A CRISPR-Cas9 screen for hepatocyte receptors for malaria parasite invasion

Juliana Jung-Mee Lee
MSc by Research in Clinical Medicine

Trinity College
Nuffield Department of Clinical Medicine
University of Oxford
Michaelmas 2020

Candidate no. 1029605
Word count: 29,120

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

More than 2 billion people are at risk of contracting malaria, and around half a million die every year. Malaria is transmitted by sporozoites from *Anopheles* female mosquitoes to human host during a blood meal. From the host dermal layer, sporozoites travel to the liver, traverse through multiple hepatocytes until stopping in one last cell. They then establish a productive cell invasion, followed by a parasitophorous vacuole formation. Inside this vacuole, sporozoites develop into thousands of merozoites which are later released by hepatocyte rupture. Then, merozoites travel to the blood vessel to invade erythrocytes where malarial clinical symptoms appear– as such, preventing the liver stage infection can protect infectees from clinical onset. However, the liver stage remains a poorly understood malaria stage today owing to its complex molecular cell invasion mechanism. Therefore, in order to understand how sporozoites establish a hepatocyte invasion and which host cell surface receptor-sporozoite surface ligand interaction is involved in this event, CRISPR-Cas9 screen technology was applied in this study. 480 genes were knocked out, and their impact on the sporozoite invasion was assessed. Initially, results showed *ITGAV, RPN1, ATP2B1, TMEM30A, FCGR2B, ITGB5, SLC35A2, MGAT1, EMC1* and *APOH* genes as hits. Amongst these, *ITGAV* and *ITGB5* (αVβ5 integrin) were investigated as αV-integrin is a known binding partner of TRAP. However, further experiments revealed that *ITGAV* and *ITGB5* hits were false positives induced by a lack of cell adhesion due to integrin disruptions rather than a direct binding activity with the sporozoites. Nevertheless, the CRISPR-Cas9 screen method used for this study remains a useful tool in studying hepatocyte-sporozoite protein interactions due to its success in identifying hepatocyte surface proteins, SR-B1 and ApoH, which are known to be important for the malaria liver stage. Therefore, this study presents an effective method to study parasite-host protein relationships by incorporating a sporozoite invasion assay with a CRISPR-Cas9 screen.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Dr. Sandy Douglas, for giving me a wonderful opportunity to conduct this research project and for sharing his in-depth knowledge and expertise over the past two years. I am also grateful to my co-supervisor, Dr. Alex Spencer, for providing her insights whenever I needed her help with experiments. I am indebted to Dr. Rameswara Segi Segireddy for his patience and mentorship through small and big experiments. A huge thank you to Dr. Adam Ritchie for listening to my stories, offering life-changing advice and revising my thesis. I would also like to thank Marta Ulaszewska for providing mosquitoes and Jon Webber for helping with FACS sort. Many, many thanks to all members of the Douglas Group – I will miss our lab outings. A big thank you to the entire Jenner CCMP family for providing moral support and free sweets. My life at Oxford would have been different without your life-changing chocolate. Thank you, Alka Ishwarbhai and Louisa Rose; I am going to miss our Oxford food adventures.

Last but not least, I would also like to thank my family and friends from around the world for their continuous support. Without them, I would not be where I am today. Thank you, Trinity College MCR family, you made my time in Oxford exciting, and I will never forget the memories we shared.

ABBREVIATIONS

WHO	World Health Organization
EEF	Exo-erythrocytic form
PV	Parasitophorous vacuole
TRAP	Thrombospondin repeat anonymous protein
CSP	Circumsporozoite protein
VWA	von Willebrand factor A
TSR	Thrombospondin type I repeat
HSPG	Heparan sulphate proteoglycan
HGF	Hepatocyte growth factor
TV	Transient vacuole
SPECT	Sporozoite microneme protein essential for cell traversal
PLP1	Perforin-like protein 1
SPECT2	Sporozoite microneme protein essential for cell traversal 2
CELTOS	Cell-traversal protein for ookinetes and sporozoites
GEST	Gamete egress and sporozoite traversal
PL	Phospholipase
CDPK6	Calcium-dependent protein kinase 6
EBL	Erythrocyte binding-like protein
RBL	Reticulocyte binding-like protein
RH5	Reticulocyte binding-like homologue 5
CyRPA	Cysteine-rich protective antigen
RIPR	RH5-interacting protein
BSG	Basigin
AMA1	Apical membrane antigen 1
RON	Rhoptry neck protein
ITN	Insecticide-treated bed nets
IRS	Indoor residual spraying
Bti	Bacillus thuringiensis israelensis
CRT	Chloroquine resistance reporter
RAS	Radiation-attenuated sporozoites
CRISPR	Clustered regularly interspaced short palindromic repeats
sgRNA	single-guide RNA
PAM	Protospacer adjacent motif
bp	base pair
DSB	Double-stranded break
NHEJ	Non-homologous end joining
NGS	Next Generation Sequencing
MAGeCK	Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout
RRA	Robust rank aggregation
MLE	Maximum likelihood estimation
MOI	Multiplicity of infection
AVEXIS	Avidity-based extracellular interaction screen
PCR	Polymerase chain reaction
QC	Quality control
FDR	False discovery rate
LFC	Log fold change
MFI	Median fluorescence intensity

RGD	Arginine-glycine-aspartate
COMP	Cartilage oligomeric matrix protein

CHAPTER 1. INTRODUCTION

1.1. MALARIA OVERVIEW

Malaria is a mosquito-borne, parasitic infection that is one of the leading causes of global morbidity and mortality. First described in China around 2700 B.C. and long associated with residing near marshes, the modern name malaria is derived from ‘mala aria’, an 18th century Italian term for ‘bad air’ (Cox, 2010; de Ridder, van der Kooy, & Verpoorte, 2008) . It was believed that malaria was caused by 'bad air' until the late 19th century when Charles Louis Alphonse Laveran observed parasites in a malaria patient's blood for the first time (Cox, 2010; Tuteja, 2007). Shortly after, Dr. Ronald Ross defined malaria as a vector-borne disease and Professor Giovanni Battista Grassi identified the *Anopheles* mosquitoes as the vector for human malaria (Cox, 2010; Tuteja, 2007).

According to the World Health Organization, around 2.5 billion people are at risk of contracting malaria, and nearly half a million die every year (World Health Organization [WHO], 2019). In 2018, 228 million malaria cases were reported worldwide – of these, 93% were in Africa and the *Plasmodium falciparum* species caused 99.7% of these African cases (WHO, 2019). The most vulnerable age group was 0-5 years, accounting for 67% of the total deaths in 2018 (WHO, 2019). Severe malaria, classified by anaemia, hypoglycemia and cerebral malaria, is seen in young children more often than adults due to their lack of acquired immunity (WHO, 2019). Among children under the age of 5, those exposed to the *Plasmodium falciparum* parasites in utero are more susceptible to malaria infection due to the development of their immunological tolerance to the parasites' soluble antigens (Malhotra et al., 2009).

1.1.1. P*LASMODIUM FALCIPARUM*

Malaria is caused by protozoan parasites of the genus Plasmodium, which is a part of a large parasitic phylum called Apicomplexa. *Plasmodium* parasites are host-specific and share similar life stages between two hosts – a mosquito vector and a vertebrate. Amongst hundreds of species identified so far, six are known to infect humans – *Plasmodium falciparum, P. vivax, P. malariae, P. ovale curtisi, P. ovale wallikeri and P. knowlesi* (Lalremruata et al., 2017). Of these six species that infect humans, *P. falciparum* is the most prevalent species. Its complex life cycle involves two parasitic stages in humans: liver and erythrocytic stages (Cowman, Healer, Marapana, & Marsh, 2016).

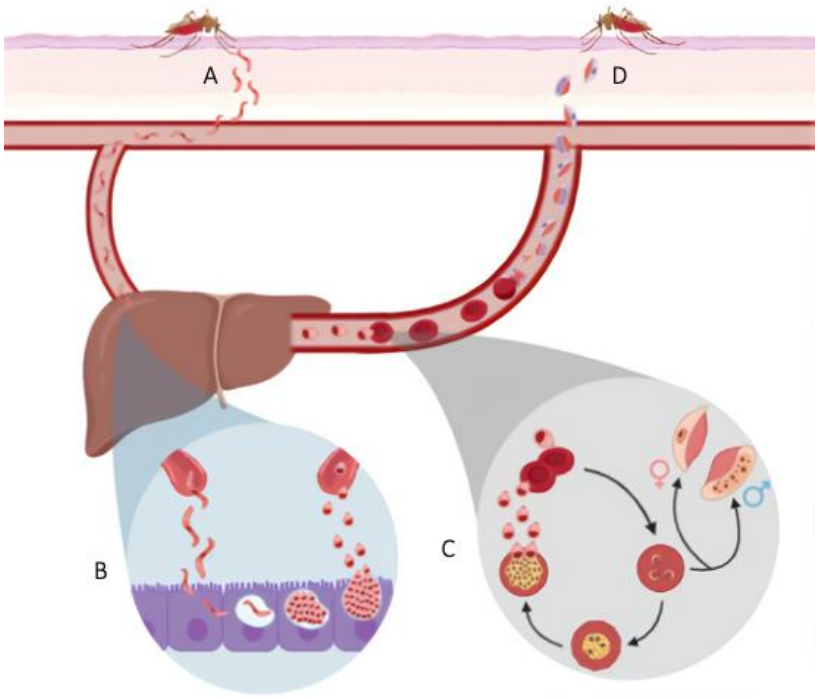


Figure 1-1. Summary of *Plasmodium falciparum* malaria infection pathway in humans. (A) An *Anopheles* female mosquito injects sporozoites into host dermal layer during a blood meal. Sporozoites traverse through the dermal layer and reach blood vessels. (B) Sporozoites arrive at liver and invade hepatocytes. They undergo schizogony to produce 20,000 merozoites. The invaded hepatocytes burst and merozoites are released into the blood stream. (C) Merozoites in the blood stream invade erythrocytes and either undergo an asexual reproduction to produce more merozoites or a sexual reproduction to develop into either microgametocytes (male) or macrogametocytes (female). (D) Microgametocytes and macrogametocytes enter the

blood stream. Mosquitoes take up these gametocytes during a blood meal. Figure created with BioRender.com.

During a blood meal, mosquitoes inject sporozoites from their salivary glands into human skin (Cowman et al., 2016) (Figure 1-1). These sporozoites migrate through the skin cells to reach blood vessels and then migrate to liver (Cowman et al., 2016). Once they reach the liver sinusoids, they arrest, traverse through Kupffer cells and invade hepatocytes, therefore marking the beginning of 'liver stage'. This stage takes approximately ten days, during which *P. falciparum* sporozoites undergo schizogony inside the hepatocytes to multiply and produce merozoites (Cowman et al., 2016).

Following the release of merozoites from the invaded hepatocytes, they infect erythrocytes, thereby entering 'erythrocytic stage' (Cowman et al., 2016). Malarial symptoms, including high body temperature, vomiting and diarrhea, appear once the merozoites begin development in the erythrocytes (Ngotho et al., 2019). The erythrocytic stage can be divided further into schizogony and asexual reproduction stages (Ngotho et al., 2019). During schizogony, merozoites undergo a cell division to produce more merozoites inside the invaded erythrocytes (Ngotho et al., 2019). These newly produced merozoites can either undergo schizogony again or asexually develop into either microgametocytes or macrogametocytes, which are characterized as male and female, respectively (Ngotho et al., 2019). When the invaded erythrocytes burst, immature gametocytes travel to bone marrow and spleen through a vascular barrier (De Niz et al., 2018). Once arrived, they develop into mature gametocytes and then subsequently exit through the vascular barrier to re-enter the peripheral blood circulation (De Niz et al., 2018).

The release of these mature gametocytes marks the end of the human host cycle, and a mosquito vector cycle begins from a blood meal (Cowman et al., 2016). When taking a blood meal, an *Anopheles* female mosquito ingests both sexes of gametocytes, and the gametocytes develop into gametes in the mosquito midgut (Cowman et al., 2016). The female gametes, called microgametes, are fertilized by male gametes, called macrogametes, to produce zygotes (Cowman et al., 2016). Within 24 hours, these zygotes become ookinetes, moving to the midgut wall to transform into oocysts (Cowman et al., 2016). Inside these oocysts, sporozoites are produced by asexual reproduction, and upon an oocyst rupture, the sporozoites are released into hemocoel (Cowman et al., 2016). Next, they travel to the salivary gland, waiting to invade the next human host (Cowman et al., 2016).

1.1.2. LIVER STAGE

On average, 123 sporozoites are injected into the host dermal layer from an infected *Anopheles* mosquito (Medica & Sinnis, 2005). They then leave the dermal layer and arrive in the space of Disse, an area inside the liver between hepatocytes and sinusoidal endothelium (Prudencio, Rodriguez, & Mota, 2006). Then, they penetrate multiple hepatocytes during cell traversal before arresting in one to establish a productive invasion, which is marked by schizogony and exoerythrocytic form (EEF) development in a parasitophorous vacuole (PV) (Meis et al., 1986). Less than 0.1% of the penetrated sporozoites reach this final stage and each successful sporozoite develops into 20,000 merozoites which are the causal agents of malarial symptoms (Meis et al., 1986). Thus, liver stage bottleneck is the key to controlling malaria (Sinnis, De La Vega, Coppi, Krzych, & Mota, 2013) and a promising target for malaria prevention as it protects infectees from clinical onset.

To understand the liver stage, studying the sporozoites' locomotion is critical as different techniques of locomotion provide more information on the stages of malaria infection and how productive hepatocyte invasions occur. Broadly, the sporozoites' locomotion can be categorized as gliding motility, cell traversal and cell invasion. They utilize these locomotion techniques interchangeably from the dermal layer to the liver sinusoids until they arrive at their final hepatocyte destination.

1.1.1.1. GLIDING MOTILITY

During a blood meal, sporozoites translocate from a mosquito salivary gland to host dermal layer by gliding motility. Soon after the injection, the host responds with two stages of inflammatory responses (Mac-Daniel et al., 2014). First, neutrophils are rapidly recruited to the injury site, and their numbers increase for up to 2 hours and then gradually decrease for the next 22 hours (Mac-Daniel et al., 2014). The second stage begins at around 24 hours post injection, where the number of inflammatory monocytes increases in both the dermal layer and the draining lymph node (Mac-Daniel et al., 2014). Owing to the rapid and sustained host immune response, the number of total sporozoites in the dermal layer starts to decrease 3 hours post infection, dropping by 80-90% (Yamauchi, Coppi, Snounou, & Sinnis, 2007). Studies have found that many sporozoites cannot successfully exit the dermal layer and instead end up lost in the lymphatic vessels (Hopp et al., 2015; Yamauchi et al., 2007). Only those that reach blood vessels in the dermis can migrate to the liver (Amino et al., 2006).

Sporozoites use gliding motility to move faster than the immune cells and survive in the dermal layer (Amino et al., 2006; Ejigiri et al., 2012). In the dermis, sporozoites can move up to 3 µm/s while phagocytes, on average, move at 0.1 µm/s (Amino et al., 2006). Observations on

an *in vitro* 2D plane showed the sporozoites' gliding motility pattern to be circular (Amino et al., 2006). However, *in vivo* experiments showed that the motility pattern varies by the local microenvironment. For example, Hellmann et al. demonstrated with a rodent parasite, *P. berghei*, that a substantial portion of the sporozoites moved linearly in the tail of a mouse but meandered in the ear pinna (Hellmann et al., 2011). Sporozoites' movements were also often observed to be interrupted by non-motile phases due to host substrates that are bound on sporozoites, momentarily halting them from moving forward (Hellmann et al., 2011).

Gliding motility can be further categorized as avascular motility and perivascular motility (Hopp et al., 2015). Avascular motility is demonstrated during the first 20 mins post dermal injection (Hopp et al., 2015). Sporozoites move linearly at high speed while maximizing coverage of the dermal cells explored (Hopp et al., 2015). At 1-2 hrs post injection, sporozoites begin to decelerate, change their movement to a circular pattern, and eventually come to an arrest (Hopp et al., 2015). These events are termed as perivascular motility in which the sporozoites glide along the capillary vessel's curvature to maximize their contact with the capillaries until the best entry region is found (Hopp et al., 2015).

Gliding motility is also observed when sporozoites exit the dermal layer and enter blood vessels. Studies have shown that this process is mostly modulated by two micronemal adhesin proteins, thrombospondin repeat anonymous protein (TRAP) and circumsporozoite protein (CSP) (Robson et al., 1995). TRAP is a type I transmembrane ligand with two ectodomains: tandem von Willebrand factor A (VWA) and thrombospondin type I repeat (TSR) (Song, Koksal, Lu, & Springer, 2012). While CSP is expressed at all parasitic pre-erythrocytic development stages, TRAP appears when sporozoites become more host-infective in the salivary gland, marked by maturing of the micronemal proteins in the apical end (Robson et al., 1995).

The cytoplasmic domain of TRAP plays essential roles in both gliding motility and cell invasion as it binds to aldolase, a filamentous F-actin-bound glycolytic enzyme, which in turn binds to myosin A (Jewett & Sibley, 2003). Furthermore, TRAP's C-terminal domain contains a conserved tryptophan at the second last residue, a site found in all parasites from the Apicomplexa phylum (Jewett & Sibley, 2003; Kappe et al., 1999). When tryptophan and asparagine (the last residue) were mutated to alanine and serine, respectively, a faulty forward movement of the sporozoites was observed (Kappe et al., 1999).

At physiological pH, TRAP's extracellular domain is positively charged and interacts with heparan sulphate proteoglycans (HSPGs), which are negatively charged glycans on a cell surface (Song et al., 2012). TRAP sticks to the HSPGs and then slips when its transmembrane domain is cleaved by a rhomboid protease and is translocated to the posterior end of the parasite (Ejigiri et al., 2012; Song et al., 2012). Song et al. described this repetitive movement as 'stick-and-slip' (Song et al., 2012). The importance of the TRAP rhomboid motif in the sporozoites' movement was demonstrated with a mutational study by Ejigiri et al. (Ejigiri et al., 2012). When the motif was mutated from AAGGIIGG to VALLIGV, the mutated sporozoites' movement speed decreased to 0.5 μm/s. Consequently, the infectivity rate was reduced (Ejigiri et al., 2012) with narrower parasite-to-dermal cell coverage (Hopp et al., 2015).

CSP is an important sporozoite surface protein involved in both motility and infectivity. CSP has a central tetrapeptide repeat region unique to each *Plasmodium* species, and two conserved domains that flank this repeat region – region I at the N-terminal and TSR at the C-terminal (Coppi et al., 2007). TSR is a cell-adhesion domain, which is usually masked by the region I domain (Coppi et al., 2007). The proteolytic cleavage of region I reveals TSR, and this conformation is referred to as adhesive conformation (Coppi et al., 2007). TSR mutation

studies showed that mutant sporozoites with their TSR consistently exposed cannot leave the dermal layer when injected intradermally; however, sporozoites were capable of infecting hepatocytes when injected intravenously (Coppi et al., 2007). These results suggest that the exposed TSR domain plays a direct role in the infection, while its hidden state supports motility (Coppi et al., 2007). This finding is consistent with the observation of CSP undergoing conformational changes only when the sporozoite has stopped and reached the hepatocytes in the space of Disse (Amino et al., 2006; Coppi et al., 2007).

1.1.1.2. CELL TRAVERSAL AND INVASION

Cell traversal is a repeated process of entering on one side, crossing, and exiting on the other side of several cells. This process is quick, and the types of cells that sporozoites traverse are epithelial cells, Kupffer cells and hepatocytes (S. Yang, Li, & Li, 2016).

In the dermal layer, sporozoites traverse through epithelial cells to reach blood vessels (A. S. Yang & Boddey, 2017). Then, these sporozoites travel to liver sinusoid and there, they traverse through Kupffer cells and endothelial cells before reaching hepatocytes (Pradel, Garapaty, & Frevert, 2004; Tavares et al., 2013). Although sporozoites traverse both Kupffer cells and endothelial cells, it is thought that passaging through Kupffer cells is the only known step necessary for subsequent hepatocyte invasion. The relationship between the Kupffer cell traversal and the hepatocyte invasion was observed by infecting *op/op* mice, which have a mutation in macrophage colony-stimulating factor (Baer et al., 2007). These mice lack Kupffer cells in the liver, and Baer et al. observed that the liver infection by *P. yoelii*, a rodent sporozoite species, is 84% lower than the wildtype when the cDNA ratio of the parasite to host hepatocyte was quantified by qRT-PCR (Baer et al., 2007). Although traversing the Kupffer

cells may be necessary for liver infection, which was recorded to occur in 68% of the sporozoites' total traversal events, Kupffer cells are not necessary for sporozoites to cross the sinusoidal barrier (Tavares et al., 2013). Sporozoites can traverse through the endothelial cells or between a Kupffer cell and an endothelial cell, or between two endothelial cells (Tavares et al., 2013). However, these traversal events are slower than the ones through Kupffer cells (Tavares et al., 2013) and it is unclear whether liver infectivity rate is affected by the absence of Kupffer cell traversal.

Sporozoites traverse through an unidentified number of hepatocytes before reaching their final destination and undergo a productive cell invasion event which is defined PV and EEF development (Tavares et al., 2013). It was previously believed that sporozoites invade hepatocytes first by penetrating through the cell membrane during traversal, leading to cell death caused by membrane rupture (Ishino, Yano, Chinzei, & Yuda, 2004). However, Risco-Castillo et al. have discovered that rather than penetration, sporozoites form transient vacuoles (TVs) inside the traversed cell, before moving through the cytoplasm and exiting at the apical end of the cell (Risco-Castillo et al., 2015). During this traversal process, hepatocyte growth factor (HGF) is upregulated in the hepatocytes, which helps in priming the sporozoites for their final productive cell invasion (Tavares et al., 2013).

Essential parasite proteins identified so far to be involved in cell traversal are sporozoite microneme protein essential for cell traversal (SPECT) and perforin-like protein 1 (PLP1), which is also known as SPECT2 (Ishino et al., 2004). Disrupting either of these genes did not affect productive hepatocyte invasion or EEF development in *in vitro* studies; however, sporozoites completely lost the ability to traverse cells prior to the productive invasion when rodents were injected intravenously (Ishino et al., 2004). Although it is unclear how sporozoites

can traverse through Kupffer cells without being phagocytosed, it has been suggested that the SPECT2 protein may play a protective role as when its gene was knocked out, sporozoites could enter but failed to exit the Kupffer cells and were subsequently destroyed (Tavares et al., 2013).

During cell traversal, both SPECT2 and lysosomal pH promote sporozoite egress from TVs (Risco-Castillo et al., 2015). Low pH activates SPECT2 to form pores in the host cell membrane for sporozoites to enter and traverse through the cell (Risco-Castillo et al., 2015). However, *in vitro* invasion experiments with SPECT2$^{-/-}$ sporozoites showed that the lack of SPECT2 did not affect productive cell invasion (Risco-Castillo et al., 2015). The purpose of traversing several hepatocytes remains puzzling as SPECT2$^{-/-}$ sporozoites demonstrated a higher infectivity rate than wildtype sporozoites (Risco-Castillo et al., 2015). Furthermore, inhibiting lysosomal acidification led to an increase in the number of productively invaded cells (Risco-Castillo et al., 2015).

Cell-traversal protein for ookinetes and sporozoites (CelTOS) was identified as another micronemal protein required for cell traversal (Kariu, Ishino, Yano, Chinzei, & Yuda, 2006). Although knocking out CelTOS inhibited the sporozoites' ability to exit the host cell and continue traversal, the CelTOS$^{-/-}$ sporozoites could enter the host cell through the TV (Kariu et al., 2006). When injected intravenously, CelTOS$^{-/-}$ sporozoites demonstrated a retardation in their EEF development and, subsequently, erythrocyte invasion (Kariu et al., 2006). However, depleting Kupffer cells with clodronate liposomes re-established liver infectivity in CelTOS$^{-/-}$ sporozoites, suggesting that CelTOS is required for sporozoites to traverse through the sinusoidal endothelium layer to reach the hepatocytes (Kariu et al., 2006). Other proteins that also play a role in cell traversal include gamete egress and sporozoite traversal (GEST) (Talman

et al., 2011) and phospholipase (PL) (Bhanot, Schauer, Coppens, & Nussenzweig, 2005). They facilitate sporozoite progression from the dermal layer to the bloodstream (Bhanot et al., 2005; Talman et al., 2011).

It is unknown why sporozoites traverse through several hepatocytes before stopping in one and how they can convert from traversal to productive invasion mechanism. There are three discussed theories:

1. Exocytosis at the apical end of the hepatocytes primes the sporozoites to become more competent for a productive invasion (Mota et al., 2001).
2. HGF is released from the hepatocytes during traversal, which induces neighbouring hepatocytes to become more susceptible to productive invasion (Carrolo et al., 2003).
3. The extracellular environment directs the sporozoites to switch their mechanism from cell traversal to cell invasion (Coppi et al., 2007).

While different experiments may point toward different theories, all three processes may be required for a productive invasion. Hepatocytes express HSPGs at a higher level than other cells, such as those in the dermis, blood vessels or the liver sinusoids. Both CSP and TRAP bind to HSPGs, and Coppi et al. suggested that interacting with HSPGs at a higher level encourages sporozoites to transition to a productive invasion mode from cell traversal (Coppi et al., 2007). Calcium-dependent protein kinase 6 (CDPK6), an enzyme activated by high calcium ion (Ca^{2+}) concentration, was also shown to be involved in this transition, which hints that Ca^{2+} also plays a role (Coppi et al., 2007). Consistent with this, increasing the Ca^{2+} level inside the traversed cell promotes the apical exocytosis of sporozoites and a higher rate of hepatocyte infection (A. S. Yang & Boddey, 2017). In parallel, the intracellular potassium level has a positive correlation with the rate of productive invasion (Kumar et al., 2007). Its high

level increases the rate of productive invasion while decreasing traversal by inhibiting the release of the micronemal proteins involved in traversal (Kumar et al., 2007). Lastly, while these are occurring, HGFs are released to the extracellular environment, consequently priming the neighbouring hepatocytes for a productive invasion (Carrolo et al., 2003).

Although both TVs and PVs serve a similar role of accompanying sporozoites, they can be distinguished by two features: 1) formation of moving junctions, and 2) membrane composition of the vacuoles (Risco-Castillo et al., 2015). PV moving junctions, a complex of AMA1 and rhoptry protein complexes, RON2/4/5, are released before the PV is formed during a cell invasion (Lamarque et al., 2011). TVs on the other hand, there is no evidence of moving junction releases during the TV formation. However, this does not rule out the possibility that rhoptry proteins are nevertheless involved in the TV formation (Risco-Castillo et al., 2015) as both TV and PV serve similar roles. Lastly, in the PV membrane, host cell membrane proteins are depleted to prevent parasite degradation; however, no such event is observed in the TV membranes (Risco-Castillo et al., 2015).

1.1.3. ERYTHROCYTIC STAGE

Sporozoites that productively invade a hepatocyte form a schizont inside a PV to undergo an asexual reproduction, known as schizogony, and produce merozoites. Then, these merozoites are released from the hepatocytes and enter the bloodstream to invade and replicate inside erythrocytes, therefore causing malaria clinical symptoms (Prudencio et al., 2006). These sequential events are mediated by proteins released from micronemes, rhoptries and dense granules, which are apical organelles found in all stages of the *Plasmodium* parasites (Baum, Gilberger, Frischknecht, & Meissner, 2008).

Low-affinity protein interactions between the merozoite surface and the erythrocyte surface initiate a reorientation of the merozoite's apical end towards the erythrocyte (Cowman et al., 2016; Miller, Baruch, Marsh, & Doumbo, 2002; Prudencio et al., 2006). This reorientation process is mediated by a release of apical organelle proteins such as erythrocyte binding-like (EBL) and reticulocyte binding-like (RBL) surface ligands (Knuepfer et al., 2019). Following the reorientation, merozoites release moving junction proteins to invade and form a PV within the erythrocyte (Baum et al., 2008; Greenbaum et al., 2002; Miller et al., 2002). It should be noted that PVs are different to phagolysosomes in that there are no host cell proteins present inside (Lingelbach & Joiner, 1998). Similar to the PVs formed during liver invasion, PVs protect merozoites from acidification and prolongs parasitic survival by acting as a barrier from the host cell's proteases in the cytosol (Lingelbach & Joiner, 1998).

During the binding of merozoite to erythrocytes, reticulocyte binding-like homologue 5 (RH5) (Baum et al., 2009), cysteine-rich protective antigen (CyRPA) (Reddy et al., 2015) and RH5-interacting protein (RIPR) are released from the merozoites (Knuepfer et al., 2019). RH5 is a rhoptry protein that binds to host basigin (BSG) receptor on the erythrocyte surface (Volz et al., 2016). Every parasite has two rhoptries at the apical end, and there are two distinct regions known as a neck and a bulb which are composed of different structures and proteins (Counihan, Kalanon, Coppel, & de Koning-Ward, 2013; Sherling et al., 2017). As the name suggests, the rhoptry neck region is long and thin, while the bulb region is broader and shorter. These compartmental rhoptry proteins have different roles related to their parasite location. Because rhoptries are connected to the apical membrane by their necks, neck proteins are involved in the initial stages of a productive invasion by forming moving junctions while bulb proteins are associated with PV membrane in the post-cell entry stages (El Hajj et al., 2007; Peixoto et al., 2010). *Plasmodium* rhoptry bulb proteins' functionalities are not yet well known; however,

their roles can be predicted by referencing a different Apicomplexan parasite, *Toxoplasma gondii*, which invade intestinal epithelial cells in humans (Pittman & Knoll, 2015). In *T. gondii*, rhoptry bulb proteins are kinases that play a role in aiding the parasite's intracellular survival by interfering with the host cell's signalling pathways (Peixoto et al., 2010) and immunity (Yamamoto & Takeda, 2012).

Although RH5 is an RBL surface ligand, it is not initially released before the merozoite reorientation stage (Weiss et al., 2015). RH5 marks *P. falciparum* parasites' host-specificity as they cannot bind to a BSG receptor of different species such as chimpanzee or gorilla (Beeson et al., 2016; Wanaguru, Liu, Hahn, Rayner, & Wright, 2013). CyPRA and RIPR are micronemal proteins that form a heterotrimeric complex with RH5 (Reddy et al., 2015), and they improve the binding affinity of RH5 to BSG (Wong et al., 2019). Unlike RH5, CyPRA and RIPR are conserved proteins found across *Plasmodium* species. Reddy et al. initially reported CyPRA to have a GPI-anchor at its C-terminal when labelled with glucosamine and that its role is to tether to the erythrocyte membrane for the RH5-BSG interaction to occur (Reddy et al., 2015). However, the anchoring mechanism of the RH5-CyPRA-RIPR complex remains unclear. Volz et al. refuted this claim by cleaving this GPI anchor with hydrofluoric acid but found no cleaved phosphoethanolamine peptides (Volz et al., 2016). Instead, they suggested a potential fourth member of the RH5-CyPRA-RIPR complex (Volz et al., 2016).

The binding of an RH5-CyPRA-RIPR complex to a BSG receptor activates a pore formation in the erythrocyte membrane, leading to the influx of Ca^{2+} along with other apical organelle proteins from the rhoptries (Volz et al., 2016). Role of this Ca^{2+} release is unknown; however, invasion failures of RIPR-, CyRPA- or RH5-deficient merozoites do not rule out the necessity of Ca^{2+} pore formation in the invasion process (Volz et al., 2016; Weiss et al., 2015).

Following the Ca^{2+} release, moving junctions are formed by AMA1, a micronemal type I transmembrane protein, and a RON2/4/5 complex from the rhoptries (Shen & Sibley, 2012). AMA1 is integrated into the erythrocyte membrane, and the hydrophobic cleft of its ectodomain binds to the C-terminal loop of RON2 within the complex in the merozoite membrane, and they are responsible for the ring formation around the invaginated merozoite (Alexander, Arastu-Kapur, Dubremetz, & Boothroyd, 2006; Beeson et al., 2016; Riglar et al., 2013; Shen & Sibley, 2012; Srinivasan et al., 2011). RON4 and RON5 bind to the cytosolic compartment of RON2 (Beeson et al., 2016). The anchored moving junction work with the actin-myosin motor; as the moving junctions move toward the merozoite's posterior end, the merozoite moves closer into the erythrocyte (Miller et al., 2002). In the end, the host cell membrane is resealed, and a PV is formed around the parasite (Greenbaum et al., 2002; O'Donnell & Blackman, 2005). The co-operative mechanism of the AMA1-RON complex and the actin-myosin motor remains unclear. The amino acid sequence of the AMA1 cytosolic tail indicates a possible interaction with aldolase, a glycolytic enzyme that forms a bridge between adhesins, such as TRAP and actin (Jewett & Sibley, 2003). However, this is a bioinformatic analysis and there is no clear experiment-based evidence of an AMA1-aldolase-TRAP interaction.

Of note, there are suggestions that some microneme and rhoptry proteins are exocytosed from merozoites to identify non-invaded, healthy erythrocytes before invasion (Lingelbach & Joiner, 1998). Furthermore, after the invagination of merozoites completes, the ligand-receptor complexes involved in the parasite entry become degraded by serine proteases in the micronemes. This marks the end of erythrocyte invasion, all occurring within a few seconds (Miller et al., 2002; Wright & Rayner, 2014).

The last organelle to release its contents during a cell invasion is dense granules, which induces PV formation and host cell remodelling by releasing their proteins into the infected erythrocyte post invasion (Baum et al., 2009). Compared to the other two secretory organelles, the role of dense granules has been less extensively studied – there is more known about *T. gondii* dense granules than *Plasmodium*. In *T. gondii*, dense granule proteins contain N-terminal signal peptides and are associated with the initial stage of PV membrane establishment (Lingelbach & Joiner, 1998). However, dense granule proteins in *P. falciparum* lack such N-terminal signal peptides (Lingelbach & Joiner, 1998). Consequently, they cannot be translocated across the ER membrane and are instead exocytosed by vesicles (Lingelbach & Joiner, 1998). Yet, *P. falciparum* dense granule proteins are associated with the PV membrane and structure after the PV is formed, similar to the *T. gondii* dense granule proteins (Lingelbach & Joiner, 1998). Precise roles of these dense granule proteins are unknown and must be studied further (Lingelbach & Joiner, 1998).

1.2. ANTIMALARIAL TREATMENTS AND VACCINES

Malaria treatments have existed as early as the 17th century where quinine was discovered from cinchona trees and used as a standard treatment in South America (Achan et al., 2011). Since the WHO announced its plans for malaria eradication in 1955, scientists have accelerated the development of more malaria treatments and vaccine, plus insecticides and other vector controls. However, these measures are being jeopardized by the development of parasite resistance against them.

Widely used malaria prevention methods include insecticide-treated bed nets (ITNs) and indoor residual spraying (IRS) (Alonso et al., 1991). ITNs are cost-effective, accessible and easy to use, making it an attractive preventive measure for many rural communities (Alonso et al., 1991). WHO recommends using insecticide pyrethroids as it is capable of killing mosquitoes and repelling them and disrupting mosquito feeding (Alonso et al., 1991). The use of pyrethroid-coated ITNs has been reported to reduce the overall mortality rate of young children who slept with the ITNs by a third compared to those who did not (Alonso et al., 1991). However, decades of pyrethroid-coated ITN usage have resulted in resistant mosquito strains (Strode, Donegan, Garner, Enayati, & Hemingway, 2014). Although this issue prompts the need to develop new insecticides, there remain concerns about developing new resistant strains due to the potential selective pressure it puts on mosquitoes (Takken & Knols, 2009).

While ITNs are used to protect sleeping areas, IRS is used to coat indoor house surfaces and walls with an insecticide (Fullman, Burstein, Lim, Medlin, & Gakidou, 2013). Although it does not directly protect humans from being bitten, IRS kills any mosquitoes that land on these coated surfaces (Fullman et al., 2013). It is often used together with ITNs to improve the

efficacy of malaria prevention strategies (Fullman et al., 2013). According to a multinational study conducted by Fullman et al., the IRS provides better protection in areas with medium to low transmission risks than ITNs (Fullman et al., 2013). On the other hand, ITNs are more effective in the high to medium transmission areas than IRS (Fullman et al., 2013). Together, ITNs and IRS provide an enhanced effect in significantly reducing parasitaemia in children living in the high to medium transmission communities (Fullman et al., 2013). Notwithstanding, IRS is not routinely used in endemic areas due to its high cost and harmful effects on the environment (Centers for Disease Control and Prevention [CDC], 2019).

Aside from using chemical tools, biological tools have been proposed to directly control mosquito larvae as a better strategy to reduce the development of resistance in *Anopheles* mosquitoes (Takken & Knols, 2009). For example, Kreogar et al. tested the efficacy of *Bacillus thuringiensis israelensis* (Bti) as a larvicide and observed a significant decrease in adult mosquito densities (Kroeger, Horstick, Riedl, Kaiser, & Becker, 1995). Bti are soil bacteria that produce endotoxins that kill certain insects, such as mosquito and blackfly larvae (Kroeger et al., 1995). Although this biological larvicide is useful, its widespread application is challenged by 1) its reduced activation capacity by anti-endotoxin proteins produced by other micro-organisms, and 2) its high density, preventing it from floating on water where the larvae reside for a prolonged period (Kroeger et al., 1995).

Chloroquine, for example, was introduced in 1934 for treating and preventing P. falciparum infections. Its less toxic version, called hydroxychloroquine, has also been widely used, and its activity is on par with chloroquine (Warhurst, Steele, Adagu, Craig, & Cullander, 2003). When ingested, chloroquine diffuses through the red blood cell membrane and into the PV (Chugh et al., 2013). Due to the acidic environment of the PV, chloroquine becomes

protonated, unable to leave the PV membrane (Chugh et al., 2013). As chloroquine accumulates, it sequesters free heme, a toxic secondary product of hemoglobin proteolysis by *Plasmodium* (Chugh et al., 2013). To detoxify the heme, *Plasmodium* crystallizes it into an insoluble complex called hemozoin through a yet unknown synthesis mechanism (Chugh et al., 2013). Hemozoin formation is essential for parasite survival and hence has been a target for antimalarial treatments (Herraiz, Guillen, Gonzalez-Pena, & Aran, 2019). Chloroquine and other quinoline variants effectively interfere with this detoxification process (Herraiz et al., 2019) and hence have been used as an antimalarial treatment for some decades. Unfortunately, chloroquine has become ineffective in treating *P. falciparum* in some areas due to the emergence of parasite resistance (Chinappi, Via, Marcatili, & Tramontano, 2010). Variants of the *P. falciparum* chloroquine resistance reporter (*Pf*CRT) have been found to bind and neutralize the protonated chloroquine enabling its release from the PV (Chinappi et al., 2010). This decreases the concentration of protonated chloroquine inside the PV, which then increases the survivability of *P. falciparum* (Chinappi et al., 2010).

Currently, there are no efficacious vaccines commercially available due to our incomplete understanding of the *Plasmodium* species' life cycles and genomics. Antimalarial tablets only reduce the risk of infection by around 90% (Lalloo & Hill, 2008), and hence vaccines are necessary for complete malaria prevention and eradication. The first report of an immune reaction against sporozoites was studied by Nussenzweig et al. in 1967 (Nussenzweig, Vanderberg, Most, & Orton, 1967). They used radiation-attenuated sporozoites (RAS) to immunize mice, which then showed resistance to challenge with wildtype sporozoites (Nussenzweig et al., 1967). Studies in humans were conducted a few years later with RAS, which also showed promising results for preventing liver-stage infection (Clyde, 1975; Spring et al., 2013). However, using RAS as a vaccine comes with several limitations regarding safety

and practicality (Doll & Harty, 2014; Hoffman et al., 2002). To have the appropriate RAS dose and form for activating a protective immune response, sporozoites must be irradiated just enough to maintain immunogenicity while remaining inactive for infection (Doll & Harty, 2014). Since the sporozoites are dissected from infected mosquitoes, it raised concerns for long-term storage and RAS upscale production for wide distribution (Doll & Harty, 2014).

In response to these problems with the RAS, scientists at GlaxoSmithKline developed a protein-based vaccine called RTS,S (Ballou, 2009). It is composed of B-cell and T-cell epitopes of *Plasmodium falciparum* CSP and hepatitis B envelope antigen (Lorenz, Karanis, & Karanis, 2014). The hepatitis B envelope antigen's role is to improve immunogenicity for the immune system to produce antibodies and elicit T cell responses against the CSP (Lorenz et al., 2014). RTS,S was initially tested with various adjuvant candidates, and AS01 was chosen as the most compatible and beneficial adjuvant for shelf life, immunogenicity and safety (Ballou, 2009).

The goal of RTS,S/AS01 is to interfere with the liver stage of the Plasmodium life cycle, which is the first of the human host cycles (Lorenz et al., 2014). The RTS,S/AS01 vaccine is currently undergoing implementation studied as the Phase III clinical trial demonstrated a declining efficacy over the years post-vaccination (Lorenz et al., 2014). Another challenge to this vaccine candidate is its inability to induce an effective immune response universally across the various *Plasmodium* species that cause human disease (Lorenz et al., 2014). Other human-infecting species such as *P. vivax* and *P. ovale* express CSP variants that differ to those expressed by *P. falciparum*, meaning immunity induced by the vaccine may be less effective (Lorenz et al., 2014). Other species also have a different life stage of dormancy within the liver that lasts for an indeterminate number of years (Lorenz et al., 2014). Another vaccine candidate specifically

targeting liver cell invasion by the *Plasmodium* species is in great need to be applicable across all *Plasmodium* species (Lorenz et al., 2014).

1.3. CRISPR-CAS9 OVERVIEW

1.3.1. ORIGIN OF CRISPR AND ITS DESIGN FOR SCREENING EXPERIMENTS

Clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 system mutates parts of a genome or a single gene by altering DNA sequences. This system can be applied from a single gene-level to a whole-genome level with single guide RNAs (sgRNA) that target genes of interest. CRISPR-Cas9 is a natural immune system originally identified in bacteria, and the Cas9 endonuclease works in conjunction with an sgRNA to cleave invading genomic materials (Koike-Yusa, Li, Tan, Velasco-Herrera Mdel, & Yusa, 2014). Upon binding to the scaffold sequence of the sgRNA, the Cas9 endonuclease is activated to bind to a specific gene sequence by recognizing the protospacer adjacent motif (PAM) sequence on the sgRNA (Koike-Yusa et al., 2014). A PAM sequence is a 3-5 nucleotide long (usually 3'-NGG-5') signal found on the target gene (Koike-Yusa et al., 2014). What makes the CRISPR-Cas9 technique revolutionary for genome editing is designing a 20 base pair (bp) spacer region for specific genes of interest (Koike-Yusa et al., 2014). Once this spacer region binds to the target gene, Cas9 generates a double-stranded break (DSB) which is then repaired by a non-homologous end joining (NHEJ) pathway, that creates either a single base substitution or deletion, resulting in a frameshift which completely stops protein expression (Koike-Yusa et al., 2014). However, this pathway is also prone to other errors, such as silent missense and nonsense mutations (Koike-Yusa et al., 2014).

Knock-out specificity of the Cas9 enzyme is directed by the PAM sequence and the 20 bp sgRNA sequence. Off-target knock-outs may occur if the sgRNA sequence has three to five mismatching nucleotides in the PAM-distal region (Zhang, Tee, Wang, Huang, & Yang, 2015).

To investigate the impact of sgRNA sequence on CRISRP-Cas9 knockouts, Wang et al. used a Cas9-expressing human cell line and *AAVS1* gene as their target (B. Wang et al., 2019). First, for a complete loss-of-function of *AAVS1*, Wang et al. observed higher genomic cleavages with continuous stable expression of sgRNAs for up to 2 weeks (B. Wang et al., 2019). Then, when they also tested off-target activities, they observed that although there was a low cleavage activity level with the sgRNAs 1-3bp different than the *AAVS1* complementarity, one of the sgRNAs with a perfect complementarity at its seed region (last ~8 bp) exhibited a high level (B. Wang et al., 2019). These results suggest that having up to 3 bp variabilities is acceptable as long as the seed region matches. Their data also found that on average, one sgRNA has on average 2.2 off-target sites in the human genome, and these sites are very likely to be located in the non-coding regions, which do not compromise overall knockout effects of genes of interest in an experiment (B. Wang et al., 2019). However, off-targets are not desired in a screening experiment as the computational programme used for this project can only read the existing sgRNA sequences in the samples without differentiating an off-target knockout effect from a specific one.

Once sgRNAs are designed, they are delivered to cells to generate a genetic knockout. For a genetic knockout to take place in a cell, both sgRNA and Cas9 enzyme must be expressed together. One way to achieve this is to generate a stable Cas9-expressing cell line by transduction with lentiviral particles carrying a Cas9 enzyme-encoding plasmid. Then, lentiviral particles carrying sgRNAs sequences can then be delivered to the Cas9-expressing cells. Another method is using lentiviral particles with plasmids encoding both Cas9 enzyme and a sgRNA of interest. At the end, these methods lead to a lentiviral plasmid integration into host cell genome for permanent knockouts and inheritability. This method can be applied to screen multiple genetic knockouts in a CRISPR-Cas9 screen. First, libraries containing a

plethora of sgRNA oligos are cloned into lentiviral plasmids to create lentivirus pools (B. Wang et al., 2019). Then, Cas9-expressing cells are transiently transduced by these lentivirus pools. Since the sgRNA-encoding lentiviral plasmids contain a drug-resistant marker, inefficiently transduced and non-transduced cells die when the selection marker is added to cell culture, leaving only successfully transduced cells that express the sgRNAs to survive and proliferate (B. Wang et al., 2019). Lastly, genomic DNA from the cells of interest is isolated and amplified to sequence the present sgRNA encoding regions by Next Generation Sequencing (NGS).

1.3.2. Application of CRISPR-Cas9 screen

Before CRISPR-Cas9 knockout screen technology was introduced, scientists used RNA interference to knock-down gene expression(s). However, as RNA interference works only on mRNA levels, its knock-down is transient and temporary. On the other hand, CRISPR-Cas9 generates knockouts at the DNA level, and the results are permanent and inheritable. Hence, CRISPR-Cas9 is useful for culturing mutated cells for a prolonged time. Combined with a drug selection system, the CRISPR-Cas9 knock-out screen technology provide an excellent tool for analyzing cells from both positive and negative selection experiments.

Positive selection is a method of enriching a cell population that can survive under selective pressure by cytotoxic drugs (Cong et al., 2013; Doench et al., 2014). If the cells of interest are appropriately transduced and express an anti-cytotoxic drug marker, they survive and proliferate (Doench et al., 2014). Meanwhile, a negative selection screening identifies essential genes that cause cells to be depleted during the selection process (Doench et al., 2014). For instance, in an experiment with pooled knockouts, if cells with gene “X” knocked-out are depleted over time, it means that this gene “X” is essential for cell survival (Doench et al.,

2014). Such methods are known as forward genetic screening approaches, where scientists can mutate multiple genes at once, select the cells by their phenotype(s) of interest and identify which mutations took place (Doench et al., 2014). As the names suggest, positive and negative selection are two independent approaches to identifying essential genes (Doench et al., 2014). However, in a CRISPR-Cas9 screening method, both selections can be analyzed simultaneously from a single experiment when a suitable computational tool is used.

CRISPR-Cas9 screen analyses come with several statistical considerations that must be applied when comparing between treatment and control samples, within one biological experiment, and between replicates. First, there are different read counts per sample. This can be due to cell culture selection with a particular group of cells, or different reagents used for preparing NGS library submissions. Second, sgRNA abundance can be different per sample and per experiment due to cloning errors or cell culture selection. Lastly, different sgRNAs targeting the same gene may have different knock-out efficiencies.

There are computational programmes capable of calculating these said considerations such as edgeR, DESeq, baySEq and NBPSeq. They are primarily used for RNA-Seq, but they can also identify hits in a CRISPR-Cas9 knock-out screen. However, they are not capable of directly identifying DNAs of interest in the cell genome. Hence, Li et al developed a programme called MAGeCK (Model-based Analysis of Genome-wide CRISPR-Cas9 Knockout), a highly sensitive computational tool that identifies positively and negatively selected hits on both DNA and sgRNA levels by analyzing sgRNA-encoding sequences in the cell genome (Li et al., 2014). From the total and individual sgRNA sequence read counts of the input samples, it can rank the genes by using robust rank aggregation (RRA; referred to as MAGeCK-RRA) or maximum likelihood estimation (MLE; referred to as MAGeCK-MLE) method (Li et al., 2014).

The MAGeCK-RRA method uses a negative binomial distribution to analyze changes in the read counts for each sgRNA between the input samples (Li et al., 2014). Hence, p-values for both negative and positive selections are calculated for the sgRNAs, ranked in order based on their calculated p-values (Li et al., 2014). Afterwards, the positively or negatively selected genes are ranked by α-robust rank aggregation (α-RRA), a method Li et al. modified from its origin, RRA (Li et al., 2014). Kolde et al. developed RRA, which assigns a significance score for each gene that ranks higher than the expected rank under the null hypothesis (Kolde, Laur, Adler, & Vilo, 2012). The RRA method provides robust data, clear from noise, thereby generating a list of statistically significant genes under either negative or positive selection (Li et al., 2014). While the standard RRA method applies to the gene level, Li et al. developed it further to rank the sgRNAs before the gene significance evaluation stage (Li et al., 2014). In α-RRA, positively and negatively selected genes are identified by the significances of their sgRNAs (Li et al., 2014). If they are continuously ranked higher than expected, their corresponding genes are deemed to have been either negatively or positively selected (Li et al., 2014).

Similar to MAGeCK-RRA, there is another, more robust, method known as MAGeCK-MLE. While MAGeCK-RRA cannot accurately compare samples between more than two conditions, MAGeCK-MLE allows the user to compare samples from multiple conditions and multiple experiments (Li et al., 2015). In addition, since the Cas9 enzyme's accessibility influences genetic knock-outs to the chromatin and the PAM sequence position on the targeting gene (Wu et al., 2014), MAGeCK-MLE also calculates knock-out efficiency scores for each gRNA, which are then incorporated in calculating β score, a significance score of each gene (Li et al., 2015). While MAGeCK-RRA assumes that all gRNAs are efficient, MAGeCK-MLE is able to label sgRNA by their knock-out efficiency scores between 0 and 1 (Li et al., 2015). In the end,

like MAGeCK-RRA, MAGeCK-MLE provides a final list of genes ranked by their β scores (Li et al., 2015).

1.3.3. CRISPR-Cas9 SCREEN REQUIREMENTS

Since the sgRNA sequence-containing lentivirus pools are delivered transiently by using toxic transduction reagents such as polybrene, excessive exposure to both toxic reagent and the lentiviral particles can compromise cell viability and growth. Hence, an optimal multiplicity of infection (MOI) must be calculated before the experiment, to minimize exposure to toxicity from multiple transductions and genomic integrations per cell. MOI is calculated by dividing the number of virions by the number of cells seeded for transduction. This way, theoretically, at the MOI of 1, one cell should be transduced by one virion. However, this does not mean that all cells will have a single viral genomic integration. A Poisson distributions model can be used to predict cell proportions transduced by a certain number of viruses (Figliozzi, Chen, Chi, & Hsia, 2016). The formula is as follows:

$$P(n) = \frac{m^n e^{-m}}{n!}$$

$P(n)$ is the proportion of cells infected by n virus particles at the MOI of m (Figliozzi et al., 2016). At the MOI of 1, according to the Poisson distribution, the fraction of uninfected cells is 37%, cells infected by one virion is 37%, and cells infected by more than one virion is 26%. A lower MOI has been observed in a pooled experiment by Shalem et al. where they used a low MOI of 0.3 for a pool of 64,751 gRNAs targeting 18,080 genes (Shalem et al., 2014). At a MOI of 0.3, 74% of the transduced cells are uninfected, 22% of them are infected by one virion, and 4% are infected by more than one virion. Overall, higher MOI provides a higher

proportion of cells transduced by one or more viruses (Figliozzi et al., 2016; Shalem et al., 2014) while lowering the proportion of uninfected cells.

The optimal number of sgRNAs per gene for high-quality genetic knockout results is an important factor for designing sgRNA libraries but it remains debatable. Hart et al. experimented with various numbers of sgRNA candidates per gene and found that 1) using more than one sgRNA showed an increase in the number of genetic screen hits, and 2) four sgRNA per gene provided adequate results and using more than four did not significantly change the results (Hart et al., 2017).

Once the number of sgRNAs per gene is decided for the sgRNA library design, a suitable computational programme must be chosen as well for high quality results of significant gene hits in a CRISPR-Cas9 screen. Li et al. explored MAGeCK's ability to detect hits using a range of gRNA candidate numbers per target gene. They first used ten sgRNA candidates to identify the top 5% of the identified genes hits and used them as a reference to see the relationship between detection of these top gene hits and the number of sgRNA candidates (Li et al., 2014). Although using more sgRNA candidates detected more top hits, they concluded that using 4-6 gRNAs was adequate to identify more than 80% of the top hits (Li et al., 2014). This result agrees with the experiment by Hart et al where they proposed that using four sgRNAs per gene was enough for producing high quality screens (Hart et al., 2017; Li et al., 2014).

Illumina Sequencing is an NGS technology used to sequence DNA regions of interest (in this case, sgRNA) encoded in a cell population expressing phenotype(s) of interest. The results from Illumina Sequencing are analyzed by a computational tool which generates sgRNA sequence read counts. Ideally, these counts exhibit a normal distribution (B. Wang et al., 2019);

however, this can be affected by several factors including sgRNA sequence to cell number ratio in a sample and polymerase chain reaction cycles used to prepare the NGS library. For an adequate sgRNA representation in a sample, Hart et al. recommended sequencing 200 times more cells than the sgRNAs (i.e., sequencing 1×10^5 cells expressing 500 sgRNAs calculates 200:1 cell to sgRNA ratio) (Hart et al., 2017). As for the PCR cycles, Koike-Yusa et al. suggested seven cycles for ligating index sequences per sample to minimize amplification errors from increased cycles and to reflect the accurate sgRNA representation in the sample (Koike-Yusa et al., 2014; Kolde et al., 2012).

1.4. OBJECTIVES

This thesis project's main objective is to identify and study surface protein-protein interactions in liver invasion using *P. berghei* rodent sporozoites and HC-04 *in vitro* grown human hepatoma cell line.

Various studies have shown that hepatocyte surface proteins such as SR-B1 (Rodrigues et al., 2008), CD81 (Silvie et al., 2003), and EphA2 (Kaushansky et al., 2015) play a role in hepatocyte invasion; however, no sporozoite proteins that bind to these host surface proteins have been identified. Likewise, CSP and 6-cysteine s48/45 domain-containing proteins such as P36 and P52 (Arredondo et al., 2018) have been identified to play a role in invasion without any counterpart hepatocyte proteins being identified. A novel binding partner of TRAP, integrin αVβ3, has been discovered by using an assay called avidity-based extracellular interaction screen (AVEXIS); however, both *in vitro* and *in vivo* knockout studies suggested that the TRAP-αVβ3 interaction is active during traversal in dermal layer, rather than hepatocyte invasion. Despite new technologies being developed to identify new interactions, the molecular mechanism of hepatocyte invasion remains an enigma.

This thesis's hypothesis of protein-protein interactions playing essential roles in a hepatocyte invasion stems from the discovery of RH5, a merozoite surface protein essential for erythrocyte invasion. Since its discovery, RH5 has become an excellent vaccine candidate for preventing blood-stage malaria. However, the liver stage is the better target for therapeutic intervention as

1. It is a bottleneck where only a few sporozoites develop into thousands of merozoites,
2. It has more prolonged exposure to the immune system than the erythrocyte stage, and

3. It is an asymptomatic stage; hence a sterilising vaccine could completely prevent clinical onset (Dundas, Shears, Sinnis, & Wright, 2019).

Studying the liver stage of malaria comes with many challenges. Unlike merozoites, sporozoites cannot multiply and be cultured easily *in vitro*. They must be dissected directly from the salivary glands of the infected and alive female mosquitoes. Moreover, studying the protein-protein interactions in their wildtype conformations *in vitro* is difficult as the interactions tend to be transient due to low binding avidity (Dundas et al., 2018). Hence, this project hypothesizes that using a CRISPR-Cas9 knockout screen method in hepatocytes provides the ability to identify essential hepatocyte proteins in both cell traversal and cell invasion. The screen results can be investigated further on the protein level to identify their sporozoite protein binding partners.

Although this study's subject is human malaria, there are safety concerns when studying its liver stage with the human-specific strain, *P. falciparum*. Since the sporozoites have to be extracted by hand from live mosquitoes, there is a high risk of exposure, and the extraction would need to be done at a CL3 laboratory. Therefore, in consideration of safe laboratory practices and practicality, a rodent specific strain, *P. berghei* was used. Although we do not fully understand *P. falciparum*'s pathological mechanisms, *P. berghei* infections in rodents exhibit similar pathology, and thus it is our best non-human model for providing insights into studying critical factors such as protein-protein interactions (Franke-Fayard et al., 2004).

HC-04, a hepatoma cell line that mimics *in vivo* human hepatocytes, was obtained through BEI Resources, NIAID, NIH: HC-04, Hepatocyte (human), MRA-975, contributed by Jetsumon Sattabongkot Prachumsri (Sattabongkot et al., 2006). These cells support development of

human malaria strains such as *P. falciparum* and *P. yoelii* and rodent strains such as *P. berghei* (Sattabongkot et al., 2006; Sinnis et al., 2013). Compared to primary human hepatocytes, which are challenging to obtain and exhibit low infection rates (0.13%), HC-04 is easy to maintain and can exhibit a 3 to 5-fold higher rate of infectivity with both humans and rodent specific strains (Tweedell et al., 2019). HC-04 is a great tool for this project in identifying and understanding essential genes for invasion by sporozoites.

Human hepatocyte receptors were ranked based on their abundance on the cell surface using proteomic and transcriptomic data. Top 470 genes were grouped into five pools (94 genes per pool) and each pool had *SCARB1* positive control gene and *CD81* negative control gene, making it a total of 480 genes. sgRNA sequences targeting these genes were cloned into a lentiviral vector that was then packaged into five lentivirus pools. Then, Cas9 expressing stable HC-04 cells were transiently transduced by these lentivirus pools and then selected by puromycin. These CRISPR-Cas9 modified cells were infected by the *P. berghei* sporozoites 15 days post-transduction. To identify infected cells, transgenic sporozoites expressing mCherry under a HSP70 promoter were used. Between 22-24 hours post incubation, 3×10^5 cells exposed to sporozoites were aliquoted to serve as a control sample (denoted as 'exposed'). Then, productively invaded cells, identified as $mCherry^{hi}$ by FACS, were sorted. Both samples were then deep sequenced by Illumina Sequencing to evaluate each genetic deletion and determine its potential influence on the sporozoite invasion. As SR-B1 is an already-known essential protein for hepatocyte invasion (Rodrigues et al., 2008), its *SCARB1* gene was used as a positive control in each CRISPR-Cas9 screen experiment. This thesis presents findings from the genetic screens and their follow-up experiments.

CHAPTER 2. MATERIALS AND METHODS

2.1. SGRNA LENTIVIRUS POOL GENERATION

Five lentivirus pools, each pool containing 500 sgRNAs that target 94 receptor genes, *SCARB1* positive control gene, *CD81* negative control gene and non-targeting sgRNAs (see Appendix A: sgRNA sequences of Pool 1, 2, 3, 4 and 5 for sgRNA sequences), were generated for CRISPR-Cas9 screen experiments as described in Methods 2.2. Details of reagents, primer sequences and kits used in this protocol are described in Table **2-3** and Table **2-4**.

2.1.1. SGRNA OLIGONUCLEOTIDE CONSTRUCTION AND RECONSTITUTION

sgRNA lists were designed by Dr. Sandy Douglas and Dr. Rameswara Segireddy. 470 genes were ranked based on their proteomic abundance on human primary hepatocyte and HepG2 hepatoma surfaces. The ranked genes were then equally grouped into five pools, and sgRNA sequences that target these genes were taken from a published library (Koike-Yusa et al., 2014). ssDNA oligonucleotide pools that encode the selected and grouped sgRNA sequences were ordered from Twist Bioscience. Each ssDNA oligonucleotide pool was reconstituted in nuclease-free ddH_2O at 1 ng/µL.

2.1.2. CLONING OF SGRNA OLIGONUCLEOTIDES TO LENTIVIRAL PLASMIDS FOR LENTIVIRUS POOL GENERATION

Five lentivirus pools were produced using methods based upon those reported by Koike-Yusa et al (Koike-Yusa et al., 2014). A graphical summary of the experimental steps taken for generating lentivirus pools is shown in Figure **2-1**.

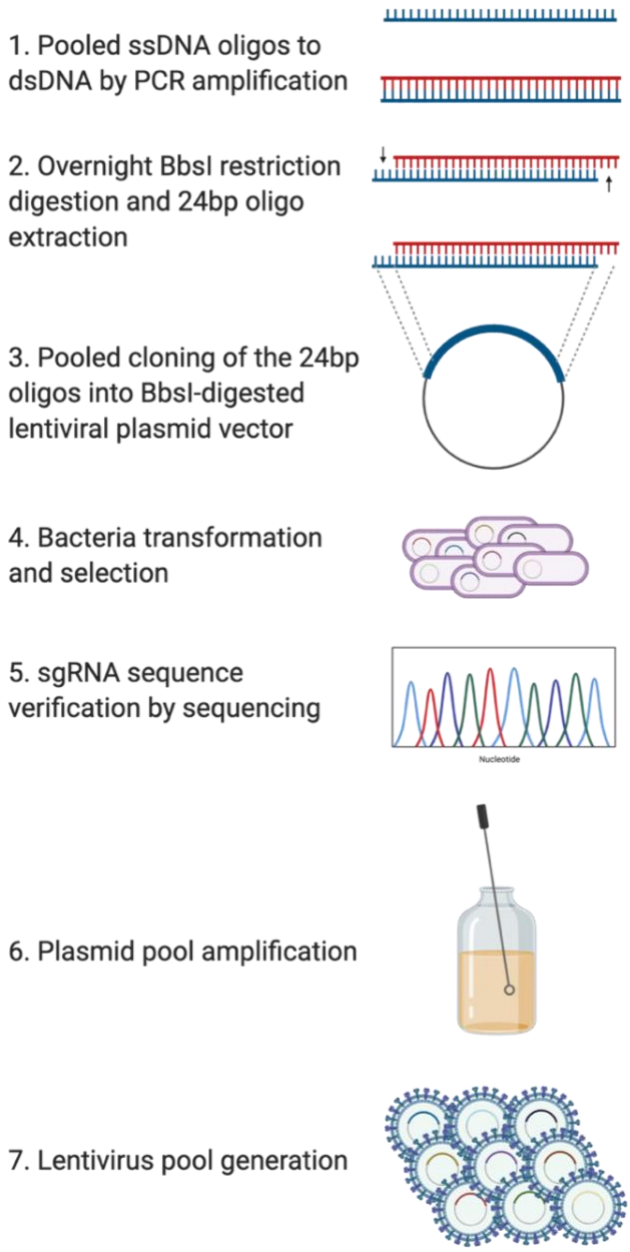


Figure 2-1 An overview of lentiviral plasmid pool cloning protocol.
Figure created with BioRender.com.

Step 1: Per lentivirus pool, 50 ng of ssDNA oligonucleotides was amplified into dsDNA by polymerase chain reaction (PCR) using a reaction mix and PCR cycling conditions as described in Table **2-1** and Table **2-2**, respectively. 5 µL of the dsDNA PCR product was mixed with 2 µL of 6X DNA Gel Loading Buffer (New England BioLabs [NEB]) and ran on a 20% Novex™ TBE gel (Life Technologies) at 150V for 1 hour. The gel was stained with Sybr™ Safe (Thermo) (4 µL of 10 mg/mL diluted in 40 mL of TAE buffer) for 30 minutes on a shaker. The stained gel was imaged by fluorescence to check a 77-bp band. Then, the remaining PCR product was column-purified using a QIAQuick Nucleotide Removal Kit (QIAGEN) as per manufacturer's instructions. The final concentration of the purified dsDNA PCR product was quantified by UV spectrophotometry.

	1 x 50 µL reaction	25 x 50 µL reactions
Q5® High-Fidelity 2X Master Mix (NEB)	25µL	625µL
100 µM 77F/77R primer mix	2.5µL	62.5µL
ddH_2O	20.5µL	512µL
ssDNA oligo pool (1ng = 1µL)	2µL (2ng)	50µL (50 ng)
Total	50µL	1250µL

Table 2-1 PCR reaction mix formula for ssDNA to dsDNA amplification in one aliquot and 25 aliquots.

77F and 77R primer sequences are in Table **2-4**.

PCR cycling step	Temperature (˚C)	Time (mm:ss)	
Initial denaturation	98	00:30	
Denaturation	98	00:10	8 cycles
Annealing	63	00:15	
Extension	72	00:15	
Final extension	72	02:00	
Incubation	4	∞	

Table 2-2 PCR cycling conditions of ssDNA to dsDNA amplification.

Step 2: The column-purified 77 bp dsDNA was digested overnight by BbsI restriction enzyme (NEB) in NEBuffer™ 2.1 (NEB) at 37˚C. The BbsI-digested dsDNA fragments were run on a

20% Novex™ TBE Gel (Life Technologies) at 150V for 1 hour. The gel was stained using the same protocol as Step 1, and 24 bp bands were cut out and crushed in TE Buffer (Invitrogen™) at 6:1 (v/v) of buffer to gel slab. To elute the 24 bp dsDNA fragment from the gel slab, first, the gel-buffer mixture was frozen for 1 hour at -80˚C, thawed in a water bath at 90˚C for 5 minutes and incubated on a rotary shaker overnight at room temperature. The following day, this mixture was centrifuged at maximum speed at 4˚C for 10 minutes, and the supernatant was transferred to an Eppendorf tube.

Then, to precipitate the dsDNA, cold 10% v/v 3M sodium acetate ($NaCH_3COO$) was added to the supernatant and vortexed for 10 seconds. To this $NaCH_3COO$/supernatant mix, cold 95% ethanol was added at 250% of the total volume. The mixture was vortexed for 10 seconds and stored at -20˚C for 2 hours. Then, 24 bp dsDNA was pelleted after centrifugation at maximum speed for one minute at 4˚C. The supernatant was discarded, and the pellet was washed twice (resuspended in 500 µL of cold 95% EtOH, centrifuged at maximum speed for one minute at 4˚C and removed supernatant) and then pelleted again by centrifugation at maximum speed. The final dsDNA pellet was resuspended in 20µL TE Buffer (Invitrogen™). The concentration of the product was quantified by fluorescence spectroscopy using QuantiFluor® dsDNA System kit (Promega).

Step 3: 3 µg of a lentiviral expression vector, pKLV2-U6gRNA5(BbsI)-PGKpuro2ABFP-W (Addgene plasmid # 67974 ; http://n2t.net/addgene:67974 ; RRID:Addgene_67974) was digested overnight with BbsI (NEB), then dephosphorylated with Antarctic phosphatase (NEB) at 37˚C for 1 hour. The linearized vector was run on a 1% agarose gel, extracted using a QIAQuick Gel Extraction Kit (QIAGEN) as per manufacturer's protocol and its concentration was quantified by UV spectrophotometry. 12 ng of the purified 24-bp dsDNA from Step 2 was

cloned into 800 ng of the linearized lentiviral vector with T4 Ligase (NEB) in a total volume of 80 µL.

Step 4: A heat-shock transformation was performed with 15 aliquots of 5 µL of the ligation mix from Step 3 per 50 µL of subcloning-efficiency competent DH5α E. coli (NEB) in 880 µL of SOC Outgrowth Medium (NEB). To quantify transformation efficiency, one transformation aliquot was serially diluted to 1:10 and 1:100 in the SOC Outgrowth Medium (NEB). This dilution step was a precaution in case the number of colonies on the undiluted agar plate was incalculable. 20 µL of the undiluted aliquot and each dilution were plated onto LB ampicillin agar plates which were then incubated overnight at 37˚C. The remaining transformation aliquots were pooled and added to 200 mL of LB ampicillin broth which was incubated overnight at 37˚C and then stored at 4˚C until Step 6.

To confirm satisfactory transformation efficiency, and hence the representation of the sgRNA oligonucleotide pools, colonies on the agar plates were counted, and the total number of colony forming units (CFU) was calculated as follows:

$$\text{Total CFU} = \text{CFU on agar plate} \times \text{dilution factor} \times \frac{1000\ \mu\text{L}}{\text{volume added to plate}} \times \text{number of aliquots}$$

Typically, the total CFU was approximately 1.5×10^5/mL (or 300 colonies per sgRNA), substantially exceeding the number of sgRNAs in a pool by 300-fold.

Step 5: Minimum of five random colonies were picked from the undiluted agar plate and cultured in LB broths overnight. Plasmids from each colony were purified by a QIAGen Plasmid Midi Kit (QIAGEN). With primer ADO34, these plasmids were sequenced by Sanger

sequencing and the sgRNA inserts encoded in the plasmids were matched to a complete list of sgRNA sequences as seen in Appendix A: sgRNA sequences of Pool 1, 2, 3, 4 and 5.

Step 6: The lentiviral plasmid pool in the 200 mL of LB broth from Step 4 was purified by a QIAGen Plasmid Plus Maxi Kit (QIAGEN) and the concentration of the purified lentiviral plasmid pool was quantified by UV spectrophotometry.

The purified lentiviral pool was used to produce a lentivirus pool. On Day 1, 5×10^6 HEK293FT cells were seeded in a T75 flask in Dulbecco's Modified Eagle Medium (DMEM)/F-12 media containing 10% heat-inactivated Fetal Bovine Serum (FBS), and incubated at 37˚C, 5% CO_2. The lentivirus production protocol was adapted from Invitrogen's Lipofectamine® LTX & PLUS™ Reagent Protocol (Invitrogen). On Day 2, 3 µg of the lentiviral plasmid pool, 90 µL of 100 ng/uL ExceLenti LTX Lentivirus Packaging Mix (OXGENE), and 12 µL of PLUS™ reagent (Invitrogen) were added to 3 mL of Opti-MEM™ I Reduced Serum Medium, no phenol red (Opti-MEM) (Life Technologies). This mixture was incubated at room temperature for five minutes. Then, 36 µL of Lipofectamine® LTX (Invitrogen) was added to this mixture and incubated at room temperature for 30 minutes. HEK293FT cells were washed once with phosphate buffered saline (PBS), and 5 mL of Opti-MEM and the DNA-lipofectamine mixture were added to the cells. Then, cells were incubated for five hours at 37˚C, 5% CO_2 to allow maximum transfection to take place. The Opti-MEM/DNA-lipofectamine mixture in the flask was replaced with 10 mL of DMEM/F-12 media containing 10% heat-inactivated FBS, 1% L-glutamine and 1% penicillin-streptomycin. On Day 3, old media was aspirated and replaced with 5 mL of DMEM/F-12 media containing 10% heat-inactivated FBS, 1% L-glutamine and 1% penicillin-streptomycin. On Day 4, the cell culture media containing the produced lentiviruses was harvested and filtered through a

0.45 µm SFCA syringe filter (Thermo). 250 µL aliquots were made for all pools and they were stored at -80˚C.

Lentiviral titres (transduction unit (TU)/mL) were calculated by flow cytometry using cells transduced with the lentivirus pool. On Day 1, 1.5×10^5 HC-04 cells were seeded per well on a 6-well plate. On Day 2, cells were transiently transduced by the lentivirus pool in a serial dilution as shown in Figure **2-2**:

A)

Dilution factor: 1X 30 µg of polybrene 150 µL of virus	Dilution factor: 0.1X 30 µg of polybrene 15 µL of virus	Dilution factor: 0.01X 30 µg of polybrene 1.5 µL of virus
Dilution factor: 0.001X 30 µg of polybrene 0.15 µL of virus	Dilution factor: 0.0001X 30 µg of polybrene 0.015 µL of virus	Dilution factor: 0.00001X 30 µg of polybrene 0.0015 µL of virus

B)

Negative control #1 30 µg of polybrene No virus	Negative control #2 No polybrene No virus	

Figure 2-2 6-well plate plan for lentivirus pool serial dilution to calculate lentiviral titres by flow cytometry.
(A) Serial dilution of lentivirus pool. (B) Negative control.

DMEM/F-12 media with 10% heat-inactivated FBS was added to each polybrene (Millipore)/lentivirus mix for a total volume of 1 mL as indicated in Figure **2-2**A. Plates were incubated at 37˚C and 5% CO_2 for 48 hours. On Day 4, cells from each well were collected and prepared following a protocol indicated in 2.6.1 for flow cytometry analysis. The transduced cells were analyzed by their intracellular BFP expression, which its marker was encoded in the lentiviral vector plasmid, pKLV2-U6gRNA5(BbsI)-PGKpuro2ABFP-W (Addgene plasmid # 67974 ; http://n2t.net/addgene:67974 ; RRID:Addgene_67974).

A dilution sample with 20-30% BFP-positive cell population was used to calculate the lentiviral titre in TU/mL as per the equation below:

$$\frac{(1.5 \times 10^5 \text{ HC} - 04 \text{ cells}) * (\text{BFP+ } \%)}{(\text{volume of virus used (mL)})} \times 1000$$

A typical lentiviral titre was between 4-7 $\times 10^6$ TU/mL.

Another lentivirus titre unit, infectious units (IFU) per mL, was calculated using Lenti-X™ GoStix™ (Takada) following the manufacturer's protocol. Unlike the flow cytometry method that detected intracellular BFP expression, Lenti-X™ GoStix™ detected concentrations of lentivirus p24 protein present in the supernatant. Both methods produced similar lentiviral titres for all pools.

2.1.3. MATERIALS

Product	Catalogue Number	Supplier
DMEM/F-12, GlutaMAX™ supplement	10565018	Life Technologies
Dulbecco's Phosphate Buffered Saline	D8537	Merck Life Science
Fetal Bovine Serum (FBS)	F9665	Sigma-Aldrich
TrypLE™ Express Enzyme (1X), no phenol red	12604013	Life Technologies
Opti-MEM™ I Reduced Serum Medium, no phenol red (Opti-MEM)	11058021	Life Technologies
Polybrene	TR-1003-G	Millipore
Q5® High-Fidelity 2X Master Mix	M0494	New England Biolabs
BbsI- HF®	R3539	New England Biolabs
NEBuffer™ 2.1	B7202S	New England Biolabs
6X DNA Gel Loading Buffer	R0611	New England Biolabs
3M CH_3COONa	N/A	N/A
95% EtOH	N/A	N/A
Antarctic phosphatase	M0289S	New England Biolabs
Antarctic phosphatase reaction buffer 10X	B0289S	New England Biolabs

T4 DNA Ligase	M0202S	New England Biolabs
T4 DNA Ligase Buffer	B0202S	New England Biolabs
Lenti-X™ GoStix™ Plus	631280	Takara Bio
NEB® 5-alpha Competent E. coli (Subcloning Efficiency)	C2988J	New England Biolabs
NEB® 5-alpha Competent E. coli (High Efficiency)	C2987I	New England Biolabs
SOC outgrowth medium	B9020S	New England Biolabs
Lipofectamine® LTX	15338-030	Invitrogen
PLUS™ Reagent	11514-015	Invitrogen
Lentivirus packaging reagent	EXL10	OXGENE
Novex™ TBE Gels, 20%, 15 well	EC63155BOX	Life Technologies
QIAquick Nucleotide Removal Kit	28304	QIAGEN
20X TE Buffer (pH 7.5)	A2651	Promega
QIAquick Gel Extraction Kit	28704	QIAGEN
0.45µm SFCA syringe filter	723-2545	Thermo Scientific™

Table 2-3 Kits and reagents used for lentivirus pool generation.

Name	Sequence (5'-3')
77F	GCAGATGGCTCTTTGTCCTA
77R	GCGACGAGAAGACTAAAAC

Table 2-4 Primer sequences used for lentivirus pool generation.

2.2. IN VITRO HC-04-CAS9 INVASION BY *P. BERGHEI* SPOROZOITES

Figure **2-3** summarizes a protocol that incorporated a sporozoite invasion experiment and a CRISPR-Cas9 screen. Reagents used in this protocol are listed in Table **2-5**.

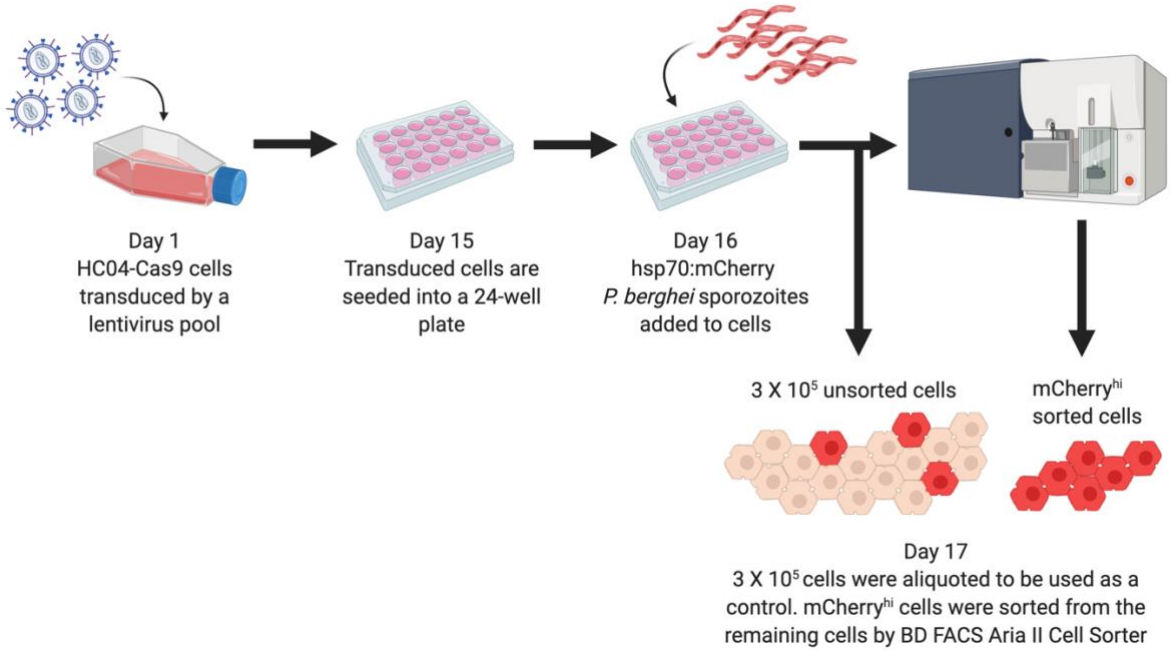


Figure 2-3 An overview of an *in vitro* invasion experiment of lentivirus pool transduced HC-04-Cas9 cells by *P. berghei* sporozoites.
Figure created with BioRender.com.

On Day 1, 5×10^5 HC-04-Cas9 cells were seeded in a T25 flask and transduced by a lentivirus pool at a multiplicity of infection (MOI) of 2 with 40 µg of polybrene (Millipore). These lentivirus pool-transduced cells were cultured in DMEM/F12 media with 10% heat-inactivated FBS, 1% penicillin-streptomycin and 0.2% of 20µg/mL blasticidin (Melford) at 37˚C and 5% CO_2.

On Day 2, the old media was aspirated and 5 mL of DMEM/F12 media with 10% heat-inactivated FBS, 1% penicillin-streptomycin, 0.2% of 20 µg/mL blasticidin (Melford) and 0.1%

of 10 mg/mL puromycin (Cambridge Bioscience) was added. Adding puromycin in the media positively selected for the successfully transduced cells as the lentiviral plasmid encoded a puromycin-resistant marker.

On Day 3, cells were transferred to a T75 flask. They were cultured in 15 mL of DMEM/F12 media with 10% heat-inactivated FBS, 1% penicillin-streptomycin, 0.2% of 20 µg/mL blasticidin (Melford) and 0.1% of 10 mg/mL puromycin (Cambridge Bioscience) and were split 1:4 once every 72 hours.

On Day 15, cells were seeded into 24-well plates at 2.5×10^5 cells per well in 1 mL of DMEM/F12 media with 10% heat-inactivated FBS and 1% penicillin-streptomycin.

On Day 16, cells were infected with 1×10^5 mCherry-fluorescent hsp70:mCherry *P. berghei* sporozoites. *Anopheles* mosquitoes were obtained from the Jenner Institute, University of Oxford at 21-28 days post-feed with blood containing the hsp70:mCherry sporozoites (strain: mCherryPbHsp70_230p 1804 cl1, a gift from Chris Janse and Shahid Khan). Mosquito pots were first incubated at 2-8 °C until the mosquitoes were motionless. Using a microscope and syringe needles, mosquito heads were removed, and the salivary glands were extracted from the body and stored in cold Roswell Park Memorial Institute (RPMI) 1640 Medium (Merck Life Science) media on ice. Then, the extracted glands were crushed using a glass Dounce homogenizer on ice, and debris was removed by centrifugation at 800 rpm for three minutes at 4˚C. Sporozoites in the supernatant were transferred to a clean Eppendorf tube. Total number of sporozoites was estimated by counting the sporozoites using a haemocytometer and a fluorescent microscope. Here, only bright sporozoites were estimated to be viable and thus dim sporozoites were excluded from the calculation. Sporozoites were diluted in DMEM/F12 media

with 10% heat-inactivated FBS at 1×10^5 sporozoites per mL (sporozoite-to-cell ratio: 0.4). Media was aspirated from the 24-well plates and replaced with the new media containing sporozoites at 1 mL per well for infection to take place. The plates exposed to the sporozoites were centrifuged at 1000 × g for one minute. Then, they were incubated for 22 hours at 37˚C and 5% CO_2.

On Day 17, at 22 hours post-infection, the cell infectivity rate was calculated using fluorescent microscopy. Under 40X magnification, bright red fluorescent dots, which were assumed to be healthy, sporozoite-invaded cells, were counted in 4 wells and averaged. Dim ones, on the other hand, were excluded from the calculation because they were thought to be either not productively invaded or in the process of cell traversal. Then, the average number of dots was divided by the total number of cells from the same 40X magnification field of view to estimate the infection rate. If the infection rate was greater than 0.5%, the experiment was considered as satisfactory.

All sporozoite-exposed cells from the wells were collected together into one Eppendorf tube. 3×10^5 of these cells were aliquoted to be used as a control sample for a Next Generation Sequencing (NGS) analysis (see Methods 0).

The remaining cells were taken to the Kennedy Institute of Rheumatology to sort a $mCherry^{hi}$ population using a BD FACS Aria II Cell Sorter. An example of the gating scheme is shown on Figure 3-11. Every invasion experiment aimed to sort fluorescently bright invaded cells ($mCherry^{hi}$) with fully grown, healthy sporozoites. The stringency of selection of the brightest parasites was balanced with the need to obtain a sufficient number of $mCherry^{hi}$ cells (≥50 per sgRNA, on average, on recommendation from Koike-Yusa, Sanger Institute) to give robust

results. In experiments with a high yield of infected cells (as determined by the visually estimated infection rate from above), more stringent gating was used, while the gate was relaxed in experiments with lower yield of infected cells. Nevertheless, all sorted cells were always clearly mCherry-positive as compared to a negative control cell sample that was not exposed to the sporozoites.

