Transcriptional Control of Maternal-Fetal Immune Tolerance

Transcriptional control of maternal-fetal immune tolerance

A dissertation presented by

Leonardo Manuel Ramos Ferreira

to

The Department of Molecular and Cellular Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biochemistry

Harvard University

Cambridge, Massachusetts

March 2016

Transcriptional control of maternal-fetal immune tolerance

ABSTRACT

Human leukocyte antigens (HLA) are important determinants of self-nonself immune recognition. HLA-G, uniquely expressed in the placenta, is believed to be key to fetus-induced immune tolerance during pregnancy. The tissue-specific expression of HLA-G, however, remains poorly understood. Using a Massively Parallel Reporter Assay (MPRA), we discovered a 121 bp sequence 12 kb upstream of *HLA-G* with enhancer activity, *Enhancer L*. Strikingly, deletion of *Enhancer L* using a CRISPR/Cas9 dual guide approach resulted in complete ablation of HLA-G expression in a trophoblast cell line. This finding was confirmed in primary extravillous trophoblasts isolated from human placenta. RNA-seq analysis demonstrated that *Enhancer L* regulates HLA-G expression specifically. Moreover, DNase-seq and Chromatin Conformation Capture (3C) defined *Enhancer L* as a cell type-specific enhancer that loops into the *HLA-G* promoter. GATA2, GATA3, and CEBP□, factors essential for placentation, associate with *Enhancer L* and regulate *HLA-G* expression levels. These results establish long-range chromatin looping as a novel mechanism controlling trophoblast-specific HLA-G expression at the maternal-fetal interface.

ABBREVIATIONS

APC, Antigen Presenting Cell

BFU-E, Burst Forming Unit-Erythrocyte

B2M, Beta-2-Microglubulin

bp, base pair

BSA, Bovine Serum Albumin

3C, Chromatin Conformation Capture

4C, Circular Chromatin Conformation Capture

5C, Carbon Copy Chromatin Conformation Capture

CAT, Chloramphenicol Acetyltransferase

CCR5, C-C Chemokine Receptor 5

CD, Cluster of Differentiation

CEBP, CCAAT Enhancer-binding Protein

CEBPB, CCAAT Enhancer-binding Protein, Beta gene

CEBP□, CCAAT Enhancer-binding Protein, Beta protein

CFC, Colony Forming Cell

CFU-E, Colony Forming Unit-Erythroid

ChIP, Chromatin Immunoprecipitation

CIITA, Class II, major histocompatibility complex, Transactivator

CMV, Cytomegalovirus

CNV, Copy Number Variant

CRISPR, Clustered Regularly Interspaced Palindromic Repeats

CTCF, CCCTC-binding Factor

DC, Dendritic Cell

DHS, DNase Hypersensitive Site

DMEM, Dulbecco's Modified Eagle's Medium

DN, Double Negative

DNA, Deoxyribonucleic Acid

DSB, Double Strand Break

DSC, Decidual Stromal Cell

EGF, Epidermal Growth Factor

EL, *Enhancer L*

ERV, Endogenous Retrovirus

EVT, Extravillous Trophoblast

FACS, Fluorescence Activated Cell Sorting

FBS, Fetal Bovine Serum

FDR, False Discovery Rate

FPKM, Fragments Per Kilobase of exon per Million fragments mapped

GAPDH, Glyceraldehyde 3-phosphate Dehydrogenase

GEMM, Granulocyte/Erythrocyte/Macrophage/Megakaryocyte

GFP, Green Fluorescent Protein

GM, Granulocyte/Megakaryocyte

GPCR, G-protein-coupled Receptor

gRNA, guide RNA

GSEA, Gene Set Enrichment Analysis

GWAS, Genome-wide Association Studies

HBB, Hemoglobin, Beta

HEK, Human Embryonic Kidney

HLA, Human Leukocyte Antigen

HR, Homologous Recombination

HSC, Hematopoietic Stem Cell

HSE, Heat Shock Element

HSPC, Hematopoietic Stem and Progenitor Cell

HSV, Herpes Simplex Virus

IDO, Indoleamine 2,3 dioxygenase

IL, Interleukin

ILC, Innate Lymphoid Cell

ILT, Ig-like Transcript

InDel, Insertion/Deletion

IRB, Institutional Review Board

IRES, Internal Ribosome Entry Site

ITGA5, Integrin, Alpha 5

kb, kilobase

KIR, Killer-cell Immunoglobulin-like Receptor

KO, Knock-out

LCR, Locus Control Region

LINE, Long Interspersed Nuclear Element

LTR, Long Terminal Repeat

MAGE, Melanoma Associated Antigen Gene

MAR, Matrix Attachment Region

Mb, Megabase

MFI, Mean Fluorescence Intensity

MHC, Major Histocompatibility Complex

mPB, mobilized Peripheral Blood

MPRA, Massively Parallel Reporter Assay

mRNA, messenger RNA

NCS, Newborn Calf Serum

NEB, New England Biolabs

NHEJ, Non-homologous End Joining

NK, Natural Killer

NLRC5, NOD-like Receptor family CARD domain-containing 5

nTreg, natural Regulatory T cell

PBS, Phosphate-buffered Saline

PCR, Polymerase Chain Reaction

PD-1, Programmed Death 1

PD-L1, Programmed Death Ligand 1

PE, Phycoerythrin

PMSF, Phenylmethane Sulfonyl Fluoride

Pol II, RNA Polymerase II

PRE, Progesterone Responsive Element

qRT-PCR, quantitative Real Time Polymerase Chain Reaction

RLU, Relative Luciferase Units

RNA, Ribonucleic Acid

RPMI, Roswell Park Memorial Institute

SAR, Scaffold Attachment Region

SDS, Sodium Dodecyl Sulfate

SEM, Standard Error of the Mean

SV40, Simian Virus 40

TALEN, Transcription Activator-like Effector Nuclease

TCR, T Cell Receptor

Th1, T helper 1

Th2, T helper 2

Treg, Regulatory T cell

UTR, Untranslated Region

VSV-G, Vesicular Stomatitis Virus G

VT, Villous Trophoblast

WT, Wild-type

ZFN, Zinc Finger Nuclease

ACKNOWLEDGMENTS

"If I have seen further, it is by standing on the shoulders of giants."

- Sir Isaac Newton

First and foremost, I would like to use this opportunity to express my utmost gratitude to my Ph.D. advisor, Jack L. Strominger. Jack truly leads by example, stopping by my bench every day, and discussing data and follow-up experiments with me on a regular basis. We had dinner several times where Jack brought my lab meeting slides to discuss them in more depth. His enthusiasm for research is relentless and has been a constant source of inspiration during the past four years. Jack's scientific accomplishments speak for themselves, and so do his trainees' accomplishments, many of which are now successful professors and chairs of departments all around the world. I feel truly honored to have been Jack's graduate student.

I was incredibly fortunate to have had Chad A. Cowan as a Ph.D. co-advisor. While still a rotation student, Chad made himself available to meet with me every week to discuss my progress, ideas for experiments and projects, and larger scientific goals. After being part of the Strominger and Cowan labs, he helped me tremendously to establish collaborations with people from all over the Greater Boston scientific community to move my projects forward.

I must also say that I could not have asked for more supportive advisors when it came to nurture my early scientific career. I was always encouraged since very early to present my work at the numerous retreats happening at Harvard, and got their support to travel to and present my work at national and international conferences.

The completion of this dissertation would also not have been possible without a third mentor, Torsten B. Meissner, a postdoctoral fellow at the Strominger and Cowan labs. He helped me design and perform experiments for several years, and has been an invaluable source of advice since day 1 on experimental design, scientific writing, and on teaching, both in the lab and in the classroom (we taught together the Harvard College course SCRB165 for two years as teaching fellows). His experimental rigor, passion for science, and holistic view of what it means to be a scientist, have helped me grow as a person and as a scientist during these formative years. I will forever be indebted to him.

The environment where I spent most of my hours was extremely supportive and stimulating. Besides my mentors, members of both the Strominger and the Cowan labs always made themselves available to help and shared reagents routinely, in particular Angela Crespo (Strominger lab), Max Friesen, and Curtis Warren (Cowan lab). I also keep many fond memories from lab retreats and other activities outside of the lab.

During the past five years, various faculty other than my direct advisors have had an important impact on my graduate student career. First, I would like to deeply thank my neutral advisor Andrew Murray for the many long and insightful chats we had during my first year at Harvard. In addition, I must thank my dissertation advisory committee members, Professors Rich Losick (Chair), John Rinn, and Hidde Ploegh, for their insightful and incisive questions about my ongoing work during our biannual meetings, as well as for their precious help and advice on my what my next steps in my career should be.

Of note, I would not be here today without the support of three professors at my college, University of Coimbra, Portugal: Professors Ana Urbano, Rui de Carvalho, and Carlos Faro. They believed in my potential and allowed me to work in their labs side by side with graduate students. It is thanks to their investment, mentorship during my first attempts at doing research, and letters of recommendation, that I became a PhD student at the Department of Molecular and Cellular Biology, Harvard University.

Such a constellation of incredible mentors made me realize how important it is to train the incoming generations of scientists. My first formal teaching experience at Harvard was as a teaching fellow for Life Sciences 1a, a large freshman biochemistry course taught by Professors Rich Losick, Dan Kahne, and Rob Lue. Being in a room facing thirteen students avid to learn biochemistry on a Thursday

night section in the Science Center building is an experience I will never forget. I think I saw myself in them, back when I was a freshman studying biochemistry. After that, I was a teaching fellow for an advanced stem cell-based lab course, SCRB165, taught by Professor Chad Cowan. Moreover, I also had the unique opportunity to teach molecular biology to high school and early college students in Bolivia, as part of the first edition of Clubes de Ciencia Bolivia, an initiative directed by my classmate and good friend Mohammed Mostajo-Radji. I would like to thank all the students I had the pleasure to interact with during these instances for reminding me of the importance of never letting the extremely curious and persistent child inside oneself die.

Yet, the students I must thank the most are the undergraduates that worked under my direct supervision in lab. I was fortunate to have as my first undergraduate student Breanna Johnson. Enthusiastic and a fast learner, Breanna worked with me for one year, in which she rapidly became an independent worker at the bench and completed her senior thesis. I had two more undergraduate students: Hannah Gomes and Andrew Mazzanti, two extremely smart and dedicated college juniors who I predict will have distinguished careers in immunology. In all three cases, the Latin proverb *docendo discitur* ("one learns by teaching") describes my experience the best, as I am not certain whether they learned more from me than I learned from them. Thanks to these three students, I have become a better scientist. Of note, both

Hannah and Andrew will be doing research in the Strominger lab this summer. That gives me a strange kind of joy I had never felt before: the feeling that, even though I am leaving the lab, my ways of designing and doing experiments will (to an extent) stay in the lab.

Graduate school was not only the time to find myself as a scientist, but also a premier opportunity to find new passions. Following the suggestions of several of my friends, and good experiences at social classes, I joined the Harvard Ballroom Dance Team. I can't imagine having learned more or having had more fun in any other team at Harvard. It definitely transformed my graduate school experience. In particular, I want to thank my partners Pearly Kim (Standard and Latin), Ma Lang (Rhythm), and Emily Sartin (Latin) for their patience, dedication and exceptional memory for routines.

Last, but not least, there are three very special people that I must thank: Kristian Herrera, Patrick Hsu and Mohammed Mostajo-Radji. We've been through more experiences together than these pages can hold. All I will say here is that I owe them a tremendous amount for making it through graduate school. They are part of the reason why Harvard will always feel like home to me. I am excited to see what the future holds for the four of us, and sincerely hope that we can keep challenging one another scientifically and personally in many years to come.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: Introduction

1.1. The pregnancy paradox

Every one of us is here today thanks to successful pregnancy. However, pregnancy has flummoxed immunologists ever since George Snell and Peter Gorer laid out the laws of transplantation (Gorer, 1937; Snell, 1948; Starzl and Zinkernagel, 2001). From an immunological perspective, a developing fetus can be viewed as a semi-allogeneic graft expressing paternally-derived antigens, yet it is nurtured by the pregnant mother for many months without suffering rejection by the maternal immune system (Medawar, 1953). This state of unresponsiveness towards cells expressing foreign antigens – immune tolerance – was initially thought to be systemic (Medawar, 1953). However, later work revealed that pregnant mothers display a normal capacity to fight infections (Kourtis et al., 2014) and to mount effective T cell responses (Lissauer et al., 2012). Instead, fetal immune tolerance is now known to be a local phenomenon established at the placenta, a transient organ consisting of fetal trophoblasts and the maternal decidua, which develops from the uterine mucosa (Arck and Hecher, 2013). During implantation, extravillous trophoblasts (EVT) arise from the tips of anchoring villi and invade the decidua, defining the boundary between mother and fetus: the maternal-fetal interface (Moffett and Loke, 2006) (Figure 1.1.1).

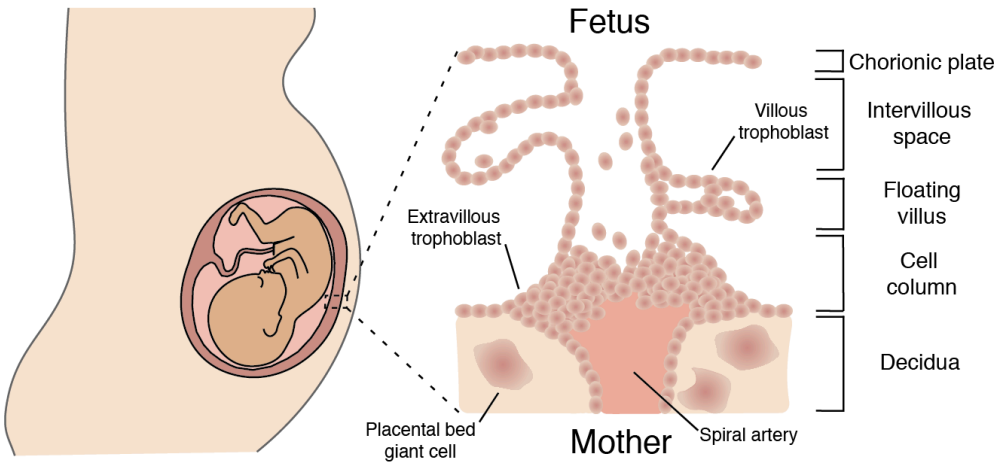


Figure 1.1.1. Anatomy of the human maternal-fetal interface.

The human placenta has an outer layer composed of extravillous trophoblasts (EVT) and floating villi containing villous trophoblasts (VT). EVT form cell columns and invade the decidua, mediating placental attachment of the fetus. Importantly, EVT progressively replace endothelial cells on the walls of uterine spiral arteries, increasing their caliber. This process ensures proper blood flow to the intervillous space to nourish the fetus. (Moffett and Loke, 2006; Moffett-King, 2002). Moreover, some invading EVT fuse and give rise to placental bed giant cells, multinucleated cells which produce hormones essential for a successful pregnancy (al-Lamki et al., 1999).

1.2. Reproductive immunology

Immunity at the maternal-fetal interface has been studied since the 1950s. Yet, the mechanisms behind immune privilege in the pregnant uterus remain somewhat mysterious. Limited access to human placental material has hampered progress in the field (Tilburgs et al., 2015a). Moreover, pregnancy in mouse is significantly different from its human counterpart. Gestation time is significantly shorter in mice (three weeks vs.

nine months), and trophoblast invasion is much less pronounced. In humans, EVT extensively invade maternal tissue, replacing the endothelial cells that form the lining of spiral arteries all the way to the myometrium. In contrast, murine trophoblasts do not deeply invade the decidual arterioles. These differences prevent some human pregnancy-related complications from occurring in the mouse, such as pre-eclampsia (gestational high blood pressure) and fetal growth restriction (Arck and Hecher, 2013; Erlebacher, 2013). Nevertheless, several aspects of reproductive immunology are shared between mouse and human, and the former has been an invaluable model in the field of reproductive immunology.

Characterization of the maternal-fetal interface has substantially progressed during the past decades. Far from being devoid of immune cells, the decidua harbors multiple populations of maternal immune cells, all of which extensively interact with fetal-derived trophoblasts. In fact, up to 40% of decidual cells are leukocytes (Vince et al., 1990). Interestingly, it has become increasingly clear from both human and mouse studies that these decidual immune cells can significantly differ from their peripheral blood counterparts in phenotype and function (Hunt and Petroff, 2013). In the next paragraphs, the different immune cell types found in the placenta, and their relative contributions to maternal-fetal immunity will be reviewed.

1.2.1. Decidual B cells and antibodies

The beginnings of reproductive immunology can be traced back to the 19th century, with the identification of the ABO blood groups by Landsteiner in 1900. After being awarded the Nobel Prize in 1930, Landsteiner's work resulted in yet another major discovery in the 1940s: the blood antigen Rh, "Rhesus factor". These discoveries truly marked the

beginning of research in pregnancy immunology. As early as 1905, Dienst described that group O pregnant mothers, who lack A or B antigens, produce antibodies against A and B antigens if the fetus they carry expresses them. As for the Rh factor, it turned out to be the basis of an already described medical condition, hemolytic disease of the newborn: during pregnancy, antibodies produced by an Rh negative mother cross the placenta and attack the red blood cells of an Rh positive fetus. This reaction can be prevented by intra-muscular injection of anti-Rh antibodies, which clear any Rh positive fetal cells that entered the mother's blood circulation before the maternal immune system itself becomes sensitized and mounts a potent immune response against the fetus. In fact, prevention of hemolytic disease of the newborn might still be the most important clinical contribution of immunology to pregnancy (Billingham and Beer, 1984; Billington, 2015).

The first reports of cellular transfer between mother and fetus described the migration of fetal erythrocytes into the maternal blood circulation (Greger and Steele, 1957; Zipursky et al., 1963). These studies explained how pregnant mothers become exposed to, and consequently mount an antibody response against, fetal ABO and Rh antigens during pregnancy. More importantly, they provided the first solid line of evidence that fetal cells are not "shielded" from maternal blood, one of Medawar's original hypotheses to explain maternal-fetal immune tolerance (Medawar, 1953). Soon after, antibodies against paternal antigens were detected in pregnant female mice bred with genetically mismatched males (Herzenberg and Gonzales, 1962). Over the following years, accrued evidence for maternal antibodies and T cells (discussed in the next section) recognizing fetal antigens have further demonstrated that the maternal immune system is "aware" of the fetus (Arck and Hecher, 2013). This is an important point to be made, as an early working hypothesis put forward by Little, a pioneer in cancer transplantation and founder

of Jackson Laboratory, as well as by Medawar himself, to explain successful pregnancy stated that the fetus was “antigenically immature” and thus escaped recognition by the maternal immune system (Little, 1924; Medawar, 1953).

Nevertheless, B cells are exceedingly rare at the maternal-fetal interface, representing only 1% and of decidual leukocytes in mice and in humans (Arenas-Hernandez et al., 2015; Tilburgs et al., 2015a; Xu et al., 2015). Moreover, it has been proposed that fetal antigen-specific maternal B cells are deleted during pregnancy (Ait-Azzouzene et al., 1998). Therefore, B cells are very unlikely to play a significant role at the maternal-fetal interface.

