.],* 0 18**,** 10.1002/0471142905.hg1802s62.
BRUDER, C. E., YAO, S., LARSON, F., CAMP, J. V., TAPP, R., MCBRAYER, A., POWERS, N., VALDIVIA GRANDA, W. & JONSSON, C. B. 2010. Transcriptome sequencing and development of an expression microarray platform for the domestic ferret. *BMC Genomics,* 11**,** 251-251.
CALLINAN, P. A. & FEINBERG, A. P. 2006. The emerging science of epigenomics. *Human Molecular Genetics,* 15**,** R95-R101.
CAMACHO, C., COULOURIS, G., AVAGYAN, V., MA, N., PAPADOPOULOS, J., BEALER, K. & MADDEN, T. L. 2009a. BLAST+: architecture and applications. *BMC bioinformatics* [Online], 10. Available: http://europepmc.org/abstract/MED/20003500

http://europepmc.org/articles/PMC2803857?pdf=render

http://europepmc.org/articles/PMC2803857

http://www.biomedcentral.com/1471-2105/10/421

http://www.biomedcentral.com/1471-2105/10/421/abstract

http://www.biomedcentral.com/content/pdf/1471-2105-10-421.pdf

http://www.pubmedcentral.nih.gov/picrender.fcgi?tool=EBI&pubmedid=20003500&action=stream&blobtype=pdf

http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=EBI&pubmedid=20003500

http://dx.doi.org/10.1186/1471-2105-10-421 [Accessed 2009].
CAMACHO, C., COULOURIS, G., AVAGYAN, V., MA, N., PAPADOPOULOS, J., BEALER, K. & MADDEN, T. L. 2009b. BLAST+: architecture and applications. *BMC Bioinformatics,* 10**,** 421.
CARNINCI, P., KASUKAWA, T., KATAYAMA, S., GOUGH, J., FRITH, M. C., MAEDA, N., OYAMA, R., RAVASI, T., LENHARD, B., WELLS, C., KODZIUS, R., SHIMOKAWA, K., BAJIC, V. B., BRENNER, S. E., BATALOV, S., FORREST, A. R., ZAVOLAN, M., DAVIS, M. J., WILMING, L. G., AIDINIS, V., ALLEN, J. E., AMBESI-IMPIOMBATO, A., APWEILER, R., ATURALIYA, R. N., BAILEY, T. L., BANSAL, M., BAXTER, L., BEISEL, K. W., BERSANO, T., BONO, H., CHALK, A. M., CHIU, K. P., CHOUDHARY, V.,

CHRISTOFFELS, A., CLUTTERBUCK, D. R., CROWE, M. L., DALLA, E., DALRYMPLE, B. P., DE BONO, B., DELLA GATTA, G., DI BERNARDO, D., DOWN, T., ENGSTROM, P., FAGIOLINI, M., FAULKNER, G., FLETCHER, C. F., FUKUSHIMA, T., FURUNO, M., FUTAKI, S., GARIBOLDI, M., GEORGII-HEMMING, P., GINGERAS, T. R., GOJOBORI, T., GREEN, R. E., GUSTINCICH, S., HARBERS, M., HAYASHI, Y., HENSCH, T. K., HIROKAWA, N., HILL, D., HUMINIECKI, L., IACONO, M., IKEO, K., IWAMA, A., ISHIKAWA, T., JAKT, M., KANAPIN, A., KATOH, M., KAWASAWA, Y., KELSO, J., KITAMURA, H., KITANO, H., KOLLIAS, G., KRISHNAN, S. P., KRUGER, A., KUMMERFELD, S. K., KUROCHKIN, I. V., LAREAU, L. F., LAZAREVIC, D., LIPOVICH, L., LIU, J., LIUNI, S., MCWILLIAM, S., MADAN BABU, M., MADERA, M., MARCHIONNI, L., MATSUDA, H., MATSUZAWA, S., MIKI, H., MIGNONE, F., MIYAKE, S., MORRIS, K., MOTTAGUI-TABAR, S., MULDER, N., NAKANO, N., NAKAUCHI, H., NG, P., NILSSON, R., NISHIGUCHI, S., NISHIKAWA, S., et al. 2005. The transcriptional landscape of the mammalian genome. *Science,* 309**,** 1559-63.

CHAIN, P. S. G., GRAFHAM, D. V., FULTON, R. S., FITZGERALD, M. G., HOSTETLER, J., MUZNY, D., ALI, J., BIRREN, B., BRUCE, D. C., BUHAY, C., COLE, J. R., DING, Y., DUGAN, S., FIELD, D., GARRITY, G. M., GIBBS, R., GRAVES, T., HAN, C. S., HARRISON, S. H., HIGHLANDER, S., HUGENHOLTZ, P., KHOURI, H. M., KODIRA, C. D., KOLKER, E., KYRPIDES, N. C., LANG, D., LAPIDUS, A., MALFATTI, S. A., MARKOWITZ, V., METHA, T., NELSON, K. E., PARKHILL, J., PITLUCK, S., QIN, X., READ, T. D., SCHMUTZ, J., SOZHAMANNAN, S., STERK, P., STRAUSBERG, R. L., SUTTON, G., THOMSON, N. R., TIEDJE, J. M., WEINSTOCK, G., WOLLAM, A., DETTER, J. C., CONSORTIUM, G. S. & JUMPSTART, H. M. P. 2009. Genome Project Standards in a New Era of Sequencing. *Science,* 326**,** 236-237.

CHAMBERS, M. A., CARTER, S. P., WILSON, G. J., JONES, G., BROWN, E., HEWINSON, R. G. & VORDERMEIER, M. 2014. Vaccination against tuberculosis in badgers and cattle: an overview of the challenges, developments and current research priorities in Great Britain. *Veterinary Record,* 175**,** 90-96.

CHAPLIN, D. D. Overview of the immune response. *Journal of Allergy and Clinical Immunology,* 125**,** S3-S23.

CHAPLIN, D. D. 2010. Overview of the immune response. *Journal of Allergy and Clinical Immunology,* 125**,** S3-S23.

CLANCY, S. 2008. DNA transcription. *Nature Education,* 1**,** 41.

CLARK, T. A., SUGNET, C. W. & ARES, M. 2002. Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science,* 296**,** 907-910.

CLEARY, G. P., CORNER, L. A. L., O'KEEFFE, J. & MARPLES, N. M. 2009. The diet of the badger Meles meles in the Republic of Ireland. *Mammalian Biology - Zeitschrift für Säugetierkunde,* 74**,** 438-447.

CLOONAN, N., FORREST, A. R., KOLLE, G., GARDINER, B. B., FAULKNER, G. J., BROWN, M. K., TAYLOR, D. F., STEPTOE, A. L., WANI, S., BETHEL, G., ROBERTSON, A. J., PERKINS, A. C., BRUCE, S. J., LEE, C. C., RANADE, S. S., PECKHAM, H. E., MANNING, J. M., MCKERNAN, K. J. & GRIMMOND, S. M. 2008. Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods,* 5**,** 613-9.

COOPER, G. M. & HAUSMAN, R. E. 2009. *The Cell: A Molecular Approach*, ASM Press.
CORNER, L. A., MURPHY, D. & GORMLEY, E. 2011. Mycobacterium bovis infection in the Eurasian badger (Meles meles): the disease, pathogenesis, epidemiology and control. *J Comp Pathol,* 144**,** 1-24.
DAVIS, C. A. & WYATT, G. R. 1989. Distribution and sequence homogeneity of an abundant satellite DNA in the beetle, Tenebrio molitor. *Nucleic Acids Research,* 17**,** 5579-5586.
DE LISLE, G. W., BENGIS, R. G., SCHMITT, S. M. & O'BRIEN, D. J. 2002. Tuberculosis in free-ranging wildlife: detection, diagnosis and management. *Rev Sci Tech,* 21**,** 317-34.
DE SOUZA, P. C., MOREY, A. T., CASTANHEIRA, G. M., BOCATE, K. P., PANAGIO, L. A., ITO, F. A., FURLANETO, M. C., YAMADA-OGATTA, S. F., COSTA, I. N., MORA-MONTES, H. M. & ALMEIDA, R. S. 2015. Tenebrio molitor (Coleoptera: Tenebrionidae) as an alternative host to study fungal infections. *Journal of Microbiological Methods,* 118**,** 182-186.
DEAN, G. S., RHODES, S. G., COAD, M., WHELAN, A. O., COCKLE, P. J., CLIFFORD, D. J., HEWINSON, R. G. & VORDERMEIER, H. M. 2005. Minimum infective dose of Mycobacterium bovis in cattle. *Infection and Immunity,* 73**,** 6467-6471.
DEFRA 2011. Bovine TB Eradication Programme for England *In:* COMMUNICATIONS, T. (ed.). LONDON: Department for Environment, Food and Rural Affairs.
DELAHAY, R. J., BROWN, J. A., MALLINSON, P. J., SPYVEE, P. D., HANDOLL, D., ROGERS, L. M. & CHEESEMAN, C. L. 2000. The use of marked bait in studies of the territorial organization of the European Badger (Meles meles). *Mammal Review,* 30**,** 73-87.
DELAHAY, R. J., CHEESEMAN, C. L. & CLIFTON-HADLEY, R. S. 2001. Wildlife disease reservoirs: the epidemiology of Mycobacterium bovis infection in the European badger (Meles meles) and other British mammals. *Tuberculosis (Edinb),* 81**,** 43-9.
DELAHAY, R. J., DE LEEUW, A. N., BARLOW, A. M., CLIFTON-HADLEY, R. S. & CHEESEMAN, C. L. 2002. The status of Mycobacterium bovis infection in UK wild mammals: a review. *Vet J,* 164**,** 90-105.
DELAHAY, R. J., SMITH, G. C., WARD, A. I. & CHEESEMAN, C. L. 2005. Options for the management of bovine tuberculosis transmission from badgers (Meles meles) to cattle: evidence from a long-term study. *Mammal Study,* 30**,** S73-S81.
DENG, T., PANG, C., LU, X., ZHU, P., DUAN, A., TAN, Z., HUANG, J., LI, H., CHEN, M. & LIANG, X. 2016. De Novo Transcriptome Assembly of the Chinese Swamp Buffalo by RNA Sequencing and SSR Marker Discovery. *PLOS ONE,* 11**,** e0147132.
DESALERMOS, A., FUCHS, B. B. & MYLONAKIS, E. 2012. Selecting an Invertebrate Model Host for the Study of Fungal Pathogenesis. *PLoS Pathog,* 8**,** e1002451.
DÉSERT, C., MERLOT, E., ZERJAL, T., BED'HOM, B., HÄRTLE, S., LE CAM, A., ROUX, P.-F., BAEZA, E., GONDRET, F., DUCLOS, M. J. & LAGARRIGUE, S. 2016. Transcriptomes of whole blood and PBMC in chickens. *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics,* 20**,** 1-9.
DIAZ-SANCHEZ, S., HANNING, I., PENDLETON, S. & D'SOUZA, D. 2013. Next-generation sequencing: The future of molecular genetics in poultry production and food safety1. *Poultry Science,* 92**,** 562-572.

DOBSON, A. J., PURVES, J. & ROLFF, J. 2014. Increased survival of experimentally evolved antimicrobial peptide-resistant Staphylococcus aureus in an animal host. *Evolutionary Applications,* 7**,** 905-912.

DONG, Z. & CHEN, Y. 2013. Transcriptomics: Advances and approaches. *Science China Life Sciences,* 56**,** 960-967.

DU, L. M., LI, W. J., FAN, Z. X., SHEN, F. J., YANG, M. Y., WANG, Z. L., JIAN, Z. Y., HOU, R., YUE, B. S. & ZHANG, X. Y. 2015. First insights into the giant panda (Ailuropoda melanoleuca) blood transcriptome: a resource for novel gene loci and immunogenetics. *Molecular Ecology Resources,* 15**,** 1001-1013.

ECKALBAR, W. L., HUTCHINS, E. D., MARKOV, G. J., ALLEN, A. N., CORNEVEAUX, J. J., LINDBLAD-TOH, K., DI PALMA, F., ALFOLDI, J., HUENTELMAN, M. J. & KUSUMI, K. 2013. Genome reannotation of the lizard Anolis carolinensis based on 14 adult and embryonic deep transcriptomes. *Bmc Genomics,* 14.

FANG, Y., MEI, H., ZHOU, B., XIAO, X., YANG, M., HUANG, Y., LONG, X., HU, S. & TANG, C. 2016. De novo Transcriptome Analysis Reveals Distinct Defense Mechanisms by Young and Mature Leaves of Hevea brasiliensis (Para Rubber Tree). *Scientific Reports,* 6**,** 33151.

FIELD, L. A., JORDAN, R. M., HADIX, J. A., DUNN, M. A., SHRIVER, C. D., ELLSWORTH, R. E. & ELLSWORTH, D. L. 2007. Functional identity of genes detectable in expression profiling assays following globin mRNA reduction of peripheral blood samples. *Clinical Biochemistry,* 40**,** 499-502.

FISH, E. N. 2008. The X-files in immunity: sex-based differences predispose immune responses. *Nat Rev Immunol,* 8**,** 737-744.

FLYNN, J. J., FINARELLI, J. A., ZEHR, S., HSU, J. & NEDBAL, M. A. 2005. Molecular Phylogeny of the Carnivora (Mammalia): Assessing the Impact of Increased Sampling on Resolving Enigmatic Relationships. *Systematic Biology,* 54**,** 317-337.

FOX, S., FILICHKIN, S. & MOCKLER, T. C. 2009. Applications of Ultra-high-Throughput Sequencing. *In:* BELOSTOTSKY, D. A. (ed.) *Plant Systems Biology.* Totowa, NJ: Humana Press.

FRANCHI, L., WARNER, N., VIANI, K. & NUÑEZ, G. 2009. Function of Nod-like Receptors in Microbial Recognition and Host Defense. *Immunological reviews,* 227**,** 106-128.

FULTON, T. L. & STROBECK, C. 2006. Molecular phylogeny of the Arctoidea (Carnivora): Effect of missing data on supertree and supermatrix analyses of multiple gene data sets. *Molecular Phylogenetics and Evolution,* 41**,** 165-181.

GALLUCCI, S., LOLKEMA, M. & MATZINGER, P. 1999. Natural adjuvants: endogenous activators of dendritic cells. *Nat Med,* 5**,** 1249-55.

GONGORA-CASTILLO, E. & BUELL, C. R. 2013. Bioinformatics challenges in de novo transcriptome assembly using short read sequences in the absence of a reference genome sequence. *Natural Product Reports,* 30**,** 490-500.

GORMLEY, E. & COLLINS, J. D. 2000. The development of wildlife control strategies for eradication of tuberculosis in cattle in Ireland. *Tuber Lung Dis,* 80**,** 229-36.

GORMLEY, E. & CORNER, L. A. L. 2013. Control Strategies for Wildlife Tuberculosis in Ireland. *Transboundary and Emerging Diseases,* 60**,** 128-135.

GREEN, E. D. & CHAKRAVARTI, A. 2001. The human genome sequence expedition: views from the "base camp". *Genome Res,* 11**,** 645-51.

GRIESBECK, M., ZIEGLER, S., LAFFONT, S., SMITH, N., CHAUVEAU, L., TOMEZSKO, P., SHAREI, A., KOURJIAN, G., PORICHIS, F., HART, M., PALMER, C. D., SIRIGNANO, M., BEISEL, C., HILDEBRANDT, H., CÉNAC, C., VILLANI, A.-C., DIEFENBACH, T. J., LE GALL, S., SCHWARTZ, O., HERBEUVAL, J.-P., AUTRAN, B., GUÉRY, J.-C., CHANG, J. J. & ALTFELD, M. 2015. Sex Differences in Plasmacytoid Dendritic Cell Levels of IRF5 Drive Higher IFN-α Production in Women. *The Journal of Immunology,* 195**,** 5327-5336.

GRIFFIN, J. M., WILLIAMS, D. H., KELLY, G. E., CLEGG, T. A., O'BOYLE, I., COLLINS, J. D. & MORE, S. J. 2005. The impact of badger removal on the control of tuberculosis in cattle herds in Ireland. *Prev Vet Med,* 67**,** 237-66.

GRIMALDI, D. A. & ENGEL, M. S. 2005. *Evolution of the insects,* Cambridge U.K. ; New York, Cambridge University Press.

HANRAHAN, S. J. & JOHNSTON, J. S. 2011. New genome size estimates of 134 species of arthropods. *Chromosome Res,* 19**,** 809-23.

HAUSEN, J., OTTE, J. C., STRÄHLE, U., HAMMERS-WIRTZ, M., HOLLERT, H., KEITER, S. H. & OTTERMANNS, R. 2015. Fold-change threshold screening: a robust algorithm to unmask hidden gene expression patterns in noisy aggregated transcriptome data. *Environmental Science and Pollution Research,* 22**,** 16384-16392.

HEBERT, P. D. N., RATNASINGHAM, S. & DEWAARD, J. R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences,* 270**,** S96-S99.

HEDGES, S. B. 2002. The origin and evolution of model organisms. *Nat Rev Genet,* 3**,** 838-849.

HENIKOFF, S. & HENIKOFF, J. G. 1992. Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America,* 89**,** 10915-10919.

HUANG, Y., ZHU, X., WANG, L., LIU, X., LU, L., GU, W. & JIAO, Y. 2014. Genome Wide Analysis of Sex Difference in Gene Expression Profiles of Bone Formations Using sfx Mice and BXD RI Strains. *The Scientific World Journal,* 2014**,** 11.

HUPFER, T., SCHICK, J., JOZEFOWSKI, K., VOEHRINGER, D., OSTROP, J. & LANG, R. 2016. Stat6-Dependent Inhibition of Mincle Expression in Mouse and Human Antigen-Presenting Cells by the Th2 Cytokine IL-4. *Frontiers in Immunology,* 7**,** 423.

HUSEMANN, P. & STOYE, J. 2010. r2cat: synteny plots and comparative assembly. *Bioinformatics,* 26**,** 570-571.

ISG 2007. The scientific evidence. Independent Scientific Group on Cattle TB. London: DEFRA.

JACQUIER, A. 2009. The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. *Nat Rev Genet,* 10**,** 833-844.

JANEWAY, C. 2001. *Immunobiology : the immune system in health and disease,* London ; New York, NY, US, Garland Pub.

JANEWAY, C. 2005. *Immunobiology : the immune system in health and disease,* New York, Garland Science.

JANSEN, R., BATISTA, S., BROOKS, A. I., TISCHFIELD, J. A., WILLEMSEN, G., VAN GROOTHEEST, G., HOTTENGA, J.-J., MILANESCHI, Y., MBAREK, H., MADAR, V., PEYROT, W., VINK, J. M., VERWEIJ, C. L., DE GEUS, E. J., SMIT, J. H., WRIGHT, F. A., SULLIVAN, P. F., BOOMSMA, D. I. & PENNINX, B. W. 2014. Sex differences in the human peripheral blood transcriptome. *BMC Genomics,* 15**,** 33.

JENKINS, H. E., WOODROFFE, R. & DONNELLY, C. A. 2010. The Duration of the Effects of Repeated Widespread Badger Culling on Cattle Tuberculosis Following the Cessation of Culling. *PLoS ONE,* 5**,** e9090.

JENNER, R. A. & WILLS, M. A. 2007. The choice of model organisms in evo-devo. *Nat Rev Genet,* 8**,** 311-9.

KANEHISA, M., ARAKI, M., GOTO, S., HATTORI, M., HIRAKAWA, M., ITOH, M., KATAYAMA, T., KAWASHIMA, S., OKUDA, S., TOKIMATSU, T. & YAMANISHI, Y. 2008. KEGG for linking genomes to life and the environment. *Nucleic Acids Research,* 36**,** D480-D484.

KANEHISA, M. & GOTO, S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research,* 28**,** 27-30.

KANEHISA, M., SATO, Y., KAWASHIMA, M., FURUMICHI, M. & TANABE, M. 2016. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research,* 44**,** D457-D462.

KENNEDY, A. D. & DELEO, F. R. 2009. Neutrophil apoptosis and the resolution of infection. *Immunologic Research,* 43**,** 25-61.

KLEIN, S. L. & FLANAGAN, K. L. 2016. Sex differences in immune responses. *Nat Rev Immunol,* 16**,** 626-638.

KOEPFLI, K.-P., DEERE, K. A., SLATER, G. J., BEGG, C., BEGG, K., GRASSMAN, L., LUCHERINI, M., VERON, G. & WAYNE, R. K. 2008a. Multigene phylogeny of the Mustelidae: Resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. *BMC Biology,* 6**,** 10-10.

KOEPFLI, K.-P., DEERE, K. A., SLATER, G. J., BEGG, C., BEGG, K., GRASSMAN, L., LUCHERINI, M., VERON, G. & WAYNE, R. K. 2008b. Multigene phylogeny of the Mustelidae: Resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. *BMC Biology,* 6**,** 10.

KOONIN, E. V. & GALPERIN, M. Y. 2003. *Sequence - evolution - function : computational approaches in comparative genomics,* Boston, Kluwer Academic.

KREBS, J. R., INDEPENDENT SCIENTIFIC REVIEW, G., GREAT BRITAIN. MINISTRY OF AGRICULTURE, F. & FOOD 1997. *Bovine tuberculosis in cattle and badgers : report to the Rt. Hon. Dr. Jack Cummingham, MP,* London, Ministry of Agriculture, Fisheries and Food.

KRUUK, H., PARISH, T., BROWN, C. A. J. & CARRERA, J. 1979. The Use of Pasture by the European Badger (Meles meles). *Journal of Applied Ecology,* 16**,** 453-459.

KUPFERSCHMIDT, K. 2015. Buzz food. *Science,* 350**,** 267-269.

LAI, Y.-F., LIN, T.-M., WANG, C.-H., SU, P.-Y., WU, J.-T., LIN, M.-C. & ENG, H.-L. 2016. Functional polymorphisms of the TLR7 and TLR8 genes contribute to Mycobacterium tuberculosis infection. *Tuberculosis,* 98**,** 125-131.

LAMB, M. J. & SMITH, J. M. 1969. Radiation-Induced Life-Shortening in Drosophila. *Radiation Research,* 40**,** 450-464.
LAMBRECHT, B. N. & HAMMAD, H. 2009. Biology of Lung Dendritic Cells at the Origin of Asthma. *Immunity,* 31**,** 412-424.
LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-2948.
LATCHMAN, D. S. 2005. *Gene regulation : a eukaryotic perspective,* New York, Taylor & Francis.
LEE, J. H., JO, Y. H., PATNAIK, B. B., PARK, K. B., TINDWA, H., SEO, G. W., CHANDRASEKAR, R., LEE, Y. S. & HAN, Y. S. 2015. Cloning, expression analysis, and RNA interference study of a HORMA domain containing autophagy-related gene 13 (ATG13) from the coleopteran beetle, Tenebrio molitor. *Frontiers in Physiology,* 6**,** 180.
LEE, S.-W., CHUANG, T.-Y., HUANG, H.-H., LIU, C.-W., KAO, Y.-H. & WU, L. S.-H. 2016. VDR and VDBP genes polymorphisms associated with susceptibility to tuberculosis in a Han Taiwanese population. *Journal of Microbiology, Immunology and Infection,* 49**,** 783-787.
LEE, Y. & RIO, D. C. 2015. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annual Review of Biochemistry,* 84**,** 291-323.
LEONELLI, S. & ANKENY, R. A. 2013. What makes a model organism? *Endeavour,* 37**,** 209-212.
LESELLIER, S., CORNER, L., COSTELLO, E., SLEEMAN, P., LYASHCHENKO, K., GREENWALD, R., ESFANDIARI, J., SINGH, M., HEWINSON, R. G., CHAMBERS, M. & GORMLEY, E. 2008. Antigen specific immunological responses of badgers (Meles meles) experimentally infected with Mycobacterium bovis. *Veterinary Immunology and Immunopathology,* 122**,** 35-45.
LI, X. C. & JEVNIKAR, A. M. 2015. *Transplant immunology*.
LINDBLAD-TOH, K. 2004. Genome sequencing: Three's company. *Nature,* 428**,** 475-476.
LIU, K., ZHANG, B., TENG, Z., WANG, Y., DONG, G., XU, C., QIN, B., SONG, C., CHAI, J., LI, Y., SHI, X., SHU, X. & ZHANG, Y. 2017. Association between SLC11A1 (NRAMP1) polymorphisms and susceptibility to tuberculosis in Chinese Holstein cattle. *Tuberculosis,* 103**,** 10-15.
LIU, L.-N. & WANG, C.-Y. 2014a. Complete mitochondrial genome of yellow meal worm(Tenebrio molitor). *Zoological Research,* 35**,** 537-545.
LIU, L. N. & WANG, C. Y. 2014b. Complete mitochondrial genome of yellow meal worm (Tenebrio molitor). *Dongwuxue Yanjiu,* 35**,** 537-45.
LIU, S. & TRAPNELL, C. 2016. Single-cell transcriptome sequencing: recent advances and remaining challenges. *F1000Research,* 5**,** F1000 Faculty Rev-182.
MAHMOOD, K. H., ROOK, G. A. W., STANFORD, J. L., STUART, F. A. & PRITCHARD, D. G. 1987. The immunological consequences of challenge with bovine tubercle bacilli in badgers (Meles meles). *Epidemiology & Infection,* 98**,** 155-163.

MARDIS, E. & MCCOMBIE, W. R. 2017. Whole-Genome Sequencing: Manual Library Preparation. *Cold Spring Harbor Protocols,* 2017**,** pdb.prot094615.
MARMI, J., LÓPEZ-GIRÁLDEZ, F., MACDONALD, D., CALAFELL, F., ZHOLNEROVSKAYA, E. & DOMINGO-ROURA, X. 2006. Mitochondrial DNA reveals a strong phylogeographic structure in the badger across Eurasia. *Molecular Ecology,* 15**,** 1007-1020.
MARTIN, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011,* 17.
MATTICK, J. S. 2005. The functional genomics of noncoding RNA. *Science,* 309**,** 1527-8.
MCGETTIGAN, P. A. 2013. Transcriptomics in the RNA-seq era. *Current Opinion in Chemical Biology,* 17**,** 4-11.
MCKENNA, D. D., SCULLY, E. D., PAUCHET, Y., HOOVER, K., KIRSCH, R., GEIB, S. M., MITCHELL, R. F., WATERHOUSE, R. M., AHN, S.-J., ARSALA, D., BENOIT, J. B., BLACKMON, H., BLEDSOE, T., BOWSHER, J. H., BUSCH, A., CALLA, B., CHAO, H., CHILDERS, A. K., CHILDERS, C., CLARKE, D. J., COHEN, L., DEMUTH, J. P., DINH, H., DODDAPANENI, H., DOLAN, A., DUAN, J. J., DUGAN, S., FRIEDRICH, M., GLASTAD, K. M., GOODISMAN, M. A. D., HADDAD, S., HAN, Y., HUGHES, D. S. T., IOANNIDIS, P., JOHNSTON, J. S., JONES, J. W., KUHN, L. A., LANCE, D. R., LEE, C.-Y., LEE, S. L., LIN, H., LYNCH, J. A., MOCZEK, A. P., MURALI, S. C., MUZNY, D. M., NELSON, D. R., PALLI, S. R., PANFILIO, K. A., PERS, D., POELCHAU, M. F., QUAN, H., QU, J., RAY, A. M., RINEHART, J. P., ROBERTSON, H. M., ROEHRDANZ, R., ROSENDALE, A. J., SHIN, S., SILVA, C., TORSON, A. S., JENTZSCH, I. M. V., WERREN, J. H., WORLEY, K. C., YOCUM, G., ZDOBNOV, E. M., GIBBS, R. A. & RICHARDS, S. 2016. Genome of the Asian longhorned beetle (Anoplophora glabripennis), a globally significant invasive species, reveals key functional and evolutionary innovations at the beetle–plant interface. *Genome Biology,* 17**,** 227.
MCWILLIAM, H., LI, W., ULUDAG, M., SQUIZZATO, S., PARK, Y. M., BUSO, N., COWLEY, A. P. & LOPEZ, R. 2013. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Research,* 41**,** W597-W600.
MERKX-JACQUES, A., COORS, A., BROUSSEAU, R., MASSON, L., MAZZA, A., TIEN, Y. C. & TOPP, E. 2013. Evaluating the pathogenic potential of environmental Escherichia coli by using the caenorhabditis elegans infection model. *Applied and Environmental Microbiology,* 79**,** 2435-2445.
METZKER, M. L. 2010. Sequencing technologies [mdash] the next generation. *Nat Rev Genet,* 11**,** 31-46.
MICHEL, A. L., MULLER, B. & VAN HELDEN, P. D. 2010. Mycobacterium bovis at the animal-human interface: a problem, or not? *Vet Microbiol,* 140**,** 371-81.
MOIANE, I., MACHADO, A., SANTOS, N., NHAMBIR, A., INLAMEA, O., HATTENDORF, J., KÄLLENIUS, G., ZINSSTAG, J. & CORREIA-NEVES, M. 2014. Prevalence of Bovine Tuberculosis and Risk Factor Assessment in Cattle in Rural Livestock Areas of Govuro District in the Southeast of Mozambique. *PLOS ONE,* 9**,** e91527.
MORET, Y. & SIVA-JOTHY, M. T. 2003. Adaptive innate immunity? Responsive-mode prophylaxis in the mealworm beetle, Tenebrio molitor. *Proceedings of the Royal Society B: Biological Sciences,* 270**,** 2475-2480.

MORIN, R. D., BAINBRIDGE, M., FEJES, A., HIRST, M., KRZYWINSKI, M., PUGH, T. J., MCDONALD, H., VARHOL, R., JONES, S. J. M. & MARRA, M. A. 2008. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *Biotechniques,* 45**,** 81-+.

MOROZOVA, O. & MARRA, M. A. 2008. Applications of next-generation sequencing technologies in functional genomics. *Genomics,* 92**,** 255-264.

MORTAZAVI, A., WILLIAMS, B. A., MCCUE, K., SCHAEFFER, L. & WOLD, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Meth,* 5**,** 621-628.

MURHEAD, R. & BURNS, K. 1974. Tuberculosis in wild badgers in Gloucestershire: epidemiology. *Veterinary Record,* 95**,** 552-555.

NAGALAKSHMI, U., WANG, Z., WAERN, K., SHOU, C., RAHA, D., GERSTEIN, M. & SNYDER, M. 2008. The Transcriptional Landscape of the Yeast Genome Defined by RNA Sequencing. *Science,* 320**,** 1344-1349.

NATER, A., BURRI, R., KAWAKAMI, T., SMEDS, L. & ELLEGREN, H. 2015. Resolving Evolutionary Relationships in Closely Related Species with Whole-Genome Sequencing Data. *Systematic Biology,* 64**,** 1000-1017.

NEAL, E. G. & CHEESEMAN, C. 1996. *Badgers,* London, Poyser.

NEYROLLES, O. & QUINTANA-MURCI, L. 2009. Sexual Inequality in Tuberculosis. *PLOS Medicine,* 6**,** e1000199.

NHAMOYEBONDE, S. & LESLIE, A. 2014. Biological Differences Between the Sexes and Susceptibility to Tuberculosis. *The Journal of Infectious Diseases,* 209**,** S100-S106.

NUGENT, G., BUDDLE, B. M. & KNOWLES, G. 2015. Epidemiology and control of Mycobacterium bovis infection in brushtail possums (Trichosurus vulpecula), the primary wildlife host of bovine tuberculosis in New Zealand. *New Zealand Veterinary Journal,* 63**,** 28-41.

OKONIEWSKI, M. & MILLER, C. 2006. Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations. *BMC Bioinformatics,* 7**,** 276.

OLEA-POPELKA, F. J., FITZGERALD, P., WHITE, P., MCGRATH, G., COLLINS, J. D., O'KEEFFE, J., KELTON, D. F., BERKE, O., MORE, S. & MARTIN, S. W. 2009. Targeted badger removal and the subsequent risk of bovine tuberculosis in cattle herds in county Laois, Ireland. *Preventive Veterinary Medicine,* 88**,** 178-184.

OLIVEIRA-CORTEZ, A., MELO, A. C., CHAVES, V. E., CONDINO-NETO, A. & CAMARGOS, P. 2016. Do HLA class II genes protect against pulmonary tuberculosis? A systematic review and meta-analysis. *European Journal of Clinical Microbiology & Infectious Diseases,* 35**,** 1567-1580.

OPPERT, B., DOWD, S. E., BOUFFARD, P., LI, L., CONESA, A., LORENZEN, M. D., TOUTGES, M., MARSHALL, J., HUESTIS, D. L., FABRICK, J., OPPERT, C. & JURAT-FUENTES, J. L. 2012a. Transcriptome profiling of the intoxication response of Tenebrio molitor larvae to Bacillus thuringiensis Cry3Aa protoxin. *PloS one,* 7.

OPPERT, B., DOWD, S. E., BOUFFARD, P., LI, L., CONESA, A., LORENZEN, M. D., TOUTGES, M., MARSHALL, J., HUESTIS, D. L., FABRICK, J., OPPERT, C. & JURAT-FUENTES, J. L. 2012b. Transcriptome Profiling of the Intoxication Response of Tenebrio molitor Larvae to Bacillus thuringiensis Cry3Aa Protoxin. *PLoS ONE,* 7**,** e34624.

ORTUTAY, C., SIERMALA, M. & VIHINEN, M. 2007. Molecular characterization of the immune system: emergence of proteins, processes, and domains. *Immunogenetics,* 59**,** 333-348.

ORTUTAY, C. & VIHINEN, M. 2009. Immunome Knowledge Base (IKB): An integrated service for immunome research. *BMC Immunology,* 10**,** 3-3.

PAMILO, P. & NEI, M. 1988. Relationships between gene trees and species trees. *Molecular Biology and Evolution,* 5**,** 568-583.

PARK, J.-W., KIM, C.-H., RUI, J., KEUN, H., RYU, K.-H., CHAI, J.-H., HWANG, H.-O., KUROKAWA, K., HA, N.-C., SÖDERHÄLL, I., SÖDERHÄLL, K. & LEE, B. 2010. Beetle Immunity. *In:* SÖDERHÄLL, K. (ed.) *Invertebrate Immunity.* Springer US.

PARK, S. H., JIANG, R., PIAO, S., ZHANG, B., KIM, E.-H., KWON, H.-M., JIN, X. L., LEE, B. L. & HA, N.-C. 2011. Structural and Functional Characterization of a Highly Specific Serpin in the Insect Innate Immunity. *The Journal of Biological Chemistry,* 286**,** 1567-1575.

PENG, R., ZENG, B., MENG, X., YUE, B., ZHANG, Z. & ZOU, F. 2007. The complete mitochondrial genome and phylogenetic analysis of the giant panda (Ailuropoda melanoleuca). *Gene,* 397**,** 76-83.

PERSSON, P. B. & MUELLER, M. 2015. Transcription. *Acta Physiologica,* 215**,** 159-160.

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

ROBINSON, P. A., CORNER, L. A. L., COURCIER, E. A., MCNAIR, J., ARTOIS, M., MENZIES, F. D. & ABERNETHY, D. A. 2012. BCG vaccination against tuberculosis in European badgers (Meles meles): A review. *Comparative Immunology, Microbiology and Infectious Diseases,* 35**,** 277-287.

ROGERS, L. M., FORRESTER, G. J., WILSON, G. J., YARNELL, R. W. & CHEESEMAN, C. L. 2003. The role of setts in badger (Meles meles) group size, breeding success and status of TB (Mycobacterium bovis). *Journal of Zoology,* 260**,** 209-215.

ROSENBLUETH, A. & WIENER, N. 1945. The Role of Models in Science. *Philosophy of Science,* 12**,** 316-321.

ROSENSTEIN, J. 2014. The Promise of Nanopore Technology: Nanopore DNA sequencing represents a fundamental change in the way that genomic information is read, with potentially big savings. *Pulse, IEEE,* 5**,** 52-54.

SCHURCH, N. J., SCHOFIELD, P., GIERLIŃSKI, M., COLE, C., SHERSTNEV, A., SINGH, V., WROBEL, N., GHARBI, K., SIMPSON, G. G., OWEN-HUGHES, T., BLAXTER, M. & BARTON, G. J. 2016. How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA,* 22**,** 839-851.

SCHWOCHOW, D., SERIEYS, L. E. K., WAYNE, R. K. & THALMANN, O. 2012. Efficient recovery of whole blood RNA - a comparison of commercial RNA extraction protocols for high-throughput applications in wildlife species. *BMC Biotechnology,* 12**,** 33-33.

SEVERI, F. & CONTICELLO, S. G. 2015. Flow-cytometric visualization of C>U mRNA editing reveals the dynamics of the process in live cells. *RNA Biology,* 12**,** 389-397.

SIEMIANOWSKA, E., KOSEWSKA, A., ALJEWICZ, M., SKIBNIEWSKA, K. A., POLAK-JUSZCZAK, L., JAROCKI, A. & J?DRAS, M. 2013. Larvae of mealworm (<i>Tenebrio molitor</i> L.) as European novel food. *Agricultural Sciences,* Vol.04No.06**,** 5.

SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SÖDING, J., THOMPSON, J. D. & HIGGINS, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology,* 7.

SIN, Y. W., DUGDALE, H. L., NEWMAN, C., MACDONALD, D. W. & BURKE, T. 2011. MHC class II genes in the European badger (Meles meles): characterization, patterns of variation, and transcription analysis. *Immunogenetics,* 64**,** 313-327.

SIN, Y. W., DUGDALE, H. L., NEWMAN, C., MACDONALD, D. W. & BURKE, T. 2012. Evolution of MHC class I genes in the European badger (Meles meles). *Ecology and Evolution,* 2**,** 1644-1662.

SKINNER, M. A., WEDLOCK, D. N. & BUDDLE, B. M. 2001. Vaccination of animals against Mycobacterium bovis. *Rev Sci Tech,* 20**,** 112-32.

SMITH, G. C., MCDONALD, R. A. & WILKINSON, D. 2012. Comparing Badger (<italic>Meles meles</italic>) Management Strategies for Reducing Tuberculosis Incidence in Cattle. *PLoS ONE,* 7**,** e39250.

SMÝKAL, P., K. VARSHNEY, R., K. SINGH, V., COYNE, C. J., DOMONEY, C., KEJNOVSKÝ, E. & WARKENTIN, T. 2016. From Mendel's discovery on pea to today's plant genetics and breeding. *Theoretical and Applied Genetics,* 129**,** 2267-2280.

SOMPAYRAC, L. 2012. *How the immune system works,* Chichester, West Sussex ; Hoboken, NJ, Wiley-Blackwell.

TANNE, A. & NEYROLLES, O. 2010. C-type lectins in immune defense against pathogens: the murine DC-SIGN homologue SIGNR3 confers early protection against Mycobacterium tuberculosis infection. *Virulence,* 1**,** 285-290.

TATUSOV, R. L., GALPERIN, M. Y., NATALE, D. A. & KOONIN, E. V. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Research,* 28**,** 33-36.

TOH, H., SHIRANE, K., MIURA, F., KUBO, N., ICHIYANAGI, K., HAYASHI, K., SAITOU, M., SUYAMA, M., ITO, T. & SASAKI, H. 2017. Software updates in the Illumina HiSeq platform affect whole-genome bisulfite sequencing. *BMC Genomics,* 18**,** 31.

TOMLINSON, A. J., CHAMBERS, M. A., WILSON, G. J., MCDONALD, R. A. & DELAHAY, R. J. 2013. Sex-related heterogeneity in the life-history correlates of Mycobacterium bovis infection in European badgers (Meles meles). *Transbound Emerg Dis,* 60 Suppl 1**,** 37-45.

TRABZUNI, D., RAMASAMY, A., IMRAN, S., WALKER, R., SMITH, C., WEALE, M. E., HARDY, J., RYTEN, M. & NORTH AMERICAN BRAIN EXPRESSION, C. 2013. Widespread sex differences in gene expression and splicing in the adult human brain. *Nat Commun,* 4**,** 2771.

TROWSDALE, J. & KNIGHT, J. C. 2013. Major Histocompatibility Complex Genomics and Human Disease. *Annual review of genomics and human genetics,* 14**,** 301-323.

TURVEY, S. E. & BROIDE, D. H. 2010a. Chapter 2: Innate Immunity. *The Journal of allergy and clinical immunology,* 125**,** S24-S32.

TURVEY, S. E. & BROIDE, D. H. 2010b. Innate immunity. *Journal of Allergy and Clinical Immunology,* 125**,** S24-S32.

VERA, J. C., WHEAT, C. W., FESCEMYER, H. W., FRILANDER, M. J., CRAWFORD, D. L., HANSKI, I. & MARDEN, J. H. 2008. Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Molecular Ecology,* 17**,** 1636-1647.

VIGNERON, N., STROOBANT, V., CHAPIRO, J., OOMS, A., DEGIOVANNI, G., MOREL, S., VAN DER BRUGGEN, P., BOON, T. & VAN DEN EYNDE, B. J. 2004. An Antigenic Peptide Produced by Peptide Splicing in the Proteasome. *Science,* 304**,** 587.

WANG, Z., GERSTEIN, M. & SNYDER, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet,* 10**,** 57-63.

WHEELER, D. L., BARRETT, T., BENSON, D. A., BRYANT, S. H., CANESE, K., CHETVERNIN, V., CHURCH, D. M., DICUCCIO, M., EDGAR, R., FEDERHEN, S., FEOLO, M., GEER, L. Y., HELMBERG, W., KAPUSTIN, Y., KHOVAYKO, O., LANDSMAN, D., LIPMAN, D. J., MADDEN, T. L., MAGLOTT, D. R., MILLER, V., OSTELL, J., PRUITT, K. D., SCHULER, G. D., SHUMWAY, M., SEQUEIRA, E., SHERRY, S. T., SIROTKIN, K., SOUVOROV, A., STARCHENKO, G., TATUSOV, R. L., TATUSOVA, T. A., WAGNER, L. & YASCHENKO, E. 2008. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research,* 36**,** D13-D21.

WHEELER, G. N. & BRANDLI, A. W. 2009. Simple vertebrate models for chemical genetics and drug discovery screens: lessons from zebrafish and Xenopus. *Dev Dyn,* 238**,** 1287-308.

WILL, C. L. & LUHRMANN, R. 2011. Spliceosome Structure and Function. *Cold Spring Harbor Perspectives in Biology,* 3.

WOODROFFE, R., DONNELLY, C. A., JOHNSTON, W. T., BOURNE, F. J., CHEESEMAN, C. L., CLIFTON-HADLEY, R. S., COX, D. R., GETTINBY, G., HEWINSON, R. G., LE FEVRE, A. M., MCINERNEY, J. P. & MORRISON, W. I. 2005. Spatial association of Mycobacterium bovis infection in cattle and badgers Meles meles. *Journal of Applied Ecology,* 42**,** 852-862.

YANDELL, M. & ENCE, D. 2012. A beginner's guide to eukaryotic genome annotation. *Nature Reviews Genetics,* 13**,** 329-342.

YANG, X., SCHADT, E. E., WANG, S., WANG, H., ARNOLD, A. P., INGRAM-DRAKE, L., DRAKE, T. A. & LUSIS, A. J. 2006. Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Research,* 16**,** 995-1004.

YE, L., HILLIER, L. W., MINX, P., THANE, N., LOCKE, D. P., MARTIN, J. C., CHEN, L., MITREVA, M., MILLER, J. R., HAUB, K. V., DOOLING, D. J., MARDIS, E. R., WILSON, R. K., WEINSTOCK, G. M. & WARREN, W. C. 2011. A vertebrate case study of the quality of assemblies derived from next-generation sequences. *Genome Biology,* 12.

ZHANG, L., WU, Q., HU, Y., WU, H. & WEI, F. 2015. Major histocompatibility complex alleles associated with parasite susceptibility in wild giant pandas. *Heredity,* 114**,** 85-93.

ZHANG, W., BLEIBEL, W. K., ROE, C. A., COX, N. J. & DOLAN, M. E. 2007. Gender-specific differences in expression in human lymphoblastoid cell lines. *Pharmacogenetics and genomics,* 17**,** 447-450.

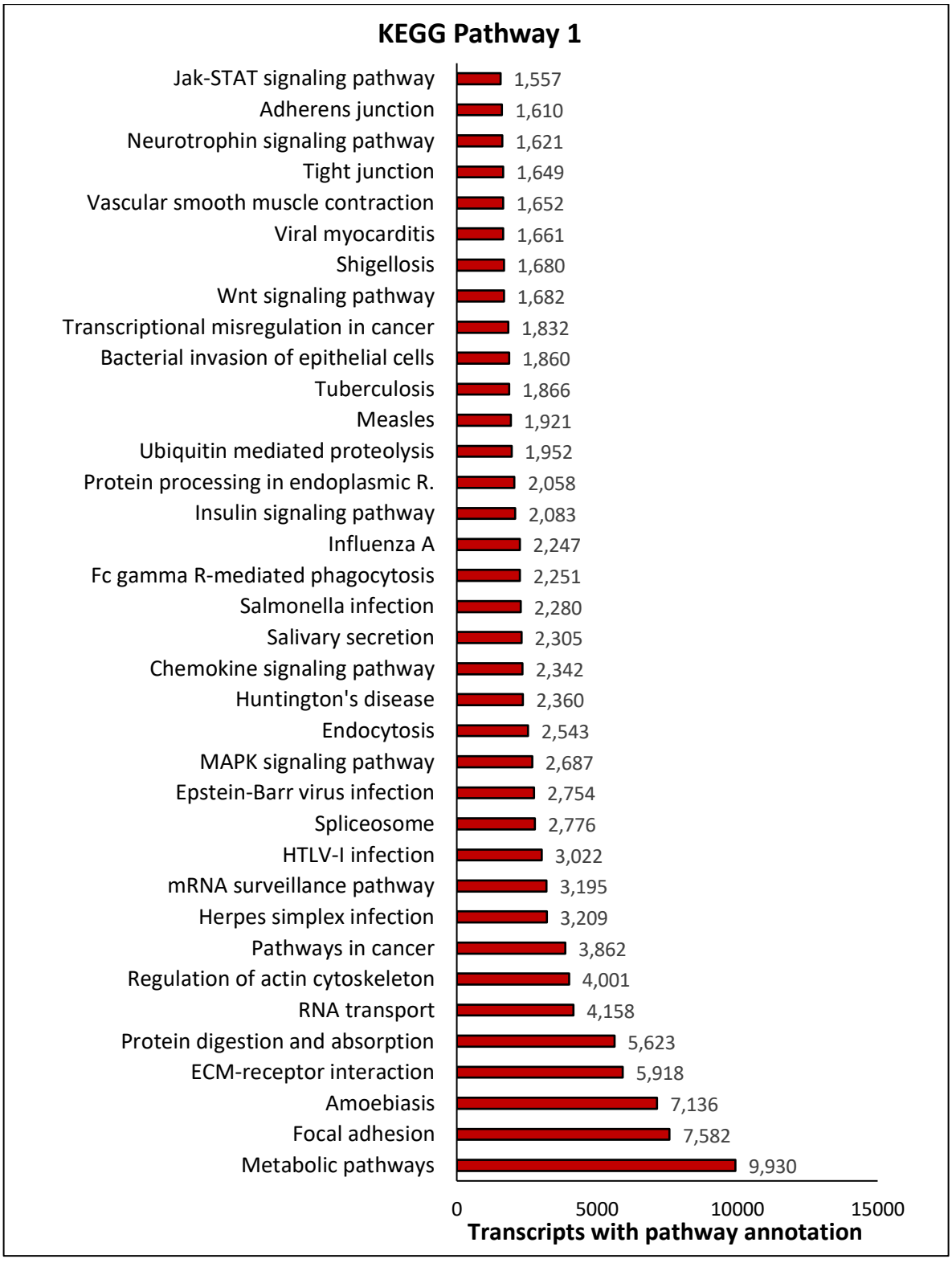
ZHANG, W., HUANG, R. S., DUAN, S. & DOLAN, M. E. 2009. Gene set enrichment analyses revealed differences in gene expression patterns between males and females. *In silico biology,* 9**,** 55-63.

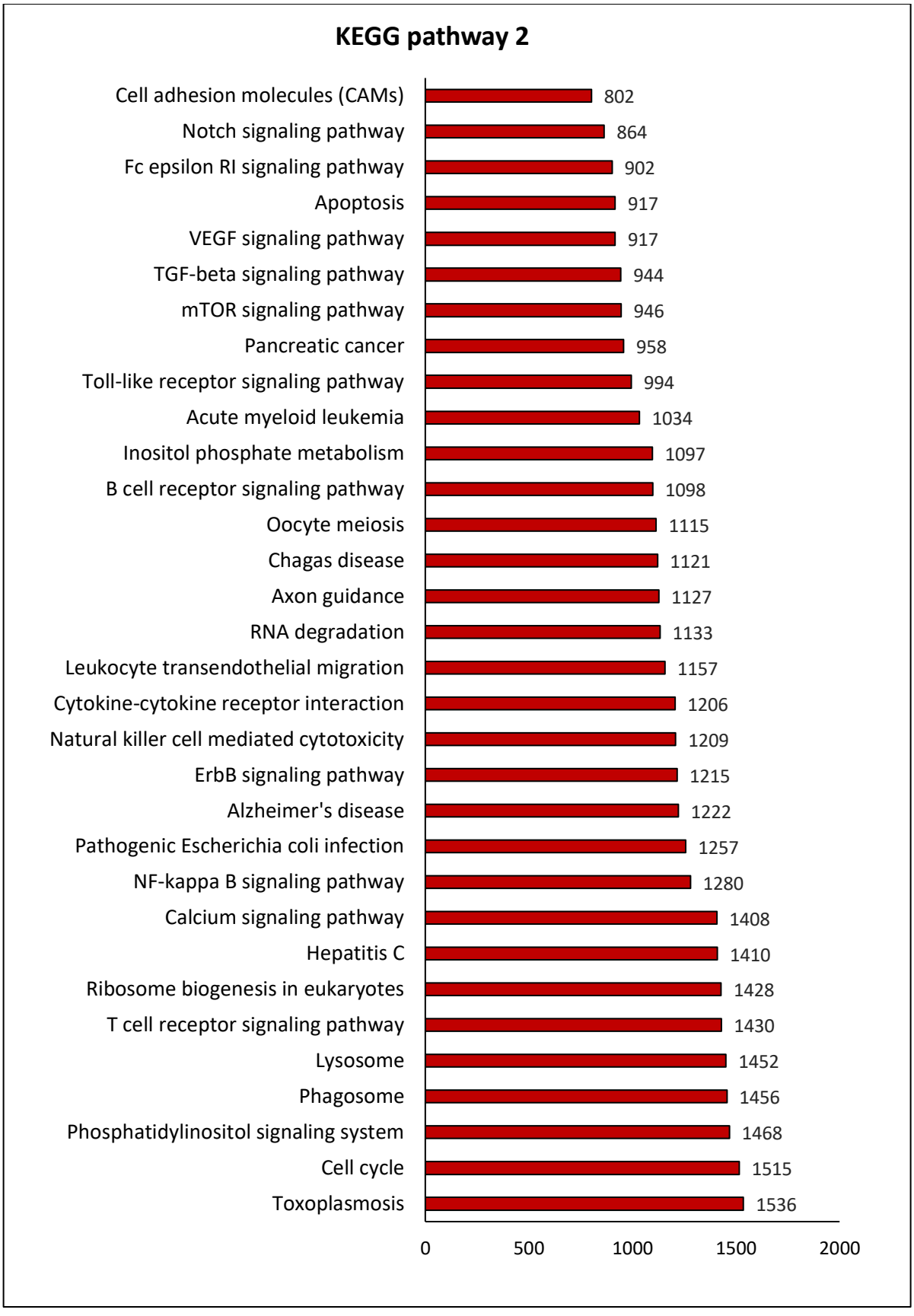
ZHOU, H., CHEN, S., WANG, M. & CHENG, A. 2014. Interferons and Their Receptors in Birds: A Comparison of Gene Structure, Phylogenetic Analysis, and Cross Modulation. *International Journal of Molecular Sciences,* 15**,** 21045-21068.

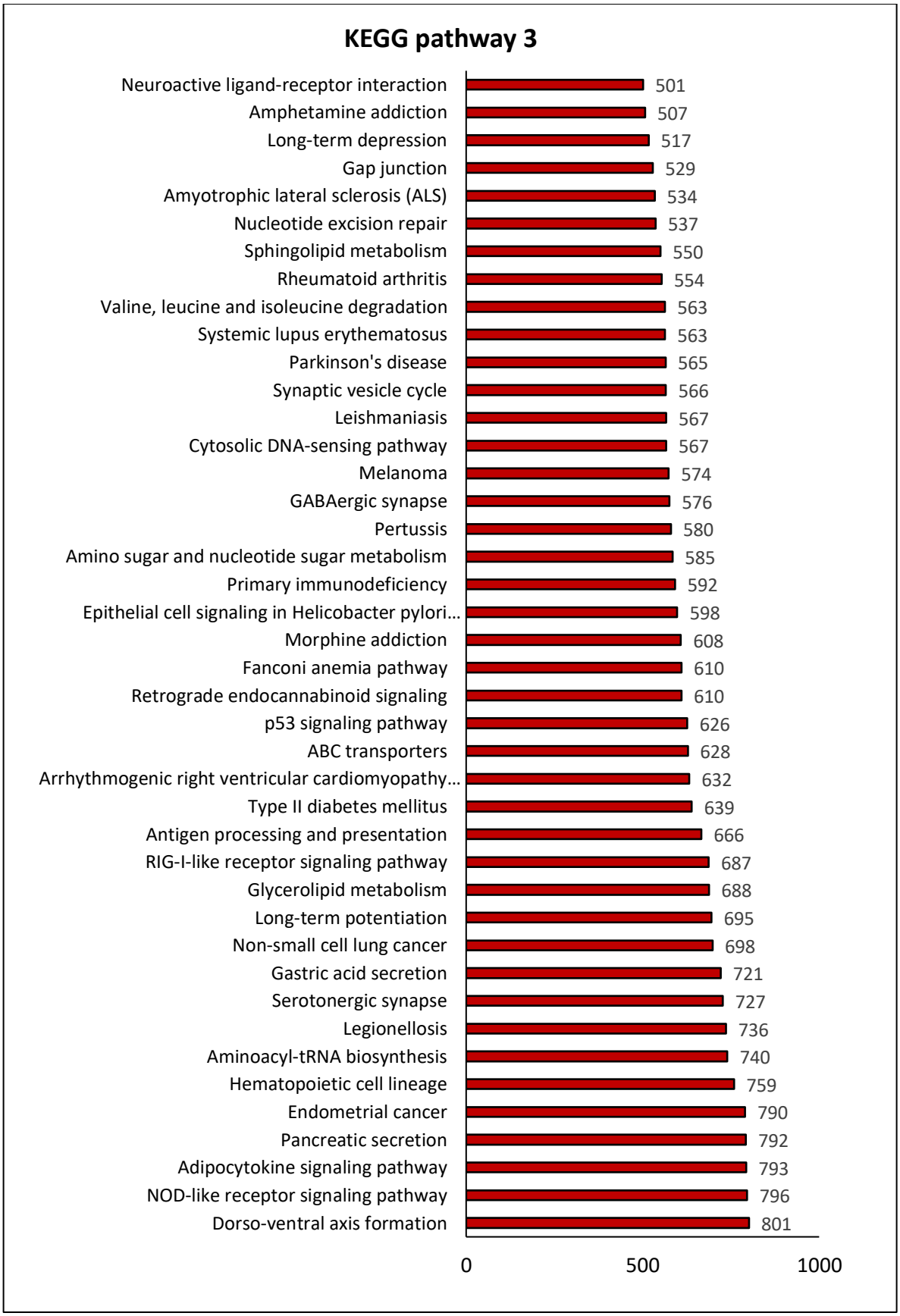
ZHU, L., RUAN, X.-D., GE, Y.-F., WAN, Q.-H. & FANG, S.-G. 2007. Low major histocompatibility complex class II DQA diversity in the Giant Panda (Ailuropoda melanoleuca). *BMC Genetics,* 8**,** 29-29.