The mCherry^{hi} population gate was calculated prior to sorting. First, the total number of sporozoite-exposed mCherry-positive cells (y) was estimated by multiplying the number of cells in one well by the total number of sporozoite-exposed wells. Then, a total number of sortable sporozoite-invaded mCherry^{hi} cells was estimated by multiplying y by the infection rate calculated earlier from a visual estimation with a fluorescent microscope. If the result was higher than 3×10^4, the minimum required percentage of singlet mCherry^{hi} cells to obtain at least 2.5×10^4 (n) was calculated by using the following method:

$$n = \frac{25000 \times 100}{y \times 0.8 \times 0.7}$$

0.8 was an estimate of a cell population gated on FSC-A vs SSC-A, and 0.7 was an estimate of singlet cells gated on FSC-W vs FSC-H (see Figure 3-11). If there were less than 3×10^4 sortable sporozoite-invaded mCherry^{hi} cells, the invasion experiment was considered as inadequate for further analysis.

2.2.1. MATERIALS

Product	Catalogue Number	Supplier
DMEM/F-12, GlutaMAX™ supplement	10565018	Life Technologies
Roswell Park Memorial Institute (RPMI) 1640 Medium	R7388	Merck Life Science
Dulbecco's Phosphate Buffered Saline	D8537	Merck Life Science
Fetal Bovine Serum (FBS)	F9665	Sigma-Aldrich
TrypLE™ Express Enzyme (1X), no phenol red	12604013	Life Technologies
Blasticidin	B12150	Melford Laboratories
Puromycin	13884	Cambridge Bioscience

Table 2-5 Reagents used for in vitro HC-04 invasion experiment.

2.3. NEXT-GENERATION SEQUENCING DATA ANALYSIS

Both sorted mCherry^{hi} cells and unsorted 3×10^5 control cells underwent two rounds of PCR to add adaptor sequences around a sgRNA-encoding region (Method 2.3.1.2) and an iPCRTag (Method 2.3.1.3) that encodes a unique identifier tag (iPCRTags), as described in Figure **2-4**. Round 2 PCR products were then multiplexed for Illumina Sequencing submission. The sequencing data were analyzed by MAGeCK-VISPR and MAGeCK-Flute (Method 2.3.2). Figure **2-5** summarizes a protocol from sample preparation to data analysis. Reagents and kits used in this protocol are described in Table **2-11**.

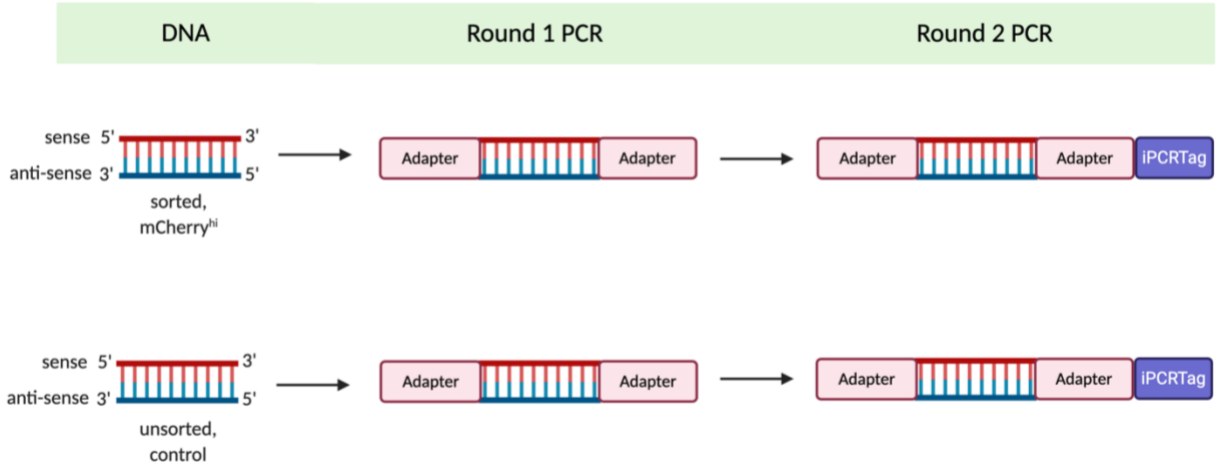


Figure 2-4 Round 1 PCR and Round 2 PCR overview for Illumina Sequencing library submission.
In Round 1 PCR, adapter sequences were added to both 3' and 5' ends around a sequence of interest (i.e., sgRNA encoding region). In Round 2 PCR, iPCRTag was added only at the 3' end of sense strand. Figure created with BioRender.com.

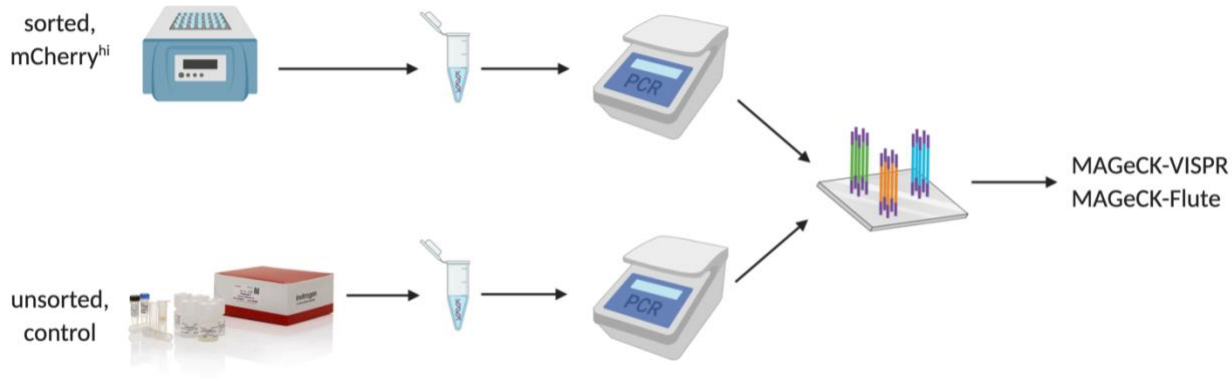


Figure 2-5 Overview of steps from gDNA isolation to Illumina Sequencing data analysis. Genomic DNA from each cell sample underwent two rounds of PCR, as illustrated in Figure **2-4**. Then, samples were multiplexed for an Illumina Sequencing submission. The sequencing results were analysed using MAGeCK-VISPR and MAGeCK-Flute. Figure created with BioRender.com.

2.3.1. SAMPLE PREPARATION

2.3.1.1. CELL LYSIS

Two protocols were used for genomic DNA (gDNA) isolation due to different cell sample sizes. Small samples were susceptible to DNA loss due to PCR-inhibitory cellular components and PCR clean-up processes. Larger samples were not at a risk of compromising the overall sample quality.

For samples with less than 1×10^5 cells, such as the sorted mCherry^{hi} cells, cells were pelleted, resuspended in 20 µL of ddH_2O and heated at 95˚C for 10 minutes. 5 µL of 2 mg/mL of Proteinase K (Sigma-Aldrich®) was added per 20 µL of the cell lysate (final concentration: 400 µg/mL) and the mixtures were incubated at 56˚C for 50 minutes for protein degradation, then again at 95˚C for 10 minutes for protease inactivation.

As for samples with >1 x 10^5 cells, such as the unsorted control sample which had 3 x 10^5 cells, their gDNA was isolated by using a PureLink Genomic DNA Mini kit (Life Technologies) and its standard manufacturer's protocol. gDNA was eluted in 30 µL Buffer EB from QIAGEN. A QIAGEN buffer product was used instead of an AE elution buffer provided with the PureLink Genomic DNA Mini kit as the AE elution buffer contained EDTA which inhibits PCR by depleting free magnesium ions in the PCR mixture.

2.3.1.2. ROUND 1 PCR

The mCherry^{hi} cell lysate and the isolated gDNA of the control cells underwent a Round 1 PCR to add adaptor sequences at both 3' and 5' ends of an sgRNA-encoding region in the gDNA (see Figure **2-4**). A Round 1 PCR Master Mix was prepared as described in Table **2-6**. 23 µL of the cell lysate or the isolated gDNA was mixed with 27 µL of the Round 1 PCR Master Mix. Then, a PCR was conducted according to the cycling conditions described in Table **2-7**. Round 1 PCR products were purified using a MinElute PCR Purification Kit (QIAGEN) following the manufacturer's protocol. The eluted DNA concentrations were quantified by fluorescence spectroscopy using a QuantiFluor® dsDNA System kit (Promega) and a Clariostar multi-mode plate reader (BMG), both according to manufacturers' instructions.

Reagent	Volume
Q5® High-Fidelity 2X Master Mix	500 µL
ADO12 (stock concentration: 20 µM)	20 µL
ADO13 (stock concentration: 20 µM)	20 µL

Table 2-6 Round 1 PCR Master Mix components and formula. See Table 2-12 for primer sequences.

PCR cycling step	Temperature (˚C)	Time (mm:ss)	
Initial denaturation	98	00:30	
Denaturation	98	00:10	25 cycles
Annealing	61	00:15	
Extension	72	00:20	
Final extension	72	02:00	
Incubation	4	∞	

Table 2-7 Round 1 PCR cycling conditions.

2.3.1.3. ROUND 2 PCR

The purified Round 1 PCR products underwent a Round 2 PCR to add iPCRTags which gave each sample its unique identity (Figure **2-4**), therefore allowing both researchers and the Illumina Sequencing machine to identify samples from another in a multiplexed library. Sequences of the primers and iPCRTags used are described in Table **2-13**. The iPCRTags were mixed with an ADO115/PE1.0 primer in ddH_2O as described in Table **2-8**. Then, a Round 2 PCR Master Mix was prepared as described in Table **2-9**. A mixture of 20 µL of the Round 2 PCR Master Mix, 25 µL of KAPA HiFi HotStart ReadyMix (KAPA Biosystems), and 1 ng of the purified Round 1 PCR products underwent a PCR according to the cycling conditions shown in Table **2-10**. Prior to Round 2 PCR product purification, 5 µL of the Round 2 PCR product was analysed on a 2% agarose gel. The expected band size was 326 bp.

Reagent	Volume
ddH_2O	360 µL
ADO115/PE1.0 (stock concentration: 20 µM)	10 µL
iPCRTag (stock concentration: 20 µM)	10 µL

Table 2-8 Index tag, ADO115/PE1.0 and ddH_2O mix formula for Round 2 PCR.

Reagent	Volume
KAPA HiFi HotStart ReadyMix	25 µL
Round 2 PCR ddH_2O-primer mix	20 µL
Cleaned up Round 1 PCR product	5 µL (1ng)

Table 2-9 Round 2 PCR reaction mix components and formula.

PCR cycling step	Temperature (˚C)	Time (mm:ss)	
Initial denaturation	98	00:30	
Denaturation	98	00:10	8-11 cycles
Annealing	66	00:15	
Extension	72	00:20	
Final extension	72	05:00	
Incubation	4	∞	

Table 2-10 Round 2 PCR cycling conditions.

After checking a 326 bp band by fluorescence, the Round 2 PCR product was purified by a solid-phase reversible immobilization (SPRI) bead technology with AMPure XP beads (Beckman Coulter) and a DynaMag-96 Side Magnet (Invitrogen™). Beads were equilibrated to room temperature and vortexed before use.

The purification process was adapted from the manufacturer's protocol. First, 31.5 µL of beads were added to 45 µL of the Round 2 PCR product and incubated at room temperature for five minutes for the DNA fragments to bind to the beads. This bead-DNA mixture was then placed on a 96-well magnetic plate and incubated for at least three minutes to capture the beads by magnet. Supernatant was removed, 150 µL of 80% ethanol was added to the bead pellet without disturbing it and incubated at room temperature for 30 seconds. 150 µL of 80% ethanol was removed and 180 µL of 80% ethanol was added immediately to the bead pellet without disturbing it. After a room temperature incubation for 30 seconds, all liquid was removed, and the bead pellet was air-dried at room temperature for five minutes. The PCR tube was removed from the magnetic plate, and the bead pellet was thoroughly resuspended in 35 µL of Buffer

EB (QIAGEN). The bead-Buffer EB mixture was incubated at room temperature for at least three minutes to release the purified DNA fragment products from the beads. The PCR tube was placed back on the magnetic plate for at least three minutes for bead capture. 30 µL of the eluate was transferred to an Eppendorf tube. The purified product was then quantified by fluorescence spectroscopy using a QuantiFluor® dsDNA System kit (Promega) and a Clariostar multi-mode plate reader (BMG).

2.3.1.4. SAMPLE MULTIPLEXING

10 ng of each purified Round 2 PCR product was pooled in an Eppendorf tube to create an Illumina Sequencing library for submission. This library's concentration was quantified using QuantiFluor® (Promega).

Illumina Sequencing was performed by Oxford Genomics Centre using a MiSeq Reagent v2 Micro Kit. Samples were loaded at 20 pM, and a 19-bp single-end sequencing was performed using a sequencing primer, ADO124, and an index primer, ADO125 (Table **2-14**).

2.3.2. SEQUENCING DATA ANALYSIS BY MAGECK PROGRAMME

Raw sequencing data were de-multiplexed by using the index tags, by Oxford Genomics Centre. Sequence data were received as bam files, which were then converted to fastq files using SAMtools (v1. 10.; Li 2009). The fastq files were analyzed by MAGeCK-VISPR, which had both RRA and MLE methods incorporated. Further details of the analysis pipeline are given in section 3.2.2.

The following files produced by MAGeCK-VISPR were used in this project for gene rank analysis as shown in Results 3.2.3: “all.countsummary.txt”, “mle.vispr.yaml”, “all.count_normalized.txt” and “mle.gene_summary.txt”.

“all.countsummary.txt”, “mle.vispr.yaml” and “all.count_normalized.txt” were used for quality control (QC) assessment of sgRNA reads and sequencing files. QC analysis criteria are listed in Results 3.2.2. The “all.count_normalized.txt” file was also used for β-score calculation, ranking negatively selected genes and MLE batch matrix by using R. 4.0.2 (R Core Team, 2020), ggpplot2 (Wickham, 2016), and MAGeCKFlute (v1.8.0.; Wang, 2020) packages. The standard deviation cut-off scale was 2 for highlighting genes as negatively selected. False discovery rate (FDR) was used as an indicator of statistical significance.

2.3.3. MATERIALS

Product	Catalogue Number	Supplier
PureLink™ Genomic DNA Mini Kit	K182002	Life Technologies Ltd
Buffer EB	19086	QIAGEN Ltd
Proteinase K	P2308	Sigma-Aldrich®
Q5® High-Fidelity 2X Master Mix	M0494	New England Biolabs
MinElute PCR Purification Kit	28004	QIAGEN
QuantiFluor® dsDNA System	E2670	Promega
KAPA HiFi HotStart ReadyMix (2X)	KK2606	KAPA Biosystems
AMPure XP	A63881	Beckman Coulter
DynaMag™-96 Side Magnet	12331D	Invitrogen™

Table 2-11 Reagents and kits used for Illumina Sequencing library submission.

Name	Sequence (5'-3')
ADO12	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTGTGGAAAGGACGAAACA
ADO13	TCGGCATTCCTGCTGAACCGCTCTTCCGATCTCTAAAGCGCATGCTCCAGAC

Table 2-12 Round 1 PCR forward and reverse primer sequences.

Name	Sequence (5'-3')	Unique 8 bp tag
ADO110/ PE1.0	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCT ACACGACGCTCTTCCGATC*T	N/A
iPCRTagT1	CAAGCAGAAGACGGCATACGAGATAACGTGATGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	ATCACGTT
iPCRTagT2	CAAGCAGAAGACGGCATACGAGATAAACATCGGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	CGATGTTT
iPCRTagT3	CAAGCAGAAGACGGCATACGAGATATGCCTAAGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	TTAGGCAT
iPCRTagT4	CAAGCAGAAGACGGCATACGAGATAGTGGTCAGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	TGACCACT
iPCRTagT5	CAAGCAGAAGACGGCATACGAGATACCACTGTGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	ACAGTGGT
iPCRTagT6	CAAGCAGAAGACGGCATACGAGATACATTGGCGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	GCCAATGT
iPCRTagT7	CAAGCAGAAGACGGCATACGAGATCAGATCTGGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	CAGATCTG

iPCRTagT8	CAAGCAGAAGACGGCATACGAGATCATCAAGTGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	ACTTGATG
iPCRTagT9	CAAGCAGAAGACGGCATACGAGATCGCTGATCGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T	GATCAGCG
iPCRTagT10	CAAGCAGAAGACGGCATACGAGATACAAGCTAGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T	TAGCTTGT
iPCRTagT11	CAAGCAGAAGACGGCATACGAGATCTGTAGCCGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	GGCTACAG
iPCRTagT12	CAAGCAGAAGACGGCATACGAGATAGTACAAGGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	CTTGTACT
iPCRTagT13	CAAGCAGAAGACGGCATACGAGATAACAACCAGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	TGGTTGTT
iPCRTagT14	CAAGCAGAAGACGGCATACGAGATAACCGAGAGAGAT CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC∗T	TCTCGGTT

Table 2-13 Sequences of iPCRTags and their unique 8-bp identifier tags.

The iPCRTag sequences were adapted from Bronner et al. (Bronner, Quail, Turner, & Swerdlow, 2014). Note that asterisks refer to phosphorothioate linkage sites. Phosphorothioate helps to protect excision of terminal thymine (T) by exonucleases. The unique 8 bp regions are used for recognizing individual samples in the multiplexed pool.

Name	Sequence (5’-3’)
ADO124	TCTTCCGATCTCTTGTGGAAAGGACGAAACACCG
ADO125	AAGAGCGGTTCAGCAGGAATGCCGAGACCGATCTC

Table 2-14 Sequences of ADO124, a sequencing primer, and ADO125, an index primer.

2.4. GENERATION OF *ITGAV*-TARGETED, *ITGB5*-TARGETED AND *SCARB1*-TARGETED HC-04 CELL LINES

ITGAV-targeted, *ITGB5*-targeted and *SCARB1*-targeted HC-04-Cas9 cell lines were generated by CRISPR-Cas9 technology with sgRNAs selected based on their Log2 fold change (LFC) and FDR results from the CRISPR-Cas9 screen experiments (Methods 2.4.1). Collagen-coated vessels (Methods 2.4.2) were used for culturing *ITGAV*-targeted and *ITGB5*-targeted cells as integrin-disruptions interfered with cell growth. *ITGAV*-targeted and *ITGB5*-targeted cells were single-cell sorted to generate monoclonal cell-lines. *SCARB1*-targeted cells, on the other hand, was produced as a polyclonal cell-line (Methods 2.4.3). Reagents and sgRNA sequences used in this protocol are described in Table **2-15** and Table **2-16**, respectively. Antibody information is listed in Table **2-19**.

2.4.1. SELECTION OF SGRNAS AND LENTIVIRUS GENERATION

To achieve the highest efficiency of genetic knock-out when creating individually gene-targeted cell lines, results of CRISPR-Cas9 screen experiments (see Results 3.2) were reviewed to identify the most effective sgRNA from a set of five sgRNAs targeting each gene. sgRNAs targeting *ITGAV*, *ITGB5* and *SCARB1* were individually assessed by FDR and LFC scores, as shown in the "rra.sgrna_summary.txt" file produced by MAGeCK-VISPR. A sgRNA sequence with the lowest LFC and FDR scores was chosen for a CRISPR-Cas9 knock out (see Table **2-16** for sgRNA details).

ssDNA oligonucleotides encoding the selected sgRNA sequences were synthesised by ThermoFisher. Then, they were cloned into lentiviral plasmids to produce lentiviral particles using described in Methods 2.1.2.

2.4.2. PRODUCTION OF COLLAGEN-COATED CELL CULTURE VESSELS

Protocols for coating cell culture vessels with collagen was adapted from Sinnis et al (Sinnis et al., 2013). 3 mg/mL of Collagen I rat tail (BD) was diluted to 0.332 mg/mL in PBS. Plates and flasks were coated at 10 µg of collagen 1cm^2. After adding the diluted collagen solution, the flasks and plates were incubated at 37°C for 3-5 hours, the collagen-PBS mixture aspirated, and then sterilized under UV light overnight. Collagen coated flasks and plates were stored at 4^{o}C for up to two weeks and washed once with PBS before each use.

2.4.3. SINGLE CELL SORTING AND CLONE GENERATION

On Day 1, 3×10^5 HC-04-Cas9 cells were seeded into 6-well plates in DMEM/F12 media with 10% heat-inactivated FBS, 1% penicillin-streptomycin and 0.2% of 20 µg/mL blasticidin (Melford Laboratories). Lentiviruses were added to the cells at MOI of 0.1, 1, and 2 with 4µL of 10 mg/mL polybrene (Millipore) in duplicate.

The transduced cells were incubated at 37 ˚C and 5% CO_2. On Day 2, media was replaced with DMEM/F12 media with 10% heat-inactivated FBS, 1% penicillin-streptomycin, 0.2% of 20 µg/mL blasticidin (Melford Laboratories) and 0.1% of 10 µg/mL puromycin (Cambridge Bioscience). On Day 4, BFP-positive cells from one well of each duplicate were sorted on a BD FACS Aria II Cell Sorter into collagen-coated 96 well plates at approximately 1 cell per

well. The second well of each duplicate was transferred to individual collagen-coated T75 flasks to create polyclonal cell-lines.

All successfully grown monoclonal colonies from the 96-well plates were picked and cultured in collagen-coated T75 flasks. To confirm loss of protein expression on the cell surface, the monoclonal cells were stained with anti-ITGAV mouse monoclonal primary antibody [Clone 272-17E6] (Sigma-Aldrich; Cat. No. MABT207) anti-ITGB5 mouse monoclonal primary antibody [Clone KN52] (eBioscience™; Cat. No. 14-0497-82), and APC-conjugated F(ab')2-Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody (Life Technologies Lt.; Cat. No. A10539) according to the protocol described in Methods 2.6.1. Loss of SCARB1 expression was not confirmed by flow cytometry but marked reduction in sporozoite invasion efficiency into the *SCARB1*-targeted HC-04 cell line (Results 4.4) provided phenotypic confirmation of likely effective gene disruption.

2.4.4. MATERIALS

Product	Catalogue Number	Supplier
Corning® Collagen I, Rat Tail	354236	BD Biosciences Discovery Labware

Table 2-15 Reagents used for culturing ITGAV-targeted and ITGB5-targeted HC-04 cells.

Target gene	sgRNA sequence	LFC	FDR
ITGAV	GCACCTCTCTTCATGGATCG	-0.98	0.038
ITGB5	GCCGAGAGGTGATGGACCGT	-1.72	0.00024
SCARB1	GGCTCTTCACGGTGTTCACG	-1.47	$< 5 \times 10^{-5}$

Table 2-16 *ITGAV*-targeting, *ITGB5*-targeting and *SCARB1*-targeting sgRNAs' sequences and their LFC and FDR scores.

2.5. ANALYSIS OF PROTEIN-PROTEIN INTERACTIONS BY AVEXIS

Human codon-optimized ectodomains of *Pf*TRAP, *Pb*TRAP, *Pf*TREP and CD200r were expressed as Bait proteins, and human integrin αV, β1, β3, and β5 as Prey proteins (Methods 2.5.1). Protein-protein interactions between Baits and Preys were quantified by AVEXIS assay (Methods 2.5.2). Reagents and antibodies used for this protocol are described in Table **2-17** and **Table 2-18**, respectively.

2.5.1. RECOMBINANT PROTEIN PRODUCTION

Plasmids including human codon-optimized ectodomain coding sequence of *Pf*TRAP, *Pb*TRAP, *Pf*TREP, CD200r and human integrin αV, β1, β3, and β5 were kind gifts from Gavin Wright (Wellcome Sanger Institute).

The ectodomain sequences were cloned into a transient mammalian expression vector based upon the pTT5 backbone (Durocher, Perret, & Kamen, 2002), as follows. 'Bait' format mammalian expression constructs (with rat CD4d3+4, human Fc and C-tags) were produced for *Pf*TRAP, *Pb*TRAP, *Pf*TREP and CD200R. *Pb*TRAP ectodomain inserts were excised from donor plasmids by overnight digestion with SacI (NEB) and AscI (NEB) restriction enzymes. Then, the fragment was run on agarose gel and extracted by using a QIAquick Gel Extraction Kit (QIAGEN) according to the manufacturer's protocol. Then, it was cloned into a SacI-AscI digested pTT5-based vector containing an CD4-human Fc-C tag. Cloning of *Pf*TRAP, *Pf*TREP and CD200R were done by Dr. Rameswara Segi Segireddy.

Human integrin αV sequence was excised from pTT-ITGAV by an overnight digestion with enzymes NotI (NEB) and AscI (NEB). Then, the fragment was cloned into a prey construct that contained 5x NanoLuc and a COMP peptide. Integrin β1, β3 and β5 sequences were not cloned into any expression vectors and remained in their pTT vector. Since their binding partner, αV, expresses a Luciferase reporter, it was redundant for the β components to express it as well. As the αV component is not expressed alone without its β counterpart, Luciferase activities observed from αVβ1, αVβ3 and αVβ5 were assumed that the corresponding β components were successfully expressed.

These recombinant proteins were produced by transient transfection of Expi293F suspension cells using an Expifectamine transfection kit (Thermo). 50 mL of Expi293F suspension cells (2×10^6/mL) were cultured on Day 1 in Expi293 Expression Medium in a shaking incubator at 37˚C and 5% CO_2. On Day 2, the Expi293F cells were transiently transfected as per the manufacturer's protocol. On Day 3, Enhancer 1 and Enhancer 2 were added to the cells. On Day 5, cells were pelleted at 1000 x g for 10 minutes, and the protein supernatant was harvested and filtered through a 0.2 µm syringe filter (Thermo). For short-term storage, it was stored at 4˚C with 0.002% sodium azide (NaN_3). For long-term storage, it was stored at -20˚C. Details on the reagent components of the ExpiFectamine™ 293 Transfection Kit (Thermo) are in Table **2-17**.

Bait protein and prey protein concentrations were quantified by ELISA and Nano-Glo® Luciferase Assay System (Promega), respectively. CD200 (plasmid provided by Dr. Rameswara Segi Segireddy, protein produced by the author) was used as a positive control for the ELISA assay and CD200r for the luciferase assay. Bait and prey proteins were normalized to 7 nM and 4×10^8 LU/mL, respectively, in Blocker™ Casein in PBS (Thermo).

2.5.2. AVEXIS

AVEXIS assays were performed according to a protocol published by Segireddy et al (Segireddy et al., 2019 PREPRINT). Figure **2-6** illustrates how the AVEXIS assay detected prey-bait protein interactions with luciferase signals.

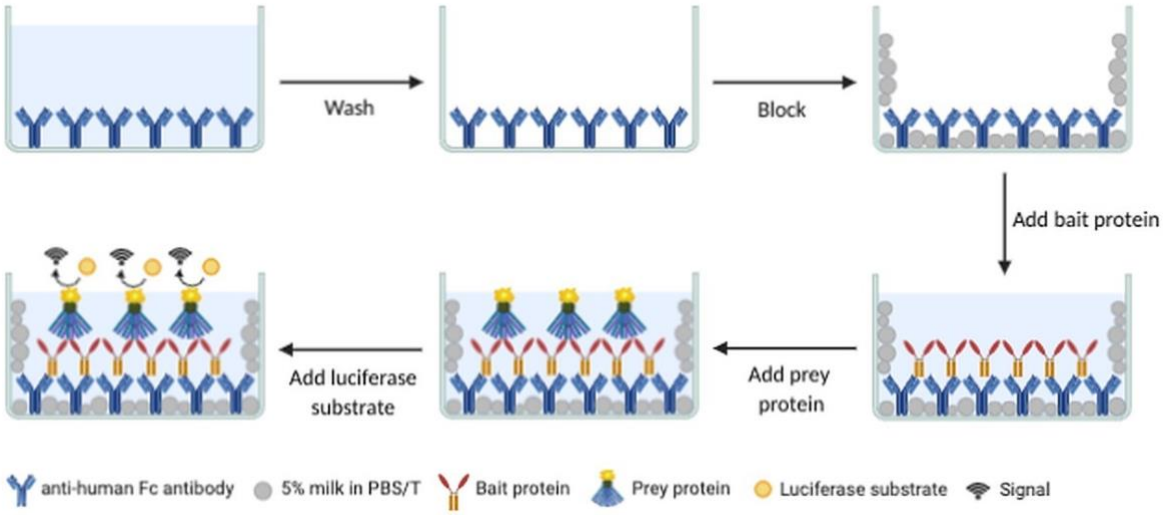


Figure 2-6 An illustrative depiction of an AVEXIS assay to detect Bait-Prey protein interactions.
ELISA wells were coated with anti-human IgG monoclonal antibody [Clone R10Z8E9] (BioRad; Cat. No. R10Z8E9) and incubated overnight. The wells were washed with PBS and 5% milk was added as a blocking agent to prevent non-specific protein binding on the surface. Then, Fc-tagged bait proteins were added to the wells which bound to the anti-human IgG antibodies. Prey proteins were added next and Luciferase substrate (Promega) was added to detect and quantify prey-bait protein interactions. Figure created with BioRender.com.

2.5.3. MATERIALS

Product	Catalogue Number	Supplier
Expi293F™ Cells	A39241	Thermo Fisher Scientific
ExpiFectamine™ 293 Transfection Kit: • ExpiFectamine™ 293 Reagent • ExpiFectamine™ 293 Transfection Enhancer 1 • ExpiFectamine™ 293 Transfection Enhancer 2	A14524	Thermo Fisher Scientific
Expi293™ Expression Medium	A1435101	Thermo Fisher Scientific
SacI	R0156S	New England Biolabs
AscI	R0558S	New England BioLabs
NotI	R0189S	New England BioLabs
NaN_3	N/A	N/A
Nano-Glo® Luciferase Assay System	N1110	Promega
Blocker™ Casein in PBS	37528	Thermo Fisher Scientific
Nalgene™ Syringe Filters	7232520	Thermo Fisher Scientific

Table 2-17 Reagents used for protein production and AVEXIS.

Product	Clone	Catalogue Number	Supplier
Anti-human IgG antibody	R10Z8E9	MCA5748G	BioRad

Table 2-18 Antibodies used for AVEXIS.

2.6. ANALYTICAL FLOW CYTOMETRY

Flow cytometry was used to analyze intracellular fluorescence (i.e., lentivirus transduction efficiency (Results 3.1.2), a mock CRISPR-Cas9 knock-out screen experiment (Results 3.1.4), sporozoite invasion experiment (Results 4.4), cell surface protein detection (i.e., single gene knock-out verification (Results 4.3.2), and cell surface-protein interaction detection (i.e. cell surface staining experiment (Results 4.3.2).

Intracellular fluorescence was detected without any staining of antibodies as BFP was encoded in the lentiviral plasmid (Methods 2.6.1). For a single gene-targeted cell line generation, protein-targeting monoclonal antibodies were used as a primary antibody (see Table **2-19** for specifications) and APC-conjugated F(ab')2-Goat anti-Mouse IgG (H+L) Cross-Adsorbed polyclonal antibody (Life Technologies Ltd.; Cat. No. A10539) as a secondary antibody (Methods 2.6.2). For analyzing cell surface-protein interactions, Fc-tagged proteins were used in lieu of a primary antibody to stain cell surfaces (Methods 2.6.3). Then, an anti-human IgG Fcγ fragment specific Alexa Fluor® 647 AffiniPure Donkey polyclonal antibody (Jackson ImmunoResearch; Cat. No. 09-005-098) was used as a secondary antibody.

Antibodies and other reagents used for flow cytometry in this project are listed in Table **2-19** and Table **2-20**, respectively. All samples were analyzed with BD™ LSR II Flow Cytometer and analysed using FlowJo™ Software (FlowJo™ Software [for Mac] Version 10. Ashland).

2.6.1. Intracellularly fluorescent cells

1×10^5 cells were pelleted at 1500 rpm for 3 minutes and supernatant was removed. The cell pellet was washed in PBS twice by repeating resuspension in 500 µL of PBS, centrifugation at 1500 rpm for 3 minutes and supernatant aspiration. This cell pellet wash method was used for all flow cytometry preparations from Methods 2.6.2 and 2.6.3. Lastly, the cell pellet was resuspended in 200 µL of PBS and filtered through 30 µm cell strainers (CellTrics®) into polystyrene FACS tubes.

2.6.2. Use of protein-targeting antibody as primary antibody

1×10^5 cells were stained with 200 µL of 1:300 diluted primary antibody in PBS and incubated at room temperature for 15 minutes. Cells were centrifuged at 1500 rpm for 3 minutes and washed using the method indicated in Methods 2.6.1. Next, cells were resuspended in 200 µL of 1:300 diluted APC-conjugated F(ab')2-Goat anti-Mouse IgG (H+L) Cross-Adsorbed polyclonal antibody (Life Technologies Ltd.; Cat. No. A10539) in PBS and incubated at room temperature for 15 minutes. Then, they were centrifuged and washed following the same method indicated in Methods 2.6.1. The final cell pellet was resuspended in 200-500 µL of PBS and filtered through 30 µm cell strainers (CellTrics®) into polystyrene FACS tubes.

2.6.3. Use of Fc-tagged proteins as primary antibody

1×10^5 cells were stained with 200 µL of normalized protein supernatant (Methods 2.5.1) and incubated at room temperature for 15 minutes. Then, cells were washed using the method indicated in Methods 2.6.1. The cell pellet was resuspended in 200 µL of PBS containing 1:300

diluted polyclonal Fc-γ fragment specific Alexa Fluor® 647 AffiniPure Donkey Anti-Human IgG antibody (Jackson ImmunoResearch; Cat. No. 09-005-098) in PBS, incubated at room temperature for 15 minutes, and washed as indicated in Methods 2.6.1. The final cell pellet was resuspended in 200-500 µL of PBS and filtered through 30 µm cell strainers (CellTrics®) into polystyrene FACS tubes.

As described in Results 4.3.2, cells were incubated with integrin-binding modulators such as cycloRGD (Santa Cruz Biotechnology) and manganese ions (Mn^{2+}) (Fluorochem). In such cases, 1×10^5 cells were incubated in either 200 µL of 10 µg/mL cycloRGD diluted in PBS or 200 µL of 2 mM $MnCl_2$ diluted in ddH_2O along with 200 µL of Fc-tagged protein supernatant at room temperature for 15 minutes. Subsequent steps were the same as described above.

2.6.4. MATERIALS

Product	Clone	Catalogue Number	Supplier
BSG	MEM-M6/6	N/A	Kind gift from Gavin Wright, Sanger Institute
F(ab')2-Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, APC	Polyclonal	A10539	Life Technologies Ltd
Alexa Fluor® 647 AffiniPure Donkey Anti-Human IgG, Fcγ fragment specific	Polyclonal	09-005-098	Jackson ImmunoResearch
Anti-Integrin alpha-V Antibody	272-17E6	MABT207	Sigma-Aldrich
Integrin beta 5 Monoclonal Antibody	KN52	14-0497-82	eBioscience™

Table 2-19 Primary and secondary antibodies used for flow cytometry analysis.

Product	Catalogue Number	Supplier
RGD peptide (GRGDNP)	sc-201176	Santa Cruz Biotechnology
Manganese(II) chloride anhydrous crystalline	494385	Fluorochem
Non-sterile CellTrics® filters (30 µm)	04-0042-2316	Sysmex

Table 2-20 Reagents and kits used for flow cytometry analysis.

CHAPTER 3. A CRISPR-CAS9 SCREEN FOR HOST DETERMINANTS OF *P. BERGHEI* SPOROZOITE INVASION AND DEVELOPMENT

470 genes were selected and ranked based on their proteomic abundance on hepatocyte (HC-04) surface as described in Methods 2.1. Then, sgRNAs targeting such genes were picked from a published library by Koike-Yusa et al. (Koike-Yusa et al., 2014) and grouped into five pools with Pool 1 containing sgRNAs targeting the top abundant genes and Pool 5 containing those targeting the least abundant. These genes were knocked out by using CRISPR-Cas9 screen technology in HC-04-Cas9 cells to identify essential genes for sporozoite invasion of HC-04-Cas9 cells (Methods 2.2).

In this chapter, I will outline 1) HC-04-Cas9 cell line generation (Results 3.1.1), 2) cloning and quality control of lentiviral plasmid pools encoding sgRNA sequences (Results 3.1.2), 3) CRISPR-Cas9 screen and Illumina Sequencing library submission optimization (Results 3.1.3 and 3.1.4), and 4) CRISPR-Cas9 screen results from sporozoite invasion experiments (Results 3.2). Overall, the following genes were identified as essential: *ITGAV, RPN1, TMEM30A, ATP2B1, ITGB5, SLC35A2, MGAT1, FCGR2B, EMC1* and *APOH*.

3.1. METHOD AND REAGENT DEVELOPMENT AND VALIDATION

3.1.1. HC-04-CAS9 CELL LINE

A Cas9-expressing monoclonal HC-04 cell-line was generated by Dr. Sandy Douglas using methods as described by Koike-Yusa et al (Koike-Yusa et al., 2014). The Cas9 activity of this cell-line was validated by Dr. Rameswara Segireddy by using a Cas9 activity reporter virus, BFP-2A-GFP-sgGFP, which was created by Koike-Yusa et al. This is a self-targeting GFP reporter virus with both GFP coding sequence and GFP-targeting sgRNA sequences encoded in the same reporter plasmid. As a result, cells transduced with this reporter virus typically lose GFP expression.

In this experiment, in parallel to the cells transduced with the BFP-2A-GFP-sgGFP reporter virus, control cells were transduced with a control virus, BFP-2A-GFP, which did not encode any sgRNA sequences targeting GFP. On day 3 post transduction, Cas9 activity was identified in the BFP-2A-GFP-sgGFP expressing cells with negative GFP expression by flow cytometry. The HC-04-Cas9 monoclonal cell-line used in this project demonstrated high Cas9 activity, as measured by the GFP expression as shown in Figure 3-1.

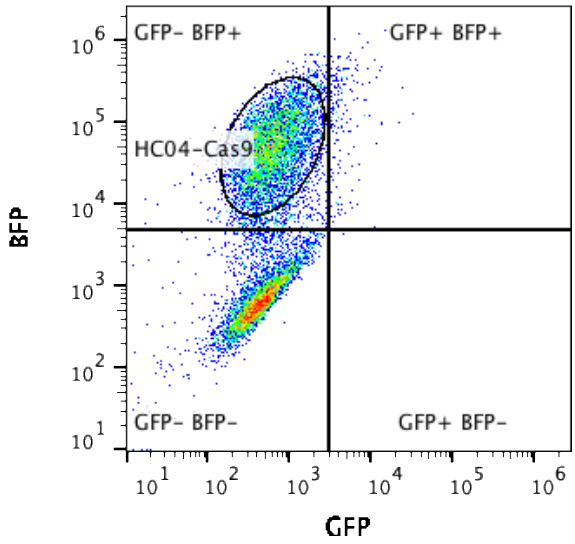


Figure 3-1 HC-04-Cas9 cell-line Cas9 activity evaluation.
More than 80 per cent of the BFP-2A-GFP-sgGFP reporter virus transduced monoclonal HC-04 cell line expressed Cas9 activity, as represented by negative GFP expression and positive BFP expression (GFP-BFP+).

3.1.2. Creation and quality control of CRISPR sgRNA-targeted cell pools

Processes of sgRNA pool design, production of pooled plasmids and lentiviral vectors, and transduction of HC-04-Cas9 cells are described in Methods 2.1.1 and 2.1.2. Illumina Sequencing sample preparation of sgRNAs present in a plasmid or cell sample, and the analysis of results, are described in Methods 2.3.1 and 2.3.2, respectively. To validate the effectiveness of these methods, samples of each plasmid pool, and of cells 17 days after transduction with the lentiviruses, were subjected to Illumina sequencing and its analysis.

Metrics provided by MAGeCK-VISPR used for quality control (QC) of the samples were as follows:

1) Gini Index:

This is a measure of income inequality in economics; however, it can be applied here to analyze the inequality of the sgRNA read counts between multiple samples per pool. An index score of 1.0, or 100%, means maximum inequality while 0.0 means perfect equal read counts across all sgRNAs (Figure 3-2).

2) Missed sgRNAs:

The number of missing sgRNAs per sample may vary due to the loss of the sgRNA-containing lentiviral plasmids during either cloning or lentivirus library production steps. The values were calculated by log of 10 (Figure 3-3).

3) Mapped reads ratio:

This data represents the proportion of the sgRNA sequences from an NGS sample that was mapped to a complete list of all sgRNA sequences for each pool. Although zero per cent unmapped reads would be ideal, zero per cent unmapped reads, this would be impossible to achieve due to mechanical or human errors. The maximum acceptable unmapped read ratio was set at 30%, following the work of Wang et al. who demonstrated high quality data with the NGS files containing up to 36.3% unmapped reads (B. Wang et al., 2019). Samples with an unmapped ratio higher than 30% were assessed for sample contamination (Figure 3-4).

4) Median quality score of each base:

This score shows whether there was any signal decay or phasing during an Illumina sequence run. The quality scores of all samples are expected to be consistently higher than 30, which represents the probability of incorrect base call at 1 in 1000 and 99.9% base call accuracy (illumina, 2011) (Figure 3-5).

5) Per sequence quality score:

High quality samples are expected to have a high peak at a value higher than 30 for the same reason as explained above (illumina, 2011). A sample with high quality reads should not have any peaks in the lower quality score region (Figure 3-6).

6) sgRNA read count frequency plot.

Each plot shows the sgRNA read count variations between two samples and a Spearman's rank correlation coefficient, r_s. A r_s value close to 1 indicates a positive, linear correlation of the sgRNA read counts between two samples. This is a useful method for determining whether the drop-out effects examined from mCherry^{hi} samples are significant or not (Figure 3-7).

sgRNA sequences in lentiviral plasmid pools (referred to as 'Plasmid') were compared against lentivirus pool transduced cells (referred to as 'Cells') to investigate if any sgRNAs were lost during transient lentiviral transduction or if there were any significant sgRNA expression biases in the cell culture at day 17 post transduction. All samples exhibited low Gini index scores (Figure 3-2), indicating high viral transfection efficiency and evenness in cell selection in culture. Pool 1, 2, 4 and 5 lost a very low number of sgRNAs (Figure 3-3); however, since they are the same sgRNA sequences in both 'Cells' and 'Plasmid' samples of their

corresponding pools, it is likely that they have been lost due to faulty cloning. Meanwhile, Pool 3 did not lose any sgRNA sequences. Nonetheless, these low missed sgRNA scores indicated that there were no lethal sgRNAs that led to them being lost in the cell culture, and hence the results of the invasion experiments were not affected as shown in Results 3.2. There were no indications of sample contamination (explained further in Figure 3-9) as the mapped reads ratios were less than 30%, a maximum accepted unmapped reads ratio (Figure 3-4). The mechanical Illumina Sequencing ran well without signs of signal decay as both median quality score (Figure 3-5) and per sequence quality scores (Figure 3-6) were high.

Next, the sgRNA read counts of the 'Plasmid' samples and the 'Cells' samples were plotted alongside each other on violin plots to visualize even representation of the sgRNA sequences in the samples (Figure 3-7). XY-scatterplots of sgRNA read counts from 'Plasmid' and 'Cells' samples were plotted to visualize background noise as represented by r_s (Figure 3-8). Similar r_s values between negative control sgRNA read counts and complete sgRNA read counts, as shown on Figure 3-8A and Figure 3-8B, indicated that the background noise observed from each pool was due to the treatment of the samples, rather than a biological effect.

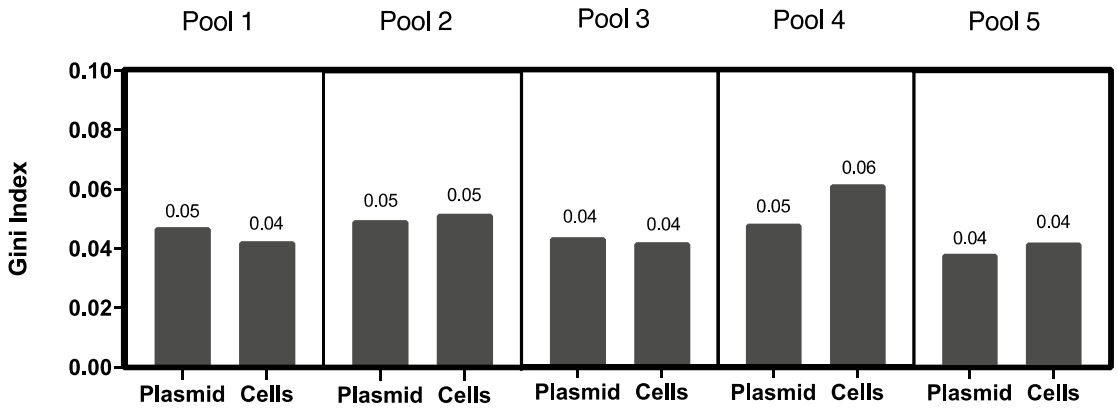


Figure 3-2 Gini index analyses of lentiviral plasmid pool ('Plasmid') and day 17 post-transduction cells ('Cells') samples show even sgRNA read counts between two samples for each pool.

Gini index scores for the plasmid and cell samples were as follows respectively: 0.05 and 0.04 for Pool 1, 0.05 and 0.05 for Pool 2, 0.04 and 0.04 for Pool 3, 0.05 and 0.06 for Pool 4, and 0.04 and 0.04 for Pool 5.

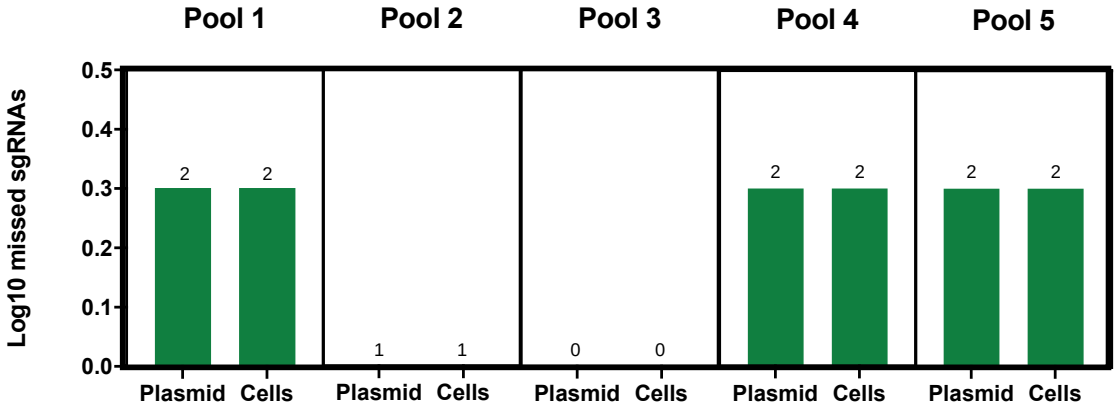


Figure 3-3 Missed sgRNA counts of lentiviral plasmid pool ('Plasmid') and day 17 post-transduction cells ('Cells') samples for all pools are low.

Numbers presented above the bars represent the missed sgRNA counts per sample as a log of 10.

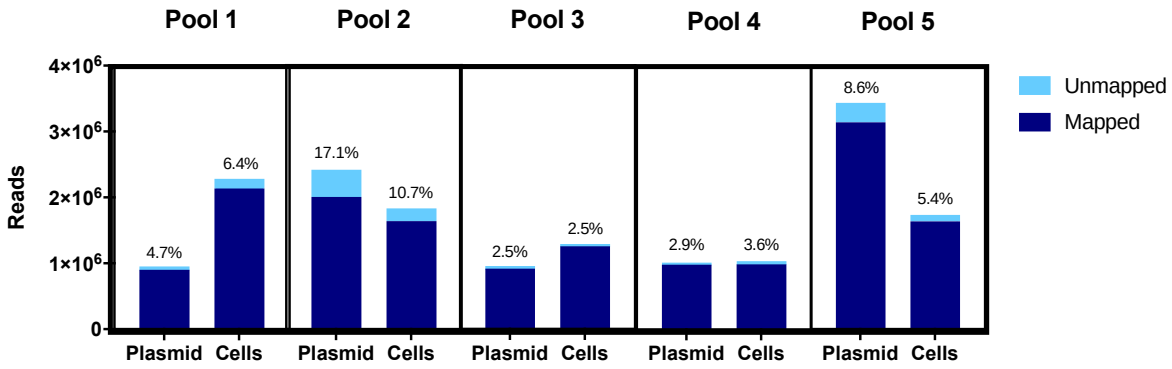


Figure 3-4 Unmapped sgRNA read count ratios of lentiviral plasmid pool ('Plasmid') and day 17 post-transduction cells ('Cells') samples for all pools.
Percentages of the unmapped reads were calculated out of the total reads.

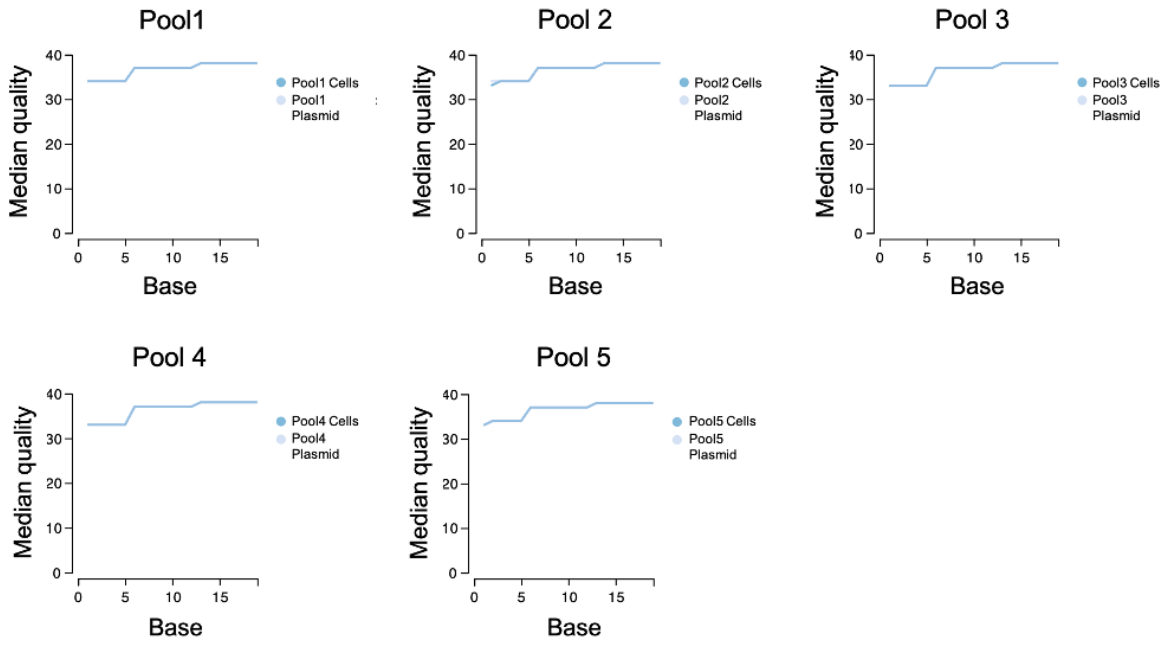


Figure 3-5 sgRNA read base median quality score of lentiviral plasmid pool ('Plasmid') and day 17 post-transduction cells ('Cells') samples for all pools.
Median quality scores were around 35 from base 0 to 19 for all samples.

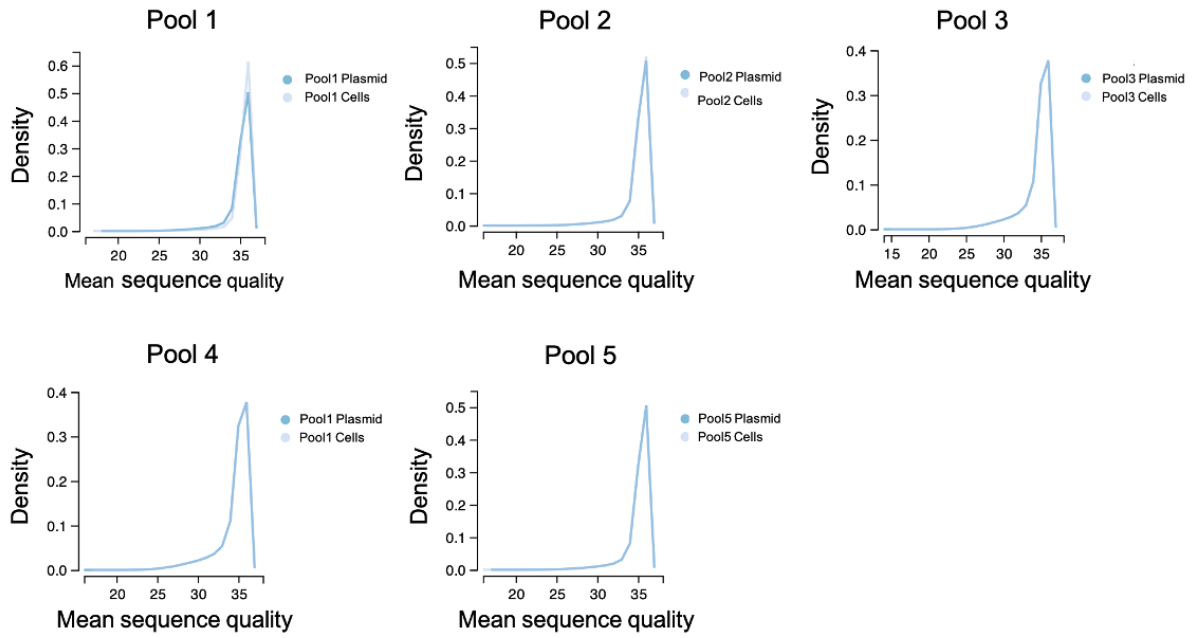


Figure 3-6 Mean sequence quality score distribution shows all pools scored a high peak at approximately 35.

Majority of reads (y-axis) scored higher than 30 for their sequence quality scores with a peak at around 35 for all samples. There were no other peaks in the lower score region.

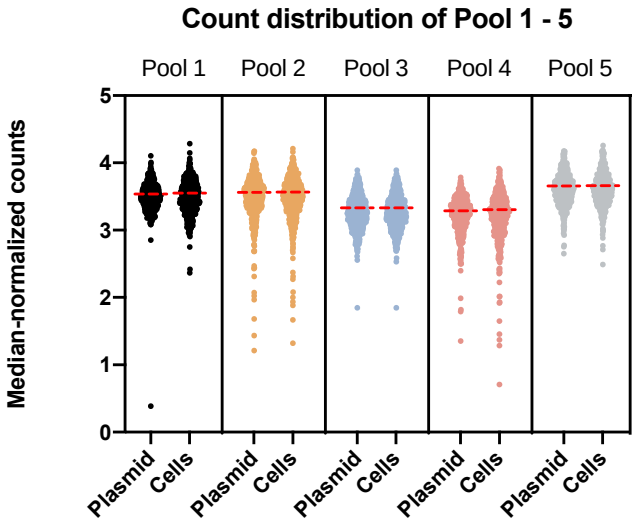


Figure 3-7 sgRNA read count distributions of Pool 1, 2, 3, 4 and 5 lentiviral plasmid pool and transduced cells.

The median (red dotted line) values of 'Plasmid' and 'Cells' samples were in the range of 1800 – 3300. Median read counts of each sample were as follows: 3095 and 3135 for Pool 1, 3139 and 3299 for Pool 2, 1822 and 1840 for Pool 3, 1802 and 1803 for Pool 4, and 2778 and 2856 for Pool 5.

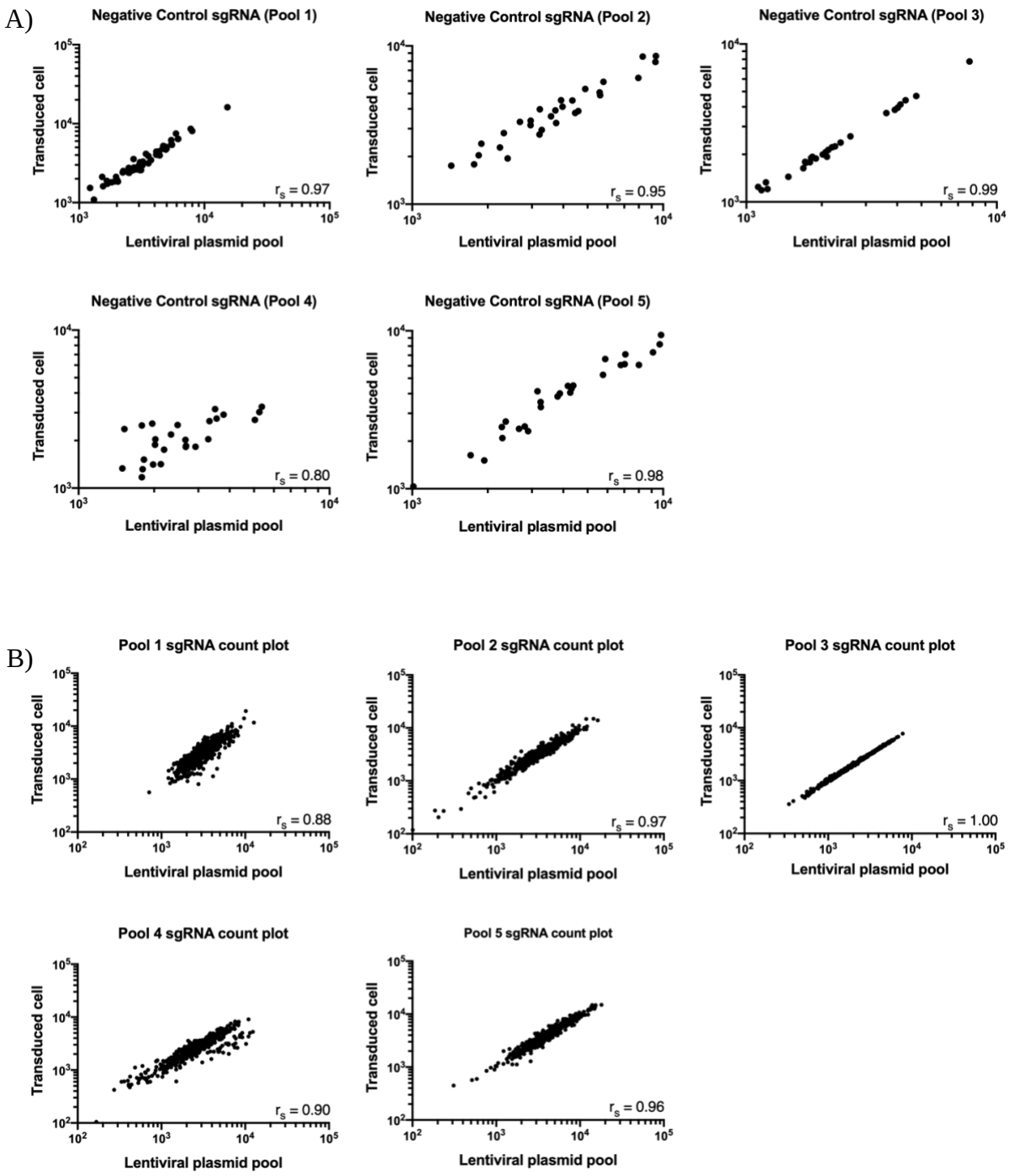


Figure 3-8 Comparison of read counts between lentiviral plasmid pools (x-axis) and transduced cells post 17 days of incubation (y-axis).

(A) Negative control sgRNA counts (non-targeting and *CD81*-targeting) of the cells at 17 days post transduction and the lentiviral plasmid libraries were plotted to visualize background noise. (B) All sgRNA read counts were compared between the cells at 17 days post transduction and the lentiviral plasmid libraries for each pool.

3.1.3. Optimisation of PCR methods to avoid contamination

Early in the project, several unsatisfactory sequencing results were obtained, in particular with high numbers of reads which did not map to the expected pool but were found to map to other pools (Figure 3-9A). In addition, significant losses of sgRNA sequences in samples were also recorded due to pipetting error during the process of transferring the sorted cells between multiple tubes prior to the PCR step.

To prevent sample contamination issues, the following measures were taken: 1) use of filtered pipette tips, 2) use of single-use aliquots of reagents, 3) aliquoting proportions of the samples at several NGS preparation intermediate steps for back-up, 4) use of DNA Away surface decontaminant to wipe bench areas and pipettes regularly, 5) changing gloves between every step, and 6) working on one sample at a time. As for preventing sample loss during a FACS sort, cells were sorted directly into an Eppendorf tube containing PBS with 10% FBS instead of a standard round-bottom polystyrene test tube so that the cells could be centrifuged as soon as the sort was finished without having to be transferred to another tube and risking loss of cells. These steps were successful in eliminating the issues, as shown in Figure 3-9B.

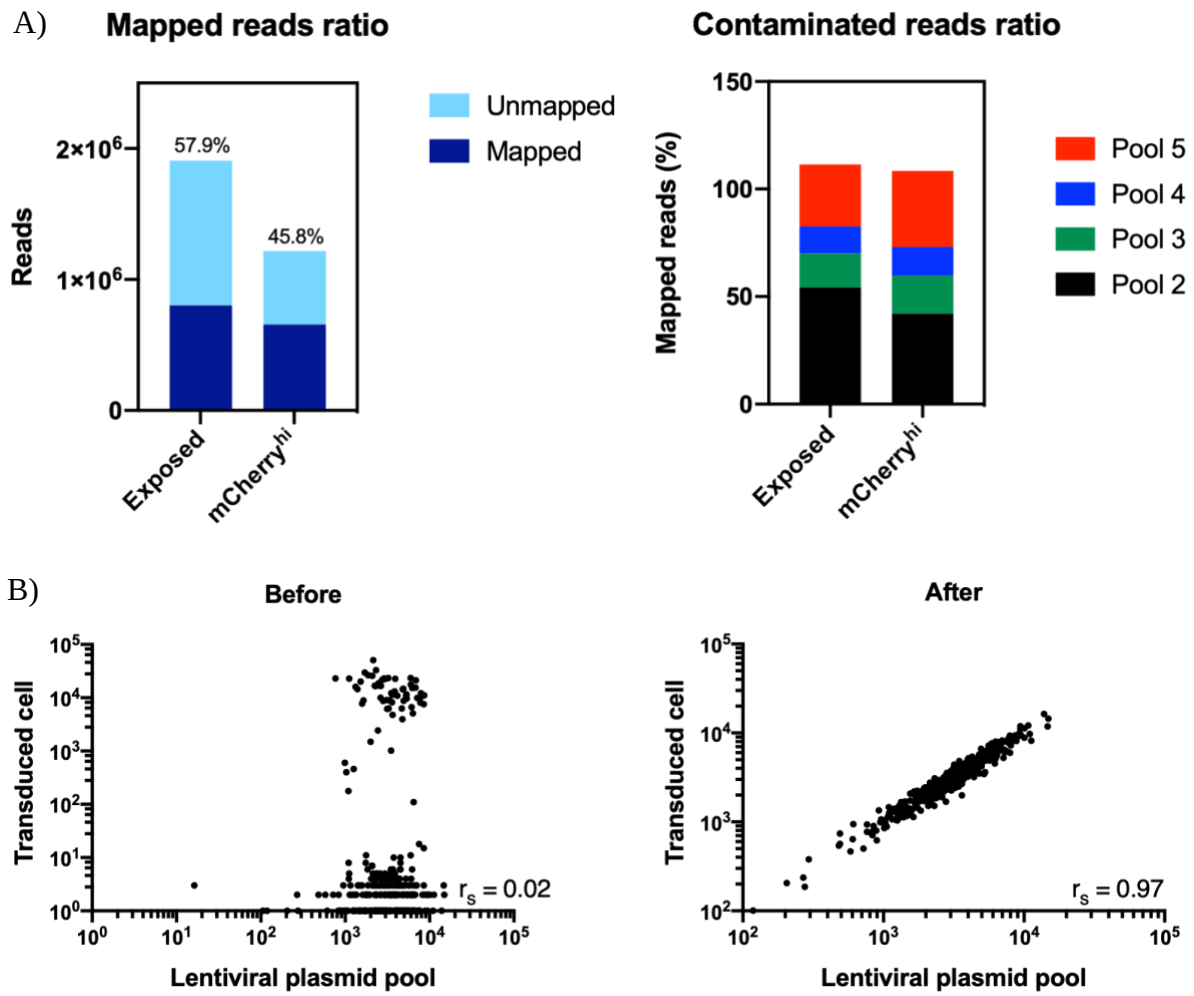


Figure 3-9 Improved sequencing data quality after PCR optimization.

(A) Left: unmapped percentages were 57.9% and 45.8% for the exposed and mCherry^{hi} samples respectively. Right: sgRNA reads were compared against Pool 3, 4 and 5. Pool 1 was excluded because all of Pool 1 experiments were completed prior to the optimization. Amongst the unmapped reads of the exposed sample, 15.9% were from Pool 3, 12.5% were from Pool 4 and 28.8% were from Pool 5. As for the mCherry^{hi} sample, 17.6% were from Pool 3, 13.4% were from Pool 4 and 35.3% were from Pool 5. Note that the total reads exceed 100% for both samples due to the overlapping control sgRNAs, which are around 8% of total reads. (B) A failed Pool 5 experiment (left) where a significant portion of the sorted cells were lost in transit. As a consequence, its r_s value was close to zero. The 'after' plot (right) showed correlation of the sgRNA read counts between the two samples from a successful Pool 5 invasion experiment after the PCR optimization.

3.1.4. Use of a mock phenotype to demonstrate ability of method to detect negative selection of sgRNAs targeting a control gene

The last step of the method development was using a readily detected phenotype to demonstrate CRISPR-Cas9 screen efficiency and reliability, and the MAGeCK programme's ability to detect under-represented sgRNAs and their target genes in a sample with a phenotype of interest, as compared to the parental control cells. Expression of the cell surface protein basigin (BSG), which its gene is targeted by a set of guides included in Pool 1, was chosen as a mock phenotype. BSG-binding antibody-stained cells (BSG+) exhibiting high median fluorescence intensity (MFI) on flow cytometry were sorted. These BSG+ cells were compared against unstained Pool 1 transduced cells (BSG-). The MAGeCK programme was able to detect *BSG* as the most negatively selected gene in this mock experiment, thus validating the robustness of this project's CRISPR-Cas9 screen and invasion experiment protocols, and the computational tool used for analyzing negatively selected genes (Figure 3-10).

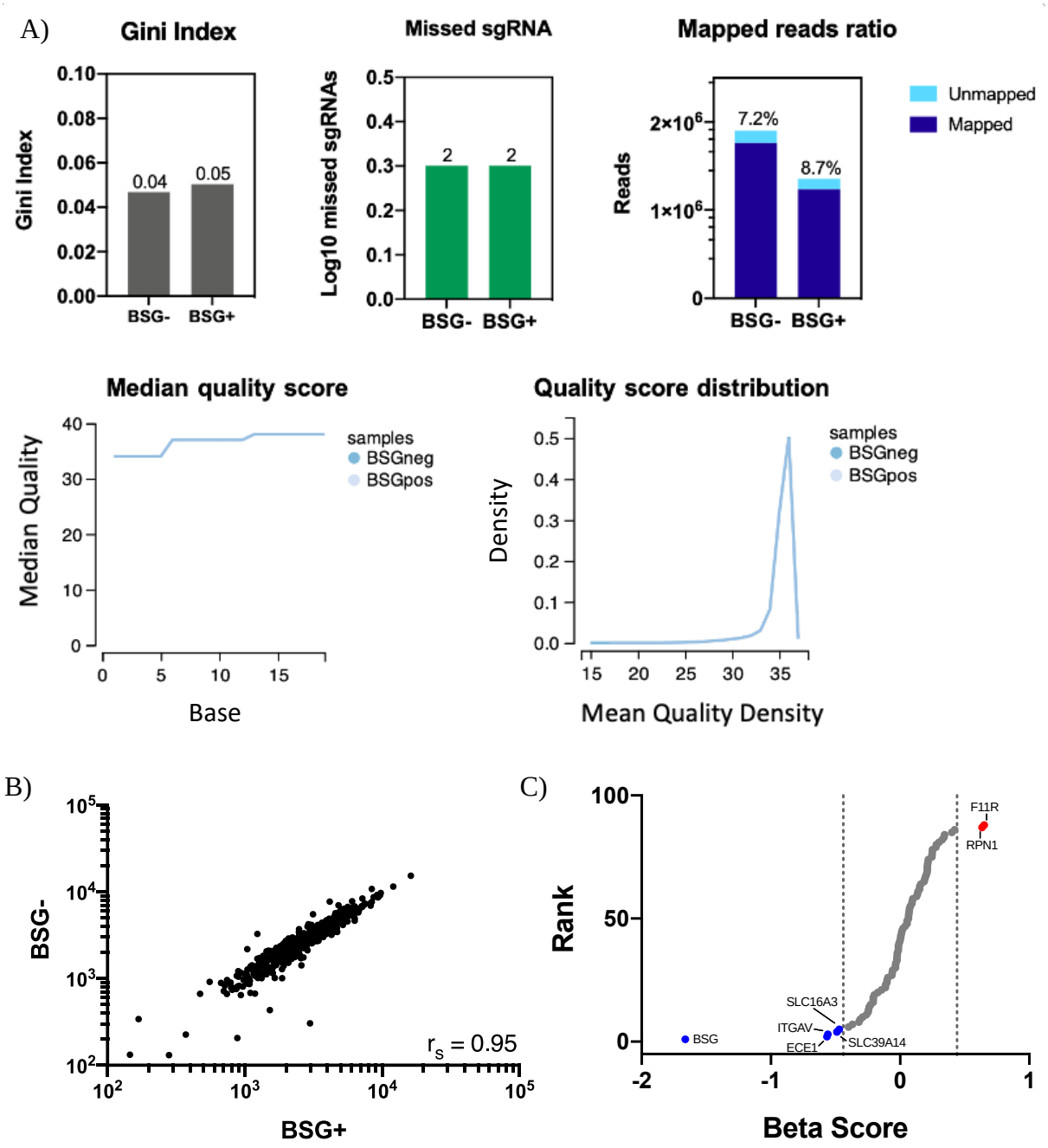


Figure 3-10 Use of a cell surface marker, basigin (BSG), to demonstrate the ability of the designed knock-out screen to detect a negatively selected gene.

(A) QC results showed that the samples were not contaminated and were sequenced at high quality. (B) sgRNA read count scatter plot of BSG+ and BSG- samples showed r_s=0.95. (C) BSG appeared as the most negatively selected gene (rank: 1; β-score: -1.66).

3.2. INVASION EXPERIMENTS

Having developed and validated a pipeline for producing pools of cells with CRISPR-Cas9 targeting genes of interest, and for detecting under-representation of sgRNAs responsible for a phenotype of interest, I proceeded to perform the screen with Pool 1-5 to identify hepatocyte genes that reduced sporozoite invasion when knocked out.

As described in Figure **2-3**, pools of lentivirus-transduced HC-04-Cas9 cells were produced and incubated for 15 days. On day 16, these cells were exposed to mCherry-expressing *P. berghei* sporozoites. 16-24 hrs later (day 17), mCherry^{hi} infected cells were sorted. As it takes approximately three days for sporozoites to asexually reproduce and produce merozoites, the parasites in the invaded hepatocytes after the said incubation time (16-24hrs) still remained as sporozoites.

Then, the sorted sample and its control, unsorted cells, underwent two stages of PCR (Figure **2-4**) for Illumina Sequencing. Results containing sgRNA reads of each sample were analyzed by several MAGeCK programmes (MAGeCK-VISPR, MAGeCK-Flute, MAGeCK-MLE and MAGeCK-RRA) to rank genes from most negatively selected to most positively selected. Overall, the CRISPR-Cas9 screens identified *ITGAV, RPN1, TMEM30A, ATP2B1, ITGB5, SLC35A2, MGAT1, FCGR2B, EMC1* and *APOH* as highly negatively selected hits. Ranks and β-scores of the genes from each pool are described in Appendix B: MAGeCK-MLE Batch Matrix gene rank .

3.2.1. FACS GATING STRATEGY FOR SORTING MCHERRY^{HI} CELLS

On day 17 of the invasion experiment, the sporozoite-exposed pool transduced cells were collected and only an mCherry^{hi} population was sorted. Calculation for determining the mCherry^{hi} population percentage is shown in Methods 2.2. Due to the infectivity rate variation between experiments, the mCherry^{hi} percentage was different for each experiment. However, other gates for selecting live (FSC-A vs SSC-A) and singlet (FSC-W vs FSC-H) cells remained the same throughout all invasion experiments (Figure 3-11).

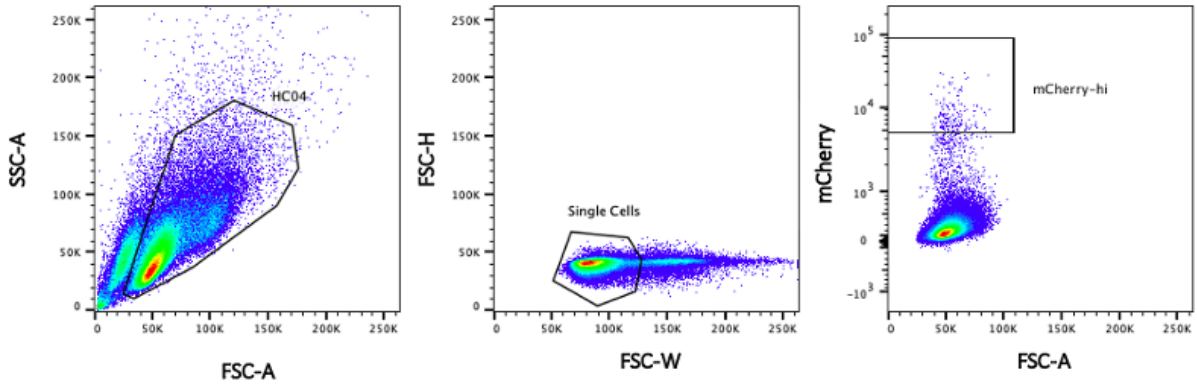


Figure 3-11 An example of flow cytometry gating strategy for sorting mCherry^{hi} cells. (Left) FSC-A vs SSC-A gate was applied around a cell population to discriminate dead cells and debris. (Middle) FSC-W and FSC-H gate was applied to differentiate singlets from doublets. (Right) mCherry^{hi} gate was applied around a cell population percentage as calculated by an equation described in Methods 2.2.

3.2.2. Overview of invasion experiments

I aimed to perform the screen in duplicate, i.e., to obtain results from two successful experiments for each of the five sgRNA pools. Due to the complexity of the experimental pipeline, careful evaluation of results was essential in order to discriminate successful, informative results from those experiments which suffered a technical flaw.

Five samples per pool were used for QC analysis: unexposed, exposed 1, exposed 2, mCherry^{hi} 1 and mCherry^{hi} 2. The unexposed samples were the transduction cells used previously for

validating the qualities of lentiviral pools and transduced cells (Results 3.1.2). The number that follows the sample names represents a replicate experiment. Once the QC data were evaluated as clean, MLE Batch Matrix analyses followed where 5 samples were used: unexposed 1, exposed 1, exposed 2, mCherry^{hi} 1 and mCherry^{hi} 2.

To be judged informative, each experiment had to fulfil the following criteria, unless otherwise justified or stated:

1) Number of sorted infected cells must be greater than 2.5×10^4 to achieve a minimum of 1:50 sgRNA to cell ratio.

2) Samples must pass QC by evaluation of their Gini index, missed sgRNAs, unmapped and mapped reads, sequencing quality score and base quality scores.

3) Low background noise as represented by $r_s > 0.8$ for analysis of correlation of negative control guide read frequency in sample vs. reference.

4) False discovery rate at <0.05 for positive control gene, *SCARB1*.

All pools contained the same positive and negative control sgRNAs. *SCARB1* gene which encodes scavenger receptor class B type 1 (SR-B1) was used as a positive control as it is known to be essential for *P. berghei* sporozoite invasion into HC-04 or HepG2 cells (Rodrigues et al., 2008). *SCARB1* was therefore expected to appear as one of the most negatively selected genes. *SCARB1* served as a measurement of a successful invasion experiment – those with *SCARB1* as one of the most negatively selected genes with significantly low FDR (<0.05) were considered informative.

Because the *SCARB1* statistical data were of the highest importance, both MLE and RRA analyses results of each experiment were referenced by evaluating its rank, p-value and FDR (Table 3-1). As for the negative control, there were non-targeting sgRNAs and *CD81* targeting sgRNAs. As the name suggests, the non-targeting sgRNAs were not designed to target any genes in the genome. These sgRNAs account for the potential off-target effects but not for the gene knockout efficiency. On the other hand, although *CD81*-targeting sgRNAs were included, because the HC-04 cell line does not express the CD81 protein, and *P. berghei* sporozoites do not need this receptor for invasion (Silvie, Franetich, Boucheix, Rubinstein, & Mazier, 2007), no gene drop-out effects were expected to be observed, thus serving as a second negative control. These negative control sgRNAs were included in sgRNA normalization process, which is a step before analyzing negatively selected genes by the MAGeCK programmes (C. H. Chen et al., 2018).

Due to variations in numbers and fitness of sporozoites, the number of mCherry^{hi} cells was different for each experiment. The exposed control samples, on the other hand, remained consistent at 3×10^5 cells with the exception of Pool 1, which had 8×10^4 cells. This was because there were several protocol optimizations attempts after the Pool 1 experiments were completed, which included increasing the number of control cells. This change, however, did not seem to affect the statistical power of *SCARB1* as its β-scores varied across the pools and there were no signs of any positive correlation between the control cell number and the *SCARB1* statistical power. Nonetheless, after Pool 1, 3×10^5 cells, rather than 8×10^4, were prepared for Illumina Sequencing as it was easier to perform PCR on a larger cell sample.

			Pool 1	Pool 2	Pool 3	Pool 4	Pool 5
Experiment 1		# of sorted cells	20,000	34,000	40,000	20,000	38,000
		r_s	0.95	0.93	0.88	0.94	0.94
	RRA	p-value	0.011	0.00039	0.0029	5×10^{-6}	5×10^{-6}
		FDR	0.00045	0.014	0.35	0.00062	0.00047
		Log2FC	-0.79	-1.23	-1.24	-0.96	-0.78
	MLE	p-value	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	0.05	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$
		FDR	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	0.11	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$
		β-Score	-0.51	-0.72	-1.00	-0.57	-0.52
Experiment 2		# of sorted cells	20,000	34,000	23,000	25,000	21,000
		r_s	0.96	0.94	0.8	0.94	0.85
	RRA	p-value	5×10^{-6}	0.00046	5×10^{-6}	0.0034	5×10^{-6}
		FDR	0.000225	0.00187	0.000599	0.091646	0.000469
		Log2FC	-0.74	-1.22	-1.55	-0.61	-2.25
	MLE	p-value	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$
		FDR	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$
		β-Score	-0.46	-0.80	-1.30	-0.42	-1.29

Table 3-1 MAGeCK-MLE and MAGeCK-RRA statistical data of *SCARB1* from each replicate experiment of Pool 1, 2, 3, 4 and 5.

3.2.3. CRISPR-Cas9 screen identified hits from Pool1, 2 and 4 but not from 3 and 5

Genes with a β-score higher or lower than 2 standard deviations calculated by MAGeCK-MLE Batch Matrix were highlighted as blue for negative selection, and red for positive selection. Since the focus of this project was on the negatively selected genes, the positively selected genes were exempt from further analysis. Then, each gene hit underwent an additional round of manual evaluation by 1) analyzing their individual p-value and FDR, 2) comparing their statistical data to *SCARB1*, and 3) referencing their decreasing sgRNA read counts between samples.

As described in Results 3.2.2, the identification of *SCARB1* as a significantly negatively selected gene was a pre-requisite for experiments to be considered informative (Table 3-1). Each hit from the MAGeCK-MLE Batch Matrix analyses was evaluated to see whether it was significant as well in each replicate experiment (Table 3-2). MAGeCK-RRA analysis was added for comparison.

		MAGeCK-MLE				MAGeCK-RRA			
Pool	Gene	β Score	P-value	FDR	Significance in replicates	Log2 FC	p-value	FDR	Significance in replicates
1	SCARB1	-0.57	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1,2	-0.64	5×10^{-6}	0.00045	1,2
	ITGAV	-0.61	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1, 2	-0.79	0.00014	0.0065	1, 2
	RPN1	-0.45	0.14	0.33	1	-0.24	0.062	0.70	0
	TMEM30A	-0.35	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1	-0.31	0.0016	0.05	2
	ATP2B1	-0.30	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	2	-0.28	0.0045	0.10	0
2	SCARB1	-0.65	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1, 2	-1.23	3×10^{-5}	0.0019	1, 2
	ITGB5	-0.85	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1, 2	-0.79	5×10^{-6}	0.00062	1, 2
	SLC35A2	-0.49	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1, 2	-0.79	5×10^{-5}	0.0019	1, 2
	MGAT1	-0.46	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1	-0.96	0.00044	0.014	1, 2
	FCGR2B	-0.39	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1	0.01	0.012	0.21	0
3	SCARB1	-1.10	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1, 2	-1.55	5×10^{-6}	0.00060	1,2
4	SCARB1	-0.43	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1, 2	-0.62	5×10^{-6}	0.00062	1,2
	EMC1	-0.34	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1,2	-0.41	0.00044	0.028	1,2
	APOH	-0.29	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1	-0.21	0.011	0.071	0
5	SCARB1	-0.91	$< 2.2 \times 10^{-308}$	$< 2.2 \times 10^{-308}$	1,2	-0.78	5×10^{-6}	0.00047	1,2

Table 3-2 MAGeCK-MLE and MAGeCK-RRA statistical analysis results of *SCARB1* and identified hits from each pool.

The identified hits were based on MAGeCK-MLE Batch Matrix analysis. For individual analysis, Unexposed, exposed and mCherry^{hi} samples from each replicate were used for MAGeCK-MLE analysis, and exposed and mCherry^{hi} for MAGeCK-RRA analysis.

3.2.3.1. POOL 1

Pool 1 contained 5 sgRNAs targeting *SCARB1* as a positive control, 5 sgRNAs targeting *CD81* as a negative control of the invasion experiment, 50 non-targeting sgRNAs as a negative control of the CRISPR-Cas9 screen, and 440 sgRNAs targeting 88 genes of interest. Based on the reads quality control data (Figure 3-12), there was a low level of sgRNA count inequality across the samples and no indication of contamination or significant sgRNA read losses.

ITGAV, RPN1, ATP2B1 and *TMEM30A* were identified as hits (Figure 3-13A). For all SCARB1-targeting sgRNAs, there was no substantial decrease from unexposed to exposed, but a sharp decrease from exposed to mCherryhi. As expected of SCARB1, this showed an under-representation of SCARB1 in the mCherryhi cells. Other hits exhibited a trend similar to that of SCARB1 as well (Figure 3-13B).