1.2.2. Decidual T cells

Similar to B cells, T cells are rare in the murine placenta, comprising only 3% of total decidual leukocytes (Croy et al., 2012; Nancy and Erlebacher, 2014). Recent work has suggested that effector T cells do not accumulate at the murine maternal-fetal interface due to epigenetic silencing of T cell-attracting chemokine genes in decidual stromal cells (Nancy et al., 2012). In humans, however, T cells represent 5-20% of total decidual leukocytes, a number that can go up to 80% at term pregnancy (Tilburgs et al., 2010). As a comparison, T cells constitute 7-24% of cells in peripheral blood. Moreover, unlike peripheral T cells, there are significantly more $CD8^+$ T cells (50%) than $CD4^+$ T cells (30%) amongst decidual T cells. The remaining 20% of decidual T cells are double negative (DN) T cells, a significantly higher percentage than the 5-10% found in blood (Sindram-Trujillo et al., 2004). One explanation for this high proportion of DN T cells is the presence of γδ T cells (Tilburgs et al., 2009). γδ T cells are innate-like tissue-

resident T cells expressing an invariant T cell receptor (TCR) composed of a γ and a δ chain, as opposed to the α and β chains found in conventional αβ T cells (Vantourout and Hayday, 2013). During pregnancy, γδ T cells accumulate in the decidua in humans and in mice (Heyborne et al., 1992; Mincheva-Nilsson et al., 1994). Interestingly, preliminary studies in the mouse, followed by work with human samples, suggest that γδ T cells recognize trophoblasts using an invariant receptor other than the γδ TCR (Barakonyi et al., 2002; Heyborne et al., 1994). The function of γδ T cells at the maternal-fetal interface, however, remains to be determined.

The initial studies on T cells in pregnancy reported a bias towards T helper 2 (Th2) fate at the expense of a T helper 1 (Th1) fate during naïve $CD4^+$ T cell differentiation at the maternal-fetal interface (Lin et al., 1993). Th1 and Th2 cells represent two polarized forms of $CD4^+$ T helper cells. Th1 cells secrete IL-2, IFN-γ, and TNF-α, potentiating inflammatory responses and sometimes tissue injury. On the other hand, Th2 cells secrete anti-inflammatory cytokines, such as IL-4, IL-10, and IL-13 and are involved in humoral responses (Zhu et al., 2010). Polarization into a Th1 or a Th2 fate is mutually exclusive: IFN-γ promotes Th1 differentiation while inhibiting Th2 cell proliferation, whereas IL-4 induces Th2 differentiation and proliferation while blocking Th1 development (O'Garra and Arai, 2000). The Th1/Th2 shift offered an attractive paradigm where suppression of T cell-mediated immunity against fetal antigens was accomplished by inhibiting Th1 responses at the maternal-fetal interface (Krasnow et al., 1996; Wegmann et al., 1993). Consistent with this view, a skew towards Th1 cytokines during pregnancy was associated with recurrent abortions and preeclampsia (Piccinni et al., 1998; Saito and Sakai, 2003). Yet, this paradigm does not explain antigen-specific tolerance during pregnancy, a phenomenon accomplished by regulatory T cells (Tregs).

Originally named “suppressor T cells” at the time of their discovery (Gershon and Kondo, 1970), Tregs are a subset of $CD4^+$ T cells specialized in suppressing effector T cell responses. Natural Tregs (nTregs) develop in the thymus and are characterized by the expression of the transcription factor Foxp3 and the high affinity IL-2 receptor CD25. Tregs have been shown to essential for physiological immune tolerance and homeostasis in numerous contexts (Vignali et al., 2008). Absence of Tregs caused by mutations in the *FOXP3* gene leads to severe multi-organ autoimmunity in mice and humans (Bennett et al., 2001; Brunkow et al., 2001). As early as the 1970s, it became clear that multiple pregnancies gradually induced immune tolerance specifically towards paternal antigens, as first shown using the model H-Y male antigen (Smith and Powell, 1977). This tolerance was then shown to be transferrable by T cells (Simpson et al., 1981). Studies using two twins who had undergone multiple pregnancies confirmed the induction of Tregs specific to the antigens of their respective husbands, confirming the occurrence of this phenomenon in humans (Engleman et al., 1978).

More recent work has demonstrated that Tregs are not only present at the maternal-fetal interface, but are also required for successful pregnancy in the mouse. Shortly after implantation, Tregs are recruited to the maternal-fetal interface via uterine draining lymph nodes and protect the fetus from maternal T cell attack (Chen et al., 2013). In fact, pre-immunization of the mother against paternal antigens does not result in high rates of fetus loss unless Tregs are depleted (Chen et al., 2013). Moreover, Treg depletion in the placenta in the context of allogeneic mating leads to fetus loss and defective spiral artery remodeling in the uterus (Samstein et al., 2012). In humans, it is harder to prove the requirement for Tregs in pregnancy. Nevertheless, there is strong circumstantial evidence that Tregs play an important role in human maternal-fetal immune tolerance.

The Treg compartment undergoes expansion during pregnancy (Somerset et al., 2004). Tregs can be detected in the decidua (Tilburgs et al., 2008), and are lower in numbers in women with pregnancy complications (Arruvito et al., 2009).

1.2.3. Decidual macrophages

Macrophages are innate immune cells whose main function is to phagocytose invading pathogens, a well as cell debris. In addition, macrophages can also function as professional antigen presenting cells (APCs), presenting peptides to T helper cells (Mosser and Edwards, 2008). At the human maternal-fetal interface, macrophages represent 20% of leukocytes. Dendritic cells (DCs), on the other hand, are virtually absent from the decidua, comprising only 1.7% of decidual leukocytes, and display an immature phenotype (Gardner and Moffett, 2003). Since both B cells and DCs are present in the decidua in vanishingly small numbers, macrophages are expected to be the main professional APCs at the maternal-fetal interface. However, there is no strong evidence to date of efficient antigen presentation at the maternal-fetal interface (Searle and Wren, 1992). Instead, decidual macrophages appear to be mostly dedicated to tissue remodeling. It is thought that, in the absence of infections, macrophages residing in the decidua are inhibited as a result of an anti-inflammatory cytokine milieu (Hunt and Petroff, 2013). Consistent with this view, gene expression profiling indicates that decidual macrophages are similar, but not identical, to M2 macrophages (Gustafsson et al., 2008), i.e. they display a phenotype associated with tissue remodeling in opposition to M1 macrophages, which are pro-inflammatory (Houser et al., 2011).

In spite of the paucity of data from mice, the current model states that decidual macrophages play an important role in remodeling the maternal-fetal interface during

implantation. Observations favoring this model include the perivascular localization of macrophages in decidua, as well as expression of high levels of factors involved in tissue remodeling by decidual macrophages (Erlebacher, 2013; Gustafsson et al., 2008).

1.2.4. NK cells are the most prevalent immune cells at the maternal-fetal interface

The most abundant immune cells at the maternal-fetal interface in early pregnancy by far are Natural Killer (NK) cells, constituting up to 90% of decidual leukocytes (Koopman et al., 2003; Moffett-King, 2002). This is very peculiar, as in peripheral blood the predominant lymphocytes are T cells, with NK cells comprising only 1-6% of leukocytes. The realization that most leukocytes present at the maternal-fetal interface are NK cells has led the field of reproductive immunology to dedicate substantial effort to understanding the role of NK cells at the maternal-fetal interface.

Unlike their peripheral blood counterparts, decidual NK cells are poorly cytotoxic, despite possessing a high content of cytotoxic granules (Kopcow et al., 2005). In addition, decidual NK cells express higher levels of cytokines, chemokines, and angiogenic factors (Hanna et al., 2006). Such striking differences in phenotype suggest that decidual NK cells are generated at the maternal-fetal interface, and are not just migrants from peripheral blood. Corroborating this hypothesis, decidual stromal cells and EVT produce large amounts of TGF-β and IL-15, cytokines that promote differentiation of hematopoietic stem cells (HSCs) residing in the decidua into NK cells. Furthermore, IL-15 is essential for NK cell proliferation and survival (Keskin et al., 2007; Vacca et al., 2011).

Similarly to decidual macrophages, pregnancy has repurposed decidual NK cells for spiral artery remodeling at the maternal-fetal interface. Spiral artery remodeling during

human pregnancy is currently though to occur in two major steps. First, decidual NK cells and macrophages destroy the vascular smooth muscle cells and endothelial cells lining the spiral arteries. Second, EVT migrate into the lumen of newly remodeled artery, replacing the lost cells with a pseudo-epithelium made of trophoblasts. This process greatly enlarges the caliber of spiral arteries, ensuring maximal blood flow through the placenta to nourish the fetus (Kam et al., 1999; Smith et al., 2009). Importantly, the spiral arteries of NK cell-deficient pregnant mice display visibly thicker walls and narrower lumens than those in wild-type controls, lending support to this model (Zhang et al., 2011).

Importantly, it has been recently unveiled that NK cells are in fact only a subset of innate lymphoid cells (ILCs). As their name suggests, ILCs are innate lymphocytes, hence devoid of a recombined TCR, that can be subdivided into various subsets mirroring T helper cell subsets in terms of cytokine secretion patterns and function in tissues (Walker et al., 2013). Recently, ILCs have been found in both mouse and human deciduas (Doisne et al., 2015; Vacca et al., 2015). The relative contribution of these previously unappreciated cell populations to immune regulation at the maternal-fetal interface is currently being explored and may yield important insights into immune privilege at this site.

1.2.5. Trophoblasts are equipped with immunomodulatory molecules

Two seminal experiments in mice carried out in the early 1960s established that trophoblasts possess intrinsic immune suppressive properties. First, while syngeneic tumors transplanted into the uterine horns of mice were able to grow indefinitely, allogeneic tumors were rejected by the recipient's immune system. Pre-sensitization of

the recipient with tumor antigens accelerated the rejection process irrespective of the recipient's pregnancy status, indicating that normal immune reactions can take place unimpeded in the uterus (Schlesinger, 1962). Hence, trophoblasts must actively induce immune tolerance in order for successful pregnancy to take place. Further evidence supporting this theory was provided by the transplantation in parallel of semiallogeneic trophoblast and embryonic tissue into the kidney capsule of recipients. While the embryonic tissues were promptly rejected, trophoblasts proliferated and recruited new vessels; results from experiments using either genetically matched or mismatched trophoblasts were undistinguishable (Simmons and Russell, 1962). If immune tolerance induction is a cell-intrinsic property of trophoblasts, how is it accomplished?

Over time it has become clear that both murine and human trophoblasts express a battery of immune inhibitory molecules predominantly targeting T cells. Fas ligand (FasL), which induces apoptosis of Fas-expressing activated T cells, is expressed by human trophoblasts. It has been demonstrated that FasL-deficient pregnant mice (*gld* mice) display extensive leukocyte infiltration and killing at the maternal-fetal interface (Balkundi et al., 2000; Hunt et al., 1997; Uckan et al., 1997). Consistent with these observations, trophoblast expression of FasL was found to be necessary and sufficient to induce deletion of T cells specific to a male specific antigen, H-Y, in mice (Vacchio and Hodes, 2005). TRAIL is another example of a T cell apoptosis-inducing molecule expressed at high levels in trophoblasts (Jeremias et al., 1998; Phillips et al., 1999). Curiously, a recent study indicates that human trophoblasts secrete exosomes containing both FasL and TRAIL molecules into the extracellular space at the maternal-fetal interface; these exosomes were indeed shown to trigger T cell death (Stenqvist et al., 2013). Yet another immune suppressive molecule expressed by trophoblasts is

indoleamine 2,3-dioxygenase (IDO), an enzyme that converts tryptophan into kynurenines. IDO provokes T cell death by two simultaneous mechanisms: depletion of tryptophan, an essential amino acid, from the microenvironment, and generation of kynurenines, which are toxic to T cells (Fallarino et al., 2002; Honig et al., 2004). In the mouse, pharmacological inhibition of IDO leads to T cell-mediated rejection of allogeneic fetuses (Munn et al., 1998). In humans, IDO deficiency has been linked to preeclampsia (Kudo et al., 2003). Consistent with this association, IDO-deficient pregnant mice exhibit symptoms of preeclampsia and intrauterine growth restriction (Santillan et al., 2015).

In addition to T cell apoptosis inducers, trophoblasts also express the "immune checkpoint" molecule PD-L1 (Brown et al., 2003). Best known in the cancer immunotherapy field, PD-L1 suppresses TCR-mediated T cell activation upon binding to its co-inhibitory receptor, PD-1, which is upregulated in activated T cells. The PD-1/PD-L1 pathway is part of a number of physiological mechanisms that limit the magnitude and duration of T cell responses, in order to prevent tissue damage (Chen and Flies, 2013). Several cancer types have been shown to overexpress PD-L1 and other negative immune checkpoint molecules in order to evade immune attack. Blocking PD-L1 using monoclonal antibodies, "immune checkpoint inhibitors", has proven to be wildly successful in cancer immunotherapy (Topalian et al., 2015). In the context of human pregnancy, PD-L1 expression in trophoblasts increases during gestation, most dramatically at the onset of the second trimester. Interestingly, this upregulation matches the onset of maternal blood flow to the placenta and increase in the numbers of maternal T cells at the maternal-fetal interface, consistent with a role for fetal PD-L1 in silencing maternal alloresponses to fetal antigens (Holets et al., 2006). In the mouse, PD-L1 is expressed in decidual cells, not in trophoblasts, exclusively in allogeneic matings.

Strikingly, blockade with anti-PD-L1 resulted in a rate of 86% spontaneous fetal resorption, compared with a baseline of 18%. PD-L1 genetic deficiency led to similar results and demonstrated that PD-L1 is required specifically to prevent T cell-mediated rejection of allogeneic fetuses (Guleria et al., 2005). In addition to inhibiting effector T cell responses, PD-L1 has also been shown to induce regulatory Treg development at the maternal-fetal interface (Zhang et al., 2015). Moreover, recent work has shown that EVT can participate in the convertion of naïve T cells to Tregs (Tilburgs et al., 2015a).

1.3. HLA-G: a key determinant of immune tolerance in pregnancy?

As discussed above, several mechanisms are in place to protect invading EVT from rejection by the maternal T cells. But how do EVT interact with the most abundant immune cell type at the maternal-fetal interface, NK cells? EVT evade the maternal immune system, while actively inducing immune tolerance at the maternal-fetal interface, by expressing of a unique set of Major Histocompatibility Complex (MHC) molecules. These cell surface glycoproteins, also known as Human Leukocyte Antigens (HLA) in humans, are the main determinants of self-nonself recognition by the immune system (Bjorkman et al., 1987a; Zinkernagel and Doherty, 1974). Classical MHC class I molecules, HLA-A, -B and -C, are encoded by highly polymorphic genes, present peptides to cytotoxic CD8 T cells, and can be found or induced in virtually all nucleated cells. EVT, in contrast, are devoid of HLA-A and HLA-B expression, expressing only HLA-C and the nonclassical nonpolymorphic MHC class I molecules HLA-E and HLA-G (Apps et al., 2009). Uniquely expressed in EVT, HLA-G has been proposed to be central to maternal-fetal immune tolerance.

After its discovery using HLA locus-specific Southern blot probes (Orr et al., 1982), followed by its cloning and sequencing (Geraghty et al., 1987; Koller et al., 1989), HLA-G was found to be specifically expressed by EVT (Ellis et al., 1990; Ellis et al., 1986; Kovats et al., 1990). Importantly, *HLA-G* displays an exceedingly low level of polymorphism for an *HLA* gene, with only 51 alleles registered in the *IMGT/HLA* database (as comparison, *HLA-A* has 3,356 alleles). The unusual lack of polymorphisms in its coding region and the trophoblast-restricted expression immediately suggested that HLA-G might play a role in immune tolerance induction at the maternal-fetal interface.

1.3.1. Determining HLA-G function

The circumstancial evidence that HLA-G was a central component of maternal-fetal tolerance prompted a period of intense investigation aimed at unraveling its function. Seminal studies showed that ectopic expression of HLA-G is sufficient to inhibit cytotoxicity of both decidual (Chumbley et al., 1994) and peripheral NK cells (Deniz et al., 1994; Pazmany et al., 1996). At the time, it had just been reported that HLA-C alleles could be divided into two groups, C1 and C2, based on their recognition by Killer Immunoglobulin-like Receptors (KIRs) on NK cells. The distinction resides in a dimorphism at residues 77 and 80 of the α1 helix; the Asn77-Lys80 combination is found in HLA-C1 alleles, while Ser77-Asn80 is present in HLA-C2 alleles (Colonna et al., 1993; Mandelboim et al., 1996). Stable transfection with an HLA-C1 molecule, HLA-Cw6, inhibited killing by an NK cell line specifically recognizing HLA-C1, while transfection with an HLA-C2 molecule, HLA-Cw7, did not produce this effect. When these stably transfected target cells were incubated with an NK cell line only recognizing HLA-C2 alleles, the opposite pattern was observed, as expected. Surprisingly, transfection with HLA-G was sufficient to inhibit killing by both HLA-C1 and HLA-C2-specific NK cell lines

(Pazmany et al., 1996). The following year, a different group found that HLA-G conferred protection against peripheral blood NK cells from 20 different donors (Rouas-Freiss et al., 1997b). In addition, antibody-mediated blocking of HLA-G abrogated protection of primary trophoblasts against both matched (semiallogeneic) and unmatched (allogeneic) decidual NK cells (Rouas-Freiss et al., 1997a), while otherwise MHC class I deficient K562 cells stably transfected with HLA-G were protected against all NK cell populations tested, further suggesting that HLA-G was a "universal" NK cell inhibitory ligand.

Altogether, these seminal studies from several groups demonstrated that HLA-G ectopic expression confers protection against both peripheral and decidual NK cells from multiple donors, arguing for the existence of one or more HLA-G receptors expressed across all subsets of NK cells. This was an intriguing possibility at the time, as, even though each NK cell receptor was known to recognize more than one MHC class I molecule, there was no single molecule described capable of inhibiting all NK cells (Yokoyama, 1997). The field then refocused on a new question: what is the NK cell receptor for HLA-G?

1.3.2. A universally expressed NK cell receptor of HLA-G

The existence of an NK cell receptor for HLA-G yet to be identified was first hinted by the fact that HLA-G transfectants were protected against killing by the YT2C2 NK cell line, which expressed no known KIR receptors (Rouas-Freiss et al., 1997b). Interaction with this putative novel receptor could not involve the α2 domain of HLA-G, as expression of a splicing isoform of HLA-G devoid of the α2 domain, HLA-G2, was as protective as the full-length HLA-G protein. In addition, NK cell inhibition mediated by HLA-G occurred

even in the presence of the anti-pan-MHC class I antibody W6/32: addition of this blocking antibody rendered the HLA-G transfectants susceptible to lysis by peripheral blood polyclonal NK cells, but not by the YT2C2 NK cell line (Rouas-Freiss et al., 1997b). The W6/32 antibody reacts with the α2 and α3 domains of HLA class I molecules associated with β2-microglobulin (Tanabe et al., 1992). From this, it can be concluded that the putative unknown HLA-G receptor should recognize the α1 domain of HLA-G.

Indeed, a novel receptor for HLA-G expressed across all peripheral NK cell clones and the majority of decidual NK cells analyzed was finally described in the late 1990s by three independent groups: KIR2DL4, also known as p49 at the time (Cantoni et al., 1998; Ponte et al., 1999; Rajagopalan and Long, 1999). KIR2DL4 thus seemed to be the "universal" NK cell receptor. Of note, it was later found that the YT2C2 cell line expresses KIR2DL4, but not KIRs recognizing HLA-A, -B, -C or -E (Le Discorde et al., 2005). Furthermore, consistent with the aforementioned observations, later studies found that, similarly to recognition of HLA-C alleles by KIR receptors (Natarajan et al., 2002), the α1 domain of HLA-G is essential for optimal KIR2DL4 recognition, as mutating residues Met76 and Gln79 abrogates HLA-G binding (Yan and Fan, 2005).

Currently, HLA-G is known to bind to KIR2DL4, expressed by NK cells (Rajagopalan and Long, 1999), as well as ILT2 and ILT4, found mainly on myeloid cells (Rajagopalan and Long, 1999; Shiroishi et al., 2003). Importantly, HLA-G has been shown to be sufficient to inhibit NK cell cytotoxicity (Pazmany et al., 1996) and necesssary to protect trophoblasts against decidual NK cells (Rouas-Freiss et al., 1997a), supporting a central

role for HLA-G in inducing immune tolerance at the maternal-fetal interface. Consistent with this model, several pregancy-related disorders, including miscarriage, recurrent fetal loss and pre-eclampsia, have been associated with polymorphisms resulting in diminished *HLA-G* expression levels (Moreau et al., 2009; O'Brien et al., 2001; Quach et al., 2014). Moreover, evidence indicating that herpes simplex virus (HSV) infection impairs HLA-G trafficking to the cell surface suggests a causal link between defects in HLA-G function and the spontaneous fetal loss observed in HSV infection (Schust et al., 1996).