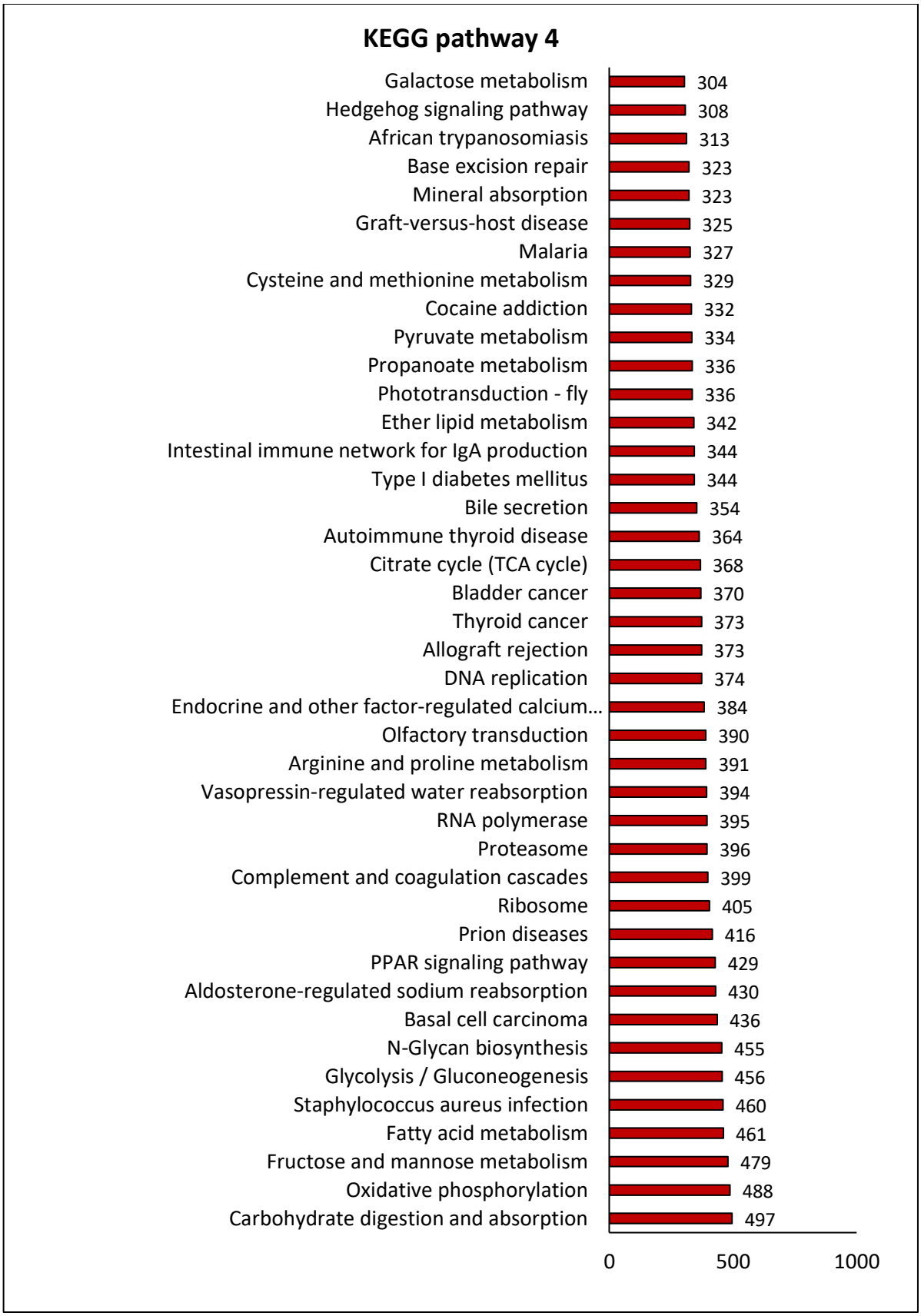
Appendix

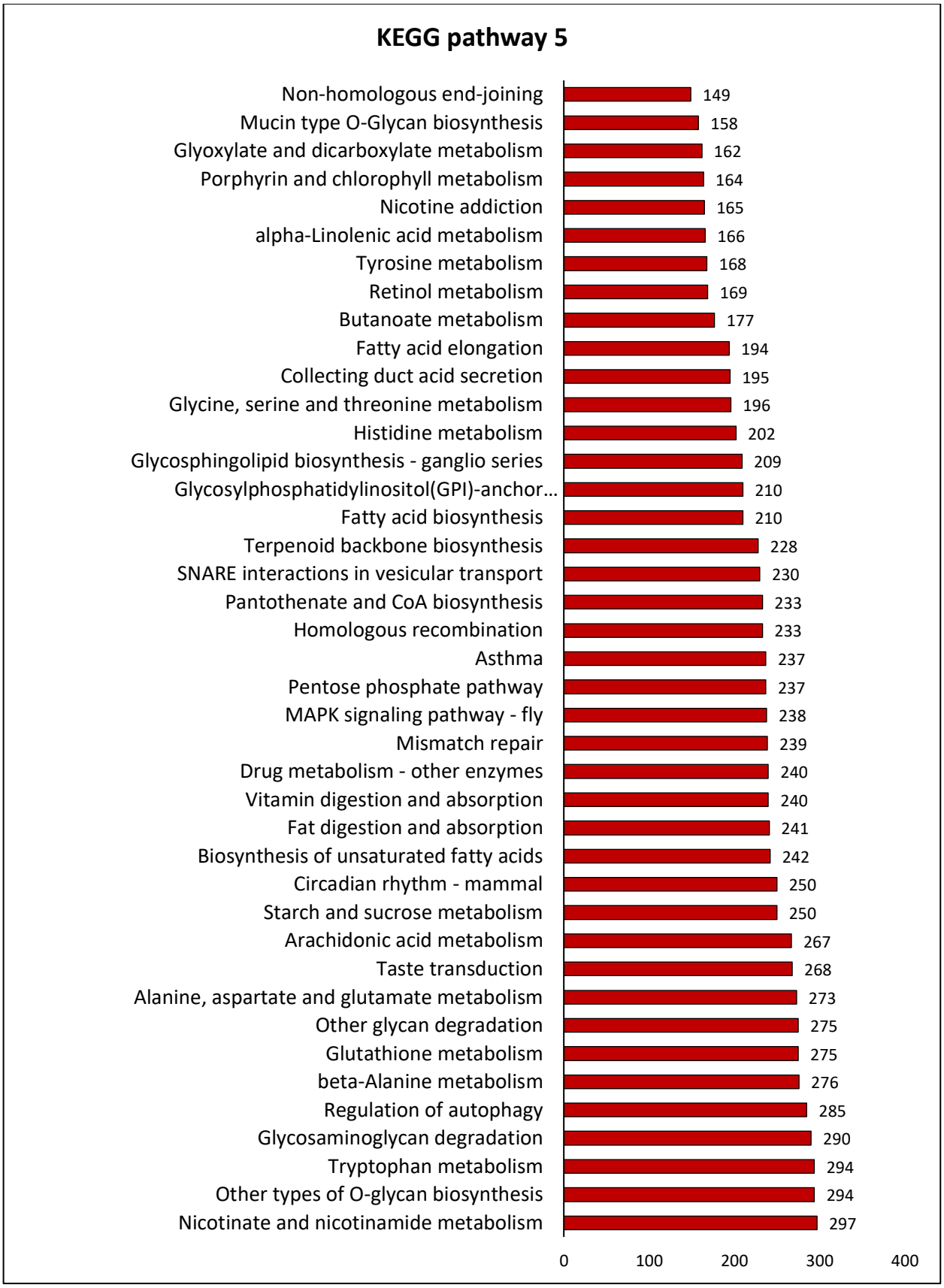
8.1 KEGG Pathways

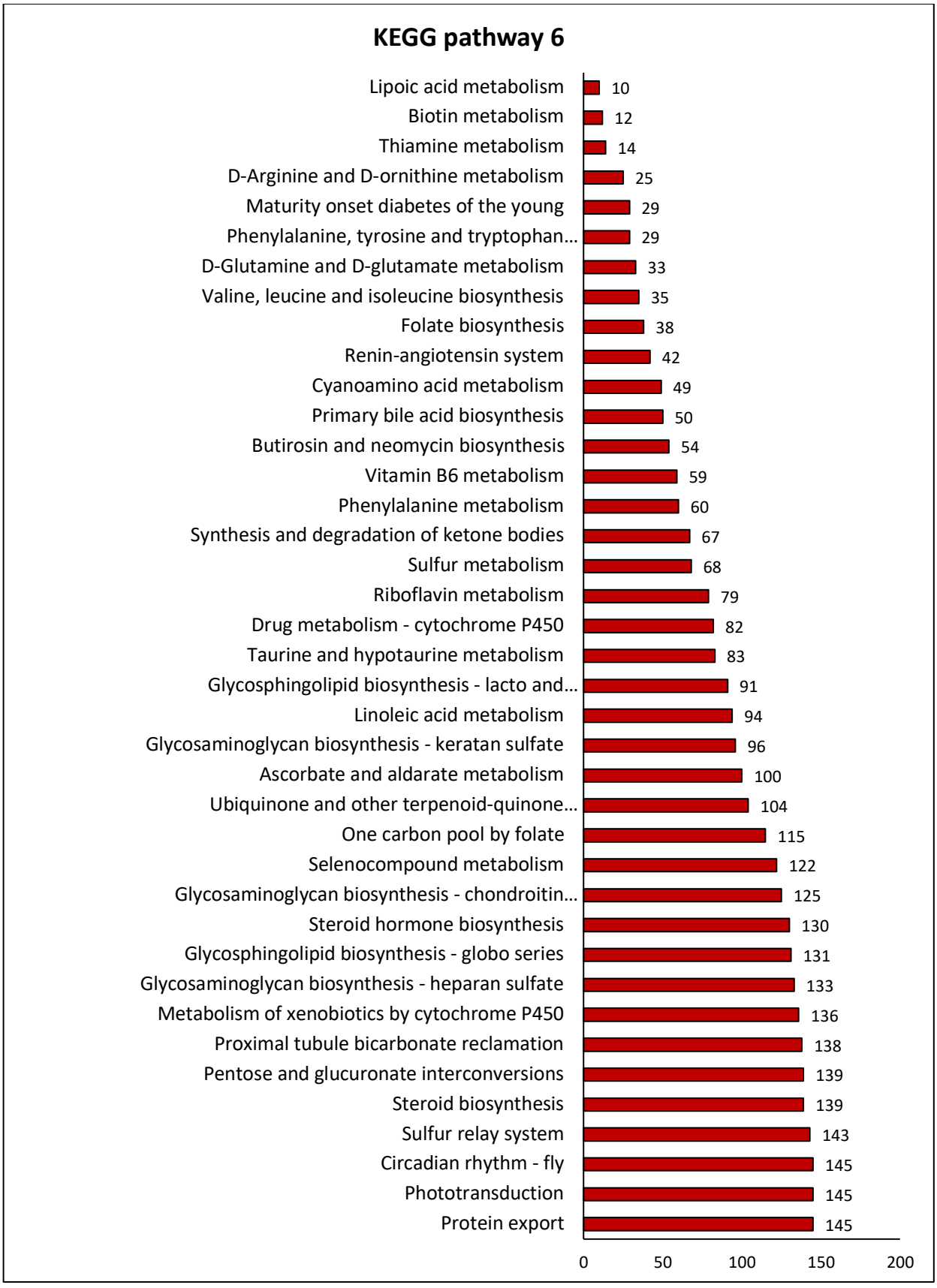



KEGG pathway 2
Cell adhesion molecules (CAMs) 802
Notch signaling pathway 864
Fc epsilon RI signaling pathway 902
Apoptosis 917
VEGF signaling pathway 917
TGF-beta signaling pathway 944
mTOR signaling pathway 946
Pancreatic cancer 958
Toll-like receptor signaling pathway 994
Acute myeloid leukemia 1034
Inositol phosphate metabolism 1097
B cell receptor signaling pathway 1098
Oocyte meiosis 1115
Chagas disease 1121
Axon guidance 1127
RNA degradation 1133
Leukocyte transendothelial migration 1157
Cytokine-cytokine receptor interaction 1206
Natural killer cell mediated cytotoxicity 1209
ErbB signaling pathway 1215
Alzheimer's disease 1222
Pathogenic Escherichia coli infection 1257
NF-kappa B signaling pathway 1280
Calcium signaling pathway 1408
Hepatitis C 1410
Ribosome biogenesis in eukaryotes 1428
T cell receptor signaling pathway 1430
Lysosome 1452
Phagosome 1456
Phosphatidylinositol signaling system 1468
Cell cycle 1515
Toxoplasmosis 1536
0
500
1000
1500
2000


KEGG pathway 3
Neuroactive ligand-receptor interaction 501
Amphetamine addiction 507
Long-term depression 517
Gap junction 529
Amyotrophic lateral sclerosis (ALS) 534
Nucleotide excision repair 537
Sphingolipid metabolism 550
Rheumatoid arthritis 554
Valine, leucine and isoleucine degradation 563
Systemic lupus erythematosus 563
Parkinson's disease 565
Synaptic vesicle cycle 566
Leishmaniasis 567
Cytosolic DNA-sensing pathway 567
Melanoma 574
GABAergic synapse 576
Pertussis 580
Amino sugar and nucleotide sugar metabolism 585
Primary immunodeficiency 592
Epithelial cell signaling in Helicobacter pylori… 598
Morphine addiction 608
Fanconi anemia pathway 610
Retrograde endocannabinoid signaling 610
p53 signaling pathway 626
ABC transporters 628
Arrhythmogenic right ventricular cardiomyopathy… 632
Type II diabetes mellitus 639
Antigen processing and presentation 666
RIG-I-like receptor signaling pathway 687
Glycerolipid metabolism 688
Long-term potentiation 695
Non-small cell lung cancer 698
Gastric acid secretion 721
Serotonergic synapse 727
Legionellosis 736
Aminoacyl-tRNA biosynthesis 740
Hematopoietic cell lineage 759
Endometrial cancer 790
Pancreatic secretion 792
Adipocytokine signaling pathway 793
NOD-like receptor signaling pathway 796
Dorso-ventral axis formation 801
0
500
1000


KEGG pathway 4
Galactose metabolism 304
Hedgehog signaling pathway 308
African trypanosomiasis 313
Base excision repair 323
Mineral absorption 323
Graft-versus-host disease 325
Malaria 327
Cysteine and methionine metabolism 329
Cocaine addiction 332
Pyruvate metabolism 334
Propanoate metabolism 336
Phototransduction - fly 336
Ether lipid metabolism 342
Intestinal immune network for IgA production 344
Type I diabetes mellitus 344
Bile secretion 354
Autoimmune thyroid disease 364
Citrate cycle (TCA cycle) 368
Bladder cancer 370
Thyroid cancer 373
Allograft rejection 373
DNA replication 374
Endocrine and other factor-regulated calcium… 384
Olfactory transduction 390
Arginine and proline metabolism 391
Vasopressin-regulated water reabsorption 394
RNA polymerase 395
Proteasome 396
Complement and coagulation cascades 399
Ribosome 405
Prion diseases 416
PPAR signaling pathway 429
Aldosterone-regulated sodium reabsorption 430
Basal cell carcinoma 436
N-Glycan biosynthesis 455
Glycolysis / Gluconeogenesis 456
Staphylococcus aureus infection 460
Fatty acid metabolism 461
Fructose and mannose metabolism 479
Oxidative phosphorylation 488
Carbohydrate digestion and absorption 497
0
500
1000


KEGG pathway 5
Non-homologous end-joining 149
Mucin type O-Glycan biosynthesis 158
Glyoxylate and dicarboxylate metabolism 162
Porphyrin and chlorophyll metabolism 164
Nicotine addiction 165
alpha-Linolenic acid metabolism 166
Tyrosine metabolism 168
Retinol metabolism 169
Butanoate metabolism 177
Fatty acid elongation 194
Collecting duct acid secretion 195
Glycine, serine and threonine metabolism 196
Histidine metabolism 202
Glycosphingolipid biosynthesis - ganglio series 209
Glycosylphosphatidylinositol(GPI)-anchor... 210
Fatty acid biosynthesis 210
Terpenoid backbone biosynthesis 228
SNARE interactions in vesicular transport 230
Pantothenate and CoA biosynthesis 233
Homologous recombination 233
Asthma 237
Pentose phosphate pathway 237
MAPK signaling pathway - fly 238
Mismatch repair 239
Drug metabolism - other enzymes 240
Vitamin digestion and absorption 240
Fat digestion and absorption 241
Biosynthesis of unsaturated fatty acids 242
Circadian rhythm - mammal 250
Starch and sucrose metabolism 250
Arachidonic acid metabolism 267
Taste transduction 268
Alanine, aspartate and glutamate metabolism 273
Other glycan degradation 275
Glutathione metabolism 275
beta-Alanine metabolism 276
Regulation of autophagy 285
Glycosaminoglycan degradation 290
Tryptophan metabolism 294
Other types of O-glycan biosynthesis 294
Nicotinate and nicotinamide metabolism 297
0
100
200
300
400


KEGG pathway 6
Lipoic acid metabolism 10
Biotin metabolism 12
Thiamine metabolism 14
D-Arginine and D-ornithine metabolism 25
Maturity onset diabetes of the young 29
Phenylalanine, tyrosine and tryptophan... 29
D-Glutamine and D-glutamate metabolism 33
Valine, leucine and isoleucine biosynthesis 35
Folate biosynthesis 38
Renin-angiotensin system 42
Cyanoamino acid metabolism 49
Primary bile acid biosynthesis 50
Butirosin and neomycin biosynthesis 54
Vitamin B6 metabolism 59
Phenylalanine metabolism 60
Synthesis and degradation of ketone bodies 67
Sulfur metabolism 68
Riboflavin metabolism 79
Drug metabolism - cytochrome P450 82
Taurine and hypotaurine metabolism 83
Glycosphingolipid biosynthesis - lacto and... 91
Linoleic acid metabolism 94
Glycosaminoglycan biosynthesis - keratan sulfate 96
Ascorbate and aldarate metabolism 100
Ubiquinone and other terpenoid-quinone... 104
One carbon pool by folate 115
Selenocompound metabolism 122
Glycosaminoglycan biosynthesis - chondroitin... 125
Steroid hormone biosynthesis 130
Glycosphingolipid biosynthesis - globo series 131
Glycosaminoglycan biosynthesis - heparan sulfate 133
Metabolism of xenobiotics by cytochrome P450 136
Proximal tubule bicarbonate reclamation 138
Pentose and glucuronate interconversions 139
Steroid biosynthesis 139
Sulfur relay system 143
Circadian rhythm - fly 145
Phototransduction 145
Protein export 145
0
50
100
150
200

8.2 BLASTN output (tabular)

Query Seq-id	Subject seq-id	Identity %	Alignment length	E-vlue	Query seq. length	Subject seq. length	Sequence classification
JQ425440.1	CL1150.Contig16_All	90.12	506	0	848	972	MHC I
JQ425440.1	CL1150.Contig16_All	87.46	295	2.00E-92	848	972	MHC I
JQ425440.1	CL1150.Contig4_All	89.33	506	2.00E-180	848	702	MHC I
JQ425440.1	CL1150.Contig15_All	92.87	421	4.00E-173	848	450	MHC I
JQ425440.1	CL1150.Contig14_All	95.42	306	5.00E-137	848	540	MHC I
JQ425440.1	CL1150.Contig14_All	96.19	105	6.00E-42	848	540	MHC I
JQ425440.1	CL1150.Contig7_All	89.56	297	1.00E-103	848	399	MHC I
JQ425440.1	CL1150.Contig18_All	88.47	295	2.00E-97	848	417	MHC I
JQ425440.1	CL1150.Contig17_All	87.46	295	2.00E-92	848	420	MHC I
JQ425440.1	CL1150.Contig13_All	85.14	296	8.00E-81	848	426	MHC I
JQ425440.1	CL1150.Contig8_All	94.51	182	1.00E-74	848	507	MHC I
JQ425440.1	CL1150.Contig8_All	96.19	105	6.00E-42	848	507	MHC I
JQ425440.1	CL1150.Contig5_All	85.8	162	6.00E-42	848	198	MHC I
JQ425440.1	Unigene117550_All	100	83	2.00E-36	848	171	MHC I
JQ425440.1	Unigene12618_All	100	83	2.00E-36	848	171	MHC I
JQ425440.1	Unigene26995_All	100	73	8.00E-31	848	126	MHC I
JQ425440.1	CL1793.Contig7_All	85.53	76	4.00E-14	848	297	MHC I
JQ425440.1	CL1793.Contig5_All	85.53	76	4.00E-14	848	297	MHC I
JQ425440.1	CL1793.Contig3_All	85.53	76	4.00E-14	848	297	MHC I
JQ425440.1	CL1793.Contig2_All	85.53	76	4.00E-14	848	297	MHC I
JQ425439.1	CL1150.Contig16_All	90.04	562	0	932	972	MHC I
JQ425439.1	CL1150.Contig16_All	87.46	295	2.00E-92	932	972	MHC I
JQ425439.1	CL1150.Contig4_All	89.09	550	0	932	702	MHC I
JQ425439.1	CL1150.Contig15_All	91.25	297	8.00E-111	932	450	MHC I
JQ425439.1	CL1150.Contig15_All	96.85	127	4.00E-54	932	450	MHC I
JQ425439.1	CL1150.Contig7_All	89.6	298	4.00E-104	932	399	MHC I
JQ425439.1	CL1150.Contig18_All	88.47	295	2.00E-97	932	417	MHC I
JQ425439.1	CL1150.Contig17_All	87.46	295	2.00E-92	932	420	MHC I
JQ425439.1	CL1150.Contig13_All	85.14	296	8.00E-81	932	426	MHC I
JQ425439.1	CL1150.Contig8_All	94.54	183	3.00E-75	932	507	MHC I
JQ425439.1	CL1150.Contig8_All	96.19	105	7.00E-42	932	507	MHC I
JQ425439.1	CL1150.Contig14_All	94.51	182	1.00E-74	932	540	MHC I
JQ425439.1	CL1150.Contig14_All	96.85	127	4.00E-54	932	540	MHC I
JQ425439.1	CL1150.Contig14_All	96.19	105	7.00E-42	932	540	MHC I
JQ425439.1	CL1150.Contig5_All	85.8	162	7.00E-42	932	198	MHC I
JQ425439.1	Unigene117550_All	100	84	7.00E-37	932	171	MHC I
JQ425439.1	Unigene12618_All	100	84	7.00E-37	932	171	MHC I

JQ425439.1	Unigene26995_All	100	73	9.00E-31	932	126	MHC I
JQ425439.1	CL1793.Contig7_All	85.53	76	4.00E-14	932	297	MHC I
JQ425439.1	CL1793.Contig5_All	85.53	76	4.00E-14	932	297	MHC I
JQ425439.1	CL1793.Contig3_All	85.53	76	4.00E-14	932	297	MHC I
JQ425439.1	CL1793.Contig2_All	85.53	76	4.00E-14	932	297	MHC I
JQ425438.1	CL1150.Contig16_All	90.8	848	0	900	972	MHC I
JQ425438.1	CL1150.Contig4_All	91.31	587	0	900	702	MHC I
JQ425438.1	CL1150.Contig15_All	96.91	324	1.00E-153	900	450	MHC I
JQ425438.1	CL1150.Contig17_All	90.17	417	2.00E-152	900	420	MHC I
JQ425438.1	CL1150.Contig18_All	89.42	416	6.00E-147	900	417	MHC I
JQ425438.1	CL1150.Contig13_All	87.5	424	7.00E-136	900	426	MHC I
JQ425438.1	CL1150.Contig7_All	88.89	387	2.00E-132	900	399	MHC I
JQ425438.1	CL1150.Contig8_All	92.72	261	1.00E-103	900	507	MHC I
JQ425438.1	CL1150.Contig14_All	93.89	180	6.00E-72	900	540	MHC I
JQ425438.1	CL1150.Contig24_All	92.26	155	2.00E-56	900	156	MHC I
JQ425438.1	CL1150.Contig23_All	92.26	155	2.00E-56	900	156	MHC I
JQ425438.1	CL1150.Contig3_All	92.26	155	2.00E-56	900	156	MHC I
JQ425438.1	CL1150.Contig1_All	92.26	155	2.00E-56	900	156	MHC I
JQ425438.1	CL1150.Contig2_All	91.67	156	3.00E-55	900	159	MHC I
JQ425438.1	CL1150.Contig22_All	99.09	110	1.00E-49	900	360	MHC I
JQ425432.1	CL1150.Contig16_All	92.08	972	0	975	972	MHC I
JQ425432.1	CL1150.Contig4_All	93.28	699	0	975	702	MHC I
JQ425432.1	CL1150.Contig15_All	96.91	324	1.00E-153	975	450	MHC I
JQ425432.1	CL1150.Contig17_All	89.69	417	4.00E-149	975	420	MHC I
JQ425432.1	CL1150.Contig18_All	89.42	416	6.00E-147	975	417	MHC I
JQ425432.1	CL1150.Contig13_All	87.5	424	8.00E-136	975	426	MHC I
JQ425432.1	CL1150.Contig7_All	88.89	387	2.00E-132	975	399	MHC I
JQ425432.1	CL1150.Contig8_All	93.1	261	3.00E-105	975	507	MHC I
JQ425432.1	CL1150.Contig14_All	94.44	180	1.00E-73	975	540	MHC I
JQ425432.1	CL1150.Contig24_All	92.9	155	5.00E-58	975	156	MHC I
JQ425432.1	CL1150.Contig23_All	92.9	155	5.00E-58	975	156	MHC I
JQ425432.1	CL1150.Contig3_All	92.9	155	5.00E-58	975	156	MHC I
JQ425432.1	CL1150.Contig1_All	92.9	155	5.00E-58	975	156	MHC I
JQ425432.1	CL1150.Contig2_All	92.31	156	7.00E-57	975	159	MHC I
JQ425432.1	CL1150.Contig22_All	100	110	3.00E-51	975	360	MHC I
JQ425433.1	CL1150.Contig16_All	94.44	972	0	975	972	MHC I
JQ425433.1	CL1150.Contig4_All	98.71	699	0	975	702	MHC I
JQ425433.1	CL1150.Contig15_All	98.15	324	3.00E-160	975	450	MHC I
JQ425433.1	CL1150.Contig17_All	89.45	417	2.00E-147	975	420	MHC I
JQ425433.1	CL1150.Contig18_All	89.18	416	3.00E-145	975	417	MHC I
JQ425433.1	CL1150.Contig13_All	87.47	423	8.00E-136	975	426	MHC I

JQ425433.1	CL1150.Contig7_All	88.37	387	4.00E-129	975	399	MHC I
JQ425433.1	CL1150.Contig8_All	91.95	261	3.00E-100	975	507	MHC I
JQ425433.1	CL1150.Contig14_All	92.78	180	1.00E-68	975	540	MHC I
JQ425433.1	CL1150.Contig2_All	94.23	156	7.00E-62	975	159	MHC I
JQ425433.1	CL1150.Contig24_All	93.55	155	1.00E-59	975	156	MHC I
JQ425433.1	CL1150.Contig23_All	93.55	155	1.00E-59	975	156	MHC I
JQ425433.1	CL1150.Contig3_All	93.55	155	1.00E-59	975	156	MHC I
JQ425433.1	CL1150.Contig1_All	93.55	155	1.00E-59	975	156	MHC I
JQ425431.1	CL1150.Contig16_All	91.98	972	0	975	972	MHC I
JQ425431.1	CL1150.Contig4_All	92.7	699	0	975	702	MHC I
JQ425431.1	CL1150.Contig15_All	96.91	324	1.00E-153	975	450	MHC I
JQ425431.1	CL1150.Contig17_All	90.17	417	2.00E-152	975	420	MHC I
JQ425431.1	CL1150.Contig18_All	89.42	416	6.00E-147	975	417	MHC I
JQ425431.1	CL1150.Contig13_All	87.5	424	8.00E-136	975	426	MHC I
JQ425431.1	CL1150.Contig7_All	88.89	387	2.00E-132	975	399	MHC I
JQ425431.1	CL1150.Contig8_All	92.72	261	1.00E-103	975	507	MHC I
JQ425431.1	CL1150.Contig14_All	93.89	180	7.00E-72	975	540	MHC I
JQ425431.1	CL1150.Contig24_All	92.26	155	3.00E-56	975	156	MHC I
JQ425431.1	CL1150.Contig23_All	92.26	155	3.00E-56	975	156	MHC I
JQ425431.1	CL1150.Contig3_All	92.26	155	3.00E-56	975	156	MHC I
JQ425431.1	CL1150.Contig1_All	92.26	155	3.00E-56	975	156	MHC I
JQ425431.1	CL1150.Contig2_All	91.67	156	3.00E-55	975	159	MHC I
JQ425431.1	CL1150.Contig22_All	99.09	110	1.00E-49	975	360	MHC I
JQ425430.1	CL1150.Contig16_All	93.44	975	0	975	972	MHC I
JQ425430.1	CL1150.Contig4_All	93.87	702	0	975	702	MHC I
JQ425430.1	CL1150.Contig17_All	90.41	417	4.00E-154	975	420	MHC I
JQ425430.1	CL1150.Contig18_All	90.14	416	6.00E-152	975	417	MHC I
JQ425430.1	CL1150.Contig15_All	96.6	324	6.00E-152	975	450	MHC I
JQ425430.1	CL1150.Contig13_All	88.42	423	2.00E-142	975	426	MHC I
JQ425430.1	CL1150.Contig7_All	89.41	387	8.00E-136	975	399	MHC I
JQ425430.1	CL1150.Contig8_All	91.95	261	3.00E-100	975	507	MHC I
JQ425430.1	CL1150.Contig14_All	92.78	180	1.00E-68	975	540	MHC I
JQ425430.1	CL1150.Contig2_All	94.23	156	7.00E-62	975	159	MHC I
JQ425430.1	CL1150.Contig24_All	93.55	155	1.00E-59	975	156	MHC I
JQ425430.1	CL1150.Contig23_All	93.55	155	1.00E-59	975	156	MHC I
JQ425430.1	CL1150.Contig3_All	93.55	155	1.00E-59	975	156	MHC I
JQ425430.1	CL1150.Contig1_All	93.55	155	1.00E-59	975	156	MHC I
JQ425430.1	CL7631.Contig3_All	91.84	98	3.00E-31	975	108	MHC I
JQ425430.1	CL1150.Contig22_All	84.68	111	6.00E-23	975	360	MHC I
JQ425429.1	CL1150.Contig16_All	95.16	972	0	975	972	MHC I
JQ425429.1	CL1150.Contig4_All	98	699	0	975	702	MHC I

JQ425429.1	CL1150.Contig15_All	98.15	324	3.00E-160	975	450	MHC I
JQ425429.1	CL1150.Contig17_All	89.45	417	2.00E-147	975	420	MHC I
JQ425429.1	CL1150.Contig18_All	89.18	416	3.00E-145	975	417	MHC I
JQ425429.1	CL1150.Contig13_All	87.47	423	8.00E-136	975	426	MHC I
JQ425429.1	CL1150.Contig7_All	88.37	387	4.00E-129	975	399	MHC I
JQ425429.1	CL1150.Contig8_All	91.95	261	3.00E-100	975	507	MHC I
JQ425429.1	CL1150.Contig14_All	92.78	180	1.00E-68	975	540	MHC I
JQ425429.1	CL1150.Contig2_All	94.23	156	7.00E-62	975	159	MHC I
JQ425429.1	CL1150.Contig24_All	93.55	155	1.00E-59	975	156	MHC I
JQ425429.1	CL1150.Contig23_All	93.55	155	1.00E-59	975	156	MHC I
JQ425429.1	CL1150.Contig3_All	93.55	155	1.00E-59	975	156	MHC I
JQ425429.1	CL1150.Contig1_All	93.55	155	1.00E-59	975	156	MHC I
JQ425428.1	CL1150.Contig16_All	94.32	546	0	543	972	MHC I
JQ425428.1	CL1150.Contig4_All	92.49	546	0	543	702	MHC I
JQ425428.1	CL1150.Contig22_All	85.59	111	7.00E-25	543	360	MHC I
JQ425427.1	CL1150.Contig4_All	93.59	546	0	543	702	MHC I
JQ425427.1	CL1150.Contig16_All	93.04	546	0	543	972	MHC I
JQ425427.1	CL1150.Contig22_All	85.59	111	7.00E-25	543	360	MHC I
JQ425427.1	CL1150.Contig22_All	97.06	68	2.00E-24	543	360	MHC I
HQ908107.1	Unigene57188_All	97.18	674	0	680	795	MHC II
HQ908107.1	Unigene67843_All	97.23	650	0	680	651	MHC II
HQ908107.1	Unigene75449_All	97.28	551	0	680	774	MHC II
HQ908107.1	Unigene75450_All	82.72	654	7.00E-160	680	774	MHC II
HQ908107.1	CL4065.Contig3_All	82.86	566	1.00E-142	680	693	MHC II
HQ908107.1	Unigene27967_All	80.07	552	3.00E-113	680	774	MHC II
HQ908107.1	Unigene4377_All	100	96	1.00E-43	680	96	MHC II
HQ908107.1	CL4065.Contig4_All	100	96	1.00E-43	680	96	MHC II
HQ908107.1	CL4065.Contig2_All	100	96	1.00E-43	680	96	MHC II
HQ908107.1	CL4065.Contig1_All	100	96	1.00E-43	680	96	MHC II
HQ908099.1	CL2981.Contig3_All	99.27	688	0	691	762	MHC II
HQ908099.1	CL2981.Contig1_All	99.27	688	0	691	762	MHC II
HQ908099.1	Unigene42913_All	99.67	603	0	691	603	MHC II
HQ908098.1	CL10050.Contig3_All	99.67	615	0	615	765	MHC II
HQ908098.1	CL10050.Contig2_All	100	582	0	615	582	MHC II
HQ908098.1	CL10050.Contig1_All	100	582	0	615	582	MHC II
HQ908097.1	CL10050.Contig3_All	91.73	617	0	615	765	MHC II
HQ908097.1	CL10050.Contig2_All	91.61	584	0	615	582	MHC II
HQ908097.1	CL10050.Contig1_All	91.61	584	0	615	582	MHC II
HQ908096.1	Unigene57188_All	99.7	674	0	697	795	MHC II
HQ908096.1	Unigene67843_All	99.85	650	0	697	651	MHC II
HQ908096.1	Unigene75449_All	97.28	551	0	697	774	MHC II

HQ908096.1	Unigene75450_All	85.32	654	0	697	774	MHC II
HQ908096.1	CL4065.Contig3_All	83.25	573	3.00E-148	697	693	MHC II
HQ908096.1	Unigene27967_All	80.07	552	4.00E-113	697	774	MHC II
HQ908096.1	Unigene4377_All	100	96	1.00E-43	697	96	MHC II
HQ908096.1	CL4065.Contig4_All	100	96	1.00E-43	697	96	MHC II
HQ908096.1	CL4065.Contig2_All	100	96	1.00E-43	697	96	MHC II
HQ908096.1	CL4065.Contig1_All	100	96	1.00E-43	697	96	MHC II
HQ908095.1	Unigene27967_All	94.83	774	0	822	774	MHC II
HQ908095.1	Unigene75450_All	94.6	741	0	822	774	MHC II
HQ908095.1	CL4065.Contig3_All	96.64	684	0	822	693	MHC II
HQ908095.1	Unigene75449_All	80.62	774	1.00E-168	822	774	MHC II
HQ908095.1	Unigene57188_All	79.89	741	2.00E-151	822	795	MHC II
HQ908095.1	Unigene67843_All	80.7	596	4.00E-128	822	651	MHC II
HQ908095.1	Unigene81593_All	93.62	94	6.00E-32	822	93	MHC II
HQ908094.1	Unigene27967_All	99.22	774	0	822	774	MHC II
HQ908094.1	Unigene75450_All	97.17	672	0	822	774	MHC II
HQ908094.1	CL4065.Contig3_All	97.17	672	0	822	693	MHC II
HQ908094.1	Unigene75449_All	84.09	773	0	822	774	MHC II
HQ908094.1	Unigene57188_All	79.88	671	5.00E-137	822	795	MHC II
HQ908094.1	Unigene67843_All	80.8	526	1.00E-113	822	651	MHC II
HQ908094.1	Unigene81593_All	93.62	94	6.00E-32	822	93	MHC II
HQ908093.1	Unigene27967_All	99.48	774	0	822	774	MHC II
HQ908093.1	Unigene75450_All	97.62	672	0	822	774	MHC II
HQ908093.1	CL4065.Contig3_All	97.62	672	0	822	693	MHC II
HQ908093.1	Unigene75449_All	83.44	773	0	822	774	MHC II
HQ908093.1	Unigene57188_All	79.28	671	2.00E-130	822	795	MHC II
HQ908093.1	Unigene67843_All	80.04	526	5.00E-107	822	651	MHC II
HQ908092.1	Unigene27967_All	96.51	774	0	822	774	MHC II
HQ908092.1	Unigene75450_All	92.31	741	0	822	774	MHC II
HQ908092.1	CL4065.Contig3_All	94.79	672	0	822	693	MHC II
HQ908092.1	Unigene75449_All	81.57	776	2.00E-180	822	774	MHC II
HQ908092.1	Unigene57188_All	76.85	743	4.00E-113	822	795	MHC II
HQ908092.1	Unigene67843_All	77.09	598	2.00E-91	822	651	MHC II
HQ908092.1	CL13896.Contig1_All	100	96	1.00E-43	822	96	MHC II
Y11647.2	Unigene54563_All	100	204	7.00E-104	501	204	Interferon Gamma