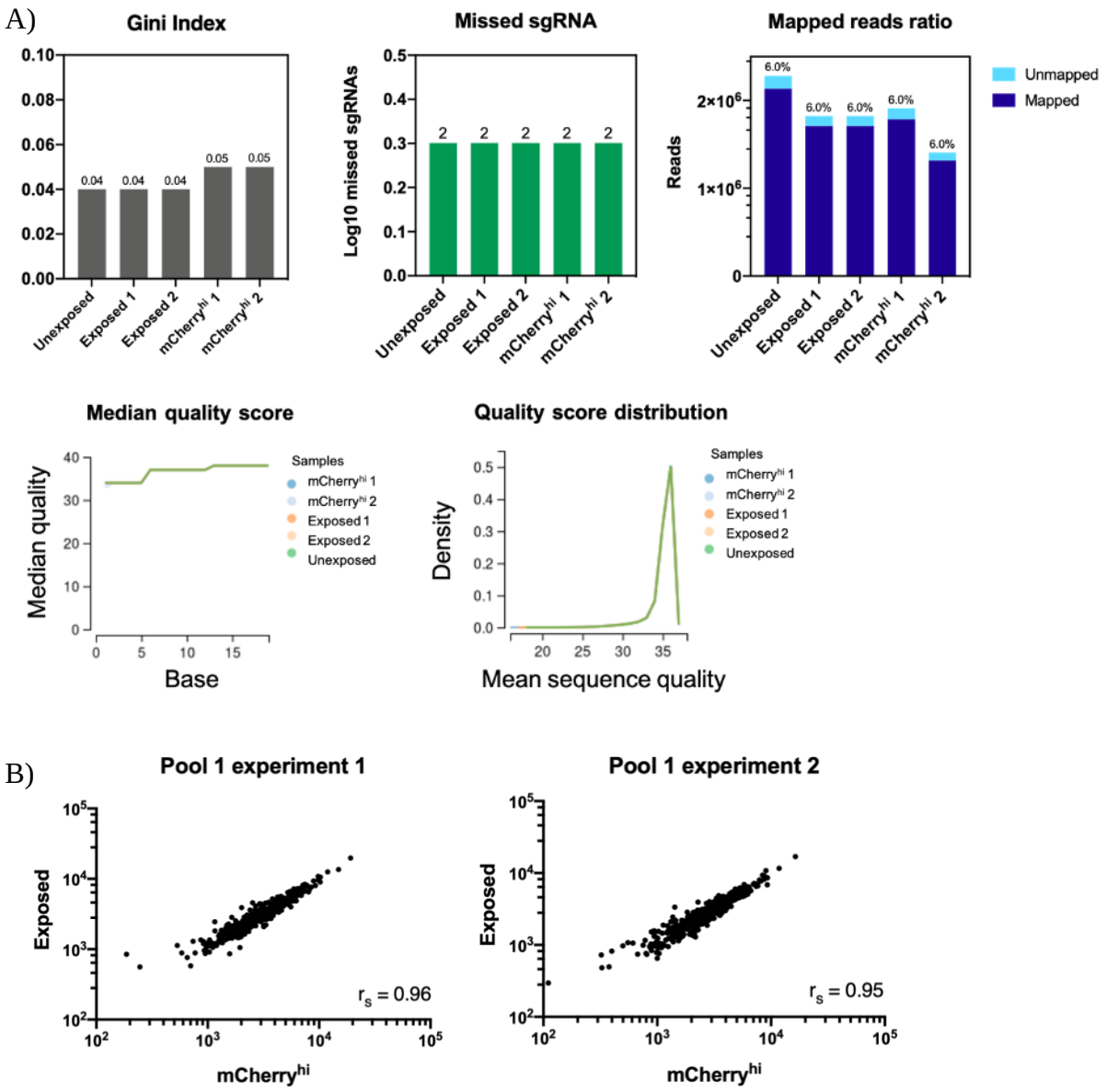


Figure 3-12 Quality Control analysis of Pool 1 replicate 1 and replicate 2 experiments. (A) Gini index scores were near zero. The unmapped reads ratios were not only lower than 30% but also similar across the samples. Both the median base quality score and sequence quality score showed a high peak at 35 and there are no signs of signal decay. (B) Spearman's coefficients of the replicates were 0.96 and 0.95, an indication of low background noise.

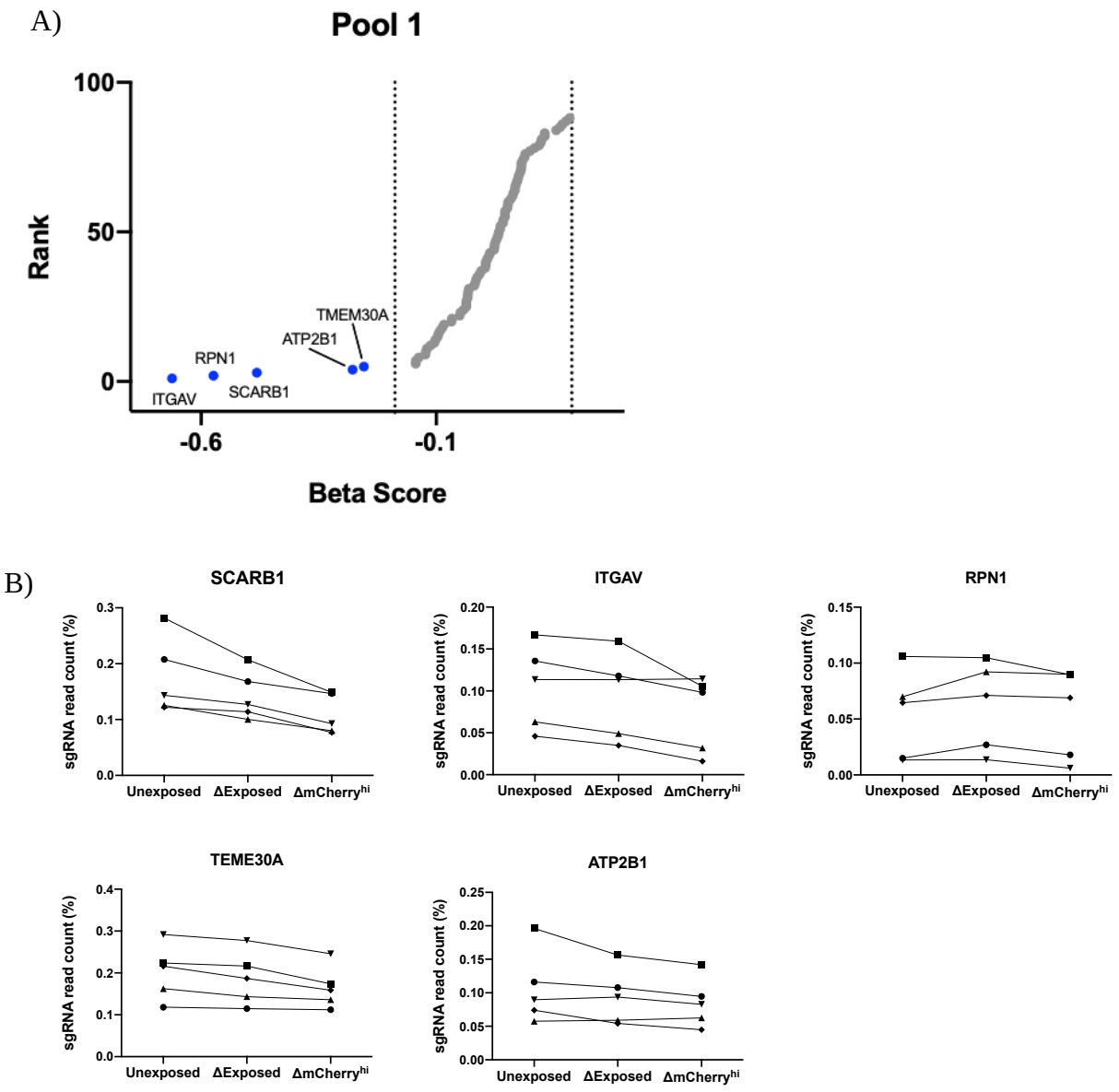


Figure 3-13 *ITGAV, RPN1, TMEM30A* and *ATP2B1* identified as hits from Pool 1.
(A) β-score standard deviation cut-offs (dotted lines) were -0.186 and 0.186. MLE batch matrix rank plot identified *ITGAV* (rank: 1; β-score: -0.66), *RPN1* (rank: 3; β-score: -0.57), *ATP2B1* (rank: 4; β-score: -0.28), and *TMEM30A* (rank: 5; β-score: -0.25) were identified as hits. (B) Read counts for the five individual sgRNAs targeting each gene were calculated as percentages of the total reads in the unexposed, exposed and mCherry^{hi} samples.

3.2.3.2. POOL 2

Pool 2 contained 5 sgRNAs targeting *SCARB1* as a positive control, 5 sgRNAs targeting *CD81* as a negative control of the invasion experiment, 25 non-targeting sgRNAs as a negative control of the CRISPR-Cas9 screen, and 465 sgRNAs targeting 98 genes of interest. Based on the sgRNA read quality control results, no substantial losses of sgRNA reads were seen and there were no signs of sample contamination (Figure 3-14).

FCGR2B, *ITGB5*, *SLC35A2* and *MGAT1* were identified as significantly negatively selected hits (Figure 3-15A). *SCARB1*-targeting sgRNAs showed a small decrease from unexposed to exposed but then a sharp decrease from exposed to mCherry^{hi}. *SLC35A2* and *MGAT1* followed a similar pattern. All of *ITGB5*-targeting sgRNAs exhibited a sharp decrease from unexposed to exposed, and then an additional one from exposed to mCherry^{hi}. Only two of the *FCGR2B*-targeting sgRNAs decreased from exposed to mCherry^{hi}, thus questioning its validity as a hit (Figure 3-15B).

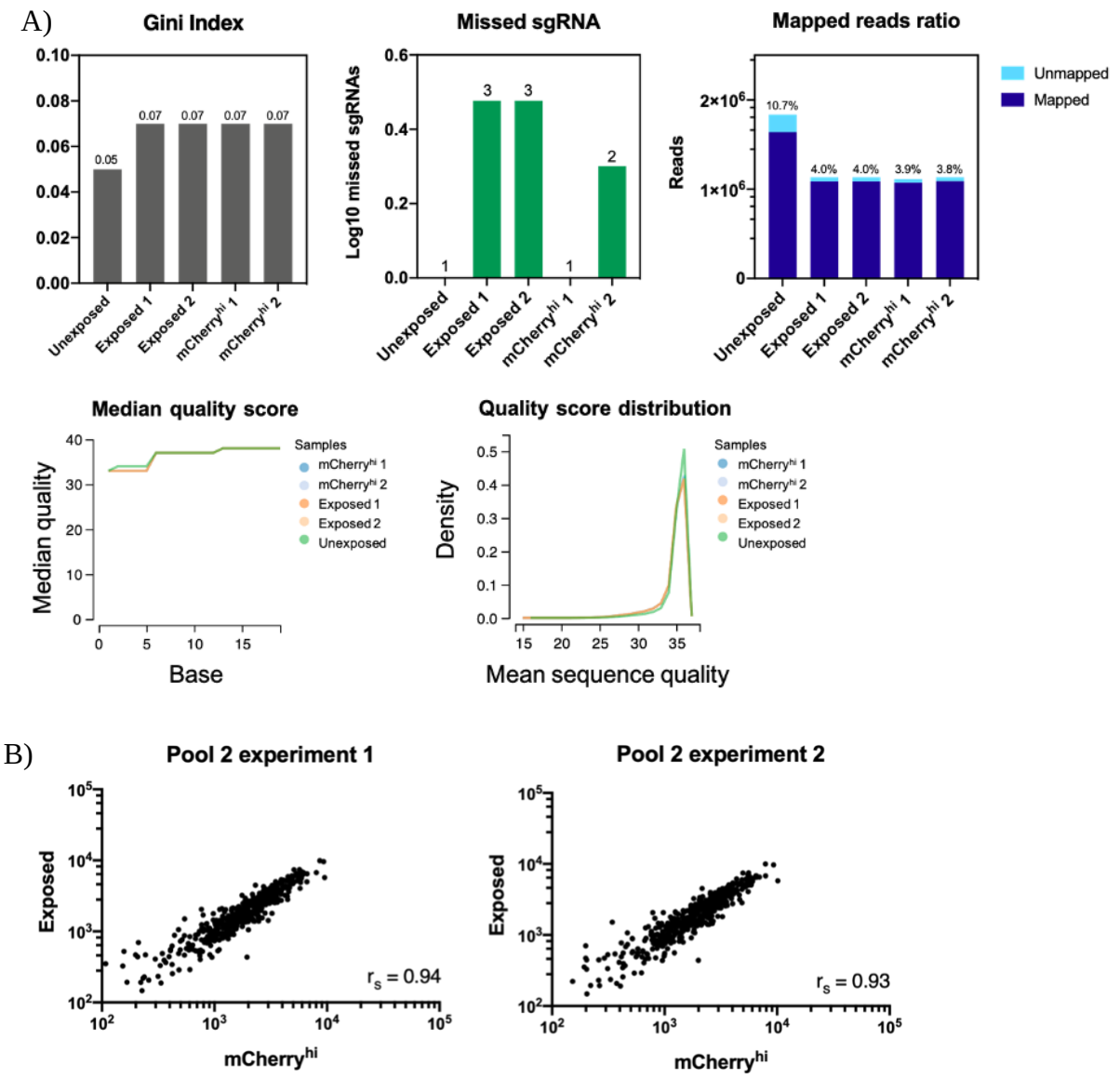


Figure 3-14 Quality Control analysis of Pool 2 replicate 1 and replicate 2 experiments. (A) The Gini index scores were near zero for all samples. The unmapped ratios of the unexposed, exposed and mCherry^{hi} replicate samples were below 30%. Both base median quality score and mean sequence quality score showed a high peak at 35. (B) Spearman's rank correlation coefficients of the replicates were 0.94 and 0.93.

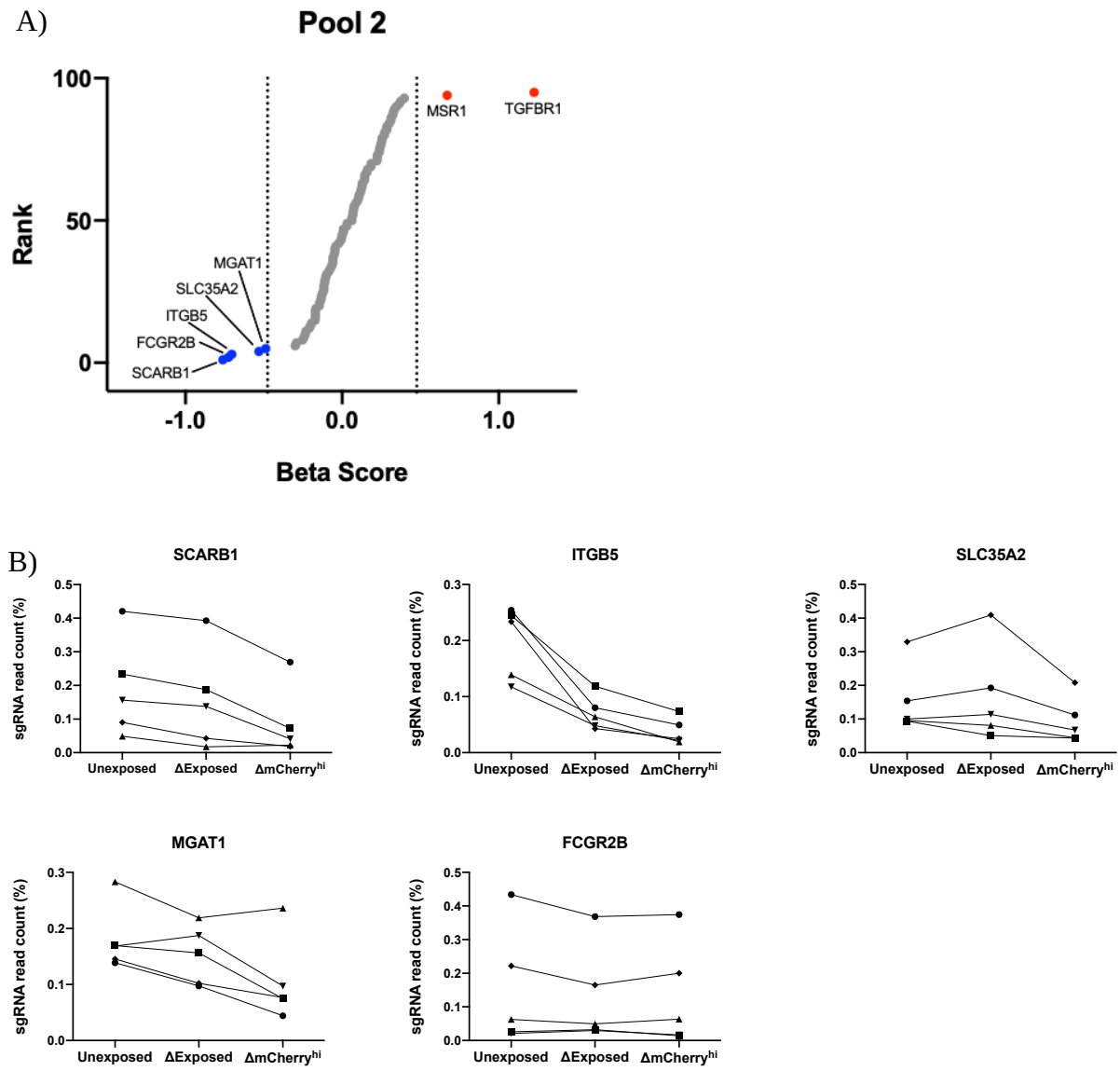


Figure 3-15 *ITGB5, SLC35A2, MGAT1* and *FCGR2B* identified as hits from Pool 2.
(A) β-score standard deviation cut-offs (dotted lines) were –0.48 and 0.48. *ITGB5* (rank: 1; β-score: -0.85), *SLC35A2* (rank: 3; β-score: -0.49), *MGAT1* (rank: 4; β-score: -0.46) and *FCGR2B* (rank: 5; β-score: -0.39) were identified as hits. (B) Read counts for the five individual sgRNAs targeting each gene were calculated as percentages of the total reads in the unexposed, exposed and mCherry^{hi} samples.

3.2.3.3. POOL 3

Pool 3 contained 5 sgRNAs targeting *SCARB1* as a positive control, 5 sgRNAs targeting *CD81* as a negative control of the invasion experiment, 25 non-targeting sgRNAs as a negative control of the CRISPR-Cas9 screen, and 465 sgRNAs targeting 93 genes of interest. Based on the sgRNA read quality control analysis (Figure 3-16), the Gini indexes of both replicates of mCherry^{hi} were nearly double that of the unexposed sample, which indicated that either there were certain sgRNA sequences that dominated in the sorted cells or some sgRNA sequences were lost during cell culture. Moreover, the 'mCherry^{hi} 2' mean sequence quality peak was lower in read density than the other samples, which indicated a low loading input of the 'mCherry^{hi} 2' sample in the sequence run. This was reflected on the Spearman's coefficient analysis, which its score was the lowest amongst all invasion experiments. The background noise was suspected to be due to its lower sequencing quality and higher Gini index score.

No significant hits were identified from Pool 3 (Figure 3-17).

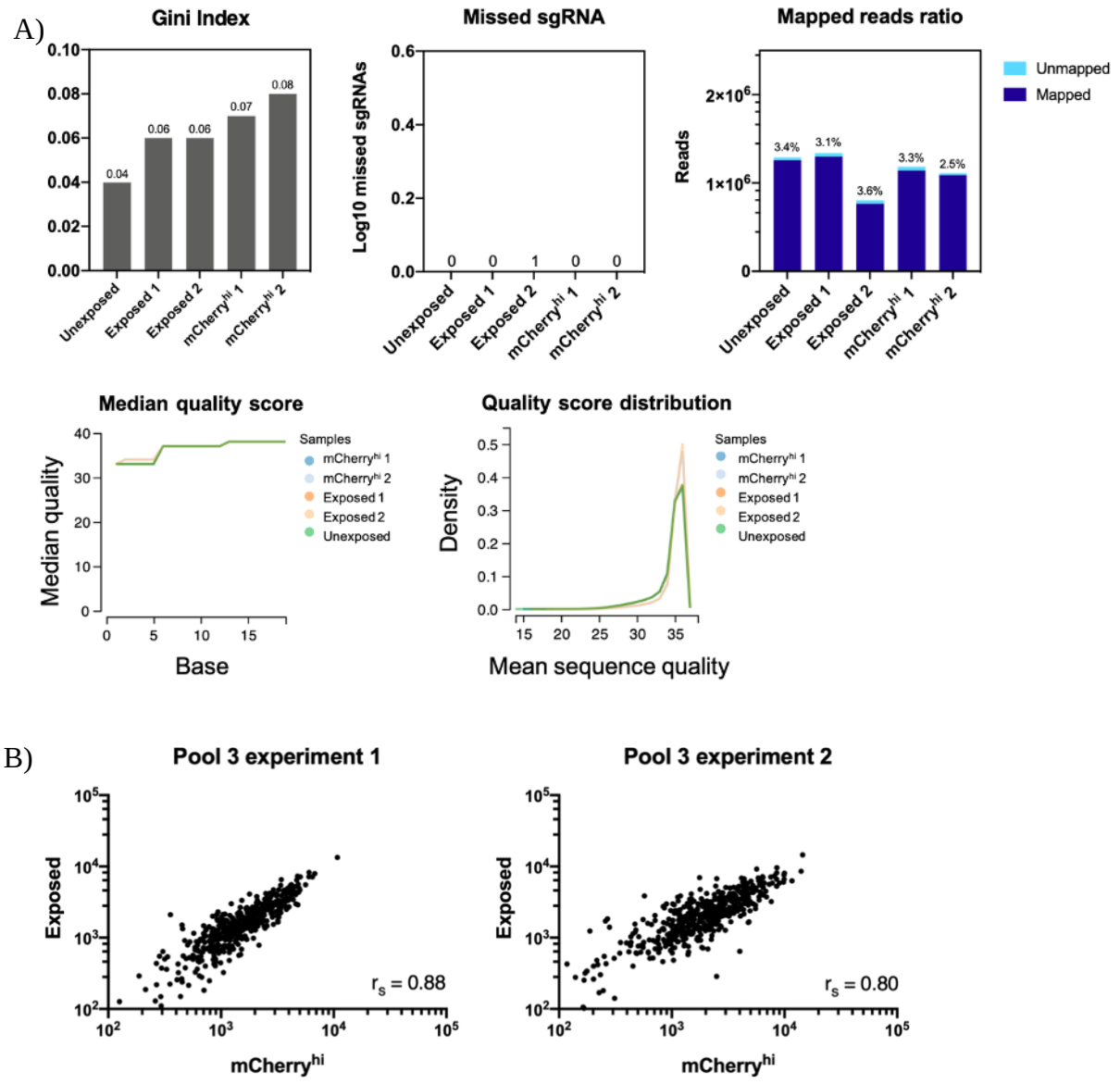
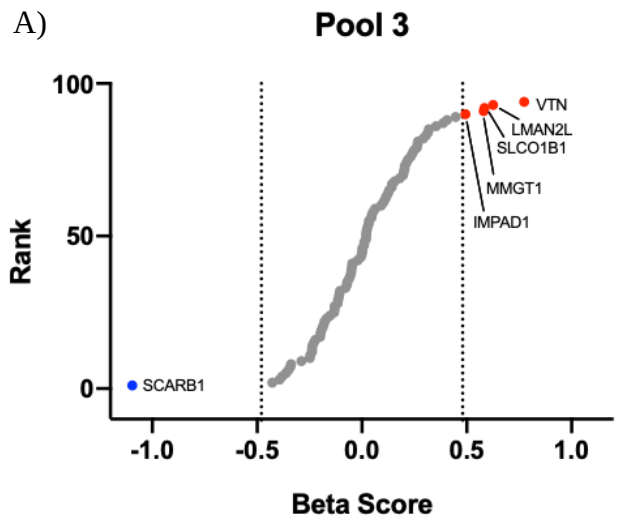


Figure 3-16 Quality Control analysis of Pool 3 replicate 1 and replicate 2 experiments. (A) Gini index scores varied among the samples – 0.04 for unexposed, 0.06 for exposed and 0.07-0.08 for mCherry^{hi} samples. Except for 'exposed 2', there were no missing sgRNAs. There were no signs of sample contamination as the unmapped ratios are extremely low. The base median quality score and the mean sequence quality scores exhibited a peak at 35, indicating that there was no signal decay. (B) Spearman's rank correlation coefficient values of the replicates were 0.88 and 0.80.

A)



B)

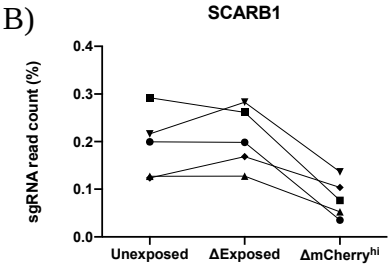


Figure 3-17 No hits identified from Pool 3.

(A) β-score standard deviation cut-offs (dotted lines) were –0.48 and 0.48. (B) SCARB1 targeting sgRNAs exhibit an expected pattern as described on the previous figures.

3.2.3.4. POOL 4

A)

Pool 4 contained 5 sgRNAs targeting *SCARB1* as a positive control, 5 sgRNAs targeting *CD81* as a negative control of the invasion experiment, 25 non-targeting sgRNAs as a negative control of the CRISPR-Cas9 screen, and 465 sgRNAs targeting 94 genes of interest. *EMC1* and *APOH* were identified as significantly negatively selected hits.

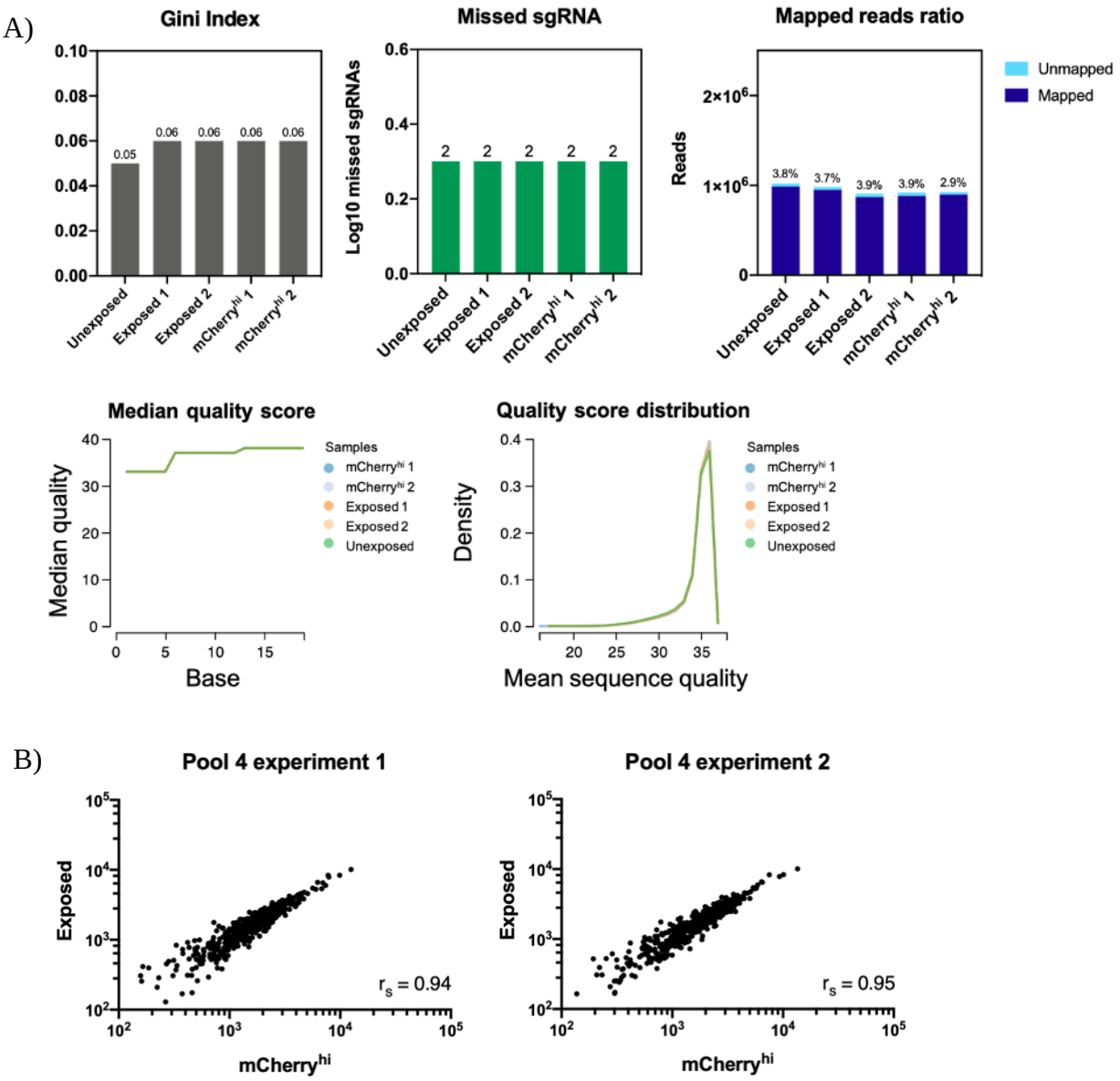


Figure 3-18 Quality Control analysis of Pool 4 replicate 1 and replicate 2 experiments. (A) Gini index scores were 0.05 for unexposed and 0.06 for the rest of the samples. The unmapped ratios were extremely low thus there were no signs of sample contamination. The

sequencing quality was high for all samples as their base median quality score and mean sequence quality scores showed a peak at 35. (B) Spearman's coefficient values were 0.94 and 0.95 for the replicates, indication of low background noise.

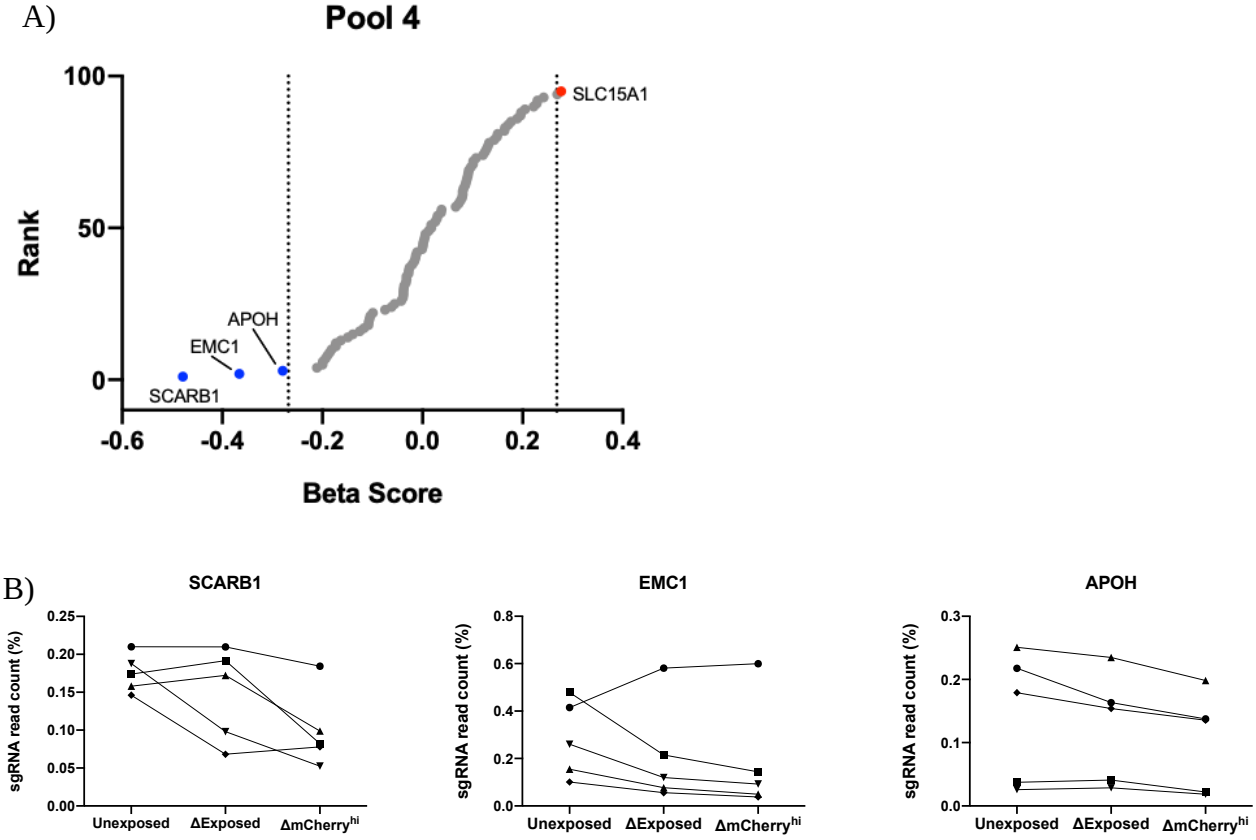


Figure 3-19 *ATP6AP2, EMC1* and *APOH* identified as hits from Pool 4.
(A) β-score standard deviation cut-offs (dotted lines) were –0.27 and 0.27. *EMC1* (rank: 2; β-score: -0.37) and *APOH* (rank: 3; β-score: -0.28) were identified as hits. (B) *SCARB1* sgRNAs exhibited a decreasing pattern as expected with the exception of two sgRNAs. With the exception of one of the EMC1 sgRNAs that exhibited an increasing trend, the other four sgRNAs showed a steady decline across the samples. For *APOH*, all sgRNAs exhibited a steady overall decreasing trend.

3.2.3.5. POOL 5

Pool 5 contained 5 sgRNAs targeting *SCARB1* as a positive control, 5 sgRNAs targeting *CD81* as a negative control of the invasion experiment, 25 non-targeting sgRNAs as a negative control of the CRISPR-Cas9 screen, and 465 sgRNAs targeting 91 genes of interest. No hits were identified.

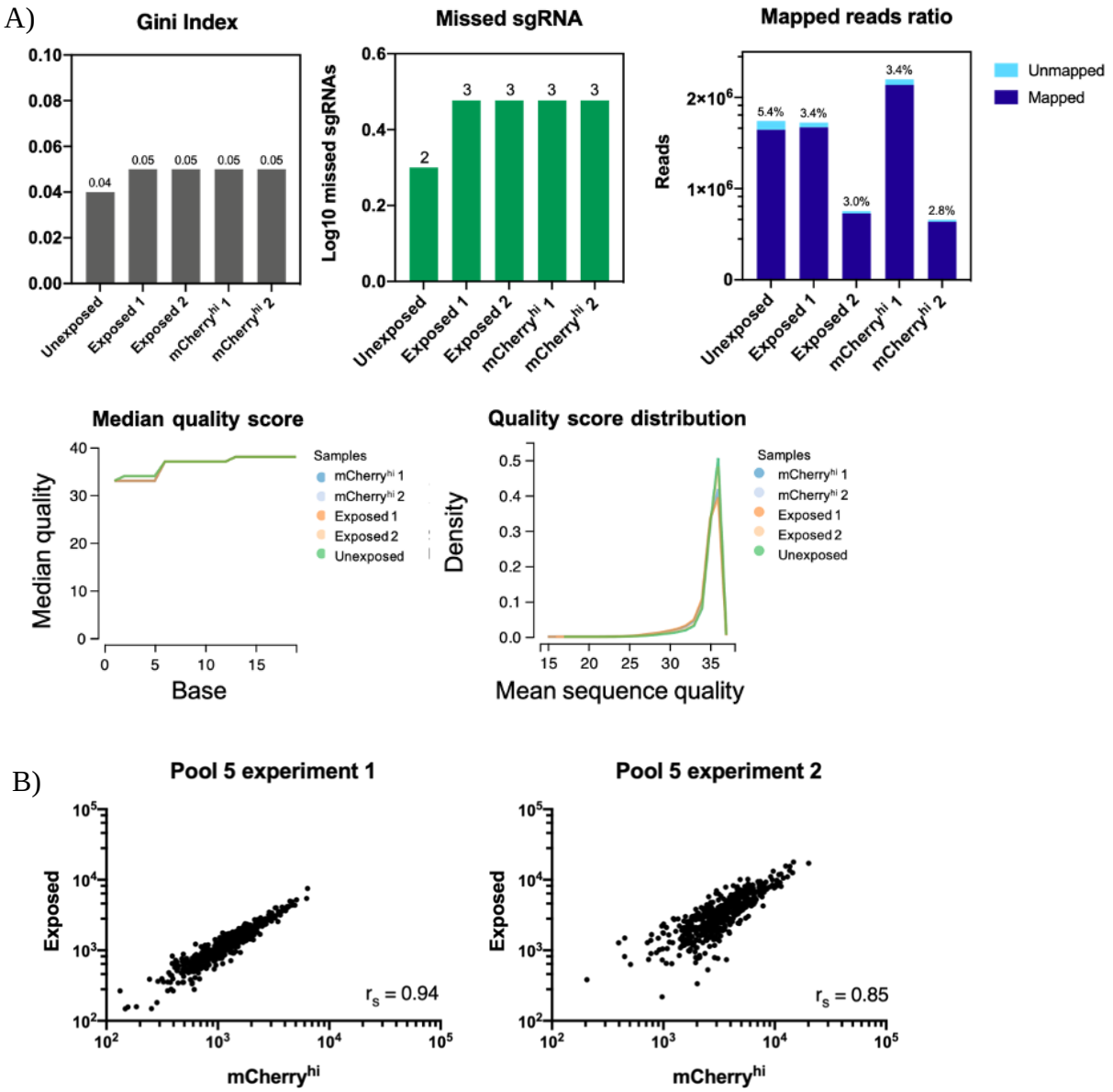
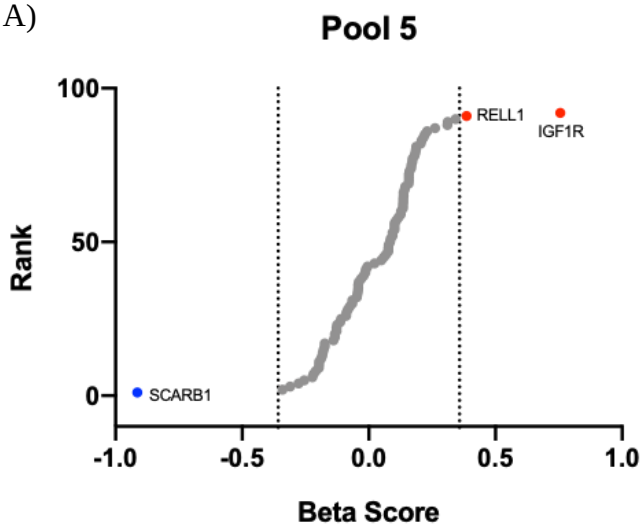


Figure 3-20 Quality Control analysis of Pool 5 replicate 1 and replicate 2 experiments. (A) Gini index scores were low at 0.04 for unexposed and 0.05 for all other samples. All samples exhibited an unmapped read ratio of less than 30%. The sequencing quality of the

samples was high as both base median quality score and mean sequence quality score exhibited a peak at 35. (B) Spearman's coefficients were 0.94 and 0.85 for the replicates, an indication of low background noise.

A)



B)

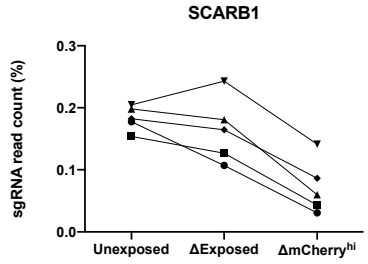


Figure 3-21 No hits were identified from Pool 5.
(A) β-score standard deviation cut-offs (dotted lines) were –0.36 and 0.36. No hits, other than SCARB1, were identified. (B) All five *SCARB1* sgRNAs exhibited an expected decreasing trend across the three samples.

3.3. DISCUSSION

The MAGeCK programme provides two gene rank analysis methods for both positive and negative selections: RRA and MLE. As the total NGS reads are different for each sample per NGS library submission, both RRA and MLE methods normalize sgRNA read counts before using them to rank genes from most negatively selected to most positively selected. Initially, only the RRA method was used for data analysis and planning follow-up experiments. However, after finding the improved version, MAGeCK-MLE Batch Matrix, all data were re-analyzed. This method sought out new valid hits and dropped statistically false positives (as opposed to biologically false positives) by implementing sgRNA knock-out efficiency scores and the ability to compare more than two conditions. Individual hits identified from the MAGeCK-MLE Batch Matrix analyses were compared to each non-Batch Matrix MLE and RRA-analyzed replicate results (Table 3-2). If a hit showed as significant in all analysis methods, it was considered to have a higher chance of being essential in the hepatocyte invasion than a hit that showed as significant in only some.

While only one control (exposed) was used for the RRA analysis, two controls (unexposed and exposed) were used for the MLE analysis. Adding another control sample removed outliers in CRISPR-Cas9 knock-out screens. For example, *ACP2* was originally presented as the most negatively selected gene in Pool 4 experiment 1 using the RRA method. However, this was because one of the *ACP2*-targeting sgRNAs' count number dropped by 98% compared to the other four that presented a neutral sgRNA read count change. Although this particular sgRNA showed a significant drop, because its read count was much less than the other four sgRNAs by approximately 95% to begin with, it is very likely that the *ACP2* hit was not due to an invasion failure. Hence, this hit was an outlier that skewed the statistical analyses, hence hiding

the true positive hits. To resolve this issue and correct the gene rank list, an additional sample, unexposed, was used for the comparison. Doing so removed *ACP2* from the negatively selected gene list, corrected the final list and uncovered hidden hits, which will be elaborated later.

MAGeCK-Flute is another version of the MAGeCK programme that features additional downstream analyses that allow comparisons between multiple replicates by using either RRA or MLE method. One of the features is called 'batch effect removal' that corrects the sgRNA read counts in samples from the experiments performed in batches. In a CRISPR screen, a batch effect occurs when invasion experiments or Illumina Sequencing library submission preparations were performed using different resources or at different times. These may alter the data with non-biologically related hits and unwanted variations between replicates that lead to inaccurate conclusions (Leek et al., 2010). MAGeCK-Flute hence removes this batch effect in the sgRNA read counts. The corrected counts can then be used by either RRA or MLE method for gene rank analysis. Another feature of MAGeCK-Flute is called 'batch matrix' which compares biological replicates to each other at equal conditions (i.e., 2 replicates of exposed and mCherry^{hi} samples for RRA; 2 replicate samples of unexposed, exposed, and mCherry^{hi} for MLE).

The MAGeCK programme identified the following genes as significantly negatively selected by using the MLE Batch Matrix method: *ITGAV*, *RPN1*, *TMEM30A* and *ATP2B1* from Pool 1 (Results 3.2.3.1), *ITGB5*, *SLC35A2*, *MGAT1* and *FCGR2B* from Pool 2 (Results 3.2.3.2), and *EMC1* and *APOH* from Pool 4 (Results 3.2.3.4). With the exception of *RPN1*, all hits' p-value and FDR analyses were deemed as significant (Table 3-2). As such, although there was a high chance that *RPN1* was a false positive, *RPN1* was nevertheless included in the individual hit literature analysis.

sgRNA read counts of the significant hits were lower in the FACS-sorted mCherry^{hi} cell population as opposed to unexposed and exposed cells. However, it is difficult to tell solely based on the CRISPR-Cas9 screen whether these results were due to unsuccessful invasion by the sporozoites or other unknown biological or mechanical factors. As the experiment was designed to observe approximately 22 hours post-invasion, some of the hits were identified to be involved in post-invasion EEF development, rather than the invasion event itself. This conclusion was further solidified by analyzing cellular location of the hits being other than plasma membrane, which included several organelles such as the endoplasmic reticulum and Golgi apparatus.

Integrin αV (*ITGAV* from Pool 1; see Figure 3-13) is located in the cell membrane as a heterodimer with a different integrin monomer. For example, αV-integrin binds to β3-integrin and forms αVβ3, which mediates cell-to-cell and cell-to-matrix adhesions by binding to molecules with an arginine-glycine-aspartate (RGD) motif (Ruoslahti, 1996). αVβ3 was identified by Dundas et al. to be a direct binding partner of an essential *P. falciparum* sporozoite surface protein for gliding motility, *Pf*TRAP, and that it plays a role in gliding motility but not in hepatocyte invasion (Dundas et al., 2018).

Similarly, integrin β5 (*ITGB5* from Pool 2; see Figure 3-15), which forms a heterodimer, αVβ5, with *ITGAV*, showed as negatively selected from Pool 2. However, unlike αVβ3, it is not known whether αVβ5 also plays an essential role in cell adhesion and binds to adhesive ligands with an RGD motif. Although β5-integrin's role in liver invasion is unknown, Dundas et al. have shown weakly positive *Pf*TRAP-αVβ5 interaction from an AVEXIS screen (Dundas et al., 2018).

Dolichyl-diphospholigosaccharide-protein glycosyltransferase subunit 1 (RPN1) is a rough endoplasmic reticulum (ER) resident protein that transfers asparagine residues to the growing high-mannose oligosaccharides, to produce N-linked glycans on proteins and lipids. RPN1 also takes part in the unfolded protein response (UPR), which is a part of the ER stress response mechanisms to clear unfolded proteins in the ER lumen and maintain cell viability. Inacio et al. have identified upregulation of 59 UPR-related ER proteins 12 hours post infection of human hepatoma cells, Huh7, by *P. berghei* sporozoites (Inacio et al., 2015). The UPR in the ER is activated during the sporozoite-to-merozoite transition stage in the PV *in vivo* studies by Inacio et al. have shown that the UPR activation and the success of EEF formation are positively correlated (Inacio et al., 2015). They hypothesized that since the sporozoites require a lot of type II fatty acids for the EEF development, they rely on the host's fatty acid synthesis mechanism (Inacio et al., 2015). In response to the accumulation of the newly formed fatty acids in the ER, protein folding is impaired and the UPR is activated (Inacio et al., 2015). This stress in the ER is exhibited 12 hours post infection and during this stage, the PV re-localizes to the host ER, and upregulates the UPR through currently unknown direct protein-protein interactions between the two membranes (Kaushansky & Kappe, 2015). The sporozoites' requirement of phosphatidylcholine lipids for their growth supports Kauchaunsky et al.,'s hypothesis. However, as glycans play an essential role in maintaining the cell structure and modulating cell-to-cell, cell-to-matrix, or cell-to-microbe signal transductions (Miura et al., 1996; Varki & Varki, 2001), knocking out RPN1 may have detrimental effects on cell viability. Therefore, these known intracellular roles of RPN1 that are critical for cell health and its predicted association in the liver invasion may explain RPN1's low β-score yet high p-value and FDR results (see Table 3-2).

Following Inacio et al.'s findings on the effect of UPR activation on EEF formation (Inacio et al., 2015), it seems likely that *EMC1* (which encodes ER membrane protein complex subunit 1) may be involved in lipid metabolism critical for the EEF development. It is a subunit of an ER membrane complex and its role in human cells is not well understood. However, by overexpressing the ER membrane protein complex in HEK293 cells, Guna et al. demonstrated that this complex plays a role in integrating tail-anchored proteins into the membrane (Guna, Volkmar, Christianson, & Hegde, 2018). Parasites use sugars to produce GPI-anchored proteins which play an essential role in both liver and erythrocyte invasions and intracellular developments (Stanway et al., 2019). As the majority of the merozoite surface proteins are GPI-anchored, and as the rapid replication process requires a lot of resources, it seems possible that the absence of ER membrane protein complex may interfere with EEF development. However, whether the sporozoites hijack this host protein building mechanism is unknown.

Cell cycle control protein 50A (encoded by *TMEM30A*) is located in the plasma membrane and the Golgi apparatus. Although its role has not been extensively studied in human cells, when it was expressed in *S. cerevisiae*, cell cycle control protein 50A was found to be involved in importing phosphatidylcholine from the extracellular matrix into the cell (R. Chen, Brady, & McIntyre, 2011). Phosphatidylcholine is an essential lipid required by the replicating sporozoites to support the growth of both PV membrane and merozoites. Sporozoites have their own fatty acid synthesis II system in the apicoplast; however, this alone cannot accommodate such high demand for the lipid (Itoe et al., 2014). Thus, *Plasmodium* sporozoites rely on the host cell's phosphatidylcholine, marking it as one of the key factors for mediating sporozoite survival inside the host liver cell post invasion (Itoe et al., 2014).

FCGR2B encodes low affinity immunoglobulin gamma Fc region receptor II-b, which is an inhibitory receptor that weakly binds to the Fc fragment of IgG. Although there have not been any studies on the role of this IgG Fc receptor II-b in the liver stage specifically, several studies have shown that a mutation or deficiency in *FCGR2B* is associated with protection against malaria infection in patients. Willcocks et al. genotyped children from endemic areas and found that those who were protected against severe malaria in East Africa exhibited a I232T mutation in *FCGR2B* (Willcocks et al., 2010). In addition to this, Gelabert et al. have found that in *FCGR2B*-deficient mice, antibody responses were enhanced, and malaria parasites were phagocytosed by macrophages at a higher rate than in the wild type (Gelabert, Olalde, de-Dios, Civit, & Lalueza-Fox, 2017). Unfortunately, this phenomenon is restricted to the blood stage in so far as no further studies have been done on the relationship between *FCGR2B*-deficiency and liver invasion.

However, as liver stage biology is disproportionately understudied compared to the blood stage, several biological events of the blood stage can be used as a reference for explaining the liver stage invasion mechanisms, owing to the similar biological makeup between sporozoites and merozoites. IgG Fc receptor II-b is expressed abundantly in the liver but the majority of its expression is found on liver sinusoidal endothelial cells (LSEC) and Kupffer cells (Ganesan et al., 2012). The role of IgG Fc receptor II-b in these cells is to take up small immune complexes and clear the blood (Ganesan et al., 2012). Although the cell-line used for this thesis project, HC-04, is a mimicry of primary hepatocytes that is more susceptible to malaria infection and supports full EEF development, it is unclear whether HC-04 is capable of expressing genes from other similar types of cells such as LSECs and Kupffer cells due to the lack of reliable HC-04 cell line transcriptomics and proteomics. Despite the fact that the role of IgG Fc receptor II-b was identified from a blood stage study, *FCGR2B* has been shown to be related to overall

susceptibility to malaria and its mutation and deficiency prevented further malaria development in both humans and mice. Prior to hepatocyte invasion, sporozoites traverse through Kupffer cells to cross the LSEC barrier; whether they penetrate the cell membrane or travel via vesicles in Kupffer cells remains unanswered. However, if the latter is true, it opens a possibility that sporozoites might hijack the immune complex clearance system by the Kupffer cells to enter them. However, it remains to be elucidated which sporozoite surface protein mimics that of the IgG Fc fragment. Considering the sporozoites' ability to detect HSPGs expressed on cell surfaces that enables recognition and binding to cells, it seems unlikely that an interaction with IgG Fc receptor II-b would be unspecific, as IgG Fc receptor II-b is not rich in HSPGs. Hence, if there is a sporozoite binding partner, it must be a IgG Fc receptor II-b -specific one.

Plasma membrane calcium-transporting ATPase 1 (encoded by *ATP2B1*) is an isoform of a plasma membrane Ca^{2+} ATPase (PMCA) protein. As the name suggests, it is located in the plasma membrane and couples ATP-hydrolysis with Ca^{2+} export from the cytosol to the extracellular space. *ATP2B1* and its 3 isoforms, *ATP2B2*, *ATP2B3* and *ATP2B4*, maintain cytosolic Ca2+ homeostasis in animal cells (Carafoli, 2002). Currently, there are no studies on investigating *ATP2B1* in liver invasion. However, *ATP2B4*, an isoform of *ATP2B1* expressed in erythrocytes, was characterized as a *P. falciparum* severe malaria resistance gene with a polymorphism at rs2334880 from a genome-wide association study by Timmann et al. (Timmann et al., 2012). In another study using ATP2B4$^{-/-}$ mice, Lessard et al. have demonstrated *ATP2B4* as a causal gene for malaria invasion of erythrocytes (Lessard et al., 2017). Although these were erythrocyte stage studies, given the similarities in the invasion mechanism between the liver and erythrocyte stages, it seems possible that a mutation in *ATP2B1* may provide some level of protection against sporozoite invasion.

UDP galactose translocator (UGT; encoded by *SLC35A2*) and α-1,3-mannosyl-glycoprotein 2-β-N-acetylglucosaminyltransferase (GNT-I; encoded by *MGAT1*) are located in the Golgi apparatus membrane. Both are involved in N-linked glycosylation but at different steps: GNT-I initiates complex N-linked carbohydrate formation by converting high-mannose to N-glycans while UGT transports galactose to N-linked glycans. Knocking out these *SLC35A2* and *MGAT1* results in the absence of N-linked glycoprotein on cell surface which negatively affects the binding partner of N-liked glycoprotein, galectin (Stewart et al., 2017). A genome-wide CRISPR screen study on N-linked glycosylation and galectin by Steward et al. has shown that galectins (Gal-1, Gal-3 and Gal-8) are only able to bind to N-glycans on the cells with their N-linked glycosylation pathways unaffected (Stewart et al., 2017). Stanway et al. have identified that N-acetylglucosamine plays a key role in the liver stage development from a genome-wide parasite metabolic study (Stanway et al., 2019). Although there are no literature studies on the role of human N-linked glycosylation in liver invasion, if the *SLC35A2* and *MGAT1* hits are biologically true, then these literature evidences indicate that 1) if the knock-out effects are due to the invaded sporozoites' early inability to develop further, then it is possible that they could not move on post invasion due to the lack of available N-acetylglucosamine from the host; and 2) if the effects are due to the failure in the host cell invasion, there must be galectin-like proteins on the sporozoite surface.

One unexpected hit that showed the MLE method as a more sensitive tool than RRA, was β-2-glycoprotein 1 (Apo-H; encoded by *APOH*). ApoH is located in the extracellular matrix and plays a role in PV formation during invasion that when it is deleted, subsequent malaria cycle onsets are delayed (Sá e Cunha et al., 2017). Its involvement in the invasion was not known at the time of the experiment design and hence it became an unanticipated second positive control in Pool 4.

Amongst the identified hits, *ITGB5* was chosen for further investigation due to the findings by Dundas et al. of *Pf*TRAP's strong interaction with αV-integrin and a weakly positive *Pf*TRAP-αVβ5 interaction from the AVEXIS screen, and *ITGB5*'s partner, *ITGAV*, being identified as a hit from a different pool. Because of the results by Dundas et al., *ITGAV* was also further studied alongside *ITGB5* to serve as a control in the follow-up experiments.

CHAPTER 4. INVESTIGATION OF INTEGRIN AVB5 AS A POSSIBLE RECEPTOR FOR *PLASMODIUM* SPOROZOITE INVASION OF HEPATOCYTES

4.1. INTRODUCTION

CRISPR-Cas9 screen results of Pool 1 (Results 3.2.3.1) and Pool 2 (Results 3.2.3.2) identified *ITGAV* and *ITGB5* as significantly negatively selected genes. Together, they express an integrin heterodimer, αVβ5, on the cell surface. These results were of particular interest as αVβ3 is a binding partner of a *P. falciparum* sporozoite surface protein, *Pf*TRAP. Both αVβ3 and αVβ5, which share the same αV component, mediate cell-to-cell and cell-to-matrix adhesions and bind to molecules with an RGD motif (Dundas et al., 2018; Nieberler et al., 2017). Ligands of αV-integrins are not restricted to RGD motif-bearing proteins, however. αV-integrins can also bind to ligands without an RGD motif such as type IV collagen (Pedchenko, Zent, & Hudson, 2004), Zika viruses (S. Wang et al., 2020) and echoviruses (Ylipaasto, Eskelinen, Salmela, Hovi, & Roivainen, 2010). Cell adhesion-mediating integrins, such as αVβ3, αVβ5, α5β1 and αIIbβ3, bind to RGD motif bearing molecules and proteins in the extracellular space (Pedchenko et al., 2004; Van Agthoven et al., 2014). Amongst these integrin pairs, only αVβ3 and αVβ5 are expressed in the liver, according to The Human Protein Atlas database (Uhlen et al., 2015).

*Pf*TRAP is an important sporozoite surface protein that mediates gliding motility and cell invasion. As αV-integrins bind to *Pf*TRAP, which expresses an RGD motif, αVβ5 was investigated to identify its possible role in a hepatocyte invasion by sporozoites. In designing experiments to investigate this, first, it was important to consider alternative explanations for αVβ5 contributing to the invasion. On the host side, as integrins mediate cell adhesions and

trigger important intracellular signalling pathways, it seemed possible that the apparent effects of *ITGAV* and *ITGB5* disruptions may have been an artefact of compromised cell health. Dissecting this from a pathogen-host ligand-receptor interaction appeared challenging.

I therefore set out to identify direct evidence of host-to-parasite protein-protein interactions involving αVβ5, either with *Pf*TRAP or *Pb*TRAP, or other possible binding partners. To further explore and validate the apparent hits from the screen, the effect of *ITGAV* and *ITGB5* single-gene disruptions upon sporozoite invasion assays were investigated. These experiments were conducted in parallel to the completion of the remainder of the CRIPSR-Cas9 screen on Pool 3, 4 and 5.

4.2 BIOINFORMATIC REVIEW OF POTENTIAL *PLASMODIUM* SPP. BINDING PARTNERS OF AVB5

Biochemical investigation of possible parasite-binding partners of αVβ5 using classical techniques such as co-immunoprecipitation was felt to be impractical, due to the possible low affinity of the interaction and the very small amounts of parasite protein which might be involved. Therefore, a bioinformatic strategy was used to identify other possible integrin-binding proteins in the *P. falciparum* and *P. berghei* proteomes.

*Pf*TRAP interacts with αV-integrins at two domain sites: RGD motif and von Willebrand factor A (VWA) (Dundas et al., 2018). Using these features as a reference, other sporozoite surface proteins that expressed either an RGD motif or a VWA domain were sought using the *Plasmodium* protein database, PlasmoDB.

A list of *P. falciparum* and *P. berghei* sporozoite surface proteins that express an RGD motif and/or a VWA domain was generated. First, a list of *P. falciparum* and *P. berghei* proteins with the RGD motif and a signal peptide was generated on PlasmoDB. Then, each one was analyzed on RMgmDB database to identify their orthologues, expressions in sporozoites, and existing literature. This resulted in the final list of six proteins: duffy binding-like merozoite surface protein 2 (DBLMSP2), erythrocyte binding antigen-140 (EBA140), TRAP-like protein (TREP), perforin-like protein 5 (PLP5), rhoptry neck protein 5 (RON5), and circumsporozoite and TRAP-related protein (CTRP) (Table 4-1).

The process of parasite invasion is conserved in many respects across *Plasmodium* species, involving orthologous proteins. The screens reported in the previous chapter were conducted

using the invasion of a rodent parasite into a human cell line (*P. berghei* rodent sporozoites were used for infecting human HC-04 cells). However, as understanding human malaria was the ultimate purpose, *P. falciparum* proteins were considered alongside the *P. berghei* proteins.

On the basis that an interaction involved in the invasion of rodent parasites into human cells was most likely to be conserved across parasite species, proteins with possible integrin-binding motifs in both *P. berghei* and *P. falciparum* orthologues were sought. Existing literatures and the RMgmDB mutant parasite database were used to exclude proteins which were already known to not play a role in hepatocyte invasion (Table 4-1).

Unfortunately, after careful analysis, it was concluded that no promising non-TRAP αV-integrin interacting candidates were identified. It was decided however that TREP, although not essential for liver invasion, was a sufficiently interesting candidate to merit some further evaluation.

Protein	Gene (top: *P. berghei*; bottom: *P. falciparum*)	Have RGD or VWA in both *P. berghei* (Pb) and *P. falciparum* (Pf)?	Expressed in sporozoites?	Studied in liver invasion?
TRAP	PBANKA_1349800 PF3D7_1335900	RGD and VWA in Pf, only VWA in Pb	Yes	Yes (Dundas et al., 2019; Matuschewski, Nunes, Nussenzweig, & Menard, 2002; Sultan et al., 1997).
DBLMSP2	N/A PF3D7_1036300	RGD only in Pf, no Pb orthologue	Yes	No.
EBA140	PBANKA_1332700 PF3D7_1301600	RGD only in Pf.	Not studied	No.
RON5	PBANKA_0713100 PF3D7_0817700	RGD only in Pb, no RGD in Pf orthologue	Yes	No.
TREP	PBANKA_1306500 PF3D7_1442600	RGD and VWA domains in both Pf and Pb	Yes	Yes. Disruption of TREP reduced hepatocyte infectivity by sporozoites in vivo but not in vitro (Moreira et al., 2008).
PLP5	PBANKA_0711600 PF3D7_0819200	RGD only in Pf, no RGD in Pb orthologue	Yes	Yes, not essential for liver invasion (Ecker, Pinto, Baker, Kafatos, & Sinden, 2007).

Table 4-1 Summary of sporozoite protein candidates of αvβ5 binding partner.

Amongst the candidates, TRAP and TREP were chosen for further investigation due to their expression of an RGD motif and a VWA domain.

4.3 INVESTIGATION OF POSSIBLE INTERACTION OF INTEGRIN AVB5 WITH PFTRAP AND PBTRAP

4.3.1 AVIDITY-BASED EXTRACELLULAR INTERACTION SCREEN (AVEXIS)

As discussed in Results 4.1, while *Pf*TRAP expresses an RGD motif that recognizes and binds to αV-integrins, *Pb*TRAP lacks this motif. Regardless, *Pb*TRAP was studied due to it being an orthologue of *Pf*TRAP and the TRAP protein's overall importance in gliding motility and invasion. Both TREP orthologues express an RGD motif and a VWA domain, which piqued an interest in TREP as a potential binding partner of αVβ5 despite previous studies that showed TREP was not essential for liver stage. *Pf*TREP was prioritised over *Pb*TREP for this study.

An AVEXIS assay was used to quantify bait-prey protein interactions following a methodology as described in Methods 2.5. *Pf*TRAP, *Pb*TRAP and *Pf*TREP were used as dimeric bait proteins and human αVβ1, αVβ3 and αVβ5 as pentameric prey proteins with a luciferase reporter (Figure 4-1). This experiment plan was similar to that of Dundas et al.; however, the designed recombinant proteins were different. Dundas et al. used monomeric bait proteins and pentameric prey proteins with a β-lactamase reporter (Dundas et al., 2018). Nonetheless, since Dundas et al. demonstrated a weak interaction between *Pf*TRAP and αVβ5, similar results were expected here as well (Dundas et al., 2018).

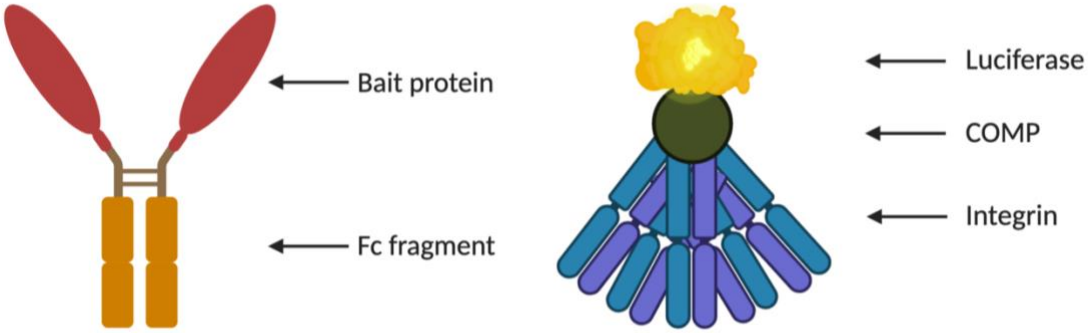


Figure 4-1 Illustrative depictions of a bait recombinant protein and an integrin recombinant prey protein design.
(Left) One bait protein was attached per one Fc fragment, hence bait proteins were expressed as a Fc fragment-mediated dimer. (Right) Integrin, a prey protein, was pentamerized by a cartilage oligomeric matrix protein (COMP) domain and fused to a luciferase tag, which was used in ELISA and AVEXIS for quantifying protein concentration (Figure **4-2**A) and protein-protein interactions (Figure **4-4**), respectively.

4.3.1.1 PROTEIN EXPRESSION

All bait and prey proteins were produced as described in Methods 2.5.1. Bait and prey protein productions were quantified by ELISA and luciferase assay, respectively (Figure 4-2). All measurements were made in duplicate. Positive controls used for quantifying protein concentrations were CD200 for the bait proteins (Figure 4-2A) and CD200R for the prey proteins (Figure 4-2B).

Target concentrations for the bait and prey proteins were 7 nM and 4 x 10^8 RLU/mL, respectively, which were shown by Segireddy et al. to be adequate for an AVEXIS assay (Segireddy et al., 2020). While CD200, PfRH5, *Pb*TRAP and *Pf*TRAP were successfully expressed above 7 nM, *Pf*TREP expression was at 3 nM (Figure 4-2A). Therefore, *Pf*TREP was excluded from further analysis. All prey proteins were expressed at > 4x10^8 RLU/mL

(Figure 4-2B). As a further quality control check, a Western blot analysis of *Pb*TRAP and *Pf*TRAP was performed. Protein expressions were detected with little degradation (Figure 4-3).

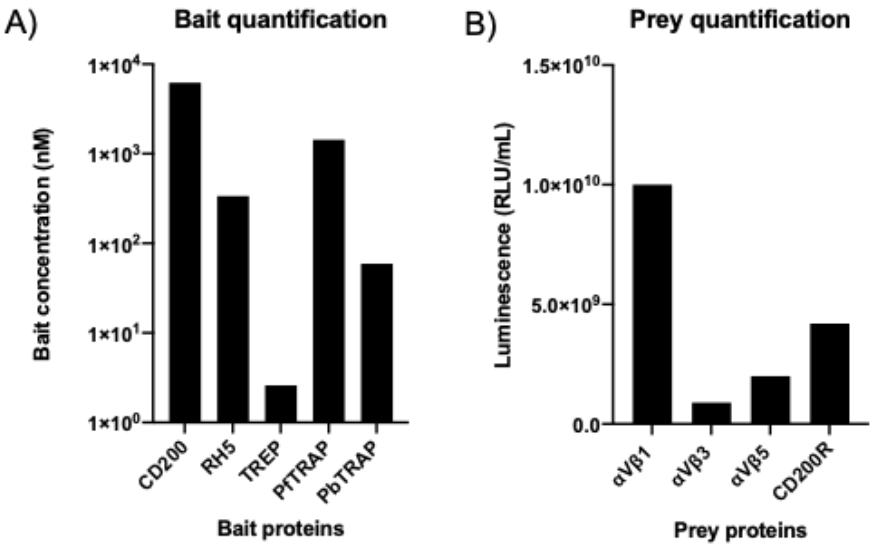


Figure 4-2 Production of recombinant bait proteins (*Pf*TRAP, *Pb*TRAP, *Pf*TREP, CD200R, *Pf*RH5) and prey proteins (αVβ1, αVβ3, αVβ5, CD200R).
(A) Average concentrations (nM) of bait proteins of two replicate measurements quantified by ELISA: 6203 nM (CD200), 337 nM (RH5), 3 nM (TREP), 1437 nM (*Pf*TRAP), 59 nM (*Pb*TRAP). (B) Prey proteins were quantified by measuring their Luciferase activities: 1 x 10^{10} RLU/mL (αVβ1), 9 x 10^{8} RLU/mL (αVβ3), 2 x 10^{9} RLU/mL (αVβ5) and 4 x 10^{9} RLU/ML (CD200R).

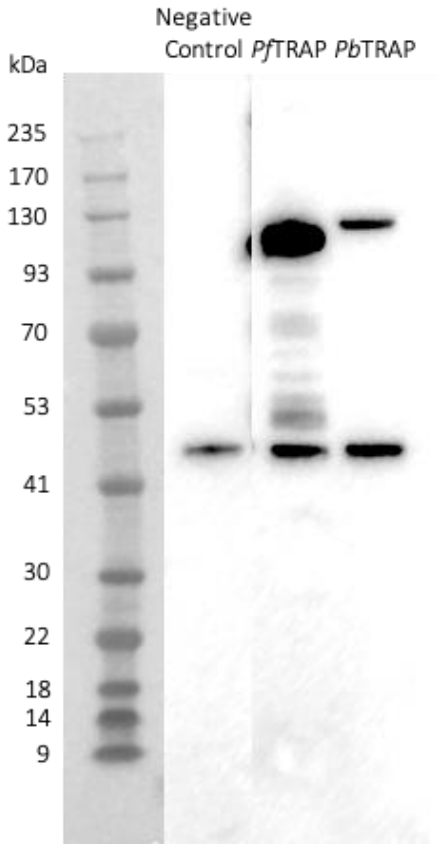


Figure 4-3 Western Blot shows *Pf*TRAP and *Pb*TRAP protein productions without signs of substantial protein degradation.
Bait protein was detected with biotinylated anti-C-tag conjugate (ThermoFisher, Cat. No. 7103252100) and a secondary antibody, Streptavidin-HRP (Pierce, Cat. No. 21130). From left to right: protein ladder, CD200R prey protein as a negative control, *Pf*TRAP and *Pb*TRAP. Band shown on negative control is an unspecific band known to be detected by anti-Ctag nanobody in any Expi293 cell supernatant. Blots have been cropped to remove irrelevant lanes for clarity.

4.3.1.2 Interaction screen between *Plasmodium* spp. TRAP and integrin heterodimers

Bait and prey proteins were diluted to 7 nM and 4 x 10^8 RLU/mL, respectively, in Blocker™ Casein (ThermoFisher) in PBS. *P. falciparum* RH5 and human BSG, which are already known to interact, were used as a positive control. Results were expressed in terms of signal-to-noise ratios, with signal being the luminescence in an individual well, and noise the luminescence of the same prey in wells coated with an irrelevant bait (CD200).

Interactions between *Pf*TRAP and αVβ5 and αVβ3 were observed, as previously reported by Dundas et al. In addition, a weaker interaction, not previously reported by Dundas et al., was observed between *Pf*TRAP and αVβ1. However, as this interaction was weaker than the RH5-BSG control, it is possible that there was an unspecific binding activity. We could not detect any interaction of *Pb*TRAP with the prey proteins (Figure 4-4).

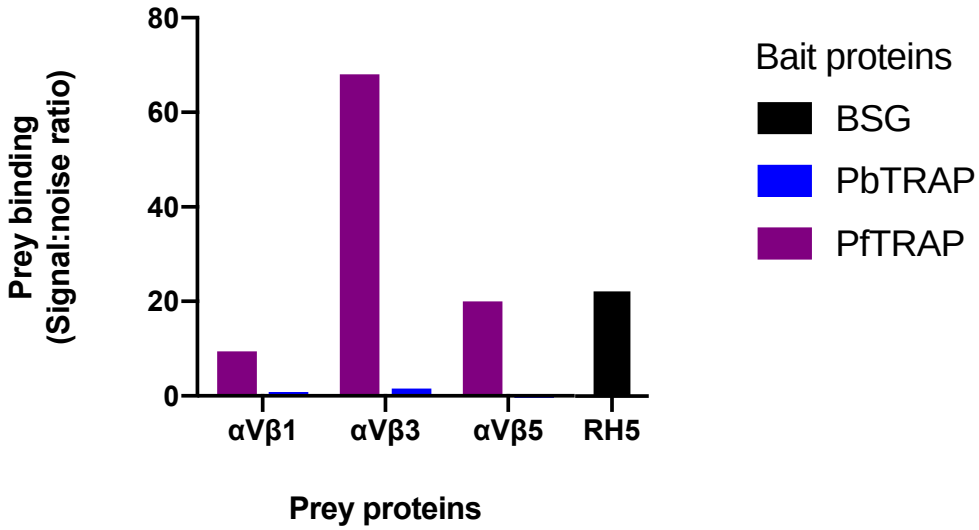


Figure 4-4 AVEXIS screen shows that *Pf*TRAP, but not *Pb*TRAP, interacts with αVβ3 and αVβ5.
Graph represents one experiment performed with triplicates per sample. 50 µL of 7 nM diluted bait proteins and 50 µL of 4 x 10^8 RLU/mL diluted prey proteins in casein were used. Luminescence signals were measured from 3 wells per protein interaction. Protein binding values were corrected to their signal-to-noise ratio. Noise signal of each prey protein was defined by luminescence signal detected when CD200 bait protein was added. The αVβ3 results were obtained by Dr Rameswara Segi Segireddy using proteins produced the author produced.

4.3.2 Cell surface staining by flow cytometry shows TRAP is able to bind to *ITGAV*-targeted and *ITGB5*-targeted HC-04 cells

Although an interaction between *Pb*TRAP and αV-integrins was not detected by AVEXIS, there was a possibility that binding to native αV-integrins on a cell surface may be detectable. Naturally expressed integrins on the cell surface may present different binding sites and conformation as compared to the recombinant, soluble proteins. Therefore, interactions of recombinant *Pf*TRAP and *Pb*TRAP bait proteins with cell-surface integrins were sought using flow cytometry. Dundas et al. had previously demonstrated an interaction of their recombinant *Pf*TRAP with αV-integrins on the surface of HepG2 cells (Dundas et al., 2018).

Both HC-04 and HepG2 cell lines were used for cell surface staining assays: these cells are genotypically indistinguishable but are phenotypically different with respect to parasite invasion. We therefore sought to investigate whether the results from Dundas et al. could be replicated in the HC-04 cells as well.

HC-04-Cas9 cells were first stained with *Pf*TRAP and *Pb*TRAP bait proteins and then stained again with anti-Fc secondary antibodies as described in Methods 2.6.3. In order to explore whether any binding detection from the wild-type HC-04-Cas9 cells was attributable to integrins, three treatments expected to modulate integrin binding were used: manganese ions (Mn^{2+}); *ITGAV*-targeting antibody (denoted as α-ITGAV); and the cyclic peptide 'cycloRGD' (see Figure 4-5). As divalent cations such as Mn^{2+} and Mg^{2+} are known to activate ligand-binding activity of αVβ3 (Stuiver, Ruggeri, & Smith, 1996), cells were incubated in 2 mM Mn^{2+} along with bait proteins to activate the integrins on the cell surface. Dundas et al. showed that the addition of Mn^{2+} improved TRAP-integrin binding activity, hence similar results were

expected to be observed. We anticipated that α-ITGAV antibodies may block TRAP binding if it was mediated by integrins. Lastly, cycloRGD peptides mimic the RGD motif on TRAP and again might compete with TRAP for integrin binding.

Protein binding activities on cell surfaces were measured using flow cytometry, with median fluorescence intensity (MFI) attributable to a secondary antibody (anti-human IgG, binding the bait Fc tag) used as a measure. Negative control cells were stained with PBS that did not contain any Fc-tagged bait proteins. Similar to the AVEXIS screen, CD200R was used as a secondary negative control as CD200R was not expected to bind to the HC-04 cell surface. The use of CD200R ensured that the stained cells were washed properly, and that the fluorescence observed on flow cytometry was not due to unspecific binding of the bait proteins (i.e., mediated by the Fc tag) or secondary antibodies.

Concentrations of *Pf*TRAP and *Pb*TRAP were adjusted to 7 µg/mL prior to cell staining. Results showed that *Pf*TRAP bound the HC-04-Cas9 surface and this interaction was improved with Mn^{2+} as shown by its higher MFI values (Figure 4-5). *Pb*TRAP, on the other hand, interacted very weakly with the cell surface and binding was unaffected by the addition of Mn^{2+}. Neither cycloRGD nor α-ITGAV efficiently blocked the TRAP-cell surface interaction.

To investigate whether any *Pf*TRAP-cell surface interaction could be detected without *ITGAV* or *ITGB5* expression, clonal *ITGAV*-targeted and *ITGB5*-targeted monoclonal HC-04-Cas9 cell lines were generated using CRISPR-Cas9 technology (Figure 4-6A) as described in Methods 2.4. Then, these cells, along with the HC-04-Cas9 wildtype control cells, were stained with Fc-tagged *Pf*TRAP and *Pb*TRAP bait proteins.

TRAP binding to *ITGAV*-targeted HepG2-Cas9 cells, a gift from Gavin Wright, was also tested. *ITGAV*-intact HepG2-Cas9 control cell line could not be revived after multiple attempts for use as a positive control. Therefore, albeit imperfect, *ITGAV*-intact HC-04-Cas9 cells were used as a comparator instead (Figure 4-6B).

It is important to note that growth of integrin-targeted cells on tissue culture plastic differed from that of wildtype cells, with reduced adhesion and cell size. Hence, all of the mutated cells, including *ITGAV*-targeted HepG2-Cas9, were cultured on collagen coated flasks as described in Methods 2.4.2 to compensate for the loss of cell-to-matrix and cell-to-cell adhesion necessary for cell growth and expansion. HC-04-Cas9 control cells were also grown on the collagen coated flasks for control sample purposes.

Overall, this work revealed that neither *Pf*TRAP nor *Pb*TRAP binding was affected by integrin disruptions in both HC-04-Cas9 (Figure 4-6B) and HepG2-Cas9 cells (Figure 4-6C). This was in contrast with the previously reported findings of Dundas et al. that *Pf*TRAP binding to HepG2-Cas9 cells was integrin-dependent.

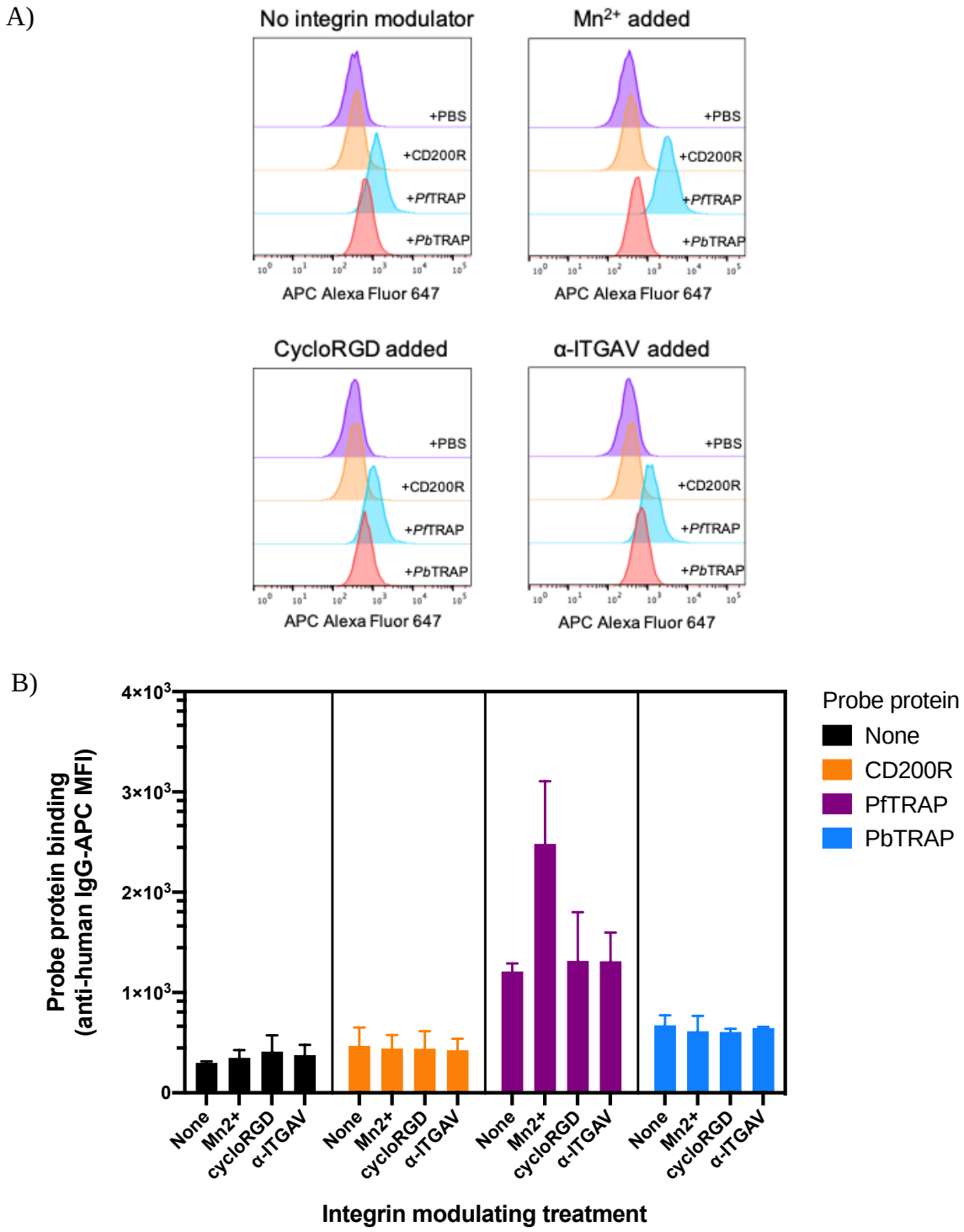
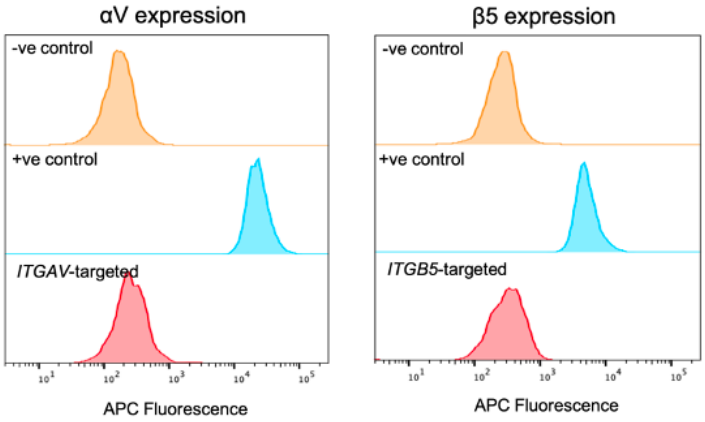


Figure 4-5 Blocking αV-integrins with cycloRGD and α-ITGAV antibody did not prevent *Pf*TRAP or *Pb*TRAP from binding to HC-04-Cas9 cells.

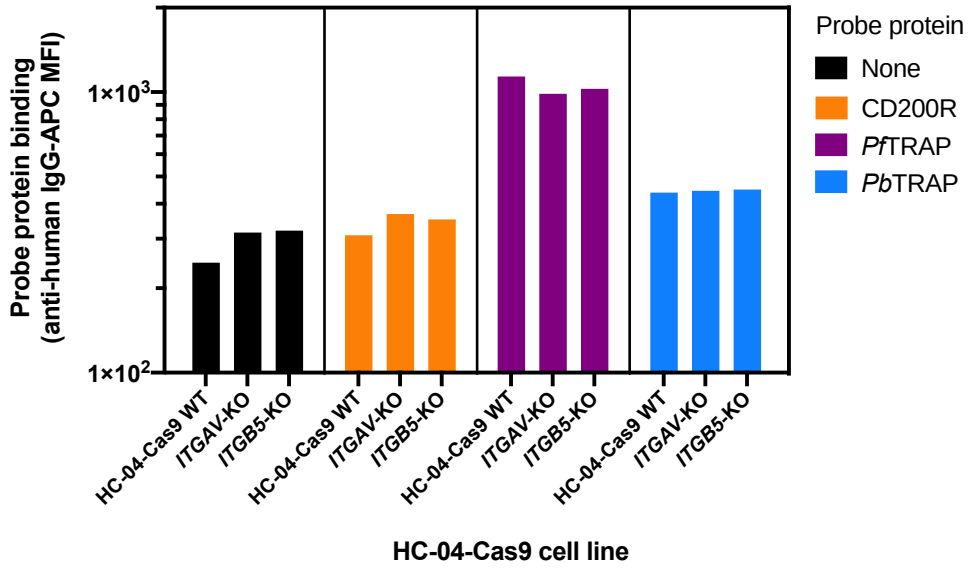
Fluorescence of bait probe proteins bound on HC-04 cell surface were measured by flow cytometry. (A) *Pf*TRAP was bound on the cell surface and its fluorescence intensified with Mn^{2+} incubation. *Pb*TRAP, on the other hand, resulted in only a marginal increase in fluorescence signal, which was unaffected by Mn^{2+}. Neither CycloRGD nor α-ITGAV antibody blocked the *Pf*TRAP-cell surface interaction. (B) Summary view of the data shown

in panel A, represented as MFI. Graph represents mean ± SEM of three independent experiments.