Have all HLA-G receptors been discovered? Many NK cell receptors remain orphan receptors, leaving open the possibility that some of them recognize HLA-G. Furthermore, the maternal-fetal interface hosts several understudied subsets of immune cells, such as γδ T cells (Heyborne et al., 1992; Mincheva-Nilsson et al., 1994) and ILCs (Doisne et al., 2015; Vacca et al., 2015), which could interact with HLA-G. In fact, a growing number of studies describe HLA-G interaction with immune cells other than NK cells, such as T cells, APCs, and even B cells (Bainbridge et al., 2000; Le Gal et al., 1999; LeMaoult et al., 2007; LeMaoult et al., 2005; Li et al., 2009; Naji et al., 2014), suggesting that those too may harbor yet-to-be-identified HLA-G receptors (Figure 1.3.2.1).

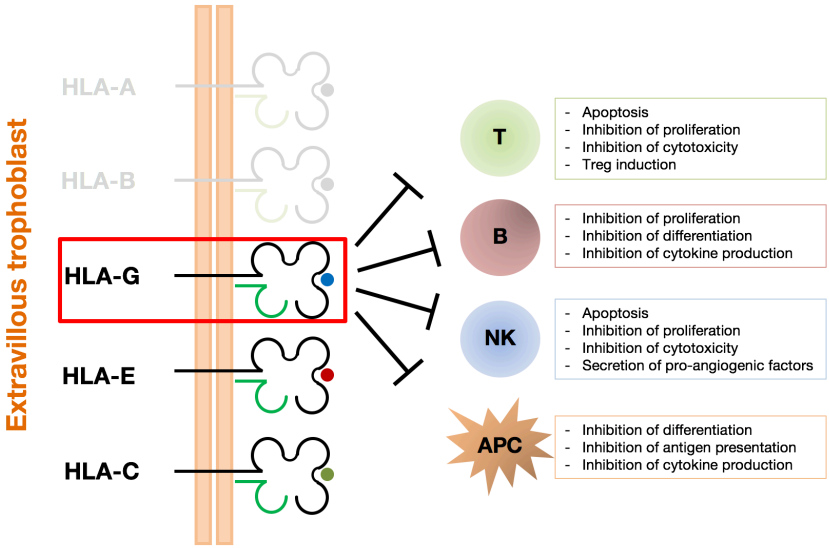


Figure 1.3.2.1. HLA-G inhibits multiple immune cell types at the maternal-fetal interface.

Extravillous trophoblasts (EVT) are devoid of HLA-A and HLA-B, the most polymorphic MHC class I molecules. Instead, they only express HLA-C, HLA-E and uniquely express HLA-G, all well-established NK cell inhibitory ligands. HLA-G, in particular, has also been shown to modulate the activity of T cells (Le Gal et al., 1999), B cells (Naji et al., 2014), and professional antigen presenting cells (APCs), such as macrophages (Li et al., 2009).

1.3.4. HLA-G expression may be co-opted by cancer

Intriguingly, HLA-G expression has also been detected in tumor lesions, where it may facilitate immune evasion (Rouas-Freiss et al., 2005; Wiendl et al., 2002). The first report of HLA-G expression in cancers, in 1996, was obtained from studies in hematopoietic cells; while no HLA-G transcript was detected in hematopoietic stem cells, thymocytes or natural killer cells, a small fraction of analyzed leukemias expressed HLA-G (Amiot et al., 1996). Two years later, detection of HLA-G expression was extended to a solid tumor

type, melanoma (Paul et al., 1998). This observation followed the realization in previous years that trophoblasts shared the expression of some genes with metastatic melanoma, such as melanoma-associated antigen genes (MAGE) and the melanoma adhesion molecule Mel-Cam (De Plaen et al., 1997; Shih and Kurman, 1996). Since then, many more parallels have been drawn between mammalian fetus immune protection and tumor immune evasion (Nehar-Belaid et al., 2016).

Given its role in inducing immune tolerance at the maternal-fetal interface, these studies suggested that HLA-G was co-opted by some malignant tumors to evade immune surveillance. Consistent with this hypothesis, the HLA-G transcript positive melanoma cell line IGR was protected against NK cell lysis, while the HLA-G transcript negative melanoma cell lines M8 and M74 were lysed. Nevertheless, the authors failed to detect full-length HLA-G protein expression in melanoma samples with HLA-G transcript: while W6/32-mediated immunoprecipitation yielded HLA-C and HLA-G in JEG3 cells, it only yielded HLA-C in HLA-G transcript positive melanoma samples. A second experiment using antibodies HCA2 and 4H84 pulled down lower molecular weight bands identified in the study as different splice isoforms of HLA-G protein that do not associate with β2-m. However, even though these bands were absent in HLA-G transcript-devoid cell lines (Paul et al., 1998), subsequent studies revealed cross-reactivity of both antibodies with HLA-A, HLA-B and HLA-E molecules (Seitz et al., 1998; Zhao et al., 2012), casting doubt on HLA-G protein expression, and thus function, in metastatic melanoma. A study by an independent group also reported HLA-G transcription in the absence of HLA-G protein, both surface and cytosolic, in hematopoietic malignancies collected from patients (Amiot et al., 1998). One exception was the monohistiocytic lymphoma cell line U937, where HLA-G surface expression could be detected with the monoclonal antibody

87G upon IFN-γ stimulation, indicating a role for HLA-G in cancer under inflammatory conditions (Amiot et al., 1998). Nonetheless, 87G binding to U937 cells stimulated with IFN-γ, monocytes and macrophages was later shown to be an artifact resulting from inappropriate Fc blocking (Blaschitz et al., 2000).

Comprehensive analysis of tumor samples from six different origins and thirty-one tumor cell lines did not find HLA-G surface expression in any of the samples, despite detecting HLA-G transcription in most samples (Real et al., 1999). While some studies found no evidence of HLA-G protein expression in melanoma, either cell lines or primary tumor samples from patients (Frumento et al., 2000), others have found HLA-G expression in a subset of melanomas characterized by high levels of HLA-G transcription (Paul et al., 1999). It should be emphasized that these independent studies were done on the same type of tumor, melanoma, using the same antibody to detect HLA-G expression, 87G, indicating that functionally meaningful HLA-G expression may be restricted to a small subset of melanoma.

Still, HLA-G protein expression has now been detected in several types of primary solid tumors in addition to melanoma, including colorectal cancer, cervical cancer, breast cancer, amongst others. In all cases, HLA-G expression was associated with disease progression and poor prognosis (de Kruijf et al., 2010; Li et al., 2012; Rouas-Freiss et al., 2005; Swets et al., 2016). Curiously, a correlation between ectopic HLA-G expression and loss of classical HLA class I expression has been observed in breast cancer (de Kruijf et al., 2010), raising the possibility that some cancers may have hijacked trophoblast-specific mechanisms of differential HLA class I gene regulation.

1.3.5. Outstanding questions in HLA-G biology

While HLA-G has a rich past in immunology research, there are still numerous questions that need to be addressed. For instance, we still do not understand the molecular events that turn on *HLA-G* expression in EVT while preventing expression of the classical *HLA-A* and *HLA-B* genes. The excitement centered on the discovery of HLA-G as a trophoblast-specific MHC molecule, along with the rapid pace at which discoveries regarding its function were published in the 1990s, is a testimony to its importance. Nevertheless, HLA-G remains the most enigmatic MHC class I molecule. Below, some of the most important unresolved questions in the field of HLA-G biology are presented.

1.3.5.1. Most wanted: better models to study HLA-G function

The most widely used *in vitro* cellular models of the human placenta are gestational choriocarcinoma-derived cell lines: JEG3 cells, which resemble EVT and express HLA-G, HLA-E and HLA-C, and JAR cells, which resemble VT and are devoid of MHC expression (Apps et al., 2009; Pattillo and Gey, 1968; Pattillo et al., 1972). Surprisingly, JEG3 and JAR have been found to be equally resistant to peripheral NK cell-induced lysis. In addition, masking MHC class I expression by JEG3 either by using blocking antibodies or acid treatment did not render them sensitive to NK cells, arguing that MHC molecules are not essential to protect trophoblasts from NK cell lysis (Avril et al., 1999). A subsequent study by the same group revealed that a mechanism through which these choriocarcinoma cell lines avoid killing by NK cells is the lack of NK cell activating ligands. The authors demonstrated that artificial activation of peripheral blood NK cells by adding PHA or agonist antibodies against CD16 resulted in killing of both JEG3 and JAR (Avril et al., 2003). A decade earlier, a different group had reported that activation of decidual NK cells, called “large granular leukocytes” at the time, with high levels of IL-2

rendered them capable of killing primary trophoblasts and the choriocarcinoma cell line JEG3 (King and Loke, 1990). However, IL-2 is absent from the maternal-fetal interface (Jokhi et al., 1994). Instead, macrophages residing in the decidua and stromal cells produce high levels of IL-15, another cytokine with a pivotal role in NK cell activation (Kitaya et al., 2000; Verma et al., 2000), suggesting that trophoblast killing by IL-2-activated decidual NK cells may not be physiologically relevant. In fact, it was later shown that, different from cells activated with IL-2, decidual NK cells activated with IL-15 did not lyse primary extravillous trophoblasts (Verma et al., 2000).

Experiments aiming to understand the contribution of HLA-G to NK cell inhibition using primary human trophoblasts have yielded conflicting results. One study found that first trimester extravillous trophoblasts are insensitive to polyclonal decidual NK cell-mediated lysis regardless of MHC class I surface expression, which was either masked by blocking antibodies or non-specifically downregulated by acid treatment (King et al., 2000). Given that better immune cell purification methods have been developed since then and genome editing in primary EVT has been accomplished (Ferreira et al., 2016), these studies warrant confirmation using better tools to rule out cell contamination or incomplete antibody blockade. Evidence that these might be the case is suggested by a study from a different group, which showed that primary human trophoblasts do become sensitive to decidual NK cell-induced lysis when their total surface MHC class I expression is masked with a pan-MHC-I antibody, yet they remain protected when only classical MHC class I (HLA-C in the case of trophoblasts) is masked (Rouas-Freiss et al., 1997a), suggesting that HLA-G is indeed essential to protect invasive trophoblasts from immune cell attack.

The observed disparity in NK cell sensitivity and a potential role for HLA-G between the trophoblast-like cancer cell lines JEG3 and JAR and primary trophoblasts may come as no surprise, as tumor cells are notable for developing multiple strategies to evade immunosurveillance – cancer immunoediting (Schreiber et al., 2011). Therefore, studies using primary EVT are indispensable to rigorously investigate whether HLA-G is necessary to achieve immune tolerance at the maternal-fetal tolerance. At present, recreation of the maternal-fetal interface with all its cell types and cytokine milieu *in vitro* has not been accomplished. One major obstacle is the fact that purified primary HLA-G$^+$ EVT are difficult to obtain in large numbers and only survive in culture for a few days, either alone or in combination with decidual leukocytes (Tilburgs et al., 2015a).

Ideally, one would use knock-out mouse models to investigate HLA-G function in more detail. However, there is no consensus on the murine ortholog of HLA-G, existing currently three candidates, all murine nonclassical MHC molecules: Qa-2 (Comiskey et al., 2003), H2-Bl (Guidry and Stroynowski, 2005) and, more recently, H2-M3 (Andrews et al., 2012). Some have argued that there simply is no ortholog of HLA-G in the mouse (Parham, 1996). In contrast, the other nonclassical HLA molecule expressed in trophoblasts, HLA-E, has a clear murine ortholog, Qa-1, which also binds MHC class I leader peptides. Nevertheless, HLA-E bearing HLA-G peptide does not bind to murine CD94/NKG2A, and Qa-1 does not bind to human CD94/NKG2A (Gays et al., 2001; Miller et al., 2003). The significant differences between mouse and human pregnancy regarding gestation time and depth of trophoblast invasion (Arck and Hecher, 2013) might preclude the need for the existence of an HLA-G ortholog to induce and maintain long-term immune tolerance at the maternal-fetal tolerance. Currently, the only

organisms besides humans where there is evidence for an HLA-G are non-human primates, much less versatile models than rodents (Golos et al., 2010).

1.3.5.2. Interplay between HLA-G and HLA-E

There is controversy in the field regarding the relative contributions of HLA-G and the other nonclassical MHC molecule expressed by EVT, HLA-E, for fetus-induced immune tolerance. The initial experiments establishing the inhibitory properties of HLA-G were performed using an MHC class I deficient lymphoblastoid cell line, LCL721.221, which does not reflect the specific MHC class I expression pattern of EVT. This discrepancy is important; HLA-C and HLA-E are well-known inhibitors of NK cell function, through binding to the inhibitory KIR2DL and CD94/NKG2A receptors, respectively (Boyington et al., 2000; Braud et al., 1998; Winter and Long, 1997). Furthermore, overexpression of HLA-G (or HLA-C) in LCL721.221 cells upregulates surface expression of HLA-E (Lee et al., 1998; Navarro et al., 1999). This phenomenon is due to the fact that HLA-E presents leader peptides of MHC class I molecules, including HLA-G (Lee et al., 1998). In fact, loading HLA-E with an HLA-G-derived nonamer peptide results in the highest interaction affinity with CD94/NKG2, when comparing across peptides derived from different HLA proteins (Llano et al., 1998). This observation suggests that, in EVT, HLA-E mainly presents HLA-G-derived peptides. Therefore, NK cell inhibition by HLA-G^+ EVT could be achieved by interaction of NK inhibitory receptors either directly with surface HLA-G, and/or with HLA-E complexed with an HLA-G-derived peptide, leading some to propose that the main role of HLA-G expression in trophoblasts is to boost HLA-E expression (Bainbridge et al., 2001; Guleria and Sayegh, 2007). It would be interesting to revisit the experiments showing HLA-G sufficiency to block NK cell responses using a wild-type and a modified HLA-G construct with a mutated leader peptide sequence that could not

be presented by HLA-E. Or, perhaps even better, test whether ectopically expressing HLA-G in cells where HLA-E has been genetically deleted still provides protection against NK cells. These experiments would bring one closer to teasing apart the contributions of these two nonclassical MHC molecules to tolerance induction by trophoblasts.

1.3.5.3. Trogocytosis and intracellular signaling

In addition to all the processes discussed above, there is a more recently proposed mechanism for HLA-G-induced immune tolerance: HLA-G trogocytosis. As hinted by its name, derived from the Greek *trogo* – to nibble – trogocytosis consists of membrane protein transfer between cells during contact. Importantly, trogocytosis involves the transfer of membrane fragments, not individual molecules, signifying that some proteins are transferred passively between donor and acquirer cells. Despite having been first observed in immune cells in the late 1990s (Huang et al., 1999; Patel et al., 1999), its prevalence, mechanism of action, and purpose remain unclear. NK cells (Caumartin et al., 2007) and T cells (LeMaoult et al., 2007) can acquire HLA-G in this manner. Strikingly, HLA-G acquisition was shown to confer the recipient cells an immune suppressive phenotype in both studies. In myeloma patients, HLA-G trogocytosis has been correlated with a poor prognosis (Brown et al., 2012). The authors showed that T cells were the preferential recipients of HLA-G from malignant myeloma cells, conferring them a suppressor phenotype (Brown et al., 2012). More recently, HLA-G trogocytosis has been confirmed in human pregnancy (Tilburgs et al., 2015b). But why evolve a system where HLA-G is transferred between cells if HLA-G is already a cell surface molecule that can bind inhibitory receptors on the surface of multiple immune cells? One possible answer may lie with its main receptor on lymphocytes, KIR2DL4. Unlike all

other KIRs, KIR2DL4 is not continuously on the cell surface of NK cells. Instead, it mostly resides in endosomes, where it signals upon binding to HLA-G. A series of studies carried out by Rajagopalan and Long revealed that HLA-G interaction with KIR2DL4 in the endosome results in pro-inflammatory and pro-angiogenic cytokine production characteristic of decidual NK cells during spiral artery remodeling (Rajagopalan et al., 2010). In fact, treatment of peripheral NK cells with a soluble form of HLA-G triggers a transition into a decidual NK cell-like global gene expression profile (Rajagopalan and Long, 2012). One can envision that HLA-G trogocytosis by NK cells at the maternal-fetal interface allows for more prolonged KIR2DL4-mediated signaling.

1.3.5.4. Trophoblast-specific HLA expression

Three decades ago, EVT were first found to express MHC molecules other than the classical HLA-A and HLA-B molecules (Redman et al., 1984). This seminal observation led to the identification of a novel unusual MHC molecule in trophoblasts (Ellis et al., 1986). Later named HLA-G, it was found to be expressed uniquely in EVT (Kovats et al., 1990). Since then, HLA-G has been regarded as a central molecule in immune tolerance induction at the maternal-fetal interface (Table 1). However, despite substantial effort, the mechanism by which the EVT-specific expression of HLA-G is obtained has remained elusive.

Table 1. Milestones in HLA-G biology

Year	Milestone	Reference
1982	Detection of a novel HLA class I gene using Southern blot	(Orr et al., 1982)
1984	EVT express HLA molecules other than HLA-A or HLA-B	(Redman et al., 1984)
1986	Novel HLA molecule with short cytoplasmic tail found in trophoblasts	(Ellis et al., 1986)
1987	6.0 kb HindIII restriction fragment cloned from HLA locus: HLA 6.0	(Geraghty et al., 1987)
1990	HLA6.0 is renamed HLA-G, the newest HLA class I gene	(Bodmer et al., 1990)
1990	HLA-G is uniquely expressed in EVT	(Kovats et al., 1990)
1994	HLA-G is sufficient to inhibit decidual NK cell killing	(Chumbley et al., 1994)
1996	HLA-G is sufficient to inhibit peripheral NK cell killing	(Pazmany et al., 1996)
1996	Detection of HLA-G expression in blood cancers	(Amiot et al., 1996)
1996	HLA-G presents peptides	(Diehl et al., 1996)
1998	Detection of HLA-G expression in solid tumors	(Paul et al., 1998)
2000	The proximal promoter of *HLA-G* is defective	(Gobin and van den Elsen, 2000)
2001	*HLA-G* polymorphisms are associated with pregnancy complications	(O'Brien et al., 2001)
2005	Crystal structure of HLA-G	(Clements et al., 2005)
2007	HLA-G can be transferred to effector immune cells via trogocytosis	(LeMaoult et al., 2007)
2012	HLA-G induces quiescence in peripheral NK cells	(Rajagopalan and Long, 2012)
2015	HLA-G trogocytosis at the maternal-fetal interface	(Tilburgs et al., 2015b)

1.4. Tissue-specific gene regulation

Tissue-specific gene expression is primarily regulated at the transcriptional level by *cis*-regulatory DNA elements – enhancers – which can be as far as 1 Mb away from the gene they regulate. Long-range enhancer-promoter interactions are crucial in the control of cell type-specific gene expression, with classical examples including the *β-globin* and *Hox* gene loci (Heinz et al., 2015; Holwerda and de Laat, 2012). Unexpectedly, genome-

wide association studies (GWAS) have revealed that the majority (90%) of disease-associated genetic variants occur in noncoding portions of the genome (Maurano et al., 2012), suggesting that enhancers may also play a role in disease.