8.3 BLASTN output (pairwise)

```
BLASTN 2.2.31+


Reference: Zheng Zhang, Scott Schwartz, Lukas Wagner, and Webb
Miller (2000), "A greedy algorithm for aligning DNA sequences", J
Comput Biol 2000; 7(1-2):203-14.



Database: badgernt
           127,401 sequences; 85,063,401 total letters



Query= JQ425440.1 Meles meles MHC class I antigen (Meme-MHCI) pseudogene
mRNA, Meme-MHCI*PS03 allele, partial sequence

Length=848
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  654     0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  632     2e-180
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  608     4e-173
CL1150.Contig14_All  3 542 minus strand hypothetical protein PAND...  488     5e-137
CL1150.Contig7_All  20 418 minus strand MHC class I antigen [Ailu...  377     1e-103
CL1150.Contig18_All  1278 1694 minus strand MHC class I antigen [...  357     2e-97
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  340     2e-92
CL1150.Contig13_All  621 1046 minus strand MHC class I antigen, p...  302     8e-81
CL1150.Contig8_All  20 526 hypothetical protein PANDA_022308 [Ail...  281     1e-74
CL1150.Contig5_All  2 199 minus strand MHC class I antigen [Bos t...  172     6e-42
Unigene117550_All  775 945 PREDICTED: patr class I histocompatibi...  154     2e-36
Unigene12618_All  419 589 PREDICTED: patr class I histocompatibil...  154     2e-36
Unigene26995_All  2 127 minus strand DLA class I histocompatibili...  135     8e-31
CL1793.Contig7_All  65 361 minus strand PREDICTED: MHC class I po...  80.5    4e-14
CL1793.Contig5_All  65 361 minus strand PREDICTED: MHC class I po...  80.5    4e-14
CL1793.Contig3_All  65 361 PREDICTED: MHC class I polypeptide-rel...  80.5    4e-14
CL1793.Contig2_All  65 361 minus strand PREDICTED: MHC class I po...  80.5    4e-14



Query_1              25   GGCGGCGAGGAGCCGCGCTTCATCTCCGTCGGCTACGTGGACTTCACGCAGTTCGTGCGG  84
CL1150.Contig16_All  58   GGCCGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGG  117
CL1150.Contig4_All   49   GGCCGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGG  108

Query_1              85   TTCGACAGCGACTCTGCCAGTCAGAGAAGGA-GGAGCCGCGGGCGCCGTAAGTGGAGCAG  143
CL1150.Contig16_All  118  TTCGACAGCGACTCTGCCAGTC-G-G-AGGATGGAGCCGCGGGCGCCGTGGATGGAGCAG  174
CL1150.Contig4_All   109  TTCGACAGCGACTCTGCCAGTC-G-GA-GGATGGAGCCGCGGGCGCCGTGGATGGAGCAG  165

Query_1              144  GAGGGGCCGGAGTATTGGGACGAGGAGACGCGGATCTGCAAGGAAACCACACAGACTTAC  203
CL1150.Contig16_All  175  GAGGGGCCGGAGTATTGGGACCGGCAGACGCAGATCTGCAAGGAAACCACACAGACTTAC  234
CL1150.Contig4_All   166  GAGGGGCCGGAGTATTGGGACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTAC  225

Query_1              204  CGAGGGAGCCTGAACATCCTGCGGGGCTACTACAACCAGAGCGAGGCCGGGTCTCACACC  263
CL1150.Contig16_All  235  CGAGGGAGCCTGAACATCCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACC  294
CL1150.Contig4_All   226  CGACGGAGCCTGAACAACCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACC  285

Query_1              264  ATCCAGCGCATGTACGGCTGTGACGTGGGGCCCGACGGGCGCCTCCTCCGCGGGTACAGT  323
CL1150.Contig16_All  295  ATCCAGAACTTGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGT  354
CL1150.Contig4_All   286  TTCCAGAACATGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGT  345

Query_1              324  CAGGACGCCTACGACGGCGCGGATTACCTCACCCTGAACGAGGACCTGCGCTCCTGGACC  383
CL1150.Contig16_All  355  CAGTTCGCCTACGACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACC  414
CL1150.Contig4_All   346  CAGTTCGCCTACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACC  405

Query_1              384  GCGGCGGATGCGTCGGCGCAGATCACCCAGCGCAAGTGGGAGGACGCGGGTGAGGCAGAG  443
CL1150.Contig16_All  415  GCGGCGGACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAG  474
CL1150.Contig4_All   406  GCGGCGGACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAG  465
CL1150.Contig15_All  1                      CAGATCTCCCGGCGCAAGTGAGAGGACGCGGGTGTGGCAGAG  42
CL1150.Contig14_All  1                      CAGATCTCCCGGCGCAAGTGAGAGGACGCGGGTGTGGCAGAG  42

Query_1              444  CTTGAGAGGGACTACCTGGAGATTACTTGCGTGAAGTGGCTCCACAGGTATCTGGAGAAC  503
CL1150.Contig16_All  475  CGCTACAGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAAC  534
CL1150.Contig4_All   466  CGCTACAGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAAC  525
CL1150.Contig15_All  43   CTTGAGAGGGACTACCTGGAGATTACTTGCGTGAAGTGGCTCCACAGGTATCTGGAGAAC  102
CL1150.Contig14_All  43   CTTGAGAGGGACTACCTGGAGATTACTTGCGTGAAGTGGCTCCACAGGTATCTGGAGAAC  102

Query_1              504  GGGAAGGAGACGCTACTGCGCACAGAGATCACCCTGACCTGGCAGAGGGATGGA-GAGGA  562
CL1150.Contig16_All  535  GGGAAGGAGTCGCTGCTGCGCGCAGA                                  560
```

```
CL1150.Contig16_All  646                          GAGATCACCCTGACCTGGCAGCGAGATGGA-GAGGA  680
CL1150.Contig4_All   526  GGGAAGGAGTCGCTGCTGCGCGCAGA                                    551
CL1150.Contig15_All  103  GGGAAGGAGACGCTACTGCGCACAGAGATCACCCTGACCTGGCAGAGGGATGGA-GAGGA  161
CL1150.Contig14_All  103  GGGAAGGAGACGCTACTGCGCACAGAGATCACCCTGACCTGGCAGAGGGATGGA-GAGGA  161
CL1150.Contig7_All   92                         CAGAGATCACCCTGACCTGGCAGAGGGATGGA-GAGGA  128
CL1150.Contig18_All  91                           GAGATCACCCTGACCTGGCAGAGGGATGGA-GAGGA  125
CL1150.Contig17_All  94                           GAGATCACCCTGACCTGGCAGCGAGATGGA-GAGGA  128
CL1150.Contig13_All  94                           GAGATCACCCTGACCTGGCACCATGA-GGAGGAGGA  128
CL1150.Contig8_All   92                         CAGAGATCACCCTGACCTGGCAGAGGGATGGA-GAGGA  128
Unigene117550_All    89                         CAGAGATCACCCTGACCTGGCAGAGGGATGGA-GAGGA  125
Unigene12618_All     89                         CAGAGATCACCCTGACCTGGCAGAGGGATGGA-GAGGA  125

Query_1              563  CCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGAGATGGAACCTTCCAGAA  622
CL1150.Contig16_All  681  CCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCAGGAGATGGAACCTTCCAGAA  740
CL1150.Contig15_All  162  CCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGAGATGGAACCTTCCAGAA  221
CL1150.Contig14_All  162  CCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGAGATGGAACCTTCCAGAA  221
CL1150.Contig7_All   129  CCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGAGATGGAACCTTCCAGAA  188
CL1150.Contig18_All  126  CCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCAGGAGATGGAACCTTCCAGAA  185
CL1150.Contig17_All  129  CCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCAGGAGATGGAACCTTCCAGAA  188
CL1150.Contig13_All  129  CCTGACCCAGGACACAGAACTTGTAGGGACCAGGCCTACAGGGAATGGAACCTTCCAGAA  188
CL1150.Contig8_All   129  CCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGAGATGGAACCTTCCAGAA  188
Unigene117550_All    126  CCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGAGAT               171
Unigene12618_All     126  CCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGAGAT               171

Query_1              623  GTGGGTGGCCATGGTGGTGCCTTCTGGACAGGAGCAGAGATACACGTGCCATGTGCAGCA  682
CL1150.Contig16_All  741  GTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATACACATGCCATGTGCAGCA  800
CL1150.Contig15_All  222  GTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAGATACACATGCTATGTGCAGCA  281
CL1150.Contig14_All  222  GTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATACACATGCCATGTGCAGCA  281
CL1150.Contig7_All   189  GTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATACACATGCCATGTGCAGCA  248
CL1150.Contig18_All  186  GTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATACACATGCCATGTGCAGCA  245
CL1150.Contig17_All  189  GTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATACACATGCCATGTGCAGCA  248
CL1150.Contig13_All  189  GTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATACACATGCCATGTGCAGCA  248
CL1150.Contig8_All   189  GTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATACACATGCCATGTGCAGCA  248
Unigene26995_All     1                          TCTGGACAGGAGCAGAGATACACGTGCCATGTGCAGCA  38

Query_1              683  TGAGGGACTGTCTGAGCCCATCACCCAGAGATGGGAGCCGCCACATCCTACCATCCCCAT  742
CL1150.Contig16_All  801  TAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCTCCTCCCACCATCCCCAT  860
CL1150.Contig15_All  282  TGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCCACCTCG-C--ACCATCCCCAT  338
CL1150.Contig14_All  282  TAAGGGGCTGCCTGAGCCCATCACC                                    306
CL1150.Contig14_All  436                                    GAGCCACCTCCTCCCACCATCCCCAT  461
CL1150.Contig7_All   249  TAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCTCCTCCCACCATCCCCAT  308
CL1150.Contig18_All  246  TAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCTCCTCCCACCATCCCCAT  305
CL1150.Contig17_All  249  TAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCTCCTCCCACCATCCCCAT  308
CL1150.Contig13_All  249  TAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCTCCTCCCACCATCCCCAT  308
CL1150.Contig8_All   249  TAAGGGGCTGCCTGAGCCCATCACC                                    273
CL1150.Contig8_All   403                                    GAGCCACCTCCTCCCACCATCCCCAT  428
CL1150.Contig5_All   1     GAGGGACTGTCTGAGCCCATTACCTTGAGATGGGAGCCTCCTCTTCCCATCGTCCTCAT  59
Unigene26995_All     39   TGAGGGACTGTCTGAGCCCATCACCCAGAGATGGG                          73

Query_1              743  CATGTGGATCATTGCTGGTCTGATTCTCCCCGTGGTCTTTGCAGTGATTAGAGCTGTGAT  802
CL1150.Contig16_All  861  CATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTGGTGGTTGGAGCTGTGAT  920
CL1150.Contig15_All  339  CACATGGATCATTGCTGGTCTGGTTCTCCTGGTGATCATTGCAGTGATTGGAGTTGCAAT  398
CL1150.Contig14_All  462  CATGTGGATCATTGCTGGTCTGATTCTCCCCGTGGTCTTTGCAGTGATTAGAGCTGTGAT  521
CL1150.Contig7_All   309  CATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTGGTGGTTGGAGCTGTGAT  368
CL1150.Contig18_All  306  CATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTGGTGGTTGGAGCTGTGAT  365
CL1150.Contig17_All  309  CATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTGGTGGTTGGAGCTGTGAT  368
CL1150.Contig13_All  309  CATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTGGTGGTTGGAGCTGTGAT  368
CL1150.Contig8_All   429  CATGTGGATCATTGCTGGTCTGATTCTCCCCGTGGTCTTTGCAGTGATTAGAGCTGTGAT  488
CL1150.Contig5_All   60   CACATGGATCATTGCTGTTCTGGCTCTCCTTGTGGTCGCTGTGGTGGTTGGAGCTGTGAT  119
CL1793.Contig7_All   1        TGGATCATTGCTGTTCTGGCTCTCCTTGTGGTCGCTGTGGTGGTTGGAGCTGTGAT  56
CL1793.Contig5_All   1        TGGATCATTGCTGTTCTGGCTCTCCTTGTGGTCGCTGTGGTGGTTGGAGCTGTGAT  56
CL1793.Contig3_All   1        TGGATCATTGCTGTTCTGGCTCTCCTTGTGGTCGCTGTGGTGGTTGGAGCTGTGAT  56
CL1793.Contig2_All   1        TGGATCATTGCTGTTCTGGCTCTCCTTGTGGTCGCTGTGGTGGTTGGAGCTGTGAT  56

Query_1              803  CTGGAGGAAGAAGCGCTCAGATGATGACAGTGCCCAGGGCTCT  845
CL1150.Contig16_All  921  CTGGAGGAGGAAGCGCTCAG                         940
CL1150.Contig15_All  399  CTGGTGGAAGAAGCGCTCAG                         418
CL1150.Contig14_All  522  CTGGAGGAAGAAGCGCTCA                          540
CL1150.Contig7_All   369  CTGGAGGAAGAAGCGCTCAG                         388
CL1150.Contig18_All  366  CTGGAGGAGGAAGCGCTCAG                         385
CL1150.Contig17_All  369  CTGGAGGAGGAAGCGCTCAG                         388
CL1150.Contig13_All  369  CTGGAGGAAGAAGCGCTCAG                         388
CL1150.Contig8_All   489  CTGGAGGAAGAAGCGCTCA                          507
CL1150.Contig5_All   120  CTGGAGGAAGAAGCGCTCAGATGATGACAGTGCCCAGGGCTCT  162
CL1793.Contig7_All   57   CTGGAGGAAGAAGCGCTCAG                         76
CL1793.Contig5_All   57   CTGGAGGAAGAAGCGCTCAG                         76
CL1793.Contig3_All   57   CTGGAGGAAGAAGCGCTCAG                         76
CL1793.Contig2_All   57   CTGGAGGAAGAAGCGCTCAG                         76


Lambda      K        H
```

```
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 67199301450


Query= JQ425439.1 Meles meles MHC class I antigen (Meme-MHCI) pseudogene
mRNA, Meme-MHCI*PS02 allele, partial sequence

Length=932
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  725     0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  680     0.0
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  401     8e-111
CL1150.Contig7_All  20 418 minus strand MHC class I antigen [Ailu...  379     4e-104
CL1150.Contig18_All  1278 1694 minus strand MHC class I antigen [...  357     2e-97
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  340     2e-92
CL1150.Contig13_All  621 1046 minus strand MHC class I antigen, p...  302     8e-81
CL1150.Contig8_All  20 526 hypothetical protein PANDA_022308 [Ail...  283     3e-75
CL1150.Contig14_All  3 542 minus strand hypothetical protein PAND...  281     1e-74
CL1150.Contig5_All  2 199 minus strand MHC class I antigen [Bos t...  172     7e-42
Unigene117550_All  775 945 PREDICTED: patr class I histocompatibi...  156     7e-37
Unigene12618_All  419 589 PREDICTED: patr class I histocompatibil...  156     7e-37
Unigene26995_All  2 127 minus strand DLA class I histocompatibili...  135     9e-31
CL1793.Contig7_All  65 361 minus strand PREDICTED: MHC class I po...  80.5    4e-14
CL1793.Contig5_All  65 361 minus strand PREDICTED: MHC class I po...  80.5    4e-14
CL1793.Contig3_All  65 361 PREDICTED: MHC class I polypeptide-rel...  80.5    4e-14
CL1793.Contig2_All  65 361 minus strand PREDICTED: MHC class I po...  80.5    4e-14


Query_2              17   GAGACCTGGGCGGGCTCCCACTCCCTGAGATATTTCGACACCGCGGTTTCCCCGCCCGGC  76
CL1150.Contig16_All  1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  60
CL1150.Contig4_All   4                GGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  51

Query_2              77   AGCGAGGAGCCGCGCTTCATCTCCGTCGGCTACGTGGACTTCACGCAGTTCGTGCGGTTC  136
CL1150.Contig16_All  61   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  120
CL1150.Contig4_All   52   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  111

Query_2              137  GACAGCGACTCTGCCAGTCAGAGAAGGA-GGAGCCGCGGGCGCCGTAAGTGGAGCAGGAG  195
CL1150.Contig16_All  121  GACAGCGACTCTGCCAGTC-G-G-AGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAG  177
CL1150.Contig4_All   112  GACAGCGACTCTGCCAGTC-G-GA-GGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAG  168

Query_2              196  GGGCCGGAGTATTGGGACGAGGAGACGCGGATCTGCAAGGAAACCACACAGACTTACCGA  255
CL1150.Contig16_All  178  GGGCCGGAGTATTGGGACCGGCAGACGCAGATCTGCAAGGAAACCACACAGACTTACCGA  237
CL1150.Contig4_All   169  GGGCCGGAGTATTGGGACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTACCGA  228

Query_2              256  GGGAGCCTGAACATCCTGCGGGGCTACTACAACCAGAGCGAGGCCGGGTCTCACACCATC  315
CL1150.Contig16_All  238  GGGAGCCTGAACATCCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCATC  297
CL1150.Contig4_All   229  CGGAGCCTGAACAACCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCTTC  288

Query_2              316  CAGCGCATGTACGGCTGTGACGTGGGGCCCGACGGGCGCCTCCTCCGCGGGTACAGTCAG  375
CL1150.Contig16_All  298  CAGAACTTGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAG  357
CL1150.Contig4_All   289  CAGAACATGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAG  348

Query_2              376  GACGCCTACGACGGCGCGGATTACCTCACCCTGAACGAGGACCTGCGCTCCTGGACCGCG  435
CL1150.Contig16_All  358  TTCGCCTACGACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCG  417
CL1150.Contig4_All   349  TTCGCCTACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCG  408

Query_2              436  GCGGATGCGTCGGCGCAGATCACCCAGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCTT  495
CL1150.Contig16_All  418  GCGGACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGC  477
CL1150.Contig4_All   409  GCGGACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGC  468
CL1150.Contig15_All  1                  CAGATCTCCCGGCGCAAGTGAGAGGACGCGGGTGTGGCAGAGCTT  45
CL1150.Contig14_All  1                  CAGATCTCCCGGCGCAAGTGAGAGGACGCGGGTGTGGCAGAGCTT  45

Query_2              496  GAGAGGGACTACCTGGAGATTACTTGCGTGAAGTGGCTCCACAGGTATCTGGAGAACGGG  555
CL1150.Contig16_All  478  TACAGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGG  537
CL1150.Contig4_All   469  TACAGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGG  528
CL1150.Contig15_All  46   GAGAGGGACTACCTGGAGATTACTTGCGTGAAGTGGCTCCACAGGTATCTGGAGAACGGG  105
CL1150.Contig14_All  46   GAGAGGGACTACCTGGAGATTACTTGCGTGAAGTGGCTCCACAGGTATCTGGAGAACGGG  105

Query_2              556  AAGGAGACGCTACTGCGCACAGGACAACGGAGGAATCCAGTCTCTCCCAGGAAGCAGAGA  615
CL1150.Contig16_All  538  AAGGAGTCGCTGCTGCGCGCAG                                      559
CL1150.Contig16_All  646                                                      GAGA  649
CL1150.Contig4_All   529  AAGGAGTCGCTGCTGCGCGCAG                                      550
CL1150.Contig15_All  125                                                   CAGAGA  130
```

```
CL1150.Contig15_All  106  AAGGAGACGCTACTGCGCACAG                                          127
CL1150.Contig7_All   91                                                    GCAGAGA  97
CL1150.Contig18_All  91                                                      GAGA  94
CL1150.Contig17_All  94                                                      GAGA  97
CL1150.Contig13_All  94                                                      GAGA  97
CL1150.Contig8_All   91                                                    GCAGAGA  97
CL1150.Contig14_All  125                                                    CAGAGA  130
CL1150.Contig14_All  106  AAGGAGACGCTACTGCGCACAG                                    127
Unigene117550_All    88                                                    GCAGAGA  94
Unigene12618_All     88                                                    GCAGAGA  94

Query_2              616  TCACCCTGACCTGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAG  674
CL1150.Contig16_All  650  TCACCCTGACCTGGCAGCGAGATGGA-GAGGACCTAACCCAGGACACAGAGCTCGTGGAG  708
CL1150.Contig15_All  131  TCACCCTGACCTGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAG  189
CL1150.Contig7_All   98   TCACCCTGACCTGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAG  156
CL1150.Contig18_All  95   TCACCCTGACCTGGCAGAGGGATGGA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAG  153
CL1150.Contig17_All  98   TCACCCTGACCTGGCAGCGAGATGGA-GAGGACCTAACCCAGGACACAGAGCTCGTGGAG  156
CL1150.Contig13_All  98   TCACCCTGACCTGGCACCATGA-GGAGGAGGACCTGACCCAGGACACAGAACTTGTAGGG  156
CL1150.Contig8_All   98   TCACCCTGACCTGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAG  156
CL1150.Contig14_All  131  TCACCCTGACCTGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAG  189
Unigene117550_All    95   TCACCCTGACCTGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAG  153
Unigene12618_All     95   TCACCCTGACCTGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAG  153

Query_2              675  ACCAGGCCTGCAGGAGATGGAACCTTCCAGAAGTGGGTGGCCATGGTGGTGCCTTCTGGA  734
CL1150.Contig16_All  709  ACCAGGCCTGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGA  768
CL1150.Contig15_All  190  ACCAGGCCTGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGA  249
CL1150.Contig7_All   157  ACCAGGCCTGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGA  216
CL1150.Contig18_All  154  ACCAGGCCTGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGA  213
CL1150.Contig17_All  157  ACCAGGCCTGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGA  216
CL1150.Contig13_All  157  ACCAGGCCTACAGGGAATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGA  216
CL1150.Contig8_All   157  ACCAGGCCTGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGA  216
CL1150.Contig14_All  190  ACCAGGCCTGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGA  249
Unigene117550_All    154  ACCAGGCCTGCAGGAGAT                                          171
Unigene12618_All     154  ACCAGGCCTGCAGGAGAT                                          171
Unigene26995_All     1                                                      TCTGGA  6

Query_2              735  CAGGAGCAGAGATACACGTGCCATGTGCAGCATGAGGGACTGTCTGAGCCCATCACCCAG  794
CL1150.Contig16_All  769  GAGGAGCAGAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTG  828
CL1150.Contig15_All  250  CAGGAGCAGAGATACACATGCTATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGG  309
CL1150.Contig7_All   217  GAGGAGCAGAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTG  276
CL1150.Contig18_All  214  GAGGAGCAGAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTG  273
CL1150.Contig17_All  217  GAGGAGCAGAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTG  276
CL1150.Contig13_All  217  GAGGAGCAGAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTG  276
CL1150.Contig8_All   217  GAGGAGCAGAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC     273
CL1150.Contig14_All  250  GAGGAGCAGAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC     306
CL1150.Contig5_All   1                                    GAGGGACTGTCTGAGCCCATTACCTTG  27
Unigene26995_All     7    CAGGAGCAGAGATACACGTGCCATGTGCAGCATGAGGGACTGTCTGAGCCCATCACCCAG  66

Query_2              795  AGATGGGAGCCGCCACATCCTACCATCCCCATCATGTGGATCATTGCTGGTCTGATTCTC  854
CL1150.Contig16_All  829  AGTTGGAAGCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTC  888
CL1150.Contig15_All  310  AGATGGGAGCCACCTCG-C--ACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTC  366
CL1150.Contig7_All   277  AGTTGGAAGCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTC  336
CL1150.Contig18_All  274  AGTTGGAAGCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTC  333
CL1150.Contig17_All  277  AGTTGGAAGCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTC  336
CL1150.Contig13_All  277  AGTTGGAAGCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTC  336
CL1150.Contig8_All   403        GAGCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGTCTGATTCTC  456
CL1150.Contig14_All  436        GAGCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGTCTGATTCTC  489
CL1150.Contig5_All   28   AGATGGGAGCCTCCTCTTCCCATCGTCCTCATCACATGGATCATTGCTGTTCTGGCTCTC  87
Unigene26995_All     67   AGATGGG                                                     73
CL1793.Contig7_All   1                                        TGGATCATTGCTGTTCTGGCTCTC  24
CL1793.Contig5_All   1                                        TGGATCATTGCTGTTCTGGCTCTC  24
CL1793.Contig3_All   1                                        TGGATCATTGCTGTTCTGGCTCTC  24
CL1793.Contig2_All   1                                        TGGATCATTGCTGTTCTGGCTCTC  24

Query_2              855  CCCGTGGTCTTTGCAGTGATTAGAGCTGTGATCTGGAGGAAGAAGCGCTCAGATGATGAC  914
CL1150.Contig16_All  889  CTGGCAGTCACTGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAG          940
CL1150.Contig15_All  367  CTGGTGATCATTGCAGTGATTGGAGTTGCAATCTGGTGGAAGAAGCGCTCAG          418
CL1150.Contig7_All   337  CTGGCAGTCACTGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAG          388
CL1150.Contig18_All  334  CTGGCAGTCACTGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAG          385
CL1150.Contig17_All  337  CTGGCAGTCACTGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAG          388
CL1150.Contig13_All  337  CTGGCAGTCACTGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAG          388
CL1150.Contig8_All   457  CCCGTGGTCTTTGCAGTGATTAGAGCTGTGATCTGGAGGAAGAAGCGCTCA           507
CL1150.Contig14_All  490  CCCGTGGTCTTTGCAGTGATTAGAGCTGTGATCTGGAGGAAGAAGCGCTCA           540
CL1150.Contig5_All   88   CTTGTGGTCGCTGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGATGATGAC  147
CL1793.Contig7_All   25   CTTGTGGTCGCTGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAG          76
CL1793.Contig5_All   25   CTTGTGGTCGCTGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAG          76
CL1793.Contig3_All   25   CTTGTGGTCGCTGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAG          76
CL1793.Contig2_All   25   CTTGTGGTCGCTGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAG          76

Query_2              915  AGTGCCCAGGGCTCT  929
CL1150.Contig5_All   148  AGTGCCCAGGGCTCT  162
```

```
Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 74066383350


Query= JQ425438.1 Meles meles MHC class I antigen (Meme-MHCI) pseudogene
mRNA, Meme-MHCI*PS01 allele, partial sequence

Length=900
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  1131    0.0
CL1150.Contig4_All   486 1187 MHC class I antigen, partial [Meles ...  802     0.0
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  544     1e-153
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  540     2e-152
CL1150.Contig18_All  1278 1694 minus strand MHC class I antigen [...  521     6e-147
CL1150.Contig13_All  621 1046 minus strand MHC class I antigen, p...  484     7e-136
CL1150.Contig7_All   20 418 minus strand MHC class I antigen [Ailu...  473     2e-132
CL1150.Contig8_All   20 526 hypothetical protein PANDA_022308 [Ail...  377     1e-103
CL1150.Contig14_All  3 542 minus strand hypothetical protein PAND...  272     6e-72
CL1150.Contig24_All  1312 1467 MHC class I antigen, partial [Mele...  220     2e-56
CL1150.Contig23_All  1371 1526 MHC class I antigen, partial [Mele...  220     2e-56
CL1150.Contig3_All   1659 1814 MHC class I antigen, partial [Meles...  220     2e-56
CL1150.Contig1_All   1600 1755 MHC class I antigen, partial [Meles...  220     2e-56
CL1150.Contig2_All   970 1128 MHC class I antigen, partial [Meles ...  217     3e-55
CL1150.Contig22_All  1 360 minus strand PREDICTED: LOW QUALITY PR...  198     1e-49



Query_3              29   GCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGGCCGG  88
CL1150.Contig16_All  125  GCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGGCCGG  184
CL1150.Contig4_All   116  GCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGGCCGG  175
CL1150.Contig22_All  1                                                  GAGGGGCCGG  10

Query_3              89   AGTATTGGGACCGGCAGACGCGGAACCTCAAGGACGCCGCACACGCTTTCCGAGTGAACC  148
CL1150.Contig16_All  185  AGTATTGGGACCGGCAGACGCAGATCTGCAAGGAAACCACACAGACTTACCGAGGGAGCC  244
CL1150.Contig4_All   176  AGTATTGGGACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTACCGACGGAGCC  235
CL1150.Contig22_All  11   AGTATTGGGACCGGCAGACGCGGAACCTCAAGGACGCCGCACACGCTTTCCGAGTGAACC  70

Query_3              149  TGAACACCCTGCGGGACTACTATAACCAGAGCGCGGCCGGGTCTCACACCATCCAGCGCA  208
CL1150.Contig16_All  245  TGAACATCCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCATCCAGAACT  304
CL1150.Contig4_All   236  TGAACAACCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCTTCCAGAACA  295
CL1150.Contig22_All  71   TGAACACCCTGCGGGACTACTACAACCAGAGCGCGGCCGG                      110

Query_3              209  TGTACGGCTGTGACGTGGGGCCCGACGGCCGCCTCCTCCGCGGGTACAGTCAGGTGGCCT  268
CL1150.Contig16_All  305  TGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTCGCCT  364
CL1150.Contig4_All   296  TGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTCGCCT  355

Query_3              269  ACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACGGTGGCGGACG  328
CL1150.Contig16_All  365  ACGACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCGGACA  424
CL1150.Contig4_All   356  ACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCGGACA  415

Query_3              329  CCACAGCGCAGATCTCCCGGCGCAAGTGGGAGGCGGCGGATGAGGCGGAGCATGAGAGGA  388
CL1150.Contig16_All  425  CGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTACAGGA  484
CL1150.Contig4_All   416  CGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTACAGGA  475

Query_3              389  ACTACCTGGAGGTGACATGCCTGGAGTGGCTCCACAGGTACCTGGAGAACGGGAAGGAGT  448
CL1150.Contig16_All  485  ACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAGGAGT  544
CL1150.Contig4_All   476  ACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAGGAGT  535

Query_3              449  CGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGACCGTG  508
CL1150.Contig16_All  545  CGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGACCGTG  604
CL1150.Contig4_All   536  CGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGACCGTG  595
CL1150.Contig17_All  4               GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGACCATG  52
CL1150.Contig18_All  2                CAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGACCATG  49
CL1150.Contig13_All  4               GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGACCATG  52
CL1150.Contig7_All   13                       CCCAACACACACGTGACCCACCACCCCATCTCTGACCATG  52
CL1150.Contig8_All   13                       CCCAACACACACGTGACCCACCACCCCATCTCTGACCATG  52
CL1150.Contig24_All  2                CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAACAATG  49
CL1150.Contig23_All  2                CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAACAATG  49
CL1150.Contig3_All   2                CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAACAATG  49
CL1150.Contig1_All   2                CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAACAATG  49
CL1150.Contig2_All   4               GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGACCATG  52
```

```
Query_3              509  ATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGC  568
CL1150.Contig16_All  605  ATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGC  664
CL1150.Contig4_All   596  ATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGA  655
CL1150.Contig15_All  127                                          GAGATCACCCTGACCTGGC  145
CL1150.Contig17_All  53   CTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGC  112
CL1150.Contig18_All  50   CTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGC  109
CL1150.Contig13_All  53   CTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGC  112
CL1150.Contig7_All   53   CTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACCTGGC  112
CL1150.Contig8_All   53   CTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACCTGGC  112
CL1150.Contig14_All  127                                          GAGATCACCCTGACCTGGC  145
CL1150.Contig24_All  50   ATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGA  109
CL1150.Contig23_All  50   ATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGA  109
CL1150.Contig3_All   50   ATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGA  109
CL1150.Contig1_All   50   ATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGA  109
CL1150.Contig2_All   53   CTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACCTGGA  112

Query_3              569  AGCGAGATGGA-GAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCAGGA  627
CL1150.Contig16_All  665  AGCGAGATGGA-GAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCAGGA  723
CL1150.Contig4_All   656  AGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC              702
CL1150.Contig15_All  146  AGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGA  204
CL1150.Contig17_All  113  AGCGAGATGGA-GAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCAGGA  171
CL1150.Contig18_All  110  AGAGGGATGGA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCAGGA  168
CL1150.Contig13_All  113  ACCATGA-GGAGGAGGACCTGACCCAGGACACAGAACTTGTAGGGACCAGGCCTACAGGG  171
CL1150.Contig7_All   113  AGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGA  171
CL1150.Contig8_All   113  AGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGA  171
CL1150.Contig14_All  146  AGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCAGGA  204
CL1150.Contig24_All  110  AGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC              156
CL1150.Contig23_All  110  AGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC              156
CL1150.Contig3_All   110  AGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC              156
CL1150.Contig1_All   110  AGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC              156
CL1150.Contig2_All   113  AGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC              159

Query_3              628  GATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAGATAC  687
CL1150.Contig16_All  724  GATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATAC  783
CL1150.Contig15_All  205  GATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAGATAC  264
CL1150.Contig17_All  172  GATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATAC  231
CL1150.Contig18_All  169  GATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATAC  228
CL1150.Contig13_All  172  AATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATAC  231
CL1150.Contig7_All   172  GATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATAC  231
CL1150.Contig8_All   172  GATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATAC  231
CL1150.Contig14_All  205  GATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGATAC  264

Query_3              688  ACATGCCATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCCACCT  747
CL1150.Contig16_All  784  ACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCT  843
CL1150.Contig15_All  265  ACATGCTATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCCACCT  324
CL1150.Contig17_All  232  ACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCT  291
CL1150.Contig18_All  229  ACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCT  288
CL1150.Contig13_All  232  ACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCT  291
CL1150.Contig7_All   232  ACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCACCT  291
CL1150.Contig8_All   232  ACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC                   273
CL1150.Contig14_All  265  ACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC                   306

Query_3              748  CG---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGGTCATTGCA  804
CL1150.Contig16_All  844  CCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTG  903
CL1150.Contig15_All  325  CG---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGATCATTGCA  381
CL1150.Contig17_All  292  CCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTG  351
CL1150.Contig18_All  289  CCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTG  348
CL1150.Contig13_All  292  CCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTG  351
CL1150.Contig7_All   292  CCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACTGTG  351

Query_3              805  GTGATTGGAGTTGCGATCTGGTGGAAGAAGCACTCAGGAGAGAAAAGGACCAGGCTACTCT  864
CL1150.Contig16_All  904  GTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAAGGACCAGGCTACTCT  963
CL1150.Contig15_All  382  GTGATTGGAGTTGCAATCTGGTGGAAGAAGCGCTCAGGAGAAAAAAGGACCAGGCTACTCT  441
CL1150.Contig17_All  352  GTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAAGGACCAGGCTACTCT  411
CL1150.Contig18_All  349  GTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAAGGACCAGGCTACTCT  408
CL1150.Contig13_All  352  GTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAAGGACCAGGCTACTCT  411
CL1150.Contig7_All   352  GTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAAGGA              399

Query_3              865  CATGCTGCACGCGAT  879
CL1150.Contig16_All  964  CATGCTGCA        972
CL1150.Contig15_All  442  CATGCTGCA        450
CL1150.Contig17_All  412  CATGCTGCA        420
CL1150.Contig18_All  409  CATGCTGCA        417
CL1150.Contig13_All  412  CATGCTGCACGCGAT  426


Lambda      K        H
    1.33    0.621     1.12

Gapped
```

```
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 71450352150


Query= JQ425432.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA,
Meme-MHCI*06 allele, partial cds

Length=975
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  1365    0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  1031    0.0
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  544     1e-153
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  529     4e-149
CL1150.Contig18_All  1278 1694 minus strand MHC class I antigen [...  521     6e-147
CL1150.Contig13_All  621 1046 minus strand MHC class I antigen, p...  484     8e-136
CL1150.Contig7_All  20 418 minus strand MHC class I antigen [Ailu...  473     2e-132
CL1150.Contig8_All  20 526 hypothetical protein PANDA_022308 [Ail...  383     3e-105
CL1150.Contig14_All  3 542 minus strand hypothetical protein PAND...  278     1e-73
CL1150.Contig24_All  1312 1467 MHC class I antigen, partial [Mele...  226     5e-58
CL1150.Contig23_All  1371 1526 MHC class I antigen, partial [Mele...  226     5e-58
CL1150.Contig3_All  1659 1814 MHC class I antigen, partial [Meles...  226     5e-58
CL1150.Contig1_All  1600 1755 MHC class I antigen, partial [Meles...  226     5e-58
CL1150.Contig2_All  970 1128 MHC class I antigen, partial [Meles ...  222     7e-57
CL1150.Contig22_All  1 360 minus strand PREDICTED: LOW QUALITY PR...  204     3e-51


Query_4              1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  60
CL1150.Contig16_All  1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  60
CL1150.Contig4_All   4                GGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  51

Query_4              61   CGCGGGGAACCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTT  120
CL1150.Contig16_All  61   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  120
CL1150.Contig4_All   52   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  111

Query_4              121  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGGTGGAGCAGGAGGGG  180
CL1150.Contig16_All  121  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  180
CL1150.Contig4_All   112  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  171
CL1150.Contig22_All  1                                                  GAGGGG  6

Query_4              181  CCGGAGTATTGGGACCGGCAGACGCGGAACCTCAAGGACGCCGCACACGCTTTCCGAGTG  240
CL1150.Contig16_All  181  CCGGAGTATTGGGACCGGCAGACGCAGATCTGCAAGGAAACCACACAGACTTACCGAGGG  240
CL1150.Contig4_All   172  CCGGAGTATTGGGACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTACCGACGG  231
CL1150.Contig22_All  7    CCGGAGTATTGGGACCGGCAGACGCGGAACCTCAAGGACGCCGCACACGCTTTCCGAGTG  66

Query_4              241  AACCTGAACACCCTGCGGGACTACTACAACCAGAGCGCGGCCGGGTCTCACACCATCCAG  300
CL1150.Contig16_All  241  AGCCTGAACATCCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCATCCAG  300
CL1150.Contig4_All   232  AGCCTGAACAACCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCTTCCAG  291
CL1150.Contig22_All  67   AACCTGAACACCCTGCGGGACTACTACAACCAGAGCGCGGCCGG                  110

Query_4              301  CGCATGTACGGCTGTGATATGGGGCCCGATGGGCGCCTCCTCCGCGGGTACAGTCAGGTG  360
CL1150.Contig16_All  301  AACTTGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTC  360
CL1150.Contig4_All   292  AACATGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTC  351

Query_4              361  GCCTACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACGGTGGCG  420
CL1150.Contig16_All  361  GCCTACGACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCG  420
CL1150.Contig4_All   352  GCCTACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCG  411

Query_4              421  GACGCCACAGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGATGAGGCGGAGCATGAG  480
CL1150.Contig16_All  421  GACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTAC  480
CL1150.Contig4_All   412  GACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTAC  471

Query_4              481  AGGAACTACCTGGAGGTCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  540
CL1150.Contig16_All  481  AGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  540
CL1150.Contig4_All   472  AGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  531

Query_4              541  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  600
CL1150.Contig16_All  541  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  600
CL1150.Contig4_All   532  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  591
CL1150.Contig17_All  4                  GCAGAACCGCCCCAACACACGCATGACCCACCACCCCTATCTCTGAC  48
CL1150.Contig18_All  2                   CAGAACCGCCCCAACACACGCATGACCCACCACCCCTATCTCTGAC  45
CL1150.Contig13_All  4                  GCAGAACCGCCCCAACACACGCATGACCCACCACCCCTATCTCTGAC  48
CL1150.Contig7_All   13                          CCCAACACACACGTGACCCACCACCCCATCTCTGAC  48
CL1150.Contig8_All   13                          CCCAACACACACGTGACCCACCACCCCATCTCTGAC  48
CL1150.Contig24_All  2                    CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig23_All  2                    CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig3_All   2                    CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
```

```
CL1150.Contig1_All   2                    CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig2_All   4                  GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGAC  48

Query_4              601  CGTGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  660
CL1150.Contig16_All  601  CGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  660
CL1150.Contig4_All   592  CGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  651
CL1150.Contig15_All  127                                              GAGATCACCCTGACC  141
CL1150.Contig17_All  49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108
CL1150.Contig18_All  46   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig13_All  49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108
CL1150.Contig7_All   49   CATGCTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACC  108
CL1150.Contig8_All   49   CATGCTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACC  108
CL1150.Contig14_All  127                                              GAGATCACCCTGACC  141
CL1150.Contig24_All  46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig23_All  46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig3_All   46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig1_All   46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig2_All   49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108

Query_4              661  TGGCAGCGAGATGGA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGC  719
CL1150.Contig16_All  661  TGGCAGCGAGATGGA-GAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGC  719
CL1150.Contig4_All   652  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          702
CL1150.Contig15_All  142  TGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGC  200
CL1150.Contig17_All  109  TGGCAGCGAGATGGA-GAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGC  167
CL1150.Contig18_All  106  TGGCAGAGGGATGGA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGC  164
CL1150.Contig13_All  109  TGGCACCATGA-GGAGGAGGACCTGACCCAGGACACAGAACTTGTAGGGACCAGGCCTAC  167
CL1150.Contig7_All   109  TGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGC  167
CL1150.Contig8_All   109  TGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGC  167
CL1150.Contig14_All  142  TGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGC  200
CL1150.Contig24_All  106  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          156
CL1150.Contig23_All  106  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          156
CL1150.Contig3_All   106  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          156
CL1150.Contig1_All   106  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          156
CL1150.Contig2_All   109  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          159

Query_4              720  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAG  779
CL1150.Contig16_All  720  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  779
CL1150.Contig15_All  201  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAG  260
CL1150.Contig17_All  168  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  227
CL1150.Contig18_All  165  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  224
CL1150.Contig13_All  168  AGGGAATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  227
CL1150.Contig7_All   168  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  227
CL1150.Contig8_All   168  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  227
CL1150.Contig14_All  201  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  260

Query_4              780  ATACACATGCCATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCC  839
CL1150.Contig16_All  780  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCC  839
CL1150.Contig15_All  261  ATACACATGCTATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCC  320
CL1150.Contig17_All  228  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCC  287
CL1150.Contig18_All  225  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCC  284
CL1150.Contig13_All  228  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCC  287
CL1150.Contig7_All   228  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCC  287
CL1150.Contig8_All   228  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC               273
CL1150.Contig14_All  261  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC               306

Query_4              840  ACCTCG---CACCATCCCCATCACATGGATCATTGCTGGTGTGGTTCTCCTGGTGGTCAT  896
CL1150.Contig16_All  840  ACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCAC  899
CL1150.Contig15_All  321  ACCTCG---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGATCAT  377
CL1150.Contig17_All  288  ACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCAC  347
CL1150.Contig18_All  285  ACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCAC  344
CL1150.Contig13_All  288  ACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCAC  347
CL1150.Contig7_All   288  ACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCAC  347

Query_4              897  TGCAGTGACTGGAGTTGCGATCTGGTGGAAGAAGCGCTCAGGAGAGAAAGGACCAGGCTA  956
CL1150.Contig16_All  900  TGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTA  959
CL1150.Contig15_All  378  TGCAGTGATTGGAGTTGCAATCTGGTGGAAGAAGCGCTCAGGAGAAAAAGGACCAGGCTA  437
CL1150.Contig17_All  348  TGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTA  407
CL1150.Contig18_All  345  TGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTA  404
CL1150.Contig13_All  348  TGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGACCAGGCTA  407
CL1150.Contig7_All   348  TGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGA          399

Query_4              957  CTCTCATGCTGCACGCGAT  975
CL1150.Contig16_All  960  CTCTCATGCTGCA        972
CL1150.Contig15_All  438  CTCTCATGCTGCA        450
CL1150.Contig17_All  408  CTCTCATGCTGCA        420
CL1150.Contig18_All  405  CTCTCATGCTGCA        417
CL1150.Contig13_All  408  CTCTCATGCTGCACGCGAT  426


Lambda      K        H
    1.33    0.621     1.12
```

```
Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 77581675275


Query= JQ425433.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA,
Meme-MHCI*07 allele, partial cds

Length=975
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  1493    0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  1242    0.0
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  566     3e-160
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  523     2e-147
CL1150.Contig18_All  1278 1694 minus strand MHC class I antigen [...  516     3e-145
CL1150.Contig13_All  621 1046 minus strand MHC class I antigen, p...  484     8e-136
CL1150.Contig7_All  20 418 minus strand MHC class I antigen [Ailu...  462     4e-129
CL1150.Contig8_All  20 526 hypothetical protein PANDA_022308 [Ail...  366     3e-100
CL1150.Contig14_All  3 542 minus strand hypothetical protein PAND...  261     1e-68
CL1150.Contig2_All  970 1128 MHC class I antigen, partial [Meles ...  239     7e-62
CL1150.Contig24_All  1312 1467 MHC class I antigen, partial [Mele...  231     1e-59
CL1150.Contig23_All  1371 1526 MHC class I antigen, partial [Mele...  231     1e-59
CL1150.Contig3_All  1659 1814 MHC class I antigen, partial [Meles...  231     1e-59
CL1150.Contig1_All  1600 1755 MHC class I antigen, partial [Meles...  231     1e-59


Query_5              1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  60
CL1150.Contig16_All  1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  60
CL1150.Contig4_All   4                GGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  51

Query_5              61   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  120
CL1150.Contig16_All  61   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  120
CL1150.Contig4_All   52   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  111

Query_5              121  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  180
CL1150.Contig16_All  121  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  180
CL1150.Contig4_All   112  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  171

Query_5              181  CCGGAGTATTGGGACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTACCGACGG  240
CL1150.Contig16_All  181  CCGGAGTATTGGGACCGGCAGACGCAGATCTGCAAGGAAACCACACAGACTTACCGAGGG  240
CL1150.Contig4_All   172  CCGGAGTATTGGGACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTACCGACGG  231

Query_5              241  AGCCTGAACAACCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCTTCCAG  300
CL1150.Contig16_All  241  AGCCTGAACATCCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCATCCAG  300
CL1150.Contig4_All   232  AGCCTGAACAACCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCTTCCAG  291

Query_5              301  AACATGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACAGTCAGCAC  360
CL1150.Contig16_All  301  AACTTGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTC  360
CL1150.Contig4_All   292  AACATGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTC  351

Query_5              361  TCCTACGACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCG  420
CL1150.Contig16_All  361  GCCTACGACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCG  420
CL1150.Contig4_All   352  GCCTACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCG  411

Query_5              421  GACACGGCGGCGCAGATCACCCAGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTGG  480
CL1150.Contig16_All  421  GACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTAC  480
CL1150.Contig4_All   412  GACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTAC  471

Query_5              481  AGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  540
CL1150.Contig16_All  481  AGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  540
CL1150.Contig4_All   472  AGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  531

Query_5              541  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  600
CL1150.Contig16_All  541  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  600
CL1150.Contig4_All   532  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  591
CL1150.Contig17_All  4                  GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGAC  48
CL1150.Contig18_All  2                   CAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGAC  45
CL1150.Contig13_All  4                  GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGAC  48
CL1150.Contig7_All   13                           CCCAACACACACGTGACCCACCACCCCATCTCTGAC  48
CL1150.Contig8_All   13                           CCCAACACACACGTGACCCACCACCCCATCTCTGAC  48
CL1150.Contig2_All   4                  GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGAC  48
CL1150.Contig24_All  2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig23_All  2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig3_All   2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig1_All   2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
```

```
Query_5              601  CGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  660
CL1150.Contig16_All  601  CGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  660
CL1150.Contig4_All   592  CGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  651
CL1150.Contig15_All  127                                              GAGATCACCCTGACC  141
CL1150.Contig17_All  49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108
CL1150.Contig18_All  46   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig13_All  49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108
CL1150.Contig7_All   49   CATGCTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACC  108
CL1150.Contig8_All   49   CATGCTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACC  108
CL1150.Contig14_All  127                                              GAGATCACCCTGACC  141
CL1150.Contig2_All   49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108
CL1150.Contig24_All  46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig23_All  46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig3_All   46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig1_All   46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105

Query_5              661  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCA  720
CL1150.Contig16_All  661  TGGCAGCGAGATGGAGAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCA  720
CL1150.Contig4_All   652  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           702
CL1150.Contig15_All  142  TGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCA  201
CL1150.Contig17_All  109  TGGCAGCGAGATGGAGAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCA  168
CL1150.Contig18_All  106  TGGCAGAGGGATGGAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCA  165
CL1150.Contig13_All  109  TGGCACCATGAGGAGGAGGACCTGACCCAGGACACAGAACTTGTAGGGACCAGGCCTACA  168
CL1150.Contig7_All   109  TGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCA  168
CL1150.Contig8_All   109  TGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCA  168
CL1150.Contig14_All  142  TGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCA  201
CL1150.Contig2_All   109  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           159
CL1150.Contig24_All  106  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           156
CL1150.Contig23_All  106  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           156
CL1150.Contig3_All   106  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           156
CL1150.Contig1_All   106  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           156

Query_5              721  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAGA  780
CL1150.Contig16_All  721  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  780
CL1150.Contig15_All  202  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAGA  261
CL1150.Contig17_All  169  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  228
CL1150.Contig18_All  166  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  225
CL1150.Contig13_All  169  GGGAATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  228
CL1150.Contig7_All   169  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  228
CL1150.Contig8_All   169  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  228
CL1150.Contig14_All  202  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  261

Query_5              781  TACACATGCTATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCCA  840
CL1150.Contig16_All  781  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCA  840
CL1150.Contig15_All  262  TACACATGCTATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCCA  321
CL1150.Contig17_All  229  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCA  288
CL1150.Contig18_All  226  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCA  285
CL1150.Contig13_All  229  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCA  288
CL1150.Contig7_All   229  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCA  288
CL1150.Contig8_All   229  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC                273
CL1150.Contig14_All  262  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC                306

Query_5              841  CCTCG---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGATCATT  897
CL1150.Contig16_All  841  CCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACT  900
CL1150.Contig15_All  322  CCTCG---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGATCATT  378
CL1150.Contig17_All  289  CCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACT  348
CL1150.Contig18_All  286  CCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACT  345
CL1150.Contig13_All  289  CCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACT  348
CL1150.Contig7_All   289  CCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACT  348

Query_5              898  GCAGTGATTGGAGTTGCAATCTGGTGGAAGAAGCGCTCAGGAGAAAAAGGACCAGGCTAC  957
CL1150.Contig16_All  901  GTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTAC  960
CL1150.Contig15_All  379  GCAGTGATTGGAGTTGCAATCTGGTGGAAGAAGCGCTCAGGAGAAAAAGGACCAGGCTAC  438
CL1150.Contig17_All  349  GTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTAC  408
CL1150.Contig18_All  346  GTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTAC  405
CL1150.Contig13_All  349  GTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGACCAGGCTAC  408
CL1150.Contig7_All   349  GTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGA           399

Query_5              958  TCTCATGCTGCACGCGAT  975
CL1150.Contig16_All  961  TCTCATGCTGCA        972
CL1150.Contig15_All  439  TCTCATGCTGCA        450
CL1150.Contig17_All  409  TCTCATGCTGCA        420
CL1150.Contig18_All  406  TCTCATGCTGCA        417
CL1150.Contig13_All  409  TCTCATGCTGCACGCGAT  426


Lambda      K        H
    1.33    0.621     1.12

Gapped
```

```
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 77581675275


Query= JQ425431.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA,
Meme-MHCI*05 allele, partial cds

Length=975
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  1360    0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  1009    0.0
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  544     1e-153
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  540     2e-152
CL1150.Contig18_All  1278 1694 minus strand MHC class I antigen [...  521     6e-147
CL1150.Contig13_All  621 1046 minus strand MHC class I antigen, p...  484     8e-136
CL1150.Contig7_All  20 418 minus strand MHC class I antigen [Ailu...  473     2e-132
CL1150.Contig8_All  20 526 hypothetical protein PANDA_022308 [Ail...  377     1e-103
CL1150.Contig14_All  3 542 minus strand hypothetical protein PAND...  272     7e-72
CL1150.Contig24_All  1312 1467 MHC class I antigen, partial [Mele...  220     3e-56
CL1150.Contig23_All  1371 1526 MHC class I antigen, partial [Mele...  220     3e-56
CL1150.Contig3_All  1659 1814 MHC class I antigen, partial [Meles...  220     3e-56
CL1150.Contig1_All  1600 1755 MHC class I antigen, partial [Meles...  220     3e-56
CL1150.Contig2_All  970 1128 MHC class I antigen, partial [Meles ...  217     3e-55
CL1150.Contig22_All  1 360 minus strand PREDICTED: LOW QUALITY PR...  198     1e-49


Query_6              1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  60
CL1150.Contig16_All  1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  60
CL1150.Contig4_All   4                GGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  51

Query_6              61   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  120
CL1150.Contig16_All  61   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  120
CL1150.Contig4_All   52   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  111

Query_6              121  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  180
CL1150.Contig16_All  121  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  180
CL1150.Contig4_All   112  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  171
CL1150.Contig22_All  1                                                  GAGGGG  6

Query_6              181  CCGGAGTATTGGGACCGGCAGACGCGGAACCTCAAGGACGCCGCACACGCTTTCCGAGTG  240
CL1150.Contig16_All  181  CCGGAGTATTGGGACCGGCAGACGCAGATCTGCAAGGAAACCACACAGACTTACCGAGGG  240
CL1150.Contig4_All   172  CCGGAGTATTGGGACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTACCGACGG  231
CL1150.Contig22_All  7    CCGGAGTATTGGGACCGGCAGACGCGGAACCTCAAGGACGCCGCACACGCTTTCCGAGTG  66

Query_6              241  AACCTGAACACCCTGCGGGACTACTATAACCAGAGCGCGGCCGGGTCTCACACCATCCAG  300
CL1150.Contig16_All  241  AGCCTGAACATCCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCATCCAG  300
CL1150.Contig4_All   232  AGCCTGAACAACCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCTTCCAG  291
CL1150.Contig22_All  67   AACCTGAACACCCTGCGGGACTACTACAACCAGAGCGCGGCCGG                  110

Query_6              301  CGCATGTACGGCTGTGACGTGGGGCCCGACGGCCGCCTCCTCCGCGGGTACAGTCAGGTG  360
CL1150.Contig16_All  301  AACTTGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTC  360
CL1150.Contig4_All   292  AACATGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTC  351

Query_6              361  GCCTACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACGGTGGCG  420
CL1150.Contig16_All  361  GCCTACGACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCG  420
CL1150.Contig4_All   352  GCCTACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCG  411

Query_6              421  GACGCCACAGCGCAGATCTCCCGGCGCAAGTGGGAGGCGGCGGATGAGGCGGAGCATGAG  480
CL1150.Contig16_All  421  GACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTAC  480
CL1150.Contig4_All   412  GACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTAC  471

Query_6              481  AGGAACTACCTGGAGGTGACATGCCTGGAGTGGCTCCACAGGTACCTGGAGAACGGGAAG  540
CL1150.Contig16_All  481  AGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  540
CL1150.Contig4_All   472  AGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  531

Query_6              541  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  600
CL1150.Contig16_All  541  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  600
CL1150.Contig4_All   532  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  591
CL1150.Contig17_All  4                  GCAGAACCGCCCCAACACACGCATGACCCACCACCCTATCTCTGAC  48
CL1150.Contig18_All  2                   CAGAACCGCCCCAACACACGCATGACCCACCACCCTATCTCTGAC  45
CL1150.Contig13_All  4                  GCAGAACCGCCCCAACACACGCATGACCCACCACCCTATCTCTGAC  48
CL1150.Contig7_All   13                          CCCAACACACACGTGACCCACCACCCCATCTCTGAC  48
CL1150.Contig8_All   13                          CCCAACACACACGTGACCCACCACCCCATCTCTGAC  48
CL1150.Contig24_All  2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig23_All  2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig3_All   2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
```

```
CL1150.Contig1_All   2                    CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig2_All   4                  GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGAC  48

Query_6              601  CGTGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  660
CL1150.Contig16_All  601  CGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  660
CL1150.Contig4_All   592  CGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  651
CL1150.Contig15_All  127                                              GAGATCACCCTGACC  141
CL1150.Contig17_All  49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108
CL1150.Contig18_All  46   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig13_All  49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108
CL1150.Contig7_All   49   CATGCTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACC  108
CL1150.Contig8_All   49   CATGCTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACC  108
CL1150.Contig14_All  127                                              GAGATCACCCTGACC  141
CL1150.Contig24_All  46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig23_All  46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig3_All   46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig1_All   46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig2_All   49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108

Query_6              661  TGGCAGCGAGATGGA-GAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGC  719
CL1150.Contig16_All  661  TGGCAGCGAGATGGA-GAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGC  719
CL1150.Contig4_All   652  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          702
CL1150.Contig15_All  142  TGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGC  200
CL1150.Contig17_All  109  TGGCAGCGAGATGGA-GAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGC  167
CL1150.Contig18_All  106  TGGCAGAGGGATGGA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGC  164
CL1150.Contig13_All  109  TGGCACCATGA-GGAGGAGGACCTGACCCAGGACACAGAACTTGTAGGGACCAGGCCTAC  167
CL1150.Contig7_All   109  TGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGC  167
CL1150.Contig8_All   109  TGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGC  167
CL1150.Contig14_All  142  TGGCAGAGGGATGGA-GAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGC  200
CL1150.Contig24_All  106  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          156
CL1150.Contig23_All  106  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          156
CL1150.Contig3_All   106  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          156
CL1150.Contig1_All   106  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          156
CL1150.Contig2_All   109  TGGAAGCGAGATGAA-GAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC          159

Query_6              720  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAG  779
CL1150.Contig16_All  720  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  779
CL1150.Contig15_All  201  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAG  260
CL1150.Contig17_All  168  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  227
CL1150.Contig18_All  165  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  224
CL1150.Contig13_All  168  AGGGAATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  227
CL1150.Contig7_All   168  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  227
CL1150.Contig8_All   168  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  227
CL1150.Contig14_All  201  AGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAG  260

Query_6              780  ATACACATGCCATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCC  839
CL1150.Contig16_All  780  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCC  839
CL1150.Contig15_All  261  ATACACATGCTATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCC  320
CL1150.Contig17_All  228  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCC  287
CL1150.Contig18_All  225  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCC  284
CL1150.Contig13_All  228  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCC  287
CL1150.Contig7_All   228  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCC  287
CL1150.Contig8_All   228  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC               273
CL1150.Contig14_All  261  ATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC               306

Query_6              840  ACCTCG---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGGTCAT  896
CL1150.Contig16_All  840  ACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCAC  899
CL1150.Contig15_All  321  ACCTCG---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGATCAT  377
CL1150.Contig17_All  288  ACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCAC  347
CL1150.Contig18_All  285  ACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCAC  344
CL1150.Contig13_All  288  ACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCAC  347
CL1150.Contig7_All   288  ACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCAC  347

Query_6              897  TGCAGTGATTGGAGTTGCGATCTGGTGGAAGAAGCACTCAGGAGAGAAAGGACCAGGCTA  956
CL1150.Contig16_All  900  TGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTA  959
CL1150.Contig15_All  378  TGCAGTGATTGGAGTTGCAATCTGGTGGAAGAAGCGCTCAGGAGAAAAAGGACCAGGCTA  437
CL1150.Contig17_All  348  TGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTA  407
CL1150.Contig18_All  345  TGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTA  404
CL1150.Contig13_All  348  TGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGACCAGGCTA  407
CL1150.Contig7_All   348  TGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGA          399

Query_6              957  CTCTCATGCTGCACGCGAT  975
CL1150.Contig16_All  960  CTCTCATGCTGCA        972
CL1150.Contig15_All  438  CTCTCATGCTGCA        450
CL1150.Contig17_All  408  CTCTCATGCTGCA        420
CL1150.Contig18_All  405  CTCTCATGCTGCA        417
CL1150.Contig13_All  408  CTCTCATGCTGCACGCGAT  426


Lambda      K        H
    1.33    0.621     1.12
```

```
Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 77581675275


Query= JQ425430.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA,
Meme-MHCI*04 allele, partial cds

Length=975
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  1437    0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  1053    0.0
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  545     4e-154
CL1150.Contig18_All  1278 1694 minus strand MHC class I antigen [...  538     6e-152
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  538     6e-152
CL1150.Contig13_All  621 1046 minus strand MHC class I antigen, p...  507     2e-142
CL1150.Contig7_All  20 418 minus strand MHC class I antigen [Ailu...  484     8e-136
CL1150.Contig8_All  20 526 hypothetical protein PANDA_022308 [Ail...  366     3e-100
CL1150.Contig14_All  3 542 minus strand hypothetical protein PAND...  261     1e-68
CL1150.Contig2_All  970 1128 MHC class I antigen, partial [Meles ...  239     7e-62
CL1150.Contig24_All  1312 1467 MHC class I antigen, partial [Mele...  231     1e-59
CL1150.Contig23_All  1371 1526 MHC class I antigen, partial [Mele...  231     1e-59
CL1150.Contig3_All  1659 1814 MHC class I antigen, partial [Meles...  231     1e-59
CL1150.Contig1_All  1600 1755 MHC class I antigen, partial [Meles...  231     1e-59
CL7631.Contig3_All  2 109 minus strand MHC class I antigen, parti...  137     3e-31
CL1150.Contig22_All  1 360 minus strand PREDICTED: LOW QUALITY PR...  110     6e-23


Query_7              1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTCCACCGCGGTGTCCCGGCCCGGC  60
CL1150.Contig16_All  1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  60
CL1150.Contig4_All   4                GGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  51

Query_7              61   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  120
CL1150.Contig16_All  61   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  120
CL1150.Contig4_All   52   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  111

Query_7              121  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGGTGGAGCAGGAGGGG  180
CL1150.Contig16_All  121  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  180
CL1150.Contig4_All   112  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  171
CL1150.Contig22_All  1                                                    GAGGGG  6

Query_7              181  CCGGAGTATTGGGACCGGCAGACGCAGATCTGCAAGGACGCCGCACAGACTTACCGAGGG  240
CL1150.Contig16_All  181  CCGGAGTATTGGGACCGGCAGACGCAGATCTGCAAGGAAACCACACAGACTTACCGAGGG  240
CL1150.Contig4_All   172  CCGGAGTATTGGGACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTACCGACGG  231
CL1150.Contig22_All  7    CCGGAGTATTGGGACCGGCAGACGCGGAACCTCAAGGACGCCGCACACGCTTTCCGAGTG  66

Query_7              241  AACCTGCAGACCGCACT-C--CG-CTACTACAACCAGAGCGCGGCCGGGTCTCACACCAT  296
CL1150.Contig16_All  241  AGCCTG-A-ACATC-CTGCGGGG-CTACTACAACCAGAGCGCGGCCGGGTCTCACACCAT  296
CL1150.Contig4_All   232  AGCCTG-A-ACAAC-CTGCGGGG-CTACTACAACCAGAGCGCGGCCGGGTCTCACACCTT  287
CL1150.Contig22_All  67   AACCTGAACACC-CTGC-G--GGACTACTACAACCAGAGCGCGGCCGG              110

Query_7              297  CCAGAACGTGTACGGCTGTGATGTGGGGCGCGACGGGCGTCTCCTCCGCGGATACAGTCA  356
CL1150.Contig16_All  297  CCAGAACTTGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCA  356
CL1150.Contig4_All   288  CCAGAACATGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCA  347

Query_7              357  GGACTCCTACGACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGC  416
CL1150.Contig16_All  357  GTTCGCCTACGACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGC  416
CL1150.Contig4_All   348  GTTCGCCTACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACCGC  407

Query_7              417  GGCGGACACGGCGGCGCAGATCACCCAGCGCAAGTGGGAGGACGCGGGTGCGGCAGAGCG  476
CL1150.Contig16_All  417  GGCGGACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCG  476
CL1150.Contig4_All   408  GGCGGACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCG  467

Query_7              477  CTGGAGGAACTACCTGGAGGTCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGG  536
CL1150.Contig16_All  477  CTACAGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGG  536
CL1150.Contig4_All   468  CTACAGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGG  527

Query_7              537  GAAGGAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTC  596
CL1150.Contig16_All  537  GAAGGAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTC  596
CL1150.Contig4_All   528  GAAGGAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTC  587
CL1150.Contig17_All  4                        GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTC  44
CL1150.Contig18_All  2                         CAGAACCGCCCAACACACGCATGACCCACCACCCTATCTC  41
CL1150.Contig13_All  4                        GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTC  44
CL1150.Contig7_All   13                               CCCAACACACACGTGACCCACCACCCCATCTC  44
CL1150.Contig8_All   13                               CCCAACACACACGTGACCCACCACCCCATCTC  44
```

```
CL1150.Contig2_All   4                      GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTC  44
CL1150.Contig24_All  2                       CAGAACCCTCCAACACACACGTGACACACCACCCGATCTC  41
CL1150.Contig23_All  2                       CAGAACCCTCCAACACACACGTGACACACCACCCGATCTC  41
CL1150.Contig3_All   2                       CAGAACCCTCCAACACACACGTGACACACCACCCGATCTC  41
CL1150.Contig1_All   2                       CAGAACCCTCCAACACACACGTGACACACCACCCGATCTC  41

Query_7              597  TGACCGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  656
CL1150.Contig16_All  597  TGACCGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  656
CL1150.Contig4_All   588  TGACCGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  647
CL1150.Contig17_All  45   TGACCATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  104
CL1150.Contig18_All  42   TGACCATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  101
CL1150.Contig15_All  127                                              GAGATCACCCT  137
CL1150.Contig13_All  45   TGACCATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  104
CL1150.Contig7_All   45   TGACCATGCTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCT  104
CL1150.Contig8_All   45   TGACCATGCTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCT  104
CL1150.Contig14_All  127                                              GAGATCACCCT  137
CL1150.Contig2_All   45   TGACCATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  104
CL1150.Contig24_All  42   TAACAATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  101
CL1150.Contig23_All  42   TAACAATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  101
CL1150.Contig3_All   42   TAACAATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  101
CL1150.Contig1_All   42   TAACAATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCT  101

Query_7              657  GACCTGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCC  716
CL1150.Contig16_All  657  GACCTGGCAGCGAGATGGAGAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCC  716
CL1150.Contig4_All   648  GACCTGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC       702
CL1150.Contig17_All  105  GACCTGGCAGCGAGATGGAGAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCC  164
CL1150.Contig18_All  102  GACCTGGCAGAGGGATGGAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCC  161
CL1150.Contig15_All  138  GACCTGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCC  197
CL1150.Contig13_All  105  GACCTGGCACCATGAGGAGGAGGACCTGACCCAGGACACAGAACTTGTAGGGACCAGGCC  164
CL1150.Contig7_All   105  GACCTGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCC  164
CL1150.Contig8_All   105  GACCTGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCC  164
CL1150.Contig14_All  138  GACCTGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCC  197
CL1150.Contig2_All   105  GACCTGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC       159
CL1150.Contig24_All  102  GACCTGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC       156
CL1150.Contig23_All  102  GACCTGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC       156
CL1150.Contig3_All   102  GACCTGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC       156
CL1150.Contig1_All   102  GACCTGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC       156

Query_7              717  TGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCA  776
CL1150.Contig16_All  717  TGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCA  776
CL1150.Contig17_All  165  TGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCA  224
CL1150.Contig18_All  162  TGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCA  221
CL1150.Contig15_All  198  TGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCA  257
CL1150.Contig13_All  165  TACAGGGAATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCA  224
CL1150.Contig7_All   165  TGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCA  224
CL1150.Contig8_All   165  TGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCA  224
CL1150.Contig14_All  198  TGCAGGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCA  257

Query_7              777  GAGATACACATGCTATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGA  836
CL1150.Contig16_All  777  GAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAA  836
CL1150.Contig17_All  225  GAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAA  284
CL1150.Contig18_All  222  GAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAA  281
CL1150.Contig15_All  258  GAGATACACATGCTATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGA  317
CL1150.Contig13_All  225  GAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAA  284
CL1150.Contig7_All   225  GAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAA  284
CL1150.Contig8_All   225  GAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC             273
CL1150.Contig14_All  258  GAGATACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC             306

Query_7              837  GCCACCTCA---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGGT  893
CL1150.Contig16_All  837  GCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGT  896
CL1150.Contig17_All  285  GCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGT  344
CL1150.Contig18_All  282  GCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGT  341
CL1150.Contig15_All  318  GCCACCTCG---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGAT  374
CL1150.Contig13_All  285  GCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGT  344
CL1150.Contig7_All   285  GCCACCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGT  344
CL7631.Contig3_All   4        CCTCC---CACCATCCCCATCACTTGGATGATTGCTGGCCTGGCTCTCCTGGTGGT  56

Query_7              894  CATTGCAGTGATTGGAGCTGTGATCTGGTGGAAGAAGCGCTCAGGAGAAAAAGGACCAGG  953
CL1150.Contig16_All  897  CACTGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGG  956
CL1150.Contig17_All  345  CACTGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGG  404
CL1150.Contig18_All  342  CACTGTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGG  401
CL1150.Contig15_All  375  CATTGCAGTGATTGGAGTTGCAATCTGGTGGAAGAAGCGCTCAGGAGAAAAAGGACCAGG  434
CL1150.Contig13_All  345  CACTGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGACCAGG  404
CL1150.Contig7_All   345  CACTGTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGA       399
CL7631.Contig3_All   57   CACTGTGGTGATTGGAGCTGTGATCTGGTGGAAGAAGCGCTCAGG                 101

Query_7              954  CTACTCTCATGCTGCACGCGAT  975
CL1150.Contig16_All  957  CTACTCTCATGCTGCA        972
CL1150.Contig17_All  405  CTACTCTCATGCTGCA        420
CL1150.Contig18_All  402  CTACTCTCATGCTGCA        417
CL1150.Contig15_All  435  CTACTCTCATGCTGCA        450
```

```
CL1150.Contig13_All  405  CTACTCTCATGCTGCACGCGAT  426


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 77581675275


Query= JQ425429.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA,
Meme-MHCI*03 allele, partial cds

Length=975
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  1531    0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  1214    0.0
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  566     3e-160
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  523     2e-147
CL1150.Contig18_All  1278 1694 minus strand MHC class I antigen [...  516     3e-145
CL1150.Contig13_All  621 1046 minus strand MHC class I antigen, p...  484     8e-136
CL1150.Contig7_All  20 418 minus strand MHC class I antigen [Ailu...  462     4e-129
CL1150.Contig8_All  20 526 hypothetical protein PANDA_022308 [Ail...  366     3e-100
CL1150.Contig14_All  3 542 minus strand hypothetical protein PAND...  261     1e-68
CL1150.Contig2_All  970 1128 MHC class I antigen, partial [Meles ...  239     7e-62
CL1150.Contig24_All  1312 1467 MHC class I antigen, partial [Mele...  231     1e-59
CL1150.Contig23_All  1371 1526 MHC class I antigen, partial [Mele...  231     1e-59
CL1150.Contig3_All  1659 1814 MHC class I antigen, partial [Meles...  231     1e-59
CL1150.Contig1_All  1600 1755 MHC class I antigen, partial [Meles...  231     1e-59



Query_8              1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  60
CL1150.Contig16_All  1    GAGACCTGGGCGGGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  60
CL1150.Contig4_All   4                GGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGC  51

Query_8              61   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACGCGCAGTTCGTGCGGTTC  120
CL1150.Contig16_All  61   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  120
CL1150.Contig4_All   52   CGCGGGGAGCCCCGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTC  111

Query_8              121  GACAGCGACTCTGCCAGTCTGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  180
CL1150.Contig16_All  121  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  180
CL1150.Contig4_All   112  GACAGCGACTCTGCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGG  171

Query_8              181  CCGGAGTATTGGGACCGGGAGACCCGGAACCTCAAGGAAACCACACAGACTTACCGAGTG  240
CL1150.Contig16_All  181  CCGGAGTATTGGGACCGGCAGACGCAGATCTGCAAGGAAACCACACAGACTTACCGAGGG  240
CL1150.Contig4_All   172  CCGGAGTATTGGGACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTACCGACGG  231

Query_8              241  AACCTGAACAACCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCATCCAG  300
CL1150.Contig16_All  241  AGCCTGAACATCCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCATCCAG  300
CL1150.Contig4_All   232  AGCCTGAACAACCTGCGGGGCTACTACAACCAGAGCGCGGCCGGGTCTCACACCTTCCAG  291

Query_8              301  AACTTGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTC  360
CL1150.Contig16_All  301  AACTTGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTC  360
CL1150.Contig4_All   292  AACATGTACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTC  351

Query_8              361  GCCTACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCG  420
CL1150.Contig16_All  361  GCCTACGACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCG  420
CL1150.Contig4_All   352  GCCTACGACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCG  411

Query_8              421  GACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTAC  480
CL1150.Contig16_All  421  GACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTAC  480