A)



B)



C)

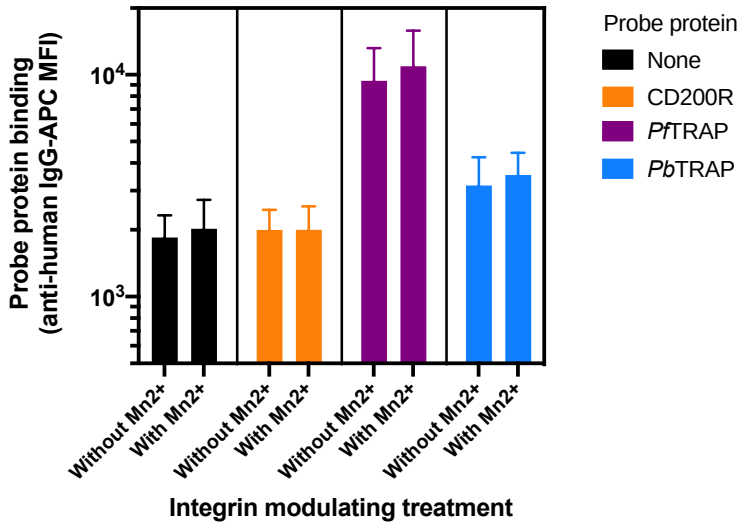


Figure 4-6 *Pf*TRAP can bind to both HC-04-Cas9 and HepG2-Cas9 cell surfaces irrespective of αV-integrins or β5-integrins.

(A) *ITGAV* and *ITGB5* disruptions in HC-04-Cas9 cell lines were analyzed by the losses of αV and β5 protein expressions, as detected by αv- and β5-specific monoclonal antibodies,

respectively, on a flow cytometer. Wildtype, integrin-intact HC-04-Cas9 cells were used as a negative control and were not stained with primary antibodies. As for a positive control, integrin-intact HC-04-Cas9 cells were stained with αV-specific and β5-specific primary antibodies for comparisons to *ITGAV*-targeted and *ITGB5*-targeted cells, respectively. (B) Graph represents one experiment. Bound *Pf*TRAP, *Pb*TRAP and a negative control recombinant protein, CD200R on the HC-04-Cas9 cell surface were analyzed by plotting the median fluorescence intensity of the bound recombinant proteins. (C) Graph represents mean ± SEM of three independent experiments. *ITGAV*-targeted HepG2-Cas9 cells' ability to bind to *Pf*TRAP, *Pb*TRAP and CD200R were measured by the median fluorescence intensity of the bound recombinant protein.

4.4 KNOCKING OUT *ITGAV* OR *ITGB5* DID NOT SIGNIFICANTLY REDUCE CELL INVASION BY *P. BERGHEI* SPOROZOITES

Although we had not been able to detect an interaction between *Pb*TRAP and αVβ5, nor identify another RGD-expressing sporozoite ligand for αVβ5, it remained possible that αVβ5 could be serving as an invasion receptor for an unknown parasite ligand. We therefore investigated the roles of *ITGAV* and *ITGB5* in sporozoite invasion of hepatocytes using CRISPR-Cas9 generated *ITGAV*-targeted, *ITGB5*-targeted and *SCARB1*-targeted HC-04-Cas9 cell lines.

Abrogation of *ITGAV* and *ITGB5* expressions were demonstrated by flow cytometry (Figure 4-6A). Loss of *SCARB1* expression on the cell surface unfortunately could not be validated by flow cytometry due to a lack of SR-B1 receptor staining antibody, and time.

Compared to the previous invasion experiments (see Results 3.2.3), this experiment was different in that it used collagen-coated flasks and plates for culturing integrin-targeted cells. To see whether culturing integrin-targeted cells on collagen would produce different results, two invasion experiments were performed in parallel using cells cultured on surfaces with and without collagen (Figure 4-7). As integrin-targeted cells cannot grow well without collagen for a long time, for the first 14 days, all cells, including HC-04-Cas9 and *SCARB1*-targeted HC-04-Cas9 cells, were cultured on collagen-coated flasks. Then on Day 15, cells were split in half and seeded on the collagen-coated and non-collagen coated 24-well plates. Surprisingly, both *ITGB5*-targeted and *ITGAV*-targeted HC-04-Cas9 cells grew in a monolayer fashion on non-collagen coated 24-well plates for the next 48 hours.

On Day 16, *SCARB1*-targeted, *ITGAV*-targeted, *ITGB5*-targeted and wildtype HC-04-Cas9 cells seeded on both non-collagen and collagen-coated 24-well plates were infected with sporozoites. This experiment investigated 1) whether disrupting *ITGAV* and *ITGB5* had a negative impact on sporozoite invasion; if yes, then 2) if these integrin-targeted cells exhibited the same reduction in sporozoite invasion when cultured with and without collagen.

Sporozoite-exposed cells cultured with and without collagen were analyzed in three populations based on their fluorescence intensities – mCherry^{hi}, mCherry^{int} and mCherry^{lo}. As expected, the proportion of *SCARB1*-targeted cells (defined as the percentage of single cells falling within the mCherry^{hi} population) was lower than the wildtype HC-04-Cas9 in both non-collagen and collagen experiments, consistent with the known role of SR-B1 as an invasion receptor in this context (Figure 4-8A).

To compare the mCherry^{hi} populations, population percentages were calculated in context of the infectivity of HC-04-Cas9 wildtype cells (Figure 4-8B). The mCherry^{hi} populations of HC-04-Cas9 cells were set as 100% and percentages of other cell lines were adjusted accordingly. *SCARB1*-targeted, *ITGAV*-targeted and *ITGB5*-targeted cells cultured without collagen showed a dramatic reduction in the infectivity rate as expected. Surprisingly, *ITGAV*-targeted cells cultured on collagen exhibited an infectivity rate higher than the wildtype cells, and *ITGB5*-targeted cells exhibited a slightly lower infectivity rate.

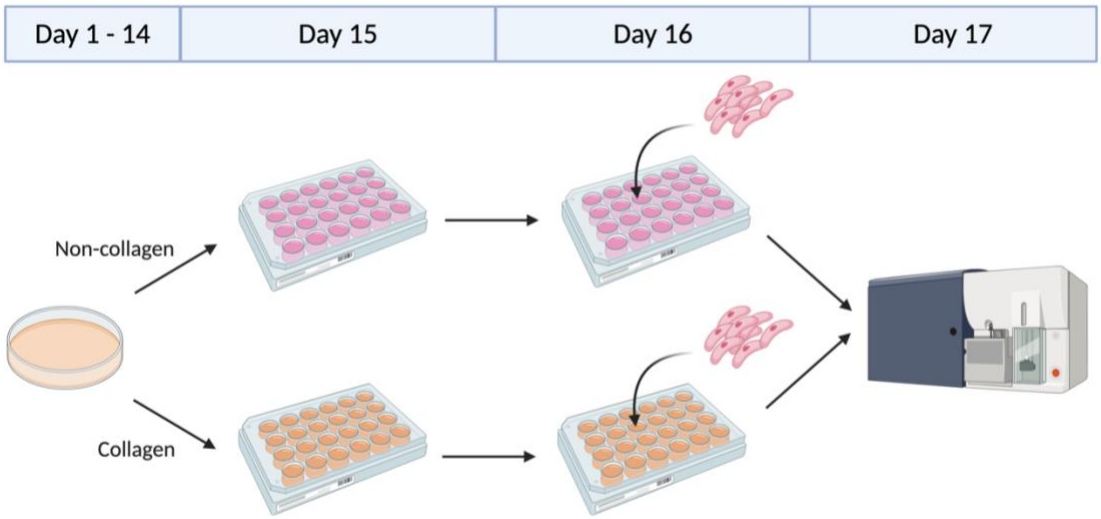


Figure 4-7 An overview of invasion experiments with cells seeded on 24-well plates coated without and with collagen.

Wildtype HC-04-Cas9, *SCARB1*-targeted, *ITGAV*-targeted and *ITGB5*-targeted cells were cultured on collagen-coated (orange) plates for 14 days. On Day 15, cells were split into two groups and cultured in different environments: non-collagen (pink) and collagen (orange). On Day 16, sporozoites were added to cells at 1×10^5 sporozoites per well. On Day 17, these sporozoite-exposed cells were analyzed with a flow cytometer.

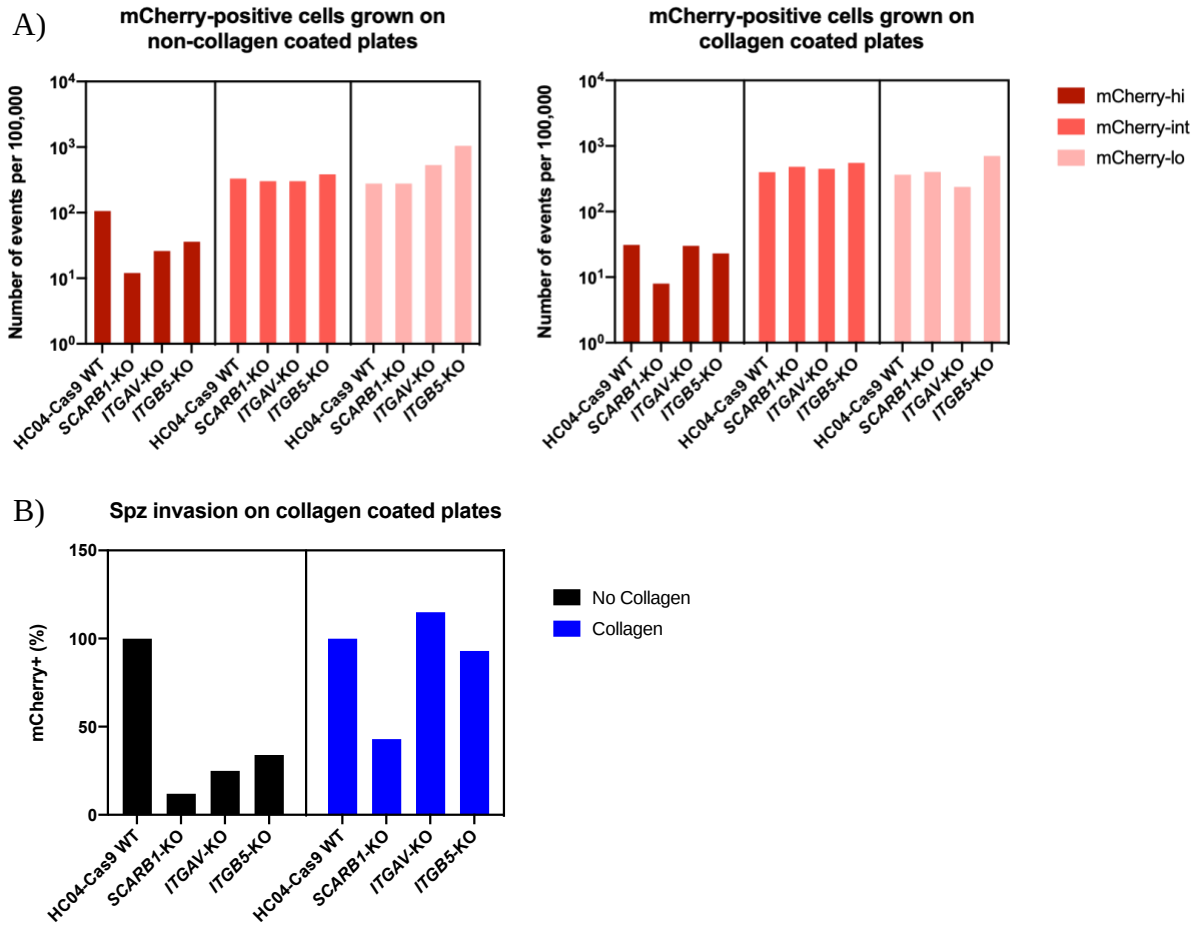


Figure 4-8 Use of collagen revealed inefficient reduction in sporozoite invasion in *ITGAV*-targeted and *ITGB5*-targeted cells.

Graph represents one experiment. (A) Left: flow cytometry analysis of sporozoite invaded cells that were not cultured on collagen-coated plates. Right: flow cytometry analysis of sporozoite invaded cells that were cultured on collagen-coated plates. (B) Comparison of mCherry^{hi} populations between samples, normalised to infection in HC-04-Cas9 WT cells (100%).

4.5 RE-ANALYSIS OF *ITGAV* AND *ITGB5* USING HC-04-CAS9 CELLS GROWN ON RAT TAIL I COLLAGEN

4.5.1 RE-VISITING THE POOL 2 SCREEN INVASION EXPERIMENT

During a flow analysis of surface protein expression losses as described in Results 4.3.2, I observed that the integrin-targeted cells were noticeably smaller compared to the HC-04-Cas9 wildtype cells. This raised suspicions that cells that expressed integrin-targeting sgRNAs may not have been included in the FSC vs SSC gate used prior to sorting of mCherry^{hi} cells in the CRISPR-Cas9 screens (as shown in Figure 3-11) and instead may have been mistaken as dead cells or debris. Moreover, it increased our suspicion that the effect of αV- and β5-integrin disruptions upon parasite invasion may be due to an effect upon cell health or mechanical adhesion strength, rather than αV- and β5-integrins serving as a receptor for a parasite protein.

As *ITGAV* and *ITGB5*-targeted cells cultured on collagen retained a larger size similar to that of wildtype HC-04-Cas9 cells, it was hypothesized that alternative integrins may be mediating adhesion to collagen, allowing normal cell health. A Pool 2 CRISPR-Cas9 screen invasion experiment was therefore repeated using collagen-coated culture vessels: persistence of the effect of αV and β5 integrin disruptions in this context would support a role as a receptor, whereas loss of the effect might suggest the previous observation had been due to a non-specific effect of loss of adherence to the culture substrate. The CRISPR-Cas9 screen invasion experiment and MAGeCK programme analysis followed the same methods used for the previous screens as described in Methods 2.2 and 0, respectively.

Exposed and mCherry^{hi} cell samples from two replicate experiments and one unexposed sample were sequenced by Illumina Sequencing (Figure 4-9A). The unexposed sample used in this experiment was different from Results 3.1.2 in that here, the unexposed cells were cultured on collagen-coated surfaces for 17 days prior to sequencing.

Quality control measures taken for evaluating the three samples are listed in Results 3.1.2. There were no losses of sgRNAs or signs of contamination in the samples. Sequencing quality was high, as indicated by the median quality score of each base and per sequence quality score, which were both higher than 35 (Figure 4-9A). The Spearman's rank correlation coefficient, r_s, of the sgRNA count comparison between unexposed and the pool 2 lentivirus pool was 0.88 for both the overall pool and the negative control non-targeting sgRNAs (Figure 4-10).

As shown in Figure 4-11, MAGeCK-MLE Batch Matrix analysis of the two replicate experiments identified *PIGS, PVRL2, SCARB1, ITGB2, SLC35A2, MGAT1, CD63, EPHB4, ITGB1, CD163* and *INS* as hits. A full list of genes from the most negatively selected to positively selected with their β-scores can be found in Appendix B: MAGeCK-MLE Batch Matrix gene rank *SCARB1* was identified as significantly negatively selected with a rank of 3, which indicated that the experiment was a success. *SLC35A2* and *MGAT1* were also identified as hits from the previous Pool 2 experiment that did not use collagen-coated cell culture vessels (Results 3.2.3.2). Surprisingly, *ITGB5* was ranked 40 with a β-score of -0.04 which indicated that *ITGB5* was neither negatively nor positively selected. Instead, different β-integrins, *ITGB2* and *ITGB1*, appeared as hits. Both *MGAT1* and *SLC35A2*, which were identified as hits in both non-collagen (Figure 3-15) and collagen (Figure 4-11B) Pool 2 experiments.

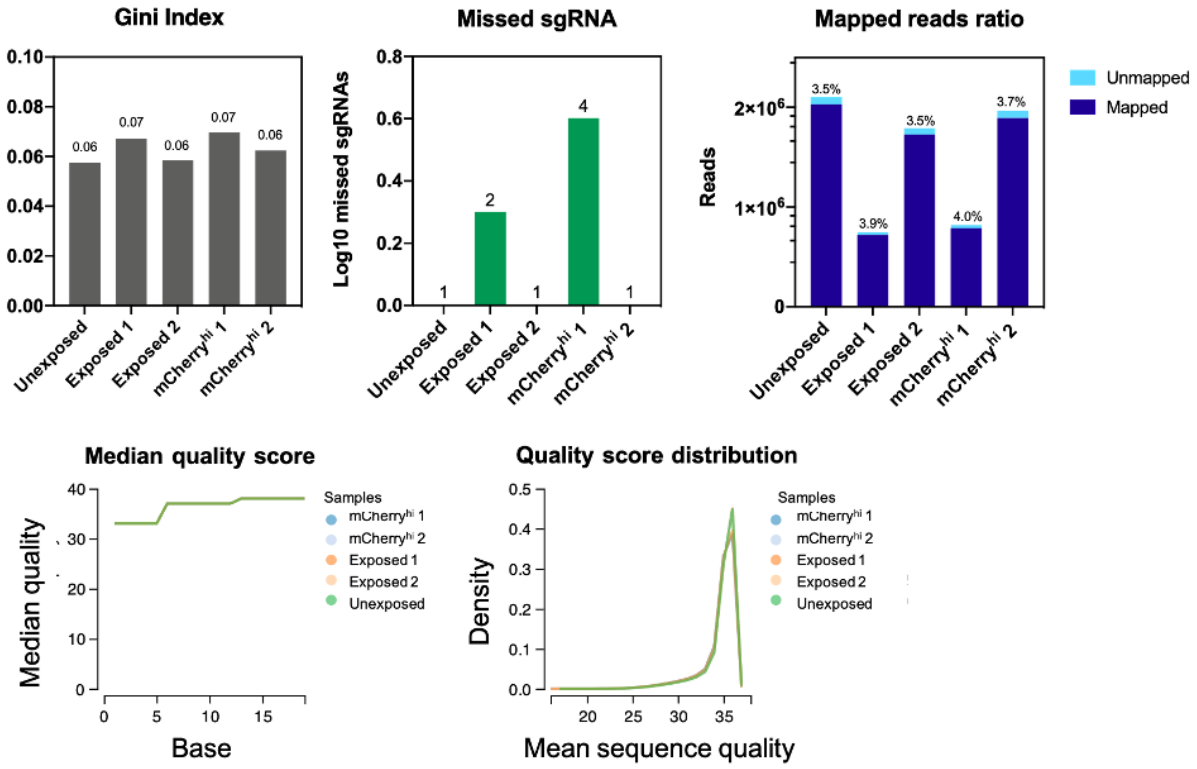


Figure 4-9 Quality control analysis of Pool 2 collagen experiment.
Gini index score of the unexposed sample was 0.06 and the rest were all 0.07. mCherry^{hi} 1 exhibited the highest number of missing sgRNAs, 4. The unmapped reads ratios were all extremely low and therefore there were no signs of sample contamination. Both base median quality score and mean sequence quality scores showed a peak at a value higher than 35, indicating there were no signal decay during the sequencing process.

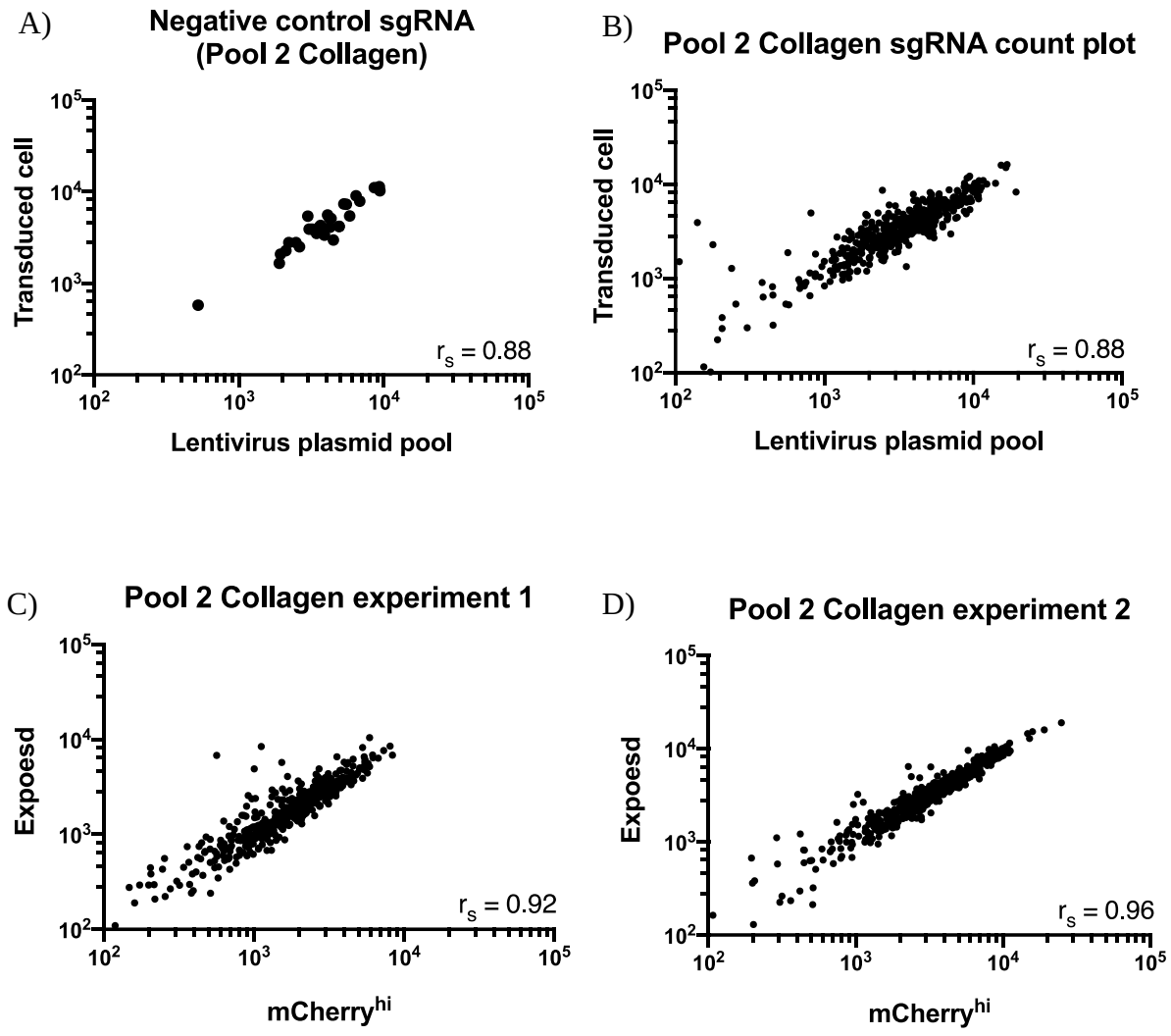
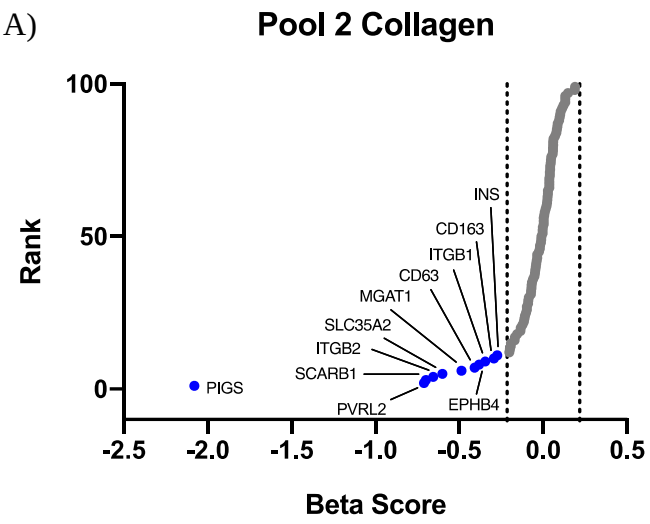


Figure 4-10 sgRNA count correlation plot of Pool 2 collagen experiment.
Spearman's rank correlation coefficient, r_s, of comparison between Pool 2 plasmid from Results 3.1.2 and unexposed sample was (A) 0.88 for total sgRNA count and (B) 0.88 for negative control sgRNAs only. (C) r_s value of total sgRNA counts was 0.92 for experiment 1 and (D) 0.96 for experiment 2.

A)


Pool 2 Collagen
Rank
Beta Score
100
50
0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
INS
CD163
ITGB1
CD63
MGAT1
SLC35A2
ITGB2
SCARB1
PIGS
PVRL2
EPHB4

B)

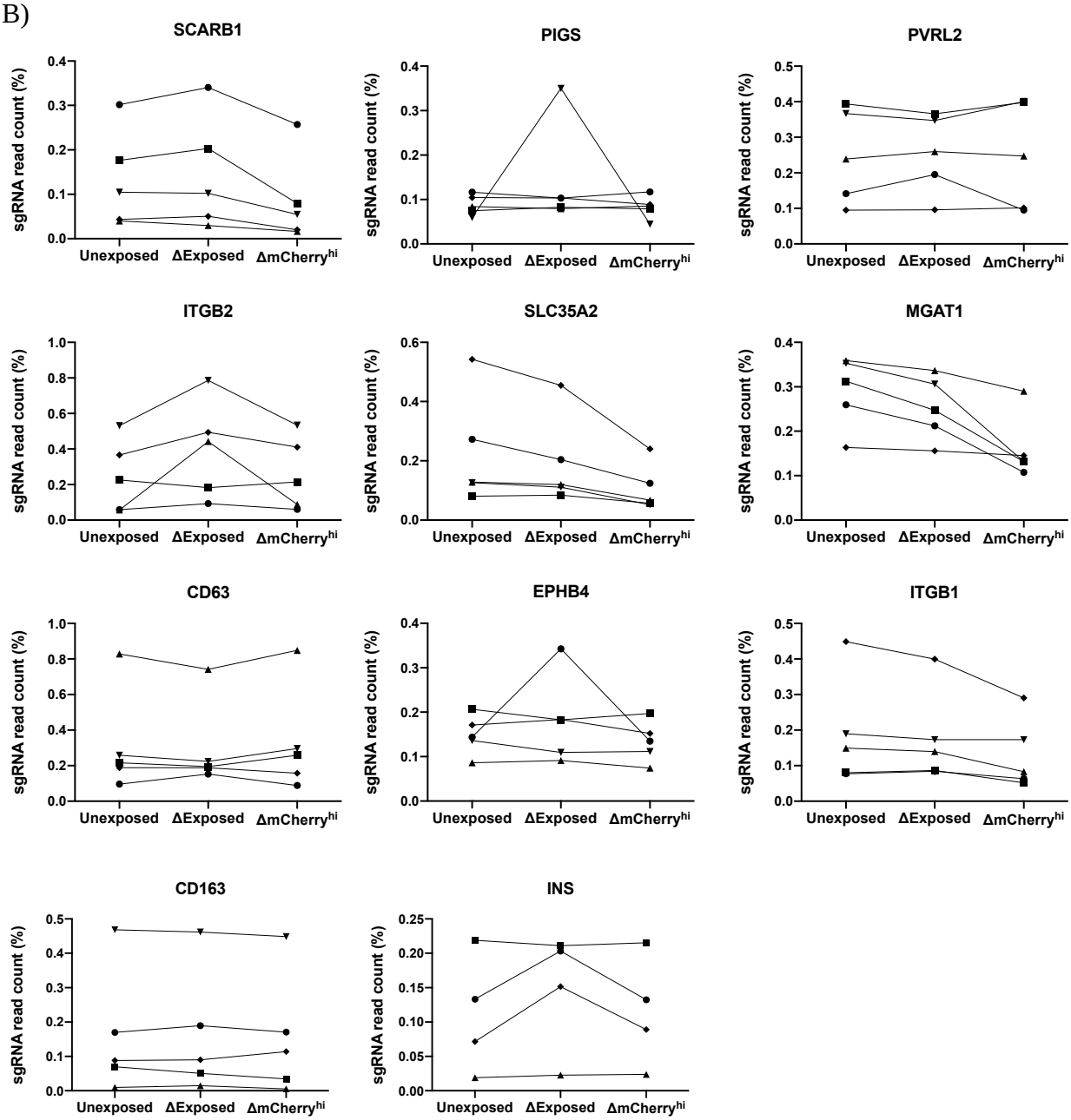

SCARB1
PIGS
PVRL2
ITGB2
SLC35A2
MGAT1
CD63
EPHB4
ITGB1
CD163
INS
sgRNA read count (%)
Unexposed
ΔExposed
ΔmCherry^hi

Figure 4-11 *PIGS, PVRL2, ITGB2, SLC35A2, MGAT1, CD63, EPHB4, ITGB1, CD163* and *INS* identified as new hits.
(A) MAGeCK-MLE Batch Matrix analysis presented by a β-score (x-axis) versus rank (y-axis) plot with standard deviation cut-offs (dotted lines) at -0.25 and 0.25. The following genes were identified as negatively selected: *PIGS* (rank: 1; β-score: -2.08), *PVRL2* (rank: 2; β-score: -0.71), *ITGB2* (rank: 4, β-score: -0.66), *SLC35A2* (rank: 5; β-score: -0.60), *MGAT1* (rank: 6; β-score: -0.49), *CD63* (rank: 7; β-score: -0.41), *EPHB4* (rank: 8; β-score: -0.38), *ITGB1* (rank: 9; β-score: -0.35), *CD163* (rank: 10; β-score: -0.29) and *INS* (rank: 11; β-score: -0.27) were identified as significant hits. (B) Read counts of the sgRNAs were calculated as percentages of the total reads of each respective sample.

4.5.2 Analysis of the effect of knocking out *ITGAV* and cell size represented by FSC and SSC on flow cytometry

Lack of cell adhesion caused by *ITGAV* and *ITGB5* knockouts which consequentially led to a smaller cell size generated a hypothesis that the transduced cells expressing the sgRNAs targeting these integrin genes may have been misidentified as dead and not included in the FSC-A vs SSC-A gate as shown on Figure 3-11 (Left). If this is true, there would be less cells carrying such sgRNAs in the FSC-A vs SSC-A gate, and the MAGeCK programme's identification of *ITGAV* and *ITGB5* as negatively selected genes would be due to the absence of cells from the FSC-A vs SSC-A gate rather than the mCherry^{hi} gate. To investigate this, 5×10^4 'FSC-low' cells sorted from the 'FSC-low' gate was compared to 3×10^5 unsorted, 'FSC-mix' cells which contained a mix of cells from 'FSC-low' gate and 'FSC-high' gate (Figure 4-13).

Illumina Sequencing data qualities of the 'FSC-low' and 'FSC-mix' samples were high with low Gini indexes, missed sgRNAs and unmapped reads ratio. Their data qualities were also clean as their median quality score was higher than 30 for all bases, and their mean sequence quality scores peaked at 35 (Figure 4-13A). Spearman's rank correlation coefficient was

calculated at 0.80, which was the lowest amongst all Illumina Sequencing experiments conducted in this project, suggesting the introduction of significant random noise during the processing of the samples. Such noise could result from, for example, inefficient cell recovery, such that the gRNA sample size is reduced. (Figure 4-13B).

MAGeCK-MLE analysis identified *ITGAV* as the most positively selected gene (Figure 4-14A), which meant that majority of the cells that expressed *ITGAV*-targeting sgRNAs were not included in the FSC-A vs SSC-A gate population in Figure 3-11. At the time of the experiment, it was difficult to determine whether the cells in the 'FSC-low' sample were dead or just physiologically smaller than the rest of the cells. This can be explored in the future by using a Live/Dead stain. However, when these results were analyzed further by their sgRNA read counts (Figure 4-14B), it appeared that the MAGeCK programme identified *ITGAV* as a positively selected gene due to two sgRNAs marked as ▲ and ◆ that were over-represented in the 'FSC-low' sample.

These results were from one replicate experiment only as a second experiment could not be performed due to the COVID-19 pandemic.

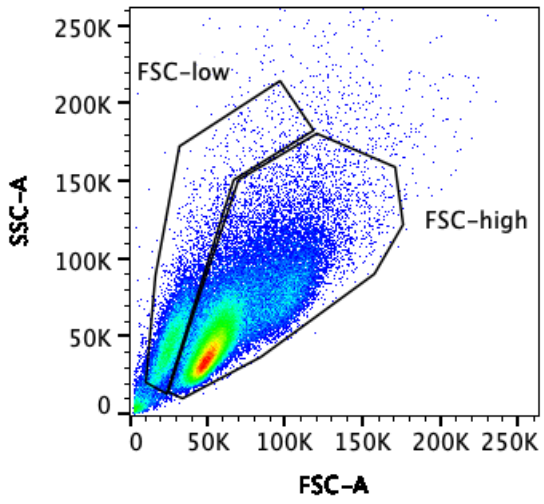


Figure 4-12 Flow cytometry gate strategy for FSC-high versus FSC-low assessment.

FSC-A vs SSC-A plot of 'FSC-low' and 'FSC-high' Pool 1 transduced cells. The 'FSC-high' gate was the same as the previous CRISPR-Cas9 screen experiments (Figure **3-11** Left).

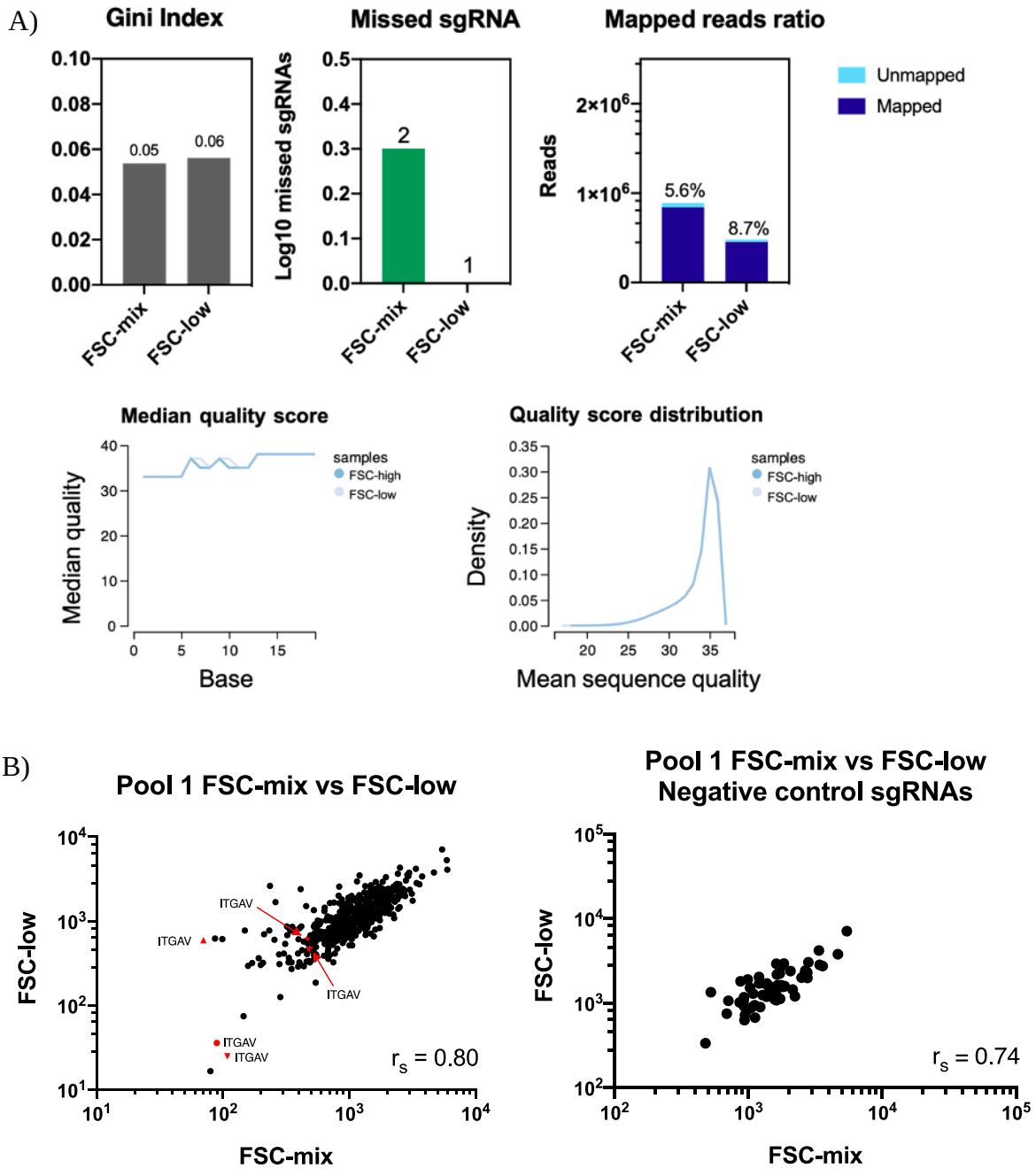


Figure 4-13 Quality control analysis FSC-mix versus FSC-low analysis.
(A) Gini index scores were 0.05 for 'FSC-mix' and 0.06 for 'FSC-low'. There were two missed sgRNAs in 'FSC-mix' but only one in 'FSC-low'. The unmapped ratios were 5.6% and 8.7% for 'FSC-mix' and 'FSC-low' respectively. (B) The correlation plot showed that its Spearman's rank correlation coefficient was 0.80. The low r_s value was due to the noise generated by significantly dropped out sgRNAs. Red symbols are *ITGAV*-targeting sgRNAs and the unique identity of each sgRNA is marked by its shape.

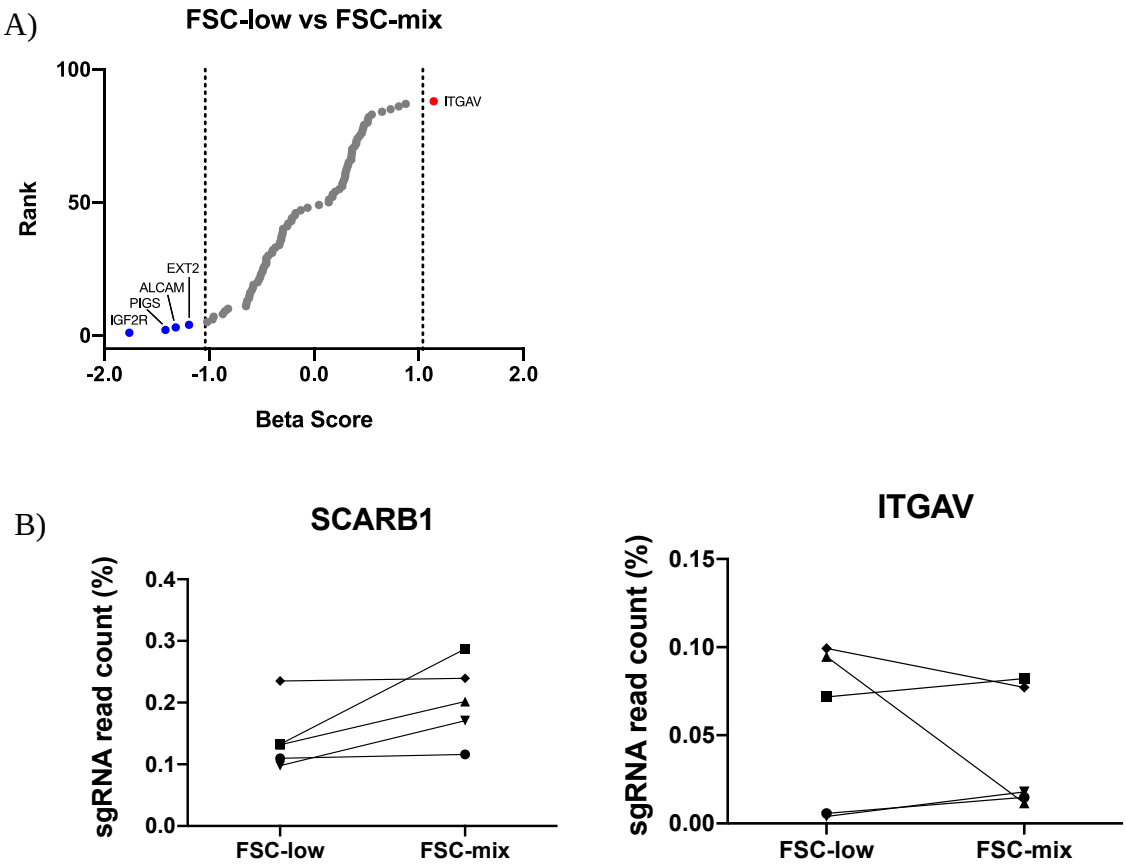


Figure 4-14 *ITGAV* shows as positively selected in FSC-low population.
(A) ITGAV was identified as the most positively selected gene in the 'FSC-low' sample based on the MLE method calculation. The standard deviation cut-offs were -1.04 and 1.04. (B) sgRNA read counts were calculated as percentages of the total reads. SCARB1-targeting sgRNA count data was included as a control.

4.6 DISCUSSION

ITGAV and *ITGB5*, which are expressed as an integrin heterodimer, αVβ5, were investigated as potential genes essential for hepatocyte invasion by sporozoites. They were chosen amongst the list of CRISPR-Cas9 screen hits due to the following reasons:

1) Both *ITGAV* and *ITGB5* resulted as significant hits despite them being from different pools. Pool 1 experiment, which identified *ITGAV*, and Pool 2 experiment, which identified *ITGB5*, were performed on different days.
2) αVβ3, another αV-integrin, directly binds to *Pf*TRAP (Dundas et al., 2018).
3) αVβ5 was shown to weakly bind to *Pf*TRAP from the findings by Dundas et al.

Despite the results from Dundas et al. that showed retainment of *Pf*TRAP binding on integrin-disrupted HepG2 cells, it is unclear whether these mutated cells were cultured on collagen-coated vessels. As integrin-disrupted cells grow poorly in culture without a reagent that aids in retaining cell-to-cell and cell-to-matrix adhesions, it seems plausible that their cell surface staining results were a false negative. From their invasion experiment, Dundas et al. showed that TRAP-integrin interaction plays no major role in hepatocyte invasion despite the AVEXIS screen results. Unlike the cell surface staining assay, *ITGAV*-targeted HepG2 cells cultured on collagen were used for the HepG2 invasion assay. However, as they assessed invaded sporozoites 2 hrs after the sporozoites were added to the cells without FITC-dextran, which aids in marking sporozoite-traversed cells, their data did not differentiate traversed cells from productively invaded cells. Our invasion experiment, on the other hand, although it did not use a FITC-dextran marker either, because of the 22hr incubation post infection, we were confident that the mCherry^{hi} populations were productively invaded cells.

In their AVEXIS screen, αVβ3, αVβ5 and αVβ6 showed weak binding while αVβ8 and αVβ1 showed none. Using their data on αVβ5 as a reference, follow-up experiments were designed to investigate the role of αVβ5 in hepatocyte invasion further. At the initial stages of the follow-up experiment design, there were suspicions around *ITGAV* and *ITGB5* being false positives, rather than true receptors for parasite proteins. Nevertheless, follow-up experiments on these genes were conducted due to the promising results by Dundas et al. Before designing the experiments, sporozoite surface protein candidates that bind to integrins were sought for using PlasmoDB, UNIPROT and RMgmDB. *Pf*TRAP and *Pb*TRAP were chosen as potential binding partners of αVβ5.

An AVEXIS assay was performed with *Pf*TRAP and *Pb*TRAP to quantify their interactions to αVβ1, αVβ3 and αVβ5. While *Pf*TRAP was bound to all three integrins as expected, *Pb*TRAP surprisingly did not. Unlike *Pf*TRAP, *Pb*TRAP does not express an RGD motif, which may explain its inability to bind to the integrins.

Next, protein-to-cell surface interactions were analyzed with HC-04-Cas9 cell surface staining assays by flow cytometry. The results demonstrated a strong interaction between *Pf*TRAP and the cell surface as expected, and this interaction was improved with Mn^{2+}. CycloRGD and α-ITGAV antibody, on the other hand, did not seem to effectively block the interaction, again consistent with Dundas' results.

To further explore the possibility that αVβ5 may be a receptor for *Pb*TRAP, flow cytometry was used to assess the binding of *Pb*TRAP to the surface of cells using *Pf*TRAP as a positive control. Very weak interaction of *Pb*TRAP with the cells was detected and this did not appear to be affected by treatments that were expected to modulate the integrin-ligand binding. Results

also showed that *Pf*TRAP interaction with the cell surface was not αv-integrin dependent, thus contradicting the previously reported results by Dundas et al. A cell surface staining assay was repeated with *ITGAV*-targeted HepG2 cells received from Gavin Wright in an attempt to replicate the results by Dundas et al. Wildtype HC-04-Cas9 cells were used in lieu of wildtype HepG2-Cas9 cells as they could not be revived after multiple attempts. As the genotype of HC-04 wildtype very closely resemble HepG2 (Tao et al., 2014), HC-04-Cas9 cells were used as a control. Surprisingly, *Pf*TRAP was able to interact with the *ITGAV*-targeted HepG2 cells strongly.

These results opened up two possibilities: 1) TRAP may bind to other RGD motif-recognizing integrins, or 2) TRAP may bind to other cell surface proteins, perhaps as specific interactions mediated by its second adhesive domain, TSR, or perhaps non-specifically. Integrins other than αVβ3 that play a role in cell-to-cell and cell-to-matrix adhesion by recognizing the RGD motif (Van Agthoven et al., 2014) are α5β1 and αIIbβ3 (Nieberler et al., 2017). Proteomic analysis of HepG2 cell line (Wisniewski, Vildhede, Noren, & Artursson, 2016) suggested that neither *ITGA5* nor *ITGA2B* are expressed in HepG2 cells, arguing against the first hypothesis that TRAP may bind to other integrins similar to αVβ3. Currently, there is no reliable data on HC-04 cell transcriptomic and proteomic.

Matuschewski et al. demonstrated a reduction in infection with *P. berghei* sporozoites expressing *Pf*TRAP with mutated TSR and VWA domains (Matuschewski et al., 2002). They also showed that mutating the VWA domain alone was not enough to effectively eliminate the infection, which hence suggested that the two domains must bind to host surface proteins for cell invasion. Although host protein receptor of TSR is unknown, Cermai et al. have discovered sulfated glycoconjugates such as heparan sulphated glycoproteins (HSPGs) as the binding

partner of TSR (Cerami et al., 1992). Then, it seems likely that *Pf*TRAP was bound on the integrin-targeted HC-04-Cas9 cell surfaces due to HSPGs. Similarly, this may also explain the contradictory result observed with the *ITGAV*-targeted HepG2 cells. One way to test this hypothesis would be to knock out *NDST* gene which encodes for an N-deacetylase/N-sulfotransferase enzyme that removes an acetyl group from N-acetylglucosamine for a sulphate group to be substituted instead (Aikawa & Esko, 1999). Consequently, integrin-targeted cells with *NDST* targeted will express a reduced level of HPSGs on the cell surface. Then, cell surface staining experiments can be conducted with these cells to detect specific and direct interactions between *Pf*TRAP and integrins without interferences caused by HPSGs.

In parallel with the above work and our ongoing screening of pools 3-5, a sporozoite invasion experiment was repeated using individually *ITGAV*-targeted, *ITGB5*-targeted and *SCARB1*-targeted HC-04-Cas9 cells, seeking to validate the hits from the pooled screen. *SCARB1*-targeted, *ITGAV*-targeted, *ITGB5*-targeted and wildtype HC-04-Cas9 cells were grown in two environments: with collagen and without collagen. Surprisingly, the detrimental effect of *ITGAV* and *ITGB5* disruptions seen from the previous CRISPR-Cas9 screen experiments was seen only in the absence of collagen. First, I considered whether collagen affected the sporozoites' ability to infect due to the different infectivity rate in the collagen-cultured wildtype cells. However, collagen is a commonly used tool for liver invasion studies from novel gene identification (i.e. discoveries of AMA1 and MAEBL proteins for hepatocyte invasion (A. S. P. Yang, Lopaticki, et al., 2017) and SPECT and PLP1 proteins for hepatocyte traversal (A. S. P. Yang, O'Neill, et al., 2017)) to quantification of sporozoite invasion and EEF development (Sinnis et al., 2013). Hence, as collagen is not known to negatively affect sporozoites from infecting cells, the results show that disrupting *ITGAV* did not prevent the invasion at all, therefore contradicting the previous Pool 1 CRISPR-Cas9 screen results.

Despite the differences in experimental designs, this finding complements the results of Dundas et al. which showed that disrupting *ITGAV* played a role in cell traversal during dermal migration but did not identify a role in liver invasion (Dundas et al., 2018). However, unlike Dundas et al., our cell surface staining results are consistent with the invasion assay which altogether demonstrated that although TRAP may bind to αV and β integrins, but they are not the sole binding partners of TRAP and thus making it possible for sporozoites to invade the integrin-targeted cells.

During a flow analysis of protein expression losses on *ITGAV*-targeted and *ITGB5*-targeted cell surfaces, suspicions arose whether the *ITGAV* and *ITGB5* results were false positives as the *ITGAV*-targeted and *ITGB5*-targeted exhibited a smaller size on the FSC vs SSC plot compared to the HC-04-Cas9 wildtype control. Moreover, as both *ITGAV* and *ITGB5* play a role in cell adhesion, a question was raised whether the sporozoites could not invade cells expressing sgRNAs targeting these integrins due to lack of cell adhesion rather than the lack of the integrin receptors.

First, collagen coated surfaces were used for cell culture in a repeated CRISPR-Cas9 screen experiment. Two replicate experiments with Pool 2 were performed. The planned testing of Pool 1 could not occur due to the COVID-19 pandemic. The following genes were identified as negatively selected: *PIGS*, *PVRL2*, *ITGB2*, *SLC35A2*, *MGAT1*, *CD63*, *EPHB4*, *ITGB1*, *CD163* and *INS*. Surprisingly, *ITGB5* fell to rank 40 from 1 and its β-score was -0.04. With collagen, the β5-integrin was revealed to not be required for sporozoite invasion; however, the *ITGB2* and *ITGB1* hits show that there are likely to be other β-integrins that may be necessary. These results are consistent with the possibility that the effects of *ITGB5* disruption from the

previous Pool 2 CRISPR-Cas9 experiments were due to lack of cell adhesion, rather than a specific protein-protein interaction.

Next, two Pool 1 transduced cell populations, 'FSC-low' and 'FSC-mix', were analyzed to determine whether the *ITGAV* hit could be an artefact of the gating strategy. Note that as the sorted cells were not stained with a Live/Dead dye, 'FSC-low' population was a mix of both small but alive and dead cells. If the *ITGAV*-targeted cells died without collagen, it seemed probable that all cells expressing *ITGAV*-targeting sgRNAs may have been in the 'FSC-low' gate, mistaken as dead cells. *ITGAV* was indeed significantly over-represented in the 'FSC-low' population, although this result was not completely conclusive as it was driven by only two guides.

To summarise, the experiments reported in this chapter gathered a number of strands of evidence against the possibility of integrin αVβ5 acting as a specific host receptor for a parasite protein such as *Pb*TRAP:

1) Inability to detect interaction of *Pb*TRAP and αV-integrins in the plate-based AVEXIS assay
2) Inability to detect integrin-mediated binding of *Pb*TRAP to HepG2-Cas9 or HC-04-Cas9 cells
3) Loss of β5-integrin, and possibly gain of other β-integrins, as 'hits' in the CRISPR-Cas9 screen when collagen was available as an adhesion substrate
4) A similar loss, in the presence of collagen, of the invasion phenotype due to αV- and β5- disruption in individually gene-targeted cell lines

5) Observation of altered growth and size characteristics of αV- and β5- disrupted cells, and possible enrichment of αV-targeting guides in a population of FSC-low cells which would have been gated out during the pooled screen

CHAPTER 5. CONCLUSION

5.1. SCREEN RESULTS AND FOLLOW-UP EXPERIMENTS

CRISPR-Cas9 screens have identified the following human hepatocyte genes as negatively selected in hepatocyte invasion assays by sporozoites: *ITGAV, RPN1, ATP2B1, TMEM30A, FCGR2B, ITGB5, SLC35A2, MGAT1, EMC1* and *APOH*. *RPN1, SLC35A2, MGAT1* and *EMC1* mediate post-translational modification of proteins and modification of cell membrane composition and so are likely to have pleotropic effects on parasite invasion rather than to mediate specific host-parasite protein-protein interactions (Guna et al., 2018; Inacio et al., 2015; Kaushansky & Kappe, 2015; Miura et al., 1996; Stewart et al., 2017). *TMEM30A* and *ATP2B1*, (which encode cell cycle control protein 50A and PMCA1, respectively) are located in plasma membrane and are presumed to play a role in phosphatidylcholine import and calcium ions export, respectively (Carafoli, 2002; R. Chen et al., 2011). Apo-H is located in extracellular space and it became an unanticipated positive control for Pool 4 due to its essential role in PV formation (Sa et al., 2017). *FCGR2B* encodes a cell surface receptor inhibitory of Fc-IgG and its mutation was identified in East Africans protected against severe malaria (Willcocks et al., 2010). Lastly, *ITGAV* and *ITGB5* are integrin genes expressed on the cell membrane as a heterodimer, αVβ5. Due to the results of Dundas et al. which demonstrated promising effects of αV-integrin-TRAP interactions in sporozoite locomotion, αVβ5 was investigated further.

Follow-up experiments seeking protein-protein interactions between αVβ5 and both *Pf*TRAP and *Pb*TRAP showed that αVβ5 did not interact detectably with *Pb*TRAP. Disrupting *ITGAV* and *ITGB5* did not reduce sporozoite invasion either. A repeated Pool 2 CRISPR-Cas9 screen

experiment using cells cultured on collagen-coated plates identified other β-integrins, such as *ITGB2* and *ITGB1*, as hits instead.

Therefore, the follow-up experiments revealed several limitations of using the CRISPR-Cas9 screen technology for studying sporozoite invasion of hepatocytes:

1. Because the sporozoite infectivity was low and variable for each experiment, the mCherry^{hi} sample size was not equal for the replicate experiments per pool. A large sample size of mCherry^{hi} cells may help in reducing false positives but is challenging to achieve because sporozoite yield and infectivity varies.
2. It was difficult to disentangle effects of gene disruption upon cell health and upon sporozoite invasion. *ITGAV* and *ITGB5*, for example, were later found to be false positives and this problem could be avoided by sorting the mCherry-negative population of the parasite-exposed cells to be used as a control rather than taking the unsorted ones. The effect of knocking out the genes essential for cell growth would be reflected in both samples. That way, the sgRNA counts of these genes would be low in both samples and thus they would not be interpreted as under-represented in the mCherry^{hi} by the MAGeCK programme.
3. Results may be quite sensitive to the exact FACS gating strategy used. As a formula was used for determining the mCherry^{hi} populations rather than gating around a population of cells that appear distinct from the cells of a lower fluorescence intensity, it is possible that some of the results included genes that may have been essential for cell traversal but not necessarily for the cell invasion event *per se*.
4. Lastly, the underlying assumption that the mCherry^{hi} cells are the truly invaded cells with healthy sporozoites inside was not tested. One way to mitigate this is to use a cell-

tracing marker, such as fluorescein isothiocyanate–dextran, along with the mCherry marker to identify and distinguish between traversed and invaded cells on FACS.

5.2. IMPACT OF COVID-19 PANDEMIC

The COVID-19 pandemic interrupted further investigations of *ITGAV* and *ITGB5*. The table below illustrates a timeline of experiments that was planned before a lock-down took place.

April 2020	2 replicate pool 1 invasion experiments with cells cultured on collagen
May 2020	1) Replicate experiment 2 of investigating if *ITGAV* was a false positive due to a wrong FSC gate 2) Investigating if *ITGB5* was also a false positive due to a wrong FSC gate
June-July 2020	Follow up experiments on newly identified hits from the new Pool 1 and 2 experiments conducted in April 2020 using the same methods as before: AVEXIS, cell surface staining and invasion experiment

5.3. IMPLICATIONS OF THIS STUDY FOR UNDERSTANDING OF HEPATOCYTE INVASION BY SPOROZOITES

As evidenced by its detection of *SCARB1* and *APOH*, the CRISPR-Cas9 screen technology has potential to identify hepatocyte receptors involved in invasion. However, no clear novel surface receptor was identified from the screens performed here. Hypotheses for the lack of robust novel hits include redundancy of receptors, other types of sporozoite-hepatocyte interactions were involved, false negative failure of the screen to detect a true essential receptor, essential genes were not on the list and use of *P. berghei* sporozoites and HC-04 cell line were not adequate models for understanding *P. falciparum* invasion into primary hepatocytes. Nevertheless, the detection of *SCARB1* as the top hit in all five pools demonstrated that CRISPR-Cas9 screen is a useful tool for studying a plethora of genes in search for the ones essential for invasion.

Redundancy of receptors is well-described for the *P. falciparum* erythrocytic stage (Stubbs et al., 2005), and given the mechanistic similarities between the liver and erythrocytic stages, it is possible that there may be redundancy of receptors in the liver stage as well. If multiple genes perform the same protein function in a cell (i.e., acting as a receptor), disrupting one gene will not have a significant effect on the biological phenotype in the cell. Hence, these genes will appear as neither positively nor negatively selected in the screen. Or it may be that a special case of redundancy might be if parasite invasion is not dependent upon specific protein-protein interactions, but instead due to interactions with host glycans, which may be displayed on multiple proteins.

It remains possible that there are false negatives in the data and there are two possibilities of the screen's failure in this matter: inefficient sgRNAs and masking of true hits due to background noise. Although the sgRNA sequences were bioinformatically derived and no individual experiment was performed on every single sgRNA to confirm their knockout efficiencies, it seems quite unlikely that all sgRNAs targeting a particular gene would be inefficient as positive control genes were detected successfully in all experiments. Background noise, which can be measured by the r_s value from the sgRNA read count XY scatter plots, may mask true hits; however, if the background noise from the screens was a considerable issue, *SCARB1* would not have been detected with clear statistical data and *APOH*, a previously unknown positive control, would not have been uncovered.

When the list of 470 hepatocyte receptor genes was designed, it was assumed that sporozoites were more likely to interact with more abundant and accessible host receptors. However, protein abundance, as measured by mass spectrometry, may not be a reliable indicator of potential as an invasion receptor: the parasite may use a receptor which is of low abundance, or poorly detected by mass spectrometry. Hence, essential genes may not have been included in the list. Repetition of the screen using a more extensive list of candidate genes may be worthwhile.

Lastly, the use of a rodent line, *P. berghei*, with the human HC-04 cell line may not have efficiently uncovered true hits that can be related to the *P. falciparum* infection. Although *P. berghei* and *P. falciparum* share several aspects of conserved invasion mechanisms, as observed by their TRAP orthologues, there may be other key surface proteins that lack key domains in one species that makes it challenging to apply observations from the *P. berghei* infection mechanism to that of *P. falciparum*. This problem may be solved with a use of human

parasites, such as *P. falciparum* and *P. vivax*, which have an infection rate of 0.07% and 0.04%, respectively, in HC-04 cells (Sattabongkot et al., 2006). However, doing so poses problems in laboratory safety during sporozoite extraction and lesser mCherry^{hi} sporozoite-invaded cells will be collected due to a lower infection rate than *P. berghei* which may reduce statistical power.

Despite these uncertainties, this has probably been the most comprehensive screen for sporozoite invasion receptors which has been performed to date. The intriguing possibility remains that SR-B1 may be the only essential receptor involved in invasion in this system (*P. berghei* invasion into HC-04 cells). The screen described here could be repeated and extended with a bigger list of genes, more sporozoites and more biological replicates. A CRISPR activation screen might be used to overexpress gene candidates and study which protein functions improved invasion, potentially overcoming genetic redundancy. More ambitiously, and subject to the development of appropriate methods, the screen could be repeated with *P. falciparum* in primary human hepatocytes.

BIBLIOGRAPHY

Achan, J., Talisuna, A. O., Erhart, A., Yeka, A., Tibenderana, J. K., Baliraine, F. N., . . . D'Alessandro, U. (2011). Quinine, an old anti-malarial drug in a modern world: role in the treatment of malaria. *Malar J, 10*, 144. doi:10.1186/1475-2875-10-144

Aikawa, J., & Esko, J. D. (1999). Molecular cloning and expression of a third member of the heparan sulfate/heparin GlcNAc N-deacetylase/ N-sulfotransferase family. *J Biol Chem, 274*(5), 2690-2695. doi:10.1074/jbc.274.5.2690

Alexander, D. L., Arastu-Kapur, S., Dubremetz, J. F., & Boothroyd, J. C. (2006). Plasmodium falciparum AMA1 binds a rhoptry neck protein homologous to TgRON4, a component of the moving junction in Toxoplasma gondii. *Eukaryot Cell, 5*(7), 1169-1173. doi:10.1128/EC.00040-06

Alonso, P. L., Lindsay, S. W., Armstrong, J. R., Conteh, M., Hill, A. G., David, P. H., . . . et al. (1991). The effect of insecticide-treated bed nets on mortality of Gambian children. *Lancet, 337*(8756), 1499-1502. doi:10.1016/0140-6736(91)93194-e

Amino, R., Thiberge, S., Martin, B., Celli, S., Shorte, S., Frischknecht, F., & Menard, R. (2006). Quantitative imaging of Plasmodium transmission from mosquito to mammal. *Nat Med, 12*(2), 220-224. doi:10.1038/nm1350

Arredondo, S. A., Swearingen, K. E., Martinson, T., Steel, R., Dankwa, D. A., Harupa, A., . . . Kappe, S. H. I. (2018). The Micronemal Plasmodium Proteins P36 and P52 Act in Concert to Establish the Replication-Permissive Compartment Within Infected Hepatocytes. *Front Cell Infect Microbiol, 8*, 413. doi:10.3389/fcimb.2018.00413

Baer, K., Roosevelt, M., Clarkson, A. B., Jr., van Rooijen, N., Schnieder, T., & Frevert, U. (2007). Kupffer cells are obligatory for Plasmodium yoelii sporozoite infection of the liver. *Cell Microbiol, 9*(2), 397-412. doi:10.1111/j.1462-5822.2006.00798.x

Ballou, W. R. (2009). The development of the RTS,S malaria vaccine candidate: challenges and lessons. *Parasite Immunol, 31*(9), 492-500. doi:10.1111/j.1365-3024.2009.01143.x

Baum, J., Chen, L., Healer, J., Lopaticki, S., Boyle, M., Triglia, T., . . . Cowman, A. F. (2009). Reticulocyte-binding protein homologue 5 - an essential adhesin involved in invasion of human erythrocytes by Plasmodium falciparum. *Int J Parasitol, 39*(3), 371-380. doi:10.1016/j.ijpara.2008.10.006

Baum, J., Gilberger, T. W., Frischknecht, F., & Meissner, M. (2008). Host-cell invasion by malaria parasites: insights from Plasmodium and Toxoplasma. *Trends Parasitol, 24*(12), 557-563. doi:10.1016/j.pt.2008.08.006

Beeson, J. G., Drew, D. R., Boyle, M. J., Feng, G., Fowkes, F. J., & Richards, J. S. (2016). Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria. *FEMS Microbiol Rev, 40*(3), 343-372. doi:10.1093/femsre/fuw001

Bhanot, P., Schauer, K., Coppens, I., & Nussenzweig, V. (2005). A surface phospholipase is involved in the migration of plasmodium sporozoites through cells. *J Biol Chem, 280*(8), 6752-6760. doi:10.1074/jbc.M411465200

Bronner, I. F., Quail, M. A., Turner, D. J., & Swerdlow, H. (2014). Improved Protocols for Illumina Sequencing. *Curr Protoc Hum Genet, 80*, 18 12 11-42. doi:10.1002/0471142905.hg1802s80

Carafoli, E. (2002). Calcium signaling: a tale for all seasons. *Proc Natl Acad Sci U S A, 99*(3), 1115-1122. doi:10.1073/pnas.032427999

Carrolo, M., Giordano, S., Cabrita-Santos, L., Corso, S., Vigario, A. M., Silva, S., . . . Mota, M. M. (2003). Hepatocyte growth factor and its receptor are required for malaria infection. *Nat Med, 9*(11), 1363-1369. doi:10.1038/nm947

Centers for Disease Control and Prevention [CDC]. (2019). Indoor Residual Spraying. *Malaria.* Retrieved from https://www.cdc.gov/malaria/malaria_worldwide/reduction/irs.html

Cerami, C., Frevert, U., Sinnis, P., Takacs, B., Clavijo, P., Santos, M. J., & Nussenzweig, V. (1992). The basolateral domain of the hepatocyte plasma membrane bears receptors for the circumsporozoite protein of Plasmodium falciparum sporozoites. *Cell, 70*(6), 1021-1033. doi:10.1016/0092-8674(92)90251-7

Chen, C. H., Xiao, T., Xu, H., Jiang, P., Meyer, C. A., Li, W., . . . Liu, X. S. (2018). Improved design and analysis of CRISPR knockout screens. *Bioinformatics, 34*(23), 4095-4101. doi:10.1093/bioinformatics/bty450

Chen, R., Brady, E., & McIntyre, T. M. (2011). Human TMEM30a promotes uptake of antitumor and bioactive choline phospholipids into mammalian cells. *J Immunol, 186*(5), 3215-3225. doi:10.4049/jimmunol.1002710

Chinappi, M., Via, A., Marcatili, P., & Tramontano, A. (2010). On the mechanism of chloroquine resistance in Plasmodium falciparum. *PLoS One, 5*(11), e14064. doi:10.1371/journal.pone.0014064

Chugh, M., Sundararaman, V., Kumar, S., Reddy, V. S., Siddiqui, W. A., Stuart, K. D., & Malhotra, P. (2013). Protein complex directs hemoglobin-to-hemozoin formation in Plasmodium falciparum. *Proc Natl Acad Sci U S A, 110*(14), 5392-5397. doi:10.1073/pnas.1218412110

Clyde, D. F. (1975). Immunization of man against falciparum and vivax malaria by use of attenuated sporozoites. *Am J Trop Med Hyg, 24*(3), 397-401. doi:10.4269/ajtmh.1975.24.397

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., . . . Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science, 339*(6121), 819-823. doi:10.1126/science.1231143

Coppi, A., Tewari, R., Bishop, J. R., Bennett, B. L., Lawrence, R., Esko, J. D., . . . Sinnis, P. (2007). Heparan sulfate proteoglycans provide a signal to Plasmodium sporozoites to stop migrating and productively invade host cells. *Cell Host Microbe, 2*(5), 316-327. doi:10.1016/j.chom.2007.10.002

Counihan, N. A., Kalanon, M., Coppel, R. L., & de Koning-Ward, T. F. (2013). Plasmodium rhoptry proteins: why order is important. *Trends Parasitol, 29*(5), 228-236. doi:10.1016/j.pt.2013.03.003

Cowman, A. F., Healer, J., Marapana, D., & Marsh, K. (2016). Malaria: Biology and Disease. *Cell, 167*(3), 610-624. doi:10.1016/j.cell.2016.07.055

Cox, F. E. (2010). History of the discovery of the malaria parasites and their vectors. *Parasit Vectors, 3*(1), 5. doi:10.1186/1756-3305-3-5

De Niz, M., Meibalan, E., Mejia, P., Ma, S., Brancucci, N. M. B., Agop-Nersesian, C., . . . Marti, M. (2018). Plasmodium gametocytes display homing and vascular transmigration in the host bone marrow. *Sci Adv, 4*(5), eaat3775. doi:10.1126/sciadv.aat3775

de Ridder, S., van der Kooy, F., & Verpoorte, R. (2008). Artemisia annua as a self-reliant treatment for malaria in developing countries. *J Ethnopharmacol, 120*(3), 302-314. doi:10.1016/j.jep.2008.09.017

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., . . . Root, D. E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol, 32*(12), 1262-1267. doi:10.1038/nbt.3026

Doll, K. L., & Harty, J. T. (2014). Correlates of protective immunity following whole sporozoite vaccination against malaria. *Immunol Res, 59*(1-3), 166-176. doi:10.1007/s12026-014-8525-0

Dundas, K., Shears, M. J., Sinnis, P., & Wright, G. J. (2019). Important Extracellular Interactions between Plasmodium Sporozoites and Host Cells Required for Infection. *Trends Parasitol, 35*(2), 129-139. doi:10.1016/j.pt.2018.11.008

Dundas, K., Shears, M. J., Sun, Y., Hopp, C. S., Crosnier, C., Metcalf, T., . . . Wright, G. J. (2018). Alpha-v-containing integrins are host receptors for the Plasmodium falciparum sporozoite surface protein, TRAP. *Proc Natl Acad Sci U S A, 115*(17), 4477-4482. doi:10.1073/pnas.1719660115

Durocher, Y., Perret, S., & Kamen, A. (2002). High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Res, 30*(2), E9. doi:10.1093/nar/30.2.e9

Ecker, A., Pinto, S. B., Baker, K. W., Kafatos, F. C., & Sinden, R. E. (2007). Plasmodium berghei: plasmodium perforin-like protein 5 is required for mosquito midgut invasion in Anopheles stephensi. *Exp Parasitol, 116*(4), 504-508. doi:10.1016/j.exppara.2007.01.015

Ejigiri, I., Ragheb, D. R., Pino, P., Coppi, A., Bennett, B. L., Soldati-Favre, D., & Sinnis, P. (2012). Shedding of TRAP by a rhomboid protease from the malaria sporozoite surface is essential for gliding motility and sporozoite infectivity. *PLoS Pathog, 8*(7), e1002725. doi:10.1371/journal.ppat.1002725

El Hajj, H., Lebrun, M., Arold, S. T., Vial, H., Labesse, G., & Dubremetz, J. F. (2007). ROP18 is a rhoptry kinase controlling the intracellular proliferation of Toxoplasma gondii. *PLoS Pathog, 3*(2), e14. doi:10.1371/journal.ppat.0030014

Figliozzi, R. W., Chen, F., Chi, A., & Hsia, S. C. (2016). Using the inverse Poisson distribution to calculate multiplicity of infection and viral replication by a high-throughput fluorescent imaging system. *Virol Sin, 31*(2), 180-183. doi:10.1007/s12250-015-3662-8

Franke-Fayard, B., Trueman, H., Ramesar, J., Mendoza, J., van der Keur, M., van der Linden, R., . . . Janse, C. J. (2004). A Plasmodium berghei reference line that constitutively expresses GFP at a high level throughout the complete life cycle. *Mol Biochem Parasitol, 137*(1), 23-33. doi:10.1016/j.molbiopara.2004.04.007

Fullman, N., Burstein, R., Lim, S. S., Medlin, C., & Gakidou, E. (2013). Nets, spray or both? The effectiveness of insecticide-treated nets and indoor residual spraying in reducing malaria morbidity and child mortality in sub-Saharan Africa. *Malar J, 12*, 62. doi:10.1186/1475-2875-12-62

Ganesan, L. P., Kim, J., Wu, Y., Mohanty, S., Phillips, G. S., Birmingham, D. J., . . . Anderson, C. L. (2012). FcgammaRIIb on liver sinusoidal endothelium clears small immune complexes. *J Immunol, 189*(10), 4981-4988. doi:10.4049/jimmunol.1202017

Gelabert, P., Olalde, I., de-Dios, T., Civit, S., & Lalueza-Fox, C. (2017). Malaria was a weak selective force in ancient Europeans. *Sci Rep, 7*(1), 1377. doi:10.1038/s41598-017-01534-5

Greenbaum, D. C., Baruch, A., Grainger, M., Bozdech, Z., Medzihradszky, K. F., Engel, J., . . . Bogyo, M. (2002). A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. *Science, 298*(5600), 2002-2006. doi:10.1126/science.1077426

Guna, A., Volkmar, N., Christianson, J. C., & Hegde, R. S. (2018). The ER membrane protein complex is a transmembrane domain insertase. *Science, 359*(6374), 470-473. doi:10.1126/science.aao3099

Hart, T., Tong, A. H. Y., Chan, K., Van Leeuwen, J., Seetharaman, A., Aregger, M., . . . Moffat, J. (2017). Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3 (Bethesda), 7*(8), 2719-2727. doi:10.1534/g3.117.041277

Hellmann, J. K., Munter, S., Kudryashev, M., Schulz, S., Heiss, K., Muller, A. K., . . . Frischknecht, F. (2011). Environmental constraints guide migration of malaria parasites during transmission. *PLoS Pathog, 7*(6), e1002080. doi:10.1371/journal.ppat.1002080

Herraiz, T., Guillen, H., Gonzalez-Pena, D., & Aran, V. J. (2019). Antimalarial Quinoline Drugs Inhibit beta-Hematin and Increase Free Hemin Catalyzing Peroxidative Reactions and Inhibition of Cysteine Proteases. *Sci Rep, 9*(1), 15398. doi:10.1038/s41598-019-51604-z

Hoffman, S. L., Goh, L. M., Luke, T. C., Schneider, I., Le, T. P., Doolan, D. L., . . . Richie, T. L. (2002). Protection of humans against malaria by immunization with radiation-attenuated Plasmodium falciparum sporozoites. *J Infect Dis, 185*(8), 1155-1164. doi:10.1086/339409

Hopp, C. S., Chiou, K., Ragheb, D. R., Salman, A. M., Khan, S. M., Liu, A. J., & Sinnis, P. (2015). Longitudinal analysis of Plasmodium sporozoite motility in the dermis reveals component of blood vessel recognition. *Elife, 4*. doi:10.7554/eLife.07789

illumina, I. (2011). Quality Scores for Next-Generation Sequencing. *Technical Note: Sequencing.* Retrieved from https://www.illumina.com/documents/products/technotes/technote_Q-Scores.pdf

Inacio, P., Zuzarte-Luis, V., Ruivo, M. T., Falkard, B., Nagaraj, N., Rooijers, K., . . . Mota, M. M. (2015). Parasite-induced ER stress response in hepatocytes facilitates Plasmodium liver stage infection. *EMBO Rep, 16*(8), 955-964. doi:10.15252/embr.201439979

Ishino, T., Yano, K., Chinzei, Y., & Yuda, M. (2004). Cell-passage activity is required for the malarial parasite to cross the liver sinusoidal cell layer. *PLoS Biol, 2*(1), E4. doi:10.1371/journal.pbio.0020004

Itoe, M. A., Sampaio, J. L., Cabal, G. G., Real, E., Zuzarte-Luis, V., March, S., . . . Mota, M. M. (2014). Host cell phosphatidylcholine is a key mediator of malaria parasite survival during liver stage infection. *Cell Host Microbe, 16*(6), 778-786. doi:10.1016/j.chom.2014.11.006

Jewett, T. J., & Sibley, L. D. (2003). Aldolase forms a bridge between cell surface adhesins and the actin cytoskeleton in apicomplexan parasites. *Mol Cell, 11*(4), 885-894. doi:10.1016/s1097-2765(03)00113-8

Kappe, S., Bruderer, T., Gantt, S., Fujioka, H., Nussenzweig, V., & Menard, R. (1999). Conservation of a gliding motility and cell invasion machinery in Apicomplexan parasites. *J Cell Biol, 147*(5), 937-944. doi:10.1083/jcb.147.5.937

Kariu, T., Ishino, T., Yano, K., Chinzei, Y., & Yuda, M. (2006). CelTOS, a novel malarial protein that mediates transmission to mosquito and vertebrate hosts. *Mol Microbiol, 59*(5), 1369-1379. doi:10.1111/j.1365-2958.2005.05024.x

Kaushansky, A., Douglass, A. N., Arang, N., Vigdorovich, V., Dambrauskas, N., Kain, H. S., . . . Kappe, S. H. (2015). Malaria parasites target the hepatocyte receptor EphA2 for successful host infection. *Science, 350*(6264), 1089-1092. doi:10.1126/science.aad3318

Kaushansky, A., & Kappe, S. H. (2015). Host ER stress during malaria parasite infection. *EMBO Rep, 16*(8), 883-884. doi:10.15252/embr.201540792

Knuepfer, E., Wright, K. E., Kumar Prajapati, S., Rawlinson, T. A., Mohring, F., Koch, M., . . . Holder, A. A. (2019). Divergent roles for the RH5 complex components, CyRPA and RIPR in human-infective malaria parasites. *PLoS Pathog, 15*(6), e1007809. doi:10.1371/journal.ppat.1007809

Koike-Yusa, H., Li, Y., Tan, E. P., Velasco-Herrera Mdel, C., & Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nat Biotechnol, 32*(3), 267-273. doi:10.1038/nbt.2800

Kolde, R., Laur, S., Adler, P., & Vilo, J. (2012). Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics, 28*(4), 573-580. doi:10.1093/bioinformatics/btr709

Kroeger, A., Horstick, O., Riedl, C., Kaiser, A., & Becker, N. (1995). The potential for malaria control with the biological larvicide Bacillus thuringiensis israelensis (Bti) in Peru and Ecuador. *Acta Trop, 60*(1), 47-57. doi:10.1016/0001-706x(95)00101-j

Kumar, K. A., Garcia, C. R., Chandran, V. R., Van Rooijen, N., Zhou, Y., Winzeler, E., & Nussenzweig, V. (2007). Exposure of Plasmodium sporozoites to the intracellular concentration of potassium enhances infectivity and reduces cell passage activity. *Mol Biochem Parasitol, 156*(1), 32-40. doi:10.1016/j.molbiopara.2007.07.004

Lalloo, D. G., & Hill, D. R. (2008). Preventing malaria in travellers. *BMJ, 336*(7657), 1362-1366. doi:10.1136/bmj.a153

Lalremruata, A., Jeyaraj, S., Engleitner, T., Joanny, F., Lang, A., Belard, S., . . . Held, J. (2017). Species and genotype diversity of Plasmodium in malaria patients from Gabon analysed by next generation sequencing. *Malar J, 16*(1), 398. doi:10.1186/s12936-017-2044-0

Lamarque, M., Besteiro, S., Papoin, J., Roques, M., Vulliez-Le Normand, B., Morlon-Guyot, J., . . . Lebrun, M. (2011). The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites. *PLoS Pathog, 7*(2), e1001276. doi:10.1371/journal.ppat.1001276

Leek, J. T., Scharpf, R. B., Bravo, H. C., Simcha, D., Langmead, B., Johnson, W. E., . . . Irizarry, R. A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat Rev Genet, 11*(10), 733-739. doi:10.1038/nrg2825

Lessard, S., Gatof, E. S., Beaudoin, M., Schupp, P. G., Sher, F., Ali, A., . . . Lettre, G. (2017). An erythroid-specific ATP2B4 enhancer mediates red blood cell hydration and malaria susceptibility. *J Clin Invest, 127*(8), 3065-3074. doi:10.1172/JCI94378

Li, W., Koster, J., Xu, H., Chen, C. H., Xiao, T., Liu, J. S., . . . Liu, X. S. (2015). Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol, 16*, 281. doi:10.1186/s13059-015-0843-6

Li, W., Xu, H., Xiao, T., Cong, L., Love, M. I., Zhang, F., . . . Liu, X. S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol, 15*(12), 554. doi:10.1186/s13059-014-0554-4

Lingelbach, K., & Joiner, K. A. (1998). The parasitophorous vacuole membrane surrounding Plasmodium and Toxoplasma: an unusual compartment in infected cells. *J Cell Sci, 111 (Pt 11)*, 1467-1475. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9580555

Lorenz, V., Karanis, G., & Karanis, P. (2014). Malaria vaccine development and how external forces shape it: an overview. *Int J Environ Res Public Health, 11*(7), 6791-6807. doi:10.3390/ijerph110706791

Mac-Daniel, L., Buckwalter, M. R., Berthet, M., Virk, Y., Yui, K., Albert, M. L., . . . Menard, R. (2014). Local immune response to injection of Plasmodium sporozoites into the skin. *J Immunol, 193*(3), 1246-1257. doi:10.4049/jimmunol.1302669

Malhotra, I., Dent, A., Mungai, P., Wamachi, A., Ouma, J. H., Narum, D. L., . . . King, C. L. (2009). Can prenatal malaria exposure produce an immune tolerant phenotype? A prospective birth cohort study in Kenya. *PLoS Med, 6*(7), e1000116. doi:10.1371/journal.pmed.1000116

Matuschewski, K., Nunes, A. C., Nussenzweig, V., & Menard, R. (2002). Plasmodium sporozoite invasion into insect and mammalian cells is directed by the same dual binding system. *EMBO J, 21*(7), 1597-1606. doi:10.1093/emboj/21.7.1597

Medica, D. L., & Sinnis, P. (2005). Quantitative dynamics of Plasmodium yoelii sporozoite transmission by infected anopheline mosquitoes. *Infect Immun, 73*(7), 4363-4369. doi:10.1128/IAI.73.7.4363-4369.2005

Meis, J. F., Rijntjes, P. J., Verhave, J. P., Ponnudurai, T., Hollingdale, M. R., Smith, J. E., . . . Yap, S. H. (1986). Fine structure of the malaria parasite Plasmodium falciparum in human hepatocytes in vitro. *Cell Tissue Res, 244*(2), 345-350. doi:10.1007/BF00219210

Miller, L. H., Baruch, D. I., Marsh, K., & Doumbo, O. K. (2002). The pathogenic basis of malaria. *Nature, 415*(6872), 673-679. doi:10.1038/415673a

Miura, N., Ishida, N., Hoshino, M., Yamauchi, M., Hara, T., Ayusawa, D., & Kawakita, M. (1996). Human UDP-galactose translocator: molecular cloning of a complementary DNA that complements the genetic defect of a mutant cell line deficient in UDP-galactose translocator. *J Biochem, 120*(2), 236-241. doi:10.1093/oxfordjournals.jbchem.a021404

Moreira, C. K., Templeton, T. J., Lavazec, C., Hayward, R. E., Hobbs, C. V., Kroeze, H., . . . Coppi, A. (2008). The Plasmodium TRAP/MIC2 family member, TRAP-Like Protein (TLP), is involved in tissue traversal by sporozoites. *Cell Microbiol, 10*(7), 1505-1516. doi:10.1111/j.1462-5822.2008.01143.x

Mota, M. M., Pradel, G., Vanderberg, J. P., Hafalla, J. C., Frevert, U., Nussenzweig, R. S., . . . Rodriguez, A. (2001). Migration of Plasmodium sporozoites through cells before infection. *Science, 291*(5501), 141-144. doi:10.1126/science.291.5501.141

Ngotho, P., Soares, A. B., Hentzschel, F., Achcar, F., Bertuccini, L., & Marti, M. (2019). Revisiting gametocyte biology in malaria parasites. *FEMS Microbiol Rev, 43*(4), 401-414. doi:10.1093/femsre/fuz010

Nieberler, M., Reuning, U., Reichart, F., Notni, J., Wester, H. J., Schwaiger, M., . . . Kessler, H. (2017). Exploring the Role of RGD-Recognizing Integrins in Cancer. *Cancers (Basel), 9*(9). doi:10.3390/cancers9090116

Nussenzweig, R. S., Vanderberg, J., Most, H., & Orton, C. (1967). Protective immunity produced by the injection of x-irradiated sporozoites of plasmodium berghei. *Nature, 216*(5111), 160-162. doi:10.1038/216160a0

O'Donnell, R. A., & Blackman, M. J. (2005). The role of malaria merozoite proteases in red blood cell invasion. *Curr Opin Microbiol, 8*(4), 422-427. doi:10.1016/j.mib.2005.06.018

Pedchenko, V., Zent, R., & Hudson, B. G. (2004). Alpha(v)beta3 and alpha(v)beta5 integrins bind both the proximal RGD site and non-RGD motifs within noncollagenous (NC1) domain of the alpha3 chain of type IV collagen: implication for the mechanism of endothelia cell adhesion. *J Biol Chem, 279*(4), 2772-2780. doi:10.1074/jbc.M311901200

Peixoto, L., Chen, F., Harb, O. S., Davis, P. H., Beiting, D. P., Brownback, C. S., . . . Roos, D. S. (2010). Integrative genomic approaches highlight a family of parasite-specific kinases that regulate host responses. *Cell Host Microbe, 8*(2), 208-218. doi:10.1016/j.chom.2010.07.004