The term “enhancer” was first coined to describe sequences derived from the SV40 virus capable of “enhancing” the expression of a heterologous gene in human cells regardless of their position or orientation relative to it (Banerji et al., 1981). Far from being readily accepted, it took over one decade for the scientific community to integrate this new concept. How could an enhancer regulate a gene regardless of being upstream or downstream thousands of base pairs away? Two main models emerged to explain these surprising results: the scanning model and the looping model. In the first, RNA polymerase or a transcription factor binds to the enhancer (the “entry site”), scans along the DNA until it reaches the closest promoter, and activates transcription. In the latter, the enhancer directly interacts with the promoter via transcription factors bound to it, forming a DNA “loop” (Ptashne, 1986). By 1989, a series of experiments to prove the looping model had been done. In one experiment, activation of two nearby promoters by an enhancer was found to be equivalent. This result refuted the scanning model, which predicted that the promoter closer to the enhancer should be more strongly activated (Heuchel et al., 1989). A second experiment further disproved the scanning model, while lending further support for the looping model. In this study, tethering of an enhancer sequence directly to the promoter via streptavidin or avidin resulted in higher transcriptional activity, suggesting that enhancers indeed loop onto their target promoters to activate transcription (Mueller-Storm et al., 1989). In light of these findings, a new question arose: how do distant enhancers discriminate between promoters and only activate the right one?

1.4.1. Seeing the genome in 3D

Already in the 1980s, it was evident that the genome was partitioned into domains inside the nucleus. It was hypothesized at the time that this organization was important not only for compaction of chromatin to fit into the nucleus, but also for the creation of isolated gene expression domains (Kellum and Schedl, 1991). In addition, it had been found that actively transcribed genes preferentially associate with the nuclear matrix, a network of "scaffold" fibers throughout the nucleus that maintains its structure along with the nuclear lamina (Ciejek et al., 1983). The next year, Laemmli and co-workers described Scaffold Attachment Regions (SARs) in the *Drosophila* genome, conserved sequences found across the genome responsible for attachment of the genome to the nuclear matrix (Mirkovitch et al., 1984). Two years later, analogous sequences were described in the mouse genome, Matrix Attachment Regions (MARs). MARs were described as being constitutive, with the same matrix contact occurring whether a given gene is active or silent (Cockerill and Garrard, 1986). These regions are collectively known today as S/MARs (Heng et al., 2004).

In 1991 the term "insulator" was introduced in the lexicon of transcriptional regulation, after the demonstration that the boundary elements flanking a *Drosophila* heat shock gene locus can "insulate" a heterologous gene from the influence of regulatory elements outside the region defined by the boundary elements, or "insulators" (Kellum and Schedl, 1991). This work established that the genome is divided into independent transcriptional units. In vertebrates, insulators bind the protein CTCF (CCCTC-binding factor) (Bell et al., 1999; Dowen et al., 2014). Of note, the orientation of CTCF binding sites determines the direction of chromatin looping. Inversion of CTCF binding sites in the genome leads to a dramatic reconfiguration of the genome topology and changes in gene expression

(Guo et al., 2015). In addition to CTCF, insulators require the cohesin complex that is recruited by CTCF. Together, CTCF and cohesins are involved in chromatin looping and creation of insulated regulatory domains. Even though the molecular details of CTCF and cohesin interactions are still not fully understood, they provide a conceptual basis to understand how enhancer-promoter communication is moderated by the 3D genome organization (Phillips and Corces, 2009; Vietri Rudan and Hadjur, 2015).

At the turn of the century, the importance of long-range chromatin interactions and chromatin architecture were well established. However, the study of genome organization was limited by the lack of tools to assess individual physical interactions between enhancers and promoters. What if one could sequence the genome in 3D? Fortunately, a new technique suited to address this problem was just being developed in the laboratory of Nancy Kleckner: Chromatin Conformation Capture (3C). The principle behind 3C is the use of chemical cross-linking to fix the all the interactions between different loci. The fixed chromatin is then isolated, digested with a restriction enzyme, ligated at low concentrations to ensure that only interacting regions ligate, and analyzed by PCR (Dekker et al., 2002). The steps of a 3C experiment are outlined in Figure 1.4.1.1.

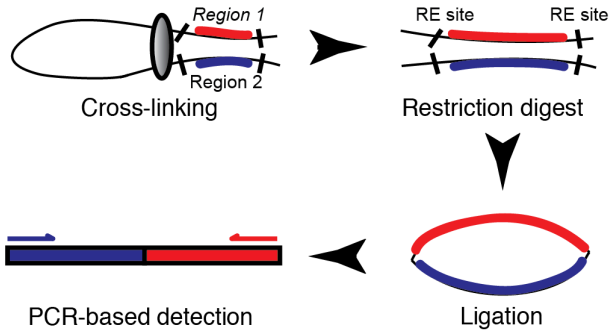


Figure 1.4.1.1. Chromatin Conformation Capture (3C).

(Continued) A 3C experiment encompasses four main steps: chromatin cross-linking, restriction digest with an enzyme recognizing sites flanking the regions of interest ("Region 1" and "Region 2"), ligation at the resulting restriction fragments at low concentrations, and PCR detection of interaction between "Region 1" and "Region 2". RE, restriction enzyme.

This conceptually simple technique was quickly revealed to be extremely powerful. Fast-forwarding almost 15 years, 3C and its variants have been used to reveal the structure of whole genomes. The variants include 4C (Circular 3C), where one region of interest can be tested for physical interactions against the rest of the genome, 5C (Carbon Copy 3C), which uses high-throughput sequencing to analyze a complex library containing all interactions within a given locus, and, finally, Hi-C, which generates lower resolution chromosome-wide maps of interactions (de Wit and de Laat, 2012). What have we learned?

1.4.2. Chromatin is a TAD more organized than we thought

In 2012, the concept of Topological Associated Domain (TAD) emerged from work in *Drosophila* and in mice using whole genome Hi-C techniques (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). These seminal studies were the first to provide evidence that, far from being random, chromatin folding at the sub-megabase level is partitioned into a succession of TADs, characterized by strong anchoring points in the absence of transcriptional output, and remarkably consistent across cell types of a given species. In fact, TAD boundaries also display conservation across species, suggesting the existence of evolutionary pressure.

What are the factors that drive chromatin folding into TADs? Interestingly, various histone modification patterns align well with TADs, namely the repressive chromatin marks dimethylation of histone H3 at lysine 9 (H3K9me2), and trimethylation of histone H3 at lysine 27 (H3K27me3), suggesting a role in TAD formation. However, analysis of mutant mice lacking each one of these marks yielded no alterations in size or position of TADs (Nora et al., 2012). Instead, TAD organization seems to be determined mostly by discrete elements at the boundaries. In agreement with this model, deletion of a multi-kb region encompassing a boundary between TADs on the X chromosome in mouse ES cells resulted in ectopic contacts between the two TADs. This control is not complete, however, as the TADs did not completely merge, indicating the involvement of elements within TADs (Nora et al., 2012). In agreement with these observations, inversion of a 2.7 Mb region spanning two TADs in the murine *HoxD* locus (chromosome 2) also did not provoke mixing of the two TADs (Lonfat et al., 2014). Whether these internal sequences function to maintain TAD integrity physiologically or only when boundaries between TADs are disrupted remains unclear.

What is the nature of inter-TAD boundaries? Are they relevant? The current paradigm states that boundaries are not involved in TAD creation. Yet, once TADs are formed, they play a key role in maintaining TADs perfectly isolated from one another. Dissecting the properties of the putative elements at the borders between TADs (and perhaps within TADs?) responsible for their formation is revealing to be difficult, as forced mixture of two adjacent TADs has not yet been accomplished. Nevertheless, some general features have arisen already: TAD boundaries display the highest concentrations of bound CTCF in the mammalian genome and correspond to highly GC-rich regions (Lonfat and

Duboule, 2015). An additional factor that may play a role in the formation of TAD boundaries is the location of Lamina Associated Domains (LADs) in the genome. Recently, Hi-C experiments have revealed that TADs form yet higher-order chromatin structures, “metaTADs”. More than half of metaTAD boundaries were found to overlap with LADs (Fraser et al., 2015). LADs are regions of the genome that contact the nuclear lamina, and are characterized by low levels of transcriptional activity, repressive histone marks and low gene density (Guelen et al., 2008; Kind et al., 2015). Indeed, association of genes with the nuclear lamina often coincides with their repression during differentiation. Conversely, loss of association with the nuclear lamina and delocalization away from the nuclear periphery often tallies with transcriptional activation (Kosak et al., 2002). Strikingly, this apparent link between transcriptional activity and association with the nuclear lamina was demonstrated by studies where forced association of a gene with the nuclear lamina via protein tethering led to silencing of that gene (Reddy et al., 2008).

A deeper understanding of how chromatin structure is formed, regulated, and how it impacts gene regulation is crucial in both physiological and disease settings. There exist various reports of enhancer hijacking in cancer, where chromosomal rearrangements bring enhancers in close proximity to proto-oncogenes, resulting in their aberrant expression (Drier et al., 2016; Groschel et al., 2014; Northcott et al., 2014). But only very recently was it found that disrupting chromatin structure could achieve the same result in tumor cells. Microdeletions disrupting CTCF-insulated chromatin neighborhoods, leading to the activation of proto-oncogenes by enhancers normally located outside of the neighborhood, have been recently described in leukemia (Hnisz et al., 2016). In addition, studies where a human regulatory region was inserted in the mouse genome at an ectopic site in *trans* (i.e. on a different chromosome) found that the regulatory region

sampled a limited space in the genome and only activated its target gene in cells where inter-chromosomal contact was established, illustrating the importance of 3D genome organization as a moderator of gene regulation (Noordermeer et al., 2011). Yet, the paramount importance of genome 3D architecture in gene regulation has been most dramatically demonstrated by recent work where forced chromatin looping was shown to be sufficient to re-activate expression of a developmentally silenced gene (Deng et al., 2012; Deng et al., 2014).

1.4.3. Impact of chromatin structure on immune gene regulation

Importantly, chromatin higher-order structure and long-range interactions have been shown to play a role in MHC gene regulation. Jeremy Boss and colleagues were the first to provide evidence of the involvement of chromatin looping in MHC gene expression. Sequence analysis of the MHC class II locus revealed the existence of intergenic RFX-binding sites in addition to the X-Y elements present in the classical promoters of MHC class II genes. One of these intergenic elements, XL9, is located between the *HLA-DRB1* and *HLA-DQA1* genes, which are separated by 44 kb. Curiously, XL9 displayed high levels of chromatin acetylation, indicative of regulatory activity, yet did not bind RFX, and did not possess either enhancer or repressor activity. Instead, the authors found that XL9 bound CTCF and associated with the nuclear matrix (Majumder et al., 2006). In subsequent studies, CTCF was found to mediate a long-range chromatin looping interaction between the promoters of *HLA-DRB1* and *HLA-DQA1*. Disrupting this interaction by depleting CTCF led to a marked reduction in the expression of these genes (Majumder et al., 2008). To date, there is no report of enhancer looping in the MHC class I locus. Could there be a distant enhancer controlling the trophoblast-restricted expression of HLA-G?

1.5. Transcriptional regulation of HLA-G

Despite decades of work delineating the transcriptional regulation of MHC genes, the mechanisms behind HLA allele-specific expression are not fully understood. In particular, how the tissue-specific expression of HLA-G is accomplished remains unknown. The first efforts to unravel HLA-G regulation focused on dissecting the "classical" promoter of *HLA-G*. Every HLA gene contains this conserved classical promoter sequence, which is responsible for both basal and induced gene expression (Kobayashi and van den Elsen, 2012). The classical promoter region of MHC class I genes harbors several well-defined regulatory motifs: Enhancer A, Interferon-Stimulated Response Element (ISRE) and the SXY module (Solier et al., 2001). Enhancer A is bound by NF-κB, downstream of TNF-α signaling, while the ISRE binds IRF1, mediating IFN-γ-induced *HLA* upregulation (Solier et al., 2001). SXY sequences are recognized by ATF1 and CREB1 transcription factors, as well as by the RFX complex (RFX5, RFXAP and RFXANK). The current paradigm states that these transcriptional regulators are assembled in an enhanceosome complex via association with the transactivator NLRC5, resulting in MHC class I gene expression (Kobayashi and van den Elsen, 2012; Meissner et al., 2012a; Meissner et al., 2012b).

The original wave of studies on *HLA-G* transcriptional regulation, performed in the laboratory of Peter van den Elsen, revealed almost one motif at a time that the classical promoter of *HLA-G* is mostly non-functional. The first report of a non-functional element within the *HLA-G* promoter focused on Enhancer A (Gobin et al., 1998). This regulatory element can be further subdivided into κB1 and κB2 sites, which bind NF-κB. NF-κB itself is a heterodimeric transcription factor composed of two Rel proteins, p50 and p65. The authors found that the *HLA-G* promoter version of the κB1 and κB2 sites bind the

p50 subunit of NF-κB, but not p65, rendering it unresponsive to TNF-α stimulation. Yet, the HLA-G κB2 site was still bound by Sp1, a ubiquitously expressed transcription factor (Gobin et al., 1998). The following year, the *HLA-G* promoter ISRE element was found to be partially deleted and unresponsive to IFN-γ stimulation (Gobin et al., 1999). Curiously, two alternative putative elements were later identified near this defective ISRE, but they were found to be equally non-functional (Chu et al., 1999; Lefebvre et al., 2001). Finally, the SXY module, comprising the S, X1, X2 and Y elements of the classical promoter, was also found to be divergent in the *HLA-G* promoter: even though the X1 element can bind to RFX factors and Sp1, the X2 and Y elements are not functional (Gobin and van den Elsen, 1999, 2000; Rousseau et al., 2000). Altogether, these data demonstrated that the classical promoter of *HLA-G* could not explain its trophoblast-restricted expression. If anything, it may be part of the reason why most cell types do not express HLA-G. Of note, EVT do not express NLRC5 (Tilburgs et al., 2015a), preventing, in theory, the assembly of the MHC enhanceosome. A schematic of the classical promoter of *HLA-G* can be found in Figure 1.5.1.

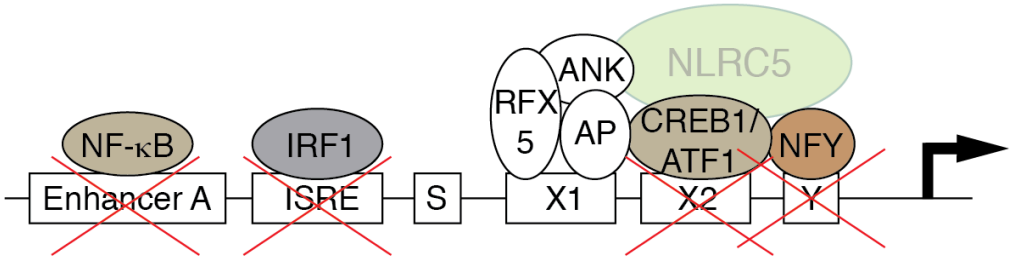


Figure 1.5.1. The classical promoter of *HLA-G* is defective.

Despite its similarity in structure with other MHC class I gene promoters, the *HLA-G* classical promoter harbors mostly non-functional versions of the regulatory elements found (Solier et al., 2001). The *HLA-G* promoter fails to be transactivated by NF-κB,

(Continued) IRF1 or CIITA (Gobin and van den Elsen, 2000). In addition, NLRC5 is not expressed in EVT (Tilburgs et al., 2015a).

Recently, post-transcriptional regulation at the 3' untranslated region (UTR) of *HLA-G* has been proposed to play a role in HLA-G repression in non-trophoblast cells. A few years after the discovery of *HLA-G*, in 1993, a 14 bp indel polymorphism was identified in the 3'UTR of *HLA-G* (Harrison et al., 1993). It then took one decade to observe that the presence of the 14 bp insertion in the genome leads to increased half-life of *HLA-G* mRNA (Rousseau et al., 2003), and yet another decade to elucidate the mechanism behind this difference in mRNA stability: the 14 bp indel overlaps with miRNA binding sites! The authors found that miR148a and miR152 downregulate *HLA-G* expression; these miRNAs are expressed at low levels in trophoblasts compared with all tissues assessed (Manaster et al., 2012). Yet, similar to the work on the *HLA-G* promoter, these data shed light on how HLA-G expression is repressed in most cells, not on how it is specifically activated in trophoblasts. The field then turned to analyzing sequences upstream of *HLA-G*. Could there be unique enhancer sequences dedicated to the transcriptional regulation of *HLA-G* that would compensate for or cooperate with the defective *HLA-G* classical promoter?

A series of elegant experiments using transgenic mice led to the identification of a Locus Control Region (LCR) located 1 kb upstream of the *HLA-G* promoter. In brief, transgenic mouse embryos carrying either the full 6 kb of the *HLA-G* gene or a 5.7 kb 5' truncated version were created. HLA-G was expressed in trophoblasts in the placenta of embryos carrying the full length *HLA-G* gene. Surprisingly, however, the truncated version of *HLA-*

G was expressed in mesenchymal cells in the placenta instead (Yelavarthi et al., 1993). In a subsequent study, replacing the 5' upstream sequence, first exon, and first intron of HLA-G with the corresponding sequences from HLA-A created a hybrid gene that was ubiquitously expressed in transgenic animals (Schmidt et al., 1993). Together, these experiments established that a 250 bp region 1 kb upstream of *HLA-G*, LCR, is critical to control tissue-specific HLA-G expression. A second pair of *in vitro* experiments found that the LCR sequence binds trophoblast-specific transcription factor complexes; their identity, however, remains to be determined (Moreau et al., 1998; Moreau et al., 1997). Nevertheless, subsequent studies identified three functional CREB/ATF binding sites within the LCR (Gobin et al., 2002). This was an interesting finding at the time, as these upstream elements might compensate for the defective X2 element in the *HLA-G* classical promoter, which is also a CREB/ATF binding site in other HLA class I genes (Solier et al., 2001). However, in our hands, addition of all three CREB/ATF binding sites upstream of the *HLA-G* classical promoter did not result in reporter gene activity in the HLA-G$^+$ JEG3 trophoblast cell line (the promoter alone has no activity). Still, the same hybrid construct was active in the HLA-G negative 293T cell line, demonstrating that the sites are indeed functional. In any scenario, this mechanism would not suffice to explain trophoblast-specific HLA-G expression, as CREB1 and ATF1 are ubiquitously expressed transcription factors. Subsequent studies focused on the *HLA-G* LCR candidate described a RREB-1 binding site. Transfection with RREB-1, a transcriptional repressor that acts via recruitment of chromatin modifiers, was shown to repress reporter gene activity driven by a construct containing the LCR and the *HLA-G* proximal promoter, leading the authors to propose that RREB-1 represses HLA-G expression in HLA-G negative cells (Flajollet et al., 2009). However, RREB-1 is highly expressed in EVT and JEG3 cells (Ferreira et al., 2016; Tilburgs et al., 2015a), casting doubt on this

hypothesis. Furthermore, in our hands, the *HLA-G* LCR was able to drive gene expression when combined with the *HLA-G* promoter in both HLA-G$^+$ and HLA-G$^-$ cell lines, indicating that it is not cell-type specific.

Recently, a negative regulatory region overlapping with a Long Interspersed Element (LINE) 4 kb upstream of *HLA-G* has been described: LINE1 (represented together with other upstream regulatory elements in Figure 1.5.2). LINE sequences are part of a group of retrotransposons, highly repetitive elements found in eukaryotic genomes originated though integration of retroviruses. The LINE1 element upstream of *HLA-G,* shown to repress gene expression, is very AT-rich (*ca.* 60%), presenting a high probability of forming hairpin loops. It has been proposed that these loops are involved in repressing HLA-G expression (Ikeno et al., 2012), but how this mechanism is prevented in HLA-G$^+$ trophoblasts remains unknown.