CL1150.Contig4_All   412  GACACGGCGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTAC  471

Query_8              481  AGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  540
CL1150.Contig16_All  481  AGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  540
CL1150.Contig4_All   472  AGGAACTATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAG  531

Query_8              541  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  600
CL1150.Contig16_All  541  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  600
CL1150.Contig4_All   532  GAGTCGCTGCTGCGCGCAGAAACCCCCAATACACACGTGACCCGCCACCCCATCTCTGAC  591
CL1150.Contig17_All  4                    GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGAC  48
CL1150.Contig18_All  2                     CAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGAC  45
CL1150.Contig13_All  4                    GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGAC  48
CL1150.Contig7_All   13                           CCCAACACACACGTGACCCACCACCCCATCTCTGAC  48
CL1150.Contig8_All   13                           CCCAACACACACGTGACCCACCACCCCATCTCTGAC  48
```

```
CL1150.Contig2_All   4                  GCAGAACCGCCCAACACACGCATGACCCACCACCCTATCTCTGAC  48
CL1150.Contig24_All  2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig23_All  2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig3_All   2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45
CL1150.Contig1_All   2                   CAGAACCCTCCAACACACACGTGACACACCACCCGATCTCTAAC  45

Query_8              601  CGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  660
CL1150.Contig16_All  601  CGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  660
CL1150.Contig4_All   592  CGTGATGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  651
CL1150.Contig15_All  127                                          GAGATCACCCTGACC  141
CL1150.Contig17_All  49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108
CL1150.Contig18_All  46   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig13_All  49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108
CL1150.Contig7_All   49   CATGCTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACC  108
CL1150.Contig8_All   49   CATGCTAACACCCTGAGGTGATGGGCCCTGGACTTCTACCCTGCAGAGATCACCCTGACC  108
CL1150.Contig14_All  127                                          GAGATCACCCTGACC  141
CL1150.Contig2_All   49   CATGCTGTCACCCTGAGGTGCTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  108
CL1150.Contig24_All  46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig23_All  46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig3_All   46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105
CL1150.Contig1_All   46   AATGATGTCACCCTGAGGTGTTGGGCCCTGGACTTCTACCCTGCGGAGATCACCCTGACC  105

Query_8              661  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCA  720
CL1150.Contig16_All  661  TGGCAGCGAGATGGAGAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCA  720
CL1150.Contig4_All   652  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           702
CL1150.Contig15_All  142  TGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCA  201
CL1150.Contig17_All  109  TGGCAGCGAGATGGAGAGGACCTAACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCA  168
CL1150.Contig18_All  106  TGGCAGAGGGATGGAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACCAGGCCTGCA  165
CL1150.Contig13_All  109  TGGCACCATGAGGAGGAGGACCTGACCCAGGACACAGAACTTGTAGGGACCAGGCCTACA  168
CL1150.Contig7_All   109  TGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCA  168
CL1150.Contig8_All   109  TGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCA  168
CL1150.Contig14_All  142  TGGCAGAGGGATGGAGAGGACCAGACCCAGGACACAGAGCTTGTGGAGACCAGGCCTGCA  201
CL1150.Contig2_All   109  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           159
CL1150.Contig24_All  106  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           156
CL1150.Contig23_All  106  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           156
CL1150.Contig3_All   106  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           156
CL1150.Contig1_All   106  TGGAAGCGAGATGAAGAGGACCTGACCCAGGACACAGAGCTCGTGGAGACC           156

Query_8              721  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAGA  780
CL1150.Contig16_All  721  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  780
CL1150.Contig15_All  202  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTGGTGCCTTCTGGACAGGAGCAGAGA  261
CL1150.Contig17_All  169  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  228
CL1150.Contig18_All  166  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  225
CL1150.Contig13_All  169  GGGAATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  228
CL1150.Contig7_All   169  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  228
CL1150.Contig8_All   169  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  228
CL1150.Contig14_All  202  GGAGATGGAACCTTCCAGAAGTGGGCGGCTGTGGTTGTGCCCTCTGGAGAGGAGCAGAGA  261

Query_8              781  TACACATGCTATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCCA  840
CL1150.Contig16_All  781  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCA  840
CL1150.Contig15_All  262  TACACATGCTATGTGCAGCATGAGGGGCTGTCTGAACCCATCACCCGGAGATGGGAGCCA  321
CL1150.Contig17_All  229  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCA  288
CL1150.Contig18_All  226  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCA  285
CL1150.Contig13_All  229  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCA  288
CL1150.Contig7_All   229  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACCTTGAGTTGGAAGCCA  288
CL1150.Contig8_All   229  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC                273
CL1150.Contig14_All  262  TACACATGCCATGTGCAGCATAAGGGGCTGCCTGAGCCCATCACC                306

Query_8              841  CCTCG---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGATCATT  897
CL1150.Contig16_All  841  CCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACT  900
CL1150.Contig15_All  322  CCTCG---CACCATCCCCATCACATGGATCATTGCTGGTCTGGTTCTCCTGGTGATCATT  378
CL1150.Contig17_All  289  CCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACT  348
CL1150.Contig18_All  286  CCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACT  345
CL1150.Contig13_All  289  CCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACT  348
CL1150.Contig7_All   289  CCTCCTCCCACCATCCCCATCATGTGGATCATTGCTGGCCTGGCTCTCCTGGCAGTCACT  348

Query_8              898  GCAGTGATTGGAGTTGCAATCTGGTGGAAGAAGCGCTCAGGAGAAAAAGGACCAGGCTAC  957
CL1150.Contig16_All  901  GTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTAC  960
CL1150.Contig15_All  379  GCAGTGATTGGAGTTGCAATCTGGTGGAAGAAGCGCTCAGGAGAAAAAGGACCAGGCTAC  438
CL1150.Contig17_All  349  GTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTAC  408
CL1150.Contig18_All  346  GTGGTGGTTGGAGCTGTGATCTGGAGGAGGAAGCGCTCAGGAGGAAAAGGACCAGGCTAC  405
CL1150.Contig13_All  349  GTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGACCAGGCTAC  408
CL1150.Contig7_All   349  GTGGTGGTTGGAGCTGTGATCTGGAGGAAGAAGCGCTCAGGAGGAAAAGGA           399

Query_8              958  TCTCATGCTGCACGCGAT  975
CL1150.Contig16_All  961  TCTCATGCTGCA        972
CL1150.Contig15_All  439  TCTCATGCTGCA        450
CL1150.Contig17_All  409  TCTCATGCTGCA        420
CL1150.Contig18_All  406  TCTCATGCTGCA        417
CL1150.Contig13_All  409  TCTCATGCTGCACGCGAT  426
```

```
Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 77581675275


Query= JQ425428.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA,
Meme-MHCI*02 allele, partial cds

Length=543
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  832     0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  776     0.0
CL1150.Contig22_All  1 360 minus strand PREDICTED: LOW QUALITY PR...  115     7e-25



Query_9              1    GGCTCCCACTCCCTGAGGTATTTCTCCACCGCGGTGTCCCGGCCCGGCCGCGGGGAGCCC  60
CL1150.Contig16_All  13   GGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGCCGCGGGGAGCCC  72
CL1150.Contig4_All   4    GGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGCCGCGGGGAGCCC  63

Query_9              61   CGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACTCT  120
CL1150.Contig16_All  73   CGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACTCT  132
CL1150.Contig4_All   64   CGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACTCT  123

Query_9              121  GCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGGTGGAGCAGGAGGGGCCGGAGTATTGG  180
CL1150.Contig16_All  133  GCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGGCCGGAGTATTGG  192
CL1150.Contig4_All   124  GCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGGCCGGAGTATTGG  183
CL1150.Contig22_All  1                                          GAGGGGCCGGAGTATTGG  18

Query_9              181  GACCGGCAGACGCAGATCTGCAAGGACGCCGCACAGACTTTCCGAGGGAACCTGCAGACC  240
CL1150.Contig16_All  193  GACCGGCAGACGCAGATCTGCAAGGAAACCACACAGACTTACCGAGGGAGCCTG-A-ACA  250
CL1150.Contig4_All   184  GACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTACCGACGGAGCCTG-A-ACA  241
CL1150.Contig22_All  19   GACCGGCAGACGCGGAACCTCAAGGACGCCGCACACGCTTTCCGAGTGAACCTGAACACC  78

Query_9              241  GCACTCC---G-CTACTACAACCAGAGCGCGGCCGGGTCTCACACCATCCAGAACGTGTA  296
CL1150.Contig16_All  251  TC-CTGCGGGG-CTACTACAACCAGAGCGCGGCCGGGTCTCACACCATCCAGAACTTGTA  308
CL1150.Contig4_All   242  AC-CTGCGGGG-CTACTACAACCAGAGCGCGGCCGGGTCTCACACCTTCCAGAACATGTA  299
CL1150.Contig22_All  79   -CTGCGG---GACTACTACAACCAGAGCGCGGCCGG                        110

Query_9              297  CGGCTGTGACGTGGGGCCCGACGGGCGTTTCCTCCGCGGGTACCGTCAGGACTCCTACGA  356
CL1150.Contig16_All  309  CGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTCGCCTACGA  368
CL1150.Contig4_All   300  CGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTCGCCTACGA  359

Query_9              357  CGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCGGACACGGC  416
CL1150.Contig16_All  369  CGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCGGACACGGC  428
CL1150.Contig4_All   360  CGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCGGACACGGC  419

Query_9              417  GGCGCAGATCACCCAGCGCAAGTGGGAGGACGCGGGTGCGGCAGAGCGCTGGAGGAACTA  476
CL1150.Contig16_All  429  GGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTACAGGAACTA  488
CL1150.Contig4_All   420  GGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTACAGGAACTA  479

Query_9              477  CCTGGAGGTCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAGGAGTCGCT  536
CL1150.Contig16_All  489  TGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAGGAGTCGCT  548
CL1150.Contig4_All   480  TGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAGGAGTCGCT  539

Query_9              537  GCTGCGC  543
CL1150.Contig16_All  549  GCTGCGC  555
CL1150.Contig4_All   540  GCTGCGC  546


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 42412998768


Query= JQ425427.1 Meles meles MHC class I antigen (Meme-MHCI) mRNA,
Meme-MHCI*01 allele, partial cds
```

```
Length=543
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  809     0.0
CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  793     0.0
CL1150.Contig22_All  1 360 minus strand PREDICTED: LOW QUALITY PR...  115     7e-25



Query_10             1    GGCTCCCACTCCCTGAGGTATTTCTCCACCGCGGTGTCCCGGCCCGGCCGCGGGGAGCCC  60
CL1150.Contig4_All   4    GGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGCCGCGGGGAGCCC  63
CL1150.Contig16_All  13   GGCTCCCACTCCCTGAGGTATTTCTACACCGGCGTGTCCCGGCCCGGCCGCGGGGAGCCC  72

Query_10             61   CGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACTCT  120
CL1150.Contig4_All   64   CGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACTCT  123
CL1150.Contig16_All  73   CGCTTCATCGCCGTCGGCTACGTGGACGACACGCAGTTCGTGCGGTTCGACAGCGACTCT  132

Query_10             121  GCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGGTGGAGCAGGAGGGGCCGGAGTATTGG  180
CL1150.Contig4_All   124  GCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGGCCGGAGTATTGG  183
CL1150.Contig16_All  133  GCCAGTCGGAGGATGGAGCCGCGGGCGCCGTGGATGGAGCAGGAGGGGCCGGAGTATTGG  192
CL1150.Contig22_All  1                                        GAGGGGCCGGAGTATTGG  18

Query_10             181  GACCGGCAGACGCGGGGGATCAAGGACGCCGCACAGACTTTCCGAGGGAACCTGCAGACC  240
CL1150.Contig4_All   184  GACCAGCAGACGCGGGGGATCAAGGAAACCACACAGACTTACCGACGGAGCCTG-A-ACA  241
CL1150.Contig16_All  193  GACCGGCAGACGCAGATCTGCAAGGAAACCACACAGACTTACCGAGGGAGCCTG-A-ACA  250
CL1150.Contig22_All  19   GACCGGCAGACGCGGAACCTCAAGGACGCCGCACACGCTTTCCGAGTGAACCTGAACACC  78

Query_10             241  GCACTCC---G-CTACTACAACCAGAGCGCGGCC-GGGTCTCACACCATCCAGAGCATGT  295
CL1150.Contig4_All   242  AC-CTGCGGGG-CTACTACAACCAGAGCGCGGCC-GGGTCTCACACCTTCCAGAACATGT  298
CL1150.Contig16_All  251  TC-CTGCGGGG-CTACTACAACCAGAGCGCGGCC-GGGTCTCACACCATCCAGAACTTGT  307
CL1150.Contig22_All  79   -CTGCGG---GACTACTACAACCAGAGCGCGGCC-GG                        110
CL1150.Contig22_All  293                                GGCCAGGGTCTCACACCATCCAGAGCATGT  322

Query_10             296  ACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACAGTCAGGACTCCTACG  355
CL1150.Contig4_All   299  ACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTCGCCTACG  358
CL1150.Contig16_All  308  ACGGCTGTGACGTGGGGCCCGACGGGCGTCTCCTCCGCGGGTACCGTCAGTTCGCCTACG  367
CL1150.Contig22_All  323  ACGGCTGTGACGTGGAGCCCGACGGGCGTCTCCTCCGC                       360

Query_10             356  ACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCGGACACGG  415
CL1150.Contig4_All   359  ACGGCGCGGATTACCTCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCGGACACGG  418
CL1150.Contig16_All  368  ACGGCGCGGATTACATCGCCCTGAACGAGGACCTGCGCTCCTGGACCGCGGCGGACACGG  427

Query_10             416  CGGCGCAGATCACCCAGCGCAAGTGGGAGGACGCGGGTGTGGCAGAGCGCTGGAGGAACT  475
CL1150.Contig4_All   419  CGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTACAGGAACT  478
CL1150.Contig16_All  428  CGGCGCAGATCTCCCGGCGCAAGTGGGAGGACGCGGGTGAGGCAGAGCGCTACAGGAACT  487

Query_10             476  ACCTGGAGGTCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAGGAGTCGC  535
CL1150.Contig4_All   479  ATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAGGAGTCGC  538
CL1150.Contig16_All  488  ATGTGGAGGGCACGTGCGTGGAGTGGCTCGGCAGGTACCTGGAGAACGGGAAGGAGTCGC  547

Query_10             536  TGCTGCGC  543
CL1150.Contig4_All   539  TGCTGCGC  546
CL1150.Contig16_All  548  TGCTGCGC  555


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 42412998768


Query= HQ908107.1 Meles meles nonfunctional MHC class II antigen (Meme-DQB)
pseudogene mRNA, Meme-DQB*PS01 allele, partial sequence

Length=680
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  1123    0.0
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  1085    0.0
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  935     0.0
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  564     7e-160
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  507     1e-142
Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  409     3e-113
Unigene4377_All  98 193 minus strand MHC class II antigen DQ beta...  178     1e-43
CL4065.Contig4_All  98 193 minus strand MHC class II antigen [Zal...  178     1e-43
```

```
CL4065.Contig2_All  98 193 minus strand MHC class II antigen [Zal...  178     1e-43
CL4065.Contig1_All  95 190 MHC class II antigen DQ beta chain, pa...  178     1e-43



Query_11            24   ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  83
Unigene57188_All    1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
Unigene67843_All    1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
Unigene75450_All    1                         GGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  39
Unigene4377_All     1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
CL4065.Contig4_All  1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
CL4065.Contig2_All  1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
CL4065.Contig1_All  1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60

Query_11            84   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCAA-----------------TTTAAG  126
Unigene57188_All    61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAG  120
Unigene67843_All    61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAG  120
Unigene75450_All    40   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAG  99
CL4065.Contig3_All  8                                    TCCAG-----------------TTTAAG  18
Unigene4377_All     61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCA                          96
CL4065.Contig4_All  61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCA                          96
CL4065.Contig2_All  61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCA                          96
CL4065.Contig1_All  61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCA                          96

Query_11            127  GGCGAGTGCTACTTCACCAACGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTAT  186
Unigene57188_All    121  GGCGAGTGCTACTTCACCAACGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTAT  180
Unigene67843_All    121  GGCGAGTGCTACTTCACCAACGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTAT  180
Unigene75449_All    103     GAGTGCCACTTCACCAACGGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTAT  159
Unigene75450_All    100  GGCGAGTGCTACTTCACCAACGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTAT  159
CL4065.Contig3_All  19   GGCGAGTGCTACTTCACCAACGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTAT  78
Unigene27967_All    103     GAGTGCCACTTCACCAACGGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTAT  159

Query_11            187  AACCGGGAGGAGTTCGTGCGCTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  246
Unigene57188_All    181  AACCGGGAGGAGTTCGTGCGCTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  240
Unigene67843_All    181  AACCGGGAGGAGTTCGTGCGCTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  240
Unigene75449_All    160  AACGGCGAGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  219
Unigene75450_All    160  AACCGGGAGGAGTTCGTGCGCTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  219
CL4065.Contig3_All  79   AACCGGGAGGAGTTCGTGCGCTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  138
Unigene27967_All    160  AACGGCGAGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  219

Query_11            247  CTGGGGCGGCCGGACGCTGAGTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAG  306
Unigene57188_All    241  CTGGGGCGGCCGGACGCTCAGTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAG  300
Unigene67843_All    241  CTGGGGCGGCCGGACGCTCAGTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAG  300
Unigene75449_All    220  CTGGGGCGGCCGGACGCTGAGTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAG  279
Unigene75450_All    220  CTGGGGCGGCCGGACGCTCAGTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGG  279
CL4065.Contig3_All  139  CTGGGGCGGCCGGACGCTCAGTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGG  198
Unigene27967_All    220  CTGGGGCGGCCGGACGCTCAGTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGG  279

Query_11            307  GCCGAGACAGACACGGTGTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAG  366
Unigene57188_All    301  GCCGAGACAGACACGGTGTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAG  360
Unigene67843_All    301  GCCGAGACAGACACGGTGTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAG  360
Unigene75449_All    280  GCCGAGACAGACACGGTGTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAG  339
Unigene75450_All    280  GCCGCGGTGGACACATACTGCAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAG  339
CL4065.Contig3_All  199  GCCGCGGTGGACACATACTGCAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAG  258
Unigene27967_All    280  GCCGCGGTGGACACATACTGCAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAG  339

Query_11            367  AGGCGAGTGGAACCTACAGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCAC  426
Unigene57188_All    361  AGGCGAGTGGAACCTACAGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCAC  420
Unigene67843_All    361  AGGCGAGTGGAACCTACAGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCAC  420
Unigene75449_All    340  AGGCGAGTGGAACCTACAGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCAC  399
Unigene75450_All    340  CGGCGAGTGGAGCCTACAGTGACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCAC  399
CL4065.Contig3_All  259  CGGCGAGTGGAGCCTACAGTGACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCAC  318
Unigene27967_All    340  CGGCGAGTGGAGCCTACAGTGACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCAC  399

Query_11            427  AACATGCTGGTCTGCTCGGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTT  486
Unigene57188_All    421  AACATGCTGGTCTGCTCGGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTT  480
Unigene67843_All    421  AACATGCTGGTCTGCTCGGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTT  480
Unigene75449_All    400  AACATGCTGGTCTGCTCGGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTT  459
Unigene75450_All    400  AACCTCCTGGTCTGCTCTGTGAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTC  459
CL4065.Contig3_All  319  AACCTCCTGGTCTGCTCTGTGAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTC  378
Unigene27967_All    400  AACCTCCTGGTCTGCTCTGTGAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTC  459

Query_11            487  CGGAATGACCAGGAGGAGAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGAC  546
Unigene57188_All    481  CGGAATGACCAGGAGGAGAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGAC  540
Unigene67843_All    481  CGGAATGACCAGGAGGAGAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGAC  540
Unigene75449_All    460  CGGAATGACCAGGAGGAGAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGAC  519
Unigene75450_All    460  CGGAATGGCCAGGAAGAGGAGTCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGAC  519
CL4065.Contig3_All  379  CGGAATGGCCAGGAAGAGGAGTCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGAC  438
Unigene27967_All    460  CGGAATGGCCAGGAAGAGGAGTCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGAC  519

Query_11            547  TGGACCTTCCAGATCCTTGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACAC  605
```

```
Unigene57188_All    541  TGGACCTTCCAGATCCTTGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACAC  599
Unigene67843_All    541  TGGACCTTCCAGATCCTTGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACAC  599
Unigene75449_All    520  TGGACCTTCCAGATCCTTGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACAC  578
Unigene75450_All    520  TGGACCTTCCAGACCCTGGTGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACAC  578
CL4065.Contig3_All  439  TGGACCTTCCAGACCCTGGTGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACAC  497
Unigene27967_All    520  TGGACCTTCCAGACCCTGGTGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACAC  578

Query_11            606  CTGCCATGTGGAGCACCCCAGCCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTC  665
Unigene57188_All    600  CTGCCATGTGGAGCACCCCAGCCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTC  659
Unigene67843_All    600  CTGCCATGTGGAGCACCCCAGCCTCCAGAGCCCCATCACAGTGGAGTGGCG           650
Unigene75449_All    579  CTGCCATGTGGAGCACCCCAGCCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTC  638
Unigene75450_All    579  CTGCCAAGTGGAGCACCCCAGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTC  638
CL4065.Contig3_All  498  CTGCCAAGTGGAGCACCCCAGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTC  557
Unigene27967_All    579  CTGCCAAGTGGAGCACCCCAGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTC  638

Query_11            666  TGAATCTGCCCAGAG  680
Unigene57188_All    660  GGAATCTGCCCAGAG  674
Unigene75449_All    639  GGAATCTGCCCAGAG  653
Unigene75450_All    639  TGGGTCTGCACAGAG  653
CL4065.Contig3_All  558  TGGGTCTGCACAGAG  572
Unigene27967_All    639  TGGGTCTGCACAGAG  653


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 53465137650


Query= HQ908099.1 Meles meles MHC class II antigen DR alpha chain
(Meme-DRA) mRNA, Meme-DRA*02 allele, partial cds

Length=691
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL2981.Contig3_All  111 872 minus strand PREDICTED: HLA class II ...  1243    0.0
CL2981.Contig1_All  111 872 minus strand PREDICTED: HLA class II ...  1243    0.0
Unigene42913_All  108 710 MHC class II antigen DR alpha chain, pa...  1103    0.0



Query_12            1    CATAAGTGGAGTCCCTGTGCTAGGATTTTTCATCATGACTTTACTGATGGGTCCCCAAGA  60
CL2981.Contig3_All  6    CATAAATGGAGTCCCAGTGCTAGGATTTTTCATCATGACTTTACTGATGGGTCCCCAAGA  65
CL2981.Contig1_All  6    CATAAATGGAGTCCCAGTGCTAGGATTTTTCATCATGACTTTACTGATGGGTCCCCAAGA  65
Unigene42913_All    1     ATAAATGGAGTCCCAGTGCTAGGATTTTTCATCATGACTTTACTGATGGGTCCCCAAGA  59

Query_12            61   ATCACAGGCTATCAAAGAGGACCATGTGATCATCCAGGCTGAGTTTTATCTGACCCCTGA  120
CL2981.Contig3_All  66   ATCACAGGCTATCAAAGAGGACCATGTGATCATCCAGGCTGAGTTTTATCTGACCCCTGA  125
CL2981.Contig1_All  66   ATCACAGGCTATCAAAGAGGACCATGTGATCATCCAGGCTGAGTTTTATCTGACCCCTGA  125
Unigene42913_All    60   ATCACAGGCTATCAAAGAGGACCATGTGATCATCCAGGCTGAGTTTTATCTGACCCCTGA  119

Query_12            121  CCCGTCAGGCGAGTTTATGTTTGACTTCGATGGTGATGAGATTTTCCACGTGGATATGGA  180
CL2981.Contig3_All  126  CCCGTCAGGCGAGTTTATGTTTGACTTCGATGGTGATGAGATTTTCCACGTGGATATGGA  185
CL2981.Contig1_All  126  CCCGTCAGGCGAGTTTATGTTTGACTTCGATGGTGATGAGATTTTCCACGTGGATATGGA  185
Unigene42913_All    120  CCCGTCAGGCGAGTTTATGTTTGACTTCGATGGTGATGAGATTTTCCACGTGGATATGGA  179

Query_12            181  AAAGAAGGAGACAGTGTGGCGGCTGGAAGAATTTGGACGCTTTGCCAGCTTTGAGGCACA  240
CL2981.Contig3_All  186  AAAGAAGGAGACAGTGTGGCGGCTGGAAGAATTTGGACGCTTTGCCAGCTTTGAGGCACA  245
CL2981.Contig1_All  186  AAAGAAGGAGACAGTGTGGCGGCTGGAAGAATTTGGACGCTTTGCCAGCTTTGAGGCACA  245
Unigene42913_All    180  AAAGAAGGAGACAGTGTGGCGGCTGGAAGAATTTGGACGCTTTGCCAGCTTTGAGGCACA  239

Query_12            241  GGGTGCCTTGGCCAACATAGCTGTGGACAAAGCTAACCTGGACATCATGATAAAGCGCTC  300
CL2981.Contig3_All  246  GGGTGCCTTGGCCAACATAGCTGTGGACAAAGCTAACCTGGACATCATGATAAAGCGCTC  305
CL2981.Contig1_All  246  GGGTGCCTTGGCCAACATAGCTGTGGACAAAGCTAACCTGGACATCATGATAAAGCGCTC  305
Unigene42913_All    240  GGGTGCCTTGGCCAACATAGCTGTGGACAAAGCTAACCTGGACATCATGATAAAGCGCTC  299

Query_12            301  CAACCACACCCCAAACACCAATGTACCTCCAGAGGTGACCGTGCTCTCAAACACCCCTGT  360
CL2981.Contig3_All  306  CAACCACACCCCAAACACCAATGTACCTCCAGAGGTGACCGTGCTCTCAAACACCCCTGT  365
CL2981.Contig1_All  306  CAACCACACCCCAAACACCAATGTACCTCCAGAGGTGACCGTGCTCTCAAACACCCCTGT  365
Unigene42913_All    300  CAACCACACCCCAAACACCAATGTACCTCCAGAGGTGACCGTGCTCTCAAACACCCCTGT  359

Query_12            361  GGAACTGGGAGAGCCCAACACCCTCATCTGCTTCATCGACAAGTTCTCCCCACCAGTGAT  420
CL2981.Contig3_All  366  GGAACTGGGAGAGCCCAACACCCTCATCTGCTTCATCGACAAGTTCTCCCCACCAGTGAT  425
CL2981.Contig1_All  366  GGAACTGGGAGAGCCCAACACCCTCATCTGCTTCATCGACAAGTTCTCCCCACCAGTGAT  425
Unigene42913_All    360  GGAACTGGGAGAGCCCAACACCCTCATCTGCTTCATCGACAAGTTCTCCCCACCAGTGAT  419
```

```
Query_12            421  CAATGTCACGTGGCTTCGAAATGGAAACCCTGTCACCACAGGAGTGTCCGAGACAGTCTT  480
CL2981.Contig3_All  426  CAATGTCACGTGGCTTCGAAATGGAAACCCTGTCACCACAGGAGTGTCCGAGACAGTCTT  485
CL2981.Contig1_All  426  CAATGTCACGTGGCTTCGAAATGGAAACCCTGTCACCACAGGAGTGTCCGAGACAGTCTT  485
Unigene42913_All    420  CAATGTCACGTGGCTTCGAAATGGAAACCCTGTCACCACAGGAGTGTCCGAGACAGTCTT  479

Query_12            481  CCTGCCCAGGGAAGACCACCTTTTCCGCAAGTTCCACTATCTCCCCTTCCTGCCCTCAGC  540
CL2981.Contig3_All  486  CCTGCCCAGGGAAGACCACCTTTTCCGCAAGTTCCACTATCTCCCCTTCCTGCCCTCAGC  545
CL2981.Contig1_All  486  CCTGCCCAGGGAAGACCACCTTTTCCGCAAGTTCCACTATCTCCCCTTCCTGCCCTCAGC  545
Unigene42913_All    480  CCTGCCCAGGGAAGACCACCTTTTCCGCAAGTTCCACTATCTCCCCTTCCTGCCCTCAGC  539

Query_12            541  CAACGATGTCTATGACTGCAAGGTGGAGCACTGGGGTCTGGATGAGCCTCTTCTCAAGCA  600
CL2981.Contig3_All  546  CAACGATGTCTATGACTGCAAGGTGGAGCACTGGGGTCTGGATGAGCCTCTTCTCAAGCA  605
CL2981.Contig1_All  546  CAACGATGTCTATGACTGCAAGGTGGAGCACTGGGGTCTGGATGAGCCTCTTCTCAAGCA  605
Unigene42913_All    540  CAACGATGTCTATGACTGCAAGGTGGAGCACTGGGGTCTGGATGAGCCTCTTCTCAAGCA  599

Query_12            601  CTGGGAGTTTGAACCACCAACTCCTCTCCCAGAGACAACCGAGAATGTGGTGTGTGCCCT  660
CL2981.Contig3_All  606  CTGGGAGTTTGAACCACCAACTCCTCTCCCAGAGACAACCGAGAATGTGGTGTGTGCCCT  665
CL2981.Contig1_All  606  CTGGGAGTTTGAACCACCAACTCCTCTCCCAGAGACAACCGAGAATGTGGTGTGTGCCCT  665
Unigene42913_All    600  CTGG                                                        603

Query_12            661  GGGCCTGGTTGTGGGTCTGGTTGGCATC  688
CL2981.Contig3_All  666  GGGCCTGGTTGTGGGTTTGGTGGGTATC  693
CL2981.Contig1_All  666  GGGCCTGGTTGTGGGTTTGGTGGGTATC  693


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 54364398375


Query= HQ908098.1 Meles meles MHC class II antigen DQ alpha chain
(Meme-DQA) mRNA, Meme-DQA*02 allele, partial cds

Length=615
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL10050.Contig3_All  75 839 minus strand PREDICTED: SLA class II ...  1125    0.0
CL10050.Contig2_All  105 686 minus strand MHC class II antigen DQ...  1075    0.0
CL10050.Contig1_All  105 686 minus strand MHC class II antigen DQ...  1075    0.0



Query_13            1    ACTCTTGCCCTGACCACCATGATGAGCCCTGGTGGCAGTGAAGACATTGTGGCTGACCAT  60
CL10050.Contig3_All 31   ACTCTTGCCCTGACCACCATGATGAGCCCTGGTGGCAGTGAAGACATTGTGGCTGACCAT  90
CL10050.Contig2_All 1    ACTCTTGCCCTGACCACCATGATGAGCCCTGGTGGCAGTGAAGACATTGTGGCTGACCAT  60
CL10050.Contig1_All 1    ACTCTTGCCCTGACCACCATGATGAGCCCTGGTGGCAGTGAAGACATTGTGGCTGACCAT  60

Query_13            61   GTTGGTGCCTATGGCGTAGAAGTCTACCAGTCTTATGGTCCCTCTGGCCAATACACTCAA  120
CL10050.Contig3_All 91   GTTGGTGCCTATGGCGTAGAAGTCTACCAGTCTTATGGTCCCTCTGGCCAATACACTCAA  150
CL10050.Contig2_All 61   GTTGGTGCCTATGGCGTAGAAGTCTACCAGTCTTATGGTCCCTCTGGCCAATACACTCAA  120
CL10050.Contig1_All 61   GTTGGTGCCTATGGCGTAGAAGTCTACCAGTCTTATGGTCCCTCTGGCCAATACACTCAA  120

Query_13            121  GAATTTGATGGAGATGAGTTGTTCTATGTGGACCTGGAGAAGAAGGAAACTGTCTGGCGG  180
CL10050.Contig3_All 151  GAATTTGATGGAGATGAGTTGTTCTATGTGGACCTGGAGAAGAAGGAAACTGTCTGGCGG  210
CL10050.Contig2_All 121  GAATTTGATGGAGATGAGTTGTTCTATGTGGACCTGGAGAAGAAGGAAACTGTCTGGCGG  180
CL10050.Contig1_All 121  GAATTTGATGGAGATGAGTTGTTCTATGTGGACCTGGAGAAGAAGGAAACTGTCTGGCGG  180

Query_13            181  CTGCCTGTGTTTAGCACATTTGCAGGTTTTGACCCACAAGGTGCACTGAGCGAAATAGCT  240
CL10050.Contig3_All 211  CTGCCTGTGTTTAGCACATTTGCAGGTTTTGACCCACAAGGTGCACTGAGCGAAATAGCT  270
CL10050.Contig2_All 181  CTGCCTGTGTTTAGCACATTTGCAGGTTTTGACCCACAAGGTGCACTGAGCGAAATAGCT  240
CL10050.Contig1_All 181  CTGCCTGTGTTTAGCACATTTGCAGGTTTTGACCCACAAGGTGCACTGAGCGAAATAGCT  240

Query_13            241  ACATCAAAACAAAACTTGAACATCCTGACTAAACGCTCCAACTATACCGCTGCTACCAAT  300
CL10050.Contig3_All 271  ACATCAAAACAAAACTTGAACATCCTGACTAAACGCTCCAACTATACCGCTGCTACCAAT  330
CL10050.Contig2_All 241  ACATCAAAACAAAACTTGAACATCCTGACTAAACGCTCCAACTATACCGCTGCTACCAAT  300
CL10050.Contig1_All 241  ACATCAAAACAAAACTTGAACATCCTGACTAAACGCTCCAACTATACCGCTGCTACCAAT  300

Query_13            301  GAGGTTCCTGAGGTGACGCTGTTTCCCAAGTCTCCTGTGATGCTGGGTCAGCCCAACACC  360
CL10050.Contig3_All 331  GAGGTTCCTGAGGTGACGCTGTTTCCCAAGTCTCCTGTGATGCTGGGTCAGCCCAACACC  390
CL10050.Contig2_All 301  GAGGTTCCTGAGGTGACGCTGTTTCCCAAGTCTCCTGTGATGCTGGGTCAGCCCAACACC  360
CL10050.Contig1_All 301  GAGGTTCCTGAGGTGACGCTGTTTCCCAAGTCTCCTGTGATGCTGGGTCAGCCCAACACC  360

Query_13            361  CTCATCTGTCTTGTGGACAACATCTTTCCTCCTGTGATCAATGTCACGTGGTTGAAGAAC  420
CL10050.Contig3_All 391  CTCATCTGTCTTGTGGACAACATCTTTCCTCCTGTGATCAATGTCACGTGGTTGAAGAAC  450
CL10050.Contig2_All 361  CTCATCTGTCTTGTGGACAACATCTTTCCTCCTGTGATCAATGTCACGTGGTTGAAGAAC  420
```

```
CL10050.Contig1_All  361  CTCATCTGTCTTGTGGACAACATCTTTCCTCCTGTGATCAATGTCACGTGGTTGAAGAAC  420

Query_13             421  AGGCACTCAGTCACAGAAGGTGTTTCTGAAACCAGCTTCCTTGCCAAGAAGGATCATTCC  480
CL10050.Contig3_All  451  AGGCACTCAGTCACAGAAGGTGTTTCTGAAACCAGCTTCCTTGCCAAGAAGGATCATTCC  510
CL10050.Contig2_All  421  AGGCACTCAGTCACAGAAGGTGTTTCTGAAACCAGCTTCCTTGCCAAGAAGGATCATTCC  480
CL10050.Contig1_All  421  AGGCACTCAGTCACAGAAGGTGTTTCTGAAACCAGCTTCCTTGCCAAGAAGGATCATTCC  480

Query_13             481  TTCTTAAAGATCAGTTACCTCACCTTCCTCCCTTCTGCTGATGATATTTATGACTGCAAG  540
CL10050.Contig3_All  511  TTCTTAAAGATCAGTTACCTCACCTTCCTCCCTTCTGCTGATGATATTTATGACTGCAAG  570
CL10050.Contig2_All  481  TTCTTAAAGATCAGTTACCTCACCTTCCTCCCTTCTGCTGATGATATTTATGACTGCAAG  540
CL10050.Contig1_All  481  TTCTTAAAGATCAGTTACCTCACCTTCCTCCCTTCTGCTGATGATATTTATGACTGCAAG  540

Query_13             541  GTGGAGCACTGGGGCCTGGATGAACCACTTCTGAAACACTGGGAACCTGAAATTCCAACC  600
CL10050.Contig3_All  571  GTGGAGCACTGGGGCCTGGATGAACCACTTCTGAAACACTGGGAACCTGAAATTCCATCC  630
CL10050.Contig2_All  541  GTGGAGCACTGGGGCCTGGATGAACCACTTCTGAAACACTGG                    582
CL10050.Contig1_All  541  GTGGAGCACTGGGGCCTGGATGAACCACTTCTGAAACACTGG                    582

Query_13             601  CCTATGTCAGAGCTG  615
CL10050.Contig3_All  631  CCCATGTCAGAGCTG  645


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 48308241840


Query= HQ908097.1 Meles meles MHC class II antigen DQ alpha chain
(Meme-DQA) mRNA, Meme-DQA*01 allele, partial cds

Length=615
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL10050.Contig3_All  75 839 minus strand PREDICTED: SLA class II ...  854     0.0
CL10050.Contig2_All  105 686 minus strand MHC class II antigen DQ...  804     0.0
CL10050.Contig1_All  105 686 minus strand MHC class II antigen DQ...  804     0.0



Query_14             1    ACTCTTGCCCTGACCACCATAATGAGCCTTGGTGGCAGTGAAGACATTGTGGCTGATCAT  60
CL10050.Contig3_All  31   ACTCTTGCCCTGACCACCATGATGAGCCCTGGTGGCAGTGAAGACATTGTGGCTGACCAT  90
CL10050.Contig2_All  1    ACTCTTGCCCTGACCACCATGATGAGCCCTGGTGGCAGTGAAGACATTGTGGCTGACCAT  60
CL10050.Contig1_All  1    ACTCTTGCCCTGACCACCATGATGAGCCCTGGTGGCAGTGAAGACATTGTGGCTGACCAT  60

Query_14             61   GTTGCTTCCTATGGCATAAGTGTCTACCAGTCTTATGGTCCCTCTGGCCAGTACACCCGT  120
CL10050.Contig3_All  91   GTTGGTGCCTATGGCGTAGAAGTCTACCAGTCTTATGGTCCCTCTGGCCAATACACTCAA  150
CL10050.Contig2_All  61   GTTGGTGCCTATGGCGTAGAAGTCTACCAGTCTTATGGTCCCTCTGGCCAATACACTCAA  120
CL10050.Contig1_All  61   GTTGGTGCCTATGGCGTAGAAGTCTACCAGTCTTATGGTCCCTCTGGCCAATACACTCAA  120

Query_14             121  GAATTTGATGGTGATGAGGAATTCTACGTGGACTTGGAGAAGAAGGAGACAGTCTGGCAG  180
CL10050.Contig3_All  151  GAATTTGATGGAGATGAGTTGTTCTATGTGGACCTGGAGAAGAAGGAAACTGTCTGGCGG  210
CL10050.Contig2_All  121  GAATTTGATGGAGATGAGTTGTTCTATGTGGACCTGGAGAAGAAGGAAACTGTCTGGCGG  180
CL10050.Contig1_All  121  GAATTTGATGGAGATGAGTTGTTCTATGTGGACCTGGAGAAGAAGGAAACTGTCTGGCGG  180

Query_14             181  CTGCCCATGTTTCAAGCAC-TTAG-ACGTTTTGACCCACAAGGTGCACTGAGAAACTTGG  238
CL10050.Contig3_All  211  CTGCCTGTGTTT--AGCACATTTGCAGGTTTTGACCCACAAGGTGCACTGAGCGAAATAG  268
CL10050.Contig2_All  181  CTGCCTGTGTTT--AGCACATTTGCAGGTTTTGACCCACAAGGTGCACTGAGCGAAATAG  238
CL10050.Contig1_All  181  CTGCCTGTGTTT--AGCACATTTGCAGGTTTTGACCCACAAGGTGCACTGAGCGAAATAG  238

Query_14             239  CAATAGCAAAACAAAACTTGAACATCCTGACTAAACGCTCCAACTATACCGCTGCTACCA  298
CL10050.Contig3_All  269  CTACATCAAAACAAAACTTGAACATCCTGACTAAACGCTCCAACTATACCGCTGCTACCA  328
CL10050.Contig2_All  239  CTACATCAAAACAAAACTTGAACATCCTGACTAAACGCTCCAACTATACCGCTGCTACCA  298
CL10050.Contig1_All  239  CTACATCAAAACAAAACTTGAACATCCTGACTAAACGCTCCAACTATACCGCTGCTACCA  298

Query_14             299  ATGAGGTTCCTGAGGTGACGCTGTTTCTAAAGACTCCTGTGATGCTGGGTCAGCCCAACA  358
CL10050.Contig3_All  329  ATGAGGTTCCTGAGGTGACGCTGTTTCCCAAGTCTCCTGTGATGCTGGGTCAGCCCAACA  388
CL10050.Contig2_All  299  ATGAGGTTCCTGAGGTGACGCTGTTTCCCAAGTCTCCTGTGATGCTGGGTCAGCCCAACA  358
CL10050.Contig1_All  299  ATGAGGTTCCTGAGGTGACGCTGTTTCCCAAGTCTCCTGTGATGCTGGGTCAGCCCAACA  358

Query_14             359  CCCTCATCTGTCTTGTGGACAACATCTTTCCTCCTGTGATCAATGTCACGTGGTTGAAGA  418
CL10050.Contig3_All  389  CCCTCATCTGTCTTGTGGACAACATCTTTCCTCCTGTGATCAATGTCACGTGGTTGAAGA  448
CL10050.Contig2_All  359  CCCTCATCTGTCTTGTGGACAACATCTTTCCTCCTGTGATCAATGTCACGTGGTTGAAGA  418
CL10050.Contig1_All  359  CCCTCATCTGTCTTGTGGACAACATCTTTCCTCCTGTGATCAATGTCACGTGGTTGAAGA  418

Query_14             419  ACAGGCATTCAGTCACAGAAGGTGTTTCTGAAACCCACTTCCTTATCAAAAAGGATTATT  478
CL10050.Contig3_All  449  ACAGGCACTCAGTCACAGAAGGTGTTTCTGAAACCAGCTTCCTTGCCAAGAAGGATCATT  508
```

```
CL10050.Contig2_All  419  ACAGGCACTCAGTCACAGAAGGTGTTTCTGAAACCAGCTTCCTTGCCAAGAAGGATCATT  478
CL10050.Contig1_All  419  ACAGGCACTCAGTCACAGAAGGTGTTTCTGAAACCAGCTTCCTTGCCAAGAAGGATCATT  478

Query_14             479  CCTTCTTAAAGTTCAGTTACCTCACCTTCCTCCCTTCTGCTGATGATATTTATGACTGCA  538
CL10050.Contig3_All  509  CCTTCTTAAAGATCAGTTACCTCACCTTCCTCCCTTCTGCTGATGATATTTATGACTGCA  568
CL10050.Contig2_All  479  CCTTCTTAAAGATCAGTTACCTCACCTTCCTCCCTTCTGCTGATGATATTTATGACTGCA  538
CL10050.Contig1_All  479  CCTTCTTAAAGATCAGTTACCTCACCTTCCTCCCTTCTGCTGATGATATTTATGACTGCA  538

Query_14             539  AGGTGGAGCACTGGGGCCTGGATGAACCACTTCTGAAACACTGGGAACCTGAAATTCCAA  598
CL10050.Contig3_All  569  AGGTGGAGCACTGGGGCCTGGATGAACCACTTCTGAAACACTGGGAACCTGAAATTCCAT  628
CL10050.Contig2_All  539  AGGTGGAGCACTGGGGCCTGGATGAACCACTTCTGAAACACTGG                  582
CL10050.Contig1_All  539  AGGTGGAGCACTGGGGCCTGGATGAACCACTTCTGAAACACTGG                  582

Query_14             599  CCCCTATGTCAGAGCTG  615
CL10050.Contig3_All  629  CCCCCATGTCAGAGCTG  645


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 48308241840


Query= HQ908096.1 Meles meles MHC class II antigen DQ beta chain (Meme-DQB)
mRNA, Meme-DQB*02 allele, partial cds

Length=697
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  1234    0.0
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  1195    0.0
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  935     0.0
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  675     0.0
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  525     3e-148
Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  409     4e-113
Unigene4377_All  98 193 minus strand MHC class II antigen DQ beta...  178     1e-43
CL4065.Contig4_All  98 193 minus strand MHC class II antigen [Zal...  178     1e-43
CL4065.Contig2_All  98 193 minus strand MHC class II antigen [Zal...  178     1e-43
CL4065.Contig1_All  95 190 MHC class II antigen DQ beta chain, pa...  178     1e-43



Query_15             24   ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  83
Unigene57188_All     1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
Unigene67843_All     1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
Unigene75450_All     1                          GGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  39
Unigene4377_All      1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
CL4065.Contig4_All   1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
CL4065.Contig2_All   1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60
CL4065.Contig1_All   1    ATGGCACTGTGGATCCCCAGAGGCCTCTGGACAGCAGCTGTGATGGCGATCTTGGTGGTG  60

Query_15             84   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAG  143
Unigene57188_All     61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAG  120
Unigene67843_All     61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAG  120
Unigene75450_All     40   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAG  99
CL4065.Contig3_All   1                                        TTCGTGTTCCAGTTTAAG  18
Unigene4377_All      61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCA                       96
CL4065.Contig4_All   61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCA                       96
CL4065.Contig2_All   61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCA                       96
CL4065.Contig1_All   61   CTGAGCGTCCCAGTGGCTGAGGGCAGAGACTCTCCA                       96

Query_15             144  GGCGAGTGCTACTTCACCAACGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTAT  203
Unigene57188_All     121  GGCGAGTGCTACTTCACCAACGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTAT  180
Unigene67843_All     121  GGCGAGTGCTACTTCACCAACGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTAT  180
Unigene75449_All     103     GAGTGCCACTTCACCAACGGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTAT  159
Unigene75450_All     100  GGCGAGTGCTACTTCACCAACGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTAT  159
CL4065.Contig3_All   19   GGCGAGTGCTACTTCACCAACGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTAT  78
Unigene27967_All     103     GAGTGCCACTTCACCAACGGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTAT  159

Query_15             204  AACCGGGAGGAGTTCGTGCGCTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  263
Unigene57188_All     181  AACCGGGAGGAGTTCGTGCGCTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  240
Unigene67843_All     181  AACCGGGAGGAGTTCGTGCGCTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  240
Unigene75449_All     160  AACGGCGAGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  219
Unigene75450_All     160  AACCGGGAGGAGTTCGTGCGCTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  219
CL4065.Contig3_All   79   AACCGGGAGGAGTTCGTGCGCTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  138
Unigene27967_All     160  AACGGCGAGGAGTACGTGCGCTTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAG  219
```

```
Query_15            264  CTGGGGCGGCCGGACGCTGAGTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAG  323
Unigene57188_All    241  CTGGGGCGGCCGGACGCTCAGTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAG  300
Unigene67843_All    241  CTGGGGCGGCCGGACGCTCAGTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAG  300
Unigene75449_All    220  CTGGGGCGGCCGGACGCTGAGTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAG  279
Unigene75450_All    220  CTGGGGCGGCCGGACGCTCAGTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGG  279
CL4065.Contig3_All  139  CTGGGGCGGCCGGACGCTCAGTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGG  198
Unigene27967_All    220  CTGGGGCGGCCGGACGCTCAGTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGG  279

Query_15            324  GCCGAGACAGACACGGTGTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAG  383
Unigene57188_All    301  GCCGAGACAGACACGGTGTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAG  360
Unigene67843_All    301  GCCGAGACAGACACGGTGTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAG  360
Unigene75449_All    280  GCCGAGACAGACACGGTGTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAG  339
Unigene75450_All    280  GCCGCGGTGGACACATACTGCAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAG  339
CL4065.Contig3_All  199  GCCGCGGTGGACACATACTGCAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAG  258
Unigene27967_All    280  GCCGCGGTGGACACATACTGCAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAG  339

Query_15            384  AGGCGAGTGGAACCTACAGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCAC  443
Unigene57188_All    361  AGGCGAGTGGAACCTACAGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCAC  420
Unigene67843_All    361  AGGCGAGTGGAACCTACAGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCAC  420
Unigene75449_All    340  AGGCGAGTGGAACCTACAGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCAC  399
Unigene75450_All    340  CGGCGAGTGGAGCCTACAGTGACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCAC  399
CL4065.Contig3_All  259  CGGCGAGTGGAGCCTACAGTGACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCAC  318
Unigene27967_All    340  CGGCGAGTGGAGCCTACAGTGACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCAC  399

Query_15            444  AACATGCTGGTCTGCTCGGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTT  503
Unigene57188_All    421  AACATGCTGGTCTGCTCGGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTT  480
Unigene67843_All    421  AACATGCTGGTCTGCTCGGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTT  480
Unigene75449_All    400  AACATGCTGGTCTGCTCGGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTT  459
Unigene75450_All    400  AACCTCCTGGTCTGCTCTGTGAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTC  459
CL4065.Contig3_All  319  AACCTCCTGGTCTGCTCTGTGAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTC  378
Unigene27967_All    400  AACCTCCTGGTCTGCTCTGTGAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTC  459

Query_15            504  CGGAATGACCAGGAGGAGAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGAC  563
Unigene57188_All    481  CGGAATGACCAGGAGGAGAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGAC  540
Unigene67843_All    481  CGGAATGACCAGGAGGAGAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGAC  540
Unigene75449_All    460  CGGAATGACCAGGAGGAGAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGAC  519
Unigene75450_All    460  CGGAATGGCCAGGAAGAGGAGTCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGAC  519
CL4065.Contig3_All  379  CGGAATGGCCAGGAAGAGGAGTCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGAC  438
Unigene27967_All    460  CGGAATGGCCAGGAAGAGGAGTCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGAC  519

Query_15            564  TGGACCTTCCAGATCCTTGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACAC  622
Unigene57188_All    541  TGGACCTTCCAGATCCTTGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACAC  599
Unigene67843_All    541  TGGACCTTCCAGATCCTTGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACAC  599
Unigene75449_All    520  TGGACCTTCCAGATCCTTGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACAC  578
Unigene75450_All    520  TGGACCTTCCAGACCCTGGTGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACAC  578
CL4065.Contig3_All  439  TGGACCTTCCAGACCCTGGTGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACAC  497
Unigene27967_All    520  TGGACCTTCCAGACCCTGGTGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACAC  578

Query_15            623  CTGCCATGTGGAGCACCCCAGCCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTC  682
Unigene57188_All    600  CTGCCATGTGGAGCACCCCAGCCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTC  659
Unigene67843_All    600  CTGCCATGTGGAGCACCCCAGCCTCCAGAGCCCCATCACAGTGGAGTGGCG            650
Unigene75449_All    579  CTGCCATGTGGAGCACCCCAGCCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTC  638
Unigene75450_All    579  CTGCCAAGTGGAGCACCCCAGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTC  638
CL4065.Contig3_All  498  CTGCCAAGTGGAGCACCCCAGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTC  557
Unigene27967_All    579  CTGCCAAGTGGAGCACCCCAGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTC  638

Query_15            683  TGAATCTGCCCAGAG  697
Unigene57188_All    660  GGAATCTGCCCAGAG  674
Unigene75449_All    639  GGAATCTGCCCAGAG  653
Unigene75450_All    639  TGGGTCTGCACAGAG  653
CL4065.Contig3_All  558  TGGGTCTGCACAGAG  572
Unigene27967_All    639  TGGGTCTGCACAGAG  653


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 54854904225


Query= HQ908095.1 Meles meles MHC class II antigen DR beta chain (Meme-DRB)
mRNA, Meme-DRB*04 allele, partial cds

Length=822
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value
```

```
Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  1208    0.0
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  1146    0.0
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  1136    0.0
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  593     1e-168
Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  536     2e-151
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  459     4e-128
Unigene81593_All  2 94 minus strand MHC class II antigen, partial...  139     6e-32


Query_16            3    GGCTCCTGGATGACAGCTTTGACACTGATACTGATGGTGCTGAGCCCTCCC-TTGGCTTG  61
Unigene27967_All    1    GGCGCCTGGATGACAGCTCTGACACTGATACTGATGGTGCTGAGCCCTCCC-TTGGCTTG  59
Unigene75450_All    35                                      TGGTGCTGAG-CGTCCCAGTGGCTGA  59
Unigene75449_All    1    GGCGCCTGGATGACAGCTCTGACACTGATACTGATGGTGCTGAGCCCTCCC-TTGGCTTG  59
Unigene57188_All    56                                      TGGTGCTGAG-CGTCCCAGTGGCTGA  80
Unigene67843_All    56                                      TGGTGCTGAG-CGTCCCAGTGGCTGA  80

Query_16            62   GGCCAGGGACACCCCACGACATTTCCTGATGCAGTTTAAGGGCGAGTGCTACTTCACCAA  121
Unigene27967_All    60   GGCCAGGGACACCCCACGACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAA  119
Unigene75450_All    60   GGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAGGGCGAGTGCTACTTCACCAA  119
CL4065.Contig3_All  10                                   CAGTTTAAGGGCGAGTGCTACTTCACCAA  38
Unigene75449_All    60   GGCCAGGGACACCCCACGACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAA  119
Unigene57188_All    81   GGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAGGGCGAGTGCTACTTCACCAA  140
Unigene67843_All    81   GGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAGGGCGAGTGCTACTTCACCAA  140

Query_16            122  TGGGACGGAGCGGGTGCGGCTCCTGGTCAGACACATCTATAACCGGGAGGAGTTCGTGCG  181
Unigene27967_All    120  CGGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCG  179
Unigene75450_All    120  CGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCG  179
CL4065.Contig3_All  39   CGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCG  98
Unigene75449_All    120  CGGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCG  179
Unigene57188_All    141  CGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCG  200
Unigene67843_All    141  CGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCG  200

Query_16            182  CTTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCCATCGCTCA  241
Unigene27967_All    180  CTTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCA  239
Unigene75450_All    180  CTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCA  239
CL4065.Contig3_All  99   CTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCA  158
Unigene75449_All    180  CTTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTGA  239
Unigene57188_All    201  CTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCA  260
Unigene67843_All    201  CTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCA  260

Query_16            242  GGGCTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGAGGTGGACACCGTGTG  301
Unigene27967_All    240  GTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGCGGTGGACACATACTG  299
Unigene75450_All    240  GTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGCGGTGGACACATACTG  299
CL4065.Contig3_All  159  GTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGCGGTGGACACATACTG  218
Unigene75449_All    240  GTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAGGCCGAGACAGACACGGTGTG  299
Unigene57188_All    261  GTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAGGCCGAGACAGACACGGTGTG  320
Unigene67843_All    261  GTACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAGGCCGAGACAGACACGGTGTG  320
Unigene81593_All    1                 CAGAAGGACATCATGGAGCGGAAGAGGTCCGAGGTGGACACCGTGTG  47

Query_16            302  CAGACACAACTACGGGGTGGTTGAGAGCTTCACGGTGCAGCGGCGAGTGGAGCCTACAGT  361
Unigene27967_All    300  CAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGT  359
Unigene75450_All    300  CAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGT  359
CL4065.Contig3_All  219  CAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGT  278
Unigene75449_All    300  CAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTACAGT  359
Unigene57188_All    321  CAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTACAGT  380
Unigene67843_All    321  CAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTACAGT  380
Unigene81593_All    48   CAGACACAACCACGGGGTGTTTGAGAGCTTC-CGGTGCAGCGGCGAG               93

Query_16            362  GACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGT  421
Unigene27967_All    360  GACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGT  419
Unigene75450_All    360  GACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGT  419
CL4065.Contig3_All  279  GACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGT  338
Unigene75449_All    360  GACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTCGGT  419
Unigene57188_All    381  GACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTCGGT  440
Unigene67843_All    381  GACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTCGGT  440

Query_16            422  GAATGGTTTCTATCCAGGCCACATTGAAGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGA  481
Unigene27967_All    420  GAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGA  479
Unigene75450_All    420  GAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGA  479
CL4065.Contig3_All  339  GAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGA  398
Unigene75449_All    420  GACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGAGAA  479
Unigene57188_All    441  GACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGAGAA  500
Unigene67843_All    441  GACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGAGAA  500

Query_16            482  GTCTGGGGTCGTGTCCACAGG-CTTGATCCCTAATGGAGACTGGACCTTCCAGACCCTGG  540
Unigene27967_All    480  GTCTGGGGTCGTGTCCACAGG-CCTGATCCGTAATGGAGACTGGACCTTCCAGACCCTGG  538
Unigene75450_All    480  GTCTGGGGTCGTGTCCACAGG-CCTGATCCGTAATGGAGACTGGACCTTCCAGACCCTGG  538
CL4065.Contig3_All  399  GTCTGGGGTCGTGTCCACAGG-CCTGATCCGTAATGGAGACTGGACCTTCCAGACCCTGG  457
Unigene75449_All    480  AGCTGGTGTGGTGTCCACTCCACTT-ATTAGGAATGGGGACTGGACCTTCCAGATCCTTG  538
```

```
Unigene57188_All    501  AGCTGGTGTGGTGTCCACTCCACTT-ATTAGGAATGGGGACTGGACCTTCCAGATCCTTG  559
Unigene67843_All    501  AGCTGGTGTGGTGTCCACTCCACTT-ATTAGGAATGGGGACTGGACCTTCCAGATCCTTG  559

Query_16            541  TGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCC  599
Unigene27967_All    539  TGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCC  597
Unigene75450_All    539  TGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCC  597
CL4065.Contig3_All  458  TGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCC  516
Unigene75449_All    539  TGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACCCC  597
Unigene57188_All    560  TGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACCCC  618
Unigene67843_All    560  TGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACCCC  618

Query_16            600  AGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAG  659
Unigene27967_All    598  AGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAG  657
Unigene75450_All    598  AGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAG  657
CL4065.Contig3_All  517  AGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAG  576
Unigene75449_All    598  AGCCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTCGGAATCTGCCCAGAGCAAG  657
Unigene57188_All    619  AGCCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTCGGAATCTGCCCAGAGCAAG  678
Unigene67843_All    619  AGCCTCCAGAGCCCCATCACAGTGGAGTGG                                648

Query_16            660  ATTCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTC  719
Unigene27967_All    658  ATTCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTC  717
Unigene75450_All    658  ATTCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTC  717
CL4065.Contig3_All  577  ATTCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTC  636
Unigene75449_All    658  ATGCTGAGTGGCATCGGAGGCTTTGTGCTGGGGCTGATCTTCCTCGGGCTGGGCCTTATC  717
Unigene57188_All    679  ATGCTGAGTGGCATCGGAGGCTTTGTGCTGGGGCTGATCTTCCTCGGGCTGGGCCTTATC  738

Query_16            720  ATCTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  776
Unigene27967_All    718  ATCTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  774
Unigene75450_All    718  ATCTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  774
CL4065.Contig3_All  637  ATCTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  693
Unigene75449_All    718  GTCCGTCACAGGAGCCAGAAAGGAC-CTCGTGGGTCTCCGCCAGCAGGGCTCCTG     771
Unigene57188_All    739  GTCCGTCACAGGAGCCAGAAAGGAC-CTCGTGGGTCTCCGCCAGCAGGGCTCCTG     792


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 65073776100


Query= HQ908094.1 Meles meles MHC class II antigen DR beta chain (Meme-DRB)
mRNA, Meme-DRB*03 allele, partial cds

Length=822
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  1397    0.0
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  1136    0.0
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  1136    0.0
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  743     0.0
Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  488     5e-137
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  411     1e-113
Unigene81593_All  2 94 minus strand MHC class II antigen, partial...  139     6e-32



Query_17            3    GGCTCCTGGATGACAGCTCTGACACTGATACTGATGGTGCTGAGCCCTCCCTTGGCTTGG  62
Unigene27967_All    1    GGCGCCTGGATGACAGCTCTGACACTGATACTGATGGTGCTGAGCCCTCCCTTGGCTTGG  60
Unigene75449_All    1    GGCGCCTGGATGACAGCTCTGACACTGATACTGATGGTGCTGAGCCCTCCCTTGGCTTGG  60

Query_17            63   GCCAGGGACACCCCACGACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAAC  122
Unigene27967_All    61   GCCAGGGACACCCCACGACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAAC  120
Unigene75450_All    103                                          GAGTGCTACTTCACCAAC  120
CL4065.Contig3_All  22                                           GAGTGCTACTTCACCAAC  39
Unigene75449_All    61   GCCAGGGACACCCCACGACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAAC  120
Unigene57188_All    124                                          GAGTGCTACTTCACCAAC  141
Unigene67843_All    124                                          GAGTGCTACTTCACCAAC  141

Query_17            123  GGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCGC  182
Unigene27967_All    121  GGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCGC  180
Unigene75450_All    121  GGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCGC  180
CL4065.Contig3_All  40   GGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCGC  99
Unigene75449_All    121  GGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCGC  180
Unigene57188_All    142  GGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCGC  201
Unigene67843_All    142  GGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCGC  201
```

```
Query_17             183  TTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  242
Unigene27967_All     181  TTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  240
Unigene75450_All     181  TACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  240
CL4065.Contig3_All   100  TACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  159
Unigene75449_All     181  TTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTGAG  240
Unigene57188_All     202  TACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  261
Unigene67843_All     202  TACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  261

Query_17             243  TACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGAGGTGGACACCGTGTGC  302
Unigene27967_All     241  TACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGCGGTGGACACATACTGC  300
Unigene75450_All     241  TACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGCGGTGGACACATACTGC  300
CL4065.Contig3_All   160  TACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGCGGTGGACACATACTGC  219
Unigene75449_All     241  TACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAGGCCGAGACAGACACGGTGTGC  300
Unigene57188_All     262  TACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAGGCCGAGACAGACACGGTGTGC  321
Unigene67843_All     262  TACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAGGCCGAGACAGACACGGTGTGC  321
Unigene81593_All     1                CAGAAGGACATCATGGAGCGGAAGAGGTCCGAGGTGGACACCGTGTGC  48

Query_17             303  AGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGTG  362
Unigene27967_All     301  AGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGTG  360
Unigene75450_All     301  AGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGTG  360
CL4065.Contig3_All   220  AGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGTG  279
Unigene75449_All     301  AGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTACAGTG  360
Unigene57188_All     322  AGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTACAGTG  381
Unigene67843_All     322  AGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTACAGTG  381
Unigene81593_All     49   AGACACAACCACGGGGTGTTTGAGAGCTTCC-GGTGCAGCGGCGAG                93

Query_17             363  ACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGTG  422
Unigene27967_All     361  ACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGTG  420
Unigene75450_All     361  ACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGTG  420
CL4065.Contig3_All   280  ACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGTG  339
Unigene75449_All     361  ACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTCGGTG  420
Unigene57188_All     382  ACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTCGGTG  441
Unigene67843_All     382  ACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTCGGTG  441

Query_17             423  AATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGAG  482
Unigene27967_All     421  AATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGAG  480
Unigene75450_All     421  AATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGAG  480
CL4065.Contig3_All   340  AATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGAG  399
Unigene75449_All     421  ACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGAGAAA  480
Unigene57188_All     442  ACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGAGAAA  501
Unigene67843_All     442  ACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGAGAAA  501

Query_17             483  TCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGACTGGACCTTCCAGACCCTGGTG  542
Unigene27967_All     481  TCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGACTGGACCTTCCAGACCCTGGTG  540
Unigene75450_All     481  TCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGACTGGACCTTCCAGACCCTGGTG  540
CL4065.Contig3_All   400  TCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGACTGGACCTTCCAGACCCTGGTG  459
Unigene75449_All     481  GCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGACTGGACCTTCCAGATCCTTGTG  540
Unigene57188_All     502  GCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGACTGGACCTTCCAGATCCTTGTG  561
Unigene67843_All     502  GCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGACTGGACCTTCCAGATCCTTGTG  561

Query_17             543  ATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCCAG  601
Unigene27967_All     541  ATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCCAG  599
Unigene75450_All     541  ATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCCAG  599
CL4065.Contig3_All   460  ATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCCAG  518
Unigene75449_All     541  ATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACCCCAG  599
Unigene57188_All     562  ATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACCCCAG  620
Unigene67843_All     562  ATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACCCCAG  620

Query_17             602  TTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAGAT  661
Unigene27967_All     600  TTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAGAT  659
Unigene75450_All     600  TTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAGAT  659
CL4065.Contig3_All   519  TTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAGAT  578
Unigene75449_All     600  CCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTCGGAATCTGCCCAGAGCAAGAT  659
Unigene57188_All     621  CCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTCGGAATCTGCCCAGAGCAAGAT  680
Unigene67843_All     621  CCTCCAGAGCCCCATCACAGTGGAGTGG                                  648

Query_17             662  TCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTCAT  721
Unigene27967_All     660  TCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTCAT  719
Unigene75450_All     660  TCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTCAT  719
CL4065.Contig3_All   579  TCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTCAT  638
Unigene75449_All     660  GCTGAGTGGCATCGGAGGCTTTGTGCTGGGGCTGATCTTCCTCGGGCTGGGCCTTATCGT  719
Unigene57188_All     681  GCTGAGTGGCATCGGAGGCTTTGTGCTGGGGCTGATCTTCCTCGGGCTGGGCCTTATCGT  740

Query_17             722  CTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  776
Unigene27967_All     720  CTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  774
Unigene75450_All     720  CTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  774
CL4065.Contig3_All   639  CTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  693
Unigene75449_All     720  CCGTCACAGGAGCCAGAAAGGAC-CTCGTGGGTCTCCGCCAGCAGGGCTCCTG     771
Unigene57188_All     741  CCGTCACAGGAGCCAGAAAGGAC-CTCGTGGGTCTCCGCCAGCAGGGCTCCTG     792
```

```
Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 65073776100


Query= HQ908093.1 Meles meles MHC class II antigen DR beta chain (Meme-DRB)
mRNA, Meme-DRB*02 allele, partial cds

Length=822
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  1408    0.0
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  1153    0.0
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  1153    0.0
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  715     0.0
Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  466     2e-130
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  388     5e-107



Query_18            3    GGCTCCTGGATGACAGCTTTGACACTGATACTGATGGTGCTGAGCCCTCCCTTGGCTTGG  62
Unigene27967_All    1    GGCGCCTGGATGACAGCTCTGACACTGATACTGATGGTGCTGAGCCCTCCCTTGGCTTGG  60
Unigene75449_All    1    GGCGCCTGGATGACAGCTCTGACACTGATACTGATGGTGCTGAGCCCTCCCTTGGCTTGG  60

Query_18            63   GCCAGGGACACCCCACGACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAAC  122
Unigene27967_All    61   GCCAGGGACACCCCACGACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAAC  120
Unigene75450_All    103                                          GAGTGCTACTTCACCAAC  120
CL4065.Contig3_All  22                                           GAGTGCTACTTCACCAAC  39
Unigene75449_All    61   GCCAGGGACACCCCACGACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAAC  120
Unigene57188_All    124                                          GAGTGCTACTTCACCAAC  141
Unigene67843_All    124                                          GAGTGCTACTTCACCAAC  141

Query_18            123  GGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCGC  182
Unigene27967_All    121  GGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCGC  180
Unigene75450_All    121  GGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCGC  180
CL4065.Contig3_All  40   GGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCGC  99
Unigene75449_All    121  GGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCGC  180
Unigene57188_All    142  GGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCGC  201
Unigene67843_All    142  GGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCGC  201

Query_18            183  TTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  242
Unigene27967_All    181  TTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  240
Unigene75450_All    181  TACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  240
CL4065.Contig3_All  100  TACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  159
Unigene75449_All    181  TTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTGAG  240
Unigene57188_All    202  TACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  261
Unigene67843_All    202  TACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCAG  261

Query_18            243  TACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGCGGTGGACACATACTGC  302
Unigene27967_All    241  TACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGCGGTGGACACATACTGC  300
Unigene75450_All    241  TACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGCGGTGGACACATACTGC  300
CL4065.Contig3_All  160  TACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGGCCGCGGTGGACACATACTGC  219
Unigene75449_All    241  TACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAGGCCGAGACAGACACGGTGTGC  300
Unigene57188_All    262  TACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAGGCCGAGACAGACACGGTGTGC  321
Unigene67843_All    262  TACTGGAACAGCCAGAAGGACATCCTGGAGAGAACGGAGGCCGAGACAGACACGGTGTGC  321

Query_18            303  AGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGTG  362
Unigene27967_All    301  AGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGTG  360
Unigene75450_All    301  AGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGTG  360
CL4065.Contig3_All  220  AGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTACAGTG  279
Unigene75449_All    301  AGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTACAGTG  360
Unigene57188_All    322  AGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTACAGTG  381
Unigene67843_All    322  AGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTACAGTG  381

Query_18            363  ACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGTG  422
Unigene27967_All    361  ACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGTG  420
Unigene75450_All    361  ACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGTG  420
CL4065.Contig3_All  280  ACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTCTGTG  339
Unigene75449_All    361  ACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTCGGTG  420
Unigene57188_All    382  ACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTCGGTG  441
Unigene67843_All    382  ACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTCGGTG  441

Query_18            423  AATGGTTTCTATCCAGGCCACATTGAAGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGAG  482
Unigene27967_All    421  AATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGAG  480
Unigene75450_All    421  AATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGAG  480
```

```
CL4065.Contig3_All  340  AATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGAGGAG  399
Unigene75449_All    421  ACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGAGAAA  480
Unigene57188_All    442  ACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGAGAAA  501
Unigene67843_All    442  ACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGAGAAA  501

Query_18            483  TCTGGGGTCGTGTCCACAGGCCTGATCCCTAATGGAGACTGGACCTTCCAGACCCTGGTG  542
Unigene27967_All    481  TCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGACTGGACCTTCCAGACCCTGGTG  540
Unigene75450_All    481  TCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGACTGGACCTTCCAGACCCTGGTG  540
CL4065.Contig3_All  400  TCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGACTGGACCTTCCAGACCCTGGTG  459
Unigene75449_All    481  GCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGACTGGACCTTCCAGATCCTTGTG  540
Unigene57188_All    502  GCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGACTGGACCTTCCAGATCCTTGTG  561
Unigene67843_All    502  GCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGACTGGACCTTCCAGATCCTTGTG  561

Query_18            543  ATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCCAG  601
Unigene27967_All    541  ATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCCAG  599
Unigene75450_All    541  ATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCCAG  599
CL4065.Contig3_All  460  ATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACCCCAG  518
Unigene75449_All    541  ATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACCCCAG  599
Unigene57188_All    562  ATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACCCCAG  620
Unigene67843_All    562  ATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACCCCAG  620

Query_18            602  TTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAGAT  661
Unigene27967_All    600  TTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAGAT  659
Unigene75450_All    600  TTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAGAT  659
CL4065.Contig3_All  519  TTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCAAGAT  578
Unigene75449_All    600  CCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTCGGAATCTGCCCAGAGCAAGAT  659
Unigene57188_All    621  CCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTCGGAATCTGCCCAGAGCAAGAT  680
Unigene67843_All    621  CCTCCAGAGCCCCATCACAGTGGAGTGG                                  648

Query_18            662  TCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTCAT  721
Unigene27967_All    660  TCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTCAT  719
Unigene75450_All    660  TCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTCAT  719
CL4065.Contig3_All  579  TCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGTTCAT  638
Unigene75449_All    660  GCTGAGTGGCATCGGAGGCTTTGTGCTGGGGCTGATCTTCCTCGGGCTGGGCCTTATCGT  719
Unigene57188_All    681  GCTGAGTGGCATCGGAGGCTTTGTGCTGGGGCTGATCTTCCTCGGGCTGGGCCTTATCGT  740

Query_18            722  CTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  776
Unigene27967_All    720  CTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  774
Unigene75450_All    720  CTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  774
CL4065.Contig3_All  639  CTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  693
Unigene75449_All    720  CCGTCACAGGAGCCAGAAAGGAC-CTCGTGGGTCTCCGCCAGCAGGGCTCCTG     771
Unigene57188_All    741  CCGTCACAGGAGCCAGAAAGGAC-CTCGTGGGTCTCCGCCAGCAGGGCTCCTG     792


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 65073776100


Query= HQ908092.1 Meles meles MHC class II antigen DR beta chain (Meme-DRB)
mRNA, Meme-DRB*01 allele, partial cds

Length=822
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  1280    0.0
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  1051    0.0
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  1048    0.0
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  632     2e-180
Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  409     4e-113
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  337     2e-91
CL13896.Contig1_All  3 98 MHC class II antigen DR beta chain, par...  178     1e-43



Query_19            3    GGCTCCTGGATGACAGCTTTGACCGTGATACTGATGGTGCTGAGCCCTCCCA-TGGCTTG  61
Unigene27967_All    1    GGCGCCTGGATGACAGCTCTGACACTGATACTGATGGTGCTGAGCCCTCCCT-TGGCTTG  59
Unigene75450_All    35                                      TGGTGCTGAG-CGTCCCAGTGGCTGA  59
Unigene75449_All    1    GGCGCCTGGATGACAGCTCTGACACTGATACTGATGGTGCTGAGCCCTCCCT-TGGCTTG  59
Unigene57188_All    56                                      TGGTGCTGAG-CGTCCCAGTGGCTGA  80
Unigene67843_All    56                                      TGGTGCTGAG-CGTCCCAGTGGCTGA  80

Query_19            62   GGCCAGGGACACCCCACCACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAA  121
Unigene27967_All    60   GGCCAGGGACACCCCACGACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAA  119
Unigene75450_All    60   GGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAGGGCGAGTGCTACTTCACCAA  119
```

```
CL4065.Contig3_All   22                                                GAGTGCTACTTCACCAA  38
Unigene75449_All     60   GGCCAGGGACACCCCACGACATTTCCTGTTCCTGACGACGTCGGAGTGCCACTTCACCAA  119
Unigene57188_All     81   GGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAGGGCGAGTGCTACTTCACCAA  140
Unigene67843_All     81   GGGCAGAGACTCTCCAAAGGATTTCGTGTTCCAGTTTAAGGGCGAGTGCTACTTCACCAA  140

Query_19             122  CGGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCG  181
Unigene27967_All     120  CGGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCG  179
Unigene75450_All     120  CGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCG  179
CL4065.Contig3_All   39   CGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCG  98
Unigene75449_All     120  CGGCACGGAGCGGGTGCGGTTCCTGGATAGGTATTTCTATAACGGCGAGGAGTACGTGCG  179
Unigene57188_All     141  CGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCG  200
Unigene67843_All     141  CGGGACGGAGCGGGTGCGGAGCGTGAACAGATACATCTATAACCGGGAGGAGTTCGTGCG  200

Query_19             182  CTTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCCATCGCTCA  241
Unigene27967_All     180  CTTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCA  239
Unigene75450_All     180  CTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCA  239
CL4065.Contig3_All   99   CTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCA  158
Unigene75449_All     180  CTTCGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTGA  239
Unigene57188_All     201  CTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCA  260
Unigene67843_All     201  CTACGACAGCGACGTGGGGGAGTACCGGCCGGTGACCGAGCTGGGGCGGCCGGACGCTCA  260

Query_19             242  GGGCTGGAACAGCCAGAAGGACATCATGGAGCAGAAGCGCG-CC-A-ACGTGGACACATA  298
Unigene27967_All     240  GTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGG-CC-G-CGGTGGACACATA  296
Unigene75450_All     240  GTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGG-CC-G-CGGTGGACACATA  296
CL4065.Contig3_All   159  GTACTGGAACAGCCAGAAGGACATCATGGAGCGGAGGCGGG-CC-G-CGGTGGACACATA  215
Unigene75449_All     240  GTACTGGAACAGCCAGAAGGACATCCTGGAG-AGAACGGAGGCCGAGACA-G-ACACGGT  296
Unigene57188_All     261  GTACTGGAACAGCCAGAAGGACATCCTGGAG-AGAACGGAGGCCGAGACA-G-ACACGGT  317
Unigene67843_All     261  GTACTGGAACAGCCAGAAGGACATCCTGGAG-AGAACGGAGGCCGAGACA-G-ACACGGT  317
CL13896.Contig1_All  1              AGCCAGAAGGACATCATGGAGCAGAAGCGCG-CC-A-ACGTGGACACATA  47

Query_19             299  CTGCAGGCACAACTACGGGGTGGGTGAGAGCTTCACGGTGCAGCGGCGAGTGGAGCCTAC  358
Unigene27967_All     297  CTGCAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTAC  356
Unigene75450_All     297  CTGCAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTAC  356
CL4065.Contig3_All   216  CTGCAGACACAACTACGGGGTGGTTGAGAGCTTCCTGGTGCAGCGGCGAGTGGAGCCTAC  275
Unigene75449_All     297  GTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTAC  356
Unigene57188_All     318  GTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTAC  377
Unigene67843_All     318  GTGCAGACACAACTACCTGACTGATGAGAGCTTCACGGTGCAGAGGCGAGTGGAACCTAC  377
CL13896.Contig1_All  48   CTGCAGGCACAACTACGGGGTGGGTGAGAGCTTCACGGTGCAGCGGCGA             96

Query_19             359  AGTGACTGTGTATCCCGCGAAGAACCAGCCCTTGCAGCACCACAGCCTCCTGGTCTGCTC  418
Unigene27967_All     357  AGTGACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTC  416
Unigene75450_All     357  AGTGACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTC  416
CL4065.Contig3_All   276  AGTGACTGTGTATCCCGCGAAGAACCAGCCCCTGCAGCACCACAACCTCCTGGTCTGCTC  335
Unigene75449_All     357  AGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTC  416
Unigene57188_All     378  AGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTC  437
Unigene67843_All     378  AGTGACCATCTCCCCATCCAGGACGGAGGTTCTGAACCACCACAACATGCTGGTCTGCTC  437

Query_19             419  TGTGAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTACCAGAATGGCCAGGAAGA  478
Unigene27967_All     417  TGTGAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGA  476
Unigene75450_All     417  TGTGAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGA  476
CL4065.Contig3_All   336  TGTGAATGGTTTCTATCCAGGCCACATTGAGGTCAGGTGGTTCCGGAATGGCCAGGAAGA  395
Unigene75449_All     417  GGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGA  476
Unigene57188_All     438  GGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGA  497
Unigene67843_All     438  GGTGACAGATTTCTACCCAGGCCAGATCAAAGTTCGGTGGTTTCGGAATGACCAGGAGGA  497

Query_19             479  GGAGTCTGGGGTCGTGTCCACAGGCCTGATCCATAATGGAGACTGGACCTTCCAGACCCT  538
Unigene27967_All     477  GGAGTCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGACTGGACCTTCCAGACCCT  536
Unigene75450_All     477  GGAGTCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGACTGGACCTTCCAGACCCT  536
CL4065.Contig3_All   396  GGAGTCTGGGGTCGTGTCCACAGGCCTGATCCGTAATGGAGACTGGACCTTCCAGACCCT  455
Unigene75449_All     477  GAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGACTGGACCTTCCAGATCCT  536
Unigene57188_All     498  GAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGACTGGACCTTCCAGATCCT  557
Unigene67843_All     498  GAAAGCTGGTGTGGTGTCCACTCCACTTATTAGGAATGGGGACTGGACCTTCCAGATCCT  557

Query_19             539  GGTGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACC  597
Unigene27967_All     537  GGTGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACC  595
Unigene75450_All     537  GGTGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACC  595
CL4065.Contig3_All   456  GGTGATGCTGGAGACTGT-TCCTCAGAGTGGAGAGGTCTACACCTGCCAAGTGGAGCACC  514
Unigene75449_All     537  TGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACC  595
Unigene57188_All     558  TGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACC  616
Unigene67843_All     558  TGTGATGCTGGAAA-TGACTCCCCAGCGAGGAGATGTCTACACCTGCCATGTGGAGCACC  616

Query_19             598  CCAGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCA  657
Unigene27967_All     596  CCAGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCA  655
Unigene75450_All     596  CCAGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCA  655
CL4065.Contig3_All   515  CCAGTTTGACGAGCCCTGTCACCGTGGAATGGAGGGCACAGTCTGGGTCTGCACAGAGCA  574
Unigene75449_All     596  CCAGCCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTCGGAATCTGCCCAGAGCA  655
Unigene57188_All     617  CCAGCCTCCAGAGCCCCATCACAGTGGAGTGGCGGGCACAGTCGGAATCTGCCCAGAGCA  676
Unigene67843_All     617  CCAGCCTCCAGAGCCCCATCACAGTGGAGTGG                              648

Query_19             658  AGATTCTTAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGT  717
Unigene27967_All     656  AGATTCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGT  715
```

```
Unigene75450_All     656  AGATTCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGT  715
CL4065.Contig3_All   575  AGATTCTGAGTGGAACTGGAGGCTTTGTTCTGGGTCTGCTCTTCCTCGTGGTGGGGCTGT  634
Unigene75449_All     656  AGATGCTGAGTGGCATCGGAGGCTTTGTGCTGGGGCTGATCTTCCTCGGGCTGGGCCTTA  715
Unigene57188_All     677  AGATGCTGAGTGGCATCGGAGGCTTTGTGCTGGGGCTGATCTTCCTCGGGCTGGGCCTTA  736

Query_19             718  TCATCTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  776
Unigene27967_All     716  TCATCTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  774
Unigene75450_All     716  TCATCTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  774
CL4065.Contig3_All   635  TCATCTACTTCAGGAATCAGAAGGGACACTC-TGGACTTCAGCCAACAGGTCTCCTGAGC  693
Unigene75449_All     716  TCGTCCGTCACAGGAGCCAGAAAGGAC-CTCGTGGGTCTCCGCCAGCAGGGCTCCTG     771
Unigene57188_All     737  TCGTCCGTCACAGGAGCCAGAAAGGAC-CTCGTGGGTCTCCGCCAGCAGGGCTCCTG     792


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 65073776100


Query= Y11647.2 M.meles mRNA for interferon gamma, partial

Length=501
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene54563_All  2 205 interferon gamma [Mustela putorius furo]      377     7e-104


Query_20          184  GAGAGTGACAAAACAATCATTCAAAGCCAAATTGTCTCCTTCTACTTGAAACTGTTTGAA  243
Unigene54563_All  1    GAGAGTGACAAAACAATCATTCAAAGCCAAATTGTCTCCTTCTACTTGAAACTGTTTGAA  60

Query_20          244  AACTTTAAAGATAACCAGATCATTCAAAGGAGCATGGATACCATCAAGGAAGACATGCTT  303
Unigene54563_All  61   AACTTTAAAGATAACCAGATCATTCAAAGGAGCATGGATACCATCAAGGAAGACATGCTT  120

Query_20          304  GTCAGGTTCTTCAATAGCAGCAGCAGTAAGCGGGAGGACTTTCTTAAGCTGATTCGAATT  363
Unigene54563_All  121  GTCAGGTTCTTCAATAGCAGCAGCAGTAAGCGGGAGGACTTTCTTAAGCTGATTCGAATT  180

Query_20          364  CCCGTGAATGATCTGCAGGTCCAG  387
Unigene54563_All  181  CCCGTGAATGATCTGCAGGTCCAG  204


Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 38974106976


  Database: badgernt
    Posted date:  Feb 23, 2017  11:56 AM
  Number of letters in database: 85,063,401
  Number of sequences in database:  127,401



Matrix: blastn matrix 1 -2
Gap Penalties: Existence: 0, Extension: 2.5
```