Pittman, K. J., & Knoll, L. J. (2015). Long-Term Relationships: the Complicated Interplay between the Host and the Developmental Stages of Toxoplasma gondii during Acute and Chronic Infections. *Microbiol Mol Biol Rev, 79*(4), 387-401. doi:10.1128/MMBR.00027-15

Pradel, G., Garapaty, S., & Frevert, U. (2004). Kupffer and stellate cell proteoglycans mediate malaria sporozoite targeting to the liver. *Comp Hepatol, 3 Suppl 1*, S47. doi:10.1186/1476-5926-2-S1-S47

Prudencio, M., Rodriguez, A., & Mota, M. M. (2006). The silent path to thousands of merozoites: the Plasmodium liver stage. *Nat Rev Microbiol, 4*(11), 849-856. doi:10.1038/nrmicro1529

Reddy, K. S., Amlabu, E., Pandey, A. K., Mitra, P., Chauhan, V. S., & Gaur, D. (2015). Multiprotein complex between the GPI-anchored CyRPA with PfRH5 and PfRipr is crucial for Plasmodium falciparum erythrocyte invasion. *Proc Natl Acad Sci U S A, 112*(4), 1179-1184. doi:10.1073/pnas.1415466112

Riglar, D. T., Rogers, K. L., Hanssen, E., Turnbull, L., Bullen, H. E., Charnaud, S. C., . . . Cowman, A. F. (2013). Spatial association with PTEX complexes defines regions for effector export into Plasmodium falciparum-infected erythrocytes. *Nat Commun, 4*, 1415. doi:10.1038/ncomms2449

Risco-Castillo, V., Topcu, S., Marinach, C., Manzoni, G., Bigorgne, A. E., Briquet, S., . . . Silvie, O. (2015). Malaria Sporozoites Traverse Host Cells within Transient Vacuoles. *Cell Host Microbe, 18*(5), 593-603. doi:10.1016/j.chom.2015.10.006

Robson, K. J., Frevert, U., Reckmann, I., Cowan, G., Beier, J., Scragg, I. G., . . . et al. (1995). Thrombospondin-related adhesive protein (TRAP) of Plasmodium falciparum: expression during sporozoite ontogeny and binding to human hepatocytes. *EMBO J, 14*(16), 3883-3894. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7664729

Rodrigues, C. D., Hannus, M., Prudencio, M., Martin, C., Goncalves, L. A., Portugal, S., . . . Mota, M. M. (2008). Host scavenger receptor SR-BI plays a dual role in the establishment of malaria parasite liver infection. *Cell Host Microbe, 4*(3), 271-282. doi:10.1016/j.chom.2008.07.012

Ruoslahti, E. (1996). RGD and other recognition sequences for integrins. *Annu Rev Cell Dev Biol, 12*, 697-715. doi:10.1146/annurev.cellbio.12.1.697

Sá e Cunha, C., Nyboer, B., Heiss, K., Sanches-Vaz, M., Fontinha, D., Wiedtke, E., . . . Mueller, A. (2017). Plasmodium berghei EXP-1 interacts with host Apolipoprotein H during Plasmodium liver-stage development. *Proceedings of the National Academy of Sciences, 115*(7), E1138-E1147. doi:10.1073/pnas.1606419114

Sa, E. C. C., Nyboer, B., Heiss, K., Sanches-Vaz, M., Fontinha, D., Wiedtke, E., . . . Mueller, A. K. (2017). Plasmodium berghei EXP-1 interacts with host Apolipoprotein H during Plasmodium liver-stage development. *Proc Natl Acad Sci U S A, 114*(7), E1138-E1147. doi:10.1073/pnas.1606419114

Sattabongkot, J., Yimamnuaychoke, N., Leelaudomlipi, S., Rasameesoraj, M., Jenwithisuk, R., Coleman, R. E., . . . Brewer, T. G. (2006). Establishment of a human hepatocyte line that supports in vitro development of the exo-erythrocytic stages of the malaria parasites Plasmodium falciparum and P. vivax. *Am J Trop Med Hyg, 74*(5), 708-715. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/16687667

Segireddy, R. R., Dundas, K., Knoeckel, J., Galaway, F., Wood, L., Wright, G. J., & Douglas, A. D. (2020). A screen for Plasmodium falciparum sporozoite surface protein binding to human hepatocyte surface receptors identifies novel host-pathogen interactions. *bioRxiv*, 2020.2002.2002.929190. doi:10.1101/2020.02.02.929190

Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelson, T., . . . Zhang, F. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science, 343*(6166), 84-87. doi:10.1126/science.1247005

Shen, B., & Sibley, L. D. (2012). The moving junction, a key portal to host cell invasion by apicomplexan parasites. *Curr Opin Microbiol, 15*(4), 449-455. doi:10.1016/j.mib.2012.02.007

Sherling, E. S., Knuepfer, E., Brzostowski, J. A., Miller, L. H., Blackman, M. J., & van Ooij, C. (2017). The Plasmodium falciparum rhoptry protein RhopH3 plays essential roles in host cell invasion and nutrient uptake. *Elife, 6*. doi:10.7554/eLife.23239

Silvie, O., Franetich, J. F., Boucheix, C., Rubinstein, E., & Mazier, D. (2007). Alternative invasion pathways for Plasmodium berghei sporozoites. *Int J Parasitol, 37*(2), 173-182. doi:10.1016/j.ijpara.2006.10.005

Silvie, O., Rubinstein, E., Franetich, J. F., Prenant, M., Belnoue, E., Renia, L., . . . Mazier, D. (2003). Hepatocyte CD81 is required for Plasmodium falciparum and Plasmodium yoelii sporozoite infectivity. *Nat Med, 9*(1), 93-96. doi:10.1038/nm808

Sinnis, P., De La Vega, P., Coppi, A., Krzych, U., & Mota, M. M. (2013). Quantification of sporozoite invasion, migration, and development by microscopy and flow cytometry. *Methods Mol Biol, 923*, 385-400. doi:10.1007/978-1-62703-026-7_27

Song, G., Koksal, A. C., Lu, C., & Springer, T. A. (2012). Shape change in the receptor for gliding motility in Plasmodium sporozoites. *Proc Natl Acad Sci U S A, 109*(52), 21420-21425. doi:10.1073/pnas.1218581109

Spring, M., Murphy, J., Nielsen, R., Dowler, M., Bennett, J. W., Zarling, S., . . . Kappe, S. H. (2013). First-in-human evaluation of genetically attenuated Plasmodium falciparum sporozoites administered by bite of Anopheles mosquitoes to adult volunteers. *Vaccine, 31*(43), 4975-4983. doi:10.1016/j.vaccine.2013.08.007

Srinivasan, P., Beatty, W. L., Diouf, A., Herrera, R., Ambroggio, X., Moch, J. K., . . . Miller, L. H. (2011). Binding of Plasmodium merozoite proteins RON2 and AMA1 triggers commitment to invasion. *Proc Natl Acad Sci U S A, 108*(32), 13275-13280. doi:10.1073/pnas.1110303108

Stanway, R. R., Bushell, E., Chiappino-Pepe, A., Roques, M., Sanderson, T., Franke-Fayard, B., . . . Heussler, V. T. (2019). Genome-Scale Identification of Essential Metabolic Processes for Targeting the Plasmodium Liver Stage. *Cell, 179*(5), 1112-1128 e1126. doi:10.1016/j.cell.2019.10.030

Stewart, S. E., Menzies, S. A., Popa, S. J., Savinykh, N., Petrunkina Harrison, A., Lehner, P. J., & Moreau, K. (2017). A genome-wide CRISPR screen reconciles the role of N-linked glycosylation in galectin-3 transport to the cell surface. *J Cell Sci, 130*(19), 3234-3247. doi:10.1242/jcs.206425

Strode, C., Donegan, S., Garner, P., Enayati, A. A., & Hemingway, J. (2014). The impact of pyrethroid resistance on the efficacy of insecticide-treated bed nets against African anopheline mosquitoes: systematic review and meta-analysis. *PLoS Med, 11*(3), e1001619. doi:10.1371/journal.pmed.1001619

Stubbs, J., Simpson, K. M., Triglia, T., Plouffe, D., Tonkin, C. J., Duraisingh, M. T., . . . Cowman, A. F. (2005). Molecular mechanism for switching of P. falciparum invasion pathways into human erythrocytes. *Science, 309*(5739), 1384-1387. doi:10.1126/science.1115257

Stuiver, I., Ruggeri, Z., & Smith, J. W. (1996). Divalent cations regulate the organization of integrins alpha v beta 3 and alpha v beta 5 on the cell surface. *J Cell Physiol, 168*(3), 521-531. doi:10.1002/(SICI)1097-4652(199609)168:3<521::AID-JCP4>3.0.CO;2-R

Sultan, A. A., Thathy, V., Frevert, U., Robson, K. J., Crisanti, A., Nussenzweig, V., . . . Menard, R. (1997). TRAP is necessary for gliding motility and infectivity of plasmodium sporozoites. *Cell, 90*(3), 511-522. doi:10.1016/s0092-8674(00)80511-5

Takken, W., & Knols, B. G. (2009). Malaria vector control: current and future strategies. *Trends Parasitol, 25*(3), 101-104. doi:10.1016/j.pt.2008.12.002

Talman, A. M., Lacroix, C., Marques, S. R., Blagborough, A. M., Carzaniga, R., Menard, R., & Sinden, R. E. (2011). PbGEST mediates malaria transmission to both mosquito and vertebrate host. *Mol Microbiol, 82*(2), 462-474. doi:10.1111/j.1365-2958.2011.07823.x

Tao, D., King, J. G., Tweedell, R. E., Jost, P. J., Boddey, J. A., & Dinglasan, R. R. (2014). The acute transcriptomic and proteomic response of HC-04 hepatoma cells to hepatocyte

growth factor and its implications for Plasmodium falciparum sporozoite invasion. *Mol Cell Proteomics, 13*(5), 1153-1164. doi:10.1074/mcp.M113.035584

Tavares, J., Formaglio, P., Thiberge, S., Mordelet, E., Van Rooijen, N., Medvinsky, A., . . . Amino, R. (2013). Role of host cell traversal by the malaria sporozoite during liver infection. *J Exp Med, 210*(5), 905-915. doi:10.1084/jem.20121130

Timmann, C., Thye, T., Vens, M., Evans, J., May, J., Ehmen, C., . . . Horstmann, R. D. (2012). Genome-wide association study indicates two novel resistance loci for severe malaria. *Nature, 489*(7416), 443-446. doi:10.1038/nature11334

Tuteja, N. (2007). Abscisic Acid and abiotic stress signaling. *Plant Signal Behav, 2*(3), 135-138. doi:10.4161/psb.2.3.4156

Tweedell, R. E., Tao, D., Hamerly, T., Robinson, T. M., Larsen, S., Gronning, A. G. B., . . . Dinglasan, R. R. (2019). The Selection of a Hepatocyte Cell Line Susceptible to Plasmodium falciparum Sporozoite Invasion That Is Associated With Expression of Glypican-3. *Front Microbiol, 10*, 127. doi:10.3389/fmicb.2019.00127

Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., . . . Ponten, F. (2015). Proteomics. Tissue-based map of the human proteome. *Science, 347*(6220), 1260419. doi:10.1126/science.1260419

Van Agthoven, J. F., Xiong, J. P., Alonso, J. L., Rui, X., Adair, B. D., Goodman, S. L., & Arnaout, M. A. (2014). Structural basis for pure antagonism of integrin alphaVbeta3 by a high-affinity form of fibronectin. *Nat Struct Mol Biol, 21*(4), 383-388. doi:10.1038/nsmb.2797

Varki, A., & Varki, N. M. (2001). P-selectin, carcinoma metastasis and heparin: novel mechanistic connections with therapeutic implications. *Braz J Med Biol Res, 34*(6), 711-717. doi:10.1590/s0100-879x2001000600003

Volz, J. C., Yap, A., Sisquella, X., Thompson, J. K., Lim, N. T., Whitehead, L. W., . . . Cowman, A. F. (2016). Essential Role of the PfRh5/PfRipr/CyRPA Complex during Plasmodium falciparum Invasion of Erythrocytes. *Cell Host Microbe, 20*(1), 60-71. doi:10.1016/j.chom.2016.06.004

Wanaguru, M., Liu, W., Hahn, B. H., Rayner, J. C., & Wright, G. J. (2013). RH5-Basigin interaction plays a major role in the host tropism of Plasmodium falciparum. *Proc Natl Acad Sci U S A, 110*(51), 20735-20740. doi:10.1073/pnas.1320771110

Wang, B., Wang, M., Zhang, W., Xiao, T., Chen, C. H., Wu, A., . . . Liu, X. S. (2019). Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. *Nat Protoc, 14*(3), 756-780. doi:10.1038/s41596-018-0113-7

Wang, S., Zhang, Q., Tiwari, S. K., Lichinchi, G., Yau, E. H., Hui, H., . . . Rana, T. M. (2020). Integrin alphavbeta5 Internalizes Zika Virus during Neural Stem Cells Infection and Provides a Promising Target for Antiviral Therapy. *Cell Rep, 30*(4), 969-983 e964. doi:10.1016/j.celrep.2019.11.020

Warhurst, D. C., Steele, J. C., Adagu, I. S., Craig, J. C., & Cullander, C. (2003). Hydroxychloroquine is much less active than chloroquine against chloroquine-resistant Plasmodium falciparum, in agreement with its physicochemical properties. *J Antimicrob Chemother, 52*(2), 188-193. doi:10.1093/jac/dkg319

Weiss, G. E., Gilson, P. R., Taechalertpaisarn, T., Tham, W. H., de Jong, N. W., Harvey, K. L., . . . Crabb, B. S. (2015). Revealing the sequence and resulting cellular morphology of receptor-ligand interactions during Plasmodium falciparum invasion of erythrocytes. *PLoS Pathog, 11*(2), e1004670. doi:10.1371/journal.ppat.1004670

WHO. (2019). *World malaria report 2019*. Retrieved from

Willcocks, L. C., Carr, E. J., Niederer, H. A., Rayner, T. F., Williams, T. N., Yang, W., . . . Smith, K. G. (2010). A defunctioning polymorphism in FCGR2B is associated with

protection against malaria but susceptibility to systemic lupus erythematosus. *Proc Natl Acad Sci U S A, 107*(17), 7881-7885. doi:10.1073/pnas.0915133107

Wisniewski, J. R., Vildhede, A., Noren, A., & Artursson, P. (2016). In-depth quantitative analysis and comparison of the human hepatocyte and hepatoma cell line HepG2 proteomes. *J Proteomics, 136*, 234-247. doi:10.1016/j.jprot.2016.01.016

Wong, W., Huang, R., Menant, S., Hong, C., Sandow, J. J., Birkinshaw, R. W., . . . Cowman, A. F. (2019). Structure of Plasmodium falciparum Rh5-CyRPA-Ripr invasion complex. *Nature, 565*(7737), 118-121. doi:10.1038/s41586-018-0779-6

World Health Organization [WHO]. (2019). *World malaria report 2019*. Retrieved from

Wright, G. J., & Rayner, J. C. (2014). Plasmodium falciparum erythrocyte invasion: combining function with immune evasion. *PLoS Pathog, 10*(3), e1003943. doi:10.1371/journal.ppat.1003943

Wu, X., Scott, D. A., Kriz, A. J., Chiu, A. C., Hsu, P. D., Dadon, D. B., . . . Sharp, P. A. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol, 32*(7), 670-676. doi:10.1038/nbt.2889

Yamamoto, M., & Takeda, K. (2012). Inhibition of ATF6beta-dependent host adaptive immune response by a Toxoplasma virulence factor ROP18. *Virulence, 3*(1), 77-80. doi:10.4161/viru.3.1.18340

Yamauchi, L. M., Coppi, A., Snounou, G., & Sinnis, P. (2007). Plasmodium sporozoites trickle out of the injection site. *Cell Microbiol, 9*(5), 1215-1222. doi:10.1111/j.1462-5822.2006.00861.x

Yang, A. S., & Boddey, J. A. (2017). Molecular mechanisms of host cell traversal by malaria sporozoites. *Int J Parasitol, 47*(2-3), 129-136. doi:10.1016/j.ijpara.2016.09.002

Yang, A. S. P., Lopaticki, S., O'Neill, M. T., Erickson, S. M., Douglas, D. N., Kneteman, N. M., & Boddey, J. A. (2017). AMA1 and MAEBL are important for Plasmodium falciparum sporozoite infection of the liver. *Cell Microbiol, 19*(9). doi:10.1111/cmi.12745

Yang, A. S. P., O'Neill, M. T., Jennison, C., Lopaticki, S., Allison, C. C., Armistead, J. S., . . . Boddey, J. A. (2017). Cell Traversal Activity Is Important for Plasmodium falciparum Liver Infection in Humanized Mice. *Cell Rep, 18*(13), 3105-3116. doi:10.1016/j.celrep.2017.03.017

Yang, S., Li, X. J., & Li, S. (2016). Molecular mechanisms underlying Spinocerebellar Ataxia 17 (SCA17) pathogenesis. *Rare Dis, 4*(1), e1223580. doi:10.1080/21675511.2016.1223580

Ylipaasto, P., Eskelinen, M., Salmela, K., Hovi, T., & Roivainen, M. (2010). Vitronectin receptors, alpha v integrins, are recognized by several non-RGD-containing echoviruses in a continuous laboratory cell line and also in primary human Langerhans' islets and endothelial cells. *J Gen Virol, 91*(Pt 1), 155-165. doi:10.1099/vir.0.012450-0

Zhang, X. H., Tee, L. Y., Wang, X. G., Huang, Q. S., & Yang, S. H. (2015). Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol Ther Nucleic Acids, 4*, e264. doi:10.1038/mtna.2015.37

APPENDIX A: SGRNA SEQUENCES OF POOL 1, 2, 3, 4 AND 5

Pool 1		
Gene	sgRNA name	sgRNA sequence (19 bp)
Non-targeting 1	nontargeting-1	GTCCATGGGTGGAGTTACG
Non-targeting 2	nontargeting-2	GGACAAGTTGGACAGTTGC
Non-targeting 3	nontargeting-3	AACAATGGCTTCGTCGACT
Non-targeting 4	nontargeting-4	ACCTTATGTTCGCCGCCTC
Non-targeting 5	nontargeting-5	AGACACCGTAACTCGAATT
Non-targeting 6	nontargeting-6	CTCGTAGGTCAATCGCGGA
Non-targeting 7	nontargeting-7	GACGGTTCCGATAACCTTA
Non-targeting 8	nontargeting-8	GGATCTCTCAATATGGCGC
Non-targeting 9	nontargeting-9	AGCGCCGTTGTTGAGAAGA
Non-targeting 10	nontargeting-10	CCGCCGGCCCTAACTGACC
Non-targeting 11	nontargeting-11	TGACTAGACCCTTACGCGG
Non-targeting 12	nontargeting-12	ATCGGAGAGCCTGAAGTTG
Non-targeting 13	nontargeting-13	TCGGGCCGCTATATAGGCG
Non-targeting 14	nontargeting-14	AGAACTCAGCGGTCTCAAG
Non-targeting 15	nontargeting-15	ATGTCCGTTGTAGTCCTCG
Non-targeting 16	nontargeting-16	CCGGTCGCAATTGTTGCTA
Non-targeting 17	nontargeting-17	TGCCTATTCAGTAGGTCGT
Non-targeting 18	nontargeting-18	GTGCGGTAAGCGTGGCTAG
Non-targeting 19	nontargeting-19	CCATGGAGCCCAATCAATC
Non-targeting 20	nontargeting-20	GTGTAGTTCGGATTGATGA
Non-targeting 21	nontargeting-21	GGAATTGGAGACGATGCCG
Non-targeting 22	nontargeting-22	AGTGCCTCGTAGTTGTGGG
Non-targeting 23	nontargeting-23	CCTCGTTTCCACGTAATTC
Non-targeting 24	nontargeting-24	ATTTGATGTGGTGCTCACA
Non-targeting 25	nontargeting-25	ATTAGGTCGTTCTGTAACA
Non-targeting 26	nontargeting-26	CCGGTAGTGACCACGAACA
Non-targeting 27	nontargeting-27	GCCACACACGCGGCTTTAC
Non-targeting 28	nontargeting-28	GGCCTCATTAGTCGTCTGG
Non-targeting 29	nontargeting-29	AAGCCGCGAATGGCTAACT
Non-targeting 30	nontargeting-30	ATATGCTAGACGTGACTCA
Non-targeting 31	nontargeting-31	TTAGCAATCGTACGGGTCA
Non-targeting 32	nontargeting-32	TCTCTTCGGATTCTTGGCA
Non-targeting 33	nontargeting-33	CGGGTCTGGAAAGGCGTAT
Non-targeting 34	nontargeting-34	AGCAGCGCGACAACCCATT
Non-targeting 35	nontargeting-35	CTCCGTAGTTCGGGACGCT
Non-targeting 36	nontargeting-36	ACTCCGGCAAGCCAGCTCT
Non-targeting 37	nontargeting-37	CCCGCTCTAGCCTATAAGA

Non-targeting 38	nontargeting-38	GAAGATTCCCACTCATCCG
Non-targeting 39	nontargeting-39	TACGGATGCTGAACTCTAG
Non-targeting 40	nontargeting-40	AACCTACGGGCTACGATAC
Non-targeting 41	nontargeting-41	CGTGTGAAGGGCGATATAC
Non-targeting 42	nontargeting-42	CTAGGCGCGGGAGTCCGCT
Non-targeting 43	nontargeting-43	GACCCAGCTGCAGTAACGA
Non-targeting 44	nontargeting-44	ATTGGTTACCTTGCGTCTC
Non-targeting 45	nontargeting-45	AGTCTGACGTCGACTATCT
Non-targeting 46	nontargeting-46	GGTTCTGCCGCTTATAATA
Non-targeting 47	nontargeting-47	GCAATTCAGTGCTATATCG
Non-targeting 48	nontargeting-48	GCGTCGGATCGATGGTTCT
Non-targeting 49	nontargeting-49	TTCGATCCGGCGACTCGAA
Non-targeting 50	nontargeting-50	TCGGTACATGTATAGACGA
CD81	CD81_CCDS7734.1.ex1_11:2411694-2411717:+	ACCACCAACCTCCTGTATC
	CD81_CCDS7734.1.ex1_11:2411744-2411767:-	GTGCACTCACCTACATAGA
	CD81_CCDS7734.1.ex2_11:2415376-2415399:+	GGCTGCTACGGGGCCATCC
	CD81_CCDS7734.1.ex3_11:2416244-2416267:-	TTGACAAAGCCCCAGATGC
	CD81_CCDS7734.1.ex4_11:2416641-2416664:-	CTTCACATCCTTGGCGATC
NPC1	NPC1_CCDS11878.1.ex15_18:21131629-21131652:+	ACCAAACGTACCCAGACAA
	NPC1_CCDS11878.1.ex16_18:21134841-21134864:+	TCGTGTTATACGGTGAAAG
	NPC1_CCDS11878.1.ex17_18:21136241-21136264:+	GGGTACATCAGCTCCCGAA
	NPC1_CCDS11878.1.ex17_18:21136427-21136450:-	ACTGCGTGTTCGTCAGGCC
	NPC1_CCDS11878.1.ex17_18:21136504-21136527:-	GGCGGCTGTTCACACGCTG
SLC10A1	SLC10A1_CCDS9797.1.ex3_14:70252846-70252869:+	GATTTGAGGACGATCCCTA
	SLC10A1_CCDS9797.1.ex3_14:70252884-70252907:-	ATCGTGATATCACTGGTCC
	SLC10A1_CCDS9797.1.ex4_14:70263552-70263575:+	CATTGGACAGGTTCCCTCC
	SLC10A1_CCDS9797.1.ex4_14:70263658-70263681:+	GATGCCATACTGTGCCACC
	SLC10A1_CCDS9797.1.ex4_14:70263693-70263716:-	CTCACTTATGGAAGCCTAA
CD55	CD55_CCDS31006.1.ex1_1:207495775-207495798:+	GGCCGTACAAGTTTTCCCG
	CD55_CCDS31006.1.ex1_1:207495832-207495855:+	GTGAAAATTCCTGGCGAGA
	CD55_CCDS31006.1.ex2_1:207497892-207497915:+	ATACTTTTAGGTAGCTGCG
	CD55_CCDS31006.1.ex2_1:207497974-207497997:-	ATTCCACAACAGTACCGAC
	CD55_CCDS31006.1.ex3_1:207499001-207499024:+	GTCAGATTGATGTACCAGG
HAVCR1	HAVCR1_CCDS43392.1.ex5_5:156479553-156479576:+	TCGTTGGAACAGTCGTCAT
	HAVCR1_CCDS43392.1.ex5_5:156479579-156479602:+	GTCGTTGTCGTTGGAACAG
	HAVCR1_CCDS43392.1.ex5_5:156479607-156479630:+	TTCGAACAGTCGTGACGGT
	HAVCR1_CCDS43392.1.ex6_5:156482241-156482264:+	CATGTCATTGAACCACCCA
	HAVCR1_CCDS43392.1.ex6_5:156482411-156482434:-	CTCTATTCACATGCCAAAA
CD68	CD68_CCDS11114.1.ex1_17:7483328-7483351:-	ACTGTGACGTTTCCATGGC
	CD68_CCDS11114.1.ex1_17:7483412-7483435:-	GCATTTCCATGACTAGTGG
	CD68_CCDS11114.1.ex1_17:7483433-7483456:-	TTGCTTGTTGGATGAACCG
	CD68_CCDS11114.1.ex1_17:7483541-7483564:+	GACCATTGGAGACTACACG

	CD68_CCDS11114.1.ex1_17:7483599-7483622:-	CATGACTCGAATCTGAATC
PIGS	PIGS_CCDS11235.1.ex5_17:26887081-26887104:+	AAGTTGCCAGCGGCACCGA
	PIGS_CCDS11235.1.ex6_17:26888583-26888606:-	CGGGGGATAATGCACCGGG
	PIGS_CCDS11235.1.ex8_17:26890836-26890859:+	TGCACACTGCCCGATGACA
	PIGS_CCDS11235.1.ex9_17:26897885-26897908:+	CTTTCATGCACAACGGTGA
	PIGS_CCDS11235.1.ex9_17:26897943-26897966:+	GCGTAAACACGACAGTGAC
CSGALNACT1	CSGALNACT1_CCDS6010.1.ex5_8:19316097-19316120:-	GTATGAGCTCACCTTCAAA
	CSGALNACT1_CCDS6010.1.ex6_8:19362867-19362890:-	CTACAGAAGGTGTACCAGC
	CSGALNACT1_CCDS6010.1.ex6_8:19362927-19362950:-	GTGAATGCTGGCGTCAAGC
	CSGALNACT1_CCDS6010.1.ex6_8:19363147-19363170:+	TCCTCCCACTCCTGAAGGA
	CSGALNACT1_CCDS6010.1.ex6_8:19363225-19363248:+	TCGTCACCTTTTGGGGTGC
EXT2	EXT2_CCDS53618.1.ex1_11:44129407-44129430:+	TATCGAGTCCTCAAATGAC
	EXT2_CCDS53618.1.ex1_11:44129635-44129658:-	TCATTATACTCCCGGGAGA
	EXT2_CCDS53618.1.ex1_11:44129724-44129747:-	GGTTAAGCACATCGATGGA
	EXT2_CCDS53618.1.ex3_11:44135759-44135782:+	CTACGTGGACTTACCGGCA
	EXT2_CCDS53618.1.ex4_11:44146473-44146496:-	GCAGCGCTTACGGACAGAA
RPN1	RPN1_CCDS3051.1.ex5_3:128348943-128348966:+	ATCCCGGTAATAAACATCC
	RPN1_CCDS3051.1.ex6_3:128350833-128350856:+	GGTAATCATAGCGTGAGAA
	RPN1_CCDS3051.1.ex6_3:128350928-128350951:+	TCAATGACTCGGGTCATGC
	RPN1_CCDS3051.1.ex7_3:128356682-128356705:+	TAGGTCCTCAGAGCGCGTG
	RPN1_CCDS3051.1.ex7_3:128356845-128356868:+	GGATACGGATGAAGCACAT
CANX	CANX_CCDS4447.1.ex1_5:179133273-179133296:-	ATACTTCCCCTGTTGGAAC
	CANX_CCDS4447.1.ex5_5:179136892-179136915:-	CCAAACATAATCGTATAAG
	CANX_CCDS4447.1.ex6_5:179143140-179143163:+	ACCAATCTGTGGTGAATAG
	CANX_CCDS4447.1.ex6_5:179143185-179143208:-	CACGTGAAGGATTTACAGG
	CANX_CCDS4447.1.ex6_5:179143217-179143240:+	CCAGAAGACCGGAAGCCCG
TMEM205	TMEM205_CCDS32909.1.ex0_19:11453699-11453722:-	ATGTGGGCCCTGCAAACCG
	TMEM205_CCDS32909.1.ex0_19:11453753-11453776:+	GTTGACAGTGGCCAGCGTA
	TMEM205_CCDS32909.1.ex1_19:11456035-11456058:+	TTGCTCTGCACTAGTCCGA
	TMEM205_CCDS32909.1.ex1_19:11456049-11456072:+	TCCGAAGGTATGTCGGGGA
	TMEM205_CCDS32909.1.ex2_19:11456190-11456213:-	GGTGACCTTCGTCTCAGGT
REEP5	REEP5_CCDS4109.2.ex2_5:112238142-112238165:+	ACACCATACACTACCCAGT
	REEP5_CCDS4109.2.ex2_5:112238169-112238192:-	CAAAGAAGATGATACCCAG
	REEP5_CCDS4109.2.ex3_5:112256851-112256874:+	CACTTACGAGATGTAGGCT
	REEP5_CCDS4109.2.ex3_5:112256878-112256901:+	AAATCCTATCAGGTTGCAG
	REEP5_CCDS4109.2.ex3_5:112256905-112256928:+	GGCTCCATAACCGAACACC
CKAP4	CKAP4_CCDS9103.1.ex0_12:106633453-106633476:-	AGCTGAAGTCCGATTCCCA
	CKAP4_CCDS9103.1.ex0_12:106633625-106633648:-	TACACCGAGGTCCGCGAGC
	CKAP4_CCDS9103.1.ex0_12:106633658-106633681:-	AGAAGTACCCTTCAGACTA
	CKAP4_CCDS9103.1.ex0_12:106633781-106633804:-	AAGGCAAAGGTTGCCTCCC
	CKAP4_CCDS9103.1.ex0_12:106633946-106633969:+	CACATGGATCCCATCCGAG
HSD17B2	HSD17B2_CCDS10936.1.ex0_16:82069186-82069209:-	AGGATCAAGCCCCAAAAAG

	HSD17B2_CCDS10936.1.ex0_16:82069205-82069228:-	GAAGCATGACACCGAGAAG
	HSD17B2_CCDS10936.1.ex0_16:82069263-82069286:-	GGACTGCCTTCTGATCCAC
	HSD17B2_CCDS10936.1.ex1_16:82101863-82101886:-	TTCGCAATTCCTCAGCTCC
	HSD17B2_CCDS10936.1.ex1_16:82101898-82101921:+	CGCCTCTCGGTGCTCCAAA
ASGR1	ASGR1_CCDS11089.1.ex3_17:7077697-7077720:-	GTCAGATGGCGGCGCTCCA
	ASGR1_CCDS11089.1.ex4_17:7080175-7080198:-	AAGTCGCTAGAGTCCCAGC
	ASGR1_CCDS11089.1.ex5_17:7080299-7080322:-	GTCAAGGGCTTGAGCACCC
	ASGR1_CCDS11089.1.ex5_17:7080326-7080349:-	AACTTCACAGCGAGCACGG
	ASGR1_CCDS11089.1.ex5_17:7080344-7080367:+	GTTGCTGAACGTCTCTCTC
LMAN2	LMAN2_CCDS4417.1.ex3_5:176764497-176764520:-	TTCCCGTACATCTCGGTGA
	LMAN2_CCDS4417.1.ex5_5:176765585-176765608:-	GCCGTGCTTCCTCAAAGAC
	LMAN2_CCDS4417.1.ex6_5:176778211-176778234:+	AGGGGTCAGACGTACGTAC
	LMAN2_CCDS4417.1.ex7_5:176778451-176778474:-	CGCTCATTAAGCCCTACCA
	LMAN2_CCDS4417.1.ex7_5:176778502-176778525:-	TGACTGCGGATATAACTGA
ANPEP	ANPEP_CCDS10356.1.ex13_15:90346918-90346941:-	GCCTCGGAGATCAACACGC
	ANPEP_CCDS10356.1.ex15_15:90347599-90347622:-	GACTTCAACGCCGGCGCCA
	ANPEP_CCDS10356.1.ex17_15:90348334-90348357:-	GTCAGTGAGTTCGACTACG
	ANPEP_CCDS10356.1.ex19_15:90349495-90349518:+	GGAGCCCTTAAAAACGTAC
	ANPEP_CCDS10356.1.ex19_15:90349547-90349570:+	AGCGTCACCCGGTAGGAAT
ATP1B1	ATP1B1_CCDS1276.1.ex1_1:169080614-169080637:-	AAATATTACGTAGAATAGA
	ATP1B1_CCDS1276.1.ex1_1:169080688-169080711:-	GTGGGCTTAAATTCACTGA
	ATP1B1_CCDS1276.1.ex1_1:169080727-169080750:-	ATGTGGATTTTACCTGGCG
	ATP1B1_CCDS1276.1.ex2_1:169094160-169094183:-	TTGGGATCATTAGGACGAA
	ATP1B1_CCDS1276.1.ex2_1:169094242-169094265:-	AAAAATCATGTCATCCCTC
MTDH	MTDH_CCDS6274.1.ex0_8:98656862-98656885:+	CCGAAACGGTACCCCGGCT
	MTDH_CCDS6274.1.ex0_8:98656924-98656947:+	GTTTCTGCTGGGCTACGGC
	MTDH_CCDS6274.1.ex0_8:98656957-98656980:+	CGCCGGCGCCCGCAAAAAG
ASPH	ASPH_CCDS34898.1.ex13_8:62550528-62550551:+	CTACAGGATTATCCTCAGG
	ASPH_CCDS34898.1.ex18_8:62557167-62557190:-	ACCGAGCATAGTTACCACG
	ASPH_CCDS34898.1.ex21_8:62566170-62566193:+	CCTCTTCTGGCGGGACTGC
	ASPH_CCDS34899.1.ex1_8:62580806-62580829:+	TTCTCTTGGCTACCCCACT
	ASPH_CCDS34899.1.ex1_8:62580820-62580843:-	GTCAACAGAAGGACCCAGT
SLC39A14	SLC39A14_CCDS47822.1.ex0_8:22262347-22262370:-	TGTATTAGATCCTGCAGGA
	SLC39A14_CCDS47822.1.ex0_8:22262360-22262383:+	CTAATACATCGGTATGGCG
	SLC39A14_CCDS47822.1.ex1_8:22265853-22265876:+	CAATTTCAGCGAGCAGTCG
	SLC39A14_CCDS47822.1.ex3_8:22272291-22272314:+	GCATTTGGTTTCAACCCTC
	SLC39A14_CCDS47822.1.ex3_8:22272329-22272352:-	CCAAACACCACTGCAGACT
STIM1	STIM1_CCDS60706.1.ex2_11:4045125-4045148:-	GTGTTTCACTGTTGGGTCA
	STIM1_CCDS60706.1.ex2_11:4045161-4045184:+	GAGGATAAGCTCATCAGCG
	STIM1_CCDS60706.1.ex3_11:4076750-4076773:+	ACAGTATACAATTGGACCG
	STIM1_CCDS60706.1.ex3_11:4076852-4076875:-	TCCTGACCTTGGCATGGCA
	STIM1_CCDS60706.1.ex5_11:4091334-4091357:+	ATCCAGAACCGTTACTCCA

DPP4	DPP4_CCDS2216.1.ex13_2:162877150-162877173:+	TAGAAGCTATTACCATCAA
	DPP4_CCDS2216.1.ex15_2:162881360-162881383:-	ATTCAGAACTATTCGGTCA
	DPP4_CCDS2216.1.ex18_2:162894857-162894880:-	TTACAGAATCACATGGACG
	DPP4_CCDS2216.1.ex19_2:162895453-162895476:+	TAACCAATTTATGACCCAC
	DPP4_CCDS2216.1.ex21_2:162902040-162902063:-	TTAGAATACAACTACGTGA
TM9SF3	TM9SF3_CCDS7450.1.ex8_10:98311063-98311086:+	CATCCAGAACCAATCAGAG
	TM9SF3_CCDS7450.1.ex9_10:98312701-98312724:-	TATGCTCGGTACAGTAAAG
	TM9SF3_CCDS7450.1.ex10_10:98319413-98319436:+	CCCGATGTTGAAAAAAGGA
	TM9SF3_CCDS7450.1.ex11_10:98321727-98321750:+	ACTTACTGAATATGACATC
	TM9SF3_CCDS7450.1.ex11_10:98321780-98321803:-	ATGTTAATCTAACTAGTGA
SLC29A1	SLC29A1_CCDS4908.1.ex3_6:44197649-44197672:-	GCCCAGGATCCGTACGGAC
	SLC29A1_CCDS4908.1.ex3_6:44197756-44197779:-	ATGAGCACGATCTTGATCA
	SLC29A1_CCDS4908.1.ex4_6:44198191-44198214:-	ATAGCGCAGATCATGGCCA
	SLC29A1_CCDS4908.1.ex5_6:44198299-44198322:-	TGATAGCTCCGAGCCACCT
	SLC29A1_CCDS4908.1.ex5_6:44198328-44198351:-	GCTGTGATAAAGTAGCCGA
TFRC	TFRC_CCDS3312.1.ex11_3:195794914-195794937:+	GAAAGTTCTGCGTTAACAA
	TFRC_CCDS3312.1.ex13_3:195798285-195798308:+	TATACGCCACATAACCCCC
	TFRC_CCDS3312.1.ex13_3:195798355-195798378:-	TTTAACAGCGCTCAAAACT
	TFRC_CCDS3312.1.ex14_3:195798984-195799007:-	CATATGTCCCTCGTGAGGC
	TFRC_CCDS3312.1.ex15_3:195800907-195800930:-	AGGAACCGAGTCTCCAGTG
PLXNB2	PLXNB2_CCDS43035.1.ex24_22:50722584-50722607:+	GCCGTAAGACTTGACGTAG
	PLXNB2_CCDS43035.1.ex25_22:50722995-50723018:+	CGTACCTTTGGGGTCCGAA
	PLXNB2_CCDS43035.1.ex31_22:50726374-50726397:+	CGACGCACCAGCCGCAGTA
	PLXNB2_CCDS43035.1.ex34_22:50727955-50727978:-	ACGGCGATATCCAGTGCGG
	PLXNB2_CCDS43035.1.ex34_22:50728329-50728352:+	ACGTAGGGGCCGTCCTCGA
FN1	FN1_CCDS2399.1.ex39_2:216289978-216290001:+	TGCACGAACATCGGTGAAG
	FN1_CCDS2399.1.ex40_2:216292990-216293013:+	CGATTATCCTTCTTGCTCC
	FN1_CCDS2399.1.ex41_2:216295518-216295541:-	TATGTGGTCGGAGAAACGT
	FN1_CCDS2399.1.ex43_2:216298079-216298102:-	TGGGACTGTACCTGCATCG
	FN1_CCDS2399.1.ex44_2:216299528-216299551:-	CAGCCGGTTGTTATGACAA
TMEM30A	TMEM30A_CCDS47453.1.ex4_6:75974987-75975010:+	AGATTTCACGTAACGACGA
	TMEM30A_CCDS47453.1.ex5_6:75994120-75994143:+	ATCTCGCGGATGTTGTTGG
	TMEM30A_CCDS47453.1.ex5_6:75994145-75994168:+	CAAAAATGCCAATGCCGAT
	TMEM30A_CCDS47453.1.ex5_6:75994200-75994223:+	TAGCACCGTGCCAGCCGTA
	TMEM30A_CCDS47453.1.ex5_6:75994260-75994283:-	ACTCGGAGACCGGATAACA
ITGAV	ITGAV_CCDS2292.1.ex2_2:187487123-187487146:+	GCATCTGTGAGGTCGAAAC
	ITGAV_CCDS2292.1.ex3_2:187490301-187490324:-	TATACATACGTGATCTACA
	ITGAV_CCDS2292.1.ex9_2:187503170-187503193:+	TACAATTTTACTGGCGAGC
	ITGAV_CCDS2292.1.ex10_2:187505637-187505660:-	TCCGAAATATGCAGCCATC
	ITGAV_CCDS2292.1.ex11_2:187506134-187506157:+	CACCTCTCTTCATGGATCG
ICAM1	ICAM1_CCDS12231.1.ex1_19:10385531-10385554:-	TATGCCCAACAACTTGGGC
	ICAM1_CCDS12231.1.ex1_19:10385552-10385575:+	GAGACCCCGTTGCCTAAAA

	ICAM1_CCDS12231.1.ex1_19:10385582-10385605:+	CTGCCTGGGAACAACCGGA
	ICAM1_CCDS12231.1.ex1_19:10385695-10385718:-	CTCAGTTACTCACAGTACA
	ICAM1_CCDS12231.1.ex2_19:10394325-10394348:+	GAGGTCACGACCACGGTGC
TMEM2	TMEM2_CCDS47979.1.ex15_9:74345025-74345048:+	ATGGAGTTGTTCCTATCGG
	TMEM2_CCDS47979.1.ex17_9:74347342-74347365:-	CCCGGAATATTGTGATCCA
	TMEM2_CCDS47979.1.ex19_9:74360029-74360052:+	CTATGGCAACAATCCGACC
	TMEM2_CCDS47979.1.ex19_9:74360232-74360255:-	ATCGTGGACGTTGTTGGCA
	TMEM2_CCDS47979.1.ex20_9:74361126-74361149:+	ATGACTATAGAATGCACGG
IGF2R	IGF2R_CCDS5273.1.ex6_6:160450658-160450681:+	ATTTGTTTGCCCGTCGGAG
	IGF2R_CCDS5273.1.ex10_6:160461705-160461728:-	TAGCGCTTCTTCCCGTCGG
	IGF2R_CCDS5273.1.ex15_6:160468321-160468344:-	GCGTCTCGATCACAGAGAA
	IGF2R_CCDS5273.1.ex16_6:160468883-160468906:-	CGGTCACTACGCATTCCAG
	IGF2R_CCDS5273.1.ex19_6:160477535-160477558:+	ACAACGACGGATACAGACC
COMT	COMT_CCDS13770.1.ex0_22:19950316-19950339:+	TGAACGTGGGCGACAAGAA
	COMT_CCDS13770.1.ex1_22:19951148-19951171:-	ACAGCTGAGTAGCCACAGT
	COMT_CCDS13770.1.ex1_22:19951211-19951234:-	CAGTCGGGGTTGATCTCGA
	COMT_CCDS13770.1.ex1_22:19951245-19951268:-	AGCGAAATCCACCATCCGC
	COMT_CCDS13770.1.ex1_22:19951260-19951283:+	TTCGCTGGCGTGAAGGACA
ASGR2	ASGR2_CCDS11088.1.ex2_17:7010359-7010382:-	GCGGAGAAGTACTGCCAGC
	ASGR2_CCDS11088.1.ex3_17:7010578-7010601:+	CCACGAAGCGCAGGTCCAC
	ASGR2_CCDS11088.1.ex5_17:7011777-7011800:+	ACCGTGGGTGCTGATTGCC
	ASGR2_CCDS11088.1.ex5_17:7011801-7011824:-	TCCTCGAGCACCCTGACGG
	ASGR2_CCDS11088.1.ex6_17:7012166-7012189:+	GAGCAGAGACGCTGTGCCA
LAMP1	LAMP1_CCDS41909.1.ex1_13:113960845-113960868:+	AACGGGACCGCGTGCATAA
	LAMP1_CCDS41909.1.ex1_13:113960891-113960914:+	TGAACTACGACACCAAGAG
	LAMP1_CCDS41909.1.ex2_13:113964041-113964064:+	TCGTGATTGCTTTTGGAAG
	LAMP1_CCDS41909.1.ex2_13:113964155-113964178:+	TCCCCAATGCGAGCTCCAA
	LAMP1_CCDS41909.1.ex3_13:113965096-113965119:-	CACGTTGTTCATGTGGACC
ANO6	ANO6_CCDS31782.1.ex4_12:45741833-45741856:+	GTAAAAGTACACGCACCAT
	ANO6_CCDS31782.1.ex4_12:45742016-45742039:-	CATCCGGTTCTTCTCAAAT
	ANO6_CCDS31782.1.ex6_12:45744468-45744491:-	ACCGTTCATTAGGGCAGCT
	ANO6_CCDS31782.1.ex7_12:45751138-45751161:-	CACTCCTACAACTGCGGCC
	ANO6_CCDS31782.1.ex10_12:45781956-45781979:+	GAGTTTTGGAAGCGACGCC
ALCAM	ALCAM_CCDS33810.1.ex1_3:105238960-105238983:-	GAGGTACGTCAAGTCGGCA
	ALCAM_CCDS33810.1.ex2_3:105243290-105243313:+	GTGTGCATGCTAGTAACTG
	ALCAM_CCDS33810.1.ex4_3:105252476-105252499:-	ACCATGTGATATTGCCATC
	ALCAM_CCDS33810.1.ex5_3:105253614-105253637:-	CCATAATATGTCACCGAGC
	ALCAM_CCDS33810.1.ex6_3:105258903-105258926:+	AATGGCAACCCTCCCCCAG
ABCC2	ABCC2_CCDS7484.1.ex2_10:101552071-101552094:-	GATTGGTATATCGAACAGC
	ABCC2_CCDS7484.1.ex3_10:101553372-101553395:-	TATCGAGAGAATCCAGAAT
	ABCC2_CCDS7484.1.ex5_10:101554194-101554217:+	GAGTAGCATTACCTACAGC
	ABCC2_CCDS7484.1.ex6_10:101556878-101556901:-	AGACATCCTCGAGTGTCAG

	ABCC2_CCDS7484.1.ex14_10:101571346-101571369:-	GCAACTCACTCTCGGACTG
CXADR	CXADR_CCDS33519.1.ex1_21:18919380-18919403:+	GATGATTGAAAAAGCCAAA
	CXADR_CCDS33519.1.ex1_21:18919395-18919418:-	AGATAGGCAGTTTCCCCTT
	CXADR_CCDS33519.1.ex1_21:18919418-18919441:-	GACTAAGCGTAAATTTGCA
	CXADR_CCDS33519.1.ex1_21:18919447-18919470:-	CTCGATGTCCAGCGGTCCC
	CXADR_CCDS33519.1.ex1_21:18919489-18919512:+	AATCAGAAGGTGGATCAAG
SCARB1	SCARB1_CCDS45008.1.ex7_12:125296469-125296492:-	GCTCTTCACGGTGTTCACG
	SCARB1_CCDS45008.1.ex8_12:125298861-125298884:+	ATGAAGGCACGTTCGCCGA
	SCARB1_CCDS45008.1.ex9_12:125299545-125299568:+	TAGTCGCTCTCCGAGCCGT
	SCARB1_CCDS45008.1.ex9_12:125299593-125299616:+	CGGTACTCGAGGAAGGACA
	SCARB1_CCDS45008.1.ex11_12:125348148-125348171:-	CCGTCGCTCATCAAGCAGC
NT5E	NT5E_CCDS5002.1.ex0_6:86160126-86160149:+	ATCTGGTTCACCGTGTACA
	NT5E_CCDS5002.1.ex0_6:86160177-86160200:-	TACCATGGCATCGTAGCGC
	NT5E_CCDS5002.1.ex1_6:86176882-86176905:-	CTGATATTTGAGATGCTAG
	NT5E_CCDS5002.1.ex2_6:86181071-86181094:+	ACTCATCGCTCAGAAAGTG
	NT5E_CCDS5002.1.ex3_6:86194975-86194998:-	TGAATGGGTACTTCCCAGC
ABCB1	ABCB1_CCDS5608.1.ex15_7:87179269-87179292:+	TCGTGGTGGCAAACAATAC
	ABCB1_CCDS5608.1.ex17_7:87179861-87179884:-	TGACAGCTATTCGAAGAGT
	ABCB1_CCDS5608.1.ex20_7:87190653-87190676:-	TCTTAGCGTATGCAAAAGC
	ABCB1_CCDS5608.1.ex22_7:87196148-87196171:-	CTATAATGCGACAGGAGAT
	ABCB1_CCDS5608.1.ex22_7:87196271-87196294:-	GGTATGCCTATTATTACAG
ENPP1	ENPP1_CCDS5150.2.ex0_6:132129318-132129341:+	TTGCTGGCCCCTATGGACG
	ENPP1_CCDS5150.2.ex0_6:132129393-132129416:+	ACCTATAAAGTACTCTCGC
	ENPP1_CCDS5150.2.ex2_6:132171229-132171252:+	ACGTGCATAGAACCAGGTA
	ENPP1_CCDS5150.2.ex5_6:132176060-132176083:-	TAGGAGGCGTTTCAAACCT
	ENPP1_CCDS5150.2.ex8_6:132182749-132182772:+	AGTATCAAGGCCTCAAGTC
ATP2B1	ATP2B1_CCDS41817.1.ex11_12:90013979-90014002:-	GTGGATTACCTCGTCACGT
	ATP2B1_CCDS41817.1.ex13_12:90017975-90017998:+	GTGAGATCGTGACTGCAAG
	ATP2B1_CCDS41817.1.ex15_12:90021453-90021476:-	CTATTGAGAATCGCAACAA
	ATP2B1_CCDS41817.1.ex17_12:90028646-90028669:+	GAATAAGTATGCCGTCAGC
	ATP2B1_CCDS41817.1.ex20_12:90049512-90049535:-	GATGCATTACGAAAAATAC
EGFR	EGFR_CCDS47587.1.ex2_7:55211103-55211126:-	GATAAGACTGCTAAGGCAT
	EGFR_CCDS47587.1.ex2_7:55211134-55211157:+	GCAAATAAAACCGGACTGA
	EGFR_CCDS47587.1.ex3_7:55214329-55214352:+	AACCCTGCCCTGTGCAACG
	EGFR_CCDS47587.1.ex3_7:55214359-55214382:-	GCTGACTATGTCCCGCCAC
	EGFR_CCDS47587.1.ex4_7:55218995-55219018:+	TGATCCAAGCTGTCCCAAT
M6PR	M6PR_CCDS8598.1.ex3_12:9096395-9096418:-	TGCAATCGACACACCCTAG
	M6PR_CCDS8598.1.ex3_12:9096446-9096469:-	TATGACAACCACTGTGGCA
	M6PR_CCDS8598.1.ex3_12:9096477-9096500:-	GATCATGCTGATCTATAAA
	M6PR_CCDS8598.1.ex4_12:9098058-9098081:-	ATCAACAAAAGTAATGGGA
	M6PR_CCDS8598.1.ex5_12:9098884-9098907:-	GCGACTTGGTAGGAGAAAA
LSR	LSR_CCDS12449.1.ex0_19:35739879-35739902:+	TACTTTGGAAGGGACGCGC

	LSR_CCDS12449.1.ex1_19:35741365-35741388:-	GGAGAAGGCATCGGCGATG
	LSR_CCDS12449.1.ex1_19:35741435-35741458:-	CAACGTAGGGGTTGTAGCC
	LSR_CCDS12449.1.ex1_19:35741537-35741560:+	GCCGGAGGATTACCATCAC
	LSR_CCDS12449.1.ex2_19:35749858-35749881:+	CCTTTGACCAGACGGCGTG
BCAM	BCAM_CCDS12644.1.ex1_19:45314500-45314523:+	TTGTCTGTACCCCCGCTGG
	BCAM_CCDS12644.1.ex1_19:45314560-45314583:-	ATGGTCGTGGGTTCCCGTA
	BCAM_CCDS12644.1.ex3_19:45315798-45315821:-	CCACCCGAGAGGTACCTCC
	BCAM_CCDS12644.1.ex4_19:45316556-45316579:+	TATCGCAACGGGCAGCGCC
	BCAM_CCDS12644.1.ex5_19:45316695-45316718:+	GGCTACATGACCAGCCGCA
CADM1	CADM1_CCDS53711.1.ex4_11:115099979-115100002:-	TGTTCCCCAGGCAAATCGG
	CADM1_CCDS53711.1.ex6_11:115109219-115109242:-	TACACCACCATCACAGTCC
	CADM1_CCDS53711.1.ex6_11:115109246-115109269:-	CTCTATACCGATCCCCCAC
CD276	CD276_CCDS10251.1.ex1_15:73994623-73994646:-	GCCCACCAGTGCCACCACT
	CD276_CCDS10251.1.ex1_15:73994694-73994717:+	GGCACAGCTCAACCTCATC
	CD276_CCDS10251.1.ex1_15:73994773-73994796:+	AGCGCCTATGCCAACCGCA
	CD276_CCDS10251.1.ex1_15:73994845-73994868:+	CGCGTGCGTGTGGCGGACG
	CD276_CCDS10251.1.ex1_15:73994877-73994900:-	TCCCGGATGCTCACGAAGC
DSG2	DSG2_CCDS42423.1.ex2_18:29099814-29099837:+	AGTGCGGCAAAAGCGCGCC
	DSG2_CCDS42423.1.ex2_18:29099888-29099911:-	TAGGAGGTACCTTGGCAAT
	DSG2_CCDS42423.1.ex5_18:29102143-29102166:-	TATTTAGGTAGAACACTGG
	DSG2_CCDS42423.1.ex7_18:29104655-29104678:+	GTTTTGCAGCTTGAAGGGA
	DSG2_CCDS42423.1.ex8_18:29111116-29111139:+	ATTTATGTTAGCGAGAGCA
MAN2A1	MAN2A1_CCDS34209.1.ex1_5:109049252-109049275:-	AGCTAGCAAACGCTCCAAA
	MAN2A1_CCDS34209.1.ex2_5:109051943-109051966:+	CTCATTCCCATAACGACCC
	MAN2A1_CCDS34209.1.ex5_5:109103236-109103259:+	GTGAAACCTCGGTCCGGCT
	MAN2A1_CCDS34209.1.ex7_5:109110574-109110597:+	CTACTGTGAATACACGGAA
	MAN2A1_CCDS34209.1.ex8_5:109117251-109117274:+	AGACCCTTTTACAAACGAA
DAG1	DAG1_CCDS2799.1.ex1_3:49568485-49568508:-	ACAGGTTCATCCGCAGCAC
	DAG1_CCDS2799.1.ex1_3:49568856-49568879:-	TCCGGACGCGTTTGGGAAG
	DAG1_CCDS2799.1.ex1_3:49569053-49569076:-	AATAATGGCGCCTCGAGTC
	DAG1_CCDS2799.1.ex1_3:49569140-49569163:-	GGTCATCGTTGGGCGAATC
	DAG1_CCDS2799.1.ex1_3:49569222-49569245:-	GTTTTGGTGTGGATACTCG
ITGA6	ITGA6_CCDS2249.1.ex0_2:173292606-173292629:+	ACAACGTGATCCGGAAATA
	ITGA6_CCDS2249.1.ex1_2:173330360-173330383:-	CAAACTCGATCCGCGTGCA
	ITGA6_CCDS2249.1.ex3_2:173333878-173333901:+	CAGCATGTTAATACGAAGC
	ITGA6_CCDS2249.1.ex3_2:173334081-173334104:+	AGCCCCGGGTACTTATAAC
	ITGA6_CCDS2249.1.ex10_2:173344739-173344762:+	AATATACTGCTAACCCCGC
EPHA2	EPHA2_CCDS169.1.ex13_1:16464769-16464792:+	TGTGCAAGGCATCGACGCT
	EPHA2_CCDS169.1.ex14_1:16475165-16475188:-	TGGGGCCGCTCACCCGCAA
	EPHA2_CCDS169.1.ex14_1:16475242-16475265:+	GTGATCTCATCGGGCGCAA
	EPHA2_CCDS169.1.ex14_1:16475310-16475333:-	TACTATGCCGAGTCGGACC
BSG	BSG_CCDS12033.1.ex1_19:577981-578004:+	ATCTCCATCGACACGCTCG

	BSG_CCDS12033.1.ex2_19:579510-579533:+	TCACTACCGTAGAAGACCT
	BSG_CCDS12032.1.ex1_19:580378-580401:+	GTGGACTCCGACGACCAGT
	BSG_CCDS12032.1.ex2_19:580670-580693:+	TCGTCAGAACACATCAACG
	BSG_CCDS12032.1.ex2_19:580732-580755:+	ACCTGTCACTGACTGGGCC
SLC3A2	SLC3A2_CCDS31588.1.ex3_11:62648592-62648615:-	TTCTTCTCGGCTCCCGCCA
	SLC3A2_CCDS31588.1.ex3_11:62648661-62648684:-	AGGCCCGTGAACTTAGCCG
	SLC3A2_CCDS31588.1.ex3_11:62648790-62648813:-	GCTCGCACGATTATGACCA
	SLC3A2_CCDS31588.1.ex3_11:62648812-62648835:+	CCGCGTTGTCGCGAGCTAC
	SLC3A2_CCDS31588.1.ex3_11:62648852-62648875:+	CGGGCGCCCTCTACCGCAT
F11R	F11R_CCDS1213.1.ex5_1:160970106-160970129:+	GGGATGTTAACTGTAGGCT
	F11R_CCDS1213.1.ex6_1:160970494-160970517:-	CCGTGACACGGGAAGACAC
	F11R_CCDS1213.1.ex6_1:160970515-160970538:+	ACTTGAAGGTGATACCAGT
	F11R_CCDS1213.1.ex6_1:160970530-160970553:-	GGGTGACCTTCTTGCCAAC
	F11R_CCDS1213.1.ex7_1:160970869-160970892:-	GGCTTTTCTTCTCCCCGTG
LAMP2	LAMP2_CCDS14599.1.ex5_X:119582841-119582864:-	TACAAGCTTTTGTCCAAAA
	LAMP2_CCDS14599.1.ex5_X:119582873-119582896:-	TGATGTTGTCCAACACTAC
	LAMP2_CCDS14599.1.ex6_X:119589306-119589329:+	TCGCAATCCAGGAAAAGCC
	LAMP2_CCDS14599.1.ex6_X:119589324-119589347:-	TAGCAGTGCAGTTCGGACC
	LAMP2_CCDS14599.1.ex7_X:119590550-119590573:-	CACTTGCCTTTATGCAAAA
CD59	CD59_CCDS7886.1.ex0_11:33731817-33731840:-	GACGTCACAACCCGCTTGA
	CD59_CCDS7886.1.ex0_11:33731872-33731895:+	TTATACACTTGTAACCCTG
	CD59_CCDS7886.1.ex1_11:33738914-33738937:-	CGTGTCTCATTACCAAAGC
	CD59_CCDS7886.1.ex1_11:33738940-33738963:+	AAATCAGATGAACAATTGA
	CD59_CCDS7886.1.ex1_11:33738962-33738985:+	CTGTTTTGCAGTCAGCAGT
ECE1	ECE1_CCDS215.1.ex10_1:21573782-21573805:+	CGGATTCATTGATCTCCAC
	ECE1_CCDS215.1.ex11_1:21582603-21582626:-	ACCGGATATCTGAACTACA
	ECE1_CCDS215.1.ex13_1:21585309-21585332:-	TCGGGGGCTGGAACATCAC
	ECE1_CCDS215.1.ex14_1:21586808-21586831:-	TGCGTGCATGAACGAGACC
	ECE1_CCDS215.1.ex15_1:21599176-21599199:+	TGCTGGATGCTTACCGAGG
ADAM10	ADAM10_CCDS10167.1.ex11_15:58957368-58957391:-	CCCATAAATACGGTCCTCA
	ADAM10_CCDS10167.1.ex12_15:58971400-58971423:+	AAATGTGCCACCACGAGTC
	ADAM10_CCDS10167.1.ex13_15:58974405-58974428:-	ATACCTCTCATATTTACAC
	ADAM10_CCDS10167.1.ex14_15:59009778-59009801:-	GTCTAGATTTCCATGCCCA
	ADAM10_CCDS10167.1.ex14_15:59009863-59009886:-	GAAGGATTATCTTACAATG
CPD	CPD_CCDS11257.1.ex1_17:28712218-28712241:+	AATCACAAACGGCGCACAT
	CPD_CCDS11257.1.ex4_17:28749710-28749733:+	CTAATGTAGTGGTGAAAGA
	CPD_CCDS11257.1.ex4_17:28749859-28749882:+	TACCAGCCAATTCAGCCAA
	CPD_CCDS11257.1.ex5_17:28750687-28750710:-	GGGACTTTTCATACCCATC
	CPD_CCDS11257.1.ex6_17:28754510-28754533:+	AACTATTGCTGTAATGAGC
CDH2	CDH2_CCDS11891.1.ex7_18:25572636-25572659:-	CGACCCAAACAGCAACGAC
	CDH2_CCDS11891.1.ex7_18:25572737-25572760:-	AATCTAACTGTGACCGATA
	CDH2_CCDS11891.1.ex10_18:25585837-25585860:-	CTTACACCAGGTTTGGAAT

	CDH2_CCDS11891.1.ex11_18:25589786-25589809:-	CACTGCGGTACAGTGTAAC
	CDH2_CCDS11891.1.ex14_18:25727666-25727689:-	TACAGTGCAGTCTTATCGA
SDC1	SDC1_CCDS1697.1.ex1_2:20402944-20402967:+	ACAGCCGTATTCTCCCCCG
	SDC1_CCDS1697.1.ex1_2:20402958-20402981:-	TTCACCTTTGAAACCTCGG
	SDC1_CCDS1697.1.ex2_2:20403694-20403717:-	ATGAGACCTCAACCCCTGC
	SDC1_CCDS1697.1.ex2_2:20403968-20403991:+	AGACGTGGGAATAGCCGTC
	SDC1_CCDS1697.1.ex2_2:20403994-20404017:+	GCGTGTCCTTCCAAGTGGA
ITGA1	ITGA1_CCDS3955.1.ex6_5:52177754-52177777:+	AAGTATTCTTCCACCGAAG
	ITGA1_CCDS3955.1.ex7_5:52183662-52183685:+	AAGCCCGGGGTGCCCGAAG
	ITGA1_CCDS3955.1.ex8_5:52189505-52189528:+	TTCTTGGCAGCTATAACCG
	ITGA1_CCDS3955.1.ex11_5:52201651-52201674:-	GGCCTGTATGATTGTACCG
	ITGA1_CCDS3955.1.ex12_5:52204798-52204821:+	GTCGGAGCCCCTATGTACA
CD99	CD99_CCDS14119.1.ex4_Y:2588451-2588474:-	ACTCTTACCGGAGGAACTA
	CD99_CCDS14119.1.ex5_Y:2590686-2590709:-	ACCTGAAACGCCATCCGCA
	CD99_CCDS14119.1.ex7_Y:2594308-2594331:-	CAATCCCGGGGATCACGCC
	CD99_CCDS14119.1.ex7_Y:2594335-2594358:+	CTGTCGTGGTCGCCGTGGC
	CD99_CCDS14119.1.ex8_Y:2606268-2606291:-	CTCTGCGTTGGCATTCCGG
BST2	BST2_CCDS12358.1.ex3_19:17516086-17516109:+	TGAGGAGCTTACCACAGTG
	BST2_CCDS12358.1.ex3_19:17516099-17516122:+	ACAGTGTGGTTGCAGGTGG
	BST2_CCDS12358.1.ex3_19:17516155-17516178:+	CTCGGTCAGCTCTTGTTGC
	BST2_CCDS12358.1.ex3_19:17516194-17516217:+	ACACTCCATCACTGCCCGA
	BST2_CCDS12358.1.ex3_19:17516207-17516230:+	GCCCGAAGGCCGTCCCGGC
OCLN	OCLN_CCDS4006.1.ex1_5:68805100-68805123:-	ACAGGATCCGAATCACTCC
	OCLN_CCDS4006.1.ex1_5:68805244-68805267:+	GCTTTGGTAGCTACGGAAG
	OCLN_CCDS4006.1.ex1_5:68805354-68805377:+	TTTTGTTTCATTGCCGCGT
	OCLN_CCDS4006.1.ex2_5:68809914-68809937:+	CCCCCCAATGTCGAGGAGT
	OCLN_CCDS4006.1.ex3_5:68830525-68830548:+	AATGTGTCTGCAGGCACAC
CEACAM1	CEACAM1_CCDS12609.1.ex4_19:43023234-43023257:+	CTCCGAGGACGGGAGACTC
	CEACAM1_CCDS12609.1.ex4_19:43023319-43023342:+	TCCTTATCTCCTGTGACTG
	CEACAM1_CCDS12609.1.ex4_19:43023356-43023379:+	TTTGGGGCTTTGCTACTAC
	CEACAM1_CCDS12609.1.ex4_19:43023370-43023393:+	ACTACTGGACTTAGCTCTG
	CEACAM1_CCDS12609.1.ex5_19:43025545-43025568:+	AATGTTCCATTGATAAGCC
HEPACAM	HEPACAM_CCDS8456.1.ex4_11:124793723-124793746:+	GGTGAGCACCTTTTGGTCG
	HEPACAM_CCDS8456.1.ex5_11:124794614-124794637:+	CGCTTTACCATCTACAGTA
	HEPACAM_CCDS8456.1.ex5_11:124794650-124794673:-	ACCGACGACACCTTCACTG
	HEPACAM_CCDS8456.1.ex5_11:124794759-124794782:+	TCTCGATAGTCAGGCCGCA
	HEPACAM_CCDS8456.1.ex5_11:124794913-124794936:+	TGCCATGGATCAGGCGCAC
DSG1	DSG1_CCDS11896.1.ex2_18:28906956-28906979:-	TGATACCTACTTTGGCGAT
	DSG1_CCDS11896.1.ex3_18:28908267-28908290:+	ACATCCATAGTTGATCGAG
	DSG1_CCDS11896.1.ex6_18:28913575-28913598:+	TAAGAGGCTCTGACCGAGA
	DSG1_CCDS11896.1.ex8_18:28916457-28916480:-	TTGAACCTGGACGAAACAC
	DSG1_CCDS11896.1.ex9_18:28918362-28918385:+	AACAGTACAATATGCTCGG

LRP1	LRP1_CCDS8932.1.ex11_12:57553652-57553675:+	TCTGTACTGGACGGACGAT
	LRP1_CCDS8932.1.ex12_12:57554848-57554871:-	GTCTCGATGCGGTCGTAGA
	LRP1_CCDS8932.1.ex16_12:57559881-57559904:-	TCGCACTTGAATCGGTCCG
	LRP1_CCDS8932.1.ex37_12:57578139-57578162:-	GCCGAGACCGCTCAATACG
	LRP1_CCDS8932.1.ex39_12:57578871-57578894:-	CCGCGCTTGATAGAGCCGT
ATP2B4	ATP2B4_CCDS1440.1.ex1_1:203667303-203667326:+	GATCTGGAGAAACGTAGGC
	ATP2B4_CCDS1440.1.ex2_1:203668767-203668790:-	AGTTGACCGTTTCGGATGA
	ATP2B4_CCDS1440.1.ex8_1:203677003-203677026:+	AATAACCTAGTACGGCACT
	ATP2B4_CCDS1440.1.ex8_1:203677129-203677152:-	GCTTGGGATTTGACGGTAA
	ATP2B4_CCDS1440.1.ex9_1:203678615-203678638:-	TACATACGGAAGCCACCGT
SLC2A2	SLC2A2_CCDS3215.1.ex4_3:170723107-170723130:+	GCATCAGTGCCACTAGAAT
	SLC2A2_CCDS3215.1.ex4_3:170723127-170723150:+	GGCTGTCGGTAGCTGGAAT
	SLC2A2_CCDS3215.1.ex5_3:170723749-170723772:-	CTTTACATCAAGTTAGATG
	SLC2A2_CCDS3215.1.ex6_3:170724953-170724976:-	CATCAGCTGGCCATCGTCA
	SLC2A2_CCDS3215.1.ex6_3:170724991-170725014:-	TTGCTCCAACCGCTCTCAG
NCSTN	NCSTN_CCDS1203.1.ex2_1:160318816-160318839:-	CTCCTCTTTCTCTACTACG
	NCSTN_CCDS1203.1.ex5_1:160321144-160321167:+	CTTCAGCATCAACCCAGGT
	NCSTN_CCDS1203.1.ex6_1:160321501-160321524:-	CCACACATTGTAATCAGAC
	NCSTN_CCDS1203.1.ex6_1:160321551-160321574:+	GACATTAAAGCCTGACGAC
	NCSTN_CCDS1203.1.ex8_1:160322699-160322722:+	TCGAGGATGGTCTACGATA
VASN	VASN_CCDS10514.1.ex0_16:4431022-4431045:+	ACGTGCCACCCGACACGGT
	VASN_CCDS10514.1.ex0_16:4431309-4431332:+	GACACGCTCGACCGCCTCC
	VASN_CCDS10514.1.ex0_16:4432064-4432087:-	ACAGGCCCTACAGTCGGTG
	VASN_CCDS10514.1.ex0_16:4432158-4432181:-	TTCGGGGCACAAGCACGCC
	VASN_CCDS10514.1.ex0_16:4432192-4432215:+	TGTACTGTGAGAGCCAGAT
ICOSLG	ICOSLG_CCDS42952.1.ex3_21:45655204-45655227:+	AGCAGCCAATGTTCACGCT
	ICOSLG_CCDS42952.1.ex3_21:45655230-45655253:-	CAGCGTGCTGAGGATCGCA
	ICOSLG_CCDS42952.1.ex3_21:45655253-45655276:-	ATGCGGGGCTTGTATGACG
	ICOSLG_CCDS42952.1.ex3_21:45655329-45655352:+	TTATTGATCCAGTACACGT
	ICOSLG_CCDS42952.1.ex3_21:45655353-45655376:-	ATCCATAAACGGCTACCCC
ROBO1	ROBO1_CCDS46872.2.ex19_3:78719414-78719437:-	CTCCCCCAGTTATTCGACA
	ROBO1_CCDS46872.2.ex24_3:78766454-78766477:+	TCCATCGTACTGTAGGTAC
	ROBO1_CCDS46872.2.ex25_3:78766983-78767006:-	ACACCCGTAAAAGTGACGC
	ROBO1_CCDS46872.2.ex27_3:78987841-78987864:-	CTTACGTATAGTACATGGA
	ROBO1_CCDS46872.2.ex27_3:78987875-78987898:-	ACCGAATGTTGCTGCCGAG
SLC11A2	SLC11A2_CCDS53791.1.ex7_12:51390677-51390700:+	GGAGTGCGACAGCCTGAAC
	SLC11A2_CCDS53791.1.ex8_12:51393008-51393031:+	CTAGCTTCCGCAAGCCTAA
	SLC11A2_CCDS53791.1.ex11_12:51398622-51398645:-	AGCGGCTTGCAGCTAGACT
	SLC11A2_CCDS53791.1.ex12_12:51399112-51399135:-	TTGAATCCGATTTGCAGTC
	SLC11A2_CCDS53791.1.ex12_12:51399172-51399195:-	TCTGGGCTTTCACCGGACC
SDC4	SDC4_CCDS13350.1.ex1_20:43959080-43959103:-	TCACCCGTTGAAGAGAGTG
	SDC4_CCDS13350.1.ex1_20:43959128-43959151:-	ACCGAACCCAAGAAACTAG

	SDC4_CCDS13350.1.ex1_20:43959185-43959208:+	ATGGTTATCTAGAGGCACC
	SDC4_CCDS13350.1.ex3_20:43964505-43964528:+	GAAGTATCGGCCTTCTAGG
	SDC4_CCDS13350.1.ex3_20:43964529-43964552:-	ACTGAGGTCATCGACCCCC
MPZL1	MPZL1_CCDS1264.1.ex1_1:167734837-167734860:-	TTTGGCGTATATACTTCCA
	MPZL1_CCDS1264.1.ex1_1:167734850-167734873:+	ACGCCAAAAGAAATCTTCG
	MPZL1_CCDS1264.1.ex1_1:167734867-167734890:+	CGTGGCAAATGGTACACAA
	MPZL1_CCDS1264.1.ex1_1:167734905-167734928:+	TCAAGTCTACTAGTACGAC
	MPZL1_CCDS1264.1.ex1_1:167734955-167734978:-	AGTGTCGGCCCCCTCTGGC
PVR	PVR_CCDS12640.1.ex1_19:45150601-45150624:+	AGCTGACTTGGGCGCGGCA
	PVR_CCDS12640.1.ex1_19:45150665-45150688:-	CGTTTGGACTCCGAATAGC
	PVR_CCDS12640.1.ex1_19:45150691-45150714:+	AATTCGTGGCAGCCAGACT
	PVR_CCDS12640.1.ex1_19:45150748-45150771:+	GGTTGCGCGTAGAGGATGA
	PVR_CCDS12640.1.ex1_19:45150783-45150806:-	CTGCGGGAACGTGACGAAC
SLC1A5	SLC1A5_CCDS12692.1.ex3_19:47281959-47281982:+	CGGCGTCACGATGCCCCAC
	SLC1A5_CCDS12692.1.ex3_19:47282040-47282063:+	ATGGATGGCGTGACCCAGC
	SLC1A5_CCDS12692.1.ex3_19:47282068-47282091:+	AGAATGTACTTGCCAAGGC
	SLC1A5_CCDS12692.1.ex3_19:47282081-47282104:-	GTTTACTCTTTGCCCGCCT
	SLC1A5_CCDS12692.1.ex3_19:47282136-47282159:+	CAGGAACATGATGCCCACA
DPEP1	DPEP1_CCDS10982.1.ex2_16:89702689-89702712:-	CTTTGTTCTGGGTGTCGCA
	DPEP1_CCDS10982.1.ex2_16:89702748-89702771:-	GTACATCCGGCACATGCGG
	DPEP1_CCDS10982.1.ex3_16:89702942-89702965:+	TTCGGCAGGCCTTCCGGGA
	DPEP1_CCDS10982.1.ex3_16:89703026-89703049:+	TGCGGGCACTCTATCAGCT
	DPEP1_CCDS10982.1.ex4_16:89703321-89703344:+	GGCTTGTCACCCTTTGGGC
ATP1B3	ATP1B3_CCDS3121.1.ex1_3:141622568-141622591:+	GTGACCAGATTCCTAGCCC
	ATP1B3_CCDS3121.1.ex2_3:141625994-141626017:+	TATGTCTTTATAGGACTCA
	ATP1B3_CCDS3121.1.ex2_3:141626028-141626051:-	TATATTCCAATGCGGTCAC
	ATP1B3_CCDS3121.1.ex2_3:141626062-141626085:+	TGATCCAACTTCGTATGCA
	ATP1B3_CCDS3121.1.ex3_3:141632561-141632584:-	ACTGACATGCAACATAAAC
SLC16A3	SLC16A3_CCDS11804.1.ex1_17:80194641-80194664:-	CCCCACAAGCATGACGGGC
	SLC16A3_CCDS11804.1.ex1_17:80194715-80194738:+	CCAGGTCTACCTCACCACT
	SLC16A3_CCDS11804.1.ex2_17:80195065-80195088:+	TTCAGCAAGCGGCGCCCCA
	SLC16A3_CCDS11804.1.ex2_17:80195320-80195343:-	GGCGTAAAGCACAAAGCCG
	SLC16A3_CCDS11804.1.ex2_17:80195461-80195484:+	ATTGACATCTTCGCGCGGC

Table S 1 Pool 1: list of genes and their 500 sgRNA sequences.

There were 50 non-targeting sgRNAs and 5 CD81-targeting sgRNAs were used as a negative control. *SCARB1*, a positive control, is ranked 41 in proteomic abundance. Genes are listed in an order of proteomic abundance.