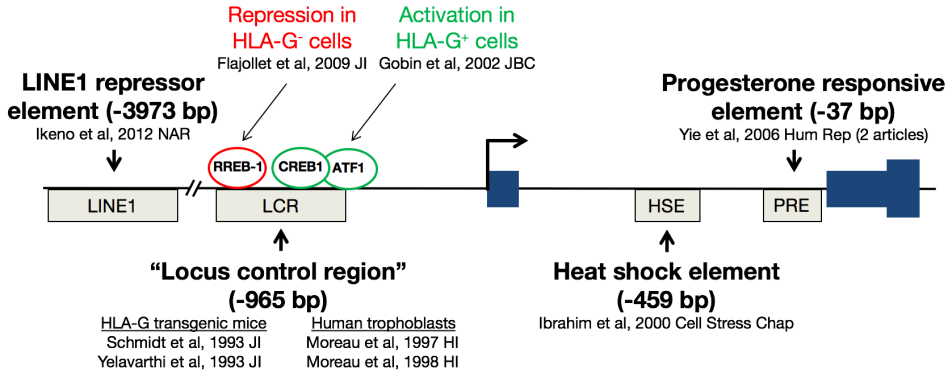


Figure 1.5.2. Nonclassical regulatory elements upstream of *HLA-G*.

Schematic summarizing the main regulatory elements upstream of *HLA-G*. In addition to the LINE1 (Long Interspersed Nuclear Element 1) and LCR (Locus Control Region) elements, a Progesterone Responsive Element (PRE) and a Heat Shock Element (HSE) are also represented. The PRE was shown to be involved in progesterone-induced upregulation of HLA-G (Yie et al., 2006a; Yie et al., 2006b), while the HSE was shown to

(Continued) mediate stress-mediated upregulation of HLA-G (Ibrahim et al., 2000). All distances, in base pairs (bp), are relative to the start of the classical promoter of *HLA-G*. Transcription factors in green are activators of *HLA-G* expression, while those in red are suppressors of *HLA-G* expression. The bent arrow denotes the transcription start site.

The question thus remains: how is HLA-G expression accomplished specifically in EVT? Traditionally, enhancer discovery has relied on examining features predictive of enhancer activity, such as chromatin accessibility, DNA and chromatin covalent modifications, and sequence conservation between species (Consortium, 2012). This approach has been successfully used to gain important insights into immune gene regulation, such as the discovery of enhancers controlling the expression of murine Foxp3, a transcription factor governing the commitment and stability of regulatory T cells (Zheng et al., 2010). However, substantial differences in regulatory sequences between species limit the ability to derive conclusions from model organisms regarding human gene regulation. In particular, the MHC locus differs significantly between mouse and humans (Yuhki et al., 2003), and *HLA-G* lacks a clear ortholog in mice.

What if one could interrogate the *HLA-G* locus for enhancer activity directly? The first enhancers described were SV40-derived strong enhancers that increased β-globin gene expression 200-fold at the mRNA level. In addition, the authors could quantify β-globin protein levels (Banerji et al., 1981). What if the enhancer being tested has a weaker effect? Or the target gene has a low expression level to start with? By the beginning of the 1980s, cell-free systems to study eukaryotic mRNA synthesis were in place (Manley

et al., 1980; Weil et al., 1979). In theory, one could utilize such systems using different regulatory DNA elements as templates and then measure the amount the mRNA produced using different regulatory DNA elements. However, it was not clear at the time whether such cell-free systems would respond to regulatory signals, and quantifying RNA was laborious and inaccurate. Ideally, the elements being tested would be linked to the expression of an invariant gene whose activity could be easily measured: a “reporter gene”. The first version of a reporter gene was chloramphenicol acetyltransferase (CAT). CAT activity was quantified by measuring the acetylation of radioactively labeled chloramphenicol using silica gel thin layer chromatography (Gorman et al., 1982). The labor-intensive CAT system was quickly replaced with firefly luciferase. In the presence of its substrate, luciferase catalyzes two reactions, which culminate in the emission of light in the form of luminescence. In addition to being faster and more facile to assay, luciferase has been estimated to be up to 1,000-fold more sensitive to changes in transcriptional output than CAT (de Wet et al., 1987). Yet, a challenge remains: what if one wants to test many regions for enhancer activity? Still today, enhancer candidates (usually resulting from epigenetic or chromatin accessibility mapping) are typically tested one by one by cloning them into a promoter-containing luciferase reporter gene plasmid.

We decided to tackle the question of trophoblast-specific HLA-G expression using a high-throughput unbiased approach: a massively parallel reporter assay (MPRA) (Melnikov et al., 2012). Building on the recent advances in DNA synthesis and sequencing, MPRA allows one to functionally test 12,000 enhancer candidates at once. In brief, enhancer candidates up to 200 bp long, in order to prevent misspellings, are synthesized and cloned into a library with 12,000 plasmids. These plasmids all share the same backbone, invariant promoter and luciferase gene. However, each one of them

has a unique DNA tag at the end of the luciferase gene, such that the mRNA produced by each one of them can be traced back to its template construct by RNA-seq. The resulting complex enhancer library is transfected into the cell line of interest, followed by mRNA purification, cDNA synthesis and high-throughput sequencing. Thanks to RNA-seq, it is possible to quantify enhancer activity in the most direct possible way: increase in the number of mRNA molecules produced by a given enhancer construct over the median of the 12,000 constructs being simultaneously tested (Melnikov et al., 2012). Unlike any other approach for enhancer discovery, MPRA allows the direct functional interrogation of a locus without any prior knowledge about its regulatory landscape. We carried out an MPRA tiling the entire *HLA-G* locus, defined by us as a 37 kb region encompassing the sequence upstream of *HLA-G* until the nearest 5' gene and the entire *HLA-G* gene, searching for previously unidentified enhancers that might explain the trophoblast-specific expression of HLA-G (Figure 1.5.3).

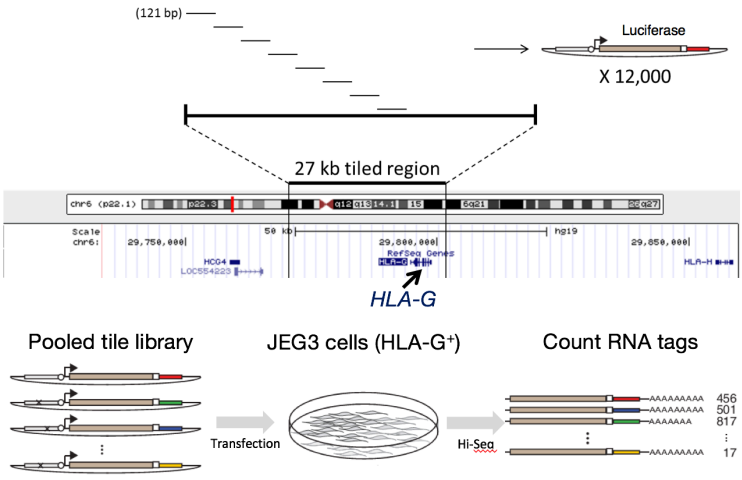


Figure 1.5.3. Strategy to functionally dissect the *HLA-G* locus using MPRA.

(Top) Partially overlapping 121 bp-long oligonucleotides covering a 27 bk region encompassing the *HLA-G* gene and the region until the upstream gene, *HCG4*, were synthesized and cloned into an invariant promoter-luciferase plasmid library (12,000

(Continued) constructs). (Bottom) Each luciferase plasmid RNA-seq tag is unique, illustrated here by a color code (red, green, blue, yellow). After transfection into JEG3 cells, expression of the library, RNA isolation, library preparation, and sequencing, the enhancer activity of each cloned element is determined by quantifying the number of mRNA molecules transcribed from each plasmid (sequencing counts), represented here by the numbers next to the mRNA molecules on the right.

In the present study, we describe a novel trophoblast-specific enhancer of HLA-G expression located 12 kb upstream of *HLA-G*. Unlike previously described regions, this 121 bp-long regulatory element, *Enhancer L*, is active specifically in HLA-G$^+$ cells. Moreover, *Enhancer L* is absolutely required for HLA-G expression in JEG3 cells and in primary EVT, as shown by CRISPR/Cas9-mediated genomic excision. Chromatin conformation capture (3C) and immunoprecipitation (ChIP) assays support a model where *Enhancer L* loops into the core promoter of *HLA*-G upon association with trophoblast CEBP and GATA transcription factors previously involved in trophoblast development and function. These findings establish chromatin looping mediated by lineage-specific transcription factors as a premier mechanism governing tissue-specific gene expression at the maternal-fetal interface.

CHAPTER 2: The discovery of *Enhancer L*, a trophoblast-specific enhancer of HLA-G expression

2.1. Identification of a trophoblast-specific Enhancer 12 kb upstream of HLA-G

In order to systematically interrogate the *HLA-G* locus for active *cis*-regulatory elements, we set up a Massively Parallel Reporter Assay (MPRA) screen (Melnikov et al., 2012). For this purpose, 12,000 partially overlapping 121 bp-long elements (tiles) spanning 27 kb of the *HLA-G* locus were synthesized, coupled to unique DNA tags, and cloned into plasmids containing an invariant promoter and a firefly luciferase reporter gene. For greater confidence, two different promoters were used in parallel libraries, a strong promoter (SV40P) and a minimal TATA box synthetic promoter (minP). The resulting libraries were co-transfected into JEG3 cells, an HLA-G$^+$ choriocarcinoma cell line commonly used to model extravillous trophoblasts (EVT) (Pattillo and Gey, 1968). To measure the relative enhancer activity of each tested element, we performed high-throughput sequencing and quantified the relative abundance of each element's tag reads in mRNA isolated from the transfected cells and in the pooled libraries. Enhancer activity was calculated as the median (cDNA count divided by the DNA count) of tags representing a tile, divided by the median ratio for all tags in a library. Nominal candidates were defined as any tile where enhancer activity measurements were >1 and p-values <0.05 for both biological replicates of each library transfection.

Our unbiased MPRA screen yielded several enhancer candidates upstream of *HLA-G* (Figure 2.1.1). We were able to retrieve 5008 tiles coupled with minP and 4992 tiles coupled with SV40P evenly distributed across the tiled 27 kb region. Of those,

approximately 50 tiles combined with each promoter reached p<0.05 in both biological replicates, represented in Figure 2.1.1, with *ca.* 20 of which displaying enhancer activity >2.

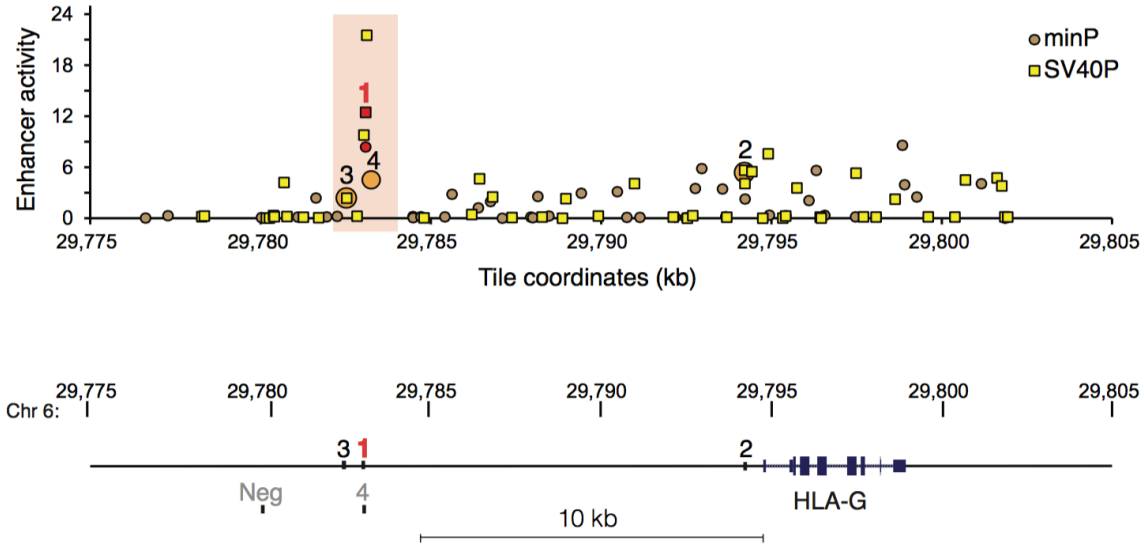


Figure 2.1.1. Massively Parallel Reporter Assay (MPRA) covering the *HLA-G* locus. (Top) Enhancer activity of tiles upstream of the minP (circles) and SV40P (squares) promoters, calculated as the median count of any tags representing a tile, divided by the median ratio for all tags in the library, plotted against genomic coordinates (genome build hg19). Only tiles with p <0.05 for both biological replicates are shown. Top-ranked tiles are numbered in decreasing order of confidence. The most confident hit (1) is in red type and the region surrounding it highlighted with a red box. (Bottom) Schematic representing the location of the most confident hits from the MPRA relative to *HLA-G*, together with a negative control region (Neg).

The four most confident hits, indicated in Figure 2.1.1, were then carried on for further analysis using classical luciferase reporter gene assays. The most confident candidate, located 12 kb upstream of the *HLA-G* gene, was the only tile with enhancer activity greater than 2 with both promoters tested, displaying the highest enhancer activity with minP (8.4) and second highest enhancer activity with SV40P (12.4) overall. This region specifically enhanced firefly luciferase activity upstream of the minimal promoter by 20-fold in HLA-G$^+$ JEG3 cells (Figure 2.1.2).

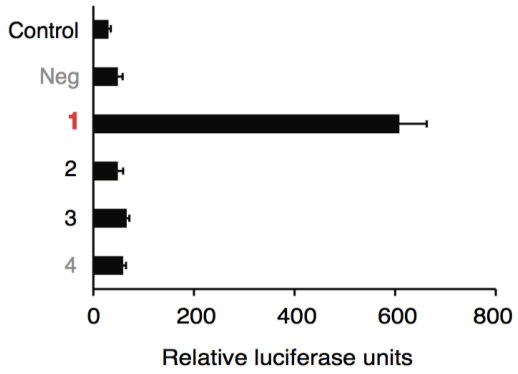


Figure 2.1.2. Enhancer activity of MPRA candidates in JEG3 cells (HLA-G$^+$). Luciferase reporter gene activity in combination with the minP promoter. The most confident candidate is marked in red. Error bars represent SEM of replicates of a representative experiment (n = 2). Control, empty vector; Neg, negative control region.

We named this novel putative regulatory element *Enhancer L*, for being a *long*-range enhancer discovered with our unbiased enhancer screen. Importantly, *Enhancer L* was not active in HEK293T cells, an HLA-G negative control cell line (Figure 2.1.3). Moreover, this cell type-specific activity pattern was maintained even when *Enhancer L* was cloned in an inverted orientation (Figure 2.1.3), a classical hallmark of an enhancer element (Banerji et al., 1981). Of note, candidates numbers 3 and 4 from our MPRA

screen, located near or even partially overlapping with *Enhancer L*, respectively, displayed negligible activity in JEG3 cells (Figure 2.1.2). Altogether, these observations suggest that *Enhancer L* corresponds to a narrowly defined regulatory region in the *HLA-G* locus that may confer tissue-specific HLA-G expression to trophoblasts.

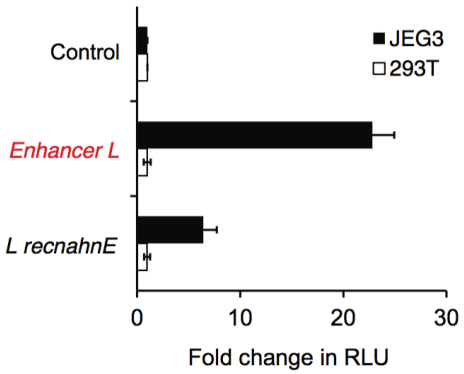


Figure 2.1.3. *Enhancer L* is a trophoblast-specific enhancer upstream of *HLA-G*.

Enhancer L remains active specifically in JEG3 cells when its direction is inverted. Control, empty vector; "*L recnahnE*", inverted *Enhancer L*; RLU, Relative Luciferase Units. Error bars represent SEM of three independent experiments.

2.2. *Enhancer L* is essential for HLA-G expression in JEG3 cells

Next, we sought to investigate whether *Enhancer L* modulates endogenous HLA-G expression. Not long ago, deleting precise DNA regions in mammalian cells was an extremely long and laborious process, requiring two rounds of targeting: one to insert recombination sites (e.g. loxP or FRT) on the same chromosome via homologous recombination (HR), followed by treatment with a recombinase (e.g. Cre or Flp, respectively) to delete the sequence flanked by the recombination sites. This strategy was plagued by very low efficiencies (HR occurs at a frequency of one in 10^4 to 10^7

cells), and was mostly limited to mouse embryonic stem cells, which are more amenable to genome engineering via HR and can be clonally expanded and injected into blastocysts to give rise to whole animals. Human cells were virtually out-of-reach (Carroll, 2011; Lee et al., 2010). Those limitations were overcome with the advent of genome editing. Celebrated as a transformative technology that allows precise modifications in the human genome, genome editing is based on engineered DNA endonucleases that can be programmed to recognize and bind to any desired site in the genome. Upon binding to the target sequence, a double strand break (DSB) is introduced, which can be repaired in one of the following two ways. Non-homologous end joining (NHEJ) is an imperfect mechanism that often leads to small insertions or deletions (InDels). In contrast, HR relies on a closely matching DNA sequence to accurately repair the DSB. The past decade has witnessed the development of several generations of engineered nucleases. Meganucleases, also known as homing endonucleases, are naturally occurring DNA endonucleases characterized by an unusually long restriction site (14-40 bp) and the capacity to induce HR in mammalian cells (Arnould et al., 2011; Rouet et al., 1994). They represent the first class of endonucleases to be successfully engineered to recognize desired DNA sequences (Arnould et al., 2011; Seligman et al., 2002). However, targeting novel sequences often resulted in very low efficiency, prompting their replacement with zinc finger nucleases (ZFNs) and, later, transcription activator-like effector nucleases (TALENs). TALENs consist of two domains: a customizable DNA-binding domain consisting of several modules, which can be linked in tandem to recognize any DNA sequence with high specificity, and a DNA-cleaving domain derived from the FokI nuclease (Ding et al., 2013a; Hockemeyer et al., 2011; Urnov et al., 2005). Yet, the most significant innovation

in the field arose with the advent of clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 technology.

The CRISPR/Cas9 system has quickly become the tool of choice to introduce targeted mutations in the genome due to its unprecedented editing efficiency and design simplicity. It consists of an endonuclease, Cas9, which introduces DSBs in any genomic sequence defined by homology with a co-expressed guide RNA (gRNA) (Hsu et al., 2014). Due to its high efficiency in generating biallelic deletions and ease of multiplexing, CRISPR/Cas9 is particularly well suited to perturb noncoding regions in the genome.

In order to directly target *Enhancer L* in JEG3 cells, we used a CRISPR/Cas9 dual guide approach (Mandal et al., 2014; Meissner et al., 2014) by targeting two gRNAs to sites flanking *Enhancer L* (Figure 2.2.1). We employed a *Streptococcus pyogenes* Cas9 linked via a self-cleaving 2A peptide to a green fluorescent protein (GFP) to facilitate identification of Cas9-expressing cells. GFP^+ cells were sorted and plated at clonal density and the emerging single cell-derived colonies were transferred 10 days post-plating into 96-well plates. PCR analysis of CRISPR/Cas9 targeted single cell-derived clones was used to identify homozygous *Enhancer L* KO clones (Figure 2.2.1).