8.4 BLASTP output (tabular)

Query Seq-id	Subject seq-id	Identity %	Alignment length	E-vlue	Query seq. length	Subject seq. length
AFR54067.1	CL1150.Contig4_All	96.11	180	7.00E-128	180	234
AFR54067.1	CL1150.Contig16_All	92.22	180	1.00E-118	180	324
AFR54067.1	CL154.Contig12_All	38.12	181	4.00E-31	180	306

AFR54067.1	CL154.Contig11_All	38.12	181	4.00E-31	180	306
AFR54067.1	CL154.Contig24_All	38.12	181	7.00E-31	180	358
AFR54067.1	CL154.Contig19_All	38.12	181	7.00E-31	180	358
AFR54067.1	CL154.Contig27_All	38.12	181	8.00E-31	180	345
AFR54067.1	CL154.Contig23_All	38.12	181	8.00E-31	180	345
AFR54067.1	CL154.Contig2_All	38.12	181	8.00E-31	180	345
AFR54067.1	CL154.Contig1_All	38.12	181	8.00E-31	180	345
AFR54066.1	CL1150.Contig4_All	85	180	6.00E-111	180	234
AFR54066.1	CL1150.Contig16_All	84.44	180	5.00E-108	180	324
AFR54066.1	CL154.Contig12_All	38.67	181	2.00E-32	180	306
AFR54066.1	CL154.Contig11_All	38.67	181	2.00E-32	180	306
AFR54066.1	CL154.Contig27_All	38.67	181	4.00E-32	180	345
AFR54066.1	CL154.Contig23_All	38.67	181	4.00E-32	180	345
AFR54066.1	CL154.Contig2_All	38.67	181	4.00E-32	180	345
AFR54066.1	CL154.Contig1_All	38.67	181	4.00E-32	180	345
AFR54066.1	CL154.Contig24_All	38.67	181	4.00E-32	180	358
AFR54066.1	CL154.Contig19_All	38.67	181	4.00E-32	180	358
AFR54065.1	CL1150.Contig4_All	84.44	180	9.00E-110	180	234
AFR54065.1	CL1150.Contig16_All	83.89	180	7.00E-107	180	324
AFR54065.1	CL154.Contig12_All	39.78	181	6.00E-34	180	306
AFR54065.1	CL154.Contig11_All	39.78	181	6.00E-34	180	306
AFR54065.1	CL154.Contig27_All	39.78	181	1.00E-33	180	345
AFR54065.1	CL154.Contig23_All	39.78	181	1.00E-33	180	345
AFR54065.1	CL154.Contig2_All	39.78	181	1.00E-33	180	345
AFR54065.1	CL154.Contig1_All	39.78	181	1.00E-33	180	345
AFR54065.1	CL154.Contig24_All	39.78	181	1.00E-33	180	358
AFR54065.1	CL154.Contig19_All	39.78	181	1.00E-33	180	358
AFR54064.1	CL1150.Contig16_All	87.22	180	4.00E-111	180	324
AFR54064.1	CL1150.Contig4_All	83.33	180	1.00E-106	180	234
AFR54064.1	CL154.Contig12_All	38.12	181	2.00E-29	180	306
AFR54064.1	CL154.Contig11_All	38.12	181	2.00E-29	180	306
AFR54064.1	CL154.Contig24_All	38.12	181	3.00E-29	180	358
AFR54064.1	CL154.Contig19_All	38.12	181	3.00E-29	180	358
AFR54064.1	CL154.Contig27_All	38.12	181	3.00E-29	180	345
AFR54064.1	CL154.Contig23_All	38.12	181	3.00E-29	180	345
AFR54064.1	CL154.Contig2_All	38.12	181	3.00E-29	180	345
AFR54064.1	CL154.Contig1_All	38.12	181	3.00E-29	180	345
AFR54063.1	CL1150.Contig4_All	94.44	180	7.00E-125	180	234
AFR54063.1	CL1150.Contig16_All	94.44	180	1.00E-121	180	324
AFR54063.1	CL154.Contig12_All	39.23	181	2.00E-34	180	306
AFR54063.1	CL154.Contig11_All	39.23	181	2.00E-34	180	306

AFR54063.1	CL154.Contig27_All	39.23	181	4.00E-34	180	345
AFR54063.1	CL154.Contig23_All	39.23	181	4.00E-34	180	345
AFR54063.1	CL154.Contig2_All	39.23	181	4.00E-34	180	345
AFR54063.1	CL154.Contig1_All	39.23	181	4.00E-34	180	345
AFR54063.1	CL154.Contig24_All	39.23	181	4.00E-34	180	358
AFR54063.1	CL154.Contig19_All	39.23	181	4.00E-34	180	358
AFR54062.1	CL1150.Contig4_All	88.89	90	3.00E-53	90	234
AFR54062.1	CL1150.Contig16_All	86.67	90	7.00E-50	90	324
AFR54062.1	CL154.Contig16_All	38.64	88	2.00E-15	90	88
AFR54062.1	CL154.Contig15_All	38.64	88	2.00E-15	90	88
AFR54062.1	CL154.Contig14_All	38.64	88	2.00E-15	90	88
AFR54062.1	CL154.Contig13_All	38.64	88	2.00E-15	90	88
AFR54062.1	CL154.Contig33_All	38.64	88	4.00E-15	90	127
AFR54062.1	CL154.Contig21_All	38.64	88	4.00E-15	90	127
AFR54062.1	CL154.Contig20_All	38.64	88	4.00E-15	90	127
AFR54062.1	CL154.Contig17_All	38.64	88	4.00E-15	90	127
AFR54061.1	CL1150.Contig4_All	74.44	90	1.00E-40	90	234
AFR54061.1	CL1150.Contig16_All	72.22	90	6.00E-38	90	324
AFR54061.1	CL154.Contig16_All	37.5	88	4.00E-13	90	88
AFR54061.1	CL154.Contig15_All	37.5	88	4.00E-13	90	88
AFR54061.1	CL154.Contig14_All	37.5	88	4.00E-13	90	88
AFR54061.1	CL154.Contig13_All	37.5	88	4.00E-13	90	88
AFR54061.1	CL154.Contig33_All	37.5	88	8.00E-13	90	127
AFR54061.1	CL154.Contig21_All	37.5	88	8.00E-13	90	127
AFR54061.1	CL154.Contig20_All	37.5	88	8.00E-13	90	127
AFR54061.1	CL154.Contig17_All	37.5	88	8.00E-13	90	127
AFR54060.1	CL1150.Contig16_All	88.58	324	0	325	324
AFR54060.1	CL1150.Contig4_All	97.01	234	1.00E-169	325	234
AFR54060.1	CL1150.Contig15_All	74.86	179	3.00E-83	325	150
AFR54060.1	CL1150.Contig17_All	79.29	140	4.00E-73	325	140
AFR54060.1	CL154.Contig12_All	42.31	312	6.00E-73	325	306
AFR54060.1	CL154.Contig11_All	42.31	312	6.00E-73	325	306
AFR54060.1	CL1150.Contig13_All	77.46	142	1.00E-72	325	142
AFR54060.1	CL154.Contig27_All	42.31	312	3.00E-72	325	345
AFR54060.1	CL154.Contig23_All	42.31	312	3.00E-72	325	345
AFR54060.1	CL154.Contig2_All	42.31	312	3.00E-72	325	345
AFR54059.1	CL1150.Contig16_All	85.19	324	0	325	324
AFR54059.1	CL1150.Contig4_All	87.61	234	1.00E-150	325	234
AFR54059.1	CL1150.Contig15_All	75.42	179	3.00E-86	325	150
AFR54059.1	CL1150.Contig17_All	81.43	140	6.00E-75	325	140
AFR54059.1	CL154.Contig12_All	42.95	312	1.00E-74	325	306

AFR54059.1	CL154.Contig11_All	42.95	312	1.00E-74	325	306
AFR54059.1	CL154.Contig27_All	42.95	312	6.00E-74	325	345
AFR54059.1	CL154.Contig23_All	42.95	312	6.00E-74	325	345
AFR54059.1	CL154.Contig2_All	42.95	312	6.00E-74	325	345
AFR54059.1	CL154.Contig1_All	42.95	312	6.00E-74	325	345
AFR54058.1	CL1150.Contig16_All	84.88	324	0	325	324
AFR54058.1	CL1150.Contig4_All	87.18	234	3.00E-149	325	234
AFR54058.1	CL1150.Contig15_All	75.42	179	2.00E-86	325	150
AFR54058.1	CL154.Contig12_All	43.27	312	3.00E-76	325	306
AFR54058.1	CL154.Contig11_All	43.27	312	3.00E-76	325	306
AFR54058.1	CL154.Contig24_All	43.27	312	2.00E-75	325	358
AFR54058.1	CL154.Contig19_All	43.27	312	2.00E-75	325	358
AFR54058.1	CL154.Contig27_All	43.27	312	2.00E-75	325	345
AFR54058.1	CL154.Contig23_All	43.27	312	2.00E-75	325	345
AFR54058.1	CL154.Contig2_All	43.27	312	2.00E-75	325	345
AFR54057.1	CL1150.Contig16_All	86.73	324	0	325	324
AFR54057.1	CL1150.Contig4_All	87.18	234	2.00E-148	325	234
AFR54057.1	CL1150.Contig15_All	73.18	179	9.00E-83	325	150
AFR54057.1	CL1150.Contig17_All	81.43	140	7.00E-75	325	140
AFR54057.1	CL1150.Contig13_All	79.58	142	2.00E-74	325	142
AFR54057.1	CL1150.Contig18_All	80.58	139	1.00E-73	325	139
AFR54057.1	CL154.Contig12_All	41.99	312	1.00E-70	325	306
AFR54057.1	CL154.Contig11_All	41.99	312	1.00E-70	325	306
AFR54057.1	CL154.Contig27_All	41.99	312	5.00E-70	325	345
AFR54057.1	CL154.Contig23_All	41.99	312	5.00E-70	325	345
AFR54056.1	CL1150.Contig16_All	89.81	324	0	325	324
AFR54056.1	CL1150.Contig4_All	95.73	234	3.00E-166	325	234
AFR54056.1	CL1150.Contig15_All	75.98	179	8.00E-85	325	150
AFR54056.1	CL154.Contig12_All	42.95	312	3.00E-76	325	306
AFR54056.1	CL154.Contig11_All	42.95	312	3.00E-76	325	306
AFR54056.1	CL154.Contig27_All	42.95	312	1.00E-75	325	345
AFR54056.1	CL154.Contig23_All	42.95	312	1.00E-75	325	345
AFR54056.1	CL154.Contig2_All	42.95	312	1.00E-75	325	345
AFR54056.1	CL154.Contig1_All	42.95	312	1.00E-75	325	345
AFR54056.1	CL154.Contig24_All	42.95	312	2.00E-75	325	358
AFR54055.1	CL1150.Contig16_All	87.29	181	5.00E-113	181	324
AFR54055.1	CL1150.Contig4_All	83.43	181	2.00E-108	181	234
AFR54055.1	CL154.Contig12_All	38.46	182	5.00E-30	181	306
AFR54055.1	CL154.Contig11_All	38.46	182	5.00E-30	181	306
AFR54055.1	CL154.Contig24_All	38.46	182	7.00E-30	181	358
AFR54055.1	CL154.Contig19_All	38.46	182	7.00E-30	181	358

AFR54055.1	CL154.Contig27_All	38.46	182	8.00E-30	181	345
AFR54055.1	CL154.Contig23_All	38.46	182	8.00E-30	181	345
AFR54055.1	CL154.Contig2_All	38.46	182	8.00E-30	181	345
AFR54055.1	CL154.Contig1_All	38.46	182	8.00E-30	181	345
AFR54054.1	CL1150.Contig4_All	85.08	181	6.00E-111	181	234
AFR54054.1	CL1150.Contig16_All	85.08	181	2.00E-108	181	324
AFR54054.1	CL154.Contig12_All	38.46	182	3.00E-30	181	306
AFR54054.1	CL154.Contig11_All	38.46	182	3.00E-30	181	306
AFR54054.1	CL154.Contig24_All	38.46	182	5.00E-30	181	358
AFR54054.1	CL154.Contig19_All	38.46	182	5.00E-30	181	358
AFR54054.1	CL154.Contig27_All	38.46	182	6.00E-30	181	345
AFR54054.1	CL154.Contig23_All	38.46	182	6.00E-30	181	345
AFR54054.1	CL154.Contig2_All	38.46	182	6.00E-30	181	345
AFR54054.1	CL154.Contig1_All	38.46	182	6.00E-30	181	345
AET36883.1	Unigene42913_All	100	81	8.00E-54	81	201
AET36883.1	CL2981.Contig3_All	100	81	3.00E-53	81	254
AET36883.1	CL2981.Contig1_All	100	81	3.00E-53	81	254
AET36883.1	CL10050.Contig3_All	56.1	82	1.00E-28	81	255
AET36883.1	CL10050.Contig2_All	56.1	82	9.00E-25	81	194
AET36883.1	CL10050.Contig1_All	56.1	82	9.00E-25	81	194
AET36883.1	CL12484.Contig1_All	52.05	73	3.00E-22	81	109
AET36883.1	CL12484.Contig3_All	52.05	73	2.00E-21	81	193
AET36881.1	Unigene67843_All	75.28	89	2.00E-43	89	217
AET36881.1	Unigene57188_All	75.28	89	3.00E-43	89	265
AET36881.1	Unigene75449_All	69.32	88	4.00E-38	89	258
AET36881.1	Unigene75450_All	69.66	89	8.00E-38	89	258
AET36881.1	CL4065.Contig3_All	69.32	88	7.00E-37	89	231
AET36881.1	CL6815.Contig5_All	58.23	79	7.00E-35	89	273
AET36881.1	Unigene27967_All	64.77	88	3.00E-33	89	258
AET36881.1	CL6815.Contig6_All	53.93	89	9.00E-31	89	215
AET36881.1	CL6815.Contig1_All	53.93	89	9.00E-31	89	215
AET36881.1	CL6815.Contig3_All	53.93	89	9.00E-31	89	252
AET36880.1	Unigene27967_All	77.38	84	5.00E-41	89	258
AET36880.1	Unigene75449_All	72.62	84	2.00E-37	89	258
AET36880.1	CL4065.Contig3_All	69.88	83	7.00E-36	89	231
AET36880.1	Unigene75450_All	69.88	83	8.00E-36	89	258
AET36880.1	Unigene67843_All	66.27	83	6.00E-33	89	217
AET36880.1	Unigene57188_All	66.27	83	9.00E-33	89	265
AET36880.1	CL6815.Contig4_All	44.58	83	8.00E-22	89	218
AET36880.1	CL6815.Contig6_All	44.58	83	8.00E-22	89	215
AET36880.1	CL6815.Contig1_All	44.58	83	8.00E-22	89	215

AET36880.1	CL6815.Contig3_All	44.58	83	1.00E-21	89	252
AET36875.1	CL2981.Contig3_All	99.13	230	7.00E-171	230	254
AET36875.1	CL2981.Contig1_All	99.13	230	7.00E-171	230	254
AET36875.1	Unigene42913_All	99.5	201	1.00E-149	230	201
AET36875.1	CL10050.Contig3_All	60.53	228	7.00E-94	230	255
AET36875.1	CL10050.Contig2_All	60.21	191	9.00E-78	230	194
AET36875.1	CL10050.Contig1_All	60.21	191	9.00E-78	230	194
AET36875.1	CL12484.Contig3_All	55.5	191	2.00E-70	230	193
AET36875.1	CL12484.Contig2_All	65.06	83	6.00E-32	230	83
AET36875.1	CL12484.Contig1_All	48.15	108	3.00E-28	230	109
AET36875.1	CL5174.Contig1_All	63.38	71	2.00E-23	230	79
AET36874.1	CL10050.Contig3_All	99.51	205	7.00E-154	205	255
AET36874.1	CL10050.Contig2_All	100	194	3.00E-146	205	194
AET36874.1	CL10050.Contig1_All	100	194	3.00E-146	205	194
AET36874.1	CL2981.Contig3_All	60.2	201	5.00E-81	205	254
AET36874.1	CL2981.Contig1_All	60.2	201	5.00E-81	205	254
AET36874.1	Unigene42913_All	60.21	191	6.00E-78	205	201
AET36874.1	CL12484.Contig3_All	59.44	180	2.00E-74	205	193
AET36874.1	CL12484.Contig2_All	66.67	81	2.00E-34	205	83
AET36874.1	CL12484.Contig1_All	54.64	97	8.00E-30	205	109
AET36874.1	CL5174.Contig1_All	60.76	79	1.00E-25	205	79
AET36873.1	CL10050.Contig3_All	86.83	205	3.00E-133	205	255
AET36873.1	CL10050.Contig2_All	86.6	194	2.00E-125	205	194
AET36873.1	CL10050.Contig1_All	86.6	194	2.00E-125	205	194
AET36873.1	CL2981.Contig3_All	58.21	201	8.00E-76	205	254
AET36873.1	CL2981.Contig1_All	58.21	201	8.00E-76	205	254
AET36873.1	Unigene42913_All	58.12	191	9.00E-73	205	201
AET36873.1	CL12484.Contig3_All	56.45	186	3.00E-69	205	193
AET36873.1	CL12484.Contig2_All	64.2	81	1.00E-31	205	83
AET36873.1	CL12484.Contig1_All	51.96	102	1.00E-27	205	109
AET36873.1	CL5174.Contig1_All	63.38	71	3.00E-24	205	79
AET36872.1	Unigene57188_All	99.55	224	4.00E-170	224	265
AET36872.1	Unigene67843_All	99.54	217	2.00E-165	224	217
AET36872.1	Unigene75449_All	87.56	217	6.00E-142	224	258
AET36872.1	Unigene75450_All	82.95	217	4.00E-134	224	258
AET36872.1	CL4065.Contig3_All	80.53	190	2.00E-112	224	231
AET36872.1	Unigene27967_All	70.51	217	1.00E-111	224	258
AET36872.1	CL6815.Contig3_All	57.6	217	8.00E-91	224	252
AET36872.1	CL6815.Contig2_All	57.6	217	1.00E-90	224	265
AET36872.1	CL6815.Contig5_All	57.6	217	2.00E-90	224	273
AET36872.1	CL6815.Contig6_All	57.01	214	7.00E-89	224	215

AET36871.1	Unigene27967_All	92.64	258	2.00E-164	258	258
AET36871.1	Unigene75450_All	89.53	258	4.00E-158	258	258
AET36871.1	CL4065.Contig3_All	94.37	231	8.00E-148	258	231
AET36871.1	Unigene75449_All	74.71	257	1.00E-133	258	258
AET36871.1	Unigene57188_All	71.98	257	4.00E-127	258	265
AET36871.1	Unigene67843_All	73.33	210	8.00E-112	258	217
AET36871.1	CL6815.Contig2_All	53.36	253	1.00E-95	258	265
AET36871.1	CL6815.Contig5_All	53.36	253	2.00E-95	258	273
AET36871.1	CL6815.Contig3_All	54.69	245	5.00E-95	258	252
AET36871.1	CL6815.Contig6_All	55.77	208	4.00E-81	258	215
AET36870.1	Unigene27967_All	98.84	258	0	258	258
AET36870.1	Unigene75450_All	88.76	258	9.00E-158	258	258
AET36870.1	Unigene75449_All	80.16	257	8.00E-148	258	258
AET36870.1	CL4065.Contig3_All	93.51	231	2.00E-147	258	231
AET36870.1	Unigene57188_All	70.43	257	3.00E-125	258	265
AET36870.1	Unigene67843_All	71.43	210	2.00E-109	258	217
AET36870.1	CL6815.Contig2_All	52.57	253	5.00E-95	258	265
AET36870.1	CL6815.Contig5_All	52.57	253	1.00E-94	258	273
AET36870.1	CL6815.Contig3_All	53.88	245	2.00E-94	258	252
AET36870.1	CL6815.Contig6_All	54.81	208	2.00E-80	258	215
AET36869.1	Unigene27967_All	99.22	258	0	258	258
AET36869.1	Unigene75450_All	89.15	258	7.00E-159	258	258
AET36869.1	CL4065.Contig3_All	93.94	231	2.00E-148	258	231
AET36869.1	Unigene75449_All	78.99	257	6.00E-145	258	258
AET36869.1	Unigene57188_All	69.26	257	2.00E-122	258	265
AET36869.1	Unigene67843_All	70	210	1.00E-106	258	217
AET36869.1	CL6815.Contig2_All	52.17	253	5.00E-94	258	265
AET36869.1	CL6815.Contig5_All	52.17	253	8.00E-94	258	273
AET36869.1	CL6815.Contig3_All	53.47	245	2.00E-93	258	252
AET36869.1	CL6815.Contig6_All	54.33	208	2.00E-79	258	215
AET36868.1	Unigene27967_All	94.19	258	2.00E-170	258	258
AET36868.1	Unigene75450_All	85.27	258	1.00E-149	258	258
AET36868.1	CL4065.Contig3_All	89.61	231	2.00E-139	258	231
AET36868.1	Unigene75449_All	75.88	257	2.00E-139	258	258
AET36868.1	Unigene57188_All	67.32	257	2.00E-118	258	265
AET36868.1	Unigene67843_All	67.62	210	8.00E-103	258	217
AET36868.1	CL6815.Contig2_All	52.17	253	1.00E-94	258	265
AET36868.1	CL6815.Contig5_All	52.17	253	3.00E-94	258	273
AET36868.1	CL6815.Contig3_All	53.47	245	6.00E-94	258	252
AET36868.1	CL6815.Contig6_All	54.33	208	4.00E-80	258	215