Pool 2		
Gene	sgRNA name	sgRNA sequence (19 bp)
Non-targeting 1	nontargeting-1	GTCCATGGGTGGAGTTACG
Non-targeting 2	nontargeting-2	GGACAAGTTGGACAGTTGC
Non-targeting 3	nontargeting-3	AACAATGGCTTCGTCGACT
Non-targeting 4	nontargeting-4	ACCTTATGTTCGCCGCCTC
Non-targeting 5	nontargeting-5	AGACACCGTAACTCGAATT
Non-targeting 6	nontargeting-6	CTCGTAGGTCAATCGCGGA
Non-targeting 7	nontargeting-7	GACGGTTCCGATAACCTTA
Non-targeting 8	nontargeting-8	GGATCTCTCAATATGGCGC
Non-targeting 9	nontargeting-9	AGCGCCGTTGTTGAGAAGA
Non-targeting 10	nontargeting-10	CCGCCGGCCCTAACTGACC
Non-targeting 11	nontargeting-11	TGACTAGACCCTTACGCGG
Non-targeting 12	nontargeting-12	ATCGGAGAGCCTGAAGTTG
Non-targeting 13	nontargeting-13	TCGGGCCGCTATATAGGCG
Non-targeting 14	nontargeting-14	AGAACTCAGCGGTCTCAAG
Non-targeting 15	nontargeting-15	ATGTCCGTTGTAGTCCTCG
Non-targeting 16	nontargeting-16	CCGGTCGCAATTGTTGCTA
Non-targeting 17	nontargeting-17	TGCCTATTCAGTAGGTCGT
Non-targeting 18	nontargeting-18	GTGCGGTAAGCGTGGCTAG
Non-targeting 19	nontargeting-19	CCATGGAGCCCAATCAATC
Non-targeting 20	nontargeting-20	GTGTAGTTCGGATTGATGA
Non-targeting 21	nontargeting-21	GGAATTGGAGACGATGCCG
Non-targeting 22	nontargeting-22	AGTGCCTCGTAGTTGTGGG
Non-targeting 23	nontargeting-23	CCTCGTTTCCACGTAATTC
Non-targeting 24	nontargeting-24	ATTTGATGTGGTGCTCACA
Non-targeting 25	nontargeting-25	ATTAGGTCGTTCTGTAACA
SCARB1	SCARB1_CCDS45008.1.ex7_12:125296469-125296492:-	GCTCTTCACGGTGTTCACG
	SCARB1_CCDS45008.1.ex8_12:125298861-125298884:+	ATGAAGGCACGTTCGCCGA
	SCARB1_CCDS45008.1.ex9_12:125299545-125299568:+	TAGTCGCTCTCCGAGCCGT
	SCARB1_CCDS45008.1.ex9_12:125299593-125299616:+	CGGTACTCGAGGAAGGACA

	SCARB1_CCDS45008.1.ex11_12:125348148-125348171:-	CCGTCGCTCATCAAGCAGC
CD81	CD81_CCDS7734.1.ex1_11:2411694-2411717:+	ACCACCAACCTCCTGTATC
	CD81_CCDS7734.1.ex1_11:2411744-2411767:-	GTGCACTCACCTACATAGA
	CD81_CCDS7734.1.ex2_11:2415376-2415399:+	GGCTGCTACGGGGCCATCC
	CD81_CCDS7734.1.ex3_11:2416244-2416267:-	TTGACAAAGCCCCAGATGC
	CD81_CCDS7734.1.ex4_11:2416641-2416664:-	CTTCACATCCTTGGCGATC
MTDH	MTDH_CCDS6274.1.ex1_8:98673301-98673324:+	AATGGGCGGACTGTTGAAG
	MTDH_CCDS6274.1.ex3_8:98699733-98699756:+	CATTGGATTCAACTATCCC
CADM1	CADM1_CCDS53711.1.ex7_11:115111021-115111044:-	TCAGCTACTGAATCCCAAC
	CADM1_CCDS53711.1.ex7_11:115111095-115111118:-	AAAGACGTGACAGTGATCG
EPHA2	EPHA2_CCDS169.1.ex10_1:16461646-16461669:-	ACAATGTGCGCCGCACCGA
SLC46A1	SLC46A1.1	CCGCCGCGCACGCACATGG
	SLC46A1.2	CGCCGCGCACGCACATGGA
	SLC46A1.3	AGCACGGCAGCCGCAGGGC
	SLC46A1.4	CAGCACGGCAGCCGCAGGG
	SLC46A1.5	GCACAGCACGGCAGCCGCA
MGAT1	MGAT1_CCDS4458.1.ex0_5:180219367-180219390:+	GCGAAACTGCCGGAAGACC
	MGAT1_CCDS4458.1.ex0_5:180219456-180219479:+	CCGCAATGCTGCTCAGGTC
	MGAT1_CCDS4458.1.ex0_5:180219647-180219670:+	ATGGGAATCACCGCCGGCG
	MGAT1_CCDS4458.1.ex0_5:180219802-180219825:+	GCGAATCACTTCCCGGGTG
	MGAT1_CCDS4458.1.ex0_5:180219836-180219859:+	CCATCGAGAGCGCTGACTG
SLC35A2	SLC35A2_CCDS14311.1.ex2_X:48763663-48763686:+	TCACCTGGAAAGTGGCAGC
	SLC35A2_CCDS14311.1.ex2_X:48763704-48763727:+	TACTGGAGGTTATTCTGCA
	SLC35A2_CCDS14311.1.ex2_X:48763719-48763742:+	TGCAAGGTGTAGATGAGAG
	SLC35A2_CCDS14311.1.ex2_X:48763757-48763780:+	CGTGTCCACATACTGCACC
	SLC35A2_CCDS14311.1.ex2_X:48763802-48763825:-	CTAGGTAACGTGAAGCACC
B4GALT7	B4GALT7_CCDS4429.1.ex1_5:177031426-177031449:-	CCTCGAAGCGTTCGCGGAA
	B4GALT7_CCDS4429.1.ex1_5:177031500-177031523:-	GAGCACGTAGATGTGGTGC
	B4GALT7_CCDS4429.1.ex1_5:177031524-177031547:-	CTACCTGAAGTGGTCCACC
	B4GALT7_CCDS4429.1.ex2_5:177034290-177034313:+	ACCCTGCACAGGTTCAACC
	B4GALT7_CCDS4429.1.ex2_5:177034309-177034332:+	GGGCAGCGCTCATCAACGT
PGAP1	PGAP1_CCDS2318.1.ex14_2:197744797-197744820:+	CCTTACTTCCACGAACAGA

	PGAP1_CCDS2318.1.ex22_2:197767361-197767384:+	TGAACGAACTTGGTAATCC
	PGAP1_CCDS2318.1.ex23_2:197777613-197777636:+	AACGATCTAATGGCATCAC
	PGAP1_CCDS2318.1.ex23_2:197777724-197777747:-	TAATTGGTCATTCTATGGG
	PGAP1_CCDS2318.1.ex25_2:197784755-197784778:+	GAACTGGAATACCCGTCAA
APOA1	APOA1_CCDS8378.1.ex0_11:116706938-116706961:+	TCTGGAAGTCGTCCAGGTA
	APOA1_CCDS8378.1.ex0_11:116707104-116707127:-	CTAAAGCTCCTTGACAACT
	APOA1_CCDS8378.1.ex1_11:116707715-116707738:+	TTTAGCTGTTTTCCCAAGG
	APOA1_CCDS8378.1.ex1_11:116707729-116707752:+	CAAGGCGGAGCCTTCAAAC
	APOA1_CCDS8378.1.ex1_11:116707786-116707809:-	GACCTGGCCACTGTGTACG
APOE	APOE_CCDS12647.1.ex1_19:45411121-45411144:+	CTTTTGGGATTACCTGCGC
	APOE_CCDS12647.1.ex2_19:45411822-45411845:+	GCCTACAAATCGGAACTGG
	APOE_CCDS12647.1.ex2_19:45411948-45411971:+	CTGGTGCAGTACCGCGGCG
	APOE_CCDS12647.1.ex2_19:45412080-45412103:-	CCCGGCCTGGTACACTGCC
	APOE_CCDS12647.1.ex2_19:45412170-45412193:-	CACAGTGGCGGCCCGCACG
PPARG	PPARG_CCDS2609.1.ex2_3:12422907-12422930:-	CGACATTCAATTGCCATGA
	PPARG_CCDS2609.1.ex3_3:12434229-12434252:+	CAGTGGGGATGTCTCATAA
	PPARG_CCDS2609.1.ex4_3:12447380-12447403:-	GGCATCCGCCCAAACCTGA
	PPARG_CCDS2609.1.ex4_3:12447450-12447473:-	AGCGGACTCTGGATTCAGC
	PPARG_CCDS2609.1.ex4_3:12447524-12447547:+	GCTGACCAAAGCAAAGGCG
ABCG1	ABCG1_CCDS13681.1.ex2_21:43691189-43691212:-	CCCGGAAATTCCTTTCAGG
	ABCG1_CCDS13681.1.ex5_21:43702417-43702440:+	TTGCGCCAACACGCGGACC
	ABCG1_CCDS13681.1.ex5_21:43702448-43702471:+	GGTGGTCAGCGCAAGCGCC
	ABCG1_CCDS13681.1.ex6_21:43704708-43704731:+	ATGAAAGGGCTCGCTCAAG
	ABCG1_CCDS13681.1.ex6_21:43704760-43704783:-	GCTCGAAGAGTTTGGCGCT
INS	INS_CCDS7729.1.ex1_11:2182059-2182082:+	GCCTCGTTCCCCGCACACT
	INS_CCDS7729.1.ex1_11:2181999-2182022:+	GGCAGTTGGCTCACCCTGC
	INS_CCDS7729.1.ex1_11:2182023-2182046:-	AAGACCCGCCGGGAGGCAG
	INS_CCDS7729.1.ex1_11:2182077-2182100:+	TAGGTAGAGAGCTTCCACC
	INS_CCDS7729.1.ex1_11:2182095-2182118:-	CACCTGTGCGGCTCACACC
PDZK1	PDZK1_CCDS55629.1.ex0_1:145747141-145747164:-	CTTCTCAACCACCCGGACC
	PDZK1_CCDS55629.1.ex0_1:145747159-145747182:+	AAGTGTAGCCCAGCAGAGA
	PDZK1_CCDS55629.1.ex0_1:145747199-145747222:+	ACAGAGTTCTTAGGATCAA
	PDZK1_CCDS55629.1.ex1_1:145748420-145748443:+	CGGGTGGACTTGAAAGAGT

	PDZK1_CCDS55629.1.ex1_1:145748523-145748546:-	TCACGAGATAGCAGAGCCG
THBS1	THBS1_CCDS32194.1.ex3_15:39876242-39876265:-	CCAATGTAGTTAGTGCGGA
	THBS1_CCDS32194.1.ex4_15:39876542-39876565:-	CTCCGTTGTGATAGCATAG
	THBS1_CCDS32194.1.ex5_15:39877751-39877774:-	GAAACTTACGCCAACAGCG
	THBS1_CCDS32194.1.ex6_15:39879550-39879573:+	CGACTCTGCGGACGATGGC
	THBS1_CCDS32194.1.ex9_15:39881176-39881199:-	ACTTCACGCCGGCAAAGCA
GRB2	GRB2_CCDS11721.1.ex1_17:73317744-73317767:+	CACCTGTTCTATGTCCCGC
	GRB2_CCDS11721.1.ex1_17:73317890-73317913:+	GCACATCGTTTCCAAACCT
	GRB2_CCDS11721.1.ex3_17:73328765-73328788:+	GGAGCTTAACTTACGGATG
	GRB2_CCDS11721.1.ex3_17:73328819-73328842:-	CAGAGCTTAATGGAAAAGA
	GRB2_CCDS11721.1.ex3_17:73328848-73328871:-	CGAAGAATGTGATCAGAAC
CBL	CBL_CCDS8418.1.ex0_11:119077261-119077284:-	GTCCACCGTCCCCGGCGGG
	CBL_CCDS8418.1.ex1_11:119103202-119103225:-	GCAGGTCTAAGATATAAGG
	CBL_CCDS8418.1.ex3_11:119144586-119144609:+	CCTTGGAAGAGCTTTCGAC
	CBL_CCDS8418.1.ex3_11:119144663-119144686:-	TTGCAGGTCAGATCAATAG
	CBL_CCDS8418.1.ex8_11:119149273-119149296:-	CAAACGGATCTACCACGAT
CAPN5	CAPN5_CCDS8248.1.ex0_11:76796047-76796070:+	CTACTATAAGGGCACGCCG
	CAPN5_CCDS8248.1.ex2_11:76823677-76823700:-	AAGATGCCCGCGTAGGCGT
	CAPN5_CCDS8248.1.ex2_11:76823804-76823827:+	GAGTTTTGGTGCGCCCTAG
	CAPN5_CCDS8248.1.ex4_11:76826469-76826492:-	CTTTACCAGGCCGCACGCC
	CAPN5_CCDS8248.1.ex4_11:76826493-76826516:-	ATCAGTGACGGCGTATGCG
GNB1	GNB1_CCDS34.1.ex4_1:1735894-1735917:-	CAATCTGAAAACTCGTGAG
	GNB1_CCDS34.1.ex4_1:1735947-1735970:+	GGCCACATAGTTCCCAGAA
	GNB1_CCDS34.1.ex5_1:1737940-1737963:+	ATAAGTTTACCATCCTGCG
	GNB1_CCDS34.1.ex6_1:1747250-1747273:-	AATCCAAATGCGCACGAGG
	GNB1_CCDS34.1.ex6_1:1747275-1747298:-	CAAACAACATCGACCCAGT
EPCAM	EPCAM_CCDS1833.1.ex2_2:47601017-47601040:+	GGGCCCTCCAGAACAATGA
	EPCAM_CCDS1833.1.ex2_2:47601048-47601071:+	TCCTGACTGCGATGAGAGC
	EPCAM_CCDS1833.1.ex2_2:47601162-47601185:-	GTTCTCACTCGCTCAGAGC
	EPCAM_CCDS1833.1.ex4_2:47604167-47604190:+	ATCACAACGCGTTATCAAC
	EPCAM_CCDS1833.1.ex5_2:47606087-47606110:-	GATAACATTATTCTCATAC
PTK7	PTK7_CCDS4884.1.ex1_6:43096856-43096879:+	ACGGAGCGGCGTTTCGCCC
	PTK7_CCDS4884.1.ex3_6:43097972-43097995:+	CTACCAATGGTTCCGAGAT

	PTK7_CCDS4884.1.ex4_6:43098382-43098405:+	TCACTAACCGCAGTCGGTA
	PTK7_CCDS4884.1.ex5_6:43099809-43099832:+	GGTCCGGCCACGCAATGCA
	PTK7_CCDS4884.1.ex6_6:43100172-43100195:-	ACACCCGTGGCTCAAATAG
ITGB1	ITGB1_CCDS7174.1.ex9_10:33212499-33212522:+	CCAGGCACTTACATAATAA
	ITGB1_CCDS7174.1.ex9_10:33212631-33212654:-	AGGAATGTTACACGGCTGC
	ITGB1_CCDS7174.1.ex11_10:33217020-33217043:-	TTACTTCGGACTTCAGAAT
	ITGB1_CCDS7174.1.ex12_10:33218748-33218771:-	TGGTTTTGCGATTAAGATC
	ITGB1_CCDS7174.1.ex12_10:33218823-33218846:-	ATGTAACCAACCGTAGCAA
APMAP	APMAP_CCDS13166.1.ex4_20:24952122-24952145:-	TTTGTGGCCGATGCATACA
	APMAP_CCDS13166.1.ex4_20:24952150-24952173:-	TATCCGTGCAGGGCCCAAT
	APMAP_CCDS13166.1.ex5_20:24954287-24954310:+	CCCGAACCAAACCGGGCAA
	APMAP_CCDS13166.1.ex6_20:24959459-24959482:-	CCAAATACGAAGCTGCGAC
	APMAP_CCDS13166.1.ex7_20:24964611-24964634:-	CGAGTGACCTTCTTGATGC
GGT1	GGT1_CCDS42992.1.ex1_22:25010867-25010890:-	CTCCCGAGGCACTCACGTG
	GGT1_CCDS42992.1.ex2_22:25011017-25011040:+	GTCATCAACGCCCGCGAGG
	GGT1_CCDS42992.1.ex3_22:25016441-25016464:-	ACGGTCCGCTTGTTTTCCA
	GGT1_CCDS42992.1.ex4_22:25016882-25016905:+	GTGTTCTGCCGGGATAGAA
	GGT1_CCDS42992.1.ex4_22:25016951-25016974:+	TACGAGACGCTGGCCATCG
SLC7A5	SLC7A5_CCDS10964.1.ex5_16:87873377-87873400:+	CGTACACCAGCGTCACGAT
	SLC7A5_CCDS10964.1.ex6_16:87874033-87874056:-	GGAAATGATCAACCCCTAC
	SLC7A5_CCDS10964.1.ex7_16:87874661-87874684:-	ACAGCGGCCTCTTTGCCTA
	SLC7A5_CCDS10964.1.ex7_16:87874676-87874699:-	TTGTGCTGGCATTATACAG
	SLC7A5_CCDS10964.1.ex7_16:87874717-87874740:+	GTGCCTTCAAATGAGAAGT
B2M	B2M_CCDS10113.1.ex1_15:45007643-45007666:+	CACGTCATCCAGCAGAGAA
	B2M_CCDS10113.1.ex1_15:45007676-45007699:+	TCCTGAATTGCTATGTGTC
	B2M_CCDS10113.1.ex1_15:45007706-45007729:-	AGTCAACTTCAATGTCGGA
	B2M_CCDS10113.1.ex1_15:45007738-45007761:+	GGAGAGAGAATTGAAAAAG
	B2M_CCDS10113.1.ex1_15:45007770-45007793:+	CTTGTCTTTCAGCAAGGAC
IL6ST	IL6ST_CCDS3971.1.ex11_5:55260070-55260093:+	GAATAATCAACAGTGCATG
	IL6ST_CCDS3971.1.ex12_5:55264131-55264154:-	GGTGGAAGGGAAACACACT
	IL6ST_CCDS3971.1.ex12_5:55264177-55264200:-	GAGTTGCATTGTGAACGAG
	IL6ST_CCDS3971.1.ex13_5:55265515-55265538:+	CTCCTTAGGAATAGTAAAA
	IL6ST_CCDS3971.1.ex13_5:55265593-55265616:-	ACTGCAGTTTGTGTGCTAA

TSPAN14	TSPAN14_CCDS44448.1.ex1_10:82271890-82271913:+	GGTTGCAGAACCAGTGCTG
	TSPAN14_CCDS44448.1.ex1_10:82271913-82271936:+	GCATATGGCCCTGAAGACT
	TSPAN14_CCDS44448.1.ex1_10:82271941-82271964:+	ACGTCTACTTCAATTGCAG
	TSPAN14_CCDS44448.1.ex1_10:82271968-82271991:+	GCTACAGCCGAGAGAAGTG
	TSPAN14_CCDS44448.1.ex1_10:82272013-82272036:-	GCCAACTCACCGCAGGATC
PLXNA1	PLXNA1_CCDS33847.2.ex0_3:126708128-126708151:-	CTTGGACAGCGTGTCCGAA
	PLXNA1_CCDS33847.2.ex1_3:126710373-126710396:+	TATTCGCCGGCACGCGAAG
	PLXNA1_CCDS33847.2.ex10_3:126733083-126733106:-	ACCATCCGCACTCGAAGCG
	PLXNA1_CCDS33847.2.ex10_3:126733155-126733178:-	GCGCGTGCATCCACGATGC
	PLXNA1_CCDS33847.2.ex11_3:126733388-126733411:+	CGATTCGAAGACGTGCGTC
FGFR4	FGFR4_CCDS4410.1.ex2_5:176517794-176517817:-	GTAACTGTGCCTATTCGAG
	FGFR4_CCDS4410.1.ex4_5:176518702-176518725:+	AGTCTCGTGATGGAGAGCG
	FGFR4_CCDS4410.1.ex5_5:176519480-176519503:-	TGCACATAGGGGAAACCGT
	FGFR4_CCDS4410.1.ex6_5:176519684-176519707:-	CTCGGCTGACACGTTCCGC
	FGFR4_CCDS4410.1.ex6_5:176519726-176519749:-	GCCGATGGAATTGCCTGCG
NAT1	NAT1_CCDS55205.1.ex2_8:18079681-18079704:-	GTCCATGGCATCCCCACAA
	NAT1_CCDS55205.1.ex2_8:18079727-18079750:+	AAGTTGTGAGAAGAAATCG
	NAT1_CCDS55205.1.ex2_8:18079794-18079817:-	GTGGTCTCAAAACCAATAG
	NAT1_CCDS55205.1.ex2_8:18079812-18079835:-	ACATACCCTCCCAACATCG
	NAT1_CCDS55205.1.ex2_8:18080003-18080026:+	TCCGTTTGACGGAAGAGAA
CD63	CD63_CCDS58242.1.ex2_12:56119973-56119996:+	GGGACTCGGTTCTTCGACA
	CD63_CCDS58242.1.ex2_12:56120009-56120032:-	GGCTGCTAACTACACAGAT
	CD63_CCDS58242.1.ex2_12:56120027-56120050:+	GCCCCACAGCACTTAAACT
	CD63_CCDS58242.1.ex3_12:56120495-56120518:-	CACTGCTTCGATCCTGGAC
	CD63_CCDS58242.1.ex3_12:56120539-56120562:-	AACAACTTCCGGCAGCAGA
FCGRT	FCGRT_CCDS12770.1.ex1_19:50017239-50017262:+	TGAGCTACAATAGCCTGCG
	FCGRT_CCDS12770.1.ex1_19:50017325-50017348:+	ACCACAGATCTGAGGATCA
	FCGRT_CCDS12770.1.ex2_19:50017533-50017556:-	TCGCCGTTCAGGGCGAACT
	FCGRT_CCDS12770.1.ex2_19:50017561-50017584:+	ATGAATTTCGACCTCAAGC
	FCGRT_CCDS12770.1.ex2_19:50017611-50017634:+	CCTGGCTATCAGTCAGCGG
DSC1	DSC1_CCDS11894.1.ex8_18:28723710-28723733:-	TAATGGAAGTGCGAGACAT

	DSC1_CCDS11894.1.ex9_18:28725686-28725709:+	GAGAGTGTCAGGTTCGTCA
	DSC1_CCDS11894.1.ex10_18:28728459-28728482:-	TGCCTGAAAATTGCCGATC
	DSC1_CCDS11894.1.ex12_18:28736004-28736027:-	TTTCCACAACACGTTCAGC
	DSC1_CCDS11894.1.ex12_18:28736067-28736090:-	AAGCGCAGCAAGAGACGAT
MME	MME_CCDS3172.1.ex0_3:154802094-154802117:+	CACTCTATGCAACCTACGA
	MME_CCDS3172.1.ex2_3:154832783-154832806:+	GCTCGACTGATCCAAAACA
	MME_CCDS3172.1.ex2_3:154832884-154832907:-	AAGTTGCCGTAACGGGAGC
	MME_CCDS3172.1.ex6_3:154836531-154836554:+	AGATTGACCAACCTCGACT
	MME_CCDS3172.1.ex6_3:154836554-154836577:-	ATAGTAATCTCTAGAAGGG
TFR2	TFR2_CCDS34707.1.ex12_7:100230716-100230739:-	GGTGTACGCCCACTACGGG
	TFR2_CCDS34707.1.ex13_7:100230870-100230893:+	TGTAGGGGCAGTAGACGTC
	TFR2_CCDS34707.1.ex14_7:100231057-100231080:-	TGGACCGACACGCACTACG
	TFR2_CCDS34707.1.ex14_7:100231108-100231131:-	CTGACTCAGGACATTCGCG
	TFR2_CCDS34707.1.ex15_7:100238463-100238486:-	GGGCTACGTCGCCTTCCGA
ENPEP	ENPEP_CCDS3691.1.ex0_4:111397659-111397682:+	TTAATAGTGGGACTTGCCG
	ENPEP_CCDS3691.1.ex0_4:111397688-111397711:-	CCGCTGGAGTCACACGATC
	ENPEP_CCDS3691.1.ex0_4:111397849-111397872:-	GGACTGGGTTGACGAAGTC
	ENPEP_CCDS3691.1.ex0_4:111398186-111398209:-	TGTCCGTTCTCCGTGTAGG
	ENPEP_CCDS3691.1.ex1_4:111409707-111409730:-	TCTGTTGGTTCATGATCGG
CD46	CD46_CCDS1479.1.ex1_1:207930371-207930394:+	CCACCAACATTTGAAGCTA
	CD46_CCDS1479.1.ex1_1:207930415-207930438:-	CGTTCACCAATCTCATAGT
	CD46_CCDS1479.1.ex1_1:207930477-207930500:-	AAATAGTATGGGTGGCAAG
	CD46_CCDS1479.1.ex5_1:207940353-207940376:-	GAAATCGACATTTGACCAC
	CD46_CCDS1479.1.ex5_1:207940368-207940391:+	GATTTCCAGTAGTCGAAAA
SLC4A1	SLC4A1_CCDS11481.1.ex10_17:42335975-42335998:+	ATCTATGCGGAACACCTAG
	SLC4A1_CCDS11481.1.ex11_17:42336562-42336585:+	AAGCTGGGTGTAATCGATG
	SLC4A1_CCDS11481.1.ex12_17:42336859-42336882:+	TTACCCACTAGCACCAACG
	SLC4A1_CCDS11481.1.ex14_17:42337835-42337858:+	GATAAACCTGTCTAGCAGT
	SLC4A1_CCDS11481.1.ex15_17:42338011-42338034:+	GACTCTACGCAGCTCTAGG
VNN1	VNN1_CCDS5159.1.ex4_6:133015210-133015233:+	GAGGATCACTGGTATCGCA
	VNN1_CCDS5159.1.ex5_6:133032846-133032869:-	CCCCTGTAATAATCGTAAC
	VNN1_CCDS5159.1.ex5_6:133032921-133032944:-	TTATGGCTGGAACTTCAAC
	VNN1_CCDS5159.1.ex6_6:133035002-133035025:-	GCATTAATGAATCGGAATC

	VNN1_CCDS5159.1.ex6_6:133035029-133035052:-	ACACCAGTGTCTCGTGAGG
ITGA5	ITGA5_CCDS8880.1.ex15_12:54798004-54798027:+	ACTAGCGGACACGATGGGG
	ITGA5_CCDS8880.1.ex18_12:54799110-54799133:+	GGGTCCGATCCATGAGCAG
	ITGA5_CCDS8880.1.ex26_12:54803077-54803100:-	AGCCACTGAGCGACCCCGT
	ITGA5_CCDS8880.1.ex29_12:54812773-54812796:-	CAGTGGAGTTTTACCGGCC
	ITGA5_CCDS8880.1.ex29_12:54812843-54812866:-	GGGGGCTTCAACTTAGACG
GGT5	GGT5_CCDS13825.1.ex5_22:24622643-24622666:-	GACGCTCAAGTTTGCCAAG
	GGT5_CCDS13825.1.ex6_22:24627350-24627373:-	TTATCCTCAACGTGCTAAG
	GGT5_CCDS13825.1.ex8_22:24628870-24628893:-	CACCATCGCGCTGCTCCGA
	GGT5_CCDS13825.1.ex10_22:24629840-24629863:-	CCATCTACAATGTGACAAC
	GGT5_CCDS13825.1.ex10_22:24629885-24629908:-	TCGTCAACCCTCAGAGCAT
PVRL2	PVRL2_CCDS12645.1.ex1_19:45368610-45368633:-	TGTACAGTCCAGGAACAGG
	PVRL2_CCDS12645.1.ex1_19:45368663-45368686:+	CCTGCGAACCACCAGAATG
	PVRL2_CCDS12645.1.ex1_19:45368684-45368707:+	GCCGCCTTCCACCCTAAGA
	PVRL2_CCDS12645.1.ex1_19:45368872-45368895:+	CCCCAAGGGGTCCGTCCGA
	PVRL2_CCDS12645.1.ex1_19:45368899-45368922:-	TCACCTATGACTCTGAGCC
NEU1	NEU1_CCDS4723.1.ex3_6:31828965-31828988:+	GAATACCAGAGCCCGGTCC
	NEU1_CCDS4723.1.ex3_6:31829171-31829194:-	TGATGGGGATGTCCCCGAT
	NEU1_CCDS4723.1.ex4_6:31829788-31829811:-	CAAGTTCATCGCCCTGCGG
	NEU1_CCDS4723.1.ex4_6:31829806-31829829:+	TTGGCCCCCTCATCGGATG
	NEU1_CCDS4723.1.ex4_6:31829836-31829859:-	TCTCGCCTTTGCTGAGGCG
IGSF8	IGSF8_CCDS1195.1.ex3_1:160063632-160063655:+	ACTACCATGCGGTACCGAT
	IGSF8_CCDS1195.1.ex3_1:160063724-160063747:-	GTGGGAATCCGGTCAGACT
	IGSF8_CCDS1195.1.ex3_1:160063753-160063776:+	GCAGAGTTGACCGCCCAAC
	IGSF8_CCDS1195.1.ex3_1:160063778-160063801:-	TTTGGGCGATCTGTGCCCG
	IGSF8_CCDS1195.1.ex4_1:160064981-160065004:+	CGCGGTACAAGGGCCCCTC
LEPR	LEPR_CCDS30740.1.ex3_1:66058510-66058533:-	AACTGACATTAGAGGTGAC
	LEPR_CCDS30740.1.ex4_1:66062135-66062158:-	GCAAACCTAATGGTGGATC
	LEPR_CCDS30740.1.ex5_1:66064368-66064391:-	AAGTATACTGTCTACTAGC
	LEPR_CCDS30740.1.ex5_1:66064398-66064421:+	TCGTATGAGGTTCAGGTGA
	LEPR_CCDS30740.1.ex6_1:66067333-66067356:-	TAATTCAGCATAGCGATGA
PIGS	PIGS_CCDS11235.1.ex5_17:26887081-26887104:+	AAGTTGCCAGCGGCACCGA
	PIGS_CCDS11235.1.ex6_17:26888583-26888606:-	CGGGGGATAATGCACCGGG

	PIGS_CCDS11235.1.ex8_17:26890836-26890859:+	TGCACACTGCCCGATGACA
	PIGS_CCDS11235.1.ex9_17:26897885-26897908:+	CTTTCATGCACAACGGTGA
	PIGS_CCDS11235.1.ex9_17:26897943-26897966:+	GCGTAAACACGACAGTGAC
PLXNB1	PLXNB1_CCDS2765.1.ex26_3:48460697-48460720:+	ACATGGACCGGCATCAACG
	PLXNB1_CCDS2765.1.ex31_3:48463186-48463209:-	CGCCGTTCTGAGTGCTCGA
	PLXNB1_CCDS2765.1.ex35_3:48465008-48465031:+	TCGCGTGCGATTAGCAAGC
	PLXNB1_CCDS2765.1.ex35_3:48465346-48465369:-	TCGTGAGTGCCTTTGCACG
	PLXNB1_CCDS2765.1.ex35_3:48465671-48465694:-	GTATGCGGGAGCGTGCACC
INSR	INSR_CCDS12176.1.ex14_19:7166210-7166233:+	GTCCGGCGTTCATCCGAAA
	INSR_CCDS12176.1.ex17_19:7172358-7172381:-	ATCCGCCGATCCTACGCTC
	INSR_CCDS12176.1.ex18_19:7174654-7174677:+	GCAGACGTCACCGAGTCGA
	INSR_CCDS12176.1.ex20_19:7267433-7267456:-	ATCTGTCCGGGTACCGCGA
	INSR_CCDS12176.1.ex20_19:7267646-7267669:-	CTGTTCTTTAACTACGCGC
LDLR	LDLR_CCDS12254.1.ex1_19:11210951-11210974:+	CCTACAAGTGGGTCTGCGA
	LDLR_CCDS12254.1.ex5_19:11218060-11218083:+	TCACAGTGACACTCTGCGA
	LDLR_CCDS12254.1.ex5_19:11218104-11218127:+	AGCGGCGAATGCATCACCC
	LDLR_CCDS12254.1.ex6_19:11221425-11221448:+	TGGCCCAGCGAAGATGCGA
	LDLR_CCDS12254.1.ex8_19:11224041-11224064:+	CTGAGGAACGTGGTCGCTC
EPHB4	EPHB4_CCDS5706.1.ex12_7:100417798-100417821:+	GAAGTACCCGACGCGGCAC
	EPHB4_CCDS5706.1.ex13_7:100419946-100419969:-	TGGGCCGAACAGCCGGTCA
	EPHB4_CCDS5706.1.ex13_7:100420069-100420092:+	AGTCTCCGGGAATCGAGTC
	EPHB4_CCDS5706.1.ex13_7:100420253-100420276:+	TTCCGGGTGAGATGCTCCG
	EPHB4_CCDS5706.1.ex14_7:100421406-100421429:+	AGCGTGGCGTACACGTGGA
ERBB2	ERBB2_CCDS32642.1.ex3_17:37865687-37865710:-	TTACAGGCCCGAGAGCGGT
	ERBB2_CCDS32642.1.ex6_17:37866685-37866708:-	CGAATGTATACCGGCCCTC
	ERBB2_CCDS32642.1.ex7_17:37868187-37868210:-	GGATCCCACGTCCGTAGAA
	ERBB2_CCDS32642.1.ex11_17:37872103-37872126:+	TTCGTGCACACGGTGCCCT
	ERBB2_CCDS32642.1.ex11_17:37872170-37872193:+	ACCGGCCAGAGGACGAGTG
IGSF1	IGSF1_CCDS14629.1.ex16_X:130419739-130419762:-	TTCTAGAGTTGGAGGCACC
	IGSF1_CCDS14629.1.ex16_X:130419792-130419815:-	TCTTTACCGGTGCTGCTAC
	IGSF1_CCDS14629.1.ex16_X:130419813-130419836:+	CCTGCATTGGACTCAGTAA
	IGSF1_CCDS14629.1.ex16_X:130419917-130419940:+	CTTGCTTGATATCCGAGAG
	IGSF1_CCDS14629.1.ex16_X:130419974-130419997:-	ATAGAGTCCAACTACCCCC

SEMA4G	SEMA4G_CCDS55724.1.ex0_10:102732766-102732789:-	TAGGGTATGGTCATCCGAG
	SEMA4G_CCDS55724.1.ex1_10:102732999-102733022:+	TGCCAACGACATAGGAGAT
	SEMA4G_CCDS55724.1.ex3_10:102737414-102737437:-	GGTAGAATTGAGCCGCTGC
	SEMA4G_CCDS55724.1.ex4_10:102738068-102738091:+	TTGCCAACCAGCTTCGAGG
	SEMA4G_CCDS55724.1.ex5_10:102738301-102738324:-	CTCCGGAATTCATACCTAG
GPC3	GPC3_CCDS14638.1.ex5_X:132887700-132887723:+	ACATTGCAGTAACCGCCAC
	GPC3_CCDS14638.1.ex5_X:132887885-132887908:-	CCCAAGCTTATTATGACCC
	GPC3_CCDS14638.1.ex5_X:132887941-132887964:-	ACATCAATGAGTGCCTCCG
	GPC3_CCDS14638.1.ex5_X:132887999-132888022:+	GGTATAGATGACTGGAAAC
	GPC3_CCDS14638.1.ex5_X:132888110-132888133:+	CTCAAAAGCTTGTGGAGTC
APLP2	APLP2_CCDS44773.1.ex1_11:129979326-129979349:+	TTGCAGCCAATGCCGGAAC
	APLP2_CCDS44773.1.ex1_11:129979406-129979429:+	AACATTCAGACTGGGAAAT
	APLP2_CCDS44773.1.ex1_11:129979422-129979445:+	AATGGGAACCTGATCCAAC
	APLP2_CCDS44773.1.ex2_11:129980475-129980498:+	GCGGGTTAGTATTGACAAC
	APLP2_CCDS44773.1.ex7_11:129996605-129996628:-	CAACATCATTGGTTGGCAG
MPZL2	MPZL2_CCDS8393.1.ex2_11:118133231-118133254:-	ACTGCAGTTCGACGACAAT
	MPZL2_CCDS8393.1.ex2_11:118133279-118133302:-	TTGGGATGGGAATCCTGAG
	MPZL2_CCDS8393.1.ex2_11:118133306-118133329:-	GAGTGGGCGGTTTAAGGAC
	MPZL2_CCDS8393.1.ex3_11:118133659-118133682:-	AATTTTCGTCCTCTAGACG
	MPZL2_CCDS8393.1.ex3_11:118133690-118133713:+	GTTAGAGCATCACCCACAG
LPHN2	LPHN2_CCDS689.1.ex1_1:82372762-82372785:+	TCTATAGATCTGCGATGCC
	LPHN2_CCDS689.1.ex1_1:82372799-82372822:+	TGATTGAGAGCGCTAACTA
	LPHN2_CCDS689.1.ex3_1:82408783-82408806:-	TTAAAGTATCGGTACGATA
	LPHN2_CCDS689.1.ex3_1:82409284-82409307:+	ATAATACCCGATTAAACCG
	LPHN2_CCDS689.1.ex8_1:82431826-82431849:-	GACGGTATTTGCGGACAGT
DSC2	DSC2_CCDS11892.1.ex8_18:28662896-28662919:+	ACGAGTAAATGTTGGCAAG
	DSC2_CCDS11892.1.ex9_18:28666618-28666641:-	AAGTACTCCATCATTGGGC
	DSC2_CCDS11892.1.ex10_18:28667732-28667755:+	GTTCTGGAGTATACCCATC
	DSC2_CCDS11892.1.ex12_18:28671021-28671044:+	AGTTTTCTAGCATCGAACA
	DSC2_CCDS11892.1.ex13_18:28672132-28672155:+	GTAAAACTTCTCTTCTCCG

ITGA2	ITGA2_CCDS3957.1.ex6_5:52347367-52347390:+	CGGAGCAATTCAATATGCA
	ITGA2_CCDS3957.1.ex7_5:52351400-52351423:+	CGAAGTGCTACGAAAGTAA
	ITGA2_CCDS3957.1.ex7_5:52351431-52351454:+	CTGACGGTGAATCACATGA
	ITGA2_CCDS3957.1.ex11_5:52356786-52356809:-	GGCCGGTATAATTTGCCCG
	ITGA2_CCDS3957.1.ex13_5:52360758-52360781:+	CAATTTCTTGAAGGCCCCG
NPTN	NPTN_CCDS10249.1.ex4_15:73879859-73879882:+	GTACCTTTCACTTCAATGG
	NPTN_CCDS10249.1.ex4_15:73879900-73879923:+	GACAAAGTGATATACGCAG
	NPTN_CCDS10249.1.ex5_15:73884365-73884388:-	CACCCTTACATACAGCTAC
	NPTN_CCDS10249.1.ex5_15:73884451-73884474:-	AGGATTGTCACCAGTGAAG
	NPTN_CCDS10249.1.ex7_15:73925454-73925477:+	CCCCGCTTACCGTTCTGAG
LPHN2	LPHN2_CCDS689.1.ex1_1:82372762-82372785:+	TCTATAGATCTGCGATGCC
	LPHN2_CCDS689.1.ex1_1:82372799-82372822:+	TGATTGAGAGCGCTAACTA
	LPHN2_CCDS689.1.ex3_1:82408783-82408806:-	TTAAAGTATCGGTACGATA
	LPHN2_CCDS689.1.ex3_1:82409284-82409307:+	ATAATACCCGATTAAACCG
	LPHN2_CCDS689.1.ex8_1:82431826-82431849:-	GACGGTATTTGCGGACAGT
SLC2A1	SLC2A1_CCDS477.1.ex4_1:43395432-43395455:+	CGCGCAGCTTCTTTAGCAC
	SLC2A1_CCDS477.1.ex6_1:43396421-43396444:+	ACCGATGATGAAGCGGCCC
	SLC2A1_CCDS477.1.ex6_1:43396520-43396543:+	CAGCATTGAATTCCGCCTG
	SLC2A1_CCDS477.1.ex7_1:43396717-43396740:+	GCCAAAGCGGTTAACGAAA
	SLC2A1_CCDS477.1.ex7_1:43396833-43396856:-	AGACATGGGTCCACCGCTA
FAS	FAS_CCDS7393.1.ex1_10:90762922-90762945:-	ACAGGGCTTATGGCAGAAT
	FAS_CCDS7393.1.ex2_10:90767450-90767473:+	GGCAGGTGAAAGGAAAGCT
	FAS_CCDS7393.1.ex2_10:90767468-90767491:+	TAGGGACTGCACAGTCAAT
	FAS_CCDS7393.1.ex2_10:90767566-90767589:+	GATGTAGATTGTGTGATGA
	FAS_CCDS7393.1.ex3_10:90768651-90768674:+	AGTGGAAATAAACTGCACC
TMPRSS6	TMPRSS6_CCDS13941.1.ex9_22:37480787-37480810:+	GTTTGGGGCGAGTAGTAGC
	TMPRSS6_CCDS13941.1.ex11_22:37485640-37485663:+	GGCCCGGCCACGTCATACA
	TMPRSS6_CCDS13941.1.ex11_22:37485803-37485826:+	AGCTGTAGCGGTAACAACC
	TMPRSS6_CCDS13941.1.ex12_22:37491576-37491599:+	AAAAATAGCGTACCCAGCG
	TMPRSS6_CCDS13941.1.ex13_22:37492065-37492088:+	CATCAGCCGGCGGTGCTCG
FCER1G	FCER1G_CCDS1225.1.ex1_1:161187818-161187841:-	GACAATTCCATACAGAAAC

	FCER1G_CCDS1225.1.ex1_1:161187848-161187871:-	TACCTTCAGTCGACAGTAG
	FCER1G_CCDS1225.1.ex2_1:161188032-161188055:-	TATAGCTGCCTTTCGCACT
	FCER1G_CCDS1225.1.ex2_1:161188055-161188078:-	AGGGTCCATACCTCATAGC
	FCER1G_CCDS1225.1.ex3_1:161188480-161188503:+	TGTCTCTGCAGAAATCAGA
CD14	CD14_CCDS4232.1.ex0_5:140012098-140012121:-	GCAACGTGTCGTGGGCGAC
	CD14_CCDS4232.1.ex0_5:140012179-140012202:-	TCGAGGACCTAAAGATAAC
	CD14_CCDS4232.1.ex0_5:140012220-140012243:+	GAGTACGCTAGCACACGCA
	CD14_CCDS4232.1.ex0_5:140012311-140012334:+	CCGTGTCAGCATACTGCCG
	CD14_CCDS4232.1.ex0_5:140012344-140012367:+	CATCGACGCGCTTTAGAAA
XPNPEP2	XPNPEP2_CCDS14613.1.ex4_X:128880244-128880267:+	CTGGACCGACAGTCGCTAC
	XPNPEP2_CCDS14613.1.ex6_X:128881640-128881663:+	ACAACCAATCTTGTGGACC
	XPNPEP2_CCDS14613.1.ex7_X:128884504-128884527:-	CGACAGAAGGACGGCAGTC
	XPNPEP2_CCDS14613.1.ex8_X:128885738-128885761:-	GGGTTATAGGGGATGTCAC
	XPNPEP2_CCDS14613.1.ex9_X:128886280-128886303:-	ATCCCATACATGGTATAGC
EFNB1	EFNB1_CCDS14391.1.ex1_X:68058478-68058501:+	TGGTGATCTATCCGAAAAT
	EFNB1_CCDS14391.1.ex1_X:68058593-68058616:-	GGGTCGAGAACTGTGCTAC
	EFNB1_CCDS14391.1.ex2_X:68059507-68059530:+	ACATCCAATGGAAGCCTGG
	EFNB1_CCDS14391.1.ex3_X:68059854-68059877:+	GCAGACAACACTGTCAAGA
	EFNB1_CCDS14391.1.ex3_X:68059909-68059932:+	GTGACTCTGATGGCAAGCA
CDH1	CDH1_CCDS10869.1.ex2_16:68835581-68835604:+	AGATTGCACCGGTCGACAA
	CDH1_CCDS10869.1.ex2_16:68835618-68835641:+	CTCGACACCCGATTCAAAG
	CDH1_CCDS10869.1.ex2_16:68835670-68835693:-	GTGGGTTATGAAACCGTAG
	CDH1_CCDS10869.1.ex7_16:68846042-68846065:-	AACCACCAGGGTATACGTA
	CDH1_CCDS10869.1.ex8_16:68847295-68847318:-	GTATACAGCCTCCCACGCT
TGFBR1	TGFBR1_CCDS47998.1.ex1_9:101891326-101891349:+	ACAACATATTGCTGCAATC
	TGFBR1_CCDS47998.1.ex1_9:101891350-101891373:-	AAGTTCTATTTTATTGCAA
	TGFBR1_CCDS47998.1.ex2_9:101900178-101900201:+	TTGTGTTACAAGAAAGCAT
	TGFBR1_CCDS47998.1.ex2_9:101900279-101900302:+	GAACGTTCGTGGTTCCGTG
	TGFBR1_CCDS47998.1.ex2_9:101900342-101900365:-	GTCTGCTGCTATAAATCCC
FCGR2B	FCGR2B_CCDS30924.1.ex2_1:161641199-161641222:+	GCTGAAACTCGAGCCCCAG

	FCGR2B_CCDS30924.1.ex2_1:161641218-161641241:+	TGGATCAACGTGCTCCAGG
	FCGR2B_CCDS30924.1.ex2_1:161641244-161641267:+	TGTGACTCTGACATGCCGG
	FCGR2B_CCDS30924.1.ex3_1:161642798-161642821:-	CACGATGGTTTCTCCCTCC
	FCGR2B.5	TGGAGCACGTTGATCCACT
ITGB2	ITGB2_CCDS13716.1.ex8_21:46318999-46319022:-	CATCTTCGCGGTGACCAGT
	ITGB2_CCDS13716.1.ex9_21:46320308-46320331:+	CTTCCCGTCGCCCGCGAAA
	ITGB2_CCDS13716.1.ex9_21:46320360-46320383:-	CGGCTGGCGCAACGTCACG
	ITGB2_CCDS13716.1.ex11_21:46323415-46323438:-	GTTCAACGTGACCTTCCGG
	ITGB2_CCDS13716.1.ex13_21:46330225-46330248:-	CCGGGAATGCATCGAGTCG
CEACAM6	CEACAM6_CCDS12585.1.ex1_19:42260561-42260584:-	TCTGCGACATTGAACGGCG
	CEACAM6_CCDS12585.1.ex1_19:42260621-42260644:+	GAATCGTATTGGTTACAGC
	CEACAM6_CCDS12585.1.ex1_19:42260846-42260869:-	CGGTATACATGGAACTGTC
	CEACAM6_CCDS12585.1.ex2_19:42265184-42265207:+	AACAACTCCAACCCCGTGG
	CEACAM6_CCDS12585.1.ex2_19:42265341-42265364:+	GCGTCAAAAGGAACGATGC
MSR1	MSR1_CCDS5995.1.ex5_8:16021675-16021698:-	GAACAAGTGCATTTGGAAC
	MSR1_CCDS5995.1.ex6_8:16026301-16026324:-	GATATAACTCAAAGTCTCA
	MSR1_CCDS5995.1.ex6_8:16026362-16026385:-	TACAGCTCAACTCCTGAAG
	MSR1_CCDS5995.1.ex7_8:16032694-16032717:-	CTCTCATTGGAATAGTGGC
	MSR1_CCDS5995.1.ex7_8:16032716-16032739:+	GATGAGAACTGCAAACACG
CD40	CD40_CCDS13393.1.ex2_20:44750917-44750940:-	TTCGCTTTCACCGCAAGGA
	CD40_CCDS13393.1.ex2_20:44750931-44750954:+	AAGCGAATTCCTAGACACC
	CD40_CCDS13393.1.ex3_20:44751233-44751256:+	CCTTCTCTTCTCAGACCTA
	CD40_CCDS13393.1.ex3_20:44751352-44751375:+	TGCACCGCTCATGCTCGCC
	CD40_CCDS13393.1.ex4_20:44751760-44751783:-	TATCAGAAACCCCTGTAGC
CLEC4M	CLEC4M_CCDS12187.1.ex1_19:7828322-7828345:+	TTCAATTCCAGCAGATACA
	CLEC4M_CCDS12187.1.ex1_19:7828344-7828367:+	CCACAAGAGCTCTACAGGT
	CLEC4M_CCDS12187.1.ex3_19:7830509-7830532:+	CTTCTTGGCCCAGTGTCCA
	CLEC4M_CCDS12187.1.ex3_19:7830535-7830558:-	TGTTCCTGACTTAGGGAGC
	CLEC4M_CCDS12187.1.ex3_19:7830587-7830610:+	ACCCAGCTTAAAGCTGCAG
CD163	CD163_CCDS53742.1.ex9_12:7640546-7640569:+	AATTCCCTGCATAGAACGC

	CD163_CCDS53742.1.ex10_12:7647692-7647715:+	GCTTCTTCATAGTGATCAC
	CD163_CCDS53742.1.ex11_12:7649541-7649564:+	CTGGCGTTAACTCGACCAA
	CD163_CCDS53742.1.ex12_12:7651700-7651723:-	ATCAAATTCCAAGGACGGT
	CD163_CCDS53742.1.ex13_12:7653951-7653974:-	AACGGTGTGTAATAATGGC
ITFG1	ITFG1_CCDS10728.1.ex9_16:47347648-47347671:+	CCAGATCTCACTAAGTAGA
	ITFG1_CCDS10728.1.ex11_16:47409804-47409827:+	AACTGGAAGGTACTAGTGG
	ITFG1_CCDS10728.1.ex14_16:47486663-47486686:+	TGAAAAGTCCTATTGAGTA
	ITFG1_CCDS10728.1.ex15_16:47488033-47488056:-	TGATAACAAGTGTAGTCCC
	ITFG1_CCDS10728.1.ex17_16:47494788-47494811:+	AGGTCCCCGAAAGCCGCAA
ITGAM	ITGAM_CCDS45470.1.ex4_16:31277412-31277435:-	CCGTAGGTTGGATCCAAAC
	ITGAM_CCDS45470.1.ex5_16:31282335-31282358:+	CATGACTTTCGGCGGATGA
	ITGAM_CCDS45470.1.ex11_16:31289296-31289319:-	CGGTTCCGTAAGATGATGG
	ITGAM_CCDS45470.1.ex11_16:31289404-31289427:-	GTGCCCTTGACATTAGCGT
	ITGAM_CCDS45470.1.ex13_16:31309224-31309247:+	ACCTGTTTCACGGAACCTC
PTPRF	PTPRF_CCDS489.2.ex2_1:44019515-44019538:-	CATCTCGCTGCACCCGCAA
	PTPRF_CCDS489.2.ex3_1:44035413-44035436:-	TTGATGCGGCCGTTGCTCG
	PTPRF_CCDS489.2.ex4_1:44044540-44044563:-	TAACGTGTGCCTGCCGAGT
	PTPRF_CCDS489.2.ex4_1:44044568-44044591:-	TCGCACATACAGGTTCGCA
	PTPRF_CCDS489.2.ex6_1:44057033-44057056:-	GTCCGGAGTATAGTAGACG
ERBB3	ERBB3_CCDS31833.1.ex1_12:56477556-56477579:-	AGCATCGCCGGTCACACTC
	ERBB3_CCDS31833.1.ex1_12:56477594-56477617:+	ACTGTACAAGCTCTACGAG
	ERBB3_CCDS31833.1.ex1_12:56477635-56477658:+	ACCTTGAGATTGTGCTCAC
	ERBB3_CCDS31833.1.ex2_12:56478775-56478798:+	AGTGGATTCGAGAAGTGAC
	ERBB3_CCDS31833.1.ex2_12:56478834-56478857:+	CCATTGCCCAACCTCCGCG
MET	MET_CCDS43636.1.ex0_7:116339268-116339291:-	GGTGTTTCCGCGGTGAAGT
	MET_CCDS43636.1.ex1_7:116371782-116371805:-	TCAACGCGCTGCAAAGCTG
	MET_CCDS43636.1.ex3_7:116381003-116381026:+	GACAAATGTGTGCGATCGG
	MET_CCDS43636.1.ex5_7:116397556-116397579:-	ACTGTATTGTGTTGTCCCG
	MET_CCDS43636.1.ex6_7:116397703-116397726:+	CAAGTATTTCGCCGAAATA
IL1R1	IL1R1_CCDS2055.1.ex1_2:102781414-102781437:+	TGGTTTGTTCCTGCTAAGG
	IL1R1_CCDS2055.1.ex3_2:102782622-102782645:+	AGGCTCATCGTGATGAATG
	IL1R1_CCDS2055.1.ex3_2:102782697-102782720:+	AAGCAATATCCTATTACCC
	IL1R1_CCDS2055.1.ex4_2:102785088-102785111:+	AGCCCAGCTAATGAGACAA

	IL1R1_CCDS2055.1.ex5_2:102788294-102788317:-	GTAAGCAATGTCACTCAAC
IFNGR1	IFNGR1_CCDS5185.1.ex2_6:137524794-137524817:-	TATAAAATACTCACGCAGA
	IFNGR1_CCDS5185.1.ex3_6:137525495-137525518:+	TTGTACACCCTAATGTAAC
	IFNGR1_CCDS5185.1.ex3_6:137525617-137525640:-	AAAATTGGACCACCTAAAC
	IFNGR1_CCDS5185.1.ex5_6:137528099-137528122:-	CCGTAGAGGTAAAGAACTA
	IFNGR1_CCDS5185.1.ex5_6:137528148-137528171:+	GTACTCCCAATATACGATA
PRNP	PRNP_CCDS13080.1.ex0_20:4679963-4679986:+	TGGGGGCAGCCGATACCCG
	PRNP_CCDS13080.1.ex0_20:4679995-4680018:-	GTGGGTAGCGGTTGCCTCC
	PRNP_CCDS13080.1.ex0_20:4680013-4680036:-	AGCCACCACCGCCCTGAGG
	PRNP_CCDS13080.1.ex0_20:4680274-4680297:-	CACTGCCGAAATGTATGAT
	PRNP_CCDS13080.1.ex0_20:4680342-4680365:-	CATGGGCCTGTAGTACACT
DCBLD1	DCBLD1_CCDS34522.1.ex2_6:117841038-117841061:-	GATCCACTCTCAAAGCGGA
	DCBLD1_CCDS34522.1.ex2_6:117841089-117841112:-	GGATGGTCGCTGCTCGCAT
	DCBLD1_CCDS34522.1.ex4_6:117846533-117846556:-	TACGTCTCTACAACCAGCT
	DCBLD1_CCDS34522.1.ex5_6:117853485-117853508:+	TCAGTGTGCTTCAGCGCAA
	DCBLD1_CCDS34522.1.ex7_6:117859794-117859817:-	CCGTCAGGTTCAAAACTCA
JAG1	JAG1_CCDS13112.1.ex19_20:10632856-10632879:-	CATCAGCCGTGTCTCAACG
	JAG1_CCDS13112.1.ex20_20:10633180-10633203:+	CGTGGACGCATCCCGGGTG
	JAG1_CCDS13112.1.ex24_20:10653465-10653488:-	GTCCCGCGTCACGGCCGGG
	JAG1_CCDS13112.1.ex24_20:10653513-10653536:+	TATGTGTCACACTCGTCGC
	JAG1_CCDS13112.1.ex24_20:10653623-10653646:-	GCCTCGGGTCAGTTCGAGT
SLC39A6	SLC39A6_CCDS42428.1.ex6_18:33703481-33703504:+	TCACCACTCAAAGTCCCAA
	SLC39A6_CCDS42428.1.ex6_18:33703497-33703520:-	TCCTTGTGGCACTGGCCGT
	SLC39A6_CCDS42428.1.ex6_18:33703543-33703566:-	TTAGTGCCTCTCATGAATC
	SLC39A6_CCDS42428.1.ex6_18:33703604-33703627:+	GCTATAAAACCACCAACCC
	SLC39A6_CCDS42428.1.ex7_18:33704481-33704504:-	AGACCTATTCATTACAAAT
PTPRM	PTPRM_CCDS11840.1.ex1_18:7774160-7774183:+	ATGAGCCGTATAGCACATG
	PTPRM_CCDS11840.1.ex6_18:7955170-7955193:-	GGCGTTGAGCTGTATCCAC
	PTPRM_CCDS11840.1.ex7_18:8069673-8069696:+	TCTAAATAGATCCCATGCG
	PTPRM_CCDS11840.1.ex7_18:8069728-8069751:+	TCGGCAAATCACTATCCGC
	PTPRM_CCDS11840.1.ex9_18:8085795-8085818:-	GTGCTAGCTCGGATGGTAA
CD58	CD58_CCDS44199.1.ex2_1:117078742-117078765:+	GTACATTATAAGTCCTCGA
	CD58_CCDS44199.1.ex2_1:117078756-117078779:-	AGCATTACAACAGCCATCG

	CD58_CCDS44199.1.ex2_1:117078807-117078830:-	TAACTTGTGCATTGACTAA
	CD58_CCDS44199.1.ex3_1:117087007-117087030:+	TGTTAAGTTGTAGATAGTG
	CD58_CCDS44199.1.ex3_1:117087145-117087168:+	AGGCACATTGCTTGGTACA
PTPRK	PTPRK_CCDS47473.1.ex19_6:128388703-128388726:-	CTCCTTTGGCTCCGCGCAA
	PTPRK_CCDS47473.1.ex19_6:128388736-128388759:-	GTGACAATCGGACCTACCA
	PTPRK_CCDS47473.1.ex19_6:128388816-128388839:+	TACGGTGCACCCCCACTCA
	PTPRK_CCDS47473.1.ex23_6:128411042-128411065:+	CGCGTAATGTTGTAACCCA
	PTPRK_CCDS47473.1.ex26_6:128561256-128561279:+	CTCTACATCCCCTAGACGG
NEO1	NEO1_CCDS10247.1.ex1_15:73408883-73408906:-	GGAGTGAACGTTCGAATGC
	NEO1_CCDS10247.1.ex1_15:73409176-73409199:+	CAGCGAAGCTCATAGTAGC
	NEO1_CCDS10247.1.ex2_15:73415064-73415087:+	GGGCTTTATCGCTGCGTAG
	NEO1_CCDS10247.1.ex7_15:73528750-73528773:+	CCGCTTCATCAAATTGACG
	NEO1_CCDS10247.1.ex13_15:73551097-73551120:+	GGTCTCAGGTCTTGATCGG
ICAM2	ICAM2_CCDS11657.1.ex1_17:62081117-62081140:-	ACAGCCACATTCAACAGCA
	ICAM2_CCDS11657.1.ex2_17:62082532-62082555:-	TCAAACATCTCCCATGACA
	ICAM2_CCDS11657.1.ex2_17:62082587-62082610:+	AGCAGAATCTTATCTAGAG
	ICAM2_CCDS11657.1.ex2_17:62082643-62082666:+	GGTGCTGCAGTTGACCTCG
	ICAM2_CCDS11657.1.ex2_17:62082668-62082691:-	GCTGGCGGTTGAGCCCAAA
ITGB5	ITGB5_CCDS3030.1.ex6_3:124527984-124528007:-	CAGAGTATCCGGTCTAAAG
	ITGB5_CCDS3030.1.ex9_3:124540245-124540268:-	GATGTGCCCCACATCGCAT
	ITGB5_CCDS3030.1.ex11_3:124567148-124567171:+	ACTTACCCAATGCACGGAT
	ITGB5_CCDS3030.1.ex11_3:124567178-124567201:-	TTTCTCCTACACGGCACCG
	ITGB5_CCDS3030.1.ex12_3:124578259-124578282:+	CCGAGAGGTGATGGACCGT

Table S 2 Pool 2: list 96 genes and 500 sgRNA sequences.

There were 25 non-targeting sgRNAs, 5 SCARB1-targeting sgRNAs, 5 CD81-targeting sgRNAs and 465 sgRNAs targeting 94 genes. Genes are listed in an order of proteomic abundance.

Pool 3		
Gene	sgRNA name	sgRNA sequence (19 bp)
Non-targeting 1	nontargeting-1	GTCCATGGGTGGAGTTACG
Non-targeting 2	nontargeting-2	GGACAAGTTGGACAGTTGC
Non-targeting 3	nontargeting-3	AACAATGGCTTCGTCGACT
Non-targeting 4	nontargeting-4	ACCTTATGTTCGCCGCCTC
Non-targeting 5	nontargeting-5	AGACACCGTAACTCGAATT
Non-targeting 6	nontargeting-6	CTCGTAGGTCAATCGCGGA
Non-targeting 7	nontargeting-7	GACGGTTCCGATAACCTTA
Non-targeting 8	nontargeting-8	GGATCTCTCAATATGGCGC
Non-targeting 9	nontargeting-9	AGCGCCGTTGTTGAGAAGA
Non-targeting 10	nontargeting-10	CCGCCGGCCCTAACTGACC
Non-targeting 11	nontargeting-11	TGACTAGACCCTTACGCGG
Non-targeting 12	nontargeting-12	ATCGGAGAGCCTGAAGTTG
Non-targeting 13	nontargeting-13	TCGGGCCGCTATATAGGCG
Non-targeting 14	nontargeting-14	AGAACTCAGCGGTCTCAAG
Non-targeting 15	nontargeting-15	ATGTCCGTTGTAGTCCTCG
Non-targeting 16	nontargeting-16	CCGGTCGCAATTGTTGCTA
Non-targeting 17	nontargeting-17	TGCCTATTCAGTAGGTCGT
Non-targeting 18	nontargeting-18	GTGCGGTAAGCGTGGCTAG
Non-targeting 19	nontargeting-19	CCATGGAGCCCAATCAATC
Non-targeting 20	nontargeting-20	GTGTAGTTCGGATTGATGA
Non-targeting 21	nontargeting-21	GGAATTGGAGACGATGCCG
Non-targeting 22	nontargeting-22	AGTGCCTCGTAGTTGTGGG
Non-targeting 23	nontargeting-23	CCTCGTTTCCACGTAATTC
Non-targeting 24	nontargeting-24	ATTTGATGTGGTGCTCACA
Non-targeting 25	nontargeting-25	ATTAGGTCGTTCTGTAACA
SCARB1	SCARB1_CCDS45008.1.ex7_12:125296469-125296492:-	GCTCTTCACGGTGTTCACG
	SCARB1_CCDS45008.1.ex8_12:125298861-125298884:+	ATGAAGGCACGTTCGCCGA
	SCARB1_CCDS45008.1.ex9_12:125299545-125299568:+	TAGTCGCTCTCCGAGCCGT
	SCARB1_CCDS45008.1.ex9_12:125299593-125299616:+	CGGTACTCGAGGAAGGACA

	SCARB1_CCDS45008.1.ex11_12:125348148-125348171:-	CCGTCGCTCATCAAGCAGC
CD81	CD81_CCDS7734.1.ex1_11:2411694-2411717:+	ACCACCAACCTCCTGTATC
	CD81_CCDS7734.1.ex1_11:2411744-2411767:-	GTGCACTCACCTACATAGA
	CD81_CCDS7734.1.ex2_11:2415376-2415399:+	GGCTGCTACGGGGCCATCC
	CD81_CCDS7734.1.ex3_11:2416244-2416267:-	TTGACAAAGCCCCAGATGC
	CD81_CCDS7734.1.ex4_11:2416641-2416664:-	CTTCACATCCTTGGCGATC
CD74	CD74_CCDS34276.1.ex3_5:149786762-149786785:+	GACTGTCAGTTTGTCCAGC
	CD74_CCDS34276.1.ex3_5:149786785-149786808:-	ACTTCCTGTACCAGCAGCA
	CD74_CCDS34276.1.ex3_5:149786857-149786880:-	GCCGCGGAGCCCTGTACAC
	CD74_CCDS34276.1.ex3_5:149786872-149786895:-	CCCACAGCAAGTGCAGCCG
	CD74_CCDS34276.1.ex4_5:149792174-149792197:+	GCTCACACATACCTCTCCG
EPHA1	EPHA1_CCDS5884.1.ex11_7:143095433-143095456:+	ATAGGTCAGGTTCGCCCCA
	EPHA1_CCDS5884.1.ex13_7:143096371-143096394:-	CATTACAGAGCTCCCGGGG
	EPHA1_CCDS5884.1.ex13_7:143096445-143096468:+	GGGGGCACGTGAGACAATG
	EPHA1_CCDS5884.1.ex13_7:143096479-143096502:+	CATCCGGTAGGAGCCGCTA
	EPHA1_CCDS5884.1.ex15_7:143098580-143098603:-	TCCAATTGGATCTACCGCG
CLEC4G	CLEC4G_CCDS12185.1.ex2_19:7794922-7794945:+	ACGATCACCAGGTGCGCGC
	CLEC4G_CCDS12185.1.ex2_19:7795012-7795035:+	TCGAAGGACAGCCACGACG
	CLEC4G_CCDS12185.1.ex3_19:7795272-7795295:+	CTCAGTGCGGACGTCCTCA
	CLEC4G_CCDS12185.1.ex3_19:7795317-7795340:-	CTTGACCCCGCAGTGACCC
	CLEC4G_CCDS12185.1.ex5_19:7795955-7795978:+	CCGCCGTCTGCTTCGAGGC
HPN	HPN_CCDS32993.1.ex2_19:35540400-35540423:-	ACCTGGGTACAGCGGCTCC
	HPN_CCDS32993.1.ex3_19:35550667-35550690:+	GTAGCCGGACTCAGCTGCG
	HPN_CCDS32993.1.ex4_19:35550789-35550812:+	TCCGAGCTGGACGTGCGAA
	HPN_CCDS32993.1.ex6_19:35551332-35551355:-	GTGTGCTCCATCATAGCGA
	HPN_CCDS32993.1.ex7_19:35551517-35551540:+	TTTCCCTGGTAGGCGGAAC
IL13RA1	IL13RA1_CCDS14573.1.ex1_X:117875097-117875120:+	TTTGGCGACAAACAAGATA
	IL13RA1_CCDS14573.1.ex2_X:117880960-117880983:+	GAGAGGATTTGTCTGCAAG
	IL13RA1_CCDS14573.1.ex2_X:117880987-117881010:-	ACTCTCATTGGTGCTACAC
	IL13RA1_CCDS14573.1.ex3_X:117883707-117883730:-	GTATAGTTAGTGTCGGGAC

	IL13RA1_CCDS14573.1.ex4_X:117892112-117892135:-	ACACTGTGTTGTTCAAAAC
CD151	CD151_CCDS7719.1.ex1_11:836235-836258:+	TGTACTGCTTGTAGCTGGC
	CD151_CCDS7719.1.ex1_11:836410-836433:-	AGGTTCCGACGCTCCTTGA
	CD151_CCDS7719.1.ex2_11:836792-836815:+	TTCTGCTGGAGATCATCGC
	CD151_CCDS7719.1.ex2_11:836818-836841:-	CTGGTAGTAGGCGTAGGCG
	CD151_CCDS7719.1.ex4_11:837503-837526:+	GAGTGGATCCGCTCACAGG
CD44	CD44_CCDS31455.1.ex1_11:35198133-35198156:-	AATACACCTGCAAAGCGGC
	CD44_CCDS31455.1.ex1_11:35198176-35198199:+	TACAGCATCTCTCGGACGG
	CD44_CCDS31455.1.ex1_11:35198206-35198229:-	GCTATTGAAAGCCTTGCAG
	CD44_CCDS31455.1.ex1_11:35198221-35198244:+	AATAGCACCTTGCCCACAA
	CD44_CCDS31455.1.ex1_11:35198265-35198288:+	CATCGGATTTGAGACCTGC
ENG	ENG_CCDS48029.1.ex9_9:130587534-130587557:+	CGATGAGCCAGGACACGTA
	ENG_CCDS48029.1.ex9_9:130587573-130587596:+	GGACGGCATCGAGATCCCC
	ENG_CCDS48029.1.ex10_9:130587962-130587985:+	AGAGCCATACCCGGCCGAG
	ENG_CCDS48029.1.ex10_9:130588074-130588097:-	CATGGGCCGCACGCTCGAG
	ENG_CCDS48029.1.ex13_9:130605436-130605459:-	CCACTAGCCAGGTCTCGAA
TXNDC15	TXNDC15_CCDS4180.1.ex0_5:134210198-134210221:+	TTCCCGTCCGCGGCGTGGA
	TXNDC15_CCDS4180.1.ex1_5:134223370-134223393:+	TGGCATGTTTTAGTTGCAG
	TXNDC15_CCDS4180.1.ex1_5:134223394-134223417:+	AGTGGTCGCTTATGGTCAG
	TXNDC15_CCDS4180.1.ex1_5:134223532-134223555:+	GACACCCAAGGCGATCACA
	TXNDC15_CCDS4180.1.ex2_5:134229228-134229251:-	GCACCACGGGGTGTAAAAC
CALR	CALR_CCDS12288.1.ex1_19:13049960-13049983:-	GTGTTTGGATTCGATCCAG
	CALR_CCDS12288.1.ex1_19:13050014-13050037:+	GGCAAGTTCTACGGTGACG
	CALR_CCDS12288.1.ex2_19:13050283-13050306:-	TTGCTGAAAGGCTCGAAAC
	CALR_CCDS12288.1.ex3_19:13050861-13050884:-	GCCACAGATGTCGGGACCT
	CALR_CCDS12288.1.ex4_19:13051141-13051164:+	CGGCTCCTTGGAAGACGAT
MYH9	MYH9_CCDS13927.1.ex27_22:36708092-36708115:-	ATTATCCACTATGCCGGCA
	MYH9_CCDS13927.1.ex29_22:36712657-36712680:-	CTATGAGCGGATGTTCCGC
	MYH9_CCDS13927.1.ex30_22:36714268-36714291:-	GGACGGGATTACGTCCAGA
	MYH9_CCDS13927.1.ex32_22:36716385-36716408:+	TACGGCTCCAACAGGAGAT

	MYH9_CCDS13927.1.ex36_22:36722638-36722661:+	CGCCACGTACGCCAGATAC
ATP1A1	ATP1A1_CCDS53351.1.ex4_1:116930773-116930796:+	CAGCCGTTGTAATCATAAC
	ATP1A1_CCDS53351.1.ex5_1:116931259-116931282:+	AAGCCCTTGTGATTCGAAA
	ATP1A1_CCDS53351.1.ex6_1:116931533-116931556:-	TCTGATTCACCAGTGAGCG
	ATP1A1_CCDS53351.1.ex7_1:116932063-116932086:-	TAGACAACAATACCACGTG
	ATP1A1_CCDS53351.1.ex7_1:116932294-116932317:-	AGCAAACCTTCCGGCACAT
HNRNPK	HNRNPK_CCDS6667.1.ex6_9:86586943-86586966:-	ACAGAATGCCTCCTGGTCG
	HNRNPK_CCDS6667.1.ex6_9:86586979-86587002:+	CCCGCATGGGAAATCCCAC
	HNRNPK_CCDS6667.1.ex9_9:86588837-86588860:+	GATTCGAGCGGGAGCTGGC
	HNRNPK_CCDS6667.1.ex11_9:86590362-86590385:+	TGTTGGGACATACCGCTCG
	HNRNPK_CCDS6667.1.ex12_9:86591958-86591981:-	TGTATCTTCTCCAGAATGC
SSR1	SSR1_CCDS4499.1.ex4_6:7301576-7301599:+	CAAATGGTCGTCCGCCCAT
	SSR1_CCDS4499.1.ex4_6:7301691-7301714:+	CTGGTAGTCCTGAGGATAA
	SSR1_CCDS4499.1.ex4_6:7301754-7301777:+	CTTGTTGGTAAAGCCTACC
	SSR1_CCDS4499.1.ex5_6:7303799-7303822:+	GTATAGTTGTATCTGCACT
	SSR1_CCDS4499.1.ex6_6:7310247-7310270:-	TCATTAGGCTTGTTAGCAG
SEC22B	SEC22B.1	CTAACAATGATCGCCCGAG
	SEC22B.2	ACAATGATCGCCCGAGTGG
	SEC22B.3	TGATCGCCCGAGTGGCGGA
	SEC22B.4	GATCGCCCGAGTGGCGGAC
	SEC22B.5	CGGGAGCCCGTCCGCCACT
TMED9	TMED9_CCDS4428.1.ex0_5:177019328-177019351:+	TACTTTCACATCGGAGAGA
	TMED9_CCDS4428.1.ex1_5:177019587-177019610:+	GCAGCTGTATGACAAGCAG
	TMED9_CCDS4428.1.ex1_5:177019629-177019652:-	ACAAACATGCCAAGCCCCG
	TMED9_CCDS4428.1.ex2_5:177020650-177020673:+	TCATCCTGGCCCGGCAGTA
	TMED9_CCDS4428.1.ex3_5:177021156-177021179:-	ATTGGCATGTTCACCTACC
DHRS7	DHRS7_CCDS9743.1.ex3_14:60619790-60619813:-	GAGCTTAACTACTTAGGGA
	DHRS7_CCDS9743.1.ex3_14:60619841-60619864:+	ATCCATGCACAGAGAACGC
	DHRS7_CCDS9743.1.ex4_14:60620720-60620743:+	ACCAGTGTCGGTCAGGTCA
	DHRS7_CCDS9743.1.ex5_14:60622808-60622831:+	TCCTCACCAATTCCACTCG

	DHRS7_CCDS9743.1.ex6_14:60631905-60631928:-	TACTATGGGCCGAGTGGCA
STT3A	STT3A_CCDS60998.1.ex0_11:125472808-125472831:-	GTAAGGTGGTACGTGACGA
	STT3A_CCDS60998.1.ex1_11:125474102-125474125:+	ATATCTCCCGATCTGTGGC
	STT3A_CCDS60998.1.ex2_11:125475586-125475609:-	CACTTAGCTGCCCAACAGA
	STT3A_CCDS60998.1.ex3_11:125476189-125476212:+	CATAGGTCTCGTCATGGGG
	STT3A_CCDS60998.1.ex5_11:125479340-125479363:-	GAGTAGAAACGCCCCGTCC
PGRMC2	PGRMC2_CCDS3739.2.ex1_4:129193653-129193676:+	ATATTCCATATGGACCCGC
	PGRMC2_CCDS3739.2.ex2_4:129208571-129208594:-	GCATCCTGCTCGCGGTCAA
	PGRMC2_CCDS3739.2.ex2_4:129208637-129208660:+	AGTCCCGCTTCTTCATGCG
	PGRMC2_CCDS3739.2.ex2_4:129208796-129208819:-	CGGCGTTGGCGCTTCTGAC
	PGRMC2_CCDS3739.2.ex2_4:129208904-129208927:-	ACGTGAAGCTAGGCACCCT
DHCR7	DHCR7_CCDS8200.1.ex2_11:71150065-71150088:-	CGAGTTTAACCCTCGGATC
	DHCR7_CCDS8200.1.ex3_11:71152341-71152364:+	TGTTGGCGCACCACAGCAG
	DHCR7_CCDS8200.1.ex3_11:71152411-71152434:+	AGCGTTTGCAAACCAGAGC
	DHCR7_CCDS8200.1.ex4_11:71153337-71153360:-	TTCTACCCGGCTACGTAGG
	DHCR7_CCDS8200.1.ex5_11:71155069-71155092:+	CGGCTTTCCTCGTTATAGG
RTN4	RTN4_CCDS1851.1.ex4_2:55209689-55209712:-	CTGCACGATAAAGGAACTC
	RTN4_CCDS1851.1.ex5_2:55214615-55214638:+	GACATCTCACCTGAATGGG
	RTN4_CCDS1851.1.ex5_2:55214638-55214661:-	CTATCCAGAAATCAGATGA
	RTN4_CCDS1851.1.ex5_2:55214671-55214694:-	TCAGCTTTAGGATATACAA
	RTN4_CCDS1851.1.ex5_2:55214747-55214770:+	TACTGTCAATGAAAGCAGC
STT3B	STT3B_CCDS2650.1.ex0_3:31574602-31574625:-	GCCGCCTTGTGCGCGCACT
	STT3B_CCDS2650.1.ex0_3:31574782-31574805:+	CATCCACGAGTTCGACCCG
	STT3B_CCDS2650.1.ex2_3:31621392-31621415:-	GCCGCTAAAAGTTGGTGCA
	STT3B_CCDS2650.1.ex2_3:31621507-31621530:+	ACATATCTCGGTCAGTAGC
	STT3B_CCDS2650.1.ex8_3:31661205-31661228:+	TGAGCATCAACCTACGACT
MIA3	MIA3_CCDS41470.1.ex1_1:222794493-222794516:+	ATACAGTGTTAATGTACCG
	MIA3_CCDS41470.1.ex3_1:222801439-222801462:-	TCTAATGAAGTAACGAGTC

	MIA3_CCDS41470.1.ex3_1:222802288-222802311:-	ATGAGTGGTGCACTACCCA
	MIA3_CCDS41470.1.ex3_1:222802388-222802411:+	TCAGTGGAGCATCAACGTG
	MIA3_CCDS41470.1.ex3_1:222803257-222803280:+	TGAAGATGACTCGTTCCAC
NCLN	NCLN_CCDS32869.1.ex0_19:3186152-3186175:-	TACACGGTGAACTCGTGCG
	NCLN_CCDS32869.1.ex1_19:3192528-3192551:-	TAGCCGCATGAGCACGCAG
	NCLN_CCDS32869.1.ex2_19:3193328-3193351:-	CTCCACGGCAAAGTACACG
	NCLN_CCDS32869.1.ex4_19:3198867-3198890:-	CACTCCAAAGGCGTCGTAG
	NCLN_CCDS32869.1.ex5_19:3201532-3201555:+	GCGCGGACTCCAACGGGAG
SERPINA1	SERPINA1_CCDS9925.1.ex3_14:94849012-94849035:-	ATCAACGATTACGTGGAGA
	SERPINA1_CCDS9925.1.ex3_14:94849045-94849068:-	AACTTCGGGGACACCGAAG
	SERPINA1_CCDS9925.1.ex3_14:94849183-94849206:+	TGGCTGGTTGAGGGTACGG
	SERPINA1_CCDS9925.1.ex3_14:94849231-94849254:+	AGCCTCCGGAATCTCCGTG
	SERPINA1_CCDS9925.1.ex3_14:94849435-94849458:+	GGTTGGGTGATCCTGATCA
C3	C3_CCDS32883.1.ex24_19:6707188-6707211:+	GATGAAACGGGTCCGGCGC
	C3_CCDS32883.1.ex29_19:6711023-6711046:-	CGAATGGACCGCGCCCACG
	C3_CCDS32883.1.ex30_19:6712259-6712282:+	GACGCACCGTGATGCTCAA
	C3_CCDS32883.1.ex33_19:6713504-6713527:+	TTCCCGTAGAGGAACCTAC
	C3_CCDS32883.1.ex39_19:6719328-6719351:+	AACATCCCCTTGCGCGTCG
CLU	CLU_CCDS47832.1.ex4_8:27462585-27462608:+	CTGCGGACGATGCGGGACT
	CLU_CCDS47832.1.ex4_8:27462704-27462727:+	GTCTATGATGCTGGACGCG
	CLU_CCDS47832.1.ex5_8:27463906-27463929:+	CGTGCGTAGAACTTCATGC
	CLU_CCDS47832.1.ex5_8:27463965-27463988:-	GTGTGCAATGAGACCATGA
	CLU_CCDS47832.1.ex5_8:27464020-27464043:-	GGATGCCCTAAATGAGACC
LGALS3BP	LGALS3BP_CCDS11759.1.ex0_17:76968538-76968561:-	AACTTCGAGGCCTTGACGC
	LGALS3BP_CCDS11759.1.ex0_17:76968767-76968790:-	CAGGTACTTCTACTCCCGA
	LGALS3BP_CCDS11759.1.ex2_17:76970853-76970876:-	GTCCAGTGCACGGGAACCG
	LGALS3BP_CCDS11759.1.ex3_17:76972073-76972096:-	GGCTTCGAGAACGCCACCC
	LGALS3BP_CCDS11759.1.ex3_17:76972106-76972129:+	GACGACGCTGGCATCAGTC

CTSD	CTSD_CCDS7725.1.ex4_11:1778624-1778647:+	ACGTTGTTGACGGAGATGC
	CTSD_CCDS7725.1.ex4_11:1778732-1778755:+	TCCACTTTGACACCGCCCA
	CTSD_CCDS7725.1.ex5_11:1780240-1780263:+	CCATAGTGGATGTCAAACG
	CTSD_CCDS7725.1.ex5_11:1780258-1780281:+	GAGGTACCATTCTTCACGT
	CTSD_CCDS7725.1.ex7_11:1782651-1782674:-	ATCCGCCGGACCATGTCGG
TF	TF_CCDS3080.1.ex1_3:133467334-133467357:-	GACGCTTTTCATATGGTCG
	TF_CCDS3080.1.ex1_3:133467420-133467443:-	CAGCAGCGACTTACCGCAA
	TF_CCDS3080.1.ex4_3:133474299-133474322:-	CCGAAGTATTGGTTAAGGG
	TF_CCDS3080.1.ex6_3:133475786-133475809:+	ACCGTCGTGGCCCGAAGTA
	TF_CCDS3080.1.ex7_3:133476762-133476785:+	CCATCCGGAATCTACGGGA
SCARB2	SCARB2_CCDS3577.1.ex6_4:77097047-77097070:-	TTATAGGTCACTTGACTGG
	SCARB2_CCDS3577.1.ex10_4:77116858-77116881:-	AGTGGGGCCATACACCTAC
	SCARB2_CCDS3577.1.ex10_4:77116888-77116911:-	CCTCAGAGGGGAGACCCCT
	SCARB2_CCDS3577.1.ex10_4:77116937-77116960:+	AATAGAACTGAGTATACAC
	SCARB2_CCDS3577.1.ex11_4:77134578-77134601:-	GTAGACCAGAGTATCGAGA
SLC25A4	SLC25A4_CCDS34114.1.ex0_4:186064605-186064628:-	AGTTTGACCCTCTCGATGG
	SLC25A4_CCDS34114.1.ex1_4:186065979-186066002:+	GTGAGAATCCCTAAGGAGC
	SLC25A4_CCDS34114.1.ex1_4:186066032-186066055:-	GGGAAGTAACGGATCACGT
	SLC25A4_CCDS34114.1.ex1_4:186066186-186066209:-	CAGCGGGTAGACAAAGCAA
	SLC25A4_CCDS34114.1.ex1_4:186066312-186066335:-	ACCCTGGTAGAGCCCCCTC
ACTN1	ACTN1_CCDS45129.1.ex9_14:69352260-69352283:+	TAGTCCTTCTGTCGCAGCA
	ACTN1_CCDS45129.1.ex13_14:69369255-69369278:-	ATCGTTGGAACTGCCCGAC
	ACTN1_CCDS45129.1.ex14_14:69371424-69371447:+	CCGTATTCAGATTTGTGAG
	ACTN1_CCDS45129.1.ex15_14:69376044-69376067:+	TCCCGTAGTCAATCAGCTC
	ACTN1_CCDS45129.1.ex18_14:69387714-69387737:+	GCCTTACCTTCGGCTCCGA
CALU	CALU_CCDS47703.1.ex0_7:128388714-128388737:-	CTGAGGCTCATGATGTACA
	CALU_CCDS56506.1.ex2_7:128394329-128394352:+	AATAGATGGCGACAAGGAC
	CALU_CCDS56506.1.ex2_7:128394415-128394438:+	ATGTAGAGCGACAGTGGAA
	CALU_CCDS56506.1.ex2_7:128394491-128394514:-	CTACCTAAAACGTAGCCGT

	CALU_CCDS47703.1.ex1_7:128394596-128394619:+	AATAGACGCGGATAAAGAT
KTN1	KTN1_CCDS41957.1.ex0_14:56078989-56079012:-	CGAGGTACACTCTCAGAGT
	KTN1_CCDS41957.1.ex2_14:56084783-56084806:+	TAAACCTGACCAAGTAGAA
	KTN1_CCDS41957.1.ex4_14:56094616-56094639:+	TCTCAATTGCAGTCAAGTA
	KTN1_CCDS41957.1.ex4_14:56094724-56094747:+	CGGTGTAAGCAGTTAACCC
	KTN1_CCDS41957.1.ex7_14:56101252-56101275:+	GAACTAAATAAACTACGCC
SLC16A1	SLC16A1_CCDS858.1.ex1_1:113460374-113460397:-	AAGCATCCCTTGAGAAAGC
	SLC16A1_CCDS858.1.ex1_1:113460431-113460454:-	GAGCCCTCATGCGACCAAT
	SLC16A1_CCDS858.1.ex1_1:113460528-113460551:-	ACTCTGGCCCCCCTCAATC
	SLC16A1_CCDS858.1.ex1_1:113460579-113460602:-	CGACCATTGGCCAACGGAC
	SLC16A1_CCDS858.1.ex3_1:113471747-113471770:+	ATCCATGACACTTCGCTGG
TMED4	TMED4_CCDS5493.1.ex2_7:44621070-44621093:-	CACTCCAATTCTACCAGGA
	TMED4_CCDS5493.1.ex2_7:44621093-44621116:+	GACAGATTTGATGGTCACC
	TMED4_CCDS5493.1.ex3_7:44621311-44621334:+	CAGCCTTACCTTGCCGTCG
	TMED4_CCDS5493.1.ex3_7:44621399-44621422:-	CAACTATCGTACCCAGATG
	TMED4_CCDS5493.1.ex4_7:44621672-44621695:-	GAGAAGCGCTGTTTCATCG
MAGT1	MAGT1_CCDS14436.2.ex7_X:77126352-77126375:+	ATCCTGTTGGTGAATGCAC
	MAGT1_CCDS14436.2.ex8_X:77130931-77130954:+	GACACACTGTCTATGCAGT
	MAGT1_CCDS14436.2.ex8_X:77130972-77130995:+	TAACGGAGTAATTTCTCGG
	MAGT1_CCDS14436.2.ex8_X:77130991-77131014:+	TGGGGCTTTCACAAGGCGA
	MAGT1_CCDS14436.2.ex9_X:77150892-77150915:-	AGCGAACATGGCAGCGCGT
ARSE	ARSE_CCDS14122.1.ex5_X:2867519-2867542:-	ATACCCGTCTCGTGGATGC
	ARSE_CCDS14122.1.ex5_X:2867645-2867668:-	ATGGGTGATTGCGCCCGCT
	ARSE_CCDS14122.1.ex6_X:2871274-2871297:+	AGAACACGGTAACCAATGC
	ARSE_CCDS14122.1.ex7_X:2873446-2873469:+	GAGACTGACCTGATCGCAC
	ARSE_CCDS14122.1.ex8_X:2876373-2876396:+	AGAAGGATGTTCGGTCGGG
GGCX	GGCX_CCDS1978.1.ex8_2:85781383-85781406:-	TAGCCTGCTGGTCGTGCAC
	GGCX_CCDS1978.1.ex10_2:85783321-85783344:+	ATAGTTCCAAAGGGGCACG
	GGCX_CCDS1978.1.ex10_2:85783352-85783375:-	CGGTCTGCTGAATGCCCAT
	GGCX_CCDS1978.1.ex11_2:85785653-85785676:+	CACATACCAGTATGGCAGC
	GGCX_CCDS1978.1.ex12_2:85786087-85786110:+	GCGTAGGGCATCCAGCAAG
IMPAD1	IMPAD1_CCDS6169.1.ex2_8:57890627-57890650:+	TTATGTATAACTCCTAGCA
	IMPAD1_CCDS6169.1.ex3_8:57892602-57892625:-	GACCCACTTGATGCTACAC

	IMPAD1_CCDS6169.1.ex3_8:57892653-57892676:-	GAAGTAACTACTCCTAAAG
	IMPAD1_CCDS6169.1.ex3_8:57892734-57892757:-	ATTAATACTGAGGAACACG
	IMPAD1_CCDS6169.1.ex4_8:57905857-57905880:-	TCCTCCACGAGAAGTCCAA
TMX2	TMX2_CCDS44601.1.ex0_11:57480234-57480257:+	TGCCCACCCAACGCGAAGA
	TMX2_CCDS44601.1.ex2_11:57505845-57505868:-	CAGGGCCCATATATAGGGG
	TMX2_CCDS44601.1.ex3_11:57506169-57506192:+	TGTGGAGTTCTTTGCCAAT
	TMX2_CCDS44601.1.ex4_11:57506438-57506461:+	TTTCAGATACAACTGTACA
	TMX2_CCDS44601.1.ex4_11:57506489-57506512:+	CTATACTGATGTTAGTACG
HM13	HM13_CCDS13182.1.ex0_20:30102441-30102464:+	AGGGCATCGCGCTGGCCTA
	HM13_CCDS13182.1.ex0_20:30102505-30102528:-	GCGCAGCGTACGGAGCGCA
	HM13_CCDS13182.1.ex2_20:30125971-30125994:+	CTGCCTCAGATATTCTCCC
	HM13_CCDS13182.1.ex2_20:30126015-30126038:-	CCCAGCACGAAGAAATACA
	HM13_CCDS13182.1.ex3_20:30132771-30132794:-	GTCGATTTGGAAAGCTGGC
APOB	APOB_CCDS1703.1.ex6_2:21237961-21237984:-	GACATGACTTTCCGGCACG
	APOB_CCDS1703.1.ex10_2:21242602-21242625:+	CCGGTCAGCGGATAGTAGG
	APOB_CCDS1703.1.ex11_2:21245824-21245847:-	TCCGGACTTCGCTAGGAGT
	APOB_CCDS1703.1.ex19_2:21255217-21255240:+	AACTCACTTGTTGACCGCG
	APOB_CCDS1703.1.ex26_2:21265319-21265342:-	GACCCGATTCAAGCACCTC
SUN2	SUN2_CCDS13978.1.ex11_22:39144716-39144739:-	TGCTGACGTGCCTGACGTA
	SUN2_CCDS13978.1.ex13_22:39146262-39146285:-	CGAAGCGCCGTCTCACGGG
	SUN2_CCDS13978.1.ex14_22:39146900-39146923:+	CCCACGTAGTCGTCCTCAG
	SUN2_CCDS13978.1.ex15_22:39147215-39147238:-	ACTGCATGGTGACGCCAAC
	SUN2_CCDS13978.1.ex15_22:39147263-39147286:-	GTCGCTGGTCCACGAGTCC
NUP210	NUP210_CCDS33704.1.ex24_3:13399874-13399897:-	GGTCATCGCCCTGTCGGTG
	NUP210_CCDS33704.1.ex32_3:13420419-13420442:-	CTATCTACGTGGTCGAACC
	NUP210_CCDS33704.1.ex33_3:13421187-13421210:+	GCAACTCGTACTGATCGGA
	NUP210_CCDS33704.1.ex34_3:13427828-13427851:-	GCCTATGACGTCTACCTGA
	NUP210_CCDS33704.1.ex39_3:13461584-13461607:-	ACGCGCGTTAACTTCACGC
TM9SF1	TM9SF1_CCDS41934.1.ex2_14:24662196-24662219:-	CCACACTTATAGCGTGCGC
	TM9SF1_CCDS41934.1.ex2_14:24662240-24662263:+	TCGTAACCCATCCAAGCTG
	TM9SF1_CCDS41934.1.ex2_14:24662268-24662291:+	ACGTCCCGCACTGAAACAT
	TM9SF1_CCDS41934.1.ex2_14:24662401-24662424:-	TAGATGACTTGCCAATCCG
	TM9SF1_CCDS41934.1.ex3_14:24663921-24663944:-	ATCCGCTTTCGGGAAAACG