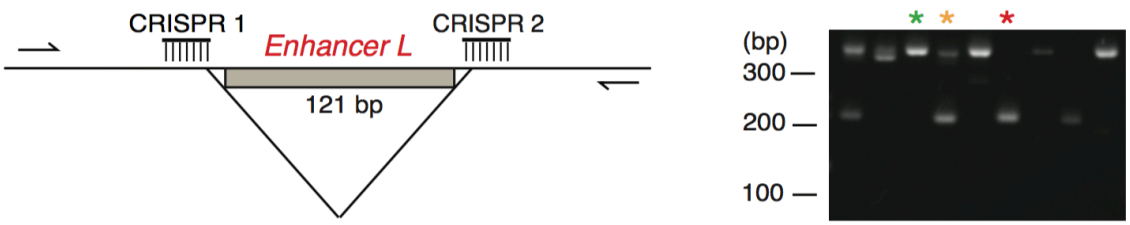


Figure 2.2.1. Dual CRISPR guide strategy to delete *Enhancer L*. (Left) Arrows represent the primers used for PCR screening. (Right) PCR screening of CRISPR/Cas9-

(Continued) targeted JEG3 single cell-derived clones. Green*, wild-type; yellow*, heterozygote; red*, null clone.

We observed a clonal targeting efficiency of 29.5%, with homozygous deletions occurring at a frequency of 8.7%. Four independent *Enhancer L* null clones and three WT clones were selected for further characterization. As expected, Sanger sequencing demonstrated excision of the DNA between the predicted Cas9 cleavage sites (3 bases 5' of the PAM sequence), with three out of four clones having the same exact deletion of 154 bp (Figure 2.2.2).

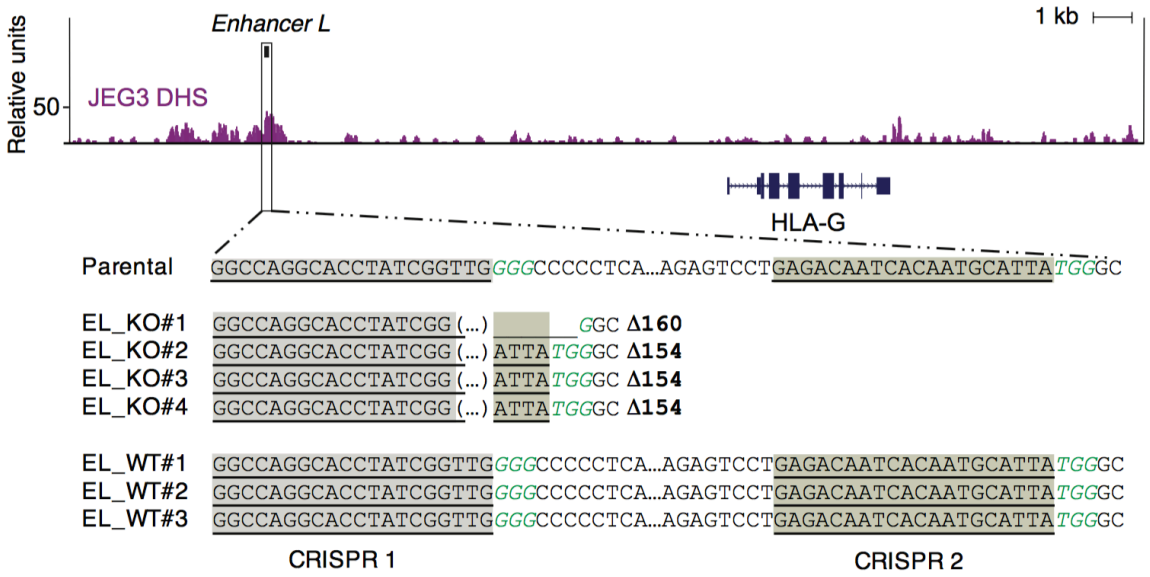


Figure 2.2.2. Successful deletion of *Enhancer L* in JEG3 cells.

Sanger sequencing of four independent homozygous *Enhancer L* KO clones and three independent WT clones resulting from CRISPR/Cas9 targeting of *Enhancer L* (black box) in JEG3 cells using a dual CRISPR guide RNA approach. Binding sites for the gRNAs targeting *Enhancer L* are underlined and shaded. PAM motifs are italicized in green type.

(Continued) *Enhancer L* is part of a DNase I hypersensitive site (DHS) in JEG3 cells, as determined by genome-wide DNase-seq. EL, *Enhancer L*.

Strikingly, deletion of *Enhancer L* resulted in complete ablation of HLA-G expression, as determined by flow cytometry and quantitative Real Time PCR (qRT-PCR) (Figure 2.2.3). Surveying the whole genome for chromatin accessibility using genome-wide DNase-seq revealed that *Enhancer L* is located within a DNase I hypersensitivity site (DHS) in JEG3 cells (Figure 2.2.2), supporting the hypothesis that *Enhancer L* is indeed an active regulatory element in its endogenous chromatin context.

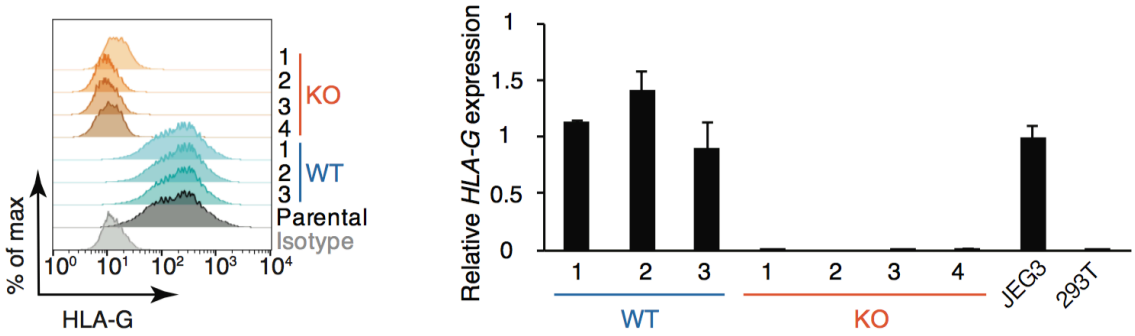


Figure 2.2.3. *Enhancer L* is required for HLA-G expression in JEG3 cells.

(Left) Combined FACS histogram demonstrating complete ablation of HLA-G surface expression in *Enhancer L* KO JEG3 clones. (Right) *HLA-G* transcript levels of *Enhancer L* KO clones, with JEG3 cells and HEK293T cells as controls. Error bars represent SEM of replicates of a representative experiment (n = 2).

2.3. Deletion of *Enhancer L* in JEG3 cells uniquely ablates HLA-G expression

Following our observation that *Enhancer L* is required for HLA-G expression, we then asked if *Enhancer L* acts specifically on HLA-G. Previous studies have identified enhancers that affect multiple genes spanning regions of hundreds of kb (Link et al., 2013; Melo et al., 2013). To investigate whether *Enhancer L* also regulates other genes in the HLA locus or elsewhere on chromosome 6, we sequenced polyA$^+$ mRNA from three *Enhancer L* KO JEG3 clones, as well as three WT clones and two independent samples of the parental JEG3 cell line as controls. RNA-seq confirmed that *HLA-G* is completely ablated across all KO clones (Figure 2.3.1).

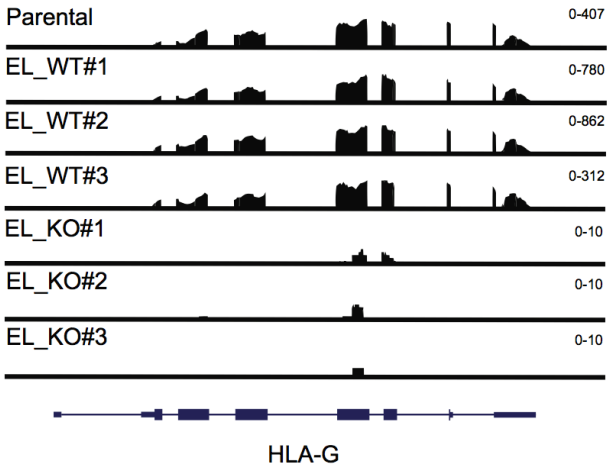


Figure 2.3.1. RNA-seq confirms complete loss of *HLA-G* transcript in three independent *Enhancer L* KO JEG3 clones. Scale in FPKM (Fragments Per Kilobase of exon per Million fragments mapped) is indicated on the top right corner of each sample. EL, *Enhancer L*.

Of note, HLA-G is the only completely ablated gene within 2 Mb of *Enhancer L* (Figures 2.3.2 and 2.3.3). Within this region, *MICB* is the only other gene with moderate expression (>10 FPKM in at least one condition) showing a more than 2-fold difference in expression levels. *MICB* is known to be induced upon genotoxic stress (Gasser et al., 2005) and is thus likely to be upregulated during the process of CRISPR targeting in a nonspecific manner, rather than as a result of *Enhancer L* deletion. The lack of other significant expression changes in the vicinity (+/- 2 Mb) of *Enhancer L* suggests that *HLA-G* is the only direct *cis* target of *Enhancer L*.

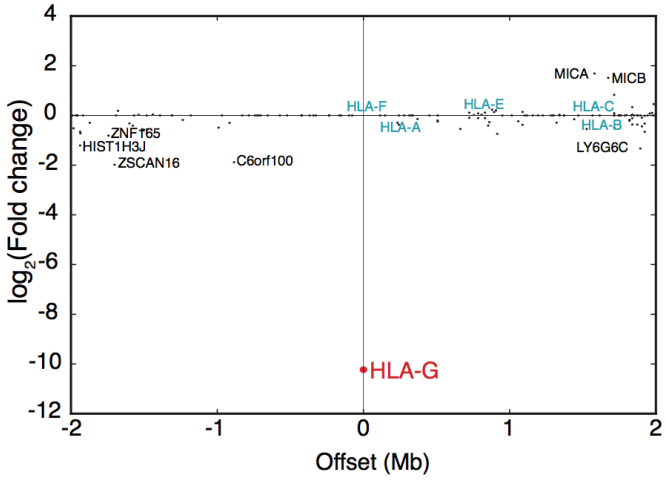


Figure 2.3.2. *Enhancer L* deletion specifically results in loss of *HLA-G* expression in a radius of 2 Mb centered on *HLA-G*.

The fold change in gene expression between combined WT and KO clones is plotted against genomic coordinates, each dot representing a gene. HLA-G is displayed in red (largest fold-change), while other HLA genes are displayed in blue (no change).

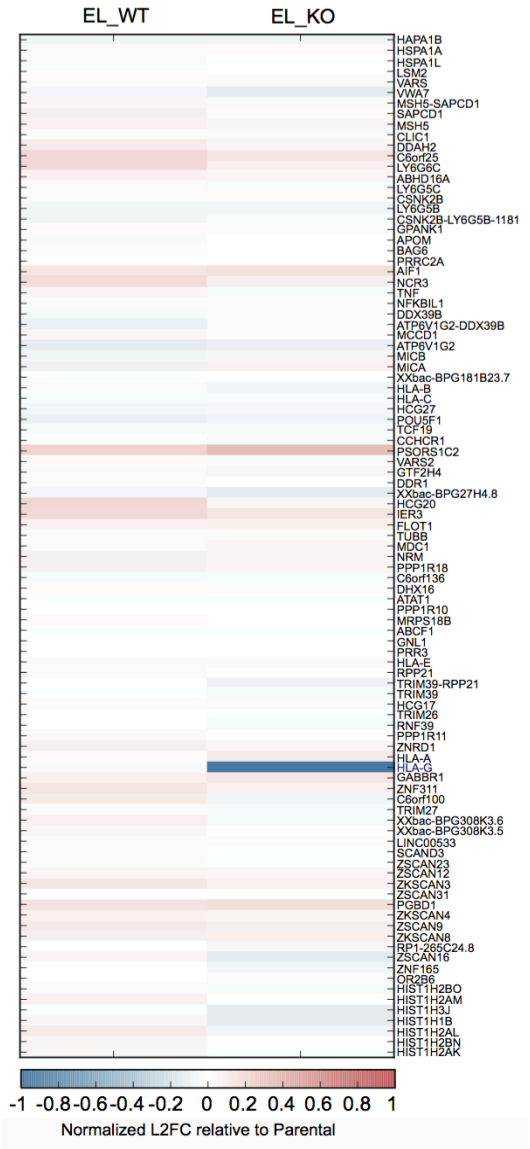


Figure 2.3.3. Heat map illustrating gene expression changes between *Enhancer L* WT and KO clones across a 2 Mb radius centered on *HLA-G*.

Ablation of *HLA-G* expression (in blue) is the most significant change observed. L2FC, logarithm base 2 of fold change in gene expression between the averages of three WT or three KO clones and the average of parental JEG3 cells. EL, *Enhancer L*; Red, gene upregulation; blue, gene downregulation; white, no change in gene expression in comparison to parental JEG3 cells.

Looking beyond chromosome 6, transcriptome-wide analysis revealed statistically significant differences in the expression of 321 genes using Cuffdiff (FDR<0.05). To rule out the possibility that these changes were caused by CRISPR/Cas9-induced off-target effects, we performed *in silico* off-target analyses of our *Enhancer L* gRNAs using the CRISPR design tool at crispr.mit.edu (Hsu et al., 2013). The top 50 predicted off-target sites yielded maximum scores of 3.3 for gRNA 1, and 0.9 for gRNA 2 (out of 100), suggesting that the observed global changes in gene expression are not likely to be a result of off-target cleavage at these sites. Gene set enrichment analysis (GSEA) of the most differentially expressed genes revealed statistically significant enrichment (FDR<0.05) for six gene sets, all of which are related to steroid hormone biosynthesis and GPCR signaling, processes expected to play a role in trophoblast physiology. Pair-wise comparison of all three experimental groups (WT, KO, parental JEG3), however, revealed that, in spite of the observed transcriptome-wide changes in gene expression, *HLA-G* was by far the most downregulated gene upon *Enhancer L* deletion at the whole transcriptome level (Figure 2.3.4), indicating that *Enhancer L* uniquely modulates HLA-G expression.

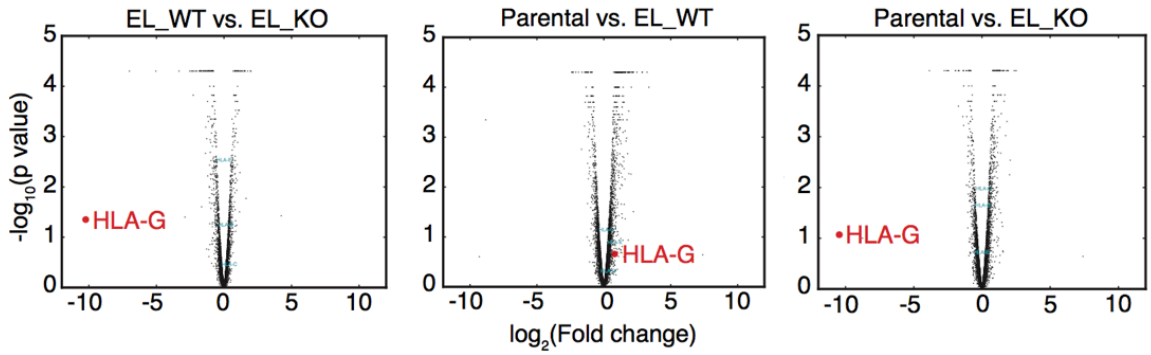


Figure 2.3.4. Deletion of *Enhancer L* specifically abrogates *HLA-G* expression. Transcriptome-wide pair-wise comparison between different genotypes, depicted as Volcano plots. X-axes represent log2 fold change in gene expression.

2.4. *Enhancer L* is required for HLA-G expression in primary extravillous trophoblasts

In order to confirm the role of *Enhancer L* in primary human trophoblasts, we obtained villi from first trimester human placental tissue and purified HLA-G$^+$ EVT by flow cytometry (Tilburgs et al., 2015a). Cas9-2A-GFP and gRNAs targeting *Enhancer L* were successfully co-delivered into primary EVT using lentiviral particles, as assessed by GFP expression (Figure 2.4.1). As expected, *Enhancer L* deletion resulted in a significant decrease in *HLA-G* mRNA levels (74.12% ± 13.61 SEM, n = 3) (Figure 2.4.1).

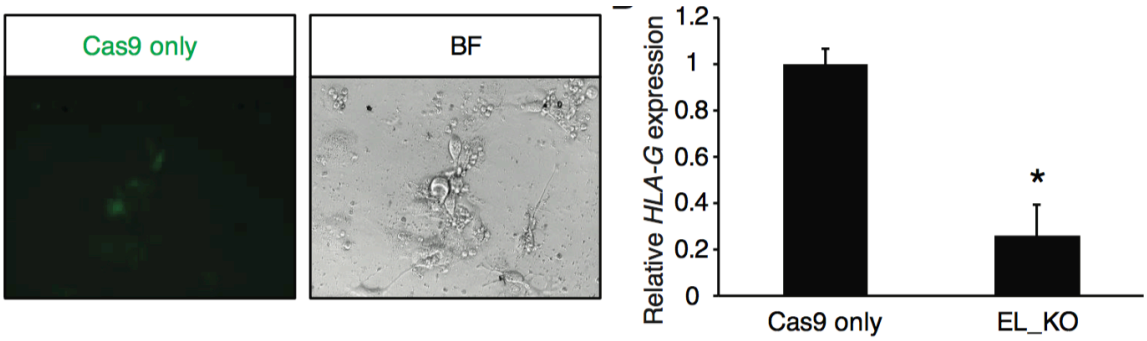


Figure 2.4.1. Figure 4. *Enhancer L* is necessary for optimal HLA-G transcription in primary EVT. (Left) Transduction of first trimester HLA-G$^+$ extravillous trophoblasts (EVT) with lentiviral Cas9 and *Enhancer L* gRNAs, assessed based on 2A peptide-linked GFP expression. BF, bright field (40X magnification). (Right) Reduction of *HLA-G* mRNA expression following *Enhancer L* deletion. Bars represent average ± SEM of three independent experiments. *, $p<0.05$, paired student's t-test. EL, *Enhancer L.*

Loss of HLA-G surface expression as a result of lentiviral CRISPR/Cas9-mediated ablation of *Enhancer L* was first evaluated in JEG3 cells, which divide rapidly in culture.

We observed complete loss of HLA-G surface expression one week post-transduction in a large percentage of transduced cells (61.9% ± 1.93 SEM, n = 3) (Figure 2.4.2).

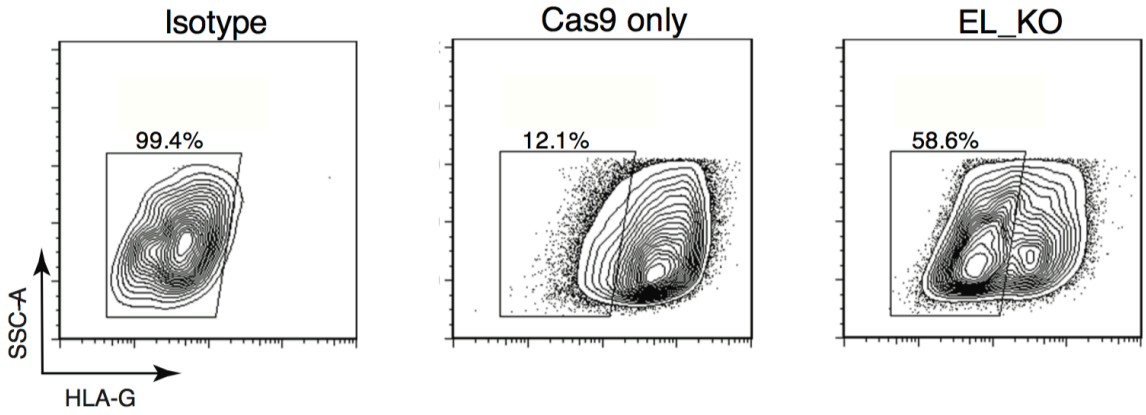


Figure 2.4.2. Lentiviral CRISPR/Cas9 deletion of *Enhancer L* in JEG3 cells.

Lentiviral transduction of JEG3 cells with Cas9-2A-GFP and two CRISPR guide RNAs targeting *Enhancer L* led to complete loss of HLA-G surface expression (n = 3). Cells successfully transduced with lentiviral particles containing the guide RNAs were selected with puromycin and analyzed one week post-transduction. EL, *Enhancer L*.