8.5 BLASTP output (pairwise)

```
BLASTP 2.2.31+


Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro A.
Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J.
Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of
protein database search programs", Nucleic Acids Res. 25:3389-3402.


Reference for composition-based statistics: Alejandro A. Schaffer,
L. Aravind, Thomas L. Madden, Sergei Shavirin, John L. Spouge, Yuri
I. Wolf, Eugene V. Koonin, and Stephen F. Altschul (2001),
"Improving the accuracy of PSI-BLAST protein database searches with
composition-based statistics and other refinements", Nucleic Acids
Res. 29:2994-3005.



Database: BLAST Database
           127,401 sequences; 28,354,467 total letters



Query= AFR54067.1 MHC class I antigen, partial [Meles meles]

Length=180
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  377     7e-128
CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  359     1e-118
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  124     4e-31
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  124     4e-31
CL154.Contig24_All  126 1199 class I histocompatibility antigen, ...  124     7e-31
CL154.Contig19_All  122 1195 class I histocompatibility antigen, ...  124     7e-31
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  124     8e-31
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  124     8e-31
CL154.Contig2_All  362 1396 class I histocompatibility antigen, G...  124     8e-31
CL154.Contig1_All  294 1328 minus strand class I histocompatibili...  124     8e-31



Query_1               1    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  60
CL1150.Contig4_All    3    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  62
CL1150.Contig16_All   6    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  65
CL154.Contig12_All    2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  59
CL154.Contig11_All    2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  59
CL154.Contig24_All    54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE
111
CL154.Contig19_All    54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE
111
CL154.Contig27_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig23_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig2_All     41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig1_All     41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98

Query_1               61   QQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYSQHSYDGA
120
CL1150.Contig4_All    63   QQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQFAYDGA
122
CL1150.Contig16_All   66   RQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQFAYDGA
125
CL154.Contig12_All    60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
117
CL154.Contig11_All    60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
117
```

```
CL154.Contig24_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
169
CL154.Contig19_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
169
CL154.Contig27_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig23_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig2_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig1_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156

Query_1              121  DYIALNEDLRSWTAADTAAQITQRKWE-DAGEAERWRNYVEGTCVEWLGRYLENGKESLL
179
CL1150.Contig4_All   123  DYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESLL
181
CL1150.Contig16_All  126  DYIALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESLL
184
CL154.Contig12_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
177
CL154.Contig11_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
177
CL154.Contig24_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
229
CL154.Contig19_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
229
CL154.Contig27_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig23_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig2_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig1_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216

Query_1              180  R  180
CL1150.Contig4_All   182  R  182
CL1150.Contig16_All  185  R  185
CL154.Contig12_All   178  R  178
CL154.Contig11_All   178  R  178
CL154.Contig24_All   230  R  230
CL154.Contig19_All   230  R  230
CL154.Contig27_All   217  R  217
CL154.Contig23_All   217  R  217
CL154.Contig2_All    217  R  217
CL154.Contig1_All    217  R  217


Lambda      K        H        a         alpha
   0.335    0.191    0.791    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 2743481178


Query= AFR54066.1 MHC class I antigen, partial [Meles meles]

Length=180
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  334     6e-111
CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  332     5e-108
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  127     2e-32
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  127     2e-32
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  127     4e-32
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  127     4e-32
```

```
CL154.Contig2_All  362 1396 class I histocompatibility antigen, G...  127     4e-32
CL154.Contig1_All  294 1328 minus strand class I histocompatibili...  127     4e-32
CL154.Contig24_All  126 1199 class I histocompatibility antigen, ...  127     4e-32
CL154.Contig19_All  122 1195 class I histocompatibility antigen, ...  127     4e-32


Query_2              1    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWVEQEGPEYWD  60
CL1150.Contig4_All   3    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  62
CL1150.Contig16_All  6    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  65
CL154.Contig12_All   2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  59
CL154.Contig11_All   2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  59
CL154.Contig27_All   41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig23_All   41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig2_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig1_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig24_All   54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE
111
CL154.Contig19_All   54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE
111

Query_2              61   RQTRNLKDAAHAFRVNLNTLRDYYNQSAAGSHTIQRMYGCDMGPDGRLLRGYSQVAYDGA
120
CL1150.Contig4_All   63   QQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQFAYDGA
122
CL1150.Contig16_All  66   RQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQFAYDGA
125
CL154.Contig12_All   60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
117
CL154.Contig11_All   60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
117
CL154.Contig27_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig23_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig2_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig1_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig24_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
169
CL154.Contig19_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
169

Query_2              121  DYLALNEDLRSWTVADATAQISRRKWE-DADEAEHERNYLEVTCVEWLGRYLENGKESLL
179
CL1150.Contig4_All   123  DYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESLL
181
CL1150.Contig16_All  126  DYIALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESLL
184
CL154.Contig12_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
177
CL154.Contig11_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
177
CL154.Contig27_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig23_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig2_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig1_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig24_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
229
CL154.Contig19_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
229

Query_2              180  R  180
CL1150.Contig4_All   182  R  182
CL1150.Contig16_All  185  R  185
CL154.Contig12_All   178  R  178
```

```
CL154.Contig11_All   178  R  178
CL154.Contig27_All   217  R  217
CL154.Contig23_All   217  R  217
CL154.Contig2_All    217  R  217
CL154.Contig1_All    217  R  217
CL154.Contig24_All   230  R  230
CL154.Contig19_All   230  R  230


Lambda      K        H        a         alpha
   0.336    0.192    0.788    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 2743481178


Query= AFR54065.1 MHC class I antigen, partial [Meles meles]

Length=180
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  331     9e-110
CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  329     7e-107
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  132     6e-34
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  132     6e-34
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  132     1e-33
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  132     1e-33
CL154.Contig2_All  362 1396 class I histocompatibility antigen, G...  132     1e-33
CL154.Contig1_All  294 1328 minus strand class I histocompatibili...  132     1e-33
CL154.Contig24_All  126 1199 class I histocompatibility antigen, ...  132     1e-33
CL154.Contig19_All  122 1195 class I histocompatibility antigen, ...  132     1e-33



Query_3              1    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  60
CL1150.Contig4_All   3    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  62
CL1150.Contig16_All  6    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  65
CL154.Contig12_All   2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  59
CL154.Contig11_All   2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  59
CL154.Contig27_All   41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig23_All   41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig2_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig1_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig24_All   54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE
111
CL154.Contig19_All   54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE
111

Query_3              61   RQTRNLKDAAHAFRVNLNTLRDYYNQSAAGSHTIQRMYGCDVGPDGRLLRGYSQVAYDGA
120
CL1150.Contig4_All   63   QQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQFAYDGA
122
CL1150.Contig16_All  66   RQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQFAYDGA
125
CL154.Contig12_All   60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
117
CL154.Contig11_All   60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
117
CL154.Contig27_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig23_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig2_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig1_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
```

```
CL154.Contig24_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
169
CL154.Contig19_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
169

Query_3              121  DYLALNEDLRSWTVADATAQISRRKWEAA-DEAEHERNYLEVTCLEWLHRYLENGKESLL
179
CL1150.Contig4_All   123  DYLALNEDLRSWTAADTAAQISRRKWEDA-GEAERYRNYVEGTCVEWLGRYLENGKESLL
181
CL1150.Contig16_All  126  DYIALNEDLRSWTAADTAAQISRRKWEDA-GEAERYRNYVEGTCVEWLGRYLENGKESLL
184
CL154.Contig12_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
177
CL154.Contig11_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
177
CL154.Contig27_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig23_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig2_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig1_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig24_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
229
CL154.Contig19_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
229

Query_3              180  R  180
CL1150.Contig4_All   182  R  182
CL1150.Contig16_All  185  R  185
CL154.Contig12_All   178  R  178
CL154.Contig11_All   178  R  178
CL154.Contig27_All   217  R  217
CL154.Contig23_All   217  R  217
CL154.Contig2_All    217  R  217
CL154.Contig1_All    217  R  217
CL154.Contig24_All   230  R  230
CL154.Contig19_All   230  R  230


Lambda      K        H        a         alpha
   0.336    0.191    0.785    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 2743481178


Query= AFR54064.1 MHC class I antigen, partial [Meles meles]

Length=180
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  340     4e-111
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  323     1e-106
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  119     2e-29
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  119     2e-29
CL154.Contig24_All  126 1199 class I histocompatibility antigen, ...  119     3e-29
CL154.Contig19_All  122 1195 class I histocompatibility antigen, ...  119     3e-29
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  119     3e-29
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  119     3e-29
CL154.Contig2_All  362 1396 class I histocompatibility antigen, G...  119     3e-29
CL154.Contig1_All  294 1328 minus strand class I histocompatibili...  119     3e-29


Query_4              1    SHSLRYFSTAVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWVEQEGPEYWD  60
```

```
CL1150.Contig16_All  6    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  65
CL1150.Contig4_All   3    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  62
CL154.Contig12_All   2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYWE  59
CL154.Contig11_All   2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYWE  59
CL154.Contig24_All   54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE
111
CL154.Contig19_All   54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE
111
CL154.Contig27_All   41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig23_All   41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig2_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98
CL154.Contig1_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYWE  98

Query_4              61   RQTQICKDAAQTYRGNLQTALRYYNQSAAGSHTIQNVYGCDVGRDGRLLRGYSQDSYDGA
120
CL1150.Contig16_All  66   RQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQFAYDGA
125
CL1150.Contig4_All   63   QQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQFAYDGA
122
CL154.Contig12_All   60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
117
CL154.Contig11_All   60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
117
CL154.Contig24_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
169
CL154.Contig19_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
169
CL154.Contig27_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig23_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig2_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig1_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156

Query_4              121  DYIALNEDLRSWTAADTAAQITQRKWEDAGA-AERWRNYLEVTCVEWLGRYLENGKESLL
179
CL1150.Contig16_All  126  DYIALNEDLRSWTAADTAAQISRRKWEDAGE-AERYRNYVEGTCVEWLGRYLENGKESLL
184
CL1150.Contig4_All   123  DYLALNEDLRSWTAADTAAQISRRKWEDAGE-AERYRNYVEGTCVEWLGRYLENGKESLL
181
CL154.Contig12_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
177
CL154.Contig11_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
177
CL154.Contig24_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
229
CL154.Contig19_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
229
CL154.Contig27_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig23_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig2_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig1_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216

Query_4              180  R  180
CL1150.Contig16_All  185  R  185
CL1150.Contig4_All   182  R  182
CL154.Contig12_All   178  R  178
CL154.Contig11_All   178  R  178
CL154.Contig24_All   230  R  230
CL154.Contig19_All   230  R  230
CL154.Contig27_All   217  R  217
CL154.Contig23_All   217  R  217
CL154.Contig2_All    217  R  217
CL154.Contig1_All    217  R  217
```

```
Lambda      K        H        a         alpha
   0.334    0.190    0.776    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 2743481178


Query= AFR54063.1 MHC class I antigen, partial [Meles meles]

Length=180
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  369     7e-125
CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  367     1e-121
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  133     2e-34
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  133     2e-34
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  133     4e-34
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  133     4e-34
CL154.Contig2_All  362 1396 class I histocompatibility antigen, G...  133     4e-34
CL154.Contig1_All  294 1328 minus strand class I histocompatibili...  133     4e-34
CL154.Contig24_All  126 1199 class I histocompatibility antigen, ...  133     4e-34
CL154.Contig19_All  122 1195 class I histocompatibility antigen, ...  133     4e-34



Query_5              1    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDAQFVRFDSDSASLRMEPRAPWMEQEGPEYWD  60
CL1150.Contig4_All   3    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  62
CL1150.Contig16_All  6    SHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYWD  65
CL154.Contig12_All   2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYWE  59
CL154.Contig11_All   2    THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYWE  59
CL154.Contig27_All   41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYWE  98
CL154.Contig23_All   41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYWE  98
CL154.Contig2_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYWE  98
CL154.Contig1_All    41   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYWE  98
CL154.Contig24_All   54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYWE
111
CL154.Contig19_All   54   THSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYWE
111

Query_5              61   RETRNLKETTQTYRVNLNNLRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQFAYDGA
120
CL1150.Contig4_All   63   QQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQFAYDGA
122
CL1150.Contig16_All  66   RQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQFAYDGA
125
CL154.Contig12_All   60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
117
CL154.Contig11_All   60   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
117
CL154.Contig27_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig23_All   99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig2_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig1_All    99   TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
156
CL154.Contig24_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
169
CL154.Contig19_All   112  TETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDGQ
169

Query_5              121  DYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESLL
179
CL1150.Contig4_All   123  DYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESLL
181
```

```
CL1150.Contig16_All  126  DYIALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESLL
184
CL154.Contig12_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
177
CL154.Contig11_All   118  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
177
CL154.Contig27_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig23_All   157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig2_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig1_All    157  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
216
CL154.Contig24_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
229
CL154.Contig19_All   170  DHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSLT
229

Query_5              180  R  180
CL1150.Contig4_All   182  R  182
CL1150.Contig16_All  185  R  185
CL154.Contig12_All   178  R  178
CL154.Contig11_All   178  R  178
CL154.Contig27_All   217  R  217
CL154.Contig23_All   217  R  217
CL154.Contig2_All    217  R  217
CL154.Contig1_All    217  R  217
CL154.Contig24_All   230  R  230
CL154.Contig19_All   230  R  230


Lambda      K        H        a         alpha
   0.335    0.193    0.791    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 2743481178


Query= AFR54062.1 MHC class I antigen, partial [Meles meles]

Length=90
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  178     3e-53
CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  171     7e-50
CL154.Contig16_All  217 480 class I histocompatibility antigen, G...  70.7    2e-15
CL154.Contig15_All  215 478 minus strand class I histocompatibili...  70.7    2e-15
CL154.Contig14_All  217 480 class I histocompatibility antigen, G...  70.7    2e-15
CL154.Contig13_All  215 478 minus strand class I histocompatibili...  70.7    2e-15
CL154.Contig33_All  343 723 class I histocompatibility antigen, G...  70.7    4e-15
CL154.Contig21_All  339 719 class I histocompatibility antigen, G...  70.7    4e-15
CL154.Contig20_All  339 719 class I histocompatibility antigen, G...  70.7    4e-15
CL154.Contig17_All  343 723 class I histocompatibility antigen, G...  70.7    4e-15



Query_6              1    GSHSLRYFYTGVSRPGRAEPRFIAVGYVDDAQFVRFDSDSASRRMEPRAPWMEQEGPEYW  60
CL1150.Contig4_All   2    GSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYW  61
CL1150.Contig16_All  5    GSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYW  64
CL154.Contig16_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig15_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig14_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig13_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig33_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97
CL154.Contig21_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97
CL154.Contig20_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97
```

```
CL154.Contig17_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97

Query_6              61   DRQTRNLKDAAQTYRRSLNNLRDYYNQSEA  90
CL1150.Contig4_All   62   DQQTRGIKETTQTYRRSLNNLRGYYNQSAA  91
CL1150.Contig16_All  65   DRQTQICKETTQTYRGSLNILRGYYNQSAA  94
CL154.Contig16_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS    86
CL154.Contig15_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS    86
CL154.Contig14_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS    86
CL154.Contig13_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS    86
CL154.Contig33_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS    125
CL154.Contig21_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS    125
CL154.Contig20_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS    125
CL154.Contig17_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS    125


Lambda      K        H        a         alpha
   0.332    0.189    0.769    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 906584538


Query= AFR54061.1 MHC class I antigen, partial [Meles meles]

Length=90
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  144     1e-40
CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  138     6e-38
CL154.Contig16_All  217 480 class I histocompatibility antigen, G...  64.4    4e-13
CL154.Contig15_All  215 478 minus strand class I histocompatibili...  64.4    4e-13
CL154.Contig14_All  217 480 class I histocompatibility antigen, G...  64.4    4e-13
CL154.Contig13_All  215 478 minus strand class I histocompatibili...  64.4    4e-13
CL154.Contig33_All  343 723 class I histocompatibility antigen, G...  64.4    8e-13
CL154.Contig21_All  339 719 class I histocompatibility antigen, G...  64.4    8e-13
CL154.Contig20_All  339 719 class I histocompatibility antigen, G...  64.4    8e-13
CL154.Contig17_All  343 723 class I histocompatibility antigen, G...  64.4    8e-13


Query_7              1    GSHSLRYFSTAVSRPGRGEARYWEVGYVDDTQFARFDSDSASPRMEPRAPWVEQAGPEYW  60
CL1150.Contig4_All   2    GSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYW  61
CL1150.Contig16_All  5    GSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYW  64
CL154.Contig16_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig15_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig14_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig13_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig33_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYW  97
CL154.Contig21_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYW  97
CL154.Contig20_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYW  97
CL154.Contig17_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVDAQYW  97

Query_7              61   DRETRGIKDAAQTYRVDLQTALGCYNQSEA  90
CL1150.Contig4_All   62   DQQTRGIKETTQTYRRSLNNLRGYYNQSAA  91
CL1150.Contig16_All  65   DRQTQICKETTQTYRGSLNILRGYYNQSAA  94
CL154.Contig16_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS    86
CL154.Contig15_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS    86
CL154.Contig14_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS    86
CL154.Contig13_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS    86
CL154.Contig33_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS    125
CL154.Contig21_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS    125
CL154.Contig20_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS    125
CL154.Contig17_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS    125


Lambda      K        H        a         alpha
   0.332    0.188    0.775    0.443     1.38
```

```
Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 906584538


Query= AFR54060.1 MHC class I antigen, partial [Meles meles]

Length=325
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  618     0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  494     1e-169
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  266     3e-83
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  239     4e-73
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  248     6e-73
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  248     6e-73
CL1150.Contig13_All  621 1046 minus strand MHC class I antigen, p...  238     1e-72
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  248     3e-72
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  248     3e-72
CL154.Contig2_All  362 1396 class I histocompatibility antigen, G...  248     3e-72



Query_8              1    ETWAGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  60
CL1150.Contig16_All  1    ETWAGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  60
CL1150.Contig4_All   1       AGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  57
CL154.Contig12_All   1        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDS--RVDRAKPQALWMATVD  54
CL154.Contig11_All   1        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDS--RVDRAKPQALWMATVD  54
CL154.Contig27_All   40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93
CL154.Contig23_All   40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93
CL154.Contig2_All    40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93

Query_8              61   PEYWDQQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYSQH
120
CL1150.Contig16_All  61   PEYWDRQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQF
120
CL1150.Contig4_All   58   PEYWDQQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQF
117
CL154.Contig12_All   55   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
112
CL154.Contig11_All   55   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
112
CL154.Contig27_All   94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig23_All   94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig2_All    94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151

Query_8              121  SYDGADYIALNEDLRSWTAADTAAQITQRKWE-DAGEAERWRNYVEGTCVEWLGRYLENG
179
CL1150.Contig16_All  121  AYDGADYIALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENG
179
CL1150.Contig4_All   118  AYDGADYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENG
176
CL1150.Contig15_All  1                              QISRRKUE-DAGVAELERDYLEITCVKWLHRYLENG  35
CL154.Contig12_All   113  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
172
CL154.Contig11_All   113  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
172
CL154.Contig27_All   152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig23_All   152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig2_All    152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
```

```
Query_8               180  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWKRDEEDLTQDTELVETRP
239
CL1150.Contig16_All   180  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP
239
CL1150.Contig4_All    177  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWKRDEEDLTQDTELVET
234
CL1150.Contig15_All   36   KETLLR----------------------------TEITLTWQRDGEDQTQDTELVETRP  66
CL1150.Contig17_All   1          RAEPPNTRMTHHPISDHAVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP  55
CL154.Contig12_All    173  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
232
CL154.Contig11_All    173  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
232
CL1150.Contig13_All   1          RAEPPNTRMTHHPISDHAVTLRCWALDFYPAEITLTWHHEEEDLTQDTELVGTRP  55
CL154.Contig27_All    212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig23_All    212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig2_All     212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271

Query_8               240  AGDGTFQKWAAVVVPSGQEQRYTCYVQHEGLSEPITRRWE-PPRTIPITWIIAGLVLLVI
298
CL1150.Contig16_All   240  AGDGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLALLAV
299
CL1150.Contig15_All   67   AGDGTFQKWAAVVVPSGQEQRYTCYVQHEGLSEPITRRWE-PPRTIPITWIIAGLVLLVI
125
CL1150.Contig17_All   56   AGDGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLALLAV
115
CL154.Contig12_All    233  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
288
CL154.Contig11_All    233  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
288
CL1150.Contig13_All   56   TGNGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLALLAV
115
CL154.Contig27_All    272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig23_All    272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig2_All     272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327

Query_8               299  IAVI---GVAIWWKKRSGEKGPGYSHAARD  325
CL1150.Contig16_All   300  TVVV---GAVIWRRKRSGGKGPGYSHAA    324
CL1150.Contig15_All   126  IAVI---GVAIWWKKRSGEKGPGYSHAA    150
CL1150.Contig17_All   116  TVVV---GAVIWRRKRSGGKGPGYSHAA    140
CL154.Contig12_All    289  IFLLLTAGVLIFIKQYS                305
CL154.Contig11_All    289  IFLLLTAGVLIFIKQYS                305
CL1150.Contig13_All   116  TVVV---GAVIWRKKRSGGKGPGYSHAARD  142
CL154.Contig27_All    328  IFLLLTAGVLIFIKQYS                344
CL154.Contig23_All    328  IFLLLTAGVLIFIKQYS                344
CL154.Contig2_All     328  IFLLLTAGVLIFIKQYS                344


Lambda      K        H        a         alpha
   0.335    0.191    0.801    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 5708182959


Query= AFR54059.1 MHC class I antigen, partial [Meles meles]

Length=325
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  596     0.0
CL1150.Contig4_All   486 1187 MHC class I antigen, partial [Meles ...  446     1e-150
```

```
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  274     3e-86
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  244     6e-75
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  253     1e-74
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  253     1e-74
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  253     6e-74
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  253     6e-74
CL154.Contig2_All  362 1396 class I histocompatibility antigen, G...  253     6e-74
CL154.Contig1_All  294 1328 minus strand class I histocompatibili...  253     6e-74


Query_9              1    ETWAGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWVEQEG  60
CL1150.Contig16_All  1    ETWAGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  60
CL1150.Contig4_All   1       AGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  57
CL154.Contig12_All   1        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDS--RVDRAKPQALWMATVD  54
CL154.Contig11_All   1        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDS--RVDRAKPQALWMATVD  54
CL154.Contig27_All   40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93
CL154.Contig23_All   40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93
CL154.Contig2_All    40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93
CL154.Contig1_All    40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93

Query_9              61   PEYWDRQTRNLKDAAHAFRVNLNTLRDYYNQSAAGSHTIQRMYGCDMGPDGRLLRGYSQV
120
CL1150.Contig16_All  61   PEYWDRQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQF
120
CL1150.Contig4_All   58   PEYWDQQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQF
117
CL154.Contig12_All   55   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
112
CL154.Contig11_All   55   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
112
CL154.Contig27_All   94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig23_All   94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig2_All    94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig1_All    94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151

Query_9              121  AYDGADYLALNEDLRSWTVADATAQISRRKWE-DADEAEHERNYLEVTCVEWLGRYLENG
179
CL1150.Contig16_All  121  AYDGADYIALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENG
179
CL1150.Contig4_All   118  AYDGADYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENG
176
CL1150.Contig15_All  1                              QISRRKUE-DAGVAELERDYLEITCVKWLHRYLENG  35
CL154.Contig12_All   113  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
172
CL154.Contig11_All   113  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
172
CL154.Contig27_All   152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig23_All   152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig2_All    152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig1_All    152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211

Query_9              180  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP
239
CL1150.Contig16_All  180  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP
239
CL1150.Contig4_All   177  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWKRDEEDLTQDTELVET
234
CL1150.Contig15_All  36   KETLLR-----------------------------TEITLTWQRDGEDQTQDTELVETRP  66
CL1150.Contig17_All  1         RAEPPNTRMTHHPISDHAVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP  55
CL154.Contig12_All   173  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
232
```

```
CL154.Contig11_All   173  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
232
CL154.Contig27_All   212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig23_All   212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig2_All    212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig1_All    212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271

Query_9              240  AGDGTFQKWAAVVVPSGQEQRYTCHVQHEGLSEPITRRWE-PPRTIPITWIIAGVV---L
295
CL1150.Contig16_All  240  AGDGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLA---L
296
CL1150.Contig15_All  67   AGDGTFQKWAAVVVPSGQEQRYTCYVQHEGLSEPITRRWE-PPRTIPITWIIAGLV---L
122
CL1150.Contig17_All  56   AGDGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLA---L
112
CL154.Contig12_All   233  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
288
CL154.Contig11_All   233  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
288
CL154.Contig27_All   272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig23_All   272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig2_All    272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig1_All    272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327

Query_9              296  LVVIAVTGVAIWWKKRSGEKGPGYSHAA  323
CL1150.Contig16_All  297  LAVTVVVGAVIWRRKRSGGKGPGYSHAA  324
CL1150.Contig15_All  123  LVIIAVIGVAIWWKKRSGEKGPGYSHAA  150
CL1150.Contig17_All  113  LAVTVVVGAVIWRRKRSGGKGPGYSHAA  140
CL154.Contig12_All   289  IFLLLTAGVLIFIKQYS             305
CL154.Contig11_All   289  IFLLLTAGVLIFIKQYS             305
CL154.Contig27_All   328  IFLLLTAGVLIFIKQYS             344
CL154.Contig23_All   328  IFLLLTAGVLIFIKQYS             344
CL154.Contig2_All    328  IFLLLTAGVLIFIKQYS             344
CL154.Contig1_All    328  IFLLLTAGVLIFIKQYS             344


Lambda      K        H        a         alpha
   0.336    0.191    0.798    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 5708182959


Query= AFR54058.1 MHC class I antigen, partial [Meles meles]

Length=325
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  593     0.0
CL1150.Contig4_All   486 1187 MHC class I antigen, partial [Meles ...  443     3e-149
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  274     2e-86
CL154.Contig12_All   217 1134 class I histocompatibility antigen, ...  257     3e-76
CL154.Contig11_All   215 1132 minus strand class I histocompatibil...  257     3e-76
CL154.Contig24_All   126 1199 class I histocompatibility antigen, ...  258     2e-75
CL154.Contig19_All   122 1195 class I histocompatibility antigen, ...  258     2e-75
CL154.Contig27_All   343 1377 class I histocompatibility antigen, ...  257     2e-75
CL154.Contig23_All   339 1373 class I histocompatibility antigen, ...  257     2e-75
CL154.Contig2_All    362 1396 class I histocompatibility antigen, G...  257     2e-75
```

```
Query_10              1    ETWAGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  60
CL1150.Contig16_All   1    ETWAGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  60
CL1150.Contig4_All    1       AGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  57
CL154.Contig12_All    1        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDS--RVDRAKPQALWMATVD  54
CL154.Contig11_All    1        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDS--RVDRAKPQALWMATVD  54
CL154.Contig24_All    53       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD
106
CL154.Contig19_All    53       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD
106
CL154.Contig27_All    40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93
CL154.Contig23_All    40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93
CL154.Contig2_All     40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93

Query_10              61   PEYWDRQTRNLKDAAHAFRVNLNTLRDYYNQSAAGSHTIQRMYGCDVGPDGRLLRGYSQV
120
CL1150.Contig16_All   61   PEYWDRQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQF
120
CL1150.Contig4_All    58   PEYWDQQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQF
117
CL154.Contig12_All    55   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
112
CL154.Contig11_All    55   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
112
CL154.Contig24_All    107  AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
164
CL154.Contig19_All    107  AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
164
CL154.Contig27_All    94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig23_All    94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig2_All     94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151

Query_10              121  AYDGADYLALNEDLRSWTVADATAQISRRKWEAAD-EAEHERNYLEVTCLEWLHRYLENG
179
CL1150.Contig16_All   121  AYDGADYIALNEDLRSWTAADTAAQISRRKWEDAG-EAERYRNYVEGTCVEWLGRYLENG
179
CL1150.Contig4_All    118  AYDGADYLALNEDLRSWTAADTAAQISRRKWEDAG-EAERYRNYVEGTCVEWLGRYLENG
176
CL1150.Contig15_All   1                              QISRRKUEDAG-VAELERDYLEITCVKWLHRYLENG  35
CL154.Contig12_All    113  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
172
CL154.Contig11_All    113  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
172
CL154.Contig24_All    165  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
224
CL154.Contig19_All    165  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
224
CL154.Contig27_All    152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig23_All    152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig2_All     152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211

Query_10              180  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP
239
CL1150.Contig16_All   180  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP
239
CL1150.Contig4_All    177  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWKRDEEDLTQDTELVET
234
CL1150.Contig15_All   36   KETLLR----------------------------TEITLTWQRDGEDQTQDTELVETRP  66
CL154.Contig12_All    173  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
232
CL154.Contig11_All    173  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
232
CL154.Contig24_All    225  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
284
```

```
CL154.Contig19_All   225  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
284
CL154.Contig27_All   212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig23_All   212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig2_All    212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271

Query_10             240  AGDGTFQKWAAVVVPSGQEQRYTCHVQHEGLSEPITRRWE-PPRTIPITWIIAGLVLLVV
298
CL1150.Contig16_All  240  AGDGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLALLAV
299
CL1150.Contig15_All  67   AGDGTFQKWAAVVVPSGQEQRYTCYVQHEGLSEPITRRWE-PPRTIPITWIIAGLVLLVI
125
CL154.Contig12_All   233  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
288
CL154.Contig11_All   233  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
288
CL154.Contig24_All   285  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
340
CL154.Contig19_All   285  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
340
CL154.Contig27_All   272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig23_All   272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig2_All    272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327

Query_10             299  IAVI---GVAIWWKKHSGEKGPGYSHAA  323
CL1150.Contig16_All  300  TVVV---GAVIWRRKRSGGKGPGYSHAA  324
CL1150.Contig15_All  126  IAVI---GVAIWWKKRSGEKGPGYSHAA  150
CL154.Contig12_All   289  IFLLLTAGVLIFIKQYS              305
CL154.Contig11_All   289  IFLLLTAGVLIFIKQYS              305
CL154.Contig24_All   341  IFLLLTAGVLIFIKQYS              357
CL154.Contig19_All   341  IFLLLTAGVLIFIKQYS              357
CL154.Contig27_All   328  IFLLLTAGVLIFIKQYS              344
CL154.Contig23_All   328  IFLLLTAGVLIFIKQYS              344
CL154.Contig2_All    328  IFLLLTAGVLIFIKQYS              344


Lambda      K        H        a         alpha
   0.336    0.191    0.799    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 5708182959


Query= AFR54057.1 MHC class I antigen, partial [Meles meles]

Length=325
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  604     0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  440     2e-148
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  265     9e-83
CL1150.Contig17_All  682 1101 minus strand MHC class I antigen, p...  243     7e-75
CL1150.Contig13_All  621 1046 minus strand MHC class I antigen, p...  243     2e-74
CL1150.Contig18_All  1278 1694 minus strand MHC class I antigen [...  240     1e-73
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  242     1e-70
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  242     1e-70
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  242     5e-70
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  242     5e-70
```

```
Query_11              1    ETWAGSHSLRYFSTAVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWVEQEG  60
CL1150.Contig16_All   1    ETWAGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  60
CL1150.Contig4_All    1       AGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  57
CL154.Contig12_All    1        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDS--RVDRAKPQALWMATVD  54
CL154.Contig11_All    1        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDS--RVDRAKPQALWMATVD  54
CL154.Contig27_All    40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93
CL154.Contig23_All    40       GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRV--DRAKPQALWMATVD  93

Query_11              61   PEYWDRQTQICKDAAQTYRGNLQTALRYYNQSAAGSHTIQNVYGCDVGRDGRLLRGYSQD
120
CL1150.Contig16_All   61   PEYWDRQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQF
120
CL1150.Contig4_All    58   PEYWDQQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQF
117
CL154.Contig12_All    55   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
112
CL154.Contig11_All    55   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
112
CL154.Contig27_All    94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig23_All    94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151

Query_11              121  SYDGADYIALNEDLRSWTAADTAAQITQRKWE-DAGAAERWRNYLEVTCVEWLGRYLENG
179
CL1150.Contig16_All   121  AYDGADYIALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENG
179
CL1150.Contig4_All    118  AYDGADYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENG
176
CL1150.Contig15_All   1                            QISRRKUE-DAGVAELERDYLEITCVKWLHRYLENG  35
CL154.Contig12_All    113  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
172
CL154.Contig11_All    113  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
172
CL154.Contig27_All    152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig23_All    152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211

Query_11              180  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWKRDEEDLTQDTELVETRP
239
CL1150.Contig16_All   180  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP
239
CL1150.Contig4_All    177  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWKRDEEDLTQDTELVET
234
CL1150.Contig15_All   36   KETLLR-----------------------------TEITLTWQRDGEDQTQDTELVETRP  66
CL1150.Contig17_All   1         RAEPPNTRMTHHPISDHAVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP  55
CL1150.Contig13_All   1         RAEPPNTRMTHHPISDHAVTLRCWALDFYPAEITLTWHHEEEDLTQDTELVGTRP  55
CL1150.Contig18_All   1          SEPPNTRMTHHPISDHAVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP  54
CL154.Contig12_All    173  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
232
CL154.Contig11_All    173  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
232
CL154.Contig27_All    212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig23_All    212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271

Query_11              240  AGDGTFQKWAAVVVPSGQEQRYTCYVQHEGLSEPITRRWE-PPHTIPITWIIAGLVLLVV
298
CL1150.Contig16_All   240  AGDGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLALLAV
299
CL1150.Contig15_All   67   AGDGTFQKWAAVVVPSGQEQRYTCYVQHEGLSEPITRRWE-PPRTIPITWIIAGLVLLVI
125
CL1150.Contig17_All   56   AGDGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLALLAV
115
CL1150.Contig13_All   56   TGNGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLALLAV
115
CL1150.Contig18_All   55   AGDGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLALLAV
114
```

```
CL154.Contig12_All   233  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
288
CL154.Contig11_All   233  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
288
CL154.Contig27_All   272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig23_All   272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327

Query_11             299  IAVI---GAVIWWKKRSGEKGPGYSHAARD  325
CL1150.Contig16_All  300  TVVV---GAVIWRRKRSGGKGPGYSHAA    324
CL1150.Contig15_All  126  IAVI---GVAIWWKKRSGEKGPGYSHAA    150
CL1150.Contig17_All  116  TVVV---GAVIWRRKRSGGKGPGYSHAA    140
CL1150.Contig13_All  116  TVVV---GAVIWRKKRSGGKGPGYSHAARD  142
CL1150.Contig18_All  115  TVVV---GAVIWRRKRSGGKGPGYSHAA    139
CL154.Contig12_All   289  IFLLLTAGVLIFIKQYS               305
CL154.Contig11_All   289  IFLLLTAGVLIFIKQYS               305
CL154.Contig27_All   328  IFLLLTAGVLIFIKQYS               344
CL154.Contig23_All   328  IFLLLTAGVLIFIKQYS               344


Lambda      K        H        a         alpha
   0.335    0.190    0.793    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 5708182959


Query= AFR54056.1 MHC class I antigen, partial [Meles meles]

Length=325
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  625     0.0
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  486     3e-166
CL1150.Contig15_All  3 452 minus strand MHC class I antigen, part...  270     8e-85
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  257     3e-76
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  257     3e-76
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  258     1e-75
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  258     1e-75
CL154.Contig2_All  362 1396 class I histocompatibility antigen, G...  258     1e-75
CL154.Contig1_All  294 1328 minus strand class I histocompatibili...  258     1e-75
CL154.Contig24_All  126 1199 class I histocompatibility antigen, ...  257     2e-75



Query_12             1    ETWAGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDAQFVRFDSDSASLRMEPRAPWMEQEG  60
CL1150.Contig16_All  1    ETWAGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  60
CL1150.Contig4_All   1       AGSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEG  57
CL154.Contig12_All   1         GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVD  54
CL154.Contig11_All   1         GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVD  54
CL154.Contig27_All   40        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVD  93
CL154.Contig23_All   40        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVD  93
CL154.Contig2_All    40        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVD  93
CL154.Contig1_All    40        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVD  93
CL154.Contig24_All   53        GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSRVD--RAKPQALWMATVD
106

Query_12             61   PEYWDRETRNLKETTQTYRVNLNNLRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQF
120
CL1150.Contig16_All  61   PEYWDRQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQF
120
CL1150.Contig4_All   58   PEYWDQQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQF
117
CL154.Contig12_All   55   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
112
```

```
CL154.Contig11_All   55   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
112
CL154.Contig27_All   94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig23_All   94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig2_All    94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig1_All    94   AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
151
CL154.Contig24_All   107  AQYWETETQKQRAWAKVQQVETWTVMGYHNQST-GMHSTQRMFGCEIREDGHT-HSFWQF
164

Query_12             121  AYDGADYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENG
179
CL1150.Contig16_All  121  AYDGADYIALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENG
179
CL1150.Contig4_All   118  AYDGADYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENG
176
CL1150.Contig15_All  1                              QISRRKUE-DAGVAELERDYLEITCVKWLHRYLENG  35
CL154.Contig12_All   113  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
172
CL154.Contig11_All   113  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
172
CL154.Contig27_All   152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig23_All   152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig2_All    152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig1_All    152  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
211
CL154.Contig24_All   165  GFDGQDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELG
224

Query_12             180  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWKRDEEDLTQDTELVETRP
239
CL1150.Contig16_All  180  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWQRDGEDLTQDTELVETRP
239
CL1150.Contig4_All   177  KESLLRAETPNTHVTRHPISDRDVTLRCWALDFYPAEITLTWKRDEEDLTQDTELVET
234
CL1150.Contig15_All  36   KETLLR-----------------------------TEITLTWQRDGEDQTQDTELVETRP  66
CL154.Contig12_All   173  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
232
CL154.Contig11_All   173  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
232
CL154.Contig27_All   212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig23_All   212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig2_All    212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig1_All    212  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
271
CL154.Contig24_All   225  GQSLTRKEPPTVQVTRHTARDGSITLKCWARGFYPRDISLSWWLGEEELVLETEYVETRP
284

Query_12             240  AGDGTFQKWAAVVVPSGQEQRYTCYVQHEGLSEPITRRWE-PPRTIPITWIIAGLVLLVI
298
CL1150.Contig16_All  240  AGDGTFQKWAAVVVPSGEEQRYTCHVQHKGLPEPITLSWKPPPPTIPIMWIIAGLALLAV
299
CL1150.Contig15_All  67   AGDGTFQKWAAVVVPSGQEQRYTCYVQHEGLSEPITRRWE-PPRTIPITWIIAGLVLLVI
125
CL154.Contig12_All   233  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
288
CL154.Contig11_All   233  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
288
CL154.Contig27_All   272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig23_All   272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
```

```
CL154.Contig2_All    272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig1_All    272  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
327
CL154.Contig24_All   285  SGDGTYQTWAAVQVPARTEDRYICHVQHSGLNHTLTVTWE-AP---PRQWITSAVVIPML
340

Query_12             299  IAVI---GVAIWWKKRSGEKGPGYSHAA  323
CL1150.Contig16_All  300  TVVV---GAVIWRRKRSGGKGPGYSHAA  324
CL1150.Contig15_All  126  IAVI---GVAIWWKKRSGEKGPGYSHAA  150
CL154.Contig12_All   289  IFLLLTAGVLIFIKQYS             305
CL154.Contig11_All   289  IFLLLTAGVLIFIKQYS             305
CL154.Contig27_All   328  IFLLLTAGVLIFIKQYS             344
CL154.Contig23_All   328  IFLLLTAGVLIFIKQYS             344
CL154.Contig2_All    328  IFLLLTAGVLIFIKQYS             344
CL154.Contig1_All    328  IFLLLTAGVLIFIKQYS             344
CL154.Contig24_All   341  IFLLLTAGVLIFIKQYS             357


Lambda      K        H        a         alpha
   0.335    0.192    0.801    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 5708182959


Query= AFR54055.1 MHC class I antigen, partial [Meles meles]

Length=181
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  345     5e-113
CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  328     2e-108
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  121     5e-30
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  121     5e-30
CL154.Contig24_All  126 1199 class I histocompatibility antigen, ...  121     7e-30
CL154.Contig19_All  122 1195 class I histocompatibility antigen, ...  121     7e-30
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  121     8e-30
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  121     8e-30
CL154.Contig2_All  362 1396 class I histocompatibility antigen, G...  121     8e-30
CL154.Contig1_All  294 1328 minus strand class I histocompatibili...  121     8e-30



Query_13             1    GSHSLRYFSTAVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWVEQEGPEYW  60
CL1150.Contig16_All  5    GSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYW  64
CL1150.Contig4_All   2    GSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYW  61
CL154.Contig12_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig11_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig24_All   53   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW
110
CL154.Contig19_All   53   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW
110
CL154.Contig27_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97
CL154.Contig23_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97
CL154.Contig2_All    40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97
CL154.Contig1_All    40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97

Query_13             61   DRQTQICKDAAQTFRGNLQTALRYYNQSAAGSHTIQNVYGCDVGPDGRFLRGYRQDSYDG
120
CL1150.Contig16_All  65   DRQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQFAYDG
124
CL1150.Contig4_All   62   DQQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQFAYDG
121
CL154.Contig12_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
116
```

```
CL154.Contig11_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
116
CL154.Contig24_All   111  ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
168
CL154.Contig19_All   111  ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
168
CL154.Contig27_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
155
CL154.Contig23_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
155
CL154.Contig2_All    98   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
155
CL154.Contig1_All    98   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
155

Query_13             121  ADYIALNEDLRSWTAADTAAQITQRKWEDAGA-AERWRNYLEVTCVEWLGRYLENGKESL
179
CL1150.Contig16_All  125  ADYIALNEDLRSWTAADTAAQISRRKWEDAGE-AERYRNYVEGTCVEWLGRYLENGKESL
183
CL1150.Contig4_All   122  ADYLALNEDLRSWTAADTAAQISRRKWEDAGE-AERYRNYVEGTCVEWLGRYLENGKESL
180
CL154.Contig12_All   117  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
176
CL154.Contig11_All   117  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
176
CL154.Contig24_All   169  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
228
CL154.Contig19_All   169  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
228
CL154.Contig27_All   156  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
215
CL154.Contig23_All   156  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
215
CL154.Contig2_All    156  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
215
CL154.Contig1_All    156  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
215

Query_13             180  LR  181
CL1150.Contig16_All  184  LR  185
CL1150.Contig4_All   181  LR  182
CL154.Contig12_All   177  TR  178
CL154.Contig11_All   177  TR  178
CL154.Contig24_All   229  TR  230
CL154.Contig19_All   229  TR  230
CL154.Contig27_All   216  TR  217
CL154.Contig23_All   216  TR  217
CL154.Contig2_All    216  TR  217
CL154.Contig1_All    216  TR  217


Lambda      K        H        a         alpha
   0.334    0.191    0.785    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 2765083392


Query= AFR54054.1 MHC class I antigen, partial [Meles meles]

Length=181
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL1150.Contig4_All  486 1187 MHC class I antigen, partial [Meles ...  334     6e-111
CL1150.Contig16_All  189 1160 minus strand MHC class I antigen, p...  333     2e-108
CL154.Contig12_All  217 1134 class I histocompatibility antigen, ...  121     3e-30
CL154.Contig11_All  215 1132 minus strand class I histocompatibil...  121     3e-30
```

```
CL154.Contig24_All  126 1199 class I histocompatibility antigen, ...  121     5e-30
CL154.Contig19_All  122 1195 class I histocompatibility antigen, ...  121     5e-30
CL154.Contig27_All  343 1377 class I histocompatibility antigen, ...  121     6e-30
CL154.Contig23_All  339 1373 class I histocompatibility antigen, ...  121     6e-30
CL154.Contig2_All  362 1396 class I histocompatibility antigen, G...  121     6e-30
CL154.Contig1_All  294 1328 minus strand class I histocompatibili...  121     6e-30


Query_14             1    GSHSLRYFSTAVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWVEQEGPEYW  60
CL1150.Contig4_All   2    GSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYW  61
CL1150.Contig16_All  5    GSHSLRYFYTGVSRPGRGEPRFIAVGYVDDTQFVRFDSDSASRRMEPRAPWMEQEGPEYW  64
CL154.Contig12_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig11_All   1    GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHY--DSRVDRAKPQALWMATVDAQYW  58
CL154.Contig24_All   53   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW
110
CL154.Contig19_All   53   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW
110
CL154.Contig27_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97
CL154.Contig23_All   40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97
CL154.Contig2_All    40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97
CL154.Contig1_All    40   GTHSLHYHYLALSEPGPDLPQFLAVGYVDDQPFIHYDSR--VDRAKPQALWMATVDAQYW  97

Query_14             61   DRQTRGIKDAAQTFRGNLQTALRYYNQSAAGSHTIQSMYGCDVGPDGRLLRGYSQDSYDG
120
CL1150.Contig4_All   62   DQQTRGIKETTQTYRRSLNNLRGYYNQSAAGSHTFQNMYGCDVGPDGRLLRGYRQFAYDG
121
CL1150.Contig16_All  65   DRQTQICKETTQTYRGSLNILRGYYNQSAAGSHTIQNLYGCDVGPDGRLLRGYRQFAYDG
124
CL154.Contig12_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
116
CL154.Contig11_All   59   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
116
CL154.Contig24_All   111  ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
168
CL154.Contig19_All   111  ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
168
CL154.Contig27_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
155
CL154.Contig23_All   98   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
155
CL154.Contig2_All    98   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
155
CL154.Contig1_All    98   ETETQKQRAWAKVQQVETWTVMGYHNQS-TGMHSTQRMFGCEIREDGH-THSFWQFGFDG
155

Query_14             121  ADYIALNEDLRSWTAADTAAQITQRKWE-DAGVAERWRNYLEVTCVEWLGRYLENGKESL
179
CL1150.Contig4_All   122  ADYLALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESL
180
CL1150.Contig16_All  125  ADYIALNEDLRSWTAADTAAQISRRKWE-DAGEAERYRNYVEGTCVEWLGRYLENGKESL
183
CL154.Contig12_All   117  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
176
CL154.Contig11_All   117  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
176
CL154.Contig24_All   169  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
228
CL154.Contig19_All   169  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
228
CL154.Contig27_All   156  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
215
CL154.Contig23_All   156  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
215
CL154.Contig2_All    156  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
215
CL154.Contig1_All    156  QDHLSLDLETLSWVSAKPAATRTKSWWETERCYAEYDKAYLEGLCLVSLRRYLELGGQSL
215

Query_14             180  LR  181
CL1150.Contig4_All   181  LR  182
```

```
CL1150.Contig16_All  184  LR  185
CL154.Contig12_All   177  TR  178
CL154.Contig11_All   177  TR  178
CL154.Contig24_All   229  TR  230
CL154.Contig19_All   229  TR  230
CL154.Contig27_All   216  TR  217
CL154.Contig23_All   216  TR  217
CL154.Contig2_All    216  TR  217
CL154.Contig1_All    216  TR  217


Lambda      K        H        a         alpha
   0.335    0.191    0.779    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 2765083392


Query= AET36883.1 MHC class II antigen, partial [Meles meles]

Length=81
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene42913_All  108 710 MHC class II antigen DR alpha chain, pa...  177     8e-54
CL2981.Contig3_All  111 872 minus strand PREDICTED: HLA class II ...  178     3e-53
CL2981.Contig1_All  111 872 minus strand PREDICTED: HLA class II ...  178     3e-53
CL10050.Contig3_All  75 839 minus strand PREDICTED: SLA class II ...  110     1e-28
CL10050.Contig2_All  105 686 minus strand MHC class II antigen DQ...  98.7    9e-25
CL10050.Contig1_All  105 686 minus strand MHC class II antigen DQ...  98.7    9e-25
CL12484.Contig1_All  50 376 minus strand HLA class II histocompat...  90.3    3e-22
CL12484.Contig3_All  47 625 minus strand MHC class II antigen DO ...  89.5    2e-21



Query_15              1    DHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDMEKKETVWRLEEFGRFASFEAQGALAN  59
Unigene42913_All      27   DHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDMEKKETVWRLEEFGRFASFEAQGALAN  85
CL2981.Contig3_All    29   DHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDMEKKETVWRLEEFGRFASFEAQGALAN  87
CL2981.Contig1_All    29   DHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDMEKKETVWRLEEFGRFASFEAQGALAN  87
CL10050.Contig3_All   29   DHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDLEKKETVWRLPVFSTFAGFDPQGALSE  88
CL10050.Contig2_All   19   DHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDLEKKETVWRLPVFSTFAGFDPQGALSE  78
CL10050.Contig1_All   19   DHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDLEKKETVWRLPVFSTFAGFDPQGALSE  78
CL12484.Contig1_All   37             FYQSYGASGQFAYEFDGEQLFSVELKKKEAVWRLPEFGNLAHFDPQNGLAS  87
CL12484.Contig3_All   38             FYQSYGASGQFAYEFDGEQLFSVELKKKEAVWRLPEFGNLAHFDPQNGLAS  88

Query_15              60   IAVDKANLDIMIKRSNHTPNTN  81
Unigene42913_All      86   IAVDKANLDIMIKRSNHTPNTN  107
CL2981.Contig3_All    88   IAVDKANLDIMIKRSNHTPNTN  109
CL2981.Contig1_All    88   IAVDKANLDIMIKRSNHTPNTN  109
CL10050.Contig3_All   89   IATSKQNLNILTKRSNYTAATN  110
CL10050.Contig2_All   79   IATSKQNLNILTKRSNYTAATN  100
CL10050.Contig1_All   79   IATSKQNLNILTKRSNYTAATN  100
CL12484.Contig1_All   88   IAVIKAHLDVLVERSNRTRATN  109
CL12484.Contig3_All   89   IAVIKAHLDVLVERSNRTRATN  110


Lambda      K        H        a         alpha
   0.336    0.192    0.769    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 733894227


Query= AET36881.1 MHC class II antigen, partial [Meles meles]
```

```
Length=89
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  151     2e-43
Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  151     3e-43
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  137     4e-38
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  136     8e-38
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  133     7e-37
CL6815.Contig5_All  88 906 minus strand PREDICTED: HLA class II h...  128     7e-35
Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  123     3e-33
CL6815.Contig6_All  88 732 minus strand major histocompatibility ...  116     9e-31
CL6815.Contig1_All  88 732 minus strand major histocompatibility ...  116     9e-31
CL6815.Contig3_All  88 843 minus strand PREDICTED: HLA class II h...  116     9e-31



Query_16            1    DFVFQFMGQCYFTNGTERVRYLTRYIYNREEYARFDSDLGKYVAVTELGRPSAQYWNSQK  60
Unigene67843_All    34   DFVFQFKGECYFTNGTERVRSVNRYIYNREEFVRYDSDVGEYRPVTELGRPDAQYWNSQK  93
Unigene57188_All    34   DFVFQFKGECYFTNGTERVRSVNRYIYNREEFVRYDSDVGEYRPVTELGRPDAQYWNSQK  93
Unigene75449_All    28    FLFLTTSECHFTNGTERVRFLDRYFYNGEEYVRFDSDVGEYRPVTELGRPDAEYWNSQK  86
Unigene75450_All    27   DFVFQFKGECYFTNGTERVRSVNRYIYNREEFVRYDSDVGEYRPVTELGRPDAQYWNSQK  86
CL4065.Contig3_All  1     FVFQFKGECYFTNGTERVRSVNRYIYNREEFVRYDSDVGEYRPVTELGRPDAQYWNSQK  59
CL6815.Contig5_All  32   DFVIQAKADCYFINGTEKVQFVVRFIFNLEEYARFDSHVGKFVALTELGKPDAELWNHRP  91
Unigene27967_All    28    FLFLTTSECHFTNGTERVRFLDRYFYNGEEYVRFDSDVGEYRPVTELGRPDAQYWNSQK  86
CL6815.Contig6_All  32   DFVIQAKADCYFINGTEKVQFVVRFIFNLEEYARFDSHVGKFVALTELGKPDAELWNHRP  91
CL6815.Contig1_All  32   DFVIQAKADCYFINGTEKVQFVVRFIFNLEEYARFDSHVGKFVALTELGKPDAELWNHRP  91
CL6815.Contig3_All  32   DFVIQAKADCYFINGTEKVQFVVRFIFNLEEYARFDSHVGKFVALTELGKPDAELWNHRP  91

Query_16            61   DIVDRTEAERDTVCRHNYKNEERTTLQRR  89
Unigene67843_All    94   DILERTEAETDTVCRHNYLTDESFTVQRR  122
Unigene57188_All    94   DILERTEAETDTVCRHNYLTDESFTVQRR  122
Unigene75449_All    87   DILERTEAETDTVCRHNYLTDESFTVQRR  115
Unigene75450_All    87   DIMERRRAAVDTYCRHNYGVVESFLVQRR  115
CL4065.Contig3_All  60   DIMERRRAAVDTYCRHNYGVVESFLVQRR  88
CL6815.Contig5_All  92   DILERSRASVDALCRHNYK            110
Unigene27967_All    87   DIMERRRAAVDTYCRHNYGVVESFLVQRR  115
CL6815.Contig6_All  92   DILERSRASVDALCRHNYKLGAPFTVGRK  120
CL6815.Contig1_All  92   DILERSRASVDALCRHNYKLGAPFTVGRK  120
CL6815.Contig3_All  92   DILERSRASVDALCRHNYKLGAPFTVGRK  120


Lambda      K        H        a         alpha
   0.336    0.192    0.778    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 884472720


Query= AET36880.1 MHC class II antigen, partial [Meles meles]

Length=89
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  145     5e-41
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  135     2e-37
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  130     7e-36
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  131     8e-36
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  122     6e-33
Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  122     9e-33
CL6815.Contig4_All  88 741 minus strand major histocompatibility ...  91.6    8e-22
CL6815.Contig6_All  88 732 minus strand major histocompatibility ...  91.6    8e-22
CL6815.Contig1_All  88 732 minus strand major histocompatibility ...  91.6    8e-22
CL6815.Contig3_All  88 843 minus strand PREDICTED: HLA class II h...  92.0    1e-21
```

```
Query_17              1    HFLLPVKPECHYCNGTERVRLLDRYFYNSEEYVHFNSDVGEYRPVTELGRPIAQGWNSQK  60
Unigene27967_All      27   HFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVRFDSDVGEYRPVTELGRPDAQYWNSQK  86
Unigene75449_All      27   HFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVRFDSDVGEYRPVTELGRPDAEYWNSQK  86
CL4065.Contig3_All    1     FVFQFKGECYFTNGTERVRSVNRYIYNREEFVRYDSDVGEYRPVTELGRPDAQYWNSQK  59
Unigene75450_All      28    FVFQFKGECYFTNGTERVRSVNRYIYNREEFVRYDSDVGEYRPVTELGRPDAQYWNSQK  86
Unigene67843_All      35    FVFQFKGECYFTNGTERVRSVNRYIYNREEFVRYDSDVGEYRPVTELGRPDAQYWNSQK  93
Unigene57188_All      35    FVFQFKGECYFTNGTERVRSVNRYIYNREEFVRYDSDVGEYRPVTELGRPDAQYWNSQK  93
CL6815.Contig4_All    33    FVIQAKADCYFINGTEKVQFVVRFIFNLEEYARFDSHVGKFVALTELGKPDAELWNHRP  91
CL6815.Contig6_All    33    FVIQAKADCYFINGTEKVQFVVRFIFNLEEYARFDSHVGKFVALTELGKPDAELWNHRP  91
CL6815.Contig1_All    33    FVIQAKADCYFINGTEKVQFVVRFIFNLEEYARFDSHVGKFVALTELGKPDAELWNHRP  91
CL6815.Contig3_All    33    FVIQAKADCYFINGTEKVQFVVRFIFNLEEYARFDSHVGKFVALTELGKPDAELWNHRP  91

Query_17              61   DIMERKRSEVDTVCRHNHGVFESF  84
Unigene27967_All      87   DIMERRRAAVDTYCRHNYGVVESF  110
Unigene75449_All      87   DILERTEAETDTVCRHNYLTDESF  110
CL4065.Contig3_All    60   DIMERRRAAVDTYCRHNYGVVESF  83
Unigene75450_All      87   DIMERRRAAVDTYCRHNYGVVESF  110
Unigene67843_All      94   DILERTEAETDTVCRHNYLTDESF  117
Unigene57188_All      94   DILERTEAETDTVCRHNYLTDESF  117
CL6815.Contig4_All    92   DILERSRASVDALCRHNYKLGAPF  115
CL6815.Contig6_All    92   DILERSRASVDALCRHNYKLGAPF  115
CL6815.Contig1_All    92   DILERSRASVDALCRHNYKLGAPF  115
CL6815.Contig3_All    92   DILERSRASVDALCRHNYKLGAPF  115


Lambda      K        H        a         alpha
   0.343    0.198    0.854    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 884472720


Query= AET36875.1 MHC class II antigen DR alpha chain, partial [Meles
meles]

Length=230
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL2981.Contig3_All  111 872 minus strand PREDICTED: HLA class II ...  492     7e-171
CL2981.Contig1_All  111 872 minus strand PREDICTED: HLA class II ...  492     7e-171
Unigene42913_All  108 710 MHC class II antigen DR alpha chain, pa...  434     1e-149
CL10050.Contig3_All  75 839 minus strand PREDICTED: SLA class II ...  295     7e-94
CL10050.Contig2_All  105 686 minus strand MHC class II antigen DQ...  250     9e-78
CL10050.Contig1_All  105 686 minus strand MHC class II antigen DQ...  250     9e-78
CL12484.Contig3_All  47 625 minus strand MHC class II antigen DO ...  230     2e-70
CL12484.Contig2_All  651 899 minus strand PREDICTED: HLA class II...  120     6e-32
CL12484.Contig1_All  50 376 minus strand HLA class II histocompat...  112     3e-28
CL5174.Contig1_All  1 237 hypothetical protein PANDA_002284 [Ailu...  97.0    2e-23



Query_18              1    ISGVPVLGFFIMTLLMGPQESQAIKEDHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDM  59
CL2981.Contig3_All    3    INGVPVLGFFIMTLLMGPQESQAIKEDHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDM  61
CL2981.Contig1_All    3    INGVPVLGFFIMTLLMGPQESQAIKEDHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDM  61
Unigene42913_All      1    INGVPVLGFFIMTLLMGPQESQAIKEDHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDM  59
CL10050.Contig3_All   6       VLILGTLALTTMMSPGGSEDIVADHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDL  62
CL10050.Contig2_All   4                LTTMMSPGGSEDIVADHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDL  52
CL10050.Contig1_All   4                LTTMMSPGGSEDIVADHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDL  52
CL12484.Contig3_All   3    LSGGLVLGLYTLMSLLSPQEIGAIKADHMGSYGPAFYQSYGASGQFAYEFDGEQLFSVEL  62
CL12484.Contig1_All   2    LSGGLVLGLYTLMSLLSPQEIGAIKADHMGSYGPAFYQSYGASGQFAYEFDGEQLFSVEL  61

Query_18              60   EKKETVWRLEEFGRFASFEAQGALANIAVDKANLDIMIKRSNHTPNTNVPPEVTVLSNTP
119
CL2981.Contig3_All    62   EKKETVWRLEEFGRFASFEAQGALANIAVDKANLDIMIKRSNHTPNTNVPPEVTVLSNTP
121
CL2981.Contig1_All    62   EKKETVWRLEEFGRFASFEAQGALANIAVDKANLDIMIKRSNHTPNTNVPPEVTVLSNTP
121
```

```
Unigene42913_All     60   EKKETVWRLEEFGRFASFEAQGALANIAVDKANLDIMIKRSNHTPNTNVPPEVTVLSNTP
119
CL10050.Contig3_All  63   EKKETVWRLPVFSTFAGFDPQGALSEIATSKQNLNILTKRSNYTAATNEVPEVTLFPKSP
122
CL10050.Contig2_All  53   EKKETVWRLPVFSTFAGFDPQGALSEIATSKQNLNILTKRSNYTAATNEVPEVTLFPKSP
112
CL10050.Contig1_All  53   EKKETVWRLPVFSTFAGFDPQGALSEIATSKQNLNILTKRSNYTAATNEVPEVTLFPKSP
112
CL12484.Contig3_All  63   KKKEAVWRLPEFGNLAHFDPQNGLASIAVIKAHLDVLVERSNRTRATNVPPRVTVLPRFR
122
CL12484.Contig2_All  1                                                 VPPRVTVLPRFR  12
CL12484.Contig1_All  62   KKKEAVWRLPEFGNLAHFDPQNGLASIAVIKAHLDVLVERSNRTRATN
109
CL5174.Contig1_All   9                    FDAWRGIGDIVVAKKNLNNLIQRSNHTRATNEPPEVTVFPKEP  51

Query_18             120  VELGEPNTLICFIDKFSPPVINVTWLRNGNPVTTGVSETVFLPREDHLFRKFHYLPFLPS
179
CL2981.Contig3_All   122  VELGEPNTLICFIDKFSPPVINVTWLRNGNPVTTGVSETVFLPREDHLFRKFHYLPFLPS
181
CL2981.Contig1_All   122  VELGEPNTLICFIDKFSPPVINVTWLRNGNPVTTGVSETVFLPREDHLFRKFHYLPFLPS
181
Unigene42913_All     120  VELGEPNTLICFIDKFSPPVINVTWLRNGNPVTTGVSETVFLPREDHLFRKFHYLPFLPS
179
CL10050.Contig3_All  123  VMLGQPNTLICLVDNIFPPVINVTWLKNRHSVTEGVSETSFLAKKDHSFLKISYLTFLPS
182
CL10050.Contig2_All  113  VMLGQPNTLICLVDNIFPPVINVTWLKNRHSVTEGVSETSFLAKKDHSFLKISYLTFLPS
172
CL10050.Contig1_All  113  VMLGQPNTLICLVDNIFPPVINVTWLKNRHSVTEGVSETSFLAKKDHSFLKISYLTFLPS
172
CL12484.Contig3_All  123  VELGQPNVLICMVDNIFPPVINITWLRNGQIVSEGVAQTSFYSQPDHLFRKFCYLTFVPS
182
CL12484.Contig2_All  13   VELGQPNVLICMVDNIFPPVINITWLRNGQIVSEGVAQTSFYSQPDHLFRKFCYLTFVPS  72
CL5174.Contig1_All   52   VELGQPNTLICHVDKFFPPVLNVTWLRN                                  79

Query_18             180  ANDVYDCKVEHWGLDEPLLKHWEFEPPTPLPETTENVVCALGLVVGLVGII  230
CL2981.Contig3_All   182  ANDVYDCKVEHWGLDEPLLKHWEFEPPTPLPETTENVVCALGLVVGLVGIV  232
CL2981.Contig1_All   182  ANDVYDCKVEHWGLDEPLLKHWEFEPPTPLPETTENVVCALGLVVGLVGIV  232
Unigene42913_All     180  ANDVYDCKVEHWGLDEPLLKHW                               201
CL10050.Contig3_All  183  ADDIYDCKVEHWGLDEPLLKHWEPEIPSPMSELTETVVCALGLAVGLVGIV  233
CL10050.Contig2_All  173  ADDIYDCKVEHWGLDEPLLKHW                               194
CL10050.Contig1_All  173  ADDIYDCKVEHWGLDEPLLKHW                               194
CL12484.Contig3_All  183  ADDMYDCKVEH                                          193
CL12484.Contig2_All  73   ADDMYDCKVEH                                          83


Lambda      K        H        a         alpha
   0.337    0.196    0.813    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 3779567088


Query= AET36874.1 MHC class II antigen DQ alpha chain, partial [Meles
meles]

Length=205
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL10050.Contig3_All  75 839 minus strand PREDICTED: SLA class II ...  447     7e-154
CL10050.Contig2_All  105 686 minus strand MHC class II antigen DQ...  423     3e-146
CL10050.Contig1_All  105 686 minus strand MHC class II antigen DQ...  423     3e-146
CL2981.Contig3_All  111 872 minus strand PREDICTED: HLA class II ...  260     5e-81
CL2981.Contig1_All  111 872 minus strand PREDICTED: HLA class II ...  260     5e-81
Unigene42913_All  108 710 MHC class II antigen DR alpha chain, pa...  249     6e-78
CL12484.Contig3_All  47 625 minus strand MHC class II antigen DO ...  239     2e-74
CL12484.Contig2_All  651 899 minus strand PREDICTED: HLA class II...  127     2e-34
CL12484.Contig1_All  50 376 minus strand HLA class II histocompat...  115     8e-30
```

```
CL5174.Contig1_All  1 237 hypothetical protein PANDA_002284 [Ailu...  102     1e-25


Query_19              1    TLALTTMMSPGGSEDIVADHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDLEKKETVWR  60
CL10050.Contig3_All  11    TLALTTMMSPGGSEDIVADHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDLEKKETVWR  70
CL10050.Contig2_All   1    TLALTTMMSPGGSEDIVADHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDLEKKETVWR  60
CL10050.Contig1_All   1    TLALTTMMSPGGSEDIVADHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDLEKKETVWR  60
CL2981.Contig3_All   14       MTLLMGPQESQAIKEDHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDMEKKETVWR  69
CL2981.Contig1_All   14       MTLLMGPQESQAIKEDHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDMEKKETVWR  69
Unigene42913_All     12       MTLLMGPQESQAIKEDHV-IIQAEFYLTPDPSGEFMFDFDGDEIFHVDMEKKETVWR  67
CL12484.Contig3_All  14       LMSLLSPQEIGAIKADHMGSYGPAFYQSYGASGQFAYEFDGEQLFSVELKKKEAVWR  70
CL12484.Contig1_All  13       LMSLLSPQEIGAIKADHMGSYGPAFYQSYGASGQFAYEFDGEQLFSVELKKKEAVWR  69

Query_19              61   LPVFSTFAGFDPQGALSEIATSKQNLNILTKRSNYTAATNEVPEVTLFPKSPVMLGQPNT
120
CL10050.Contig3_All   71   LPVFSTFAGFDPQGALSEIATSKQNLNILTKRSNYTAATNEVPEVTLFPKSPVMLGQPNT
130
CL10050.Contig2_All   61   LPVFSTFAGFDPQGALSEIATSKQNLNILTKRSNYTAATNEVPEVTLFPKSPVMLGQPNT
120
CL10050.Contig1_All   61   LPVFSTFAGFDPQGALSEIATSKQNLNILTKRSNYTAATNEVPEVTLFPKSPVMLGQPNT
120
CL2981.Contig3_All    70   LEEFGRFASFEAQGALANIAVDKANLDIMIKRSNHTPNTNVPPEVTVLSNTPVELGEPNT
129
CL2981.Contig1_All    70   LEEFGRFASFEAQGALANIAVDKANLDIMIKRSNHTPNTNVPPEVTVLSNTPVELGEPNT
129
Unigene42913_All      68   LEEFGRFASFEAQGALANIAVDKANLDIMIKRSNHTPNTNVPPEVTVLSNTPVELGEPNT
127
CL12484.Contig3_All   71   LPEFGNLAHFDPQNGLASIAVIKAHLDVLVERSNRTRATNVPPRVTVLPRFRVELGQPNV
130
CL12484.Contig2_All   3                                                PRVTVLPRFRVELGQPNV  20
CL12484.Contig1_All   70   LPEFGNLAHFDPQNGLASIAVIKAHLDVLVERSNRTRATN
109
CL5174.Contig1_All    1     PDFIHAFDFDAWRGIGDIVVAKKNLNNLIQRSNHTRATNEPPEVTVFPKEPVELGQPNT  59

Query_19              121  LICLVDNIFPPVINVTWLKNRHSVTEGVSETSFLAKKDHSFLKISYLTFLPSADDIYDCK
180
CL10050.Contig3_All   131  LICLVDNIFPPVINVTWLKNRHSVTEGVSETSFLAKKDHSFLKISYLTFLPSADDIYDCK
190
CL10050.Contig2_All   121  LICLVDNIFPPVINVTWLKNRHSVTEGVSETSFLAKKDHSFLKISYLTFLPSADDIYDCK
180
CL10050.Contig1_All   121  LICLVDNIFPPVINVTWLKNRHSVTEGVSETSFLAKKDHSFLKISYLTFLPSADDIYDCK
180
CL2981.Contig3_All    130  LICFIDKFSPPVINVTWLRNGNPVTTGVSETVFLPREDHLFRKFHYLPFLPSANDVYDCK
189
CL2981.Contig1_All    130  LICFIDKFSPPVINVTWLRNGNPVTTGVSETVFLPREDHLFRKFHYLPFLPSANDVYDCK
189
Unigene42913_All      128  LICFIDKFSPPVINVTWLRNGNPVTTGVSETVFLPREDHLFRKFHYLPFLPSANDVYDCK
187
CL12484.Contig3_All   131  LICMVDNIFPPVINITWLRNGQIVSEGVAQTSFYSQPDHLFRKFCYLTFVPSADDMYDCK
190
CL12484.Contig2_All   21   LICMVDNIFPPVINITWLRNGQIVSEGVAQTSFYSQPDHLFRKFCYLTFVPSADDMYDCK  80
CL5174.Contig1_All    60   LICHVDKFFPPVLNVTWLRN                                       79

Query_19              181  VEHWGLDEPLLKHWEPEIPTPMSEL  205
CL10050.Contig3_All   191  VEHWGLDEPLLKHWEPEIPSPMSEL  215
CL10050.Contig2_All   181  VEHWGLDEPLLKHW             194
CL10050.Contig1_All   181  VEHWGLDEPLLKHW             194
CL2981.Contig3_All    190  VEHWGLDEPLLKHWEFEPPTPLPE   213
CL2981.Contig1_All    190  VEHWGLDEPLLKHWEFEPPTPLPE   213
Unigene42913_All      188  VEHWGLDEPLLKHW             201
CL12484.Contig3_All   191  VEH                        193
CL12484.Contig2_All   81   VEH                        83


Lambda      K        H        a         alpha
   0.333    0.190    0.770    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3
```

```
Effective search space used: 3242696763


Query= AET36873.1 MHC class II antigen DQ alpha chain, partial [Meles
meles]

Length=205
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

CL10050.Contig3_All  75 839 minus strand PREDICTED: SLA class II ...  394     3e-133
CL10050.Contig2_All  105 686 minus strand MHC class II antigen DQ...  370     2e-125
CL10050.Contig1_All  105 686 minus strand MHC class II antigen DQ...  370     2e-125
CL2981.Contig3_All  111 872 minus strand PREDICTED: HLA class II ...  246     8e-76
CL2981.Contig1_All  111 872 minus strand PREDICTED: HLA class II ...  246     8e-76
Unigene42913_All  108 710 MHC class II antigen DR alpha chain, pa...  235     9e-73
CL12484.Contig3_All  47 625 minus strand MHC class II antigen DO ...  226     3e-69
CL12484.Contig2_All  651 899 minus strand PREDICTED: HLA class II...  119     1e-31
CL12484.Contig1_All  50 376 minus strand HLA class II histocompat...  109     1e-27
CL5174.Contig1_All  1 237 hypothetical protein PANDA_002284 [Ailu...  98.3    3e-24



Query_20             1    TLALTTIMSLGGSEDI---VADHVASYGISVYQSYGPSGQYTREFDGDEEFYVDLEKKET  57
CL10050.Contig3_All  11   TLALTTMMSPGGSEDI---VADHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDLEKKET  67
CL10050.Contig2_All  1    TLALTTMMSPGGSEDI---VADHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDLEKKET  57
CL10050.Contig1_All  1    TLALTTMMSPGGSEDI---VADHVGAYGVEVYQSYGPSGQYTQEFDGDELFYVDLEKKET  57
CL2981.Contig3_All   14      MTLLMGPQESQAI---KEDHVIIQA-EFYLTPDPSGEFMFDFDGDEIFHVDMEKKET  66
CL2981.Contig1_All   14      MTLLMGPQESQAI---KEDHVIIQA-EFYLTPDPSGEFMFDFDGDEIFHVDMEKKET  66
Unigene42913_All     12      MTLLMGPQESQAI---KEDHVIIQA-EFYLTPDPSGEFMFDFDGDEIFHVDMEKKET  64
CL12484.Contig3_All  8    VLGLYTLMSLLSPQEIGAIKADHMGSYGPAFYQSYGASGQFAYEFDGEQLFSVELKKKEA  67
CL12484.Contig1_All  8     LGLYTLMSLLSPQEIGAIKADHMGSYGPAFYQSYGASGQFAYEFDGEQLFSVELKKKEA  66

Query_20             58   VWQLPMFQALRRFDPQGALRNLAIAKQNLNILTKRSNYTAATNEVPEVTLFLKTPVMLGQ
117
CL10050.Contig3_All  68   VWRLPVFSTFAGFDPQGALSEIATSKQNLNILTKRSNYTAATNEVPEVTLFPKSPVMLGQ
127
CL10050.Contig2_All  58   VWRLPVFSTFAGFDPQGALSEIATSKQNLNILTKRSNYTAATNEVPEVTLFPKSPVMLGQ
117
CL10050.Contig1_All  58   VWRLPVFSTFAGFDPQGALSEIATSKQNLNILTKRSNYTAATNEVPEVTLFPKSPVMLGQ
117
CL2981.Contig3_All   67   VWRLEEFGRFASFEAQGALANIAVDKANLDIMIKRSNHTPNTNVPPEVTVLSNTPVELGE
126
CL2981.Contig1_All   67   VWRLEEFGRFASFEAQGALANIAVDKANLDIMIKRSNHTPNTNVPPEVTVLSNTPVELGE
126
Unigene42913_All     65   VWRLEEFGRFASFEAQGALANIAVDKANLDIMIKRSNHTPNTNVPPEVTVLSNTPVELGE
124
CL12484.Contig3_All  68   VWRLPEFGNLAHFDPQNGLASIAVIKAHLDVLVERSNRTRATNVPPRVTVLPRFRVELGQ
127
CL12484.Contig2_All  3                                                PRVTVLPRFRVELGQ  17
CL12484.Contig1_All  67   VWRLPEFGNLAHFDPQNGLASIAVIKAHLDVLVERSNRTRATN
109
CL5174.Contig1_All   9                FDAWRGIGDIVVAKKNLNNLIQRSNHTRATNEPPEVTVFPKEPVELGQ  56

Query_20             118  PNTLICLVDNIFPPVINVTWLKNRHSVTEGVSETHFLIKKDYSFLKFSYLTFLPSADDIY
177
CL10050.Contig3_All  128  PNTLICLVDNIFPPVINVTWLKNRHSVTEGVSETSFLAKKDHSFLKISYLTFLPSADDIY
187
CL10050.Contig2_All  118  PNTLICLVDNIFPPVINVTWLKNRHSVTEGVSETSFLAKKDHSFLKISYLTFLPSADDIY
177
CL10050.Contig1_All  118  PNTLICLVDNIFPPVINVTWLKNRHSVTEGVSETSFLAKKDHSFLKISYLTFLPSADDIY
177
CL2981.Contig3_All   127  PNTLICFIDKFSPPVINVTWLRNGNPVTTGVSETVFLPREDHLFRKFHYLPFLPSANDVY
186
CL2981.Contig1_All   127  PNTLICFIDKFSPPVINVTWLRNGNPVTTGVSETVFLPREDHLFRKFHYLPFLPSANDVY
186
Unigene42913_All     125  PNTLICFIDKFSPPVINVTWLRNGNPVTTGVSETVFLPREDHLFRKFHYLPFLPSANDVY
184
CL12484.Contig3_All  128  PNVLICMVDNIFPPVINITWLRNGQIVSEGVAQTSFYSQPDHLFRKFCYLTFVPSADDMY
187
```

```
CL12484.Contig2_All  18   PNVLICMVDNIFPPVINITWLRNGQIVSEGVAQTSFYSQPDHLFRKFCYLTFVPSADDMY  77
CL5174.Contig1_All   57   PNTLICHVDKFFPPVLNVTWLRN                                    79

Query_20             178  DCKVEHWGLDEPLLKHWEPEIPTPMSEL  205
CL10050.Contig3_All  188  DCKVEHWGLDEPLLKHWEPEIPSPMSEL  215
CL10050.Contig2_All  178  DCKVEHWGLDEPLLKHW             194
CL10050.Contig1_All  178  DCKVEHWGLDEPLLKHW             194
CL2981.Contig3_All   187  DCKVEHWGLDEPLLKHWEFEPPTPLPE   213
CL2981.Contig1_All   187  DCKVEHWGLDEPLLKHWEFEPPTPLPE   213
Unigene42913_All     185  DCKVEHWGLDEPLLKHW             201
CL12484.Contig3_All  188  DCKVEH                        193
CL12484.Contig2_All  78   DCKVEH                        83


Lambda      K        H        a         alpha
   0.334    0.192    0.775    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 3242696763


Query= AET36872.1 MHC class II antigen DQ beta chain, partial [Meles meles]

Length=224
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  490     4e-170
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  475     2e-165
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  418     6e-142
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  398     4e-134
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  341     2e-112
Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  341     1e-111
CL6815.Contig3_All  88 843 minus strand PREDICTED: HLA class II h...  286     8e-91
CL6815.Contig2_All  88 882 minus strand PREDICTED: HLA class II h...  287     1e-90
CL6815.Contig5_All  88 906 minus strand PREDICTED: HLA class II h...  287     2e-90
CL6815.Contig6_All  88 732 minus strand major histocompatibility ...  279     7e-89



Query_21             1    MALWIPRGLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIY  60
Unigene57188_All     1    MALWIPRGLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIY  60
Unigene67843_All     1    MALWIPRGLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIY  60
Unigene75449_All     1           GAWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFY  53
Unigene75450_All     1           GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIY  53
CL4065.Contig3_All   1                                       FVFQFKGECYFTNGTERVRSVNRYIY  26
Unigene27967_All     1           GAWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFY  53
CL6815.Contig3_All   5        WVP---WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIF  58
CL6815.Contig2_All   5        WVP---WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIF  58
CL6815.Contig5_All   5        WVP---WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIF  58
CL6815.Contig6_All   5        WVP---WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIF  58

Query_21             61   NREEFVRYDSDVGEYRPVTELGRPDAEYWNSQKDILERTEAETDTVCRHNYLTDESFTVQ  120
Unigene57188_All     61   NREEFVRYDSDVGEYRPVTELGRPDAQYWNSQKDILERTEAETDTVCRHNYLTDESFTVQ  120
Unigene67843_All     61   NREEFVRYDSDVGEYRPVTELGRPDAQYWNSQKDILERTEAETDTVCRHNYLTDESFTVQ  120
Unigene75449_All     54   NGEEYVRFDSDVGEYRPVTELGRPDAEYWNSQKDILERTEAETDTVCRHNYLTDESFTVQ  113
Unigene75450_All     54   NREEFVRYDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQ  113
CL4065.Contig3_All   27   NREEFVRYDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQ  86
Unigene27967_All     54   NGEEYVRFDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQ  113
CL6815.Contig3_All   59   NLEEYARFDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVG  118
CL6815.Contig2_All   59   NLEEYARFDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVG  118
CL6815.Contig5_All   59   NLEEYARFDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVG  118
CL6815.Contig6_All   59   NLEEYARFDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVG  118

Query_21             121  RRVEPTVTISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGD  180
Unigene57188_All     121  RRVEPTVTISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGD  180
Unigene67843_All     121  RRVEPTVTISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGD  180
Unigene75449_All     114  RRVEPTVTISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGD  173
```

```
Unigene75450_All    114  RRVEPTVTVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGD  173
CL4065.Contig3_All  87   RRVEPTVTVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGD  146
Unigene27967_All    114  RRVEPTVTVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGD  173
CL6815.Contig3_All  119  RKVQPEVAVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGD  178
CL6815.Contig2_All  119  RKVQPEVAVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGD  178
CL6815.Contig5_All  119  RKVQPEVAVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGD  178
CL6815.Contig6_All  119  RKVQPEVAVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGD  178

Query_21            181  WTFQILVMLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQ  224
Unigene57188_All    181  WTFQILVMLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQ  224
Unigene67843_All    181  WTFQILVMLEMTPQRGDVYTCHVEHPSLQSPITVEWR         217
Unigene75449_All    174  WTFQILVMLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQ  217
Unigene75450_All    174  WTFQTLVMLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQ  217
CL4065.Contig3_All  147  WTFQTLVMLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQ  190
Unigene27967_All    174  WTFQTLVMLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQ  217
CL6815.Contig3_All  179  WTFQTMVMLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQS      218
CL6815.Contig2_All  179  WTFQTMVMLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQS      218
CL6815.Contig5_All  179  WTFQTMVMLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQS      218
CL6815.Contig6_All  179  WTFQTMVMLEMTPALGDVYTCLVNHVSLLSPVSVEWR         215


Lambda      K        H        a         alpha
   0.336    0.191    0.786    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 3650718210


Query= AET36871.1 MHC class II antigen DR beta chain, partial [Meles meles]

Length=258
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  478     2e-164
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  462     4e-158
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  434     8e-148
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  399     1e-133
Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  383     4e-127
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  341     8e-112
CL6815.Contig2_All  88 882 minus strand PREDICTED: HLA class II h...  302     1e-95
CL6815.Contig5_All  88 906 minus strand PREDICTED: HLA class II h...  302     2e-95
CL6815.Contig3_All  88 843 minus strand PREDICTED: HLA class II h...  300     5e-95
CL6815.Contig6_All  88 732 minus strand major histocompatibility ...  261     4e-81



Query_22            1    GSWMTALTLILMVLSPPLAWARDTPRHFLMQFKGECYFTNGTERVRLLVRHIYNREEFVR  60
Unigene27967_All    1    GAWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
Unigene75450_All    1    GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  60
CL4065.Contig3_All  1                              FVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  33
Unigene75449_All    1    GAWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
Unigene57188_All    8    GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  67
Unigene67843_All    8    GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  67
CL6815.Contig2_All  8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig5_All  8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig3_All  8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig6_All  8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65

Query_22            61   FDSDVGEYRPVTELGRPIAQGWNSQKDIMERRRAEVDTVCRHNYGVVESFTVQRRVEPTV  120
Unigene27967_All    61   FDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  120
Unigene75450_All    61   YDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  120
CL4065.Contig3_All  34   YDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  93
Unigene75449_All    61   FDSDVGEYRPVTELGRPDAEYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  120
Unigene57188_All    68   YDSDVGEYRPVTELGRPDAQYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  127
Unigene67843_All    68   YDSDVGEYRPVTELGRPDAQYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  127
CL6815.Contig2_All  66   FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig5_All  66   FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
```

```
CL6815.Contig3_All  66   FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig6_All  66   FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125

Query_22            121  TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIPNGDWTFQTLV  180
Unigene27967_All    121  TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGDWTFQTLV  180
Unigene75450_All    121  TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGDWTFQTLV  180
CL4065.Contig3_All  94   TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGDWTFQTLV  153
Unigene75449_All    121  TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  180
Unigene57188_All    128  TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  187
Unigene67843_All    128  TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  187
CL6815.Contig2_All  126  AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig5_All  126  AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig3_All  126  AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig6_All  126  AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185

Query_22            181  MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
Unigene27967_All    181  MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
Unigene75450_All    181  MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
CL4065.Contig3_All  154  MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  213
Unigene75449_All    181  MLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQSKMLSGIGGFVLGLIFLGLGLIV  240
Unigene57188_All    188  MLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQSKMLSGIGGFVLGLIFLGLGLIV  247
Unigene67843_All    188  MLEMTPQRGDVYTCHVEHPSLQSPITVEWR                              217
CL6815.Contig2_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig5_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig3_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig6_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWR                              215

Query_22            241  YFRNQKGHSGLQPTGLLS  258
Unigene27967_All    241  YFRNQKGHSGLQPTGLLS  258
Unigene75450_All    241  YFRNQKGHSGLQPTGLLS  258
CL4065.Contig3_All  214  YFRNQKGHSGLQPTGLLS  231
Unigene75449_All    241  RHRSQKGPRGSPPAGLL   257
Unigene57188_All    248  RHRSQKGPRGSPPAGLL   264
CL6815.Contig2_All  246  CLRAQKGYAETRLSG     260
CL6815.Contig5_All  246  CLRAQKGYAETRLSG     260
CL6815.Contig3_All  246  CLRAQKG             252


Lambda      K        H        a         alpha
   0.339    0.194    0.800    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 4333524636


Query= AET36870.1 MHC class II antigen DR beta chain, partial [Meles meles]

Length=258
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  554     0.0
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  461     9e-158
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  436     8e-148
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  433     2e-147
Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  379     3e-125
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  335     2e-109
CL6815.Contig2_All  88 882 minus strand PREDICTED: HLA class II h...  301     5e-95
CL6815.Contig5_All  88 906 minus strand PREDICTED: HLA class II h...  300     1e-94
CL6815.Contig3_All  88 843 minus strand PREDICTED: HLA class II h...  298     2e-94
CL6815.Contig6_All  88 732 minus strand major histocompatibility ...  260     2e-80



Query_23            1    GSWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
Unigene27967_All    1    GAWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
Unigene75450_All    1    GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  60
Unigene75449_All    1    GAWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
```

```
CL4065.Contig3_All  1                                 FVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  33
Unigene57188_All    8     GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  67
Unigene67843_All    8     GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  67
CL6815.Contig2_All  8       WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig5_All  8       WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig3_All  8       WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig6_All  8       WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65

Query_23            61    FDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAEVDTVCRHNYGVVESFLVQRRVEPTV  120
Unigene27967_All    61    FDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  120
Unigene75450_All    61    YDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  120
Unigene75449_All    61    FDSDVGEYRPVTELGRPDAEYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  120
CL4065.Contig3_All  34    YDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  93
Unigene57188_All    68    YDSDVGEYRPVTELGRPDAQYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  127
Unigene67843_All    68    YDSDVGEYRPVTELGRPDAQYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  127
CL6815.Contig2_All  66    FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig5_All  66    FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig3_All  66    FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig6_All  66    FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125

Query_23            121   TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGDWTFQTLV  180
Unigene27967_All    121   TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGDWTFQTLV  180
Unigene75450_All    121   TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGDWTFQTLV  180
Unigene75449_All    121   TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  180
CL4065.Contig3_All  94    TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGDWTFQTLV  153
Unigene57188_All    128   TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  187
Unigene67843_All    128   TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  187
CL6815.Contig2_All  126   AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig5_All  126   AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig3_All  126   AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig6_All  126   AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185

Query_23            181   MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
Unigene27967_All    181   MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
Unigene75450_All    181   MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
Unigene75449_All    181   MLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQSKMLSGIGGFVLGLIFLGLGLIV  240
CL4065.Contig3_All  154   MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  213
Unigene57188_All    188   MLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQSKMLSGIGGFVLGLIFLGLGLIV  247
Unigene67843_All    188   MLEMTPQRGDVYTCHVEHPSLQSPITVEWR                              217
CL6815.Contig2_All  186   MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig5_All  186   MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig3_All  186   MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig6_All  186   MLEMTPALGDVYTCLVNHVSLLSPVSVEWR                              215

Query_23            241   YFRNQKGHSGLQPTGLLS  258
Unigene27967_All    241   YFRNQKGHSGLQPTGLLS  258
Unigene75450_All    241   YFRNQKGHSGLQPTGLLS  258
Unigene75449_All    241   RHRSQKGPRGSPPAGLL   257
CL4065.Contig3_All  214   YFRNQKGHSGLQPTGLLS  231
Unigene57188_All    248   RHRSQKGPRGSPPAGLL   264
CL6815.Contig2_All  246   CLRAQKGYAETRLSG     260
CL6815.Contig5_All  246   CLRAQKGYAETRLSG     260
CL6815.Contig3_All  246   CLRAQKG             252


Lambda      K        H        a         alpha
   0.339    0.195    0.804    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 4333524636


Query= AET36869.1 MHC class II antigen DR beta chain, partial [Meles meles]

Length=258
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value
```

```
Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  557     0.0
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  464     7e-159
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  435     2e-148
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  428     6e-145
Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  371     2e-122
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  327     1e-106
CL6815.Contig2_All  88 882 minus strand PREDICTED: HLA class II h...  298     5e-94
CL6815.Contig5_All  88 906 minus strand PREDICTED: HLA class II h...  298     8e-94
CL6815.Contig3_All  88 843 minus strand PREDICTED: HLA class II h...  296     2e-93
CL6815.Contig6_All  88 732 minus strand major histocompatibility ...  257     2e-79



Query_24            1    GSWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
Unigene27967_All    1    GAWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
Unigene75450_All    1    GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  60
CL4065.Contig3_All  1                                FVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  33
Unigene75449_All    1    GAWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
Unigene57188_All    8    GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  67
Unigene67843_All    8    GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  67
CL6815.Contig2_All  8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig5_All  8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig3_All  8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig6_All  8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65

Query_24            61   FDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  120
Unigene27967_All    61   FDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  120
Unigene75450_All    61   YDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  120
CL4065.Contig3_All  34   YDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  93
Unigene75449_All    61   FDSDVGEYRPVTELGRPDAEYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  120
Unigene57188_All    68   YDSDVGEYRPVTELGRPDAQYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  127
Unigene67843_All    68   YDSDVGEYRPVTELGRPDAQYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  127
CL6815.Contig2_All  66   FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig5_All  66   FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig3_All  66   FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig6_All  66   FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125

Query_24            121  TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEEESGVVSTGLIPNGDWTFQTLV  180
Unigene27967_All    121  TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEEESGVVSTGLIRNGDWTFQTLV  180
Unigene75450_All    121  TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEEESGVVSTGLIRNGDWTFQTLV  180
CL4065.Contig3_All  94   TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEEESGVVSTGLIRNGDWTFQTLV  153
Unigene75449_All    121  TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  180
Unigene57188_All    128  TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  187
Unigene67843_All    128  TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  187
CL6815.Contig2_All  126  AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig5_All  126  AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig3_All  126  AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig6_All  126  AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185

Query_24            181  MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
Unigene27967_All    181  MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
Unigene75450_All    181  MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
CL4065.Contig3_All  154  MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  213
Unigene75449_All    181  MLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQSKMLSGIGGFVLGLIFLGLGLIV  240
Unigene57188_All    188  MLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQSKMLSGIGGFVLGLIFLGLGLIV  247
Unigene67843_All    188  MLEMTPQRGDVYTCHVEHPSLQSPITVEWR                              217
CL6815.Contig2_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig5_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig3_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig6_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWR                              215

Query_24            241  YFRNQKGHSGLQPTGLLS  258
Unigene27967_All    241  YFRNQKGHSGLQPTGLLS  258
Unigene75450_All    241  YFRNQKGHSGLQPTGLLS  258
CL4065.Contig3_All  214  YFRNQKGHSGLQPTGLLS  231
Unigene75449_All    241  RHRSQKGPRGSPPAGLL   257
Unigene57188_All    248  RHRSQKGPRGSPPAGLL   264
CL6815.Contig2_All  246  CLRAQKGYAETRLSG     260
CL6815.Contig5_All  246  CLRAQKGYAETRLSG     260
CL6815.Contig3_All  246  CLRAQKG             252
```

```
Lambda      K        H        a         alpha
   0.338    0.194    0.805    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 4333524636


Query= AET36868.1 MHC class II antigen DR beta chain, partial [Meles meles]

Length=258
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Unigene27967_All  155 928 minus strand MHC class II antigen DR be...  493     2e-170
Unigene75450_All  119 892 minus strand MHC class II antigen DR be...  440     1e-149
CL4065.Contig3_All  1603 2295 minus strand MHC class II antigen D...  412     2e-139
Unigene75449_All  155 928 minus strand MHC class II antigen [Zalo...  414     2e-139
Unigene57188_All  98 892 minus strand MHC class II antigen [Zalop...  361     2e-118
Unigene67843_All  95 745 minus strand MHC class II antigen DQ bet...  317     8e-103
CL6815.Contig2_All  88 882 minus strand PREDICTED: HLA class II h...  299     1e-94
CL6815.Contig5_All  88 906 minus strand PREDICTED: HLA class II h...  299     3e-94
CL6815.Contig3_All  88 843 minus strand PREDICTED: HLA class II h...  297     6e-94
CL6815.Contig6_All  88 732 minus strand major histocompatibility ...  258     4e-80



Query_25              1    GSWMTALTVILMVLSPPMAWARDTPPHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
Unigene27967_All      1    GAWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
Unigene75450_All      1    GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  60
CL4065.Contig3_All    1                                  FVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  33
Unigene75449_All      1    GAWMTALTLILMVLSPPLAWARDTPRHFLFLTTSECHFTNGTERVRFLDRYFYNGEEYVR  60
Unigene57188_All      8    GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  67
Unigene67843_All      8    GLWTAAVMAILVVLSVPVAEGRDSPKDFVFQFKGECYFTNGTERVRSVNRYIYNREEFVR  67
CL6815.Contig2_All    8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig5_All    8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig3_All    8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65
CL6815.Contig6_All    8      WTVVLLVSVIRLDSSRTQGRDSPEDFVIQAKADCYFINGTEKVQFVVRFIFNLEEYAR  65

Query_25             61    FDSDVGEYRPVTELGRPIAQGWNSQKDIMEQKRANVDTYCRHNYGVGESFTVQRRVEPTV  120
Unigene27967_All     61    FDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  120
Unigene75450_All     61    YDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  120
CL4065.Contig3_All   34    YDSDVGEYRPVTELGRPDAQYWNSQKDIMERRRAAVDTYCRHNYGVVESFLVQRRVEPTV  93
Unigene75449_All     61    FDSDVGEYRPVTELGRPDAEYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  120
Unigene57188_All     68    YDSDVGEYRPVTELGRPDAQYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  127
Unigene67843_All     68    YDSDVGEYRPVTELGRPDAQYWNSQKDILERTEAETDTVCRHNYLTDESFTVQRRVEPTV  127
CL6815.Contig2_All   66    FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig5_All   66    FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig3_All   66    FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125
CL6815.Contig6_All   66    FDSHVGKFVALTELGKPDAELWNHRPDILERSRASVDALCRHNYKLGAPFTVGRKVQPEV  125

Query_25            121    TVYPAKNQPLQHHSLLVCSVNGFYPGHIEVRWYQNGQEEESGVVSTGLIHNGDWTFQTLV  180
Unigene27967_All    121    TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGDWTFQTLV  180
Unigene75450_All    121    TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGDWTFQTLV  180
CL4065.Contig3_All   94    TVYPAKNQPLQHHNLLVCSVNGFYPGHIEVRWFRNGQEEESGVVSTGLIRNGDWTFQTLV  153
Unigene75449_All    121    TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  180
Unigene57188_All    128    TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  187
Unigene67843_All    128    TISPSRTEVLNHHNMLVCSVTDFYPGQIKVRWFRNDQEEKAGVVSTPLIRNGDWTFQILV  187
CL6815.Contig2_All  126    AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig5_All  126    AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig3_All  126    AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185
CL6815.Contig6_All  126    AVHPERTPSLQHRSLLFCSVTGFYPGDIKIRWFRNGQEQRVGVLSTGLVRNGDWTFQTMV  185

Query_25            181    MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
Unigene27967_All    181    MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
Unigene75450_All    181    MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  240
CL4065.Contig3_All  154    MLETVPQSGEVYTCQVEHPSLTSPVTVEWRAQSGSAQSKILSGTGGFVLGLLFLVVGLFI  213
Unigene75449_All    181    MLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQSKMLSGIGGFVLGLIFLGLGLIV  240
```

```
Unigene57188_All    188  MLEMTPQRGDVYTCHVEHPSLQSPITVEWRAQSESAQSKMLSGIGGFVLGLIFLGLGLIV  247
Unigene67843_All    188  MLEMTPQRGDVYTCHVEHPSLQSPITVEWR                              217
CL6815.Contig2_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig5_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig3_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWRAQSAYSWRKMLSGIAAFLIGLIFVLVGTVI  245
CL6815.Contig6_All  186  MLEMTPALGDVYTCLVNHVSLLSPVSVEWR                              215

Query_25            241  YFRNQKGHSGLQPTGLLS  258
Unigene27967_All    241  YFRNQKGHSGLQPTGLLS  258
Unigene75450_All    241  YFRNQKGHSGLQPTGLLS  258
CL4065.Contig3_All  214  YFRNQKGHSGLQPTGLLS  231
Unigene75449_All    241  RHRSQKGPRGSPPAGLL   257
Unigene57188_All    248  RHRSQKGPRGSPPAGLL   264
CL6815.Contig2_All  246  CLRAQKGYAETRLSG     260
CL6815.Contig5_All  246  CLRAQKGYAETRLSG     260
CL6815.Contig3_All  246  CLRAQKG             252


Lambda      K        H        a         alpha
   0.338    0.193    0.800    0.443     1.38

Gapped
Lambda      K        H        a         alpha    sigma
   0.290   0.0750    0.280     1.04     11.5     12.3

Effective search space used: 4333524636


  Database: BLAST Database
    Posted date:  Feb 23, 2017  2:08 PM
  Number of letters in database: 28,354,467
  Number of sequences in database:  127,401



Matrix: BLOSUM90
Gap Penalties: Existence: 10, Extension: 1
Neighboring words threshold: 11
Window for multiple hits: 40
```