SLC38A2	SLC38A2_CCDS8749.1.ex9_12:46760656-46760679:+	TTCAATGTTCGTTAATGCC
	SLC38A2_CCDS8749.1.ex9_12:46760712-46760735:+	GGTAGCTTGACATAGCTGG
	SLC38A2_CCDS8749.1.ex12_12:46764402-46764425:-	TTCCCTCTATGTAGCATCC
	SLC38A2_CCDS8749.1.ex13_12:46764617-46764640:+	AGGATCTACATCTGCATAA
	SLC38A2_CCDS8749.1.ex14_12:46764955-46764978:-	CAAGCTGCTCTGAAAAGGT
GLG1	GLG1_CCDS32485.1.ex18_16:74519765-74519788:+	CAATGCGGTAATCTGCACC
	GLG1_CCDS32485.1.ex19_16:74524970-74524993:-	TTCCGGATTACATCGAAAA
	GLG1_CCDS32485.1.ex19_16:74525009-74525032:-	GATCATCCTAAGCTGTCGG
	GLG1_CCDS32485.1.ex20_16:74526899-74526922:+	GGCTTCACGCGATCGCGGA
	GLG1_CCDS32485.1.ex20_16:74527021-74527044:+	TAAGTGCTTCTCGACACTG
VDAC1	VDAC1_CCDS4168.1.ex3_5:133316496-133316519:+	TCAAAATTCATCTGGTAGC
	VDAC1_CCDS4168.1.ex3_5:133316534-133316557:+	ACCTAGCACCAGAGCACCC
	VDAC1_CCDS4168.1.ex3_5:133316582-133316605:-	ATTAACCTGGGCTGCGACA
	VDAC1_CCDS4168.1.ex5_5:133326717-133326740:-	GGAATACCGACAATACACT
	VDAC1_CCDS4168.1.ex5_5:133326770-133326793:-	TCTGGAAACCAAGTACAGA
TMEM97	TMEM97_CCDS11226.2.ex0_17:26646367-26646390:-	CGACTGGGTAGAGCTCGCG
	TMEM97_CCDS11226.2.ex1_17:26652526-26652549:+	GTTTAGAAACCTGCTGAAG
	TMEM97_CCDS11226.2.ex1_17:26652560-26652583:+	TTCAAAGACCCACTGCTAC
	TMEM97_CCDS11226.2.ex1_17:26652585-26652608:-	AGGACTTAAACCAGGCTGG
	TMEM97_CCDS11226.2.ex1_17:26652651-26652674:+	CAACGTATGCCTTCCTCAA
GPI	GPI_CCDS12437.1.ex1_19:34857279-34857302:-	GTTCTTGGAGTAATCCACC
	GPI_CCDS12437.1.ex3_19:34859483-34859506:-	GTGCAGCACGGCTCGACCC
	GPI_CCDS12437.1.ex3_19:34859519-34859542:+	CGGTCAAACACACCCATCC
	GPI_CCDS12437.1.ex5_19:34868687-34868710:-	TGGAGACATACCAGACGCG
	GPI_CCDS12437.1.ex6_19:34869861-34869884:+	ACGAATGCAGAGACGGCGA
PON1	PON1_CCDS5638.1.ex3_7:94937414-94937437:-	GATGTATTTGGGTTTAGCG
	PON1_CCDS5638.1.ex4_7:94940850-94940873:+	ATCTGGATGGTTCACCACC
	PON1_CCDS5638.1.ex5_7:94944702-94944725:-	GATCCAACAGTGTTGGAAT
	PON1_CCDS5638.1.ex5_7:94944755-94944778:-	GCTTCAACCCCAACAGTCC

	PON1_CCDS5638.1.ex6_7:94946059-94946082:-	GAGATACTGCCTAATGGAC
YBX1	YBX1_CCDS470.1.ex0_1:43148349-43148372:+	CCGGGCGGCCTCACATCGG
	YBX1_CCDS470.1.ex1_1:43149079-43149102:+	GGTTTTGGGAACAGTAAAA
	YBX1_CCDS470.1.ex1_1:43149115-43149138:+	CGGATATGGTTTCATCAAC
	YBX1_CCDS470.1.ex4_1:43162321-43162344:+	CAGCAAATGTTACAGGTCC
	YBX1_CCDS470.1.ex4_1:43162417-43162440:-	TTTGCTGGTAATTGCGTGG
FGB	FGB_CCDS3786.1.ex1_4:155486973-155486996:-	GTCAAGGGGTCGATGACCA
	FGB_CCDS3786.1.ex1_4:155487031-155487054:+	CACCGCCCATCAGTGGAGG
	FGB_CCDS3786.1.ex2_4:155487637-155487660:+	AGGGGGTGTTGTGTCCTAC
	FGB_CCDS3786.1.ex2_4:155487669-155487692:+	CAAGAGGCTTTGCTACAAC
	FGB_CCDS3786.1.ex3_4:155488838-155488861:+	CGTGTGCTTCGTTCAATCC
SLC13A5	SLC13A5_CCDS11079.1.ex7_17:6606283-6606306:+	TCACTCGTTCATCTGGCCC
	SLC13A5_CCDS11079.1.ex7_17:6606314-6606337:+	TTGGGTCCCGTCCCGGTCA
	SLC13A5_CCDS11079.1.ex8_17:6607307-6607330:-	ATCAGTAACACGGCAACCA
	SLC13A5_CCDS11079.1.ex9_17:6609950-6609973:+	CGTAATTACCGTGCAGGCT
	SLC13A5_CCDS11079.1.ex9_17:6609973-6609996:-	CGCACGCTCCTCTGGGTGG
STEAP3	STEAP3_CCDS2125.1.ex0_2:120003194-120003217:+	TCCCTGGCCACACGCCTGG
	STEAP3_CCDS2125.1.ex0_2:120003263-120003286:+	GCCAGGCTGTTTCCCTCAG
	STEAP3_CCDS2125.1.ex0_2:120003437-120003460:-	ATTGGACTCACGATGCTGA
	STEAP3_CCDS2125.1.ex1_2:120005282-120005305:-	ATCTCCGAGACAGCACGCT
	STEAP3_CCDS2125.1.ex1_2:120005296-120005319:+	CGGAGATGGCGCTCGCCAT
CCDC90B	CCDC90B_CCDS66190.1.ex3_11:82984834-82984857:-	AGGAGCAGAGTAACAGATA
	CCDC90B_CCDS66190.1.ex3_11:82984853-82984876:-	ACTGGATATCAACTTAGAA
	CCDC90B_CCDS66190.1.ex3_11:82984877-82984900:+	TTATCTGCTCTGATTCGAC
	CCDC90B_CCDS66190.1.ex4_11:82984992-82985015:+	CATTAGTTGTTGCTTAACT
	CCDC90B_CCDS66190.1.ex5_11:82985680-82985703:-	AATCTGAGAGCAGAGAATG
SIGMAR1	SIGMAR1_CCDS6562.1.ex2_9:34637253-34637276:+	AGCACATACTCGGACAGCG
	SIGMAR1_CCDS6562.1.ex2_9:34637301-34637324:-	GTTCGTGAATGCGGGTGGC
	SIGMAR1_CCDS6562.1.ex2_9:34637370-34637393:-	TCGTCTGATCGTGGAGCTG
	SIGMAR1_CCDS6562.1.ex3_9:34637542-34637565:-	AGTTGGCGCGGCAGTACGC

	SIGMAR1_CCDS6562.1.ex3_9:34637561-34637584:+	CTGCGCTATCTCTTCGCGC
VTN	VTN_CCDS11229.1.ex3_17:26695988-26696011:+	GATATTTCGGGGGTAATCA
	VTN_CCDS11229.1.ex4_17:26696344-26696367:+	GGTGAAGGCGGCATCGATG
	VTN_CCDS11229.1.ex5_17:26696540-26696563:+	AGGGAACCGTTCTTGAGGT
	VTN_CCDS11229.1.ex5_17:26696640-26696663:+	GCCTTGAGTCTATCCCCTC
	VTN_CCDS11229.1.ex5_17:26696809-26696832:-	ACGGTCTATGACGATGGCG
GOLIM4	GOLIM4_CCDS3204.1.ex6_3:167747807-167747830:+	TGGCTGAAGGGTACACCTA
	GOLIM4_CCDS3204.1.ex8_3:167754711-167754734:+	GATCAGGTTTTCGAAGGCT
	GOLIM4_CCDS3204.1.ex11_3:167761304-167761327:-	TCTTTGCAGAGCCAACACG
	GOLIM4_CCDS3204.1.ex15_3:167812874-167812897:+	GTTAGCCGTACCTTGTAAC
	GOLIM4_CCDS3204.1.ex15_3:167812940-167812963:-	ACGCAGCTGCGGAAAGCCG
FNDC3A	FNDC3A_CCDS41886.1.ex3_13:49705389-49705412:-	GCGGCATCATAGTTGGGAC
	FNDC3A_CCDS41886.1.ex3_13:49705416-49705439:-	CGGGTGAGTACATATGACG
	FNDC3A_CCDS41886.1.ex8_13:49742822-49742845:+	TGATATACCTAATCCACCA
	FNDC3A_CCDS41886.1.ex10_13:49748666-49748689:+	TATCGGCCAGAAATGACTA
	FNDC3A_CCDS41886.1.ex11_13:49749599-49749622:+	TACAATGGAGTAAGCCCTC
LMAN2L	LMAN2L_CCDS2023.1.ex4_2:97399298-97399321:-	TGGACAAATTTGTGGGGCT
	LMAN2L_CCDS2023.1.ex5_2:97400144-97400167:-	CAAAGGATCGGATGCAGCC
	LMAN2L_CCDS2023.1.ex5_2:97400207-97400230:-	ACTTCAAAATCCATGGACA
	LMAN2L_CCDS2023.1.ex5_2:97400240-97400263:+	CCCAGTCTCTCAGGAAACA
	LMAN2L_CCDS2023.1.ex7_2:97405618-97405641:-	AACGTTCGAGTACTTGAAA
SDC2	SDC2_CCDS6272.1.ex1_8:97605791-97605814:+	ATGACTACGCTTCTGCGTC
	SDC2_CCDS6272.1.ex2_8:97614645-97614668:-	GTCGAGATGTTGTCAGCTC
	SDC2_CCDS6272.1.ex2_8:97614666-97614689:-	TCAACAGTATCTTTGGAAG
	SDC2_CCDS6272.1.ex2_8:97614741-97614764:-	AACGCACCTTTGTCTGAGC
	SDC2_CCDS6272.1.ex3_8:97620612-97620635:+	AAAATGGACCCAGCCGAAG
CLPTM1	CLPTM1_CCDS12651.1.ex3_19:45477771-45477794:+	CGATCTTGTGTATGGCGAC

	CLPTM1_CCDS12651.1.ex4_19:45480703-45480726:-	CGACTTACTCCGGGACATG
	CLPTM1_CCDS12651.1.ex6_19:45489776-45489799:+	CGTGGACGACCACACGCCG
	CLPTM1_CCDS12651.1.ex7_19:45490438-45490461:+	TGAAGTTCGACGCCGTGAG
	CLPTM1_CCDS12651.1.ex7_19:45490515-45490538:+	TACCCCATCAACGAGAGCC
PTTG1IP	PTTG1IP_CCDS68221.1.ex0_21:46271402-46271425:+	GGAGACCGCAATGGCCACG
	PTTG1IP_CCDS68221.1.ex0_21:46271452-46271475:+	TGCACCTCACAGGAAGCGT
	PTTG1IP_CCDS13715.1.ex0_21:46271518-46271541:+	CGGGTTTTCTTCTTTAAAC
	PTTG1IP_CCDS13715.1.ex2_21:46276237-46276260:-	ATCACCATGTCGGTAGTCG
	PTTG1IP_CCDS13715.1.ex3_21:46281089-46281112:-	AAATTGAGCTCTGCACGCT
ITIH2	ITIH2_CCDS31141.1.ex3_10:7751025-7751048:+	TCTACTATTACTTCTCGGA
	ITIH2_CCDS31141.1.ex4_10:7755191-7755214:+	GGCCGAGCTCTTTATGCAC
	ITIH2_CCDS31141.1.ex5_10:7759654-7759677:+	TTCGAACTTCACTACCAGG
	ITIH2_CCDS31141.1.ex6_10:7762823-7762846:+	GTGTGGGTTATCGAACCAC
	ITIH2_CCDS31141.1.ex6_10:7762883-7762906:-	GACCGGAACACCATCGAAA
SLC44A1	SLC44A1_CCDS6763.1.ex2_9:108072052-108072075:+	TGTCAGGATACGACAGCTA
	SLC44A1_CCDS6763.1.ex2_9:108072138-108072161:-	CAAGTCTACTCACTTCCGC
	SLC44A1_CCDS6763.1.ex3_9:108097908-108097931:+	GTGTGTAGCAGCGTGTCCA
	SLC44A1_CCDS6763.1.ex4_9:108110718-108110741:-	AAATTTACCTCGCTGGAAC
	SLC44A1_CCDS6763.1.ex8_9:108125152-108125175:+	TCGCCTTGTTCCACGTAGC
SLC1A4	SLC1A4_CCDS1879.1.ex1_2:65228607-65228630:+	TGCAGCTTTCCGTACGGTA
	SLC1A4_CCDS1879.1.ex3_2:65237834-65237857:-	GGAATTGAAGAAACGGATG
	SLC1A4_CCDS1879.1.ex3_2:65237852-65237875:+	TCCCTCAACGAGGCGACGA
	SLC1A4_CCDS1879.1.ex5_2:65245218-65245241:-	TCAATGCACTTCATCATAG
	SLC1A4_CCDS1879.1.ex5_2:65245268-65245291:+	GCAGGTTTATTCTCCCCAT
SLC31A1	SLC31A1_CCDS6789.1.ex1_9:116019392-116019415:-	AGCCAAAGTAGAAGGTCAT
	SLC31A1_CCDS6789.1.ex1_9:116019440-116019463:+	GTTTGGTGATCAATACAGC

	SLC31A1_CCDS6789.1.ex2_9:116021002-116021025:+	TACTAGCAATGTTCTATGA
	SLC31A1_CCDS6789.1.ex2_9:116021077-116021100:+	ACAATTCCATGCCTGTCCC
	SLC31A1_CCDS6789.1.ex2_9:116021094-116021117:-	GATGGTTCCATTTGGTCCT
F2	F2_CCDS31476.1.ex4_11:46744744-46744767:+	TCTGGGTACGAACTACCGA
	F2_CCDS31476.1.ex5_11:46744946-46744969:+	CATCCTGGGGCCGACCTAC
	F2_CCDS31476.1.ex6_11:46747428-46747451:+	CGATGACTCCACGCTCCGA
	F2_CCDS31476.1.ex6_11:46747490-46747513:-	GTACTGCTGCCCCCGATCA
	F2_CCDS31476.1.ex6_11:46747686-46747709:-	CGCAGTACCCAAAGTCGCC
CLDND1	CLDND1_CCDS2930.1.ex1_3:98235916-98235939:-	TTATATCCCACCATTGCCA
	CLDND1_CCDS2930.1.ex1_3:98235984-98236007:+	TCAAACCTAAACTCACAAA
	CLDND1_CCDS2930.1.ex1_3:98235997-98236020:+	CACAAAAGGTAAAAGGAAC
	CLDND1_CCDS2930.1.ex2_3:98237716-98237739:+	AAAGTACTCACAGGTCCTA
	CLDND1_CCDS2930.1.ex2_3:98237748-98237771:-	ATCCCGGAAACCACAATAG
ITFG3	ITFG3_CCDS10402.1.ex0_16:304552-304575:-	AAGAAAAAACGCCGCCGCT
	ITFG3_CCDS10402.1.ex0_16:304588-304611:-	GACGACGAACACCACAAAA
	ITFG3_CCDS10402.1.ex0_16:304614-304637:+	CGTCATCCCGTGTCCAGAC
	ITFG3_CCDS10402.1.ex0_16:304658-304681:+	GGATAGACTACAGTGCCGC
	ITFG3_CCDS10402.1.ex1_16:309576-309599:+	GCCGATCCTGTGTGGACGA
MMGT1	MMGT1_CCDS14653.1.ex0_X:135047290-135047313:+	CGATGATTAAATACATAAA
	MMGT1_CCDS14653.1.ex1_X:135049573-135049596:-	GCAGGAGAGTTTAAAGACA
	MMGT1_CCDS14653.1.ex1_X:135049595-135049618:+	ATATGAACTATACCGTAAC
	MMGT1_CCDS14653.1.ex1_X:135049630-135049653:-	ATAGTTCTTCAGACACTTC
	MMGT1_CCDS14653.1.ex2_X:135053202-135053225:+	ATGACCAACTTACATCTAT
BRI3BP	BRI3BP_CCDS9262.1.ex0_12:125478446-125478469:-	GTTCTTCTCCGCGCCGCCG
	BRI3BP_CCDS9262.1.ex0_12:125478476-125478499:-	GGAGAAGGTGTTGACCGTG
	BRI3BP_CCDS9262.1.ex0_12:125478512-125478535:-	CACGTTGTCCTCGCCGAAC
	BRI3BP_CCDS9262.1.ex0_12:125478527-125478550:+	AACGTGCGCGCCGCTCAGA

	BRI3BP_CCDS9262.1.ex2_12:125509549-125509572:-	CGAGGCTGGGCTGAAATAC
ABHD12	ABHD12_CCDS13172.1.ex9_20:25300835-25300858:+	GGTACCTGCGTTCCCATGC
	ABHD12_CCDS13172.1.ex11_20:25319897-25319920:+	GACATAGTTTGATGAGAAA
	ABHD12_CCDS13172.1.ex11_20:25319971-25319994:-	TTGCAGGCGAAAGGGCGTG
	ABHD12_CCDS13172.1.ex12_20:25371156-25371179:+	GCCCGCTTCATTCCCGCGT
	ABHD12_CCDS13172.1.ex12_20:25371209-25371232:-	AAGCAGAACCTACGCCTGA
FNDC3B	FNDC3B_CCDS3217.1.ex3_3:171965333-171965356:+	AGTACTGGAGTCCGCCGGG
	FNDC3B_CCDS3217.1.ex3_3:171965487-171965510:-	TATCTCCAGGTCCGGTAAC
	FNDC3B_CCDS3217.1.ex8_3:172025160-172025183:-	CCCTTTACGGAATTGTACA
	FNDC3B_CCDS3217.1.ex10_3:172046837-172046860:+	TGGCCGCTCGAAACGACAT
	FNDC3B_CCDS3217.1.ex13_3:172052814-172052837:+	CTACCAGACCGCTTGTCAA
RAB5C	RAB5C_CCDS11419.1.ex2_17:40280374-40280397:-	ATTTGCACGGGCCAAGAAC
	RAB5C_CCDS11419.1.ex3_17:40280705-40280728:-	GGCCCCCATGTACTATCGG
	RAB5C_CCDS11419.1.ex3_17:40280725-40280748:-	CAGGAGCGGTATCACAGCC
	RAB5C_CCDS11419.1.ex3_17:40280770-40280793:+	CTTGACTGTTGTGTCATCC
	RAB5C_CCDS11419.1.ex4_17:40282389-40282412:-	TCGTCCTCCGCTTTGTCAA
RPN2	RPN2_CCDS13291.1.ex3_20:35827592-35827615:+	ACTGCTCGTCTCAGCAAGG
	RPN2_CCDS13291.1.ex5_20:35833175-35833198:+	TTGCTCGCCTGGATGAACT
	RPN2_CCDS13291.1.ex5_20:35833289-35833312:-	TAGGTACCTCCTTAATGGA
	RPN2_CCDS13291.1.ex6_20:35835686-35835709:-	GAAGATCGCGTTCATCAGC
	RPN2_CCDS13291.1.ex6_20:35835818-35835841:+	GCTTCCGACACTCATGAAC
CERS2	CERS2_CCDS973.1.ex4_1:150939923-150939946:+	AATCATGTAGTACCAATAC
	CERS2_CCDS973.1.ex6_1:150940312-150940335:+	GCCGGCAATGAAGGCAATC
	CERS2_CCDS973.1.ex8_1:150940971-150940994:+	TGGTGTAGCCACGTACCTG
	CERS2_CCDS973.1.ex9_1:150941397-150941420:+	AAAGAAGTATCGAACGATG
	CERS2_CCDS973.1.ex9_1:150941430-150941453:-	CTCTATATCACGCTGCCCC
A2M	A2M_CCDS44827.1.ex18_12:9246130-9246153:+	AACATGCACCAGGCGTGCA
	A2M_CCDS44827.1.ex22_12:9252012-9252035:+	GAATCCCCAATCACGTCCC

	A2M_CCDS44827.1.ex22_12:9252049-9252072:+	AGATGAGCAACCGAGCGAC
	A2M_CCDS44827.1.ex24_12:9254225-9254248:+	ACCCACTGGTAGCCGTAAC
	A2M_CCDS44827.1.ex25_12:9256894-9256917:+	TGCTCATCCGTGGTAGCAT
CASC4	CASC4_CCDS10108.1.ex0_15:44581379-44581402:-	CACTTCGGTGCGCTGGACC
	CASC4_CCDS10108.1.ex0_15:44581457-44581480:+	AAGAAACAGATCGACCAGA
	CASC4_CCDS10108.1.ex0_15:44581470-44581493:+	ACCAGAAGGAGGCCGACTA
	CASC4_CCDS10108.1.ex0_15:44581526-44581549:-	ATCCTCGCATCTCTTCCCG
	CASC4_CCDS10108.1.ex1_15:44615173-44615196:+	AACAACATATCGTATCAGA
TMX4	TMX4_CCDS13101.1.ex4_20:7980371-7980394:+	AGTTACGTTAGAGAAGCCG
	TMX4_CCDS13101.1.ex4_20:7980463-7980486:+	GATTCCTGGGCCACGATAA
	TMX4_CCDS13101.1.ex6_20:7990845-7990868:-	TAGATGTCATTCAAGAACC
	TMX4_CCDS13101.1.ex6_20:7990949-7990972:-	CTCTTATAGTTACGCCCCA
	TMX4_CCDS13101.1.ex7_20:8000112-8000135:-	ACCGCCTCCAACTGGACGC
PTGFRN	PTGFRN_CCDS890.1.ex1_1:117484552-117484575:+	GGGCGAGATCCTGTTAAGG
	PTGFRN_CCDS890.1.ex1_1:117484571-117484594:+	CGGACTGCCAACGACGCCG
	PTGFRN_CCDS890.1.ex2_1:117487629-117487652:+	GCGAGTGGATCGCCGAGCA
	PTGFRN_CCDS890.1.ex3_1:117491888-117491911:+	AGACCGAGCCGATGACGTC
	PTGFRN_CCDS890.1.ex3_1:117491963-117491986:-	TCAAGCCGCGCCAACACGC
ATP6AP1	ATP6AP1_CCDS35451.1.ex1_X:153657406-153657429:+	CTGCGGCCGACACTCATGA
	ATP6AP1_CCDS35451.1.ex1_X:153657456-153657479:+	ACCTACTTAGATCCCGCCC
	ATP6AP1_CCDS35451.1.ex3_X:153660639-153660662:+	GGTGCTTCCTGCCGTCGAC
	ATP6AP1_CCDS35451.1.ex5_X:153662017-153662040:+	GAAGATGTCCCATACACAG
	ATP6AP1_CCDS35451.1.ex6_X:153662555-153662578:+	GCCCGTGATGTAGCCGTGG
GDF15	GDF15_CCDS12376.1.ex0_19:18497096-18497119:-	AAACTTGCGCGGCTCGCCT
	GDF15_CCDS12376.1.ex0_19:18497160-18497183:+	GAGTTGCGGAAACGCTACG
	GDF15_CCDS12376.1.ex0_19:18497223-18497246:+	TCGAACACCGACCTCGTCC
	GDF15_CCDS12376.1.ex0_19:18497254-18497277:+	TCCGGATACTCACGCCAGA
	GDF15_CCDS12376.1.ex1_19:18499117-18499140:-	CCGAGAGATACGCAGGTGC
HSD11B1	HSD11B1_CCDS1489.1.ex1_1:209879178-209879201:+	TTGTCACAGGGGCCAGCAA

	HSD11B1_CCDS1489.1.ex1_1:209879216-209879239:+	GCTTATCATCTGGCGAAGA
	HSD11B1_CCDS1489.1.ex2_1:209880123-209880146:-	GCAACAAATTGCTCTGCGA
	HSD11B1_CCDS1489.1.ex2_1:209880142-209880165:+	GCCCAAGCAGGAAAGCTCA
	HSD11B1_CCDS1489.1.ex3_1:209880357-209880380:-	TTCCATGCTTTTGCGCACA
SLCO1B1	SLCO1B1_CCDS8685.1.ex3_12:21329762-21329785:-	TTAATTAAACAAGTGGATA
	SLCO1B1_CCDS8685.1.ex5_12:21331884-21331907:+	CAATCATTGGCTTTACCCT
	SLCO1B1_CCDS8685.1.ex6_12:21349891-21349914:+	AACTCCTACTGATTCTCGA
	SLCO1B1_CCDS8685.1.ex8_12:21355438-21355461:-	ACATTCCACTTGCAAAAAT
	SLCO1B1_CCDS8685.1.ex8_12:21355483-21355506:+	AATTCAAACTGAACACCGT
ORM1	ORM1_CCDS6803.1.ex1_9:117085950-117085973:+	AAGTGGTTTTATATCGCAT
	ORM1_CCDS6803.1.ex2_9:117086303-117086326:-	GGTGGTGTTATAGATGCAC
	ORM1_CCDS6803.1.ex3_9:117087065-117087088:-	AATGCTCTTGGCCTCCCAC
	ORM1_CCDS6803.1.ex3_9:117087088-117087111:+	GCTCACTTGCTGATCCTCA
	ORM1.5	CACCTACCTGAATGTCCAG
IFITM3	IFITM3_CCDS41585.1.ex1_11:320575-320598:+	GGCGAATGCTATGAAGCCC
	IFITM3_CCDS41585.1.ex1_11:320608-320631:+	CATGAAGAGGGTGTTGAAC
	IFITM3_CCDS41585.1.ex1_11:320633-320656:-	CGTGCCCGACCATGTCGTC
	IFITM3_CCDS41585.1.ex1_11:320653-320676:+	GGAGGTCTCGCTGCGGATG
	IFITM3_CCDS41585.1.ex1_11:320673-320696:+	GGATCACGGTGGACGTCGG
GJB1	GJB1_CCDS14408.1.ex0_X:70443771-70443794:-	AGGGACCACAGCCGCACAT
	GJB1_CCDS14408.1.ex0_X:70443871-70443894:+	CTACGGCTTGAGGGCCATG
	GJB1_CCDS14408.1.ex0_X:70443920-70443943:+	GGCACAAGGTCCACATCTC
	GJB1_CCDS14408.1.ex0_X:70443952-70443975:+	TGGACCTATGTCATCAGCG
	GJB1_CCDS14408.1.ex0_X:70444030-70444053:+	GGCTATGCCATGGTGCGGC
ABCB11	ABCB11_CCDS46444.1.ex16_2:169833142-169833165:-	TACAAGTTGGATCGAATCA
	ABCB11_CCDS46444.1.ex18_2:169842760-169842783:-	TCTTGTGTTCGCCCAGCGT
	ABCB11_CCDS46444.1.ex20_2:169850327-169850350:+	GGTCATGCGCTGAATGAAA

	ABCB11_CCDS46444.1.ex22_2:169853154-169853177:-	CTGTCGCAGTACTTATCAC
	ABCB11_CCDS46444.1.ex23_2:169869923-169869946:+	AATGAGTAGCACGCCTGGC
TSPAN8	TSPAN8_CCDS8999.1.ex4_12:71531954-71531977:-	AGGTGGCGACAGGTATCCT
	TSPAN8_CCDS8999.1.ex5_12:71533516-71533539:+	TATAGCACCGCAGCATCCC
	TSPAN8_CCDS8999.1.ex5_12:71533538-71533561:+	AAGCCCAGAATCATGATGA
	TSPAN8_CCDS8999.1.ex5_12:71533574-71533597:+	AATATGTCCACAGCAACGT
	TSPAN8_CCDS8999.1.ex6_12:71537957-71537980:-	GATCCTAGCATTAGCAATA

Table S 3 Pool 3: list of 94 genes and 500 sgRNA sequences.

There were 25 non-targeting sgRNAs, 5 SCARB1-targeting sgRNAs, 5 CD81-targeting sgRNAs and 465 sgRNAs targeting 94 genes. Genes are listed in an order of proteomic abundance.

Pool 4		
Gene	sgRNA name	sgRNA sequence (19 bp)
Non-targeting 1	nontargeting-1	GTCCATGGGTGGAGTTACG
Non-targeting 2	nontargeting-2	GGACAAGTTGGACAGTTGC
Non-targeting 3	nontargeting-3	AACAATGGCTTCGTCGACT
Non-targeting 4	nontargeting-4	ACCTTATGTTCGCCGCCTC
Non-targeting 5	nontargeting-5	AGACACCGTAACTCGAATT
Non-targeting 6	nontargeting-6	CTCGTAGGTCAATCGCGGA
Non-targeting 7	nontargeting-7	GACGGTTCCGATAACCTTA
Non-targeting 8	nontargeting-8	GGATCTCTCAATATGGCGC
Non-targeting 9	nontargeting-9	AGCGCCGTTGTTGAGAAGA
Non-targeting 10	nontargeting-10	CCGCCGGCCCTAACTGACC
Non-targeting 11	nontargeting-11	TGACTAGACCCTTACGCGG
Non-targeting 12	nontargeting-12	ATCGGAGAGCCTGAAGTTG
Non-targeting 13	nontargeting-13	TCGGGCCGCTATATAGGCG
Non-targeting 14	nontargeting-14	AGAACTCAGCGGTCTCAAG
Non-targeting 15	nontargeting-15	ATGTCCGTTGTAGTCCTCG
Non-targeting 16	nontargeting-16	CCGGTCGCAATTGTTGCTA
Non-targeting 17	nontargeting-17	TGCCTATTCAGTAGGTCGT
Non-targeting 18	nontargeting-18	GTGCGGTAAGCGTGGCTAG
Non-targeting 19	nontargeting-19	CCATGGAGCCCAATCAATC
Non-targeting 20	nontargeting-20	GTGTAGTTCGGATTGATGA
Non-targeting 21	nontargeting-21	GGAATTGGAGACGATGCCG
Non-targeting 22	nontargeting-22	AGTGCCTCGTAGTTGTGGG
Non-targeting 23	nontargeting-23	CCTCGTTTCCACGTAATTC
Non-targeting 24	nontargeting-24	ATTTGATGTGGTGCTCACA
Non-targeting 25	nontargeting-25	ATTAGGTCGTTCTGTAACA
SCARB1	SCARB1_CCDS45008.1.ex7_12:125296469-125296492:-	GCTCTTCACGGTGTTCACG
	SCARB1_CCDS45008.1.ex8_12:125298861-125298884:+	ATGAAGGCACGTTCGCCGA
	SCARB1_CCDS45008.1.ex9_12:125299545-125299568:+	TAGTCGCTCTCCGAGCCGT
	SCARB1_CCDS45008.1.ex9_12:125299593-125299616:+	CGGTACTCGAGGAAGGACA

	SCARB1_CCDS45008.1.ex11_12:125348148-125348171:-	CCGTCGCTCATCAAGCAGC
CD81	CD81_CCDS7734.1.ex1_11:2411694-2411717:+	ACCACCAACCTCCTGTATC
	CD81_CCDS7734.1.ex1_11:2411744-2411767:-	GTGCACTCACCTACATAGA
	CD81_CCDS7734.1.ex2_11:2415376-2415399:+	GGCTGCTACGGGGCCATCC
	CD81_CCDS7734.1.ex3_11:2416244-2416267:-	TTGACAAAGCCCCAGATGC
	CD81_CCDS7734.1.ex4_11:2416641-2416664:-	CTTCACATCCTTGGCGATC
TSPAN9	TSPAN9_CCDS8520.1.ex1_12:3387626-3387649:-	GCAAAGTTGCCTTGGGACA
	TSPAN9_CCDS8520.1.ex1_12:3387696-3387719:+	GCCATAGGCACCATTGTCA
	TSPAN9_CCDS8520.1.ex1_12:3387768-3387791:-	GACAGAACTTACGCTGAGG
	TSPAN9_CCDS8520.1.ex3_12:3389592-3389615:+	ACCACACCGAGAACAACGT
	TSPAN9_CCDS8520.1.ex4_12:3390376-3390399:+	TGTCACTGACTACACAGAC
CD47	CD47_CCDS43125.1.ex6_3:107789962-107789985:+	ACAGGAGTATAGCAAAAAT
	CD47_CCDS43125.1.ex7_3:107798908-107798931:-	GTGATGCTGTCTCACACAC
	CD47_CCDS43125.1.ex7_3:107798928-107798951:+	CTCTTATCCATCTTCAAAG
	CD47_CCDS43125.1.ex7_3:107798982-107799005:+	TTTGCACTACTAAAGTCAG
	CD47_CCDS43125.1.ex7_3:107799063-107799086:-	TACTGAAGTATACGTAAAG
SERPING1	SERPING1_CCDS7962.1.ex1_11:57367458-57367481:+	CTATTCGTTGAACCCATCC
	SERPING1_CCDS7962.1.ex1_11:57367695-57367718:+	TTGGAGAGTCATTCAACAG
	SERPING1_CCDS7962.1.ex1_11:57367835-57367858:-	GTCTTACCGAGCAGGACCT
	SERPING1_CCDS7962.1.ex2_11:57369584-57369607:+	TGAAGGGCTTCACGACCAA
	SERPING1_CCDS7962.1.ex3_11:57373500-57373523:-	GTCCGAGAGGCATTCACAA
FXYD1	FXYD1_CCDS12445.1.ex1_19:35631519-35631542:-	CCGCTCACCGTAAGTGAAC
	FXYD1_CCDS12445.1.ex2_19:35632037-35632060:+	ACCAGTCCCTGCAGATCGG
	FXYD1_CCDS12445.1.ex2_19:35632053-35632076:+	CGGAGGCCTCGTCATCGCC
	FXYD1_CCDS12445.1.ex2_19:35632071-35632094:-	CCCAGGATGAAGAGGATCC
	FXYD1_CCDS12445.1.ex2_19:35632087-35632110:-	CAGCACGATGAGGATGCCC

ABCC6	ABCC6_CCDS10568.1.ex15_16:16276732-16276755:-	GGCTGTTGTCGGTCCAGTG
	ABCC6_CCDS10568.1.ex17_16:16281043-16281066:-	CGGGTGTCCTTTGACCGTC
	ABCC6_CCDS10568.1.ex18_16:16282818-16282841:-	ACCCCCCAGGTCGCACTGG
	ABCC6_CCDS10568.1.ex21_16:16291900-16291923:-	CTGCCTCTCGTCTGGATCG
	ABCC6_CCDS10568.1.ex23_16:16297388-16297411:+	TCTTGCCGTAGGAAGGGCT
ABCA6	ABCA6_CCDS11683.1.ex22_17:67106962-67106985:-	CTTGTATATACTTTGCCAC
	ABCA6_CCDS11683.1.ex27_17:67111582-67111605:-	GGTCAAATCACGGCAATCC
	ABCA6_CCDS11683.1.ex30_17:67121163-67121186:-	GAATCTTAGATTATCAAAC
	ABCA6_CCDS11683.1.ex32_17:67125796-67125819:-	CCCAAATTATAGTCATGAC
	ABCA6_CCDS11683.1.ex35_17:67132250-67132273:+	AAGTGGACTGTTATATCCC
SLC22A7	SLC22A7_CCDS4892.1.ex0_6:43266233-43266256:-	ACCCGGCAGGGCACATCGG
	SLC22A7_CCDS4892.1.ex1_6:43267177-43267200:+	TTCTTCGCCGGTGTGCTGG
	SLC22A7_CCDS4892.1.ex1_6:43267213-43267236:+	GGATATCTGTCCGACAGGT
	SLC22A7_CCDS4892.1.ex2_6:43267388-43267411:+	GTAGCCTACGTGAGTACCC
	SLC22A7_CCDS4892.1.ex2_6:43267432-43267455:-	AACATTACATAGCTGACGG
TOR1AIP2	TOR1AIP2_CCDS1334.1.ex1_1:179816683-179816706:-	TGAGACAGATTAATAAAAA
	TOR1AIP2_CCDS1334.1.ex2_1:179819998-179820021:-	AGAGGGTGAGGATACACTG
	TOR1AIP2_CCDS1334.1.ex2_1:179820095-179820118:-	TCCCTAAGGAAGCGAGTGA
	TOR1AIP2_CCDS1334.1.ex2_1:179820192-179820215:-	AATCTGGGTAAAGAACCCT
	TOR1AIP2_CCDS1334.1.ex2_1:179820373-179820396:+	TGGTCTTTGCTAAGACCAC
APOH	APOH_CCDS11663.1.ex3_17:64216793-64216816:-	GTGTTTATAAGCCATCAGC
	APOH_CCDS11663.1.ex3_17:64216816-64216839:+	AGTGTTGCAAACGTAGGTA
	APOH_CCDS11663.1.ex4_17:64219805-64219828:+	CAGACTTACGAGCACAGAC
	APOH_CCDS11663.1.ex5_17:64222180-64222203:+	TCAAAAGTCGTATAGCGTA

	APOH_CCDS11663.1.ex6_17:64224199-64224222:-	CGGGCTATGTGTCCCGAGG
CD82	CD82_CCDS31469.1.ex1_11:44621763-44621786:+	TTCATCTCTGTCCTGCGTA
	CD82_CCDS31469.1.ex2_11:44626616-44626639:-	GCCCCCATCCTAAGCGAGC
	CD82_CCDS31469.1.ex2_11:44626637-44626660:-	CCCACGCCGATGAAGACAT
	CD82_CCDS31469.1.ex2_11:44626689-44626712:+	TGCATCGGCGCCGTCAACG
	CD82_CCDS31469.1.ex3_11:44636862-44636885:+	GACTACAACAGCAGTCGCG
AMBP	AMBP_CCDS6800.1.ex6_9:116836304-116836327:+	GAGCCTTTACCGTAGAGCT
	AMBP_CCDS6800.1.ex6_9:116836341-116836364:-	AGAAATTCAGCCGCCATCA
	AMBP_CCDS6800.1.ex8_9:116838870-116838893:+	GCTCACCGCCAACGAGTGC
	AMBP_CCDS6800.1.ex8_9:116838977-116839000:+	GCAGGTGGAACCGATGGCC
	AMBP_CCDS6800.1.ex8_9:116838995-116839018:-	TATGGGAAGTGGTACAACC
AHSG	AHSG_CCDS3278.1.ex1_3:186333493-186333516:+	GAGATTGAAATAGACACCC
	AHSG_CCDS3278.1.ex2_3:186334254-186334277:-	GCCATCTAGTTTCAACAGC
	AHSG_CCDS3278.1.ex2_3:186334283-186334306:-	TCACATTTTGCGTATACCA
	AHSG_CCDS3278.1.ex2_3:186334307-186334330:-	TAGTCATCTGTACCTGGAC
	AHSG_CCDS3278.1.ex3_3:186335024-186335047:+	CCGCTGAACGACACCAGGG
CLRN3	CLRN3_CCDS7656.1.ex1_10:129681992-129682015:+	CCCCAGGAATGTCTGGTAA
	CLRN3_CCDS7656.1.ex1_10:129682052-129682075:+	CAGCGACGTGATCAAACTC
	CLRN3_CCDS7656.1.ex1_10:129682076-129682099:-	TCGGTGACTATCCTGTTCC
	CLRN3_CCDS7656.1.ex2_10:129690848-129690871:-	GTAGTGAAGAATTGAGTCA
	CLRN3_CCDS7656.1.ex2_10:129690922-129690945:+	TCTCTAACAGCAATTGTAC
KIAA0319L	KIAA0319L_CCDS390.1.ex12_1:35925920-35925943:-	AACTAAGTAAACTCGTCCC
	KIAA0319L_CCDS390.1.ex13_1:35928297-35928320:-	TGCAGAGCCCCGTAAGAAT
	KIAA0319L_CCDS390.1.ex14_1:35932207-35932230:-	TGAACGTGACAGTCAAGCC
	KIAA0319L_CCDS390.1.ex15_1:35936462-35936485:+	CTTCGATAGTTTGAGGATC
	KIAA0319L_CCDS390.1.ex15_1:35936538-35936561:-	AGAAACCTACACCTACGAC
ACP2	ACP2_CCDS44583.1.ex1_11:47269188-47269211:+	ACCTCTTGCCGGTGATAAG
	ACP2_CCDS44583.1.ex1_11:47269221-47269244:+	CCGTGATAGCGCTGCCGCA

	ACP2_CCDS44583.1.ex2_11:47269590-47269613:+	GGTTAACTGACCAAACCCC
	ACP2_CCDS44583.1.ex2_11:47269638-47269661:-	CCAGTGAAGACATATCCCA
	ACP2_CCDS44583.1.ex2_11:47269653-47269676:+	CACTGGTGAACGGTCTCCA
CREG1	CREG1_CCDS1262.1.ex3_1:167522623-167522646:-	CTCTCCGTGAGCAACCTGC
	CREG1_CCDS1262.1.ex3_1:167522697-167522720:-	ACGTCCTCTCGCTCAGCGA
	CREG1_CCDS1262.1.ex3_1:167522772-167522795:-	TGACGCACGTCTCCGACTG
	CREG1_CCDS1262.1.ex3_1:167522814-167522837:+	CGTCCTCGCGGGGTGGTAG
	CREG1_CCDS1262.1.ex3_1:167522886-167522909:-	TGCTCGTGTCGCCCGCGCG
COL6A1	COL6A1_CCDS13727.1.ex1_21:47402649-47402672:+	GCGCTTCATCGACAACCTG
	COL6A1_CCDS13727.1.ex2_21:47404338-47404361:+	ACCGACTGCGCTATCAAGA
	COL6A1_CCDS13727.1.ex4_21:47406879-47406902:-	CGCCGGTACGTGTGGTCCG
	COL6A1_CCDS13727.1.ex12_21:47410299-47410322:+	GGCAAGCGTGGCATCGACG
	COL6A1_CCDS13727.1.ex20_21:47417345-47417368:-	GTAGCCTTTAGGTCCGATA
ATP6AP2	ATP6AP2_CCDS14252.1.ex1_X:40448293-40448316:+	ATGGAAATTGGCCTATACC
	ATP6AP2_CCDS14252.1.ex2_X:40450479-40450502:+	TTTAGGACCTTTCTTGGCC
	ATP6AP2_CCDS14252.1.ex2_X:40450526-40450549:+	CGTCCTCGGGCTACCGTCA
	ATP6AP2_CCDS14252.1.ex2_X:40450566-40450589:+	ACAAACTGGCTCTACCCCC
	ATP6AP2_CCDS14252.1.ex4_X:40456885-40456908:-	CCTACTCAGAGAATTGAGG
SLC30A10	SLC30A10_CCDS31026.1.ex3_1:220101141-220101164:-	CCGTGTTCGCAAACGTAGC
	SLC30A10_CCDS31026.1.ex3_1:220101184-220101207:+	TTCCACCGAGGTCCCCCGG
	SLC30A10_CCDS31026.1.ex3_1:220101207-220101230:+	TTACGGCCGAGTCCGAGCC
	SLC30A10_CCDS31026.1.ex3_1:220101436-220101459:-	CGCATCGATGACCCCGAGC
	SLC30A10_CCDS31026.1.ex3_1:220101547-220101570:-	TACGGCTACGCCCGCGCCG

PLD3	PLD3_CCDS33027.1.ex2_19:40872702-40872725:-	GCCCACAACCGCCAGAATG
	PLD3_CCDS33027.1.ex2_19:40872802-40872825:-	ACTCGCAAGGGTCATAGCA
	PLD3_CCDS33027.1.ex3_19:40873643-40873666:-	GGGTTCCCCGTGGAGGCAT
	PLD3_CCDS33027.1.ex4_19:40875843-40875866:-	GTTCACGCCCTTTGGTGCC
	PLD3_CCDS33027.1.ex5_19:40876052-40876075:-	TTGGTATGCAGGACGCCAT
LMF2	LMF2_CCDS14093.2.ex7_22:50943895-50943918:+	TAGGCCAGAAGCCCGTAGA
	LMF2_CCDS14093.2.ex9_22:50944479-50944502:+	CCAAGCGCAGGCGTCGAAT
	LMF2_CCDS14093.2.ex9_22:50944520-50944543:+	ACAGCGATCTCAATTAGGA
	LMF2_CCDS14093.2.ex9_22:50944573-50944596:-	TGGTTCGCACACCACCTGC
	LMF2_CCDS14093.2.ex10_22:50944755-50944778:-	CTCATGTTCGCCTCAGGCG
TMX3	TMX3_CCDS32840.1.ex5_18:66351663-66351686:-	TACATCAGTTGAACATACC
	TMX3_CCDS32840.1.ex6_18:66354924-66354947:+	GAGGAAGCCATCCATAGCA
	TMX3_CCDS32840.1.ex6_18:66354963-66354986:-	CTGTCATCATGGATCAACA
	TMX3_CCDS32840.1.ex10_18:66367705-66367728:-	TAAAGATTAAAAGGGGACT
	TMX3_CCDS32840.1.ex12_18:66377319-66377342:+	GACCAACTTCATTCCAAAT
CRTAP	CRTAP_CCDS2657.1.ex0_3:33155676-33155699:-	CGGCATCAGCTCGTCCCGT
	CRTAP_CCDS2657.1.ex0_3:33155695-33155718:-	GCCGGTAGGCCGACTCGAG
	CRTAP_CCDS2657.1.ex0_3:33155920-33155943:+	GCCTCAAGCGCTGCAAGCA
	CRTAP_CCDS2657.1.ex1_3:33161937-33161960:-	TAATGTAGTCCTCGGCACC
	CRTAP_CCDS2657.1.ex2_3:33165993-33166016:-	GCTGCGAGACACTCGTAAA
GALNT1	GALNT1_CCDS11915.1.ex2_18:33257647-33257670:-	TGAGCGATTAATGACACTA
	GALNT1_CCDS11915.1.ex3_18:33263400-33263423:+	CCAGTTCATGTAATTCGAA
	GALNT1_CCDS11915.1.ex3_18:33263500-33263523:+	CCCATTGTGAGTGTACAGT
	GALNT1_CCDS11915.1.ex3_18:33263543-33263566:-	TACCTGTCATGTTTGATCC
	GALNT1_CCDS11915.1.ex4_18:33267050-33267073:+	GACCTATGGTGGGTTCAAC
CNNM3	CNNM3_CCDS2025.1.ex0_2:97482147-97482170:+	CGCGGGTTGGGTACGCGGA
	CNNM3_CCDS2025.1.ex0_2:97482333-97482356:-	GAGTGCGGGCGCACGGCCT
	CNNM3_CCDS2025.1.ex0_2:97482812-97482835:+	TGCCCGTCGCGCTGCCCGT
	CNNM3_CCDS2025.1.ex0_2:97483219-97483242:-	ACCTCGCTTGAATTCCTCC

	CNNM3_CCDS2025.1.ex2_2:97492558-97492581:+	TCCATCTTAGGAGACACCG
CD36	CD36_CCDS34673.1.ex1_7:80285994-80286017:+	AAGAGGTCCTTATACGTAC
	CD36_CCDS34673.1.ex2_7:80290455-80290478:-	GATAGTGAAGGTTCGAAGA
	CD36_CCDS34673.1.ex3_7:80292452-80292475:-	GACCAACTGTGGTAGTAAC
	CD36_CCDS34673.1.ex5_7:80295766-80295789:-	TCGCAGTGACTTTCCCAAT
	CD36_CCDS34673.1.ex5_7:80295783-80295806:+	GCGACATGATTAATGGTAC
FXYD2	FXYD2_CCDS8385.1.ex2_11:117693105-117693128:+	GCCACCCCACTTACTGAGG
	FXYD2_CCDS8385.1.ex2_11:117693138-117693161:-	GCTGGACTGGCCTTCATCG
	FXYD2_CCDS8385.1.ex2_11:117693155-117693178:-	ATGGGGGCCTGATCTTCGC
	FXYD2_CCDS8385.1.ex2_11:117693171-117693194:-	TATGAGACCGTTCGCAATG
	FXYD2.5	ACGTGGACCCGTTCTACTA
COL6A2	COL6A2_CCDS13728.1.ex1_21:47532171-47532194:-	GCCAGCGCGCAGTCGGTGA
	COL6A2_CCDS13728.1.ex1_21:47532439-47532462:+	GACTCCACCGAGATCGACC
	COL6A2_CCDS13728.1.ex7_21:47536573-47536596:+	GAGCCGACGGTCGCAAGGT
	COL6A2_CCDS13728.1.ex10_21:47537791-47537814:-	GCTTCCCCCGGGTAACCGT
	COL6A2_CCDS13728.1.ex11_21:47538558-47538581:-	TTGGCCCCGATTTCTCCCG
LRBA	LRBA_CCDS3773.1.ex35_4:151773034-151773057:+	CCTCAAGCACATGTCGATG
	LRBA_CCDS3773.1.ex44_4:151821289-151821312:-	ATAACACCATTCGGAGAGT
	LRBA_CCDS3773.1.ex48_4:151829898-151829921:-	TGCATTCACGTACAATCCA
	LRBA_CCDS3773.1.ex50_4:151836840-151836863:-	AGTACACATCTATAACCGA
	LRBA_CCDS3773.1.ex56_4:151935621-151935644:-	TGACCGGTTTGGTTGAAGT
CLPTM1L	CLPTM1L_CCDS3862.1.ex11_5:1335163-1335186:+	CACATACCGAACTGCTGCA
	CLPTM1L_CCDS3862.1.ex13_5:1339054-1339077:-	AGTGTCCCACTGGCGACCG
	CLPTM1L_CCDS3862.1.ex14_5:1341784-1341807:-	ACCGGGGAGTCTGATACAC
	CLPTM1L_CCDS3862.1.ex15_5:1344532-1344555:-	ACCACGACGAGGTCCCACC

	CLPTM1L_CCDS3862.1.ex16_5:1344832-1344855:+	GTTGGCGTCGCCGGAGCAC
LAMB2	LAMB2_CCDS2789.1.ex15_3:49163496-49163519:+	CATTGGTAGCGTTCAAAGG
	LAMB2_CCDS2789.1.ex21_3:49167052-49167075:-	AACGTCTAGTGACTGGACG
	LAMB2_CCDS2789.1.ex25_3:49168504-49168527:-	AACCTACTCGACCCACGGA
	LAMB2_CCDS2789.1.ex25_3:49168531-49168554:+	CAACGTGTGTAGACGAGTC
	LAMB2_CCDS2789.1.ex27_3:49169081-49169104:+	AAATATCGGTACACATGCC
CANT1	CANT1_CCDS11760.1.ex2_17:76993093-76993116:+	CGTCGGACAGAATCACCCA
	CANT1_CCDS11760.1.ex2_17:76993129-76993152:-	GGGTCGTCTACCAGATCGA
	CANT1_CCDS11760.1.ex2_17:76993149-76993172:-	CTCCGTGGATGACCGGACG
	CANT1_CCDS11760.1.ex2_17:76993364-76993387:-	ATCGCAGTTATCGCAGACC
	CANT1_CCDS11760.1.ex2_17:76993384-76993407:+	TTCGATACCGAATCCCAGC
CLDN3	CLDN3_CCDS5559.1.ex0_7:73183964-73183987:+	CGACCAGGACACCGGCACG
	CLDN3_CCDS5559.1.ex0_7:73184021-73184044:-	AAGGCCAAGATCACCATCG
	CLDN3_CCDS5559.1.ex0_7:73184045-73184068:-	AACTGCGTGCAGGACGACA
	CLDN3_CCDS5559.1.ex0_7:73184090-73184113:-	GCCGCCTTCGGGCTGCTAG
	CLDN3_CCDS5559.1.ex0_7:73184241-73184264:-	CATCACGTCGCAGAACATC
ALPL	ALPL_CCDS217.1.ex3_1:21889702-21889725:-	TTGCACCGGGAACGCTCAG
	ALPL_CCDS217.1.ex4_1:21890554-21890577:-	GCATGGTTCACTCTCGTGG
	ALPL_CCDS217.1.ex4_1:21890596-21890619:-	CAGTCCCGGTCAGCCGAGT
	ALPL_CCDS217.1.ex4_1:21890677-21890700:+	CCAGCTCATGCATAACATC
	ALPL_CCDS217.1.ex5_1:21894587-21894610:+	TCTTTTAGGTGATCATGGG
TSPAN6	TSPAN6_CCDS14470.1.ex2_X:99888500-99888523:-	TGTCACCGATTATAGAGAT
	TSPAN6_CCDS14470.1.ex4_X:99890180-99890203:+	CTGAAAACAAATCCTACGA
	TSPAN6_CCDS14470.1.ex5_X:99890627-99890650:-	CCTTCGTGCTCATTGCTAC
	TSPAN6_CCDS14470.1.ex5_X:99890714-99890737:-	GCGTTATCCTTCTTGCAGT
	TSPAN6_CCDS14470.1.ex5_X:99890735-99890758:-	GAATTCCCACATAGATCAC

ABCA1	ABCA1_CCDS6762.1.ex28_9:107581892-107581915:+	ATATTCCCCTGCGGGAGTA
	ABCA1_CCDS6762.1.ex28_9:107581948-107581971:+	ACCTTAGATCCCCCGACAA
	ABCA1_CCDS6762.1.ex29_9:107582304-107582327:-	GTTCTATGCCCGCTTGAAA
	ABCA1_CCDS6762.1.ex36_9:107593327-107593350:-	TGAGGACATGCGGTACGTC
	ABCA1_CCDS6762.1.ex43_9:107620822-107620845:-	CGAGTACTTCGTTCCAACA
QSOX2	QSOX2_CCDS35178.1.ex4_9:139110636-139110659:+	CCGTGTACAGCTTCGACCT
	QSOX2_CCDS35178.1.ex6_9:139113668-139113691:+	GGTAGATCAGGTAACACGA
	QSOX2_CCDS35178.1.ex6_9:139113735-139113758:-	ATCGTGGTGACCCGAGCAC
	QSOX2_CCDS35178.1.ex10_9:139118604-139118627:+	CGCGTGATACTCACCCGGA
	QSOX2_CCDS35178.1.ex11_9:139137346-139137369:-	CATCGGCTACGCGCCCACT
SLC2A9	SLC2A9_CCDS3406.1.ex6_4:9943629-9943652:+	GACCACAATCACTCCAAAC
	SLC2A9_CCDS3406.1.ex7_4:9982257-9982280:+	ACGCCAATGCAGATAAAGA
	SLC2A9_CCDS3406.1.ex7_4:9982300-9982323:-	TCTCACCCAAGGAGATCCG
	SLC2A9_CCDS3406.1.ex8_4:9987304-9987327:-	ATCGTGGGACGCTTCATCA
	SLC2A9_CCDS3406.1.ex9_4:9998448-9998471:+	AGTCCACCGATGGCGAATA
SLC5A6	SLC5A6_CCDS1740.1.ex9_2:27427674-27427697:-	CGGTACCTCAGTTCCCGCA
	SLC5A6_CCDS1740.1.ex9_2:27427707-27427730:-	TCCTTATACGGGGTGAACC
	SLC5A6_CCDS1740.1.ex9_2:27427762-27427785:+	CAGAAGGTGTGCCGCACAA
	SLC5A6_CCDS1740.1.ex11_2:27428905-27428928:+	ACGGTACAGACAATGCCCA
	SLC5A6_CCDS1740.1.ex14_2:27430365-27430388:-	CTACCATGCTTGTCGTGGC
REEP6	REEP6_CCDS12070.1.ex0_19:1491361-1491384:+	AGAAGCGGTATCTGGCTGC
	REEP6_CCDS12070.1.ex1_19:1495371-1495394:-	CACTCACGAGGCATATGCG
	REEP6_CCDS12070.1.ex2_19:1495527-1495550:+	GTGTACGCCCTGTTTGGGC
	REEP6_CCDS12070.1.ex2_19:1495550-1495573:-	AGTAGATCGCTGAAGAACT
	REEP6_CCDS12070.1.ex2_19:1495585-1495608:-	CCTTGCCCACGTAGTAGAA
NPC1	NPC1_CCDS11878.1.ex15_18:21131629-21131652:+	ACCAAACGTACCCAGACAA
	NPC1_CCDS11878.1.ex16_18:21134841-21134864:+	TCGTGTTATACGGTGAAAG

	NPC1_CCDS11878.1.ex17_18:21136241-21136264:+	GGGTACATCAGCTCCCGAA
	NPC1_CCDS11878.1.ex17_18:21136427-21136450:-	ACTGCGTGTTCGTCAGGCC
	NPC1_CCDS11878.1.ex17_18:21136504-21136527:-	GGCGGCTGTTCACACGCTG
SORT1	SORT1_CCDS55618.1.ex10_1:109883361-109883384:+	GTTATGTAGACGCCGCGGA
	SORT1_CCDS55618.1.ex10_1:109883410-109883433:-	ATCTCTACACTACCACAGG
	SORT1_CCDS55618.1.ex10_1:109883461-109883484:-	TCTTTACCTCAGATGATCG
	SORT1_CCDS55618.1.ex11_1:109884740-109884763:+	CCCTTGATCTGTTGAAACG
	SORT1_CCDS55618.1.ex16_1:109897992-109898015:-	CTATTGGTCCTGAGAACTC
MCAM	MCAM_CCDS31690.1.ex11_11:119185265-119185288:-	GGTCGCTACCTGTGTAGGG
	MCAM_CCDS31690.1.ex13_11:119185530-119185553:+	GGGTCACGCACTGTAGACG
	MCAM_CCDS31690.1.ex13_11:119185560-119185583:+	GCGGTACTCCTGGGACCGA
	MCAM_CCDS31690.1.ex13_11:119185677-119185700:+	GTACTCCCCAGGTTCGCTC
	MCAM_CCDS31690.1.ex13_11:119185706-119185729:-	TCATCTTCCGTGTGCGCCA
ATG9A	ATG9A_CCDS42820.1.ex8_2:220089200-220089223:+	GTTAGCGATGCCAATCCAC
	ATG9A_CCDS42820.1.ex8_2:220089261-220089284:-	GGCCGAGTACAAACGTGGG
	ATG9A_CCDS42820.1.ex8_2:220089431-220089454:-	CGTTTCCAGAACTACATGG
	ATG9A_CCDS42820.1.ex8_2:220089549-220089572:+	TGCCACGTGCAATACGGAA
	ATG9A_CCDS42820.1.ex10_2:220090178-220090201:-	CACCCTACTGAACCCGTCA
CP	CP_CCDS3141.1.ex11_3:148917536-148917559:-	GATTCAATAAGAACAACGA
	CP_CCDS3141.1.ex11_3:148917563-148917586:-	CCCTCAGTATTGAGCCGAT
	CP_CCDS3141.1.ex11_3:148917593-148917616:-	GAGTAACCTTCCATAACAA
	CP_CCDS3141.1.ex12_3:148919924-148919947:-	GCCTCCTTCACAAATCGAA
	CP_CCDS3141.1.ex13_3:148924059-148924082:-	CATCAAAGGATAATATCCG
GPC6	GPC6_CCDS9469.1.ex0_13:93879807-93879830:+	CGCCAGGCGTACGGTGCCA
	GPC6_CCDS9469.1.ex2_13:94482450-94482473:+	ATATGTTTGTACGGACCTA

	GPC6_CCDS9469.1.ex2_13:94482520-94482543:+	GCTGAAAAGGTACTACACT
	GPC6_CCDS9469.1.ex3_13:94679993-94680016:-	GAGGGCACGGATACACCCT
	GPC6_CCDS9469.1.ex3_13:94680025-94680048:+	GTACTGCCCATACTGTCGG
GPC1	GPC1_CCDS2534.1.ex0_2:241375482-241375505:-	CACCCGAGATCTCCGCCTG
	GPC1_CCDS2534.1.ex2_2:241401677-241401700:+	CTGTACACGCAGAACGCGA
	GPC1_CCDS2534.1.ex2_2:241401728-241401751:-	GTTGGCACCGCGGTAGTAC
	GPC1_CCDS2534.1.ex2_2:241401971-241401994:+	AGCGACGTGGTCCGGAAAG
	GPC1_CCDS2534.1.ex3_2:241402855-241402878:+	TATTGCCGAAATGTGCTCA
SLC2A3	SLC2A3_CCDS8586.1.ex5_12:8083183-8083206:-	GAGCTATGGCCGCTGCTAC
	SLC2A3_CCDS8586.1.ex6_12:8084001-8084024:-	CTGTGTAAAGTAGCTAAGT
	SLC2A3_CCDS8586.1.ex7_12:8085590-8085613:+	CGGTTGACGAAGAGTCCGA
	SLC2A3_CCDS8586.1.ex7_12:8085626-8085649:+	ATACCCCCGACGGAAAATA
	SLC2A3_CCDS8586.1.ex7_12:8085702-8085725:-	ATAAAACTTTGACGGACAA
EMC1	EMC1_CCDS190.1.ex12_1:19564536-19564559:-	ACCACGATAACATTTAGCC
	EMC1_CCDS190.1.ex12_1:19564581-19564604:-	ACCATTAACCTATACCTCG
	EMC1_CCDS190.1.ex16_1:19566919-19566942:-	GGTTAGGGTTTCAACTCCG
	EMC1_CCDS190.1.ex17_1:19567594-19567617:-	TGGTGTATTCTTACGGCTC
	EMC1_CCDS190.1.ex21_1:19571407-19571430:-	CAGCATTAAATTCCCGAAC
TMED7	TMED7_CCDS4120.1.ex1_5:114956363-114956386:-	TCTTACAGGTGATTACTGG
	TMED7_CCDS4120.1.ex1_5:114956324-114956347:-	GTCGATTAGAAGATCCTGA
	TMED7_CCDS4120.1.ex1_5:114956257-114956280:-	CTTCACAGCCTCCAAAAAT
	TMED7_CCDS4120.1.ex2_5:114961322-114961345:-	TCTACGAGGACATCGCTCA
	TMED7_CCDS4120.1.ex0_5:114952109-114952132:+	GCTTCGTGAATTGAAACAC
IL1RAP	IL1RAP_CCDS3298.1.ex1_3:190321993-190322016:-	TCAAGAAGTGTTCAAAGAG
	IL1RAP_CCDS3298.1.ex1_3:190322060-190322083:+	GTATTGGACTAGGCAGGAC
	IL1RAP_CCDS3298.1.ex1_3:190322109-190322132:-	ACTAATGCGGTTCTCGGGG
	IL1RAP_CCDS3298.1.ex2_3:190326865-190326888:-	CTATATACAGTTTATGCAC
	IL1RAP_CCDS3298.1.ex2_3:190326938-190326961:-	GTGATAGTCGGTTTGACAC

SORL1	SORL1_CCDS8436.1.ex0_11:121323251-121323274:+	CGCGGACGAGAAGCCGCTC
	SORL1_CCDS8436.1.ex2_11:121348934-121348957:-	TTACCCGCTTGTTGTCCGC
	SORL1_CCDS8436.1.ex11_11:121403207-121403230:-	AGCCGTGATGATTCCGCCG
	SORL1_CCDS8436.1.ex18_11:121428031-121428054:-	TGAGTCGGAAGTCGCCATC
	SORL1_CCDS8436.1.ex22_11:121440871-121440894:-	CGATACTGGTTGCGAAGAC
B4GALT1	B4GALT1_CCDS6535.1.ex4_9:33135269-33135292:-	ATTCCATTCCGCAACCGGC
	B4GALT1_CCDS6535.1.ex4_9:33135393-33135416:+	ATGTTAAACTCAATCAGCA
	B4GALT1_CCDS6535.1.ex5_9:33166816-33166839:+	ACCGAGGTCAAGTTGCTAG
	B4GALT1_CCDS6535.1.ex5_9:33166992-33167015:+	TGTGGAGACTCCGACCAGT
	B4GALT1_CCDS6535.1.ex5_9:33167046-33167069:+	GTAAACGAGGGTGACGCCA
PIGT	PIGT_CCDS13353.1.ex0_20:44044935-44044958:+	ATTCCAGTTCCGCACGCGC
	PIGT_CCDS13353.1.ex0_20:44044954-44044977:+	TGGGATTCGGAGCTTCAGC
	PIGT_CCDS13353.1.ex3_20:44047941-44047964:-	AGCATAGCGCAGAAAGTAG
	PIGT_CCDS13353.1.ex4_20:44048218-44048241:-	CATGACTTACTCTGCAAAC
	PIGT_CCDS13353.1.ex5_20:44048815-44048838:-	GCATCAAATACAACTGACA
FAM171A1	FAM171A1_CCDS31154.1.ex0_10:15256362-15256385:-	GATGTCTCCGGGCGGCGAA
	FAM171A1_CCDS31154.1.ex1_10:15258091-15258114:-	CTAGGTCCCGTTGTAACAC
	FAM171A1_CCDS31154.1.ex3_10:15290650-15290673:+	AACCGCCACGCCGCGACAT
	FAM171A1_CCDS31154.1.ex3_10:15290733-15290756:-	AGTAATGGAACGCCGGTGC
	FAM171A1_CCDS31154.1.ex4_10:15296777-15296800:+	CTGGCGGCCGTGAGAAACG
AGRN	AGRN_CCDS30551.1.ex5_1:977005-977028:+	TGTGTCATGGGCCGATCGG
	AGRN_CCDS30551.1.ex7_1:978789-978812:-	ATCTCCCGCCCGAGGGTAC
	AGRN_CCDS30551.1.ex11_1:979785-979808:+	GGCTCTACGTAGCGGCCCA
	AGRN_CCDS30551.1.ex13_1:980756-980779:-	GTGCCGCCGTAAGAGCCAT
	AGRN_CCDS30551.1.ex13_1:980872-980895:+	TCGTCACCGATGGCCGGAG

TGFBR3	TGFBR3_CCDS30770.1.ex8_1:92185679-92185702:-	AACCCGCCCATCCGGGGAG
	TGFBR3_CCDS30770.1.ex8_1:92185727-92185750:-	CCTGAGCTACGGATCCTGC
	TGFBR3_CCDS30770.1.ex10_1:92193257-92193280:-	GTGCAAAAAGTCTGTCAAC
	TGFBR3_CCDS30770.1.ex10_1:92193316-92193339:-	GATATAAGACCTTCTCAAG
	TGFBR3_CCDS30770.1.ex11_1:92195498-92195521:-	CCCTCCAAAGTGCAACATA
SIRPA	SIRPA_CCDS13022.1.ex1_20:1895785-1895808:+	AGTCCGTGTTGGTTGCAGC
	SIRPA_CCDS13022.1.ex1_20:1895928-1895951:-	GTCTGAAACAGTTGTTACC
	SIRPA_CCDS13022.1.ex1_20:1896048-1896071:-	CTTAAACTCCACGTCATCG
	SIRPA_CCDS13022.1.ex2_20:1902053-1902076:-	CGCAGGGCCCGATACCACG
	SIRPA_CCDS13022.1.ex2_20:1902202-1902225:-	TAGGACACGCTCTCTCCTA
SLC22A1	SLC22A1_CCDS5274.1.ex0_6:160543262-160543285:-	AGGGGGTCTACACAGCTGA
	SLC22A1_CCDS5274.1.ex0_6:160543355-160543378:-	TCAGTGACGATGGAAGAGC
	SLC22A1_CCDS5274.1.ex1_6:160551181-160551204:-	AAGAAGCCCGCATTCAAAC
	SLC22A1_CCDS5274.1.ex2_6:160553328-160553351:-	GACATGTAGTTGGGCGAGA
	SLC22A1_CCDS5274.1.ex3_6:160555022-160555045:+	GTGGCGATCATGTACCAGA
SLC19A3	SLC19A3_CCDS2468.1.ex3_2:228563919-228563942:+	AAAGTACGACATGTTCGCC
	SLC19A3_CCDS2468.1.ex3_2:228564039-228564062:-	TACGCCTACATATACAGCG
	SLC19A3_CCDS2468.1.ex3_2:228564058-228564081:+	TAGTAGGCCACCTCGGCGG
	SLC19A3_CCDS2468.1.ex3_2:228564192-228564215:+	GTAGCGGACATAATCGGTG
	SLC19A3_CCDS2468.1.ex3_2:228564240-228564263:+	GTAGGAGTATGTCCAAACG
CA14	CA14_CCDS947.1.ex2_1:150233884-150233907:-	TTTCCACACTCAGGGTAAG
	CA14_CCDS947.1.ex2_1:150233912-150233935:-	CTGAATATCGATGGGCGAC
	CA14_CCDS947.1.ex2_1:150233978-150234001:-	GCCAGGCTGGTCATATCCG
	CA14_CCDS947.1.ex3_1:150234566-150234589:+	CTGCCCTCTACCCTGTATC
	CA14_CCDS947.1.ex6_1:150235535-150235558:-	AGAGCACACTCTGGTAGCA

SLCO2B1	SLCO2B1_CCDS44683.1.ex3_11:74880434-74880457:+	TCGCCCCTCTACCTCGGTG
	SLCO2B1_CCDS44683.1.ex4_11:74880765-74880788:-	CACATAAAGGCGCAGCATG
	SLCO2B1_CCDS44683.1.ex5_11:74883439-74883462:+	ATAAAGGACCCCCGATGGG
	SLCO2B1_CCDS44683.1.ex5_11:74883559-74883582:+	GAGCTTCAGTTTCGGCGAA
	SLCO2B1_CCDS44683.1.ex6_11:74899242-74899265:-	AGTCAGGTTTGGTGCAATC
SLC22A9	SLC22A9_CCDS8043.1.ex0_11:63137845-63137868:+	AACACAAGTGACGCAGACA
	SLC22A9_CCDS8043.1.ex1_11:63138688-63138711:+	AGGCGGTCATTTATCAGAC
	SLC22A9_CCDS8043.1.ex2_11:63141196-63141219:-	GCGTAGTGAGCAGTAAATG
	SLC22A9_CCDS8043.1.ex3_11:63141361-63141384:-	GTGTTGCCCACTCGGCTAC
	SLC22A9_CCDS8043.1.ex3_11:63141464-63141487:-	AGGATATGCCAGTCTCGAA
CD9	CD9_CCDS8540.1.ex1_12:6334678-6334701:+	ATAATTCCAGCTTCTACAC
	CD9_CCDS8540.1.ex2_12:6341799-6341822:+	TCTATATTCTGATCGGAGC
	CD9_CCDS8540.1.ex3_12:6342588-6342611:-	AATGGCGAATATCACCAAG
	CD9_CCDS8540.1.ex3_12:6342608-6342631:+	TGAAATAGCTGCGGCCATC
	CD9_CCDS8540.1.ex3_12:6342623-6342646:-	TTGTGGGAATATCCCCAGA
GPRC5C	GPRC5C_CCDS11699.1.ex1_17:72436032-72436055:+	ACAACCTGTGTGACCGCTC
	GPRC5C_CCDS11699.1.ex1_17:72436141-72436164:-	TTCTTGGTGTCCTGCACAA
	GPRC5C_CCDS11699.1.ex1_17:72436343-72436366:-	CCCGTGGTTCTTCCGGGCC
	GPRC5C_CCDS11699.1.ex1_17:72436604-72436627:-	GCGCTTGTAGCGGCCACAC
	GPRC5C_CCDS11699.1.ex1_17:72436617-72436640:+	ACAAGCGCTGGCGTAAGCA
DPP7	DPP7_CCDS7030.1.ex6_9:140007550-140007573:-	GCCTACGACACGGTCCGCT
	DPP7_CCDS7030.1.ex7_9:140007675-140007698:-	GAAGCGTTCCGACAGATCA
	DPP7_CCDS7030.1.ex8_9:140007814-140007837:+	CGTGACGTCCCGGAAGAAC
	DPP7_CCDS7030.1.ex9_9:140008350-140008373:-	GCGCTACGACGCGACCTCG
	DPP7_CCDS7030.1.ex9_9:140008436-140008459:+	GCGTGGACTGCGCACCGAA

TMBIM1	TMBIM1_CCDS2412.1.ex3_2:219142157-219142180:+	GCTTTGGTTTGGTACATAC
	TMBIM1_CCDS2412.1.ex5_2:219142681-219142704:-	CTTCAGACGCCGTTTCCCA
	TMBIM1_CCDS2412.1.ex7_2:219143787-219143810:+	TCCTCACAAAGGCGCTGAC
	TMBIM1_CCDS2412.1.ex9_2:219144791-219144814:-	TGAGTGATAGCTTCGGGCC
	TMBIM1_CCDS2412.1.ex10_2:219146662-219146685:+	GTAGTTCATGGGCATCGGG
CYBB	CYBB_CCDS14242.1.ex2_X:37642845-37642868:-	ACATTTTCTCTTACCGCAC
	CYBB_CCDS14242.1.ex3_X:37651240-37651263:+	AGTTCGAAGACAACTGGAC
	CYBB_CCDS14242.1.ex3_X:37651290-37651313:+	GGATGATTGCACTTCACTC
	CYBB_CCDS14242.1.ex6_X:37658199-37658222:+	CACCCAGACGAATTGTACG
	CYBB_CCDS14242.1.ex6_X:37658328-37658351:-	GAATTGTACATACCATAGG
CD99L2	CD99L2_CCDS14697.1.ex3_X:149945911-149945934:-	TGACGACCCTGGATCTGGT
	CD99L2_CCDS14697.1.ex3_X:149945925-149945948:+	GTCGTCATTGCTGCCGTAC
	CD99L2_CCDS14697.1.ex3_X:149945939-149945962:-	CTTCAGGTGATGGCCGGTA
	CD99L2_CCDS14697.1.ex6_X:149999692-149999715:+	TACTACTCACGCTTTACTG
	CD99L2_CCDS14697.1.ex6_X:149999736-149999759:-	GACTTTGATGATTTTAACC
CRB3	CRB3_CCDS12166.1.ex1_19:6465593-6465616:-	AGCTGGAGCTGGTGGATGA
	CRB3_CCDS12166.1.ex1_19:6465607-6465630:+	TCCAGCTCCGATGGCAACC
	CRB3_CCDS12166.1.ex2_19:6466484-6466507:+	GCCATCACTGCTATCATCG
	CRB3_CCDS12166.1.ex2_19:6466514-6466537:+	CTCTTGGCTGCCTTGCTCC
	CRB3_CCDS12166.1.ex2_19:6466532-6466555:-	CAGTGCCAGCCCCACAGCC
TMEM123	TMEM123_CCDS41702.1.ex2_11:102272743-102272766:+	TTGGGTGTAGACTTTAAGG
	TMEM123_CCDS41702.1.ex2_11:102272792-102272815:-	CATCTAATACAACAACACC
	TMEM123_CCDS41702.1.ex2_11:102272812-102272835:+	GCCGCTGTAGGTTTCATGG
	TMEM123_CCDS41702.1.ex2_11:102272838-102272861:-	TCAGACTCCAGTAATACAA
	TMEM123_CCDS41702.1.ex3_11:102319541-102319564:-	ACTCCAGTGCTAACTCAAC
ABCB4	ABCB4_CCDS5605.1.ex13_7:87060823-87060846:+	ATTTCGGACCGTAGACAGT

	ABCB4_CCDS5605.1.ex17_7:87072977-87073000:+	CTTGACGTTAGCTCGAGAA
	ABCB4_CCDS5605.1.ex22_7:87082254-87082277:-	AATACGCGGCTAACAGAGT
	ABCB4_CCDS5605.1.ex23_7:87083877-87083900:-	CCTTGTCGCTGCTAAATCC
	ABCB4_CCDS5605.1.ex24_7:87092206-87092229:-	TTAGTTTCGATACTCCGAT
SLC10A1	SLC10A1_CCDS9797.1.ex3_14:70252846-70252869:+	GATTTGAGGACGATCCCTA
	SLC10A1_CCDS9797.1.ex3_14:70252884-70252907:-	ATCGTGATATCACTGGTCC
	SLC10A1_CCDS9797.1.ex4_14:70263552-70263575:+	CATTGGACAGGTTCCCTCC
	SLC10A1_CCDS9797.1.ex4_14:70263658-70263681:+	GATGCCATACTGTGCCACC
	SLC10A1_CCDS9797.1.ex4_14:70263693-70263716:-	CTCACTTATGGAAGCCTAA
SLC47A1	SLC47A1_CCDS11209.1.ex0_17:19437380-19437403:-	GGCCACCTTACTCACCGCG
	SLC47A1_CCDS11209.1.ex2_17:19449741-19449764:+	TGCAGGTTATCAATGTCAC
	SLC47A1_CCDS11209.1.ex3_17:19451293-19451316:-	GTTCTGGCTCCCGTACGTC
	SLC47A1_CCDS11209.1.ex6_17:19458538-19458561:-	TGCAACTCCAGTTACGATC
	SLC47A1_CCDS11209.1.ex6_17:19458565-19458588:-	GTTGGCGAGGGCATTGACA
CDHR5	CDHR5_CCDS7707.1.ex8_11:621074-621097:+	TTACCAGTATGTGCCCCGT
	CDHR5_CCDS7707.1.ex9_11:621372-621395:+	GCCGCTCGTAGAAGTCCAG
	CDHR5_CCDS7707.1.ex9_11:621404-621427:-	TGTAAACCGTCCCGCCCTG
	CDHR5_CCDS7707.1.ex10_11:621617-621640:-	ATCCCCGAGACGCAACTGC
	CDHR5_CCDS7707.1.ex13_11:624648-624671:-	GTGGACATCCACGTCCCGG
SLC3A1	SLC3A1_CCDS1819.1.ex0_2:44502843-44502866:-	GGCTGGACGCCCTTGAAGT
	SLC3A1_CCDS1819.1.ex0_2:44502910-44502933:+	GCCCGCTACCGCATACCTC
	SLC3A1_CCDS1819.1.ex2_2:44508584-44508607:+	AATTGAGTCGGACACGGAC
	SLC3A1_CCDS1819.1.ex3_2:44513189-44513212:-	ACTTCGTCAAAGTGCCAAC
	SLC3A1_CCDS1819.1.ex4_2:44527115-44527138:+	TACGGTTCTGGCTCACAAA
TIMD4	TIMD4_CCDS4332.1.ex6_5:156378712-156378735:+	GTTGTTGTCATTTGTCGGG
	TIMD4_CCDS4332.1.ex7_5:156381516-156381539:+	CTGGGGTTTAAGATGGTCA
	TIMD4_CCDS4332.1.ex7_5:156381547-156381570:-	TTCAGGGGACTATCCCGAG

	TIMD4_CCDS4332.1.ex7_5:156381625-156381648:+	CCTCCTTGCAACCGGAGTA
	TIMD4_CCDS4332.1.ex7_5:156381719-156381742:-	ACGGAGGTTTTGGGTCACC
GYPC	GYPC_CCDS2136.1.ex1_2:127447865-127447888:+	TGCATACTACCACCATTGC
	GYPC_CCDS2136.1.ex2_2:127451441-127451464:-	ATCCAGACATCCCTGGATC
	GYPC_CCDS2136.1.ex2_2:127451465-127451488:-	AGGTCTCCATTCTGCCATC
	GYPC_CCDS2136.1.ex2_2:127451492-127451515:-	CGACAATGTCCATTATGGT
	GYPC_CCDS2136.1.ex3_2:127453516-127453539:+	GCAGGTGTGATTGCTGCTG
NAALADL2	NAALADL2_CCDS46960.1.ex1_3:174814716-174814739:+	TTGACCAATTCCAGCTAGA
	NAALADL2_CCDS46960.1.ex2_3:174951961-174951984:+	ATGCAGCCTATTCTGCCAA
	NAALADL2_CCDS46960.1.ex3_3:174974207-174974230:+	ATCGATGTGAGTTATGGAA
	NAALADL2_CCDS46960.1.ex4_3:175042095-175042118:-	TACCGACACTTGGGTAACC
	NAALADL2_CCDS46960.1.ex5_3:175165079-175165102:-	GAGATCAGTTTTGCAACGA
ATP13A3	ATP13A3_CCDS43187.1.ex16_3:194159640-194159663:-	AATCGAATTATGCCCACAG
	ATP13A3_CCDS43187.1.ex18_3:194165463-194165486:-	CTTTGGGGGATTCAACGAG
	ATP13A3_CCDS43187.1.ex21_3:194169217-194169240:-	TTCAGACTCGTTTCTACAC
	ATP13A3_CCDS43187.1.ex22_3:194170968-194170991:-	TTTCTACCGACCTTGTGCC
	ATP13A3_CCDS43187.1.ex29_3:194181451-194181474:+	CCACTCAGGCATCCAATAG
LILRB5	LILRB5_CCDS12885.1.ex10_19:54760425-54760448:-	CCACGGTGTATGACAGTGC
	LILRB5_CCDS12885.1.ex10_19:54760489-54760512:+	CTGTCTCTTCCGGGCCCAT
	LILRB5_CCDS12885.1.ex10_19:54760504-54760527:-	GATAAGGAGGGACTCCCAT
	LILRB5_CCDS12885.1.ex10_19:54760529-54760552:+	CGGTACTCCTCAGTCTCCA
	LILRB5_CCDS12885.1.ex10_19:54760577-54760600:+	TTCCCCCGAGCTATCACAG
DSC3	DSC3_CCDS32810.1.ex7_18:28598155-28598178:-	CGAATACCTATAGAAGATA
	DSC3_CCDS32810.1.ex9_18:28602453-28602476:-	TGTGTAGGTACTACAGTGG
	DSC3_CCDS32810.1.ex10_18:28604385-28604408:+	CCTCTACCCTGATGGGTAG

	DSC3_CCDS32810.1.ex11_18:28605751-28605774:-	CTATTTTGCACTCGGCCTG
	DSC3_CCDS32810.1.ex13_18:28611065-28611088:+	CATTTAGAACTCTGAAATC
LNPEP	LNPEP_CCDS4087.1.ex1_5:96314969-96314992:+	GTTCCCGACTGCTGGTGCG
	LNPEP_CCDS4087.1.ex1_5:96315326-96315349:-	GTAGTGGCACAACGGCAGT
	LNPEP_CCDS4087.1.ex1_5:96315561-96315584:+	CCCCGAAGCCCTTCTAGCA
	LNPEP_CCDS4087.1.ex3_5:96322245-96322268:+	CATCAGTCGTTCTAGATGA
	LNPEP_CCDS4087.1.ex5_5:96329617-96329640:+	TCAATGGCGGATAGAAAGC
STEAP4	STEAP4_CCDS43611.1.ex3_7:87913127-87913150:-	GCACTGGATGCAAGTCGGC
	STEAP4_CCDS43611.1.ex3_7:87913178-87913201:+	GTTAAATGCTTTTACCACG
	STEAP4_CCDS43611.1.ex3_7:87913202-87913225:+	TCCTGGCACCAAATGAGCA
	STEAP4_CCDS43611.1.ex3_7:87913319-87913342:+	AAAATCATAATGCTCTCTG
	STEAP4_CCDS43611.1.ex3_7:87913395-87913418:+	TCTGCACCACTGGGCAGTA
SLC15A1	SLC15A1_CCDS9489.1.ex9_13:99361877-99361900:+	CGGGACCATGATCACGATC
	SLC15A1_CCDS9489.1.ex13_13:99364801-99364824:-	AGATTTAGGCATCGGAGTA
	SLC15A1_CCDS9489.1.ex14_13:99368122-99368145:+	ATACTTACACCGATGCACT
	SLC15A1_CCDS9489.1.ex19_13:99378390-99378413:-	AGCTCTTATCGCCGACTCG
	SLC15A1_CCDS9489.1.ex19_13:99378438-99378461:+	GCCACAAACGTATGGTAGA
ABCG8	ABCG8_CCDS1815.1.ex3_2:44078875-44078898:-	GTCTCTCGCACAGTCAAGT
	ABCG8_CCDS1815.1.ex3_2:44078944-44078967:-	GTTACCCTTTTGTCACGCT
	ABCG8_CCDS1815.1.ex4_2:44079544-44079567:+	GGGCAACATGTACGTGCGG
	ABCG8_CCDS1815.1.ex5_2:44079891-44079914:-	GGTGCCAGACGTCATCAGG
	ABCG8_CCDS1815.1.ex5_2:44079912-44079935:-	GGCCGCCCCTAAGTAGATG
ITM2B	ITM2B_CCDS9409.1.ex0_13:48807579-48807602:+	ATCCCCCCCGACGCCGTCG
	ITM2B_CCDS9409.1.ex2_13:48830312-48830335:-	CACAGTAGTACACGTCATC
	ITM2B_CCDS9409.1.ex2_13:48830462-48830485:-	TATCTGCAAACTCTGGGAC
	ITM2B_CCDS9409.1.ex3_13:48832269-48832292:+	GCCTATTTAGATCTTAACC
	ITM2B_CCDS9409.1.ex3_13:48832333-48832356:-	ACTCCAGTAGGTTTCTGGG
ADCY9	ADCY9_CCDS32382.1.ex9_16:4163799-4163822:-	CCGGTACGAAATGGAAGAT

	ADCY9_CCDS32382.1.ex9_16:4164921-4164944:-	GTACGCCCGGCATTACGCG
	ADCY9_CCDS32382.1.ex9_16:4165002-4165025:+	GCGACCATGACGATCAGTC
	ADCY9_CCDS32382.1.ex9_16:4165086-4165109:+	GCATACCGGAACCGGCGCT
	ADCY9_CCDS32382.1.ex9_16:4165142-4165165:-	AAGTTCGACTCGGTGAACC
CNNM4	CNNM4_CCDS2024.2.ex0_2:97427113-97427136:+	AACACGTCCGGCGTGCTGG
	CNNM4_CCDS2024.2.ex0_2:97427293-97427316:+	CTCCTAATTACGGTGCTGC
	CNNM4_CCDS2024.2.ex0_2:97427429-97427452:+	TTGAGCCCATCCGGCGCAA
	CNNM4_CCDS2024.2.ex0_2:97427512-97427535:-	CCCGATGAGGTTGTCTAGA
	CNNM4_CCDS2024.2.ex0_2:97427727-97427750:+	GATTCGCACTGTTTACAAC
CNNM2	CNNM2_CCDS44474.1.ex0_10:104678483-104678506:+	ACGACGTGTCGTTCATGGA
	CNNM2_CCDS44474.1.ex0_10:104678742-104678765:-	TCGATCTCGATGATGCCCG
	CNNM2_CCDS44474.1.ex0_10:104678764-104678787:+	ATCAAACCGCTACGCAAGA
	CNNM2_CCDS44474.1.ex0_10:104678853-104678876:+	GTCCACGGGTGGCGCCGTC
	CNNM2_CCDS44474.1.ex0_10:104678928-104678951:-	CCGTCGTGGTAAATCCAGG
TGFBR2	TGFBR2_CCDS2648.1.ex3_3:30713146-30713169:-	ATATGACTAGCAACAAGTC
	TGFBR2_CCDS2648.1.ex3_3:30713190-30713213:-	GGCAACTCCCAGTGGTGGC
	TGFBR2_CCDS2648.1.ex3_3:30713360-30713383:-	TTGATGTTGTTGGCACACG
	TGFBR2_CCDS2648.1.ex3_3:30713684-30713707:+	GACGCGGCATGTCATCAGC
	TGFBR2_CCDS2648.1.ex3_3:30713829-30713852:-	GGTTAGGTCGTTCTTCACG
FLVCR1	FLVCR1_CCDS1510.1.ex0_1:213032089-213032112:-	CGCGGGGAGAGCGCCGTAA
	FLVCR1_CCDS1510.1.ex0_1:213032215-213032238:+	CACCTTGCTGCACATCGAC
	FLVCR1_CCDS1510.1.ex0_1:213032459-213032482:-	AAACCACACTGAGGCGATG
	FLVCR1_CCDS1510.1.ex0_1:213032510-213032533:+	ACCGCCGTGCTGGGCAATC
	FLVCR1_CCDS1510.1.ex1_1:213037096-213037119:-	GTGTGTTGGGTACTAAAAC
KLB	KLB_CCDS3451.1.ex1_4:39435882-39435905:+	GGTTGGTTATCGATCACGT
	KLB_CCDS3451.1.ex1_4:39435924-39435947:+	AACCGGTCGGAAAACACGA
	KLB_CCDS3451.1.ex1_4:39436010-39436033:+	GGATGGCGACTATCCAGAG
	KLB_CCDS3451.1.ex2_4:39439367-39439390:+	AGTGTTTGGTTATACTGCC

	KLB_CCDS3451.1.ex2_4:39439427-39439450:-	TAAAATAATCCTCGGCGGA

Table S 4 Pool 4: list of 94 genes and 500 sgRNA sequences.