Detecting changes in HLA-G surface expression in primary EVT, however, is hampered by the unusually long half-life of HLA-G protein on the cell membrane (Davis et al., 1997), and the fact that primary EVT can only be cultured *ex vivo* for a short period of time (< 5 days). Despite these technical limitations, we were able to detect a significant reduction in HLA-G surface expression 5 days after targeting *Enhancer L* in primary EVT (60.71% ± 10.68 SEM, n = 3) (Figure 2.4.3).

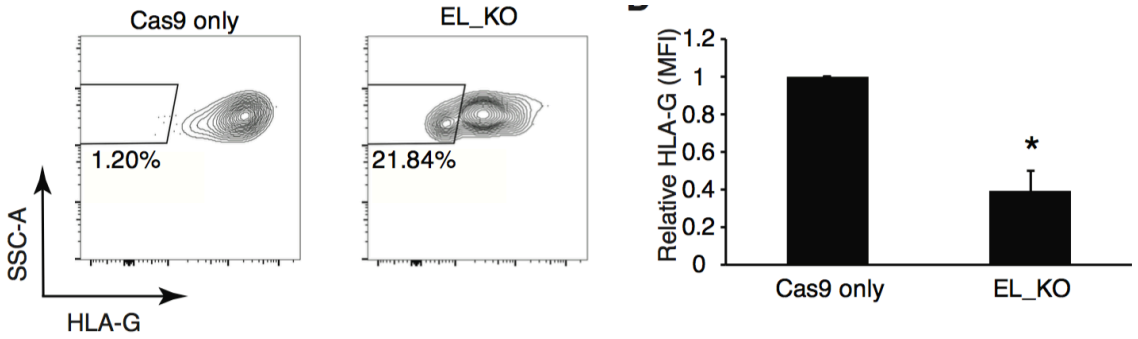


Figure 2.4.3. *Enhancer L* is necessary for HLA-G surface expression in EVT.

Enhancer L deletion leads to significant reduction in HLA-G surface expression in primary EVT. One representative experiment is shown (n = 3). **(D)** Significant reduction in HLA-G surface expression upon *Enhancer L* deletion in primary EVT. Bars represent average ± SEM of three independent experiments. *, p<0.05, paired student's t-test. MFI, Mean Fluorescence Intensity.

Successful genomic deletion of *Enhancer L* was confirmed by PCR sequencing (Figure 2.4.4). Altogether, our results demonstrate that *Enhancer L* is necessary for HLA-G expression in primary human EVT.

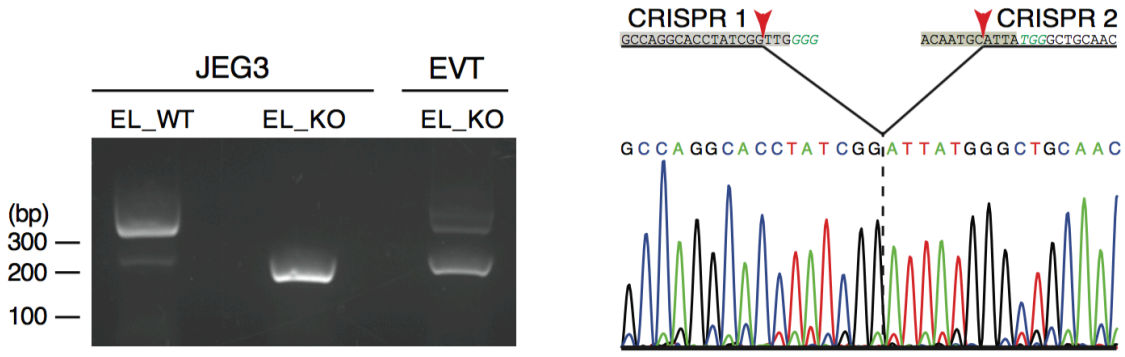


Figure 2.4.4. Confirmation of successful *Enhancer L* deletion in primary EVT.

(Left) PCR demonstrating successful *Enhancer L* genomic deletion in EVT using lentivirally delivered CRISPR/Cas9. An *Enhancer L* WT JEG3 cell clone (EL_WT) and a

(Continued) KO clone (EL_KO) were included as controls. (Right) Sanger sequencing confirmation of Cas9-mediated genomic deletion of *Enhancer L* in primary EVT. CRISPR gRNA binding sites are shaded, PAM motifs italicized in green type, and Cas9 cutting sites indicated with red arrowheads.

CHAPTER 3: Mechanistic characterization of trophoblast-specific HLA-G expression

3.1. *Enhancer L* is a distant regulatory element that loops into the HLA-G proximal promoter

Next, we aimed to characterize the mechanism by which *Enhancer L* activates HLA-G expression at a distance. The current model of long-range gene regulation postulates that remote *cis*-regulatory elements come into close proximity to the promoters of the genes they regulate via chromatin looping (Sexton and Cavalli, 2015). To test for the involvement of looping in *Enhancer L–HLA-G* promoter long-range communication, we carried out Chromatin Conformation Capture (3C) assays in JEG3 and HLA-G negative HEK293T cells (Figure 3.1.1).

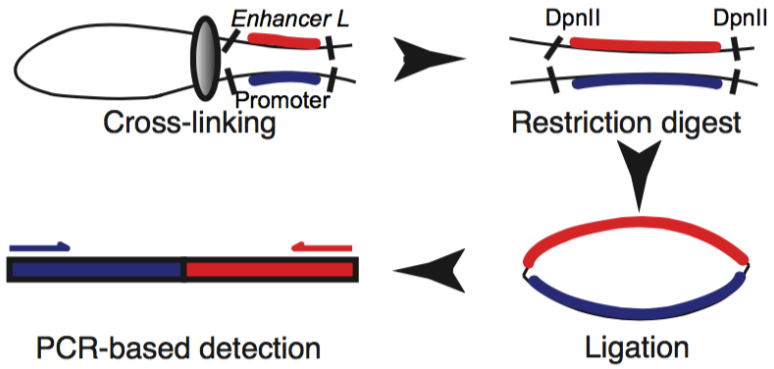


Figure 3.1.1. Schematic representing the main steps in a 3C experiment.

These steps are: chromatin cross-linking, restriction digest, ligation at low DNA concentrations, and PCR-based detection of looping interactions. Our regions of interest (*Enhancer L* and the *HLA-G* promoter) and chosen restriction enzyme (DpnII) are used as examples.

We outlined a 3C strategy focused on these two regulatory regions by designing primers pointing towards DpnII restriction sites that flanked *Enhancer L* and the classical promoter of *HLA-G*, as depicted in Figure 3.1.2.

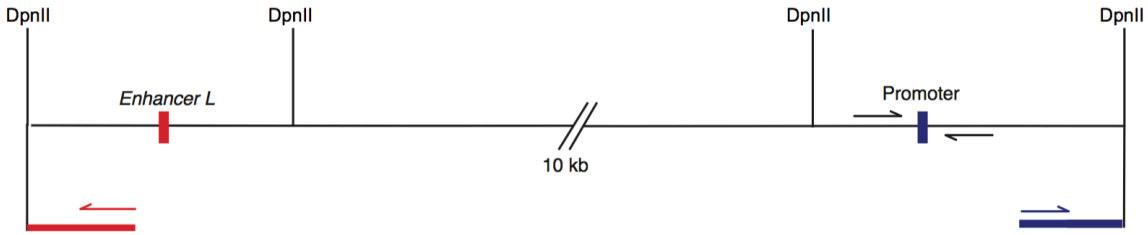


Figure 3.1.2. Strategy for 3C analysis of the *HLA-G* locus.

DpnII restriction sites flank the two regions of interest: *Enhancer L*, labeled in red, and the *HLA-G* classical promoter, given in blue. Primers in black amplify a product that serves as "loading control", and primers in red and in blue were used to detect loop formation.

Following chromatin cross-linking, DpnII digestion and ligation of interacting DNA segments at dilute concentrations, we detected a looping interaction between *Enhancer L* and the classical promoter of *HLA-G* specifically in JEG3 cells (Figure 3.1.3).

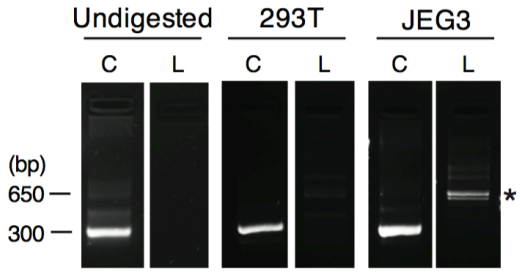


Figure 3.1.3. *Enhancer L* loops into the classical promoter of *HLA-G*.

Enhancer L physically interacts with the classical promoter of *HLA-G* in JEG3 cells, but not in HLA-G negative HEK293T cells. C, Loading Control PCR product (322 bp); L,

(Continued) *Enhancer L*-classical promoter Looping interaction PCR product (640 bp), marked with a star.

We confirmed the nature of the resulting hybrid DNA molecule consisting of *Enhancer L* and the proximal promoter by TOPO cloning and Sanger sequencing of the PCR product (Figure 3.1.4).

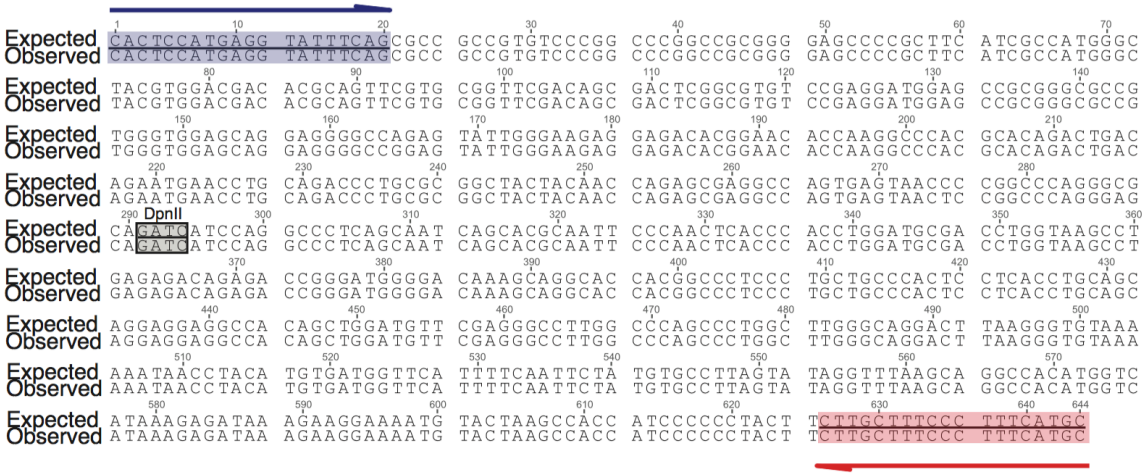


Figure 3.1.4. Sequence confirmation of the physical interaction between *Enhancer L* and the classical promoter of *HLA-G* detected by 3C.

PCR amplicon shown in Figure 5B. Primer pointing away from the classical promoter is depicted in blue, primer pointing away from *Enhancer L* in red, and the DpnII restriction site where ligation of the contact regions occurred is enclosed by a gray-shaded box.

Of note, this looping interaction was absent in HEK293T cells (Figure 3.1.3), in agreement with the lack of *Enhancer L* activity in these cells (Figure 2.1.3).

3.2. MPRA-based scanning mutagenesis reveals motifs controlling *Enhancer L* activity

Having established *Enhancer L* as a *bona fide* enhancer upstream of *HLA-G*, we sought to identify the transcriptional regulators that mediate its action. To our surprise, truncation of *Enhancer L* invariably led to loss of enhancer activity in luciferase reporter gene assays, suggesting multiple active motifs spread across its length (Figure 3.2.1).

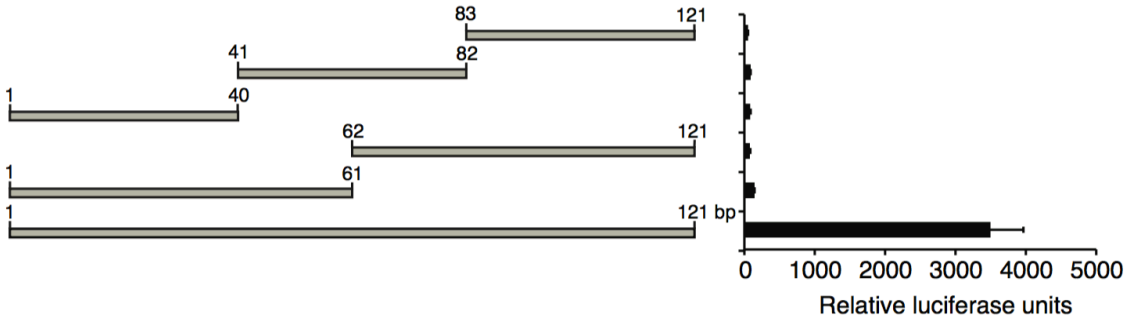


Figure 3.2.1. Systematic truncation of *Enhancer L* invariably results in loss of reporter gene activity.

Luciferase reporter gene assays performed in JEG3 cells with constructs containing a strong SV40 promoter (n = 2).

To fine map the active regulatory motifs responsible for *Enhancer L* activity, we carried out an MPRA-based scanning mutagenesis at the single base pair resolution (Melnikov et al., 2012). In brief, we generated a total of 12,000 *Enhancer L* variants, representing all possible single substitutions, as well as small insertions or deletions at all positions. To reduce experimental noise, each variant was coupled to 16 tags on average, for a total of 200,000 distinct variant-tag combinations. As before, this complex library was co-transfected into JEG3 cells, followed by RNA harvesting and sequencing analysis. This

fine mapping of *Enhancer L* led to the identification of five putative regulatory motifs, consistent across both promoters tested (SV40P and minP). Subsequent *in silico* analysis using the TRANSFAC database (Matys et al., 2006) predicted binding of CEBP and GATA family transcription factors within these five motifs (Figures 3.2.2 and 3.2.3).

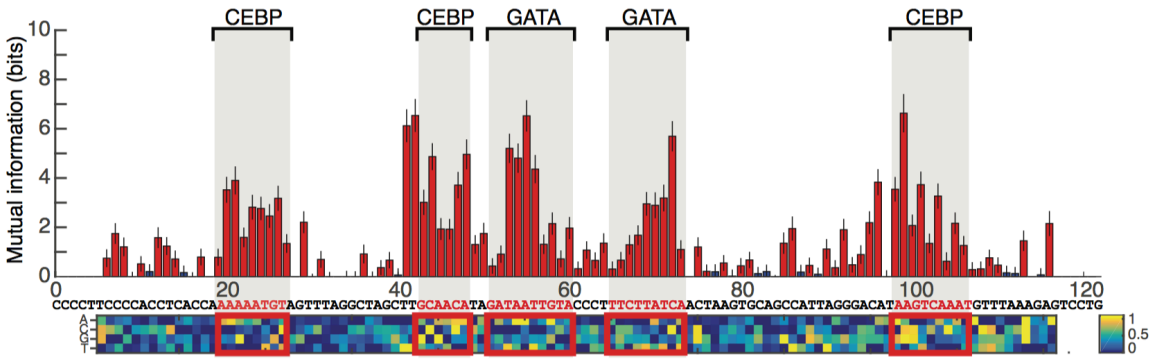


Figure 3.2.2. Identification of five putative regulatory motifs required for *Enhancer L* activity using MPRA-based scanning mutagenesis with an SV40 promoter.

Red bars indicate a significant change from original *Enhancer L* activity (Mann-Whitney U-test, 5% FDR); blue bars, not significant. The matrix represents the estimated additive contribution of each nucleotide to *Enhancer L* activity. Transcription factor binding site prediction was performed using the TRANSFAC database.

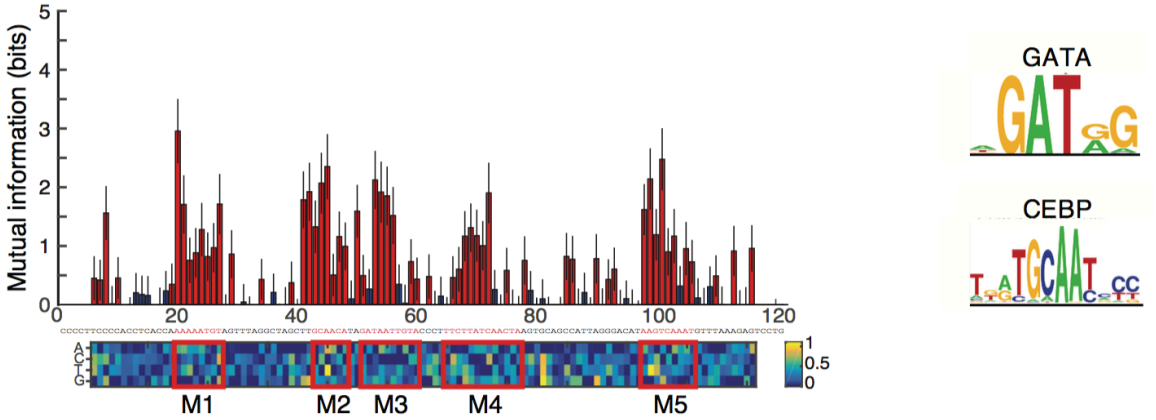


Figure 3.2.3. Identification of five putative regulatory motifs involved in *Enhancer L* activity using MPRA-based scanning mutagenesis with a minP promoter.

Red indicates a significant change from the original *Enhancer L* sequence (Mann-Whitney U-test, 5% False Discovery Rate – FDR); blue means not significant. The matrix represents the estimated additive contribution of each nucleotide to *Enhancer L* activity. Putative motifs (in red boxes) are numbered M1 through M5. CEBP and GATA consensus motifs according to the TRANSFAC database are shown on the right.

Reporter gene assays with truncated versions of *Enhancer L* lacking each one of these motifs (M1 through M5) showed that each one of them is essential for optimal *Enhancer L* activity in JEG3 cells (Figure 3.2.4).

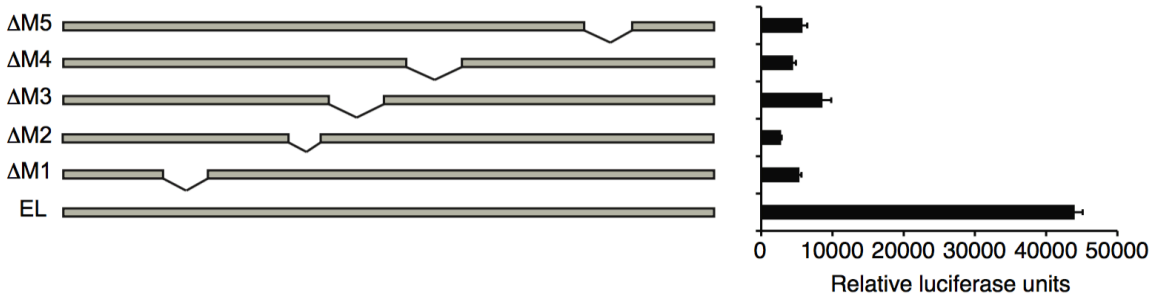


Figure 3.2.4. Systematic deletion of the five putative motifs within *Enhancer L*.

(Continued) Each motif (M1-M5) was required for maximum *Enhancer L* activity in JEG3 cells, as measured by luciferase reporter gene activity (strong SV40 promoter) (n = 2). EL, *Enhancer L*.

3.3. CEBP and GATA factors regulate trophoblast-specific HLA-G expression

Motif sequence analysis alone does not allow discrimination between different members of transcription factor families. We reasoned that the transcription factors controlling HLA-G expression via *Enhancer L* must be highly expressed specifically in HLA-G$^+$ trophoblasts. Microarray analysis of primary cells isolated from human placental tissue, and JEG3 cells, (Tilburgs et al., 2015a) revealed that *CEBPA*, *CEBPB*, *GATA2*, and *GATA3* are the most highly expressed genes within their respective transcription factor families (Figure 3.3.1).