8.6 Number of annotations per gene of IKB database

#	Gene abbreviation	Full name	Number of repeated annotations
1	WAS	Wiskott-Aldrich syndrome protein.	421
2	MME	membrane metallo-endopeptidase.	295
3	CD22	CD22 molecule.	295
4	CD2	CD2 molecule.	273
5	TMC8	EVIN2.	187
6	IL12RB1	interleukin 12 receptor, beta 1 isoform 1 precursor.	171
7	TRAF3	TNF receptor-associated factor 3 isoform 1.	151
8	LYST	lysosomal trafficking regulator.	128

9	KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog isoform 1 precursor.	126
10	CD19	CD19 antigen.	112
11	CD44	CD44 antigen isoform 5 precursor.	109
12	CD46	CD46 antigen, complement regulatory protein isoform 12 precursor.	107
13	ADAM17	ADAM metallopeptidase domain 17 preproprotein.	106
14	C4B	complement component 4B preproprotein.	97
15	RFX1	regulatory factor X1.	97
16	ITGA5	integrin alpha 5 precursor.	88
17	DKC1	dyskerin.	88
18	CD5	CD5 molecule.	87
19	BLNK	B-cell linker.	86
20	STAT3	signal transducer and activator of transcription 3 isoform 2.	84
21	STAT4	signal transducer and activator of transcription 4.	82
22	FGFR1	fibroblast growth factor receptor 1 isoform 2 precursor.	80
23	NFATC1	nuclear factor of activated T-cells, cytosolic component 1 isoform B.	77
24	NCF1C	neutrophil cytosolic factor 1 isoform 1.	76
25	STAT6	signal transducer and activator of transcription 6.	76
26	UNC13D	unc-13 homolog D.	74
27	MR1	major histocompatibility complex, class I-related.	69
28	TCIRG1	T-cell, immune regulator 1 isoform a.	67
29	TCF7	transcription factor 7 (T-cell specific, HMG-box) isoform 1.	64
30	CD4	CD4 antigen precursor.	63
31	JAK3	Janus kinase 3.	59
32	CD69	CD69 molecule.	58
33	TCF3	transcription factor 3.	58
34	IKBKE	IKK-related kinase epsilon.	58
35	CLCN7	chloride channel 7.	55
36	CD2BP2	CD2 antigen (cytoplasmic tail) binding protein 2.	53
37	CD300A	leukocyte membrane antigen.	53
38	TNFRSF14	tumor necrosis factor receptor superfamily, member 14 precursor.	53
39	HAMP	hepcidin antimicrobial peptide.	53
40	CD80	CD80 antigen precursor.	52
41	CD226	CD226 molecule.	52
42	IL21R	interleukin 21 receptor precursor.	52
43	CD151	CD151 antigen.	51
44	CCR1	chemokine (C-C motif) receptor 1.	51

45	CIITA	class II transactivator.	51
46	CD40	CD40 antigen isoform 2 precursor.	50
47	TMC6	EVIN1.	50
48	SP110	SP110 nuclear body protein isoform a.	49
49	MVK	mevalonate kinase.	49
50	IGSF8	immunoglobulin superfamily, member 8.	48
51	CCR2	chemokine (C-C motif) receptor 2 isoform A.	48
52	NCF2	neutrophil cytosolic factor 2.	48
53	WIPF1	WAS/WASL interacting protein family, member 1.	48
54	CD1C	CD1C antigen precursor.	47
55	ITGA6	integrin alpha chain, alpha 6 isoform b precursor.	47
56	CD97	CD97 antigen isoform 3 precursor.	47
57	CD33	CD33 antigen (gp67) isoform 1 precursor.	45
58	TRAF5	TNF receptor-associated factor 5.	45
59	CASP8	caspase 8 isoform A precursor.	44
60	CRLF3	cytokine receptor-like factor 3.	44
61	IL3RA	interleukin 3 receptor, alpha precursor.	44
62	TAP1	transporter 1, ATP-binding cassette, sub-family B.	43
63	TBX21	T-box 21.	43
64	ABCF1	ATP-binding cassette, sub-family F, member 1 isoform a.	43
65	IGF2R	insulin-like growth factor 2 receptor precursor.	42
66	PPP3CC	protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform (calcineurin A gamma).	42
67	IRF5	interferon regulatory factor 5 isoform a.	41
68	IFNGR1	interferon gamma receptor 1.	40
69	IKBKB	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta.	40
70	SIGIRR	single Ig IL-1R-related molecule.	39
71	CD7	CD7 antigen precursor.	38
72	JAK2	Janus kinase 2.	38
73	CEACAM1	carcinoembryonic antigen-related cell adhesion molecule 1 isoform 1 precursor.	37
74	CD86	CD86 antigen isoform 2 precursor.	37
75	FYN	protein-tyrosine kinase fyn isoform a.	37
76	PSIP1	PC4 and SFRS1 interacting protein 1 isoform 1.	37
77	LAX1	lymphocyte transmembrane adaptor 1.	37
78	STAT2	signal transducer and activator of transcription 2.	36
79	LPIN2	lipin 2.	36
80	CD48	CD48 molecule.	35

81	IRF2	interferon regulatory factor 2.	34
82	FOXK2	forkhead box K2.	34
83	PLK3	polo-like kinase 3.	34
84	SEMA7A	semaphorin 7A.	33
85	TNFRSF4	tumor necrosis factor receptor superfamily, member 4 precursor.	33
86	PPIA	peptidylprolyl isomerase A.	33
87	HPS3	Hermansky-Pudlak syndrome 3 protein.	33
88	MYO5A	myosin VA (heavy polypeptide 12, myoxin).	33
89	FAS	tumor necrosis factor receptor superfamily, member 6 isoform 1 precursor.	32
90	DCLRE1C	artemis protein isoform a.	32
91	DOCK2	dedicator of cytokinesis 2.	32
92	STAT5B	signal transducer and activator of transcription 5B.	32
93	SEMA4D	semaphorin 4D.	31
94	CD163	CD163 antigen isoform a.	31
95	LY9	lymphocyte antigen 9 isoform b.	31
96	FOXP3	forkhead box P3.	31
97	IL12RB2	interleukin 12 receptor, beta 2 precursor.	31
98	IRAK1	interleukin-1 receptor-associated kinase 1 isoform 2.	30
99	SH2B2	SH2B adaptor protein 2.	30
100	RELB	reticuloendotheliosis viral oncogene homolog B.	30
101	TNFRSF1A	tumor necrosis factor receptor 1 precursor.	29
102	IL6R	interleukin 6 receptor isoform 1 precursor.	29
103	JMJD6	jumonji domain containing 6 isoform 2.	29
104	TYK2	tyrosine kinase 2.	29
105	RIPK1	receptor (TNFRSF)-interacting serine-threonine kinase 1.	29
106	ZAP70	zeta-chain associated protein kinase 70kDa isoform 1.	29
107	IRAK3	interleukin-1 receptor-associated kinase 3.	29
108	DDR1	discoidin domain receptor family, member 1 isoform b.	28
109	IRF8	interferon regulatory factor 8.	28
110	CD27	tumor necrosis factor receptor superfamily, member 7 precursor.	27
111	FLT3	fms-related tyrosine kinase 3.	27
112	STAT1	signal transducer and activator of transcription 1 isoform alpha.	27
113	MX1	myxovirus resistance protein 1.	27
114	ITGAV	integrin alpha-V precursor.	26
115	PLAUR	plasminogen activator, urokinase receptor isoform 1 precursor.	26
116	CCR5	chemokine (C-C motif) receptor 5.	26

117	CD200	CD200 antigen isoform b.	26
118	C1QBP	complement component 1, q subcomponent binding protein precursor.	26
119	IL18	interleukin 18 proprotein.	26
120	ITGAX	PREDICTED: similar to integrin alpha X precursor.	25
121	CD83	CD83 antigen isoform b.	25
122	CD99	CD99 molecule.	25
123	ALCAM	activated leukocyte cell adhesion molecule.	25
124	PRNP	prion protein preproprotein.	25
125	CLIP1	restin isoform a.	25
126	PAFAH1B1	platelet-activating factor acetylhydrolase, isoform Ib, alpha subunit (45kD).	25
127	PPP3CB	protein phosphatase 3 (formerly 2B), catalytic subunit, beta isoform (calcineurin A beta).	25
128	TRAF2	TNF receptor-associated factor 2.	25
129	FANCD2	Fanconi anemia complementation group D2 isoform b.	25
130	NRAS	neuroblastoma RAS viral (v-ras) oncogene homolog.	25
131	CSF2RB	colony stimulating factor 2 receptor, beta, low-affinity (granulocyte-macrophage).	24
132	SDC1	syndecan 1 precursor.	24
133	ATM	ataxia telangiectasia mutated protein isoform 1.	24
134	MAPK14	mitogen-activated protein kinase 14 isoform 1.	24
135	CD1A	CD1A antigen precursor.	23
136	ICAM3	intercellular adhesion molecule 3 precursor.	23
137	PTPRJ	protein tyrosine phosphatase, receptor type, J isoform 1 precursor.	23
138	IL17RA	interleukin 17A receptor precursor.	23
139	CLIP2	CAP-GLY domain containing linker protein 2 isoform 1.	23
140	IL16	interleukin 16 isoform 1 precursor.	23
141	PSME1	proteasome activator subunit 1 isoform 1.	23
142	TRAF4	TNF receptor-associated factor 4.	23
143	CD53	CD53 antigen.	22
144	XRCC5	ATP-dependent DNA helicase II.	22
145	CSF1	colony stimulating factor 1 isoform a precursor.	22
146	FGFR3	fibroblast growth factor receptor 3 isoform 1 precursor.	22
147	IL28RA	interleukin 28 receptor, alpha isoform 1.	22
148	NFATC2	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2 isoform B.	21
149	CASP10	caspase 10 isoform a preproprotein.	21
150	BANK1	B-cell scaffold protein with ankyrin repeats 1 isoform 1.	21
151	CLU	clusterin isoform 1.	21

152	CSF2RA	colony stimulating factor 2 receptor alpha chain isoform a precursor.	21
153	IL15RA	interleukin 15 receptor, alpha isoform 1 precursor.	21
154	TNFRSF18	tumor necrosis factor receptor superfamily, member 18 isoform 1 precursor.	21
155	CD1E	CD1E antigen isoform a precursor.	20
156	CD6	CD6 molecule.	20
157	CD36	CD36 antigen.	20
158	CD37	CD37 antigen isoform B.	20
159	CD47	CD47 antigen isoform 1 precursor.	20
160	IL1R1	interleukin 1 receptor, type I precursor.	20
161	CD244	CD244 natural killer cell receptor 2B4.	20
162	TLR2	toll-like receptor 2.	20
163	PLAA	phospholipase A2-activating protein isoform 1.	20
164	LIG1	DNA ligase I.	20
165	CHUK	conserved helix-loop-helix ubiquitous kinase.	20
166	ILF3	interleukin enhancer binding factor 3 isoform b.	20
167	ZEB1	zinc finger E-box binding homeobox 1.	20
168	TLR3	toll-like receptor 3.	20
169	MLPH	melanophilin isoform 1.	20
170	CD28	CD28 antigen.	19
171	ICAM2	intercellular adhesion molecule 2 precursor.	19
172	CD300C	CD300C antigen.	19
173	MRE11A	meiotic recombination 11 homolog A isoform 2.	19
174	MYD88	myeloid differentiation primary response gene (88).	19
175	GFI1	growth factor independent 1.	19
176	CD1B	CD1B antigen precursor.	18
177	ST6GAL1	sialyltransferase 1 isoform a.	18
178	CD84	CD84 molecule.	18
179	IGF1R	insulin-like growth factor 1 receptor precursor.	18
180	EMR3	egf-like module-containing mucin-like receptor 3.	18
181	RFXANK	regulatory factor X-associated ankyrin-containing protein isoform a.	18
182	FCRL5	Fc receptor-like 5.	18
183	FANCA	Fanconi anemia, complementation group A isoform a.	18
184	PTPN22	protein tyrosine phosphatase, non-receptor type 22 (lymphoid) isoform 2.	17
185	ITGA1	integrin, alpha 1 precursor.	17
186	CD9	CD9 antigen.	16
187	ENTPD1	ectonucleoside triphosphate diphosphohydrolase 1 isoform 1.	16

188	NT5E	5` nucleotidase, ecto.	16
189	TNFRSF1B	tumor necrosis factor receptor 2 precursor.	16
190	MUC1	mucin 1 isoform 2 precursor.	16
191	CXCL16	chemokine (C-X-C motif) ligand 16.	16
192	PSMB8	proteasome beta 8 subunit isoform E1 proprotein.	16
193	RAG1	recombination activating gene 1.	16
194	PAFAH2	platelet-activating factor acetylhydrolase 2.	16
195	ANP32B	acidic (leucine-rich) nuclear phosphoprotein 32 family, member B.	16
196	LCK	lymphocyte-specific protein tyrosine kinase precursor.	16
197	PIK3CG	phosphoinositide-3-kinase, catalytic, gamma polypeptide.	16
198	SLAMF6	activating NK receptor precursor.	16
199	PSMF1	proteasome inhibitor subunit 1.	16
200	HAX1	HCLS1 associated protein X-1 isoform b.	16
201	CD1D	CD1D antigen precursor.	15
202	SIGLEC5	sialic acid binding Ig-like lectin 5.	15
203	SPN	sialophorin.	15
204	LY75	lymphocyte antigen 75.	15
205	CCR6	chemokine (C-C motif) receptor 6.	15
206	TRADD	TNFRSF1A-associated via death domain.	15
207	IRF9	interferon-stimulated transcription factor 3, gamma 48kDa.	15
208	IL1B	interleukin 1, beta proprotein.	15
209	IL1F10	interleukin 1 family, member 10.	15
210	ITGAD	integrin, alpha D precursor.	15
211	CD8A	CD8 antigen alpha polypeptide isoform 1 precursor.	14
212	FCER2	Fc fragment of IgE, low affinity II, receptor for (CD23A).	14
213	IL2RA	interleukin 2 receptor, alpha chain precursor.	14
214	ITGA2	integrin alpha 2 precursor.	14
215	SELP	selectin P precursor.	14
216	CD72	CD72 molecule.	14
217	CD81	CD81 antigen.	14
218	CSF1R	colony stimulating factor 1 receptor precursor.	14
219	C2	complement component 2 precursor.	14
220	SIRPA	signal-regulatory protein alpha precursor.	14
221	NFATC3	cytoplasmic nuclear factor of activated T-cells 3 isoform 2.	14
222	EBF1	early B-cell factor.	14
223	MBP	myelin basic protein isoform 1.	14

224	IKBKG	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase gamma.	14
225	PARP1	poly (ADP-ribose) polymerase family, member 1.	14
226	IL17RB	interleukin 17B receptor precursor.	14
227	LITAF	lipopolysaccharide-induced TNF factor.	14
228	SOCS7	suppressor of cytokine signaling 7.	14
229	TRAF7	ring finger and WD repeat domain 1.	14
230	TSPYL2	TSPY-like 2.	14
231	MEFV	Mediterranean fever protein.	14
232	UNC93B1	unc-93 homolog B1.	14
233	SIGLEC6	sialic acid binding Ig-like lectin 6 isoform 1 precursor.	13
234	CD164	CD164 molecule, sialomucin.	13
235	MASP2	mannan-binding lectin serine protease 2 isoform 1 precursor.	13
236	LTB	lymphotoxin-beta isoform a.	13
237	ADA	adenosine deaminase.	13
238	G6PD	glucose-6-phosphate dehydrogenase isoform a.	13
239	TNFRSF13C	BAFF receptor.	13
240	MICB	MHC class I polypeptide-related sequence B.	13
241	LIG4	DNA ligase IV.	13
242	C1R	complement component 1, r subcomponent.	13
243	PSMB7	proteasome beta 7 subunit proprotein.	13
244	IL27RA	class I cytokine receptor.	13
245	ADAM10	ADAM metallopeptidase domain 10.	13
246	BTLA	B and T lymphocyte associated isoform 1.	13
247	CASP1	caspase 1 isoform beta precursor.	13
248	CD2AP	CD2-associated protein.	13
249	CKLF	chemokine-like factor isoform e.	13
250	IL1RAP	interleukin 1 receptor accessory protein isoform 1.	13
251	TRAF1	TNF receptor-associated factor 1.	13
252	TAZ	tafazzin isoform 1.	13
253	TFRC	transferrin receptor.	12
254	KLRD1	killer cell lectin-like receptor subfamily D, member 1 isoform 1.	12
255	TNFRSF13B	tumor necrosis factor receptor 13B.	12
256	RFX5	regulatory factor X, 5.	12
257	UNG	uracil-DNA glycosylase isoform UNG1 precursor.	12
258	HLA-DMB	major histocompatibility complex, class II, DM beta precursor.	12
259	CD248	tumor endothelial marker 1 precursor.	12