There were 25 non-targeting sgRNAs, 5 *SCARB1*-targeting sgRNAs, 5 *CD81*-targeting sgRNAs and 465 sgRNAs targeting 94 genes. Genes are listed in an order of proteomic abundance.

Pool 5		
Gene	sgRNA name	sgRNA sequence (19 bp)
Non-targeting 1	nontargeting-1	GTCCATGGGTGGAGTTACG
Non-targeting 2	nontargeting-2	GGACAAGTTGGACAGTTGC
Non-targeting 3	nontargeting-3	AACAATGGCTTCGTCGACT
Non-targeting 4	nontargeting-4	ACCTTATGTTCGCCGCCTC
Non-targeting 5	nontargeting-5	AGACACCGTAACTCGAATT
Non-targeting 6	nontargeting-6	CTCGTAGGTCAATCGCGGA
Non-targeting 7	nontargeting-7	GACGGTTCCGATAACCTTA
Non-targeting 8	nontargeting-8	GGATCTCTCAATATGGCGC
Non-targeting 9	nontargeting-9	AGCGCCGTTGTTGAGAAGA
Non-targeting 10	nontargeting-10	CCGCCGGCCCTAACTGACC
Non-targeting 11	nontargeting-11	TGACTAGACCCTTACGCGG
Non-targeting 12	nontargeting-12	ATCGGAGAGCCTGAAGTTG
Non-targeting 13	nontargeting-13	TCGGGCCGCTATATAGGCG
Non-targeting 14	nontargeting-14	AGAACTCAGCGGTCTCAAG
Non-targeting 15	nontargeting-15	ATGTCCGTTGTAGTCCTCG
Non-targeting 16	nontargeting-16	CCGGTCGCAATTGTTGCTA
Non-targeting 17	nontargeting-17	TGCCTATTCAGTAGGTCGT
Non-targeting 18	nontargeting-18	GTGCGGTAAGCGTGGCTAG
Non-targeting 19	nontargeting-19	CCATGGAGCCCAATCAATC
Non-targeting 20	nontargeting-20	GTGTAGTTCGGATTGATGA
Non-targeting 21	nontargeting-21	GGAATTGGAGACGATGCCG
Non-targeting 22	nontargeting-22	AGTGCCTCGTAGTTGTGGG
Non-targeting 23	nontargeting-23	CCTCGTTTCCACGTAATTC
Non-targeting 24	nontargeting-24	ATTTGATGTGGTGCTCACA
Non-targeting 25	nontargeting-25	ATTAGGTCGTTCTGTAACA
SCARB1	SCARB1_CCDS45008.1.ex7_12:125296469-125296492:-	GCTCTTCACGGTGTTCACG
	SCARB1_CCDS45008.1.ex8_12:125298861-125298884:+	ATGAAGGCACGTTCGCCGA
	SCARB1_CCDS45008.1.ex9_12:125299545-125299568:+	TAGTCGCTCTCCGAGCCGT
	SCARB1_CCDS45008.1.ex9_12:125299593-125299616:+	CGGTACTCGAGGAAGGACA

	SCARB1_CCDS45008.1.ex11_12:125348148-125348171:-	CCGTCGCTCATCAAGCAGC
CD81	CD81_CCDS7734.1.ex1_11:2411694-2411717:+	ACCACCAACCTCCTGTATC
	CD81_CCDS7734.1.ex1_11:2411744-2411767:-	GTGCACTCACCTACATAGA
	CD81_CCDS7734.1.ex2_11:2415376-2415399:+	GGCTGCTACGGGGCCATCC
	CD81_CCDS7734.1.ex3_11:2416244-2416267:-	TTGACAAAGCCCCAGATGC
	CD81_CCDS7734.1.ex4_11:2416641-2416664:-	CTTCACATCCTTGGCGATC
NAGPA	NAGPA_CCDS10527.1.ex7_16:5081801-5081824:-	TCGTGTGGCTGATTCGTAA
	NAGPA_CCDS10527.1.ex8_16:5083288-5083311:-	AGTTCGGGATCCGCCGCGA
	NAGPA_CCDS10527.1.ex8_16:5083517-5083540:+	CTCAACGGCCCGCGTCAGG
	NAGPA_CCDS10527.1.ex8_16:5083623-5083646:-	CAACCGCGAGCACGAGAGT
	NAGPA_CCDS10527.1.ex8_16:5083645-5083668:-	ACTGCACACGGGTGCGCGC
ADAM9	ADAM9_CCDS6112.1.ex4_8:38873645-38873668:+	ATTATCGGGGCTATGTGGA
	ADAM9_CCDS6112.1.ex4_8:38873691-38873714:+	TAGCGACTGTTTTGGACTC
	ADAM9_CCDS6112.1.ex5_8:38874916-38874939:+	GACTCAGCTACTTCGAGTA
	ADAM9_CCDS6112.1.ex8_8:38880694-38880717:+	CGAATTGTGCTAGTTGGAC
	ADAM9_CCDS6112.1.ex10_8:38884255-38884278:+	TTGGAATGAATCACGATGA
TSPAN3	TSPAN3_CCDS10292.1.ex2_15:77345235-77345258:-	AATTCACAACTACTCAGAC
	TSPAN3_CCDS10292.1.ex3_15:77346524-77346547:+	CTGTACATAATCAATAGCC
	TSPAN3_CCDS10292.1.ex3_15:77346543-77346566:-	CAACCCTGATGCTGCTAGC
	TSPAN3_CCDS10292.1.ex3_15:77346568-77346591:-	AAGTGTATAAGACCTACAA
	TSPAN3_CCDS10292.1.ex4_15:77348156-77348179:-	GTCACAGAAGTTGTTGTAG
RELL1	RELL1_CCDS43221.1.ex1_4:37636620-37636643:-	GTGGGCGGTGTTGTCGAGA
	RELL1_CCDS43221.1.ex1_4:37636638-37636661:-	GGCCATCATCTGCATACGG
	RELL1_CCDS43221.1.ex4_4:37650951-37650974:+	ATAGCCTTTCTTCTTAAGC
	RELL1_CCDS43221.1.ex4_4:37650996-37651019:+	ACCCATGATAAAGAACACA
	RELL1_CCDS43221.1.ex4_4:37651057-37651080:+	GTATCGTTGCTGGGCGACG
TYRO3	TYRO3_CCDS10080.1.ex2_15:41853743-41853766:+	CAGTGGAGCGCTCTGACGC
	TYRO3_CCDS10080.1.ex3_15:41854858-41854881:+	GAGGAACTACGAAGATCGG

	TYRO3_CCDS10080.1.ex8_15:41861134-41861157:-	GACGCACACACGTACGATC
	TYRO3_CCDS10080.1.ex8_15:41861203-41861226:+	TCTCATGACCGTGCAGGTG
	TYRO3_CCDS10080.1.ex9_15:41862270-41862293:+	CTTGGTGTGCTAACGGCCC
NAALAD	NAALADL1_CCDS31604.1.ex10_11:64820790-64820813:+	CATACAGCACGTAGCGATC
	NAALADL1_CCDS31604.1.ex13_11:64822138-64822161:-	CCCTGCCGACATCAACGAT
	NAALADL1_CCDS31604.1.ex14_11:64824857-64824880:+	ACACCCCCATATCGAGTCA
	NAALADL1_CCDS31604.1.ex14_11:64824940-64824963:+	CCGGTTGGCATAGACGAGG
	NAALADL1_CCDS31604.1.ex16_11:64825548-64825571:-	AGCCCAACGTCGTGGACAT
IGF1R	IGF1R_CCDS10378.1.ex1_15:99250922-99250945:-	TCGGTAATGACCGTGAGCT
	IGF1R_CCDS10378.1.ex1_15:99251001-99251024:-	CCAGCCGCGGATGACCGTG
	IGF1R_CCDS10378.1.ex3_15:99440123-99440146:-	GGAATACTTACTCCCCCGT
	IGF1R_CCDS10378.1.ex6_15:99454576-99454599:-	ATGATGCGATTCTTCGACG
	IGF1R_CCDS10378.1.ex7_15:99456407-99456430:+	CAGTACGCCGTTTACGTCA
BTN2A1	BTN2A1_CCDS4613.1.ex1_6:26459990-26460013:-	TCTTGGAAGTAACAGCGGT
	BTN2A1_CCDS4613.1.ex2_6:26463482-26463505:+	AGCCCCTCATTTCAATGAG
	BTN2A1_CCDS4613.1.ex2_6:26463496-26463519:+	ATGAGGGGCCATGAAGACG
	BTN2A1_CCDS4613.1.ex2_6:26463570-26463593:+	GTGGAGGGACCCCTACGGT
	BTN2A1_CCDS4613.1.ex2_6:26463614-26463637:+	TCTCCATGCCTGATGCAGA
F3	F3_CCDS53345.1.ex1_1:94998732-94998755:-	TGAACGGACTTTAGTCAGA
	F3_CCDS53345.1.ex2_1:95001510-95001533:+	GCCACTTACTCTCCAGGTA
	F3_CCDS53345.1.ex2_1:95001614-95001637:-	GAAGCAGACGTACTTGGCA
	F3_CCDS53345.1.ex2_1:95001701-95001724:-	TAGCACTAAGTCAGGAGAT
	F3_CCDS53345.1.ex4_1:95007090-95007113:+	CCTGAAGCGCCGGCCACCT
PTPRG	PTPRG_CCDS2895.1.ex1_3:61734576-61734599:-	GCCGTGTCTATTCTCGTGC
	PTPRG_CCDS2895.1.ex1_3:61734623-61734646:+	GGCTTCAGGCGACCCGTAC
	PTPRG_CCDS2895.1.ex6_3:62142832-62142855:-	CCACTATTTCGCTACACGG
	PTPRG_CCDS2895.1.ex9_3:62180790-62180813:-	CGCATGTCGTTCCGACACA
	PTPRG_CCDS2895.1.ex11_3:62189152-62189175:-	ACGAATCACCATCGGGCCC

SLC7A1	SLC7A1_CCDS9333.1.ex6_13:30097571-30097594:+	GGACGCCACGATCCCCACG
	SLC7A1_CCDS9333.1.ex9_13:30107000-30107023:+	TTTTCAGCCAGCACGCCGG
	SLC7A1_CCDS9333.1.ex9_13:30107040-30107063:+	TCCGTGAGAACTCCCCGAT
	SLC7A1_CCDS9333.1.ex9_13:30107066-30107089:+	CCTATCAGCTCGTCGAAGG
	SLC7A1_CCDS9333.1.ex9_13:30107098-30107121:-	GGTACTTCAAGCGTAGCGA
SLC6A1	SLC6A11_CCDS2602.1.ex0_3:10858063-10858086:+	CCGCGCGTCAAGCGCGACA
	SLC6A11_CCDS2602.1.ex0_3:10858096-10858119:+	CGCGGCCACTGGAACAACA
	SLC6A11_CCDS2602.1.ex1_3:10861156-10861179:-	AAACACCACGTAGGGAATC
	SLC6A11_CCDS2602.1.ex4_3:10885918-10885941:+	TGACGGGATCGAGCACATC
	SLC6A11_CCDS2602.1.ex5_3:10916689-10916712:-	CGTGACCCCTCGTATCAGG
SLC6A9	SLC6A9_CCDS30695.1.ex7_1:44468204-44468227:-	GTACTACCTAACCCCGCAG
	SLC6A9_CCDS30695.1.ex7_1:44468255-44468278:+	ACTCCGCGGACAAACAGAA
	SLC6A9_CCDS30695.1.ex9_1:44474141-44474164:-	GGAACACGCATGACTGCGC
	SLC6A9_CCDS30695.1.ex11_1:44476399-44476422:-	ACCTCTGCTATCGCAACGG
	SLC6A9_CCDS30695.1.ex11_1:44476459-44476482:-	AGTTTGTACTGACGAGCGT
SLC22A	SLC22A5_CCDS4154.1.ex0_5:131705780-131705803:+	GTGTTCCTGATAGCGACCC
	SLC22A5_CCDS4154.1.ex2_5:131719833-131719856:-	CATTCTTCCGGCCAAACCT
	SLC22A5_CCDS4154.1.ex2_5:131719851-131719874:+	ATGTGCTGTTCGTGACCAT
	SLC22A5_CCDS4154.1.ex2_5:131719943-131719966:-	GATCTGGCCCATGCCTACA
	SLC22A5_CCDS4154.1.ex4_5:131722820-131722843:-	TCACTCGGGTCAAAGATAG
TTYH3	TTYH3_CCDS34588.1.ex2_7:2686786-2686809:+	AGTGGGATTCTACGGCAAC
	TTYH3_CCDS34588.1.ex2_7:2686851-2686874:+	ACGCCAACCGCACGGTGGC
	TTYH3_CCDS34588.1.ex3_7:2687136-2687159:-	AGCCTCTGTACGGCTCGCA
	TTYH3_CCDS34588.1.ex3_7:2687201-2687224:-	CCGCCGTGTTCCTCCAAAA
	TTYH3_CCDS34588.1.ex3_7:2687255-2687278:+	TCTACGACTGGTACAGGTG

IFNAR2	IFNAR2_CCDS13621.1.ex2_21:34617336-34617359:-	GTATAGTGAGTTGGTACAA
	IFNAR2_CCDS13621.1.ex3_21:34619075-34619098:+	TTGTGACCTCACAGATGAG
	IFNAR2_CCDS13621.1.ex3_21:34619091-34619114:+	GAGTGGAGAAGCACACACG
	IFNAR2_CCDS13621.1.ex3_21:34619123-34619146:-	CCGCTGAATCCTTCTAGGA
	IFNAR2_CCDS13621.1.ex5_21:34624963-34624986:+	AGCATAAACCCGAAATAAA
TNFRSF	TNFRSF10B_CCDS6035.1.ex2_8:22884784-22884807:-	TCCAGAGCTCACAACGACC
	TNFRSF10B_CCDS6035.1.ex6_8:22888283-22888306:+	GGTGCAGCGCAAGCAGAAA
	TNFRSF10B_CCDS6035.1.ex6_8:22888314-22888337:-	ACAGGACTATAGCACTCAC
	TNFRSF10B_CCDS6035.1.ex6_8:22888363-22888386:-	GACACCATATCTCAGAAGA
	TNFRSF10B_CCDS6035.1.ex8_8:22926274-22926297:+	CGCGGCGACAACGAGCACA
CADM4	CADM4_CCDS12627.1.ex4_19:44130365-44130388:-	CGGTTTCGTGTGGACCGTA
	CADM4_CCDS12627.1.ex4_19:44130384-44130407:-	GAGCGTGGCAAGCACAGTA
	CADM4_CCDS12627.1.ex5_19:44130977-44131000:+	CGGACGGGACCGCGGAACG
	CADM4_CCDS12627.1.ex6_19:44131320-44131343:+	TACCGTGAGCGTGGCAATC
	CADM4_CCDS12627.1.ex7_19:44131843-44131866:+	AACTATGGACCCATCATAC
FGFRL1	FGFRL1_CCDS3344.1.ex1_4:1016136-1016159:+	GCTTCCGCGTGCTGCCGCA
	FGFRL1_CCDS3344.1.ex1_4:1016240-1016263:+	AACTACACCCTCGTCGTGC
	FGFRL1_CCDS3344.1.ex3_4:1017642-1017665:+	GGGTGATCGCACGGCCCGT
	FGFRL1_CCDS3344.1.ex3_4:1017701-1017724:-	CCACGTGATGTCGGGCCGA
	FGFRL1_CCDS3344.1.ex3_4:1017829-1017852:-	GCGCGGTTCGACACGCGGC
MPEG1	MPEG1_CCDS41650.1.ex0_11:58979516-58979539:-	CTCAAACCGAACCAACTCC
	MPEG1_CCDS41650.1.ex0_11:58979702-58979725:-	CACCACCAGTGTCGACGCT
	MPEG1_CCDS41650.1.ex0_11:58979774-58979797:-	TCTAGAGAACAACCAGACG
	MPEG1_CCDS41650.1.ex0_11:58979829-58979852:-	CAACTTTAGAGCTAAGCTC

	MPEG1_CCDS41650.1.ex0_11:58979854-58979877:+	GATTTTGACTGTGTAGACG
ICAM3	ICAM3_CCDS12235.1.ex3_19:10445864-10445887:-	GCGACAGTCATGAACCACG
	ICAM3_CCDS12235.1.ex4_19:10446343-10446366:+	ACCAAAGGTTCGGAGCTGG
	ICAM3_CCDS12235.1.ex4_19:10446414-10446437:+	GTTCTGTGCGGCATGAGAA
	ICAM3_CCDS12235.1.ex4_19:10446556-10446579:+	GGGCGACCCATCCTCCACT
	ICAM3_CCDS12235.1.ex5_19:10449480-10449503:-	ACGTCCCTATCAAAGGAGC
APP	APP_CCDS13576.1.ex9_21:27354737-27354760:+	GTCTCGAGATACTTGTCAA
	APP_CCDS13576.1.ex13_21:27423437-27423460:+	CGGAACTTGTCAATTCCGC
	APP_CCDS13576.1.ex13_21:27423467-27423490:+	ATGCCGTAGTCATGCAAGT
	APP_CCDS13576.1.ex14_21:27425550-27425573:-	TGGCACACCGTCGCCAAAG
	APP_CCDS13576.1.ex15_21:27462250-27462273:+	GGCTCACCTAAGCAGCGGT
VCAM1	VCAM1_CCDS55617.1.ex2_1:101188594-101188617:+	ATTCATTTGAGTGGCCCTC
	VCAM1_CCDS55617.1.ex2_1:101188663-101188686:-	TATCTCCAGCCTGTCAAAT
	VCAM1_CCDS55617.1.ex2_1:101188699-101188722:+	GATCATCTCATGAAGAGTC
	VCAM1_CCDS55617.1.ex3_1:101190211-101190234:+	CATCCACAAAGCTGCAAGA
	VCAM1_CCDS55617.1.ex5_1:101196830-101196853:-	GGTCAAGGGGGTACACGCT
TMEM21	TMEM219_CCDS42145.1.ex1_16:29974757-29974780:-	GTCCCATTCCTGGGACACA
	TMEM219_CCDS42145.1.ex1_16:29974823-29974846:-	CCGTCTCCGAAGTTCAAGG
	TMEM219_CCDS42145.1.ex1_16:29974878-29974901:+	ACAGTCCTGGGAAGTCAGA
	TMEM219_CCDS42145.1.ex2_16:29979408-29979431:-	ACAGCACTAAAATATAGGC
	TMEM219_CCDS42145.1.ex2_16:29979451-29979474:-	TGAGCAGGATATGGGTGGC
C11orf	C11orf24_CCDS8180.1.ex0_11:68029808-68029831:+	GCTACAGTCTGAGCACGTG
	C11orf24_CCDS8180.1.ex0_11:68029868-68029891:+	CTCTTTGGCACTTGTGCGA
	C11orf24_CCDS8180.1.ex0_11:68029910-68029933:-	CCCGTCCACTACCGCCACT

	C11orf24_CCDS8180.1.ex0_11:68030290-68030313:-	ACGTCTGAGGATGTAACCA
	C11orf24_CCDS8180.1.ex0_11:68030356-68030379:-	GTCCCTAACAAAATGTGGA
PTPRC	PTPRC_CCDS1397.2.ex1_1:198661480-198661503:+	GCCCAACACCTTCCCCCAC
	PTPRC_CCDS1397.2.ex1_1:198661493-198661516:-	TATTAATTCTTACCAGTGG
	PTPRC_CCDS44291.2.ex2_1:198663147-198663170:-	TTGCTCCTCAGCTTGCAGA
	PTPRC_CCDS44291.2.ex2_1:198663174-198663197:-	GAGGTTTGGTGACTTGGAT
	PTPRC_CCDS44291.2.ex2_1:198663209-198663232:+	TGACAGTTCAGCCTTCAGT
STAB1	STAB1_CCDS33768.1.ex8_3:52537869-52537892:-	CAAGTGGCAGACCGGTCGC
	STAB1_CCDS33768.1.ex17_3:52540749-52540772:+	TAGACGTGATGGCCGCCAA
	STAB1_CCDS33768.1.ex19_3:52541982-52542005:-	TTAGCACATTGAGCCCCGT
	STAB1_CCDS33768.1.ex31_3:52547896-52547919:+	GGTCTTACTGCCCCCCCGA
	STAB1_CCDS33768.1.ex34_3:52548766-52548789:-	CGACAGACCAATCAGCGAG
MFAP3L	MFAP3L_CCDS34103.1.ex0_4:170912929-170912952:-	TTTGTGAGGCATACTCCAG
	MFAP3L_CCDS34103.1.ex0_4:170913005-170913028:-	GCCTCTCATTATGAACTGC
	MFAP3L_CCDS34103.1.ex0_4:170913068-170913091:+	AACTCCATGGTTTTGAACT
	MFAP3L_CCDS34103.1.ex0_4:170913283-170913306:-	TACTACATGGTCGTGTGCC
	MFAP3L_CCDS34103.1.ex0_4:170913346-170913369:-	TACGGCACCGTGAACAACA
CDHR2	CDHR2_CCDS34297.1.ex2_5:175995762-175995785:+	CGCTGTCACTCCGAAAACT
	CDHR2_CCDS34297.1.ex10_5:176003160-176003183:-	TCCGGGTCGTAGACCACCA
	CDHR2_CCDS34297.1.ex11_5:176004500-176004523:+	GTGAGAGTATCCGCGCTGG
	CDHR2_CCDS34297.1.ex13_5:176005039-176005062:+	GACCCAGACACGGGCGCGT
	CDHR2_CCDS34297.1.ex14_5:176005524-176005547:+	ATCAACGACAATGCACCCG
IL17RB	IL17RB_CCDS2874.1.ex2_3:53883701-53883724:+	TACAACATGATCTAATCCC
	IL17RB_CCDS2874.1.ex2_3:53883753-53883776:+	TACAACTAGTGTTGCAACA

	IL17RB_CCDS2874.1.ex3_3:53886025-53886048:-	GCCTTCAACAAGCGGATGC
	IL17RB_CCDS2874.1.ex3_3:53886133-53886156:-	ACTTTACCACCAGAGGGTC
	IL17RB_CCDS2874.1.ex4_3:53886920-53886943:-	GACTGTGTTCAGCTCTACA
NOTCH2	NOTCH2_CCDS908.1.ex14_1:120480507-120480530:+	ACATTGGGCACGTCACAAT
	NOTCH2_CCDS908.1.ex21_1:120497749-120497772:-	TCCGCTGTATATGCCCCGA
	NOTCH2_CCDS908.1.ex21_1:120497790-120497813:+	CATGTTGCACCCTTGCGAC
	NOTCH2_CCDS908.1.ex23_1:120506247-120506270:+	GCGACCATCGTTCAGGCAA
	NOTCH2_CCDS908.1.ex24_1:120508123-120508146:-	AGTACTCCGTGTCTGAATG
MMP15	MMP15_CCDS10792.1.ex1_16:58071496-58071519:-	GTCTCTTCGTCGAGCACAC
	MMP15_CCDS10792.1.ex2_16:58072202-58072225:-	TTTCACTCGTACCCCGAAC
	MMP15_CCDS10792.1.ex2_16:58072242-58072265:+	GGAAGCGCTACGCCCTCAC
	MMP15_CCDS10792.1.ex3_16:58073811-58073834:-	GCGCACCGCCTCCATCGAG
	MMP15_CCDS10792.1.ex3_16:58073965-58073988:+	CGCCGTTTGATGGCACCGG
LRP10	LRP10_CCDS9578.1.ex3_14:23344283-23344306:-	GTTAAGCGCTCTGAGCCAC
	LRP10_CCDS9578.1.ex4_14:23344896-23344919:+	GCATGTGTATGACGGCCCT
	LRP10_CCDS9578.1.ex4_14:23344919-23344942:-	GTAGTCGGGAGCTCTCAGG
	LRP10_CCDS9578.1.ex4_14:23344999-23345022:-	GTGGTAGGACACAACAGCC
	LRP10_CCDS9578.1.ex4_14:23345160-23345183:+	ACAGCGCTGTGACGGCTCA
CXCL16	CXCL16_CCDS11052.1.ex2_17:4641701-4641724:-	TGAGCTGTCTTGATCTCAA
	CXCL16_CCDS11052.1.ex2_17:4641720-4641743:+	CATCAATTCCTGAACCCAT
	CXCL16_CCDS11052.1.ex3_17:4642075-4642098:-	TCGGTGTCTATACTACACG
	CXCL16_CCDS11052.1.ex3_17:4642129-4642152:+	ATGAACTGAACCGATGGCG
	CXCL16_CCDS11052.1.ex4_17:4642546-4642569:+	CGGCTAACCTGGCTGAGTC
CLSTN3	CLSTN3_CCDS8575.1.ex1_12:7285724-7285747:-	TTACCTGCATAGCGCAGCG
	CLSTN3_CCDS8575.1.ex3_12:7287985-7288008:+	TATCGTGCGGCTGTGACAG
	CLSTN3_CCDS8575.1.ex3_12:7288015-7288038:+	TACGATCGCATCCTGCGGG
	CLSTN3_CCDS8575.1.ex8_12:7293857-7293880:-	CTCGAGGTTCAGAGCGTAG
	CLSTN3_CCDS8575.1.ex8_12:7293885-7293908:+	CAGTCACACTCTATACCGA

LY6E	LY6E_CCDS6394.1.ex1_8:144102788-144102811:+	CCGACCATCTGCTCCGACC
	LY6E_CCDS6394.1.ex1_8:144102802-144102825:-	ACGCAGTAGTTGTCCTGGT
	LY6E_CCDS6394.1.ex1_8:144102828-144102851:+	TGTCTGCTAGTGCCGGCAT
	LY6E_CCDS6394.1.ex2_8:144103001-144103024:-	ACAGGTCTTGCTCAGGCTG
	LY6E_CCDS6394.1.ex2_8:144103036-144103059:-	ACATTGACGCCTTCTGGGA
IFNGR2	IFNGR2_CCDS33544.1.ex2_21:34793782-34793805:-	AACCATTTACTGTCGGTGC
	IFNGR2_CCDS33544.1.ex2_21:34793807-34793830:+	GCCGACATCATGTCCATAG
	IFNGR2_CCDS33544.1.ex2_21:34793821-34793844:-	TGTGTACAATTCACCCCTA
	IFNGR2_CCDS33544.1.ex2_21:34793921-34793944:-	TGCTCCCAGCTCAGCTCGA
	IFNGR2_CCDS33544.1.ex2_21:34793935-34793958:+	GGGAGCACTCCATTCTGCC
FGFR2	FGFR2_CCDS31298.1.ex12_10:123298140-123298163:-	GTGTAGTGGAGAATGAATA
	FGFR2_CCDS31298.1.ex13_10:123310813-123310836:-	TTAAGCAGGAGCATCGCAT
	FGFR2_CCDS31298.1.ex13_10:123310843-123310866:-	CCATGCGGTGGCTGAAAAA
	FGFR2_CCDS31298.1.ex13_10:123310959-123310982:-	CAAAACAGGAGCACCATAC
	FGFR2_CCDS31298.1.ex16_10:123353221-123353244:-	ATACCACATTAGAGCCAGA
NRP1	NRP1_CCDS31179.1.ex6_10:33545287-33545310:-	ACAGCGCGATAGCAAAAGA
	NRP1_CCDS31179.1.ex6_10:33545338-33545361:+	ATGAGGATCGGATTCGACC
	NRP1_CCDS31179.1.ex7_10:33552601-33552624:-	GATGTTCTGTCGCTACGAC
	NRP1_CCDS31179.1.ex8_10:33559770-33559793:-	TGTATTAGGTATGACTACG
	NRP1_CCDS31179.1.ex9_10:33619644-33619667:+	TGTCCTCCAAATCGAAGTG
TNFRSF	TNFRSF14_CCDS44046.1.ex1_1:2489221-2489244:+	AGGAGGACGAGTACCCAGT
	TNFRSF14_CCDS44046.1.ex1_1:2489258-2489281:-	CACCTACCTGGACTGCACT
	TNFRSF14_CCDS44046.1.ex2_1:2489783-2489806:+	ATCGTGTGAAGGAGGCCTG
	TNFRSF14_CCDS44046.1.ex3_1:2491257-2491280:-	TCGCGCGCAGGCCCATGGC
	TNFRSF14_CCDS44046.1.ex3_1:2491330-2491353:+	CTTCTGCATCGTCCAGGAC
CDH5	CDH5_CCDS10804.1.ex0_16:66413364-66413387:+	CCGGCGCCAAAAGAGAGAT

	CDH5_CCDS10804.1.ex1_16:66420922-66420945:+	TCATGACGTGAACGACAAC
	CDH5_CCDS10804.1.ex2_16:66422328-66422351:-	CACAGACCAGAATTATCGA
	CDH5_CCDS10804.1.ex4_16:66424411-66424434:+	ATCTTGCGGGGCGACTACC
	CDH5_CCDS10804.1.ex5_16:66426096-66426119:-	CTCATGTATCGGAGGTCGA
TYROBP	TYROBP_CCDS12482.1.ex1_19:36398150-36398173:+	GTTTCCGGGTCGCTGCTGG
	TYROBP_CCDS12482.1.ex2_19:36398359-36398382:+	TCGCCCCCGAGGGACCAGC
	TYROBP_CCDS12482.1.ex2_19:36398383-36398406:-	GCCCTGGCCGTGTACTTCC
	TYROBP_CCDS12482.1.ex2_19:36398404-36398427:+	AATGAGCACTGTCAGCACC
	TYROBP_CCDS12482.1.ex2_19:36398431-36398454:-	GTGCTGGCAGGGATCGTGA
CD164	CD164_CCDS47462.1.ex4_6:109700839-109700862:+	CAGCTGTTTCGACCTTCAC
	CD164_CCDS47462.1.ex5_6:109703402-109703425:-	TCACCACTCCGGCACCAGG
	CD164_CCDS47462.1.ex5_6:109703424-109703447:-	CCGGTGACGTCCCTCCCGC
	CD164_CCDS47462.1.ex5_6:109703442-109703465:-	TCCAACGTAACCTCGGCGC
	CD164_CCDS47462.1.ex5_6:109703475-109703498:+	AGTCGTCACGTTCGGGTGC
CRIM1	CRIM1_CCDS1783.1.ex0_2:36583673-36583696:+	TTACGGAACCTGCGACCGG
	CRIM1_CCDS1783.1.ex0_2:36583733-36583756:-	CCCGCTTCGTACTCGGTGA
	CRIM1_CCDS1783.1.ex3_2:36669746-36669769:+	CCCCTTTCAGTTTTCGGCG
	CRIM1_CCDS1783.1.ex4_2:36691731-36691754:-	CGCCACGAGAGACTATGCG
	CRIM1_CCDS1783.1.ex5_2:36704168-36704191:+	AGAGGTACTACGTGCCCGA
ATRAID	ATRAID_CCDS1741.1.ex2_2:27438218-27438241:-	CTGATGAAAGTTTGGACCA
	ATRAID_CCDS1741.1.ex3_2:27438369-27438392:+	ACTTGGCCAACACCTTCCG
	ATRAID_CCDS1741.1.ex3_2:27438382-27438405:-	AGCTGAGTAAAGCCACGGA
	ATRAID_CCDS1741.1.ex3_2:27438404-27438427:-	TACCCTTACTCACAGAGTC
	ATRAID_CCDS1741.1.ex4_2:27438522-27438545:+	TCCTGGAGGAATTAATGCC
RAMP1	RAMP1_CCDS2522.1.ex1_2:238785893-238785916:-	GGCACCGTAGTTAGCCTCC
	RAMP1_CCDS2522.1.ex1_2:238785947-238785970:+	GTAGACATGGAGGCCGTCG
	RAMP1_CCDS2522.1.ex1_2:238785969-238785992:+	AGACGCTGTGGTGTGACTG

	RAMP1_CCDS2522.1.ex1_2:238785982-238786005:+	TGACTGGGGCAGGACCATC
	RAMP1_CCDS2522.1.ex2_2:238820157-238820180:+	TCCATCCGCAGGAGCTACA
TNFRSF	TNFRSF12A_CCDS10489.1.ex1_16:3071242-3071265:-	TTGTCCAGGTCCGCGCTCC
	TNFRSF12A_CCDS10489.1.ex1_16:3071267-3071290:+	ATGGACTGCGCGTCTTGCA
	TNFRSF12A_CCDS10489.1.ex1_16:3071294-3071317:+	CCGCACAGCGACTTCTGCC
	TNFRSF12A_CCDS10489.1.ex2_16:3071575-3071598:-	TGGGCCAAAGCAGCCGGAA
	TNFRSF12A_CCDS10489.1.ex2_16:3071593-3071616:-	TCAGAGCGCCCCCAAGGAT
VSIG4	VSIG4_CCDS14383.1.ex6_X:65253347-65253370:+	TCACGACTTGGTTGCCATC
	VSIG4_CCDS14383.1.ex6_X:65253376-65253399:-	CTACACGTGTGAAGTCACC
	VSIG4_CCDS14383.1.ex6_X:65253417-65253440:-	TCCCTCCAATTGAGCACCC
	VSIG4_CCDS14383.1.ex6_X:65253489-65253512:-	TCTGGAGACCATATCCAGC
	VSIG4_CCDS14383.1.ex6_X:65253545-65253568:-	TGAAGTGGCTGGTACAACG
PDGFRA	PDGFRA_CCDS3495.1.ex1_4:55127480-55127503:-	AACCCTGTGTGGGCCGCCG
	PDGFRA_CCDS3495.1.ex4_4:55133511-55133534:+	TACACTTTGACGGTCCCCG
	PDGFRA_CCDS3495.1.ex7_4:55138577-55138600:+	ACTTGGTCGATGATCACCA
	PDGFRA_CCDS3495.1.ex8_4:55139823-55139846:+	ACCATCGCCGTGCGATGCC
	PDGFRA_CCDS3495.1.ex9_4:55140698-55140721:-	CACCGTGAGTTCAGAACGC
LRP5	LRP5_CCDS8181.1.ex1_11:68115336-68115359:+	AACCGCCGGGACGTACGGC
	LRP5_CCDS8181.1.ex1_11:68115552-68115575:+	GACGGCCTCGCCTGCGACT
	LRP5_CCDS8181.1.ex5_11:68153898-68153921:+	ATCGACTACGACCCGCTAG
	LRP5_CCDS8181.1.ex8_11:68173991-68174014:+	AACCAACCCGTGTGCGGAC
	LRP5_CCDS8181.1.ex8_11:68174193-68174216:+	ATCCCGCTCACGGGCGTCA
LTBR	LTBR_CCDS59233.1.ex2_12:6494199-6494222:+	AGCTAAATGTAGCCGCATC
	LTBR_CCDS59233.1.ex2_12:6494235-6494258:-	TAGGAATTCTCGGCACATG
	LTBR_CCDS59233.1.ex3_12:6494479-6494502:-	TCGCAGTGTGTACACTCGA
	LTBR_CCDS59233.1.ex3_12:6494523-6494546:+	GCACTGAAGCCGAGCTCAA
	LTBR_CCDS59233.1.ex5_12:6495554-6495577:-	TTTGCAGGTTGTGTCGGAC
CD52	CD52_CCDS30647.1.ex1_1:26646692-26646715:-	GAGGGGCTGCTGGTTTGGC

	CD52.2	GAGGCTGATGGTGAGTAGG
	CD52.3	CAGGAGGCTGATGGTGAGT
	CD52.4	CTACTCACCATCAGCCTCC
	CD52.5	ACCATCAGCCTCCTGGTTA
IFNAR1	IFNAR1_CCDS13624.1.ex4_21:34715965-34715988:+	TGTATAAAGACCACAGGTA
	IFNAR1_CCDS13624.1.ex6_21:34721416-34721439:+	TGGAAACCATTTGTATAAA
	IFNAR1_CCDS13624.1.ex6_21:34721522-34721545:-	TCCATCAGATGCTTGTACG
	IFNAR1_CCDS13624.1.ex7_21:34721753-34721776:+	TCGGTGCTCCAAAACAGTC
	IFNAR1_CCDS13624.1.ex7_21:34721783-34721806:-	GTGGATAATCCTGGATCAC
IL6R	IL6R_CCDS1067.1.ex1_1:154401682-154401705:+	GCGTGCTGACCAGTCTGCC
	IL6R_CCDS1067.1.ex1_1:154401862-154401885:+	ACTATTCATGCTACCGGGC
	IL6R_CCDS1067.1.ex2_1:154402999-154403022:-	CACAAACAACATTGCTGAG
	IL6R_CCDS1067.1.ex3_1:154406985-154407008:+	TCTAACAGTCAGAACAGTC
	IL6R_CCDS1067.1.ex3_1:154407061-154407084:+	GCCAGTTAGCAGTCCCGGA
IL4R	IL4R_CCDS10629.1.ex2_16:27356204-27356227:+	CCTGAGAACAACGGAGGCG
	IL4R_CCDS10629.1.ex3_16:27357823-27357846:-	AGCAGAGTGTCGGAGACAT
	IL4R_CCDS10629.1.ex3_16:27357860-27357883:-	GTAATTGTCAGGGGGATAC
	IL4R_CCDS10629.1.ex6_16:27370275-27370298:-	TATAGCCACGAGGCGGCTG
	IL4R_CCDS10629.1.ex7_16:27372126-27372149:-	TCAGATACATACGGGCACT
TMPRSS2	TMPRSS2_CCDS33564.1.ex8_21:42852427-42852450:-	TGGAACGAGAACTACGGGC
	TMPRSS2_CCDS33564.1.ex8_21:42852502-42852525:+	GATGAAGTTTGGTCCGTAG
	TMPRSS2_CCDS33564.1.ex9_21:42860324-42860347:+	CGATTCTCGTCCTCCCCGC
	TMPRSS2_CCDS33564.1.ex11_21:42866285-42866308:+	GGTGCACACTGTCCCGGAT
	TMPRSS2_CCDS33564.1.ex11_21:42866354-42866377:-	GTGCCCCAGTACGCCCCGA
SEMA4B	SEMA4B_CCDS45347.1.ex1_15:90760754-90760777:-	TCTCGAGCACCCACGTACA
	SEMA4B_CCDS45347.1.ex3_15:90763020-90763043:-	GTAGTTTTGACAGTCGCGC
	SEMA4B_CCDS45347.1.ex3_15:90763106-90763129:-	CTCACGATGTAGGTACACA

	SEMA4B_CCDS45347.1.ex4_15:90764286-90764309:-	GACTTGAAATTCGGGTCGA
	SEMA4B_CCDS45347.1.ex5_15:90764645-90764668:-	TTTGGCTCCGCGAGATGGC
TCTN3	TCTN3_CCDS31258.2.ex7_10:97446269-97446292:+	GTCATGCTTCTTGGAACCT
	TCTN3_CCDS31258.2.ex8_10:97446806-97446829:+	ACTGTGAAGTTATAGTAAG
	TCTN3_CCDS31258.2.ex11_10:97452701-97452724:-	CATGGATTCTAATGGAATC
	TCTN3_CCDS31258.2.ex12_10:97453174-97453197:+	CAGCAGCAATTTATATCGC
	TCTN3_CCDS31258.2.ex13_10:97453481-97453504:-	GAGGCGACTGCAACTCGCC
TNFRSF1A	TNFRSF1A_CCDS8542.1.ex4_12:6440030-6440053:+	GCCCTTAACATTCTCAATC
	TNFRSF1A_CCDS8542.1.ex6_12:6442592-6442615:+	ACTCCAATAATGCCGGTAC
	TNFRSF1A_CCDS8542.1.ex7_12:6442902-6442925:-	AGCTGCTCCAAATGCCGAA
	TNFRSF1A_CCDS8542.1.ex7_12:6443005-6443028:+	GGACAGTCATTGTACAAGT
	TNFRSF1A_CCDS8542.1.ex8_12:6443343-6443366:-	GGACTGGTCCCTCACCTAG
GFRA1	GFRA1_CCDS44481.1.ex5_10:117884894-117884917:-	CTCCGGCAGTTCTTTGACA
	GFRA1_CCDS44481.1.ex5_10:117884984-117885007:-	ATTTGCAAGAAGTACAGGT
	GFRA1_CCDS44481.1.ex7_10:118029004-118029027:+	TCTGCTTACCTGATATGAA
	GFRA1_CCDS44481.1.ex7_10:118029029-118029052:-	TTGTCAGATATATTCCGGG
	GFRA1_CCDS44481.1.ex8_10:118030507-118030530:-	CGCACGCTAAGGCAGTGCG
SDC3	SDC3_CCDS30661.1.ex2_1:31349599-31349622:-	GACCACAGTGGCTACGGCA
	SDC3_CCDS30661.1.ex2_1:31349650-31349673:+	GTGGTCCTTATAACAGCAG
	SDC3_CCDS30661.1.ex2_1:31349885-31349908:+	CTTCAAATGGTGTGCCCAC
	SDC3_CCDS30661.1.ex3_1:31351472-31351495:+	CGAGCCCGACCCCGAGTAG
	SDC3_CCDS30661.1.ex3_1:31351541-31351564:-	TTCGAGAGACCCGTGGACC
HFE2	HFE2_CCDS910.1.ex2_1:145416408-145416431:+	TCAATGGAGGTGACCGACC
	HFE2_CCDS910.1.ex2_1:145416422-145416445:-	CAAACTGGATCCCCCAGGT
	HFE2_CCDS910.1.ex2_1:145416444-145416467:+	CGATTCAAACTGCTAACCC

	HFE2_CCDS910.1.ex2_1:145416482-145416505:-	TGTGCCAATGTAGGCAGCT
	HFE2_CCDS910.1.ex2_1:145416507-145416530:+	TAATCATTCGGCAGACAGC
OSMR	OSMR_CCDS3928.1.ex1_5:38876350-38876373:-	AAACTCTGACGCGTAGAAT
	OSMR_CCDS3928.1.ex1_5:38876453-38876476:+	ACATCCAATGTCATCTGGG
	OSMR_CCDS3928.1.ex2_5:38881821-38881844:+	GCCAAATTTCTGGAGCAAC
	OSMR_CCDS3928.1.ex3_5:38883926-38883949:+	GTACAAGATTCTACTGGAC
	OSMR_CCDS3928.1.ex3_5:38884119-38884142:-	TCCCTTTATTCCTAATGAA
LRP6	LRP6_CCDS8647.1.ex13_12:12315125-12315148:-	CGCGTTGGACCCTGCCGAA
	LRP6_CCDS8647.1.ex13_12:12315202-12315225:-	CGAATTGAGGTGTCAAAGT
	LRP6_CCDS8647.1.ex17_12:12334179-12334202:+	TCTATAAATGAACGGCGTA
	LRP6_CCDS8647.1.ex17_12:12334234-12334257:-	TAGATTACGATCCTGTGGA
	LRP6_CCDS8647.1.ex21_12:12397343-12397366:-	TTATTGTCCCCCGATGGGC
MARCO	MARCO_CCDS2124.1.ex1_2:119726771-119726794:-	TAGATGACCACCACAGCTA
	MARCO_CCDS2124.1.ex2_2:119727685-119727708:-	GCGCCTGCAGATTCAGAAC
	MARCO_CCDS2124.1.ex2_2:119727851-119727874:-	TCATGGCTGACGCGGACCC
	MARCO_CCDS2124.1.ex2_2:119727892-119727915:+	ACAACTTCACTCAGAACCC
	MARCO_CCDS2124.1.ex4_2:119731913-119731936:+	AAGGTCACAAGGGGGCCAT
CD55	CD55_CCDS31006.1.ex1_1:207495775-207495798:+	GGCCGTACAAGTTTTCCCG
	CD55_CCDS31006.1.ex1_1:207495832-207495855:+	GTGAAAATTCCTGGCGAGA
	CD55_CCDS31006.1.ex2_1:207497892-207497915:+	ATACTTTTAGGTAGCTGCG
	CD55_CCDS31006.1.ex2_1:207497974-207497997:-	ATTCCACAACAGTACCGAC
	CD55_CCDS31006.1.ex3_1:207499001-207499024:+	GTCAGATTGATGTACCAGG
ACVR1	ACVR1_CCDS2206.1.ex4_2:158626934-158626957:+	CATGACTTCTCATCACGGG
	ACVR1_CCDS2206.1.ex5_2:158630670-158630693:-	TGGATCATTCGTGTACATC
	ACVR1_CCDS2206.1.ex6_2:158634678-158634701:-	GGAGTATGGCACTATCGAA
	ACVR1_CCDS2206.1.ex6_2:158634758-158634781:-	TGTCTTTTAGCCTGCCTGC

	ACVR1_CCDS2206.1.ex7_2:158636891-158636914:-	AGCCGTGGAGTGCTGCCAA
ACVR1B	ACVR1B_CCDS44893.2.ex1_12:52370099-52370122:+	CTTGACTCAGGTCACCTCA
	ACVR1B_CCDS44893.2.ex1_12:52370216-52370239:-	CTGATGATAGTTAATGACA
	ACVR1B_CCDS44893.2.ex1_12:52370243-52370266:+	TATCACAACCGCCAGAGAC
	ACVR1B_CCDS44893.2.ex2_12:52374758-52374781:-	ACTGTGCGCTGGACAAAGA
	ACVR1B_CCDS44893.2.ex2_12:52374832-52374855:+	TATGGCGGGGCCGCTGGAG
CSF1R	CSF1R_CCDS4302.1.ex13_5:149449841-149449864:+	TATGACGCTTACCTCTGGG
	CSF1R_CCDS4302.1.ex15_5:149452927-149452950:-	GGTTTTAACTGGACCTACC
	CSF1R_CCDS4302.1.ex17_5:149457704-149457727:+	AAGTTAACATCAACGCTGC
	CSF1R_CCDS4302.1.ex18_5:149459613-149459636:-	TCCGGCTGAAAGTGCAGAA
	CSF1R_CCDS4302.1.ex18_5:149459860-149459883:-	AACGTGCTAGCACAGGAGG
EFNA1	EFNA1_CCDS1091.1.ex1_1:155103810-155103833:-	TAGTCCTCATTCCGGAACC
	EFNA1_CCDS1091.1.ex1_1:155103834-155103857:-	TCATTCAGCTGCACATGTA
	EFNA1_CCDS1091.1.ex1_1:155103877-155103900:+	CACTATGAAGATCACTCTG
	EFNA1_CCDS1091.1.ex1_1:155103965-155103988:-	GGACTTGGTCCTTGGACTG
	EFNA1_CCDS1091.1.ex1_1:155104005-155104028:-	TCCGGGCCATGCTTGGCAC
TP53I13	TP53I13_CCDS42289.1.ex2_17:27896350-27896373:+	CCTACACACGAGTGAGCCC
	TP53I13_CCDS42289.1.ex3_17:27898678-27898701:-	GGGTTAGCCATACAAGCAT
	TP53I13_CCDS42289.1.ex4_17:27899025-27899048:+	CGAGCCAACTCCCGGTAGA
	TP53I13_CCDS42289.1.ex5_17:27899202-27899225:+	CACAGAGGTGACATCTCAA
	TP53I13_CCDS42289.1.ex5_17:27899235-27899258:+	CTCTTCTAGTGGTGCCAAG
FLT1	FLT1_CCDS53860.1.ex9_13:29007982-29008005:+	ACTCTCGTGTTCAAGGGAG
	FLT1_CCDS53860.1.ex10_13:29008313-29008336:+	AGATTATGCGTTTTCCATC

	FLT1_CCDS53860.1.ex11_13:29012379-29012402:+	TAGGTGACGTAACCCGGCA
	FLT1_CCDS53860.1.ex12_13:29041182-29041205:-	GCATAACTAAATCTGCCTG
	FLT1_CCDS53860.1.ex13_13:29041656-29041679:-	GACACTGCATCTCCAATGC
STOM	STOM_CCDS6830.1.ex0_9:124103614-124103637:+	GATTCAGTGATGACCATGG
	STOM_CCDS6830.1.ex0_9:124103628-124103651:-	GCTCTGAAAGAAGCCTCCA
	STOM_CCDS6830.1.ex0_9:124103669-124103692:-	CATAGGTTATTGCAGCCGA
	STOM_CCDS6830.1.ex2_9:124111490-124111513:+	CGGGTTGCTGAGTCAGCGT
	STOM_CCDS6830.1.ex2_9:124111528-124111551:-	CGCGTTCAGAATGCAACCC
CCDC47	CCDC47_CCDS11643.1.ex7_17:61838268-61838291:+	ATACCACTTGATCACTCAC
	CCDC47_CCDS11643.1.ex8_17:61838618-61838641:+	CACTCGACCAGAACACCAC
	CCDC47_CCDS11643.1.ex8_17:61838693-61838716:+	GTTAGTTCCATCATCCCCT
	CCDC47_CCDS11643.1.ex9_17:61841403-61841426:+	TCCCTATGAGTGTTAAACC
	CCDC47_CCDS11643.1.ex11_17:61843366-61843389:-	GTTACTGAATCTCCTCAAC
TM9SF2	TM9SF2_CCDS9493.1.ex0_13:100154013-100154036:+	AGAGCGACGAGTGCAAGGT
	TM9SF2_CCDS9493.1.ex4_13:100188866-100188889:+	GGATAATATGCCTGTAACG
	TM9SF2_CCDS9493.1.ex5_13:100190074-100190097:-	GCCACTAATCTTGCTCCCA
	TM9SF2_CCDS9493.1.ex6_13:100191741-100191764:+	CCCATGGACATAAGTAACA
	TM9SF2_CCDS9493.1.ex7_13:100192977-100193000:+	GATCAGATGGGCGTCTAGA
VAPA	VAPA_CCDS11847.2.ex1_18:9931850-9931873:-	AACACACTTTTCTATCCGA
	VAPA_CCDS11847.2.ex1_18:9931890-9931913:+	ACCTCGCCGGTACTGTGTG
	VAPA_CCDS11847.2.ex2_18:9936118-9936141:-	TCATTCGGATCATAGTCAA
	VAPA_CCDS11847.2.ex4_18:9944914-9944937:+	CTCAGGGTATAACTCCACC
	VAPA_CCDS11847.2.ex4_18:9944929-9944952:-	CAGTCGGAGCATTCCCTGG
SCAMP2	SCAMP2_CCDS10271.1.ex3_15:75142850-75142873:+	CACCTAAAGGCCTTATAGA
	SCAMP2_CCDS10271.1.ex3_15:75142995-75143018:+	ACAGAGTCACTGAATGCAC
	SCAMP2_CCDS10271.1.ex4_15:75143724-75143747:-	GATCCCTGCCGACTACCAG

	SCAMP2_CCDS10271.1.ex6_15:75146371-75146394:+	GCTGGGTTGGTTCCACTGA
	SCAMP2_CCDS10271.1.ex6_15:75146416-75146439:+	ACCCAGGGAGTTGGGTGAC
ESYT2	ESYT2_CCDS34791.1.ex14_7:158560333-158560356:+	ACACGCTTACCTTTGGTAC
	ESYT2_CCDS34791.1.ex19_7:158590640-158590663:-	AGTTTCACGAAGGTCGACG
	ESYT2_CCDS34791.1.ex19_7:158590688-158590711:-	ACTATAGAACCAGCCGTGC
	ESYT2_CCDS34791.1.ex21_7:158621941-158621964:+	AGGTAGCCCAGCGCGTACA
	ESYT2_CCDS34791.1.ex21_7:158622029-158622052:-	CGGCGCCTGAGAACCCCGG
ATP6V0A1	ATP6V0A1_CCDS11426.1.ex2_17:40620056-40620079:+	ACATTCCGATTATGGACAC
	ATP6V0A1_CCDS11426.1.ex2_17:40620101-40620124:-	CTAAGTCAATCATGTCCCG
	ATP6V0A1_CCDS11426.1.ex5_17:40630529-40630552:+	TGCTTTGGCGGGTATGCCG
	ATP6V0A1_CCDS11426.1.ex6_17:40632679-40632702:-	TTGTGCACGTAGTCGCCCT
	ATP6V0A1_CCDS11426.1.ex9_17:40642633-40642656:+	CTTACCGAGAGATAAATCC
FLNA	FLNA_CCDS44021.1.ex24_X:153588136-153588159:+	TCGCCACGGTCCTGAACGT
	FLNA_CCDS44021.1.ex38_X:153594434-153594457:+	CCGGCAAACGTGACGTGCA
	FLNA_CCDS44021.1.ex40_X:153594946-153594969:-	TCCGTCTGGTACGTCCCCG
	FLNA_CCDS44021.1.ex43_X:153596046-153596069:-	CCCGTTACCAATGCGCGAG
	FLNA_CCDS44021.1.ex45_X:153599249-153599272:-	GACCGCGAGAGCATCAAAC
CLIC1	CLIC1_CCDS4719.1.ex1_6:31700169-31700192:-	TCCCTCTCTTGCACAGATC
	CLIC1_CCDS4719.1.ex2_6:31701435-31701458:-	GCTTTCAGGTACCCCAAGC
	CLIC1_CCDS4719.1.ex3_6:31701597-31701620:+	GAGTGCCCCTATACCTGGG
	CLIC1_CCDS4719.1.ex3_6:31701673-31701696:+	AGTGCCATACAGCAGGAAT
	CLIC1_CCDS4719.1.ex4_6:31701929-31701952:-	TACCACCGTTGACACCAAA
TMED1	TMED1_CCDS12249.1.ex2_19:10945961-10945984:+	TGACCAACAGCACGCCCTG
	TMED1_CCDS12249.1.ex2_19:10945976-10945999:-	TCACGCTGGAGAGCCCTCA

	TMED1_CCDS12249.1.ex3_19:10946683-10946706:-	AGCCTCGAGACCGAATACC
	TMED1_CCDS12249.1.ex3_19:10946699-10946722:+	AGGCTTGCGTTGGCCGGCG
	TMED1_CCDS12249.1.ex3_19:10946744-10946767:-	CACGTTCCTGTTGCCGGCG
TGOLN2	TGOLN2_CCDS46351.1.ex2_2:85554297-85554320:-	CTAATAAGTCGGGTGCGGA
	TGOLN2_CCDS46351.1.ex2_2:85554469-85554492:-	AGCACTAGTAAGTCGCATC
	TGOLN2_CCDS46351.1.ex2_2:85554593-85554616:+	TCCGCACCCGACTTGTTGG
	TGOLN2_CCDS46351.1.ex2_2:85554684-85554707:+	GCGACTTGGTAGAGCCTCC
	TGOLN2_CCDS46351.1.ex2_2:85554701-85554724:+	CCAGGCCGTTGGCTCAAGC
SCAMP1	SCAMP1.1	TAGAAACAGTGTTACTGCA
	SCAMP1.2	CACTGTTTCTAAATATCTT
	SCAMP1.3	CTAAATATCTTCGGATGCT
	SCAMP1.4	TATCTTCGGATGCTTGGCT
	SCAMP1.5	TGTGTTGATTCTGCAAGAG
STX7	STX7_CCDS5153.1.ex3_6:132791112-132791135:+	TCTCTCATGAATAAGACGG
	STX7_CCDS5153.1.ex4_6:132791711-132791734:-	GAATCTTGTATCCTGGGAA
	STX7_CCDS5153.1.ex5_6:132792657-132792680:-	TTCCAGAAGGTCCAGAGGC
	STX7_CCDS5153.1.ex5_6:132792717-132792740:-	CGTCAAAGGAAAATACAGA
	STX7_CCDS5153.1.ex6_6:132793409-132793432:+	AGCTTGATACCTGTTCACT
SCAMP3	SCAMP3_CCDS1105.1.ex3_1:155227432-155227455:+	GAGCCAGCGTGCTGCCTAA
	SCAMP3_CCDS1105.1.ex5_1:155230159-155230182:-	GCAGAGGAGTTGGACCGAA
	SCAMP3_CCDS1105.1.ex6_1:155230326-155230349:-	TATGGCTCATACAGCACTC
	SCAMP3_CCDS1105.1.ex6_1:155230343-155230366:-	CCACAGAACCTAAGAACTA
	SCAMP3_CCDS1105.1.ex6_1:155230361-155230384:+	TGGGGCTGAGCTTTCTCGA
ESYT1	ESYT1_CCDS53801.1.ex0_12:56522378-56522401:-	TTCTTTCTCGTCGCGGACC
	ESYT1_CCDS53801.1.ex2_12:56524640-56524663:-	GGTTAGATCCCCTAACAGC
	ESYT1_CCDS53801.1.ex5_12:56525340-56525363:-	GTTTTCCCTTACCGGGCGT
	ESYT1_CCDS53801.1.ex6_12:56525577-56525600:-	GTCCTTGATACCTAAGTCC

	ESYT1_CCDS53801.1.ex7_12:56526034-56526057:-	TAATCGGTTGGGCAACACG
FAM134C	FAM134C_CCDS11432.1.ex3_17:40737220-40737243:+	ATATGCTCGATCCCACAGT
	FAM134C_CCDS11432.1.ex4_17:40738059-40738082:+	TCATACTTACGCATCAAGT
	FAM134C_CCDS11432.1.ex5_17:40738838-40738861:+	AACCCAGACTTCAGCTACA
	FAM134C_CCDS11432.1.ex5_17:40738870-40738893:+	CGGGCACGCTGAGCAACCG
	FAM134C_CCDS11432.1.ex6_17:40739850-40739873:-	CGACGCATTAGACAATGAG
CLIC4	CLIC4_CCDS256.1.ex1_1:25124273-25124296:-	AAGAATCATGAAGAGCCTC
	CLIC4_CCDS256.1.ex2_1:25140613-25140636:-	ATAAATGGTGGGTGGGTCC
	CLIC4_CCDS256.1.ex2_1:25140701-25140724:-	CCTTGATACTCACTTGGGA
	CLIC4_CCDS256.1.ex3_1:25153522-25153545:+	ACCCAGAATCAAATACTGC
	CLIC4_CCDS256.1.ex3_1:25153591-25153614:-	TTTTACCTTCATTAGCCTC
AIG1	AIG1_CCDS5198.1.ex0_6:143382150-143382173:+	CTCACACCAGACCTACGGA
	AIG1_CCDS5198.1.ex0_6:143382184-143382207:-	TACCAGATCAATGAACGTC
	AIG1_CCDS5198.1.ex1_6:143457958-143457981:+	GTCCCCCTACAGGTTATCC
	AIG1_CCDS5198.1.ex1_6:143458071-143458094:+	GCTCATCTCTCTCCGGGAC
	AIG1_CCDS5198.1.ex1_6:143458084-143458107:-	CACAGCTAACATCCAGTCC
MAL2	MAL2.1	AGCGGCAGCGGCAGCATGT
	MAL2.2	GCAGCGGCAGCATGTCGGC
	MAL2.3	GCGGCAGCATGTCGGCCGG
	MAL2.4	GGCGGGACTGACGCTCCGC
	MAL2.5	GCGGCGGGAAGGACACGGC
SLC30A1	SLC30A1_CCDS1499.1.ex0_1:211749513-211749536:-	CTGGACAACTTAACATGCG
	SLC30A1_CCDS1499.1.ex1_1:211751336-211751359:+	GGGTCCAATTTCAGCCCGT
	SLC30A1_CCDS1499.1.ex1_1:211751445-211751468:+	GGCGGGTGCTCTTAACGCG
	SLC30A1_CCDS1499.1.ex1_1:211751493-211751516:-	GCCACTCGCACGGGGGTCA
	SLC30A1_CCDS1499.1.ex1_1:211751719-211751742:-	TTCGGCTGGATCCGAGCCG
SLC7A2	SLC7A2_CCDS34852.1.ex0_8:17401158-17401181:+	CGTGACTGTCGGAGAGCTG

	SLC7A2_CCDS34852.1.ex1_8:17401980-17402003:-	TCATCAAAGGTGCCACTCC
	SLC7A2_CCDS34852.1.ex1_8:17402093-17402116:+	GCCTTATATTACTTCTAGC
	SLC7A2_CCDS34852.1.ex3_8:17407861-17407884:-	ACGTTCCCGTAAAGCCATA
	SLC7A2_CCDS34852.1.ex4_8:17409304-17409327:-	AAGACGTCACAATTCCAAT

Table S 5 Pool 5: list of 94 genes and 500 sgRNA sequences.

There were 25 non-targeting sgRNAs, 5 SCARB1-targeting sgRNAs, 5 CD81-targeting sgRNAs and 465 sgRNAs targeting 94 genes. Genes are listed in an order of proteomic abundance.

APPENDIX B: MAGECK-MLE BATCH MATRIX GENE RANK

Highlighted genes are negatively selected genes with β-scores higher than 2-fold standard deviation, as presented in Results 3.2.3.

Pool 1:

Rank	Gene	β-Score
1	ITGAV	-0.66186
2	RPN1	-0.57372
3	SCARB1	-0.481264
4	ATP2B1	-0.277989
5	TMEM30A	-0.253921
6	ICOSLG	-0.143745
7	CDH2	-0.142979
8	TM9SF3	-0.137126
9	CXADR	-0.1233887
10	ASGR1	-0.1218487
11	SLC3A2	-0.119478
12	SLC29A1	-0.112032
13	ASGR2	-0.103736
14	NCSTN	-0.101632
15	SLC1A5	-0.097144
16	F11R	-0.09519
17	CANX	-0.0916702
18	MAN2A1	-0.087298
19	BST2	-0.0835588
20	TMEM205	-0.067981
21	DPEP1	-0.066733
22	CPD	-0.050207
23	COMT	-0.0486833
24	OCLN	-0.0423298
25	ABCB1	-0.036203
26	ATP1B1	-0.03601
27	SLC16A3	-0.0357003
28	LAMP1	-0.033371
29	CD99	-0.032958
30	BSG	-0.0322663
31	EXT2	-0.030475
32	DAG1	-0.0192386
33	PIGS	-0.016877

34	SLC39A14	-0.015054
35	MTDH	-0.012271
36	VASN	-0.007738
37	SLC2A2	-0.0043407
38	CD55	0.0040117
39	LMAN2	0.005076
40	HEPACAM	0.005918
41	ICAM1	0.008832
42	CD59	0.0118896
43	CADM1	0.01440784
44	NT5E	0.0228952
45	SDC1	0.023682
46	DSG1	0.0245711
47	EGFR	0.02670623
48	MPZL1	0.029486
49	PLXNB2	0.0321427
50	ENPP1	0.03302
51	ALCAM	0.036046
52	ANPEP	0.03682
53	ROBO1	0.041653
54	ATP2B4	0.042373
55	TMEM2	0.045729
56	ECE1	0.0459005
57	CD68	0.04667397
58	SLC10A1	0.050975
59	CEACAM1	0.0523145
60	HSD17B2	0.053151
61	LSR	0.0584103
62	HAVCR1	0.062425
63	PVR	0.0638017
64	ITGA1	0.067402
65	BCAM	0.067556
66	FN1	0.07016272
67	ITGA6	0.072037
68	LRP1	0.074176
69	DPP4	0.0758247
70	M6PR	0.078589
71	STIM1	0.080306
72	TFRC	0.080795
73	DSG2	0.081136
74	LAMP2	0.083881
75	ANO6	0.087823

76	SLC11A2	0.089025
77	ABCC2	0.09876714
78	IGF2R	0.108805
79	CKAP4	0.118975
80	CSGALNACT1	0.122493
81	SDC4	0.1238954
82	ATP1B3	0.130361
83	ADAM10	0.130747
84	REEP5	0.155191
85	EPHA2	0.163129
86	NPC1	0.168672
87	CD276	0.176065
88	ASPH	0.183517

Pool 2:

Rank	Gene	β-Score
1	SCARB1	-0.7615
2	FCGR2B	-0.72372
3	ITGB5	-0.7042
4	SLC35A2	-0.5316172
5	MGAT1	-0.48696
6	SLC46A1	-0.298575
7	APOA1	-0.29049
8	DSC1	-0.252113
9	B4GALT7	-0.242711
10	APMAP	-0.2363395
11	PTPRF	-0.230036
12	TFR2	-0.211307
13	CADM1	-0.203211
14	PLXNB1	-0.193684
15	SLC4A1	-0.17176
16	LDLR	-0.171206
17	PRNP	-0.170584
18	ITGB2	-0.170272
19	EPHA2	-0.167454
20	SLC39A6	-0.145916
21	ICAM2	-0.1447488
22	GGT5	-0.135086
23	CBL	-0.132569
24	CD14	-0.12671
25	IGSF8	-0.120683
26	ABCG1	-0.11659
27	NPTN	-0.116211
28	FGFR4	-0.11279
29	NEO1	-0.1089881
30	PTK7	-0.104455
31	ERBB2	-0.099627
32	XPNPEP2	-0.08503
33	CDH1	-0.07588
34	JAG1	-0.06508
35	EPHB4	-0.060146
36	PLXNA1	-0.0595487
37	CAPN5	-0.05833
38	PVRL2	-0.0514375
39	INS	-0.04593

40	CD63	-0.045037
41	GRB2	-0.03864
42	THBS1	-0.02113
43	INSR	-0.00698
44	FCGRT	-0.00187
45	SLC2A1	0.00144
46	CEACAM6	0.00834
47	TMPRSS6	0.009
48	IL1R1	0.02886
49	MTDH	0.0332867
50	ITFG1	0.061773
51	CD40	0.06328
52	EFNB1	0.068439
53	CLEC4M	0.07006
54	FAS	0.074
55	MPZL2	0.077756
56	PTPRK	0.0867005
57	DCBLD1	0.0993629
58	GNB1	0.106644
59	LEPR	0.10912
60	ITGA5	0.117853
61	IGSF1	0.123144
62	PDZK1	0.126871
63	FCER1G	0.128475
64	VNN1	0.143165
65	ENPEP	0.144234
66	APLP2	0.145789
67	NEU1	0.15876
68	APOE	0.162628
69	B2M	0.18479
70	ITGAM	0.185392
71	ERBB3	0.223191
72	IFNGR1	0.22393
73	PIGS	0.22596
74	EPCAM	0.237106
75	GGT1	0.2389
76	MME	0.243653
77	CD163	0.249309
78	CD46	0.253846
79	TSPAN14	0.255229
80	IL6ST	0.268926
81	DSC2	0.27323

82	PPARG	0.285969
83	SLC7A5	0.2866
84	NAT1	0.299564
85	CD58	0.310945
86	LPHN2	0.3118
87	PGAP1	0.32322
88	MET	0.32691
89	SEMA4G	0.33349
90	PTPRM	0.346333
91	ITGA2	0.363538
92	GPC3	0.376069
93	ITGB1	0.397806
94	MSR1	0.67108
95	TGFBR1	1.22633

Pool 3:

Rank	Gene	β-Score
1	SCARB1	-1.09528
2	GJB1	-0.427109
3	MAGT1	-0.39297
4	SLC1A4	-0.38068
5	MYH9	-0.363
6	STT3B	-0.3516229
7	MIA3	-0.341647
8	SUN2	-0.337356
9	A2M	-0.28928
10	CERS2	-0.24817
11	C3	-0.245516
12	RPN2	-0.2383
13	CLU	-0.238197
14	PGRMC2	-0.236577
15	SSR1	-0.230115
16	NCLN	-0.22163
17	PTGFRN	-0.198749
18	CALR	-0.198529
19	CD74	-0.194535
20	ABHD12	-0.18793
21	CLEC4G	-0.18178
22	NUP210	-0.178259
23	STEAP3	-0.167755
24	FGB	-0.15178
25	BRI3BP	-0.132607
26	LGALS3BP	-0.130213
27	GDF15	-0.1300554
28	ATP6AP1	-0.1153
29	ENG	-0.11415
30	GOLIM4	-0.109545
31	TXNDC15	-0.106993
32	SIGMAR1	-0.10527
33	SLC44A1	-0.0785
34	GGCX	-0.07104
35	SLC16A1	-0.069974
36	APOB	-0.062444
37	SLC13A5	-0.05485
38	CLPTM1	-0.05077
39	GPI	-0.0495

40	PTTG1IP	-0.0469
41	TMX2	-0.0463
42	FNDC3B	-0.02008
43	HPN	-0.00607
44	RAB5C	-0.00025
45	TM9SF1	0.00172
46	HSD11B1	0.00199
47	CD151	0.00943
48	F2	0.011215
49	IFITM3	0.0193
50	TMED4	0.020303
51	HM13	0.02081
52	HNRNPK	0.0238
53	SLC38A2	0.0275
54	CD44	0.03307
55	FNDC3A	0.03368
56	ARSE	0.04537
57	ACTN1	0.04772
58	CTSD	0.053743
59	ORM1	0.063966
60	SEC22B	0.0915
61	SLC25A4	0.1018291
62	CALU	0.109613
63	TSPAN8	0.116734
64	ABCB11	0.124633
65	TMED9	0.130245
66	DHRS7	0.142565
67	SCARB2	0.143241
68	SERPINA1	0.156417
69	TMX4	0.181032
70	CASC4	0.196134
71	RTN4	0.199253
72	KTN1	0.203467
73	ITIH2	0.204424
74	SDC2	0.21273
75	STT3A	0.2199
76	GLG1	0.233114
77	CCDC90B	0.241306
78	DHCR7	0.246769
79	ITFG3	0.26234
80	EPHA1	0.265
81	TMEM97	0.266819

82	TF	0.29232
83	PON1	0.30544
84	ATP1A1	0.3151
85	SLC31A1	0.31973
86	YBX1	0.3543
87	CLDND1	0.38749
88	IL13RA1	0.405423
89	VDAC1	0.44739
90	IMPAD1	0.49431
91	MMGT1	0.58097
92	SLCO1B1	0.58453
93	LMAN2L	0.62585
94	VTN	0.77351

Pool 4:

Rank	Gene	β-Score
1	SCARB1	-0.48
2	EMC1	-0.37
3	APOH	-0.28
4	B4GALT1	-0.21
5	SLC2A3	-0.20
6	COL6A2	-0.20
7	FXYD1	-0.19
8	ATP13A3	-0.19
9	CD9	-0.19
10	NPC1	-0.18
11	LRBA	-0.17
12	GPRC5C	-0.17
13	ABCA1	-0.16
14	SERPING1	-0.15
15	SIRPA	-0.14
16	ABCA6	-0.13
17	SLC47A1	-0.12
18	SORL1	-0.11
19	ITM2B	-0.11
20	GYPC	-0.11
21	PIGT	-0.10
22	SLC30A10	-0.10
23	ATP6AP2	-0.07
24	CD47	-0.06
25	FXYD2	-0.06
26	ATG9A	-0.04
27	CDHR5	-0.04
28	SLC10A1	-0.04
29	TGFBR2	-0.04
30	FLVCR1	-0.04
31	LNPEP	-0.04
32	SLC5A6	-0.03
33	CRTAP	-0.03
34	NAALADL2	-0.03
35	SLC19A3	-0.03
36	CLRN3	-0.03
37	ACP2	-0.03
38	TMEM123	-0.02
39	TGFBR3	-0.02

40	CNNM4	-0.01
41	ADCY9	-0.01
42	COL6A1	-0.01
43	CYBB	0.00
44	KLB	0.00
45	TMBIM1	0.00
46	DSC3	0.00
47	TSPAN9	0.00
48	ABCG8	0.01
49	TMX3	0.01
50	CD82	0.02
51	SLC2A9	0.02
52	SLCO2B1	0.03
53	AMBP	0.03
54	LILRB5	0.03
55	CNNM3	0.04
56	TMED7	0.04
57	GPC1	0.07
58	SLC3A1	0.07
59	GPC6	0.08
60	DPP7	0.08
61	AHSG	0.08
62	CNNM2	0.08
63	IL1RAP	0.08
64	LAMB2	0.09
65	CP	0.09
66	TSPAN6	0.09
67	FCGR2B	0.09
68	CLDN3	0.09
69	LMF2	0.09
70	FAM171A1	0.10
71	GALNT1	0.10
72	CLPTM1L	0.10
73	AGRN	0.11
74	SLC22A9	0.12
75	TIMD4	0.12
76	SORT1	0.13
77	ABCC6	0.13
78	CREG1	0.13
79	PLD3	0.14
80	ALPL	0.15
81	TOR1AIP2	0.15

82	CA14	0.16
83	CANT1	0.16
84	KIAA0319L	0.17
85	MCAM	0.18
86	STEAP4	0.19
87	CRB3	0.20
88	ABCB4	0.20
89	SLC22A1	0.20
90	QSOX2	0.22
91	SLC22A7	0.23
92	REEP6	0.23
93	CD99L2	0.24
94	CD36	0.27
95	SLC15A1	0.28

Pool 5:

Rank	Gene	β-Score
1	SCARB1	-0.91
2	MARCO	-0.34
3	ESYT1	-0.31
4	TM9SF2	-0.28
5	VSIG4	-0.25
6	IFNAR2	-0.22
7	SCAMP2	-0.22
8	IL4R	-0.21
9	CDH5	-0.20
10	IFNGR2	-0.20
11	CRIM1	-0.20
12	FGFRL1	-0.19
13	LY6E	-0.18
14	ATP6V0A1	-0.18
15	MAL2	-0.18
16	SLC22A	-0.17
17	EFNA1	-0.17
18	ESYT2	-0.14
19	TP53I13	-0.13
20	NAALAD	-0.13
21	TNFRSF1A	-0.13
22	PTPRG	-0.13
23	TNFRSF	-0.13
24	IL17RB	-0.11
25	ACVR1B	-0.11
26	CD164	-0.09
27	NAGPA	-0.09
28	SLC30A1	-0.08
29	RAMP1	-0.07
30	TGOLN2	-0.07
31	LRP5	-0.06
32	HFE2	-0.05
33	CLIC1	-0.04
34	CD52	-0.04
35	TMED1	-0.04
36	STOM	-0.04
37	NOTCH2	-0.04
38	VAPA	-0.03
39	ACVR1	-0.02

40	CLSTN3	-0.01
41	AIG1	-0.01
42	SCAMP3	0.00
43	FAM134C	0.02
44	APP	0.05
45	LRP10	0.06
46	FLNA	0.07
47	F3	0.08
48	MMP15	0.08
49	MPEG1	0.08
50	CSF1R	0.08
51	SLC7A1	0.09
52	SLC6A1	0.09
53	ATRAID	0.09
54	CDHR2	0.10
55	VCAM1	0.10
56	CLIC4	0.10
57	SEMA4B	0.11
58	GFRA1	0.12
59	TMPRSS2	0.13
60	STAB1	0.13
61	MFAP3L	0.14
62	SLC7A2	0.14
63	BTN2A1	0.14
64	IFNAR1	0.14
65	NRP1	0.14
66	FLT1	0.14
67	LTBR	0.14
68	SLC6A9	0.14
69	FGFR2	0.16
70	TMEM21	0.16
71	C11orf	0.16
72	TYROBP	0.16
73	SDC3	0.16
74	CD55	0.17
75	SCAMP1	0.17
76	STX7	0.17
77	OSMR	0.17
78	CCDC47	0.18
79	PDGFRA	0.18
80	CXCL16	0.19
81	TCTN3	0.19

82	ADAM9	0.20
83	IL6R	0.21
84	TSPAN3	0.22
85	TTYH3	0.22
86	CADM4	0.23
87	ICAM3	0.26
88	LRP6	0.31
89	PTPRC	0.31
90	TYRO3	0.34
91	RELL1	0.39
92	IGF1R	0.76

Pool 2 with collagen:

Rank	Gene	β-Score
1	PIGS	-2.07974
2	PVRL2	-0.71082
3	SCARB1	-0.699781
4	ITGB2	-0.655848
5	SLC35A2	-0.60115
6	MGAT1	-0.48705
7	CD63	-0.40977
8	EPHB4	-0.38256
9	ITGB1	-0.3464075
10	CD163	-0.294087
11	INS	-0.274131
12	MPZL2	-0.201809
13	EPHA2	-0.199938
14	CEACAM6	-0.194229
15	MET	-0.1932087
16	PLXNA1	-0.1704
17	CDH1	-0.1697574
18	SLC46A3	-0.155247
19	FCGR2B	-0.135776
20	EFNB1	-0.134283
21	ICAM2	-0.124475
22	APOA1	-0.116496
23	CADM1	-0.112514
24	GNB1	-0.10926
25	JAG1	-0.102786
26	NEU1	-0.098209
27	PTPRM	-0.097561
28	ABCG1	-0.091521
29	FAS	-0.089955
30	TSPAN14	-0.08586
31	CLEC4M	-0.0743403
32	B2M	-0.072551
33	SLC39A6	-0.071205
34	GGT1	-0.069146
35	PGAP1	-0.068201
36	LDLR	-0.06107
37	DSC1	-0.053045
38	SLC46A2	-0.049731
39	MME	-0.0476964

40	ITGB5	-0.04433
41	DCBLD1	-0.043253
42	ENPEP	-0.039637
43	SLC7A5	-0.03633
44	FCER1G	-0.0359836
45	CD58	-0.026035
46	SLC4A1	-0.0200696
47	PPARG	-0.0188154
48	XPNPEP2	-0.0147761
49	PDZK1	-0.01272
50	PRNP	-0.011567
51	INSR	-0.003732
52	IL1R1	-0.00335
53	NAT1	-0.002091
54	NPTN	0.0020416
55	EPCAM	0.0023
56	SLC46A1	0.003101
57	GPC3	0.0060145
58	CAPN5	0.009158
59	TGFBR1	0.014972
60	LPHN2	0.0185853
61	ITGA5	0.023193
62	LEPR	0.024988
63	ITFG1	0.0261124
64	ERBB2	0.02675607
65	SLC46A4	0.02783
66	GRB2	0.035862
67	NEO1	0.03611
68	IGSF8	0.03619
69	THBS1	0.03646
70	VNN1	0.038715
71	PTPRF	0.0388498
72	CD40	0.040537
73	MSR1	0.041596
74	TMPRSS6	0.0441946
75	PTK7	0.0458611
76	GGT5	0.0469903
77	APOE	0.055466
78	IL6ST	0.05687582
79	ITGAM	0.05739
80	APMAP	0.058428
81	CD46	0.059988

82	FGFR4	0.0610375
83	DSC2	0.07013521
84	CBL	0.077453
85	CD14	0.080196
86	FCGRT	0.083682
87	SLC2A1	0.0856832
88	ERBB3	0.094051
89	IGSF1	0.097354
90	MTDH	0.100034
91	SLC46A5	0.103796
92	B4GALT7	0.111812
93	SEMA4G	0.12039
94	ITGA2	0.131031
95	PTPRK	0.131076
96	TFR2	0.131184
97	APLP2	0.147254
98	PLXNB1	0.186746
99	IFNGR1	0.188904