260	ERGIC2	PTX1 protein.	12
261	IL17RD	interleukin 17 receptor D isoform hSef-b.	12
262	MRC2	mannose receptor, C type 2.	12
263	DNMT3B	DNA cytosine-5 methyltransferase 3 beta isoform 1.	12
264	CD55	decay accelerating factor for complement.	11
265	CD96	CD96 antigen isoform 2 precursor.	11
266	IL4R	interleukin 4 receptor alpha chain isoform a precursor.	11
267	SMARCAL1	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin a-like 1.	11
268	CLEC12A	myeloid inhibitory C-type lectin-like receptor isoform alpha.	11
269	SH2D1A	SH2 domain protein 1A.	11
270	LAIR1	leukocyte-associated immunoglobulin-like receptor 1 isoform a precursor.	11
271	AP3B1	adaptor-related protein complex 3, beta 1 subunit.	11
272	CR1	complement receptor 1 isoform S precursor.	10
273	ITGB3	integrin beta chain, beta 3 precursor.	10
274	CD79A	CD79A antigen isoform 2 precursor.	10
275	CD82	CD82 antigen isoform 2.	10
276	SLC44A1	CDW92 antigen.	10
277	ITGB4	integrin beta 4 isoform 1 precursor.	10
278	LAMP2	lysosomal-associated membrane protein 2 precursor.	10
279	IL7R	interleukin 7 receptor precursor.	10
280	IL6ST	interleukin 6 signal transducer isoform 1 precursor.	10
281	PDGFRB	platelet-derived growth factor receptor beta precursor.	10
282	BLM	Bloom syndrome protein.	10
283	LILRB2	leukocyte immunoglobulin-like receptor, subfamily B, member 2 isoform 1.	10
284	CAMP	cathelicidin antimicrobial peptide.	10
285	TLR1	toll-like receptor 1.	10
286	LTF	lactotransferrin.	10
287	IFI27	interferon, alpha-inducible protein 27.	10
288	NFKBIA	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha.	10
289	TOLLIP	toll interacting protein.	10
290	CD99L2	CD99 antigen-like 2 isoform E3`-E4`-E3-E4.	10
291	IL18RAP	interleukin 18 receptor accessory protein precursor.	10
292	IL32	interleukin 32 isoform B.	10
293	ILF2	interleukin enhancer binding factor 2.	10
294	LILRB3	leukocyte immunoglobulin-like receptor, subfamily B, member 3 isoform 2.	10

295	NCR3	natural cytotoxicity triggering receptor 3.	10
296	CD8B	CD8b antigen isoform 5 precursor.	9
297	MS4A1	membrane-spanning 4-domains, subfamily A, member 1.	9
298	CD68	CD68 antigen isoform B.	9
299	CD74	CD74 antigen isoform a.	9
300	IKBKAP	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex-associated protein.	9
301	CTSS	cathepsin S preproprotein.	9
302	A2M	alpha-2-macroglobulin precursor.	9
303	ICOS	inducible T-cell co-stimulator precursor.	9
304	ULBP1	UL16 binding protein 1.	9
305	NOD2	nucleotide-binding oligomerization domain containing 2.	9
306	CD24	CD24 antigen precursor.	9
307	CD276	CD276 antigen isoform a.	9
308	CD302	CD302 molecule.	9
309	CLCF1	cardiotrophin-like cytokine factor 1.	9
310	IL17RE	interleukin 17 receptor E isoform 1.	9
311	ISG20	interferon stimulated exonuclease gene 20kDa.	9
312	NFIL3	nuclear factor, interleukin 3 regulated.	9
313	RAB27A	Ras-related protein Rab-27A.	9
314	SBDS	Shwachman-Bodian-Diamond syndrome protein.	9
315	NLRP3	NLR family, pyrin domain containing 3 isoform a.	9
316	ITGAM	integrin alpha M precursor.	8
317	PTPRC	protein tyrosine phosphatase, receptor type, C isoform 4.	8
318	SELL	selectin L precursor.	8
319	CD63	CD63 antigen isoform B.	8
320	ITGAE	integrin, alpha E precursor.	8
321	ENG	endoglin precursor.	8
322	CSF3R	colony stimulating factor 3 receptor isoform a precursor.	8
323	SLAMF1	signaling lymphocytic activation molecule family member 1.	8
324	MFI2	melanoma-associated antigen p97 isoform 1, precursor.	8
325	CCR3	CC chemokine receptor 3.	8
326	WASF1	Wiskott-Aldrich syndrome protein family member 1.	8
327	JAK1	janus kinase 1.	8
328	NDUFS3	NADH dehydrogenase (ubiquinone) Fe-S protein 3, 30kDa (NADH-coenzyme Q reductase).	8
329	SDF2	stromal cell-derived factor 2 precursor.	8
330	CCR4	chemokine (C-C motif) receptor 4.	8

331	HLA-DQB1	major histocompatibility complex, class II, DQ beta 1 precursor.	8
332	IFIT1	interferon-induced protein with tetratricopeptide repeats 1 isoform 2.	8
333	IRF1	interferon regulatory factor 1.	8
334	PSMB6	proteasome beta 6 subunit.	8
335	CD300LG	CD300 molecule-like family member g.	8
336	CD320	8D6 antigen.	8
337	SCARB1	scavenger receptor class B, member 1 isoform 1.	7
338	CD40LG	CD40 ligand.	7
339	LRP1	low density lipoprotein-related protein 1.	7
340	IL2RB	interleukin 2 receptor beta precursor.	7
341	ACE	angiotensin I converting enzyme isoform 1 precursor.	7
342	CTLA4	cytotoxic T-lymphocyte-associated protein 4 isoform b precursor.	7
343	CD160	CD160 antigen.	7
344	L1CAM	L1 cell adhesion molecule isoform 1 precursor.	7
345	VPREB1	immunoglobulin iota chain preproprotein.	7
346	MSR1	macrophage scavenger receptor 1 isoform type 2.	7
347	ABCB1	ATP-binding cassette sub-family B member 1.	7
348	CCRN4L	CCR4 carbon catabolite repression 4-like.	7
349	IL10	interleukin 10 precursor.	7
350	NP	purine nucleoside phosphorylase.	7
351	C8G	complement component 8, gamma polypeptide.	7
352	RAC2	ras-related C3 botulinum toxin substrate 2.	7
353	CFH	complement factor H isoform a precursor.	7
354	HLA-DOA	major histocompatibility complex, class II, DO alpha precursor.	7
355	IFI35	interferon-induced protein 35.	7
356	IL1A	interleukin 1, alpha proprotein.	7
357	IL15	interleukin 15 preproprotein.	7
358	IL18R1	interleukin 18 receptor 1 precursor.	7
359	BATF	basic leucine zipper transcription factor, ATF-like.	7
360	IFI44L	histocompatibility 28.	7
361	TAP2	transporter 2, ATP-binding cassette, sub-family B isoform 1.	7
362	C6	Complement component 6 precursor.	7
363	CCRL2	chemokine (C-C motif) receptor-like 2.	7
364	CD200R1	CD200 receptor 1 isoform a.	7
365	CD300E	CD300e molecule.	7
366	CD300LF	NK inhibitory receptor precursor.	7

367	CX3CR1	chemokine (C-X3-C motif) receptor 1.	7
368	IL17D	interleukin 17D precursor.	7
369	SOCS2	suppressor of cytokine signaling-2.	7
370	SOCS4	suppressor of cytokine signaling 4.	7
371	CD3E	CD3E antigen, epsilon polypeptide (TiT3 complex).	6
372	PTGFRN	prostaglandin F2 receptor negative regulator.	6
373	ITGAL	integrin alpha L precursor.	6
374	CD70	tumor necrosis factor ligand superfamily, member 7.	6
375	TNFSF8	tumor necrosis factor (ligand) superfamily, member 8.	6
376	PECAM1	platelet/endothelial cell adhesion molecule (CD31 antigen).	6
377	ITGA4	integrin alpha 4 precursor.	6
378	ICAM1	intercellular adhesion molecule 1 precursor.	6
379	CD59	CD59 antigen p18-20.	6
380	SLC7A5	solute carrier family 7 (cationic amino acid transporter, y+ system), member 5.	6
381	PVRL2	poliovirus receptor related 2 isoform alpha precursor.	6
382	IL5RA	interleukin 5 receptor, alpha isoform 1 precursor.	6
383	MST1R	macrophage stimulating 1 receptor.	6
384	CD180	CD180 molecule.	6
385	CD209	CD209 molecule.	6
386	INSR	insulin receptor isoform Long precursor.	6
387	PLXNC1	plexin C1.	6
388	GYPA	glycophorin A precursor.	6
389	CXCL1	chemokine (C-X-C motif) ligand 1.	6
390	IL8	interleukin 8 precursor.	6
391	KLRK1	NKG2-D type II integral membrane protein.	6
392	IFNAR2	interferon alpha/beta receptor 2 isoform b precursor.	6
393	ICOSLG	inducible T-cell co-stimulator ligand.	6
394	NPTN	neuroplastin isoform b precursor.	6
395	EBI2	EBV-induced G protein-coupled receptor 2.	6
396	HLA-DOB	major histocompatibility complex, class II, DO beta precursor.	6
397	HRH2	histamine receptor H2.	6
398	IL19	interleukin 19 isoform 2 precursor.	6
399	IL20	interleukin 20 precursor.	6
400	HSP90B1	tumor rejection antigen (gp96) 1.	6
401	FCRLA	Fc receptor-like and mucin-like 1.	6
402	IL11RA	interleukin 11 receptor, alpha isoform 1 precursor.	6

403	IL18BP	interleukin 18 binding protein precursor.	6
404	SOCS6	suppressor of cytokine signaling 6.	6
405	TNFSF10	tumor necrosis factor (ligand) superfamily, member 10.	6
406	TRAF3IP1	TNF receptor-associated factor 3 interacting protein 1.	6
407	CD5L	CD5 molecule-like.	5
408	CD14	CD14 antigen precursor.	5
409	ITGB2	integrin, beta 2 precursor.	5
410	CD34	CD34 antigen isoform a.	5
411	SCARB2	scavenger receptor class B, member 2.	5
412	CEACAM8	carcinoembryonic antigen-related cell adhesion molecule 8.	5
413	CD93	CD93 antigen precursor.	5
414	FASLG	fas ligand.	5
415	HMMR	hyaluronan-mediated motility receptor isoform a.	5
416	CXCR4	chemokine (C-X-C motif) receptor 4 isoform a.	5
417	CCR9	chemokine (C-C motif) receptor 9 isoform B.	5
418	CCL3	chemokine (C-C motif) ligand 3.	5
419	ENC1	ectodermal-neural cortex (with BTB-like domain).	5
420	CFB	complement factor B preproprotein.	5
421	CASP2	caspase 2 isoform 1 preproprotein.	5
422	CASP7	caspase 7 isoform alpha precursor.	5
423	TNFSF13B	tumor necrosis factor (ligand) superfamily, member 13b.	5
424	PSMB9	proteasome beta 9 subunit isoform 1 proprotein.	5
425	C1QB	complement component 1, q subcomponent, B chain precursor.	5
426	AICDA	activation-induced cytidine deaminase.	5
427	PLA2R1	phospholipase A2 receptor 1 isoform 2 precursor.	5
428	PAFAH1B2	platelet-activating factor acetylhydrolase, isoform Ib, beta subunit 30kDa.	5
429	CFP	complement factor properdin.	5
430	CYSLTR1	cysteinyl leukotriene receptor 1.	5
431	GUSB	glucuronidase, beta.	5
432	FCER1G	Fc fragment of IgE, high affinity I, receptor for, gamma polypeptide precursor.	5
433	IFIT2	interferon-induced protein with tetratricopeptide repeats 2.	5
434	PSME3	proteasome activator subunit 3 isoform 1.	5
435	CLEC5A	C-type lectin, superfamily member 5.	5
436	IL12A	interleukin 12A precursor.	5
437	PSME2	proteasome activator subunit 2.	5
438	TAL1	T-cell acute lymphocytic leukemia 1.	5

439	C1RL	complement component 1, r subcomponent-like precursor.	5
440	CCRL1	chemokine (C-C motif) receptor-like 1.	5
441	CXCR6	G protein-coupled receptor TYMSTR.	5
442	IL17C	interleukin 17C.	5
443	IRAK1BP1	interleukin-1 receptor-associated kinase 1 binding protein 1.	5
444	IRAK2	interleukin-1 receptor-associated kinase 2.	5
445	LILRA5	leukocyte immunoglobulin-like receptor subfamily A member 5 isoform 1.	5
446	SLAMF7	SLAM family member 7.	5
447	TICAM2	toll-like receptor adaptor molecule 2.	5
448	TLR4	toll-like receptor 4 precursor.	5
449	TLR5	toll-like receptor 5.	5
450	TNFSF11	tumor necrosis factor ligand superfamily, member 11 isoform 1.	5
451	CTSC	cathepsin C isoform a preproprotein.	5
452	FANCG	Fanconi anemia, complementation group G.	5
453	IFNGR2	interferon-gamma receptor beta chain precursor.	5
454	FCGR3B	low affinity immunoglobulin gamma Fc region receptor III-B precursor.	4
455	CR2	complement component (3d/Epstein Barr virus) receptor 2 isoform 2.	4
456	ITGB1	integrin beta 1 isoform 1B precursor.	4
457	GP1BA	platelet glycoprotein Ib alpha polypeptide precursor.	4
458	NCAM1	neural cell adhesion molecule 1 isoform 1.	4
459	CD58	CD58 molecule.	4
460	CD109	CD109.	4
461	MPL	myeloproliferative leukemia virus oncogene.	4
462	PVRL1	poliovirus receptor-related 1 isoform 1.	4
463	IL2RG	interleukin 2 receptor, gamma precursor.	4
464	KLRB1	killer cell lectin-like receptor subfamily B, member 1.	4
465	SIGLEC1	sialoadhesin precursor.	4
466	LAMP3	lysosomal-associated membrane protein 3.	4
467	IL13RA1	interleukin 13 receptor, alpha 1 precursor.	4
468	KEL	Kell blood group, metallo-endopeptidase.	4
469	CCL4	chemokine C-C motif ligand 4 isoform 1 precursor.	4
470	CCL27	small inducible cytokine A27 precursor.	4
471	CCL28	chemokine (C-C motif) ligand 28 precursor.	4
472	IL9	interleukin 9 precursor.	4
473	CASP3	caspase 3 preproprotein.	4
474	CRADD	CASP2 and RIPK1 domain containing adaptor with death domain.	4

475	YWHAZ	tyrosine 3/tryptophan 5 -monooxygenase activation protein, zeta polypeptide.	4
476	CFI	complement factor I.	4
477	MYLK	myosin light chain kinase isoform 1.	4
478	TCN2	transcobalamin II precursor.	4
479	PDGFB	platelet-derived growth factor beta isoform 1, preproprotein.	4
480	SOD3	superoxide dismutase 3, extracellular precursor.	4
481	TNF	tumor necrosis factor alpha.	4
482	RELA	v-rel reticuloendotheliosis viral oncogene homolog A, nuclear factor of kappa light polypeptide gene enhancer in B-cells 3, p65.	4
483	PPP3CA	protein phosphatase 3 (formerly 2B), catalytic subunit, alpha isoform (calcineurin A alpha).	4
484	IGJ	immunoglobulin J chain.	4
485	LCP2	lymphocyte cytosolic protein 2.	4
486	PTAFR	platelet-activating factor receptor.	4
487	IL24	interleukin 24 isoform 1 precursor.	4
488	IL1RN	interleukin 1 receptor antagonist isoform 3.	4
489	POU2AF1	POU domain, class 2, associating factor 1.	4
490	LY86	MD-1, RP105-associated.	4
491	C4BPB	complement component 4 binding protein, beta chain isoform 1 precursor.	4
492	CISH	cytokine-inducible SH2-containing protein.	4
493	CMTM7	CKLF-like MARVEL transmembrane domain containing 7 isoform a.	4
494	IL10RA	interleukin 10 receptor, alpha precursor.	4
495	IL9R	interleukin 9 receptor isoform 1 precursor.	4
496	IRAK4	interleukin-1 receptor-associated kinase 4.	4
497	ITFG1	integrin alpha FG-GAP repeat containing 1.	4
498	LILRA6	leukocyte immunoglobulin-like receptor, subfamily A, member 6.	4
499	SOCS1	suppressor of cytokine signaling 1.	4
500	SOCS3	suppressor of cytokine signaling 3.	4
501	SOCS5	suppressor of cytokine signaling 5.	4
502	MAPBPIP	mitogen-activated protein-binding protein-interacting protein.	4
503	ANPEP	membrane alanine aminopeptidase precursor.	3
504	CD79B	CD79B antigen isoform 1 precursor.	3
505	IGSF2	immunoglobulin superfamily, member 2.	3
506	THBD	thrombomodulin precursor.	3
507	CDH5	cadherin 5, type 2 preproprotein.	3
508	BSG	basigin isoform 1.	3
509	TSPAN7	tetraspanin 7.	3

510	GYPC	glycophorin C isoform 1.	3
511	PF4	platelet factor 4 (chemokine (C-X-C motif) ligand 4).	3
512	PRF1	perforin 1 precursor.	3
513	GZMB	granzyme B precursor.	3
514	FADD	Fas-associated via death domain.	3
515	NBN	nibrin isoform 2.	3
516	C1QA	complement component 1, q subcomponent, A chain precursor.	3
517	NCR1	natural cytotoxicity triggering receptor 1.	3
518	PAFAH1B3	platelet-activating factor acetylhydrolase, isoform Ib, gamma subunit 29kDa.	3
519	C3	complement component 3 precursor.	3
520	HLA-B	major histocompatibility complex, class I, B.	3
521	HLA-DRA	major histocompatibility complex, class II, DR alpha precursor.	3
522	IFI16	interferon, gamma-inducible protein 16.	3
523	IFIT3	interferon-induced protein with tetratricopeptide repeats 3.	3
524	MIF	macrophage migration inhibitory factor (glycosylation-inhibiting factor).	3
525	IL1F5	interleukin 1 family, member 5.	3
526	IL1F8	interleukin 1 family, member 8 isoform 1.	3
527	IL1F6	interleukin 1 family, member 6 (epsilon).	3
528	CLEC4D	C-type lectin domain family 4, member D.	3
529	IL1F9	interleukin 1 family, member 9.	3
530	CLEC7A	dendritic cell-associated C-type lectin 1 isoform b.	3
531	C1QL1	complement component 1, q subcomponent-like 1.	3
532	C1QTNF6	C1q and tumor necrosis factor related protein 6.	3
533	CMKLR1	chemokine-like receptor 1.	3
534	CMTM6	CKLF-like MARVEL transmembrane domain containing 6.	3
535	IFITM1	interferon induced transmembrane protein 1 (9-27).	3
536	IL1F7	interleukin 1 family, member 7 isoform 1 proprotein.	3
537	IL20RA	interleukin 20 receptor, alpha.	3
538	IL23A	interleukin 23, alpha subunit p19 precursor.	3
539	IL27	interleukin 27.	3
540	IL4I1	interleukin 4 induced 1 isoform 1 precursor.	3
541	PDCD1	programmed cell death 1 precursor.	3
542	TLR9	toll-like receptor 9 isoform A precursor.	3
543	TNFSF12	tumor necrosis factor (ligand) superfamily, member 12 precursor.	3
544	FANCC	Fanconi anemia, complementation group C.	3
545	FANCE	Fanconi anemia, complementation group E.	3

546	STX11	syntaxin 11.	3
547	NHEJ1	nonhomologous end-joining factor 1.	3
548	PSTPIP1	proline-serine-threonine phosphatase interacting protein 1.	2
549	CD3D	CD3D antigen, delta polypeptide isoform A precursor.	2
550	CD3G	CD3G gamma precursor.	2
551	LTBR	lymphotoxin beta receptor.	2
552	TNFRSF8	tumor necrosis factor receptor superfamily, member 8 isoform 2.	2
553	FCGR2B	Fc fragment of IgG, low affinity IIb, receptor for (CD32) isoform 3.	2
554	ITGA2B	integrin alpha 2b preproprotein.	2
555	GP9	glycoprotein IX (platelet).	2
556	PTPRCAP	protein tyrosine phosphatase, receptor type, C-associated protein.	2
557	CD52	CD52 antigen.	2
558	SELE	selectin E precursor.	2
559	FCGR1A	Fc fragment of IgG, high affinity Ia, receptor (CD64).	2
560	CEACAM6	carcinoembryonic antigen-related cell adhesion molecule 6 (non-specific cross reacting antigen).	2
561	FCAR	Fc alpha receptor isoform a precursor.	2
562	LAMP1	lysosomal-associated membrane protein 1.	2
563	IL1R2	interleukin 1 receptor, type II precursor.	2
564	TNFSF4	tumor necrosis factor (ligand) superfamily, member 4.	2
565	ADAM8	ADAM metallopeptidase domain 8 precursor.	2
566	BST1	bone marrow stromal cell antigen 1 precursor.	2
567	SELPLG	selectin P ligand.	2
568	CXCR3	chemokine (C-X-C motif) receptor 3.	2
569	TEK	TEK tyrosine kinase, endothelial precursor.	2
570	SLC4A1	solute carrier family 4, anion exchanger, member 1.	2
571	RHCE	Rhesus blood group, CcEe antigens isoform 1.	2
572	RHD	Rh blood group D antigen.	2
573	BLR1	Burkitt lymphoma receptor 1 isoform 1.	2
574	CCR10	CC chemokine receptor 10.	2
575	CCL5	small inducible cytokine A5 precursor.	2
576	CCL14	chemokine (C-C motif) ligand 14 isoform 1 precursor.	2
577	CCL23	small inducible cytokine A23 isoform CKbeta8-1 precursor.	2
578	XCL1	chemokine (C motif) ligand 1.	2
579	SCYE1	small inducible cytokine subfamily E, member 1.	2
580	LIF	leukemia inhibitory factor (cholinergic differentiation factor).	2
581	CASP6	caspase 6 isoform alpha preproprotein.	2

582	GZMK	granzyme K precursor.	2
583	PSMB10	proteasome beta 10 subunit proprotein.	2
584	PSMB5	proteasome beta 5 subunit.	2
585	CANX	calnexin precursor.	2
586	TRAF6	TNF receptor-associated factor 6.	2
587	LY96	MD-2 protein.	2
588	HRH4	histamine H4 receptor.	2
589	RFXAP	regulatory factor X-associated protein.	2
590	C4A	complement component 4A preproprotein.	2
591	MPO	myeloperoxidase.	2
592	TIMP1	tissue inhibitor of metalloproteinase 1 precursor.	2
593	CEBPE	CCAAT/enhancer binding protein epsilon.	2
594	HLA-DMA	major histocompatibility complex, class II, DM alpha precursor.	2
595	SOD1	superoxide dismutase 1, soluble.	2
596	TNFSF14	tumor necrosis factor ligand superfamily, member 14 isoform 1 precursor.	2
597	C1QC	complement component 1, q subcomponent, gamma polypeptide.	2
598	FCER1A	Fc fragment of IgE, high affinity I, receptor for; alpha polypeptide precursor.	2
599	HLA-DPB1	major histocompatibility complex, class II, DP beta 1 precursor.	2
600	BCAP31	B-cell receptor-associated protein 31.	2
601	NCF4	neutrophil cytosolic factor 4 (40kD) isoform 1.	2
602	B2M	beta-2-microglobulin precursor.	2
603	FCAMR	Fc receptor, IgA, IgM, high affinity.	2
604	PRDX6	peroxiredoxin 6.	2
605	BST2	bone marrow stromal cell antigen 2.	2
606	C1QL3	complement component 1, q subcomponent-like 3.	2
607	C1QL4	complement component 1, q subcomponent-like 4.	2
608	C1QTNF4	C1q and tumor necrosis factor related protein 4.	2
609	C7	complement component 7 precursor.	2
610	CD274	CD274 molecule.	2
611	CD3EAP	CD3E antigen, epsilon polypeptide associated protein.	2
612	CXCR7	chemokine orphan receptor 1.	2
613	CMTM3	chemokine-like factor superfamily 3.	2
614	IL10RB	interleukin 10 receptor, beta precursor.	2
615	IL17RC	interleukin 17 receptor C isoform 3 precursor.	2
616	IL1RAPL1	interleukin 1 receptor accessory protein-like 1.	2
617	IL1RAPL2	interleukin 1 receptor accessory protein-like 2.	2

618	LILRA2	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 2.	2
619	LTB4R	leukotriene B4 receptor.	2
620	PDCD1LG2	programmed cell death 1 ligand 2.	2
621	SIVA1	CD27-binding (Siva) protein isoform 1.	2
622	TNFRSF12A	tumor necrosis factor receptor superfamily, member 12A.	2
623	TNFSF15	tumor necrosis factor (ligand) superfamily, member 15.	2
624	TTRAP	TRAF and TNF receptor-associated protein.	2
625	FANCB	Fanconi anemia complementation group B.	2
626	SPINK5	serine peptidase inhibitor, Kazal type 5 precursor.	2
627	SLC35C1	solute carrier family 35, member C1.	2
628	ORAI1	hypothetical protein LOC84876.	2
629	FCGR3A	Fc fragment of IgG, low affinity IIIa, receptor for (CD16).	1
630	DPP4	dipeptidylpeptidase IV.	1
631	ITGA3	integrin alpha 3 isoform b, precursor.	1
632	C5AR1	complement component 5 receptor 1 (C5a ligand).	1
633	SLC3A2	solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform a.	1
634	IL8RB	interleukin 8 receptor beta.	1
635	PVR	poliovirus receptor.	1
636	KLRC1	killer cell lectin-like receptor subfamily C, member 1 isoform NKG2-A.	1
637	CCR7	chemokine (C-C motif) receptor 7 precursor.	1
638	PROCR	endothelial protein C receptor precursor.	1
639	MRC1	mannose receptor C type 1 precursor.	1
640	LAG3	lymphocyte-activation protein 3 precursor.	1
641	DARC	Duffy blood group.	1
642	ICAM4	intercellular adhesion molecule 4 isoform 3 precursor.	1
643	CD247	T-cell receptor zeta chain isoform 2 precursor.	1
644	CCR8	chemokine (C-C motif) receptor 8.	1
645	CXCL2	chemokine (C-X-C motif) ligand 2.	1
646	CXCL3	chemokine (C-X-C motif) ligand 3.	1
647	PPBP	pro-platelet basic protein precursor.	1
648	CXCL12	chemokine (C-X-C motif) ligand 12 (stromal cell-derived factor 1) isoform beta.	1
649	CX3CL1	chemokine (C-X3-C motif) ligand 1.	1
650	MBL2	soluble mannose-binding lectin precursor.	1
651	IL4	interleukin 4 isoform 1 precursor.	1
652	GZMA	granzyme A precursor.	1

653	TNFSF13	tumor necrosis factor ligand superfamily, member 13 isoform alpha proprotein.	1
654	TAPBP	tapasin isoform 1 precursor.	1
655	TIRAP	Toll-interleukin 1 receptor domain-containing adaptor protein isoform a.	1
656	PGLYRP1	peptidoglycan recognition protein 1.	1
657	TNFRSF17	tumor necrosis factor receptor superfamily, member 17.	1
658	RAG2	recombination activating gene 2.	1
659	BTK	Bruton agammaglobulinemia tyrosine kinase.	1
660	WASF3	WAS protein family, member 3.	1
661	AIRE	autoimmune regulator isoform 1.	1
662	CYBB	cytochrome b-245, beta polypeptide (chronic granulomatous disease).	1
663	CYBA	cytochrome b, alpha polypeptide.	1
664	C8A	complement component 8, alpha polypeptide precursor.	1
665	SERPING1	complement component 1 inhibitor precursor.	1
666	LYZ	lysozyme precursor.	1
667	MARCO	macrophage receptor with collagenous structure.	1
668	CFD	complement factor D preproprotein.	1
669	IFNG	interferon, gamma.	1
670	LTB4R2	leukotriene B4 receptor 2.	1
671	SDF2L1	stromal cell-derived factor 2-like 1 precursor.	1
672	PPP3R1	protein phosphatase 3, regulatory subunit B, alpha isoform 1.	1
673	PPP3R2	protein phosphatase 3 regulatory subunit B, beta isoform.	1
674	FCGRT	Fc fragment of IgG, receptor, transporter, alpha.	1
675	HLA-A	major histocompatibility complex, class I, A precursor.	1
676	HLA-C	major histocompatibility complex, class I, C precursor.	1
677	HLA-DPA1	major histocompatibility complex, class II, DP alpha 1 precursor.	1
678	HLA-F	major histocompatibility complex, class I, F isoform 2 precursor.	1
679	PRSS16	protease, serine, 16.	1
680	A2ML1	alpha-2-macroglobulin-like 1.	1
681	IL23R	interleukin 23 receptor precursor.	1
682	CADM2	immunoglobulin superfamily, member 4D.	1
683	PILRA	paired immunoglobulin-like type 2 receptor alpha isoform 1 precursor.	1
684	S100A8	S100 calcium-binding protein A8.	1
685	C5	complement component 5.	1
686	C9	complement component 9.	1
687	COLEC12	collectin sub-family member 12.	1

688	C4BPA	complement component 4 binding protein, alpha chain precursor.	1
689	CR1L	complement component (3b/4b) receptor 1-like.	1
690	CRLF1	cytokine receptor-like factor 1.	1
691	KIR2DL4	killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 4 isoform a.	1
692	KIR3DL2	killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 2 precursor.	1
693	TNFRSF11A	tumor necrosis factor receptor superfamily, member 11a precursor.	1
694	XCR1	XC chemokine receptor 1.	1
695	FANCF	Fanconi anemia, complementation group F.	1
696	IGHG2	immunoglobulin gamma-2 heavy chain.	1
697	OSTM1	osteopetrosis associated transmembrane protein 1.	1
698	FANCL	Fanconi anemia, complementation group L.	1
699	MS4A3	membrane-spanning 4-domains, subfamily A, member 3 isoform c.	0
700	MS4A5	membrane-spanning 4-domains, subfamily A, member 5.	0
701	FCGR2A	PREDICTED: similar to Low affinity immunoglobulin gamma Fc region receptor II-a precursor (Fc-gamma RII-a) (FcRII-a) (IgG Fc receptor II-a) (Fc-gamma-RIIa) (CD32 antigen) (CDw32).	0
702	CD38	CD38 antigen.	0
703	GP1BB	glycoprotein Ib, beta polypeptide precursor.	0
704	GP5	glycoprotein V (platelet).	0
705	CEACAM3	carcinoembryonic antigen-related cell adhesion molecule 3 precursor.	0
706	CEACAM5	carcinoembryonic antigen-related cell adhesion molecule 5 preproprotein.	0
707	PSG1	pregnancy specific beta-1-glycoprotein 1.	0
708	LILRB1	leukocyte immunoglobulin-like receptor, subfamily B, member 1 isoform 1.	0
709	THY1	Thy-1 cell surface antigen.	0
710	VCAM1	vascular cell adhesion molecule 1 isoform a precursor.	0
711	IL8RA	interleukin 8 receptor alpha.	0
712	PROM1	prominin 1.	0
713	TNFRSF9	tumor necrosis factor receptor superfamily, member 9 precursor.	0
714	PDGFRA	platelet-derived growth factor receptor alpha precursor.	0
715	F3	coagulation factor III precursor.	0
716	MCAM	melanoma cell adhesion molecule.	0
717	FUT3	fucosyltransferase 3.	0
718	CD177	CD177 molecule.	0
719	IGLL1	immunoglobulin lambda-like polypeptide 1 isoform a precursor.	0
720	CD207	CD207 antigen, langerin.	0
721	CLEC4M	C-type lectin domain family 4, member M isoform 1.	0

722	IL13RA2	interleukin 13 receptor, alpha 2 precursor.	0
723	GYPB	glycophorin B precursor.	0
724	BCAM	basal cell adhesion molecule isoform 2 precursor.	0
725	RHAG	Rh-associated glycoprotein.	0
726	ALK	anaplastic lymphoma kinase Ki-1.	0
727	CCBP2	chemokine binding protein 2.	0
728	CCL1	small inducible cytokine A1 precursor.	0
729	CCL2	small inducible cytokine A2 precursor.	0
730	CCL7	chemokine (C-C motif) ligand 7 precursor.	0
731	CCL8	small inducible cytokine A8 precursor.	0
732	CCL11	small inducible cytokine A11 precursor.	0
733	CCL13	small inducible cytokine A13 precursor.	0
734	CCL15	chemokine (C-C motif) ligand 15 precursor.	0
735	CCL16	small inducible cytokine A16 precursor.	0
736	CCL17	small inducible cytokine A17 precursor.	0
737	CCL18	small inducible cytokine A18 precursor.	0
738	CCL19	small inducible cytokine A19 precursor.	0
739	CCL20	chemokine (C-C motif) ligand 20.	0
740	CCL21	small inducible cytokine A21 precursor.	0
741	CCL22	small inducible cytokine A22 precursor.	0
742	CCL24	small inducible cytokine A24 precursor.	0
743	CCL25	small inducible cytokine A25 precursor.	0
744	CCL26	chemokine (C-C motif) ligand 26 precursor.	0
745	CXCL5	chemokine (C-X-C motif) ligand 5 precursor.	0
746	CXCL6	chemokine (C-X-C motif) ligand 6 (granulocyte chemotactic protein 2).	0
747	CXCL9	small inducible cytokine B9 precursor.	0
748	CXCL10	small inducible cytokine B10 precursor.	0
749	CXCL11	small inducible cytokine B11 precursor.	0
750	CXCL13	chemokine (C-X-C motif) ligand 13 (B-cell chemoattractant).	0
751	XCL2	chemokine (C motif) ligand 2.	0
752	CXCL14	small inducible cytokine B14 precursor.	0
753	MASP1	mannan-binding lectin serine protease 1 isoform 3.	0
754	IL2	interleukin 2 precursor.	0
755	IL3	interleukin 3 precursor.	0
756	IL11	interleukin 11 precursor.	0
757	CSF3	colony stimulating factor 3 isoform a precursor.	0
758	NOS2A	nitric oxide synthase 2A.	0

759	LTA	lymphotoxin alpha precursor.	0
760	GZMM	granzyme M precursor.	0
761	LPO	lactoperoxidase.	0
762	DEFA3	defensin, alpha 3 preproprotein.	0
763	DEFA5	defensin, alpha 5 preproprotein.	0
764	DEFA6	defensin, alpha 6 preproprotein.	0
765	POMC	proopiomelanocortin preproprotein.	0
766	PGLYRP2	peptidoglycan recognition protein 2 precursor.	0
767	IL12B	interleukin 12B precursor.	0
768	C8B	complement component 8, beta polypeptide preproprotein.	0
769	LILRB4	leukocyte immunoglobulin-like receptor, subfamily B, member 4 isoform 1.	0
770	GYPE	glycophorin E precursor.	0
771	GNLY	granulysin isoform NKG5.	0
772	EPO	erythropoietin precursor.	0
773	EPX	eosinophil peroxidase.	0
774	MICA	MHC class I chain-related gene A protein.	0
775	PLA2G7	phospholipase A2, group VII.	0
776	NFATC4	cytoplasmic nuclear factor of activated T-cells 4.	0
777	CHL1	cell adhesion molecule with homology to L1CAM precursor.	0
778	IFNB1	interferon, beta 1, fibroblast.	0
779	CTSG	cathepsin G preproprotein.	0
780	PRG2	proteoglycan 2 preproprotein.	0
781	EBF2	early B-cell factor 2.	0
782	CSF2	colony stimulating factor 2 precursor.	0
783	HRB	HIV-1 Rev binding protein.	0
784	DEFB4	defensin, beta 4 precursor.	0
785	HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1 precursor.	0
786	HLA-DQB2	major histocompatibility complex, class II, DQ beta 2.	0
787	HLA-DRB3	major histocompatibility complex, class II, DR beta 3 precursor.	0
788	HLA-DRB4	major histocompatibility complex, class II, DR beta 4 precursor.	0
789	HLA-DRB5	major histocompatibility complex, class II, DR beta 5 precursor.	0
790	HLA-E	major histocompatibility complex, class I, E precursor.	0
791	IL5	interleukin 5 precursor.	0
792	IL13	interleukin 13 precursor.	0
793	IL17A	interleukin 17A precursor.	0
794	KIR2DL1	killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 1.	0

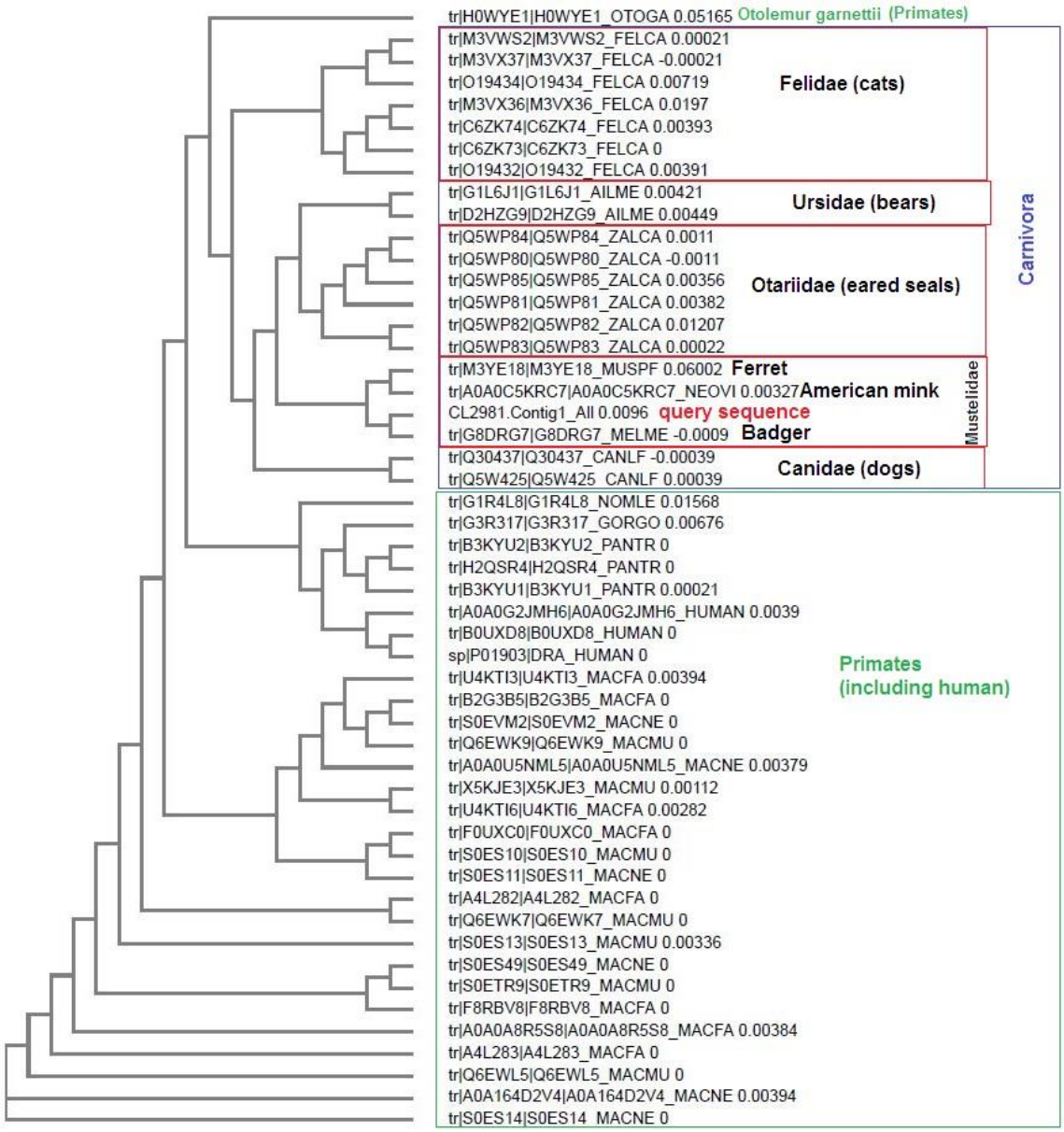
795	KIR2DL3	killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 3 isoform 1.	0
796	KIR2DS1	killer cell immunoglobulin-like receptor, two domains, short cytoplasmic tail, 1 precursor.	0
797	KIR2DS2	killer cell immunoglobulin-like receptor, two domains, short cytoplasmic tail, 2 precursor.	0
798	KIR2DS5	killer cell immunoglobulin-like receptor, two domains, short cytoplasmic tail, 5.	0
799	KIR3DL1	killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 1 precursor.	0
800	IL17B	interleukin 17B precursor.	0
801	CLEC10A	C-type lectin, superfamily member 14 isoform 2.	0
802	PGLYRP3	peptidoglycan recognition protein 3 precursor.	0
803	LEAP2	liver-expressed antimicrobial peptide 2 precursor.	0
804	DCD	dermcidin preproprotein.	0
805	WFDC12	WAP four-disulfide core domain 12 precursor.	0
806	CRP	C-reactive protein, pentraxin-related.	0
807	DEFA1	defensin, alpha 1 preproprotein.	0
808	DEFA4	defensin, alpha 4 preproprotein.	0
809	DEFB1	defensin, beta 1 preproprotein.	0
810	DEFB105A	defensin, beta 105A precursor.	0
811	DEFB106A	defensin, beta 106A precursor.	0
812	DEFB119	defensin, beta 119 isoform a precursor.	0
813	DEFB123	beta defensin 123 precursor.	0
814	LYG2	lysozyme G-like 2.	0
815	CLEC4E	C-type lectin domain family 4, member E.	0
816	HTN3	histatin 3.	0
817	IL6	interleukin 6 (interferon, beta 2).	0
818	IL7	interleukin 7 precursor.	0
819	INDO	indoleamine-pyrrole 2,3 dioxygenase.	0
820	IFIT1L	interferon-induced protein with tetratricopeptide repeats 1-like.	0
821	IL22	interleukin 22.	0
822	CLEC4A	C-type lectin, superfamily member 6 isoform 1.	0
823	DEFB103A	defensin, beta 103B precursor.	0
824	CCL3L1	chemokine (C-C motif) ligand 3-like 1 precursor.	0
825	IL25	interleukin 25 isoform 1 precursor.	0
826	FOXN1	forkhead box N1.	0
827	RNASE7	ribonuclease, RNase A family, 7.	0
828	CLEC6A	dectin-2.	0
829	C1QL2	complement component 1, q subcomponent-like 2.	0

830	C1QTNF2	C1q and tumor necrosis factor related protein 2.	0
831	C1QTNF3	C1q and tumor necrosis factor related protein 3 isoform a.	0
832	C1QTNF5	C1q and tumor necrosis factor related protein 5.	0
833	C1QTNF7	C1q and tumor necrosis factor related protein 7.	0
834	C1S	complement component 1, s subcomponent.	0
835	C3AR1	complement component 3a receptor 1.	0
836	CCL3L3	chemokine (C-C motif) ligand 3-like 3 precursor.	0
837	CCL4L1	chemokine (C-C motif) ligand 4-like 1 precursor.	0
838	CCL4L2	chemokine (C-C motif) ligand 4-like 2 precursor.	0
839	CD164L2	CD164 sialomucin-like 2.	0
840	CD200R2	CD200 cell surface glycoprotein receptor isoform 2.	0
841	CD300LB	CD300 molecule-like family member b.	0
842	SIGLEC15	sialic acid binding Ig-like lectin 15.	0
843	CFHR1	complement factor H-related 1.	0
844	CFHR2	H factor (complement)-like 3.	0
845	CFHR3	complement factor H-related 3.	0
846	CFHR4	complement factor H-related 4.	0
847	CFHR5	complement factor H-related 5.	0
848	CLEC4C	C-type lectin domain family 4, member C isoform 1.	0
849	CMTM1	chemokine-like factor superfamily 1 isoform 13.	0
850	CMTM2	chemokine-like factor superfamily 2.	0
851	CMTM4	chemokine-like factor superfamily 4 isoform 1.	0
852	CMTM5	chemokine-like factor superfamily 5 isoform c.	0
853	CMTM8	CKLF-like MARVEL transmembrane domain containing 8.	0
854	CRLF2	cytokine receptor-like factor 2 isoform 2.	0
855	CYTL1	cytokine-like 1.	0
856	FCGR2C	Fc fragment of IgG, low affinity IIc, receptor for isoform 2.	0
857	FGFR2	fibroblast growth factor receptor 2 isoform 1 precursor.	0
858	FGFR4	fibroblast growth factor receptor 4 isoform 1 precursor.	0
859	IL17F	interleukin 17F precursor.	0
860	IL1RL1	interleukin 1 receptor-like 1 isoform 2 precursor.	0
861	IL1RL2	interleukin 1 receptor-like 2 precursor.	0
862	IL21	interleukin 21.	0
863	IL22RA1	interleukin 22 receptor, alpha 1.	0
864	IL22RA2	interleukin 22-binding protein isoform 1.	0
865	IL26	interleukin 26 precursor.	0
866	IL28A	interleukin 28A.	0

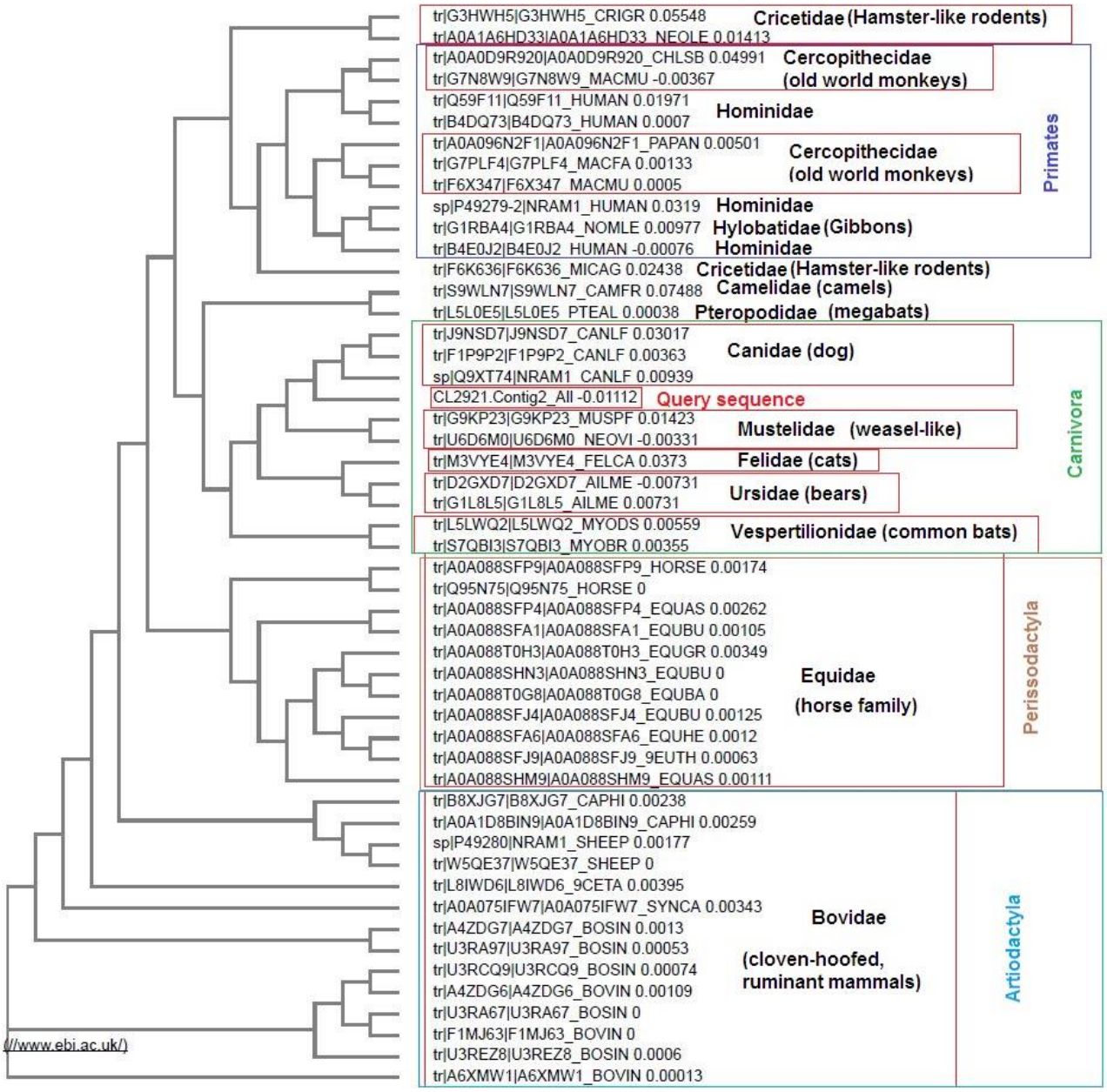
867	IL28B	interleukin 28B.	0
868	IL29	interleukin 29.	0
869	IL31	interleukin 31.	0
870	IL31RA	gp130-like monocyte receptor.	0
871	KDR	kinase insert domain receptor (a type III receptor tyrosine kinase).	0
872	KIR2DL2	killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 2 precursor.	0
873	KIR2DL5A	killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 5A.	0
874	KIR2DS4	killer cell immunoglobulin-like receptor, two domains, short cytoplasmic tail, 4.	0
875	KIR3DL3	killer cell immunoglobulin-like receptor, three domains, long cytoplasmic tail, 3.	0
876	KLRC2	killer cell lectin-like receptor subfamily C, member 2.	0
877	LAIR2	leukocyte-associated immunoglobulin-like receptor 2 isoform a.	0
878	LIFR	leukemia inhibitory factor receptor precursor.	0
879	LILRA1	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 1.	0
880	LILRA3	leukocyte immunoglobulin-like receptor, subfamily A (without TM domain), member 3.	0
881	LILRA4	leukocyte immunoglobulin-like receptor subfamily A member 4.	0
882	LILRB5	leukocyte immunoglobulin-like receptor, subfamily B, member 5 isoform 2.	0
883	NCR2	natural cytotoxicity triggering receptor 2.	0
884	SCGB3A1	secretoglobin, family 3A, member 1.	0
885	TNFRSF10A	tumor necrosis factor receptor superfamily, member 10a.	0
886	TNFRSF10B	tumor necrosis factor receptor superfamily, member 10b isoform 1 precursor.	0
887	TNFRSF10C	tumor necrosis factor receptor superfamily, member 10c precursor.	0
888	TNFRSF10D	tumor necrosis factor receptor superfamily, member 10d precursor.	0
889	CA2	carbonic anhydrase II.	0
890	ELA2	elastase 2, neutrophil preproprotein.	0
891	F12	coagulation factor XII precursor.	0
892	IGHM	FLJ00385 protein.	0
893	NLRP7	NACHT, leucine rich repeat and PYD containing 7 isoform 1.	0

8.7 Phylogenetic trees

8.7.1 HLA class II gene



8.7.2 NRAMP1


tr|G3HWH5|G3HWH5_CRIGR 0.05548
tr|A0A1A6HD33|A0A1A6HD33_NEOLE 0.01413
Cricetidae (Hamster-like rodents)
tr|A0A0D9R920|A0A0D9R920_CHLSB 0.04991
tr|G7N8W9|G7N8W9_MACMU -0.00367
Cercopithecidae (old world monkeys)
tr|Q59F11|Q59F11_HUMAN 0.01971
tr|B4DQ73|B4DQ73_HUMAN 0.0007
Hominidae
tr|A0A096N2F1|A0A096N2F1_PAPAN 0.00501
tr|G7PLF4|G7PLF4_MACFA 0.00133
tr|F6X347|F6X347_MACMU 0.0005
Cercopithecidae (old world monkeys)
sp|P49279-2|NRAM1_HUMAN 0.0319
Hominidae
tr|G1RBA4|G1RBA4_NOMLE 0.00977
Hylobatidae (Gibbons)
tr|B4E0J2|B4E0J2_HUMAN -0.00076
Hominidae
Primates
tr|F6K636|F6K636_MICAG 0.02438
Cricetidae (Hamster-like rodents)
tr|S9WLN7|S9WLN7_CAMFR 0.07488
Camelidae (camels)
tr|L5L0E5|L5L0E5_PTEAL 0.00038
Pteropodidae (megabats)
tr|J9NSD7|J9NSD7_CANLF 0.03017
tr|F1P9P2|F1P9P2_CANLF 0.00363
sp|Q9XT74|NRAM1_CANLF 0.00939
Canidae (dog)
CL2921.Contig2_All -0.01112
Query sequence
tr|G9KP23|G9KP23_MUSPF 0.01423
tr|U6D6M0|U6D6M0_NEOVI -0.00331
Mustelidae (weasel-like)
tr|M3VYE4|M3VYE4_FELCA 0.0373
Felidae (cats)
tr|D2GXD7|D2GXD7_AILME -0.00731
tr|G1L8L5|G1L8L5_AILME 0.00731
Ursidae (bears)
tr|L5LWQ2|L5LWQ2_MYODS 0.00559
tr|S7QBI3|S7QBI3_MYOBR 0.00355
Vespertilionidae (common bats)
Carnivora
tr|A0A088SFP9|A0A088SFP9_HORSE 0.00174
tr|Q95N75|Q95N75_HORSE 0
tr|A0A088SFP4|A0A088SFP4_EQUAS 0.00262
tr|A0A088SFA1|A0A088SFA1_EQUBU 0.00105
tr|A0A088T0H3|A0A088T0H3_EQUGR 0.00349
tr|A0A088SHN3|A0A088SHN3_EQUBU 0
tr|A0A088T0G8|A0A088T0G8_EQUBA 0
tr|A0A088SFJ4|A0A088SFJ4_EQUBU 0.00125
tr|A0A088SFA6|A0A088SFA6_EQUHE 0.0012
tr|A0A088SFJ9|A0A088SFJ9_9EUTH 0.00063
tr|A0A088SHM9|A0A088SHM9_EQUAS 0.00111
Equidae (horse family)
Perissodactyla
tr|B8XJG7|B8XJG7_CAPHI 0.00238
tr|A0A1D8BIN9|A0A1D8BIN9_CAPHI 0.00259
sp|P49280|NRAM1_SHEEP 0.00177
tr|W5QE37|W5QE37_SHEEP 0
tr|L8IWD6|L8IWD6_9CETA 0.00395
tr|A0A075IFW7|A0A075IFW7_SYNCA 0.00343
tr|A4ZDG7|A4ZDG7_BOSIN 0.0013
tr|U3RA97|U3RA97_BOSIN 0.00053
tr|U3RCQ9|U3RCQ9_BOSIN 0.00074
tr|A4ZDG6|A4ZDG6_BOVIN 0.00109
tr|U3RA67|U3RA67_BOSIN 0
tr|F1MJ63|F1MJ63_BOVIN 0
tr|U3REZ8|U3REZ8_BOSIN 0.0006
tr|A6XMW1|A6XMW1_BOVIN 0.00013
Bovidae (cloven-hoofed, ruminant mammals)
Artiodactyla
//www.ebi.ac.uk/)

8.7.3 TLR8 gene