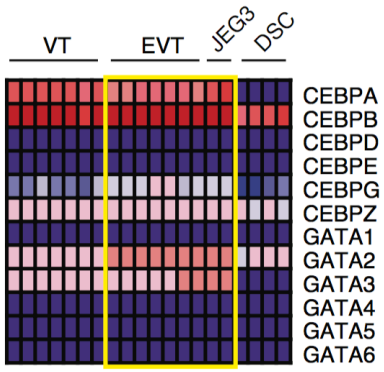


Figure 3.3.1. Expression levels of genes belonging to the CEBP and GATA transcription factor families in primary trophoblasts and JEG3 cells.

Heat map generated with published microarray data using *GenePattern*, with dark blue representing lowest expression and dark red highest expression. VT, villous trophoblasts; EVT, extravillous trophoblasts; DSC, decidual stromal cells.

Our whole transcriptome RNA-seq analysis, confirmed high expression levels of *CEBPB*, *GATA2*, and *GATA3* in JEG3 cells (Figure 3.3.2). In addition, a survey of publicly available gene expression profiles (*BioGPS*) revealed that these three transcription factors are highly co-expressed in human placenta, and also more restricted in expression to this tissue than any other CEBP or GATA transcription factor family member. Importantly, CEBPβ, GATA2, and GATA3 have been implicated in murine placental development and trophoblast-specific gene regulation (Begay et al., 2004; Cheng and Handwerger, 2005; Ma and Linzer, 2000), making them strong candidates for transcriptional regulators of HLA-G expression in human trophoblasts.

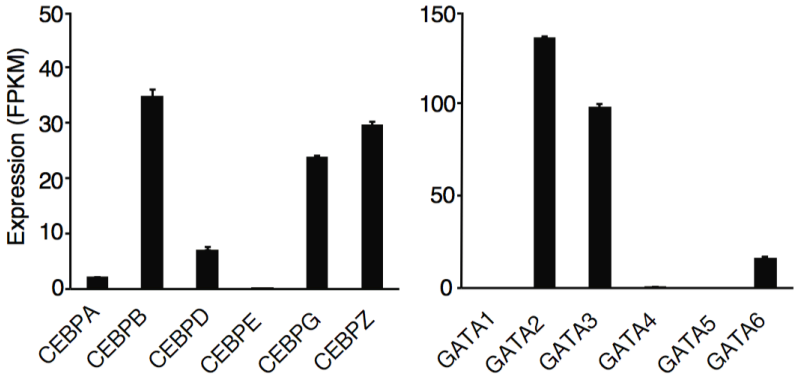


Figure 3.3.2. CEBP and GATA gene expression levels in JEG3 cells, as determined by whole-transcriptome RNA-seq.

FPKM, Fragments Per Kilobase of exon per Million fragments mapped.

In order to test our prediction, we sought to determine whether CEBPβ, GATA2, and GATA3 bind to *Enhancer L*. Indeed, chromatin immunoprecipitation (ChIP) using validated ChIP-grade antibodies, followed by qPCR analysis (ChIP-qPCR), revealed a 40-fold enrichment for CEBPβ on *Enhancer L.* Similarly, a significant enrichment for GATA2 and GATA3 (5-fold) was detected on *Enhancer L*, indicating that, in JEG3 cells, endogenous CEBPβ, GATA2, and GATA3 associate with *Enhancer L* (Figure 3.3.3). In addition, all three factors were found to bind to the proximal promoter of *HLA-G* (Figure 3.3.3), providing further evidence for the existence of a chromatin loop between *Enhancer L* and the core promoter of *HLA-G*, possibly established by GATA2 and GATA3 (Chen et al., 2012; Deng et al., 2012).

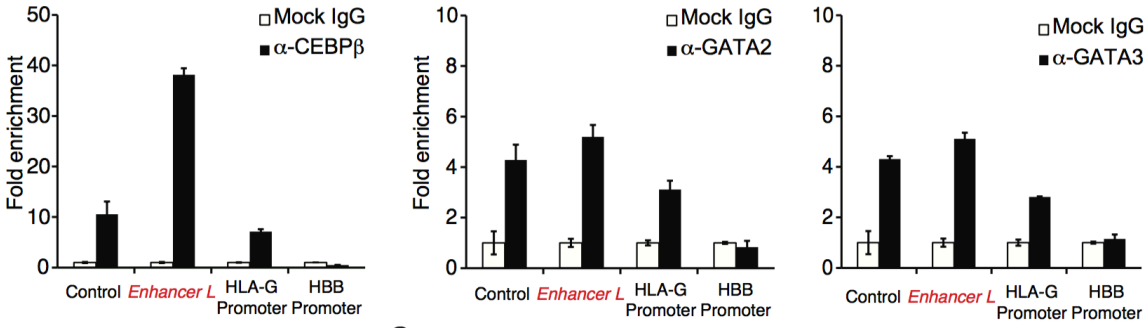


Figure 3.3.3. CEBPβ, GATA2, and GATA3 associate with *Enhancer L* and with the classical promoter of *HLA-G*.

Assessed by ChIP-qPCR (n = 2). Control, positive control region predicted to be bound by the respective transcription factor according to *ENCODE* data; *HBB* promoter, negative control.

Of note, RNA Polymerase II (Pol II) associated with both *Enhancer L* and the *HLA-G* core promoter (Figure 3.3.4), suggesting that active transcription is involved in the formation of this long-range chromatin loop.

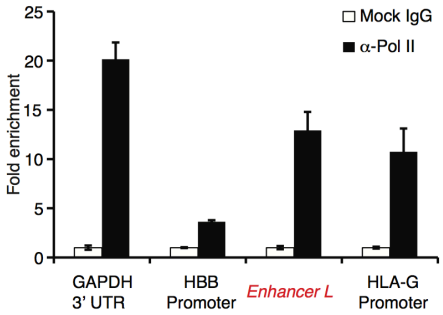


Figure 3.3.4. RNA Polymerase II associates with *Enhancer L* and with the classical promoter of *HLA-G*.

Assessed by ChIP-qPCR (n = 2). *GAPDH* 3'UTR was used as a positive control, and the *HBB* promoter as a negative control for Pol II binding.

Consistent with a role in *HLA-G* transcriptional activation, transient overexpression of CEBPβ, GATA2, and GATA3 individually in JEG3 cells led to an up to 8-fold increase in *HLA-G* expression, indicating that these three factors are transcriptional activators of *HLA-G* expression (Figure 3.3.5).

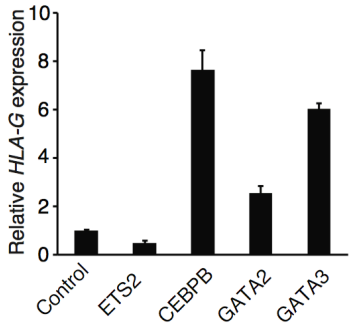


Figure 3.3.5. Trophoblast CEBP and GATA factors regulate HLA-G expression.

Ectopic expression of CEBPβ, GATA2, or GATA3 upregulate HLA-G expression in JEG3 cells, as measured by qPCR. ETS2, a transcription factor expressed in trophoblasts, was used as a negative control. Control, empty vector. Error bars represent SEM of replicates of a representative experiment (n = 2).

Taken together, our data supports a model where CEBPβ and GATA2/3 mediate long-range chromatin interactions between *Enhancer L* and the classical promoter of *HLA-G* (Figure 3.3.6), driving HLA-G expression specifically in extravillous trophoblasts at the maternal-fetal interface.

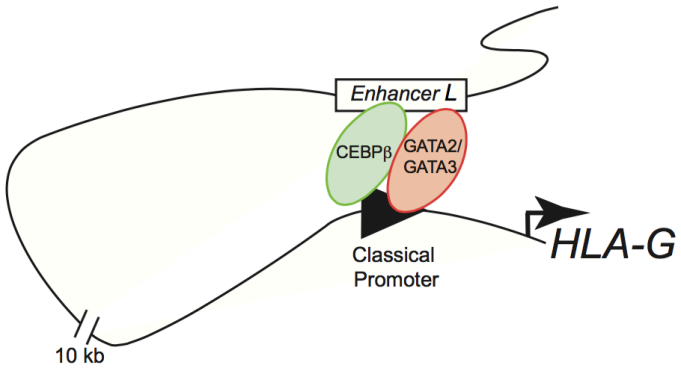


Figure 3.3.6. Proposed model of trophoblast-specific HLA-G transcriptional regulation by CEBPβ, GATA2, and GATA3 via *Enhancer L*.

CHAPTER 4: Conclusions and future perspectives

Noncoding regulatory elements are increasingly recognized as important determinants of development and disease. Altered enhancer function has been implicated in multiple syndromes, termed “enhanceropathies” (Smith and Shilatifard, 2014), indicating that targeting enhancers may open additional therapeutic avenues for monogenic diseases. For instance, abrogation of BCL11A expression by means of deleting an enhancer mapped by GWAS significantly increases fetal globin levels, an important therapeutic option for patients with hemoglobin disorders (Bauer et al., 2013).

GWAS have uncovered an astonishing number of disease-associated noncoding loci (Maurano et al., 2012), posing a challenge to functionally validate and characterize putative regulatory elements. Massively Parallel Reporter Assay (MPRA) represents an unbiased high-throughput method for *de novo* discovery and validation of *cis*-regulatory regions. In this study, the most confident candidate from our MPRA screen, located 12 kb upstream of *HLA-G,* was found to be active specifically in the HLA-G^{+} JEG3 choriocarcinoma cell line (Figure 2.1.3), suggesting that it may be involved in tissue-specific *HLA-G* transcriptional regulation. Indeed, CRISPR/Cas9 genome editing revealed that this novel enhancer, *Enhancer L*, is essential for trophoblast expression of HLA-G (Figures 2.2.3 and 2.4.3).

Previous studies established that MHC gene expression is mainly controlled at the level of a conserved proximal promoter. Upon interaction with a transcriptional activator – CIITA for class II and NLRC5 for class I genes – a multiprotein transcription factor

complex is assembled, forming the MHC enhanceosome (Kobayashi and van den Elsen, 2012; Meissner et al., 2010; Steimle et al., 1994). Even though the enhanceosome is essential for basal and induced expression of MHC class I genes, its relevance in trophoblasts is uncertain: EVT do not express NLRC5 or CIITA (Tilburgs et al., 2015a) and the *HLA-G* proximal promoter harbors several non-functional motifs (Solier et al., 2001), suggesting that tissue-specific HLA-G expression is mediated by a distinct mechanism. While several studies have described *cis*-regulatory regions involved in *HLA-G* transcriptional regulation (Gobin et al., 2002; Ikeno et al., 2012; Moreau et al., 2009), the present study is the first to report a noncoding sequence, *Enhancer L*, absolutely required for the tissue-specific expression of HLA-G in trophoblasts.

Interestingly, *Enhancer L* is contained within a long terminal repeat (LTR) sequence, LTR7 (Kelley and Rinn, 2012), associated with a human endogenous retroviral element (ERV), ERV1, as indicated in Figure 4.1. LTR sequences have been co-opted by mammalian genomes as regulatory elements, especially in the placenta (Chuong et al., 2013). Well-known examples include the placenta-specific promoter of *CYP19* (van de Lagemaat et al., 2003) and MER20, regulatory sequences found upstream of progesterone-responsive genes essential for decidualization (Lynch et al., 2011).

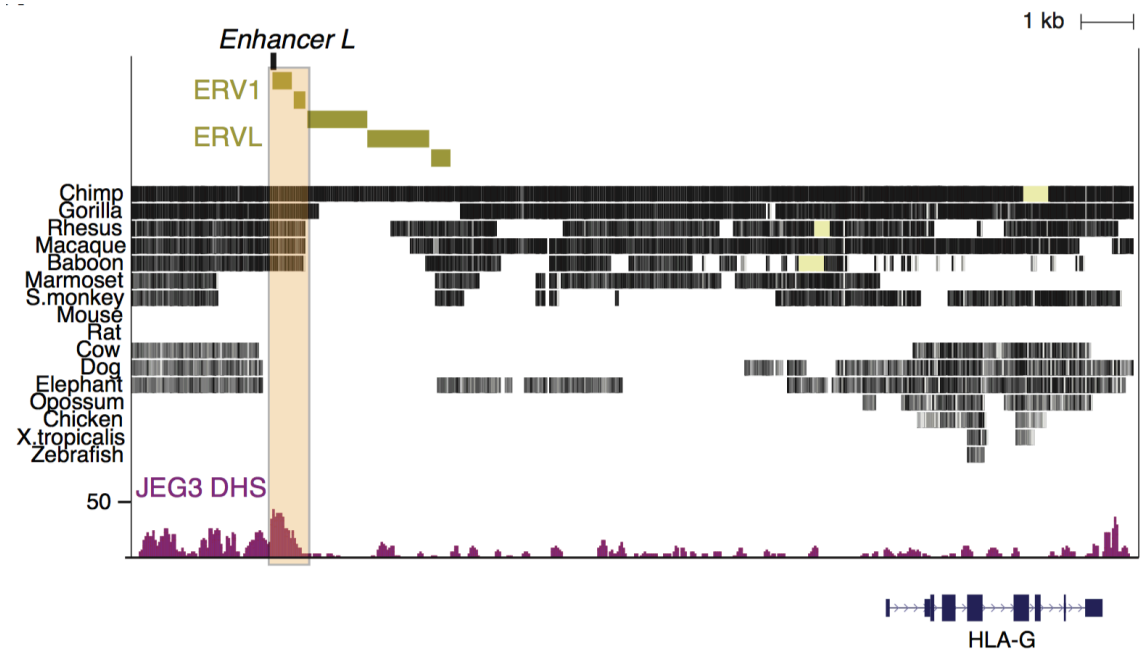


Figure 4.1. *Enhancer L* is part of a lineage-specific retrotransposon element found in primates.

The region containing *Enhancer L* (beige box) is conserved in apes (with the exception of orangutans) and Old World monkeys, but absent in New World monkeys and other placental mammals. This region is part of a long terminal repeat (LTR) retrotransposon element mostly found in the human genome, ERV1. A different class of retrotransposon elements, ERVL, is also present in the vicinity.

Enhancer L sequence is unique in the human genome and well conserved across apes and Old World monkeys, yet absent in New World monkeys (Figures 4.1 and 4.2), where HLA-G appears to be a classical MHC molecule (Adams and Parham, 2001; Arnaiz-Villena et al., 1999; Slukvin et al., 2000).

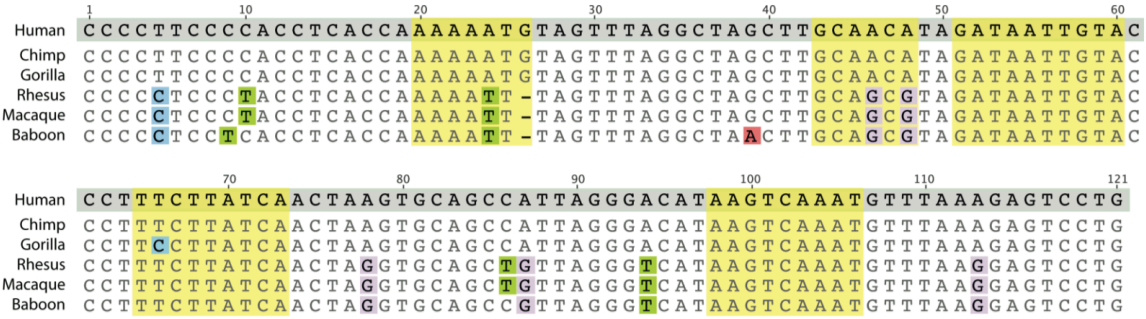


Figure 4.2. *Enhancer L* is conserved among apes and Old World monkeys.

Alignment of *Enhancer L* sequences across apes and Old World monkeys indicates strong conservation across the analyzed species. *Enhancer L* regulatory motifs are highlighted in yellow and mismatches in blue (C), green (T), pink (A) or purple (G).

Intriguingly, the orangutan genome, the only ape genome containing a functional *HLA-G* promoter (Figure 4.3), does not harbor the *Enhancer L* sequence. In addition, similar to New World monkeys, the orangutan HLA-G ortholog is a polymorphic MHC molecule.

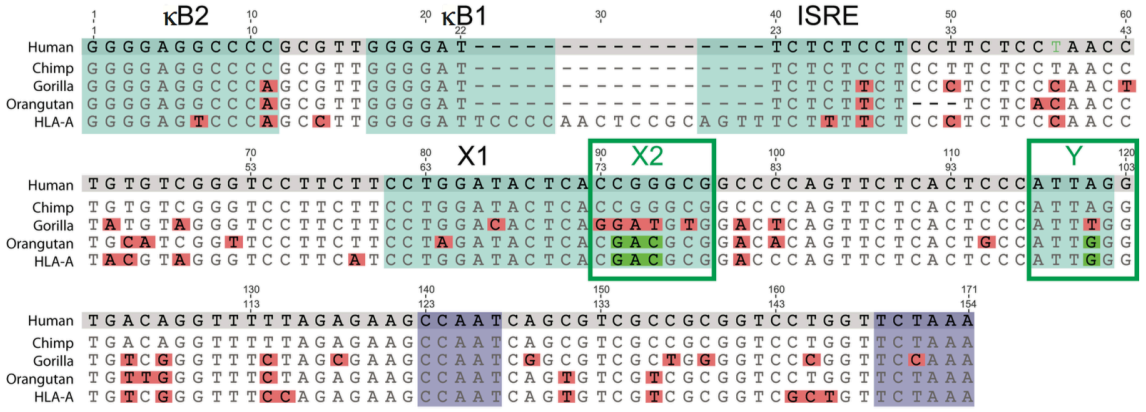


Figure 4.3. The orangutan genome, which lacks *Enhancer L*, contains a potentially functional *HLA-G* promoter.

Alignment of the *HLA-G* classical promoter sequences across apes reveals that orangutans, which lack an *Enhancer L* sequence, possess intact X2 and Y motifs

(highlighted in green), unlike any other ape. Mismatches are colored pink; kB2, kB1, ISRE, X1, X2 and Y boxes characteristic of an HLA classical promoter are highlighted in blue; the ubiquitous CAAT an TATA boxes necessary for transcription are highlighted in purple.

Perhaps in orangutans, because they are predominantly monogamous and thus less exposed to allogeneic fetuses (Arnaiz-Villena et al., 1999), *HLA-G* functions as a classical antigen-presenting molecule. The observation that *Enhancer L* is only found in genomes that lack a functional *HLA-G* classical promoter raises the possibility that a retroviral element was co-opted during evolution to function in trophoblast-specific tolerogenic MHC expression.

Previous literature suggests that differential expression of transcription factors plays a role in cell type-specific HLA-G transcription. The identity of such factors, however, has remained elusive (Moreau et al., 1998; Moreau et al., 1997). In our study, MPRA-based saturation mutagenesis allowed us to fine map the regulatory elements responsible for *Enhancer L* activity, ultimately pointing towards CEBP and GATA factors as candidates for transcriptional activators of HLA-G expression in trophoblasts (Figures 3.2.2 and 3.2.3). Indeed, chromatin immunoprecipitation (ChIP) and transient transfection studies revealed that CEBPβ, GATA2, and GATA3 associate with *Enhancer L* (Figure 3.3.3) and are positive regulators of *HLA-G* expression (Figure 3.3.5).

Chromatin Conformation Capture (3C) revealed that *Enhancer L* loops across a 12 kb distance into the classical promoter of *HLA-G* (Figures 3.1.3 and 3.1.4). Consistent with

this long-range chromatin interaction, genome-wide DNase-seq demonstrated that *Enhancer L* is part of a DNase hypersensitive site (DHS) specifically in HLA-G$^+$ JEG3 cells (Figure 4.4). Publicly available ChIP-seq data indicates CTCF binding flanking *Enhancer L* and the *HLA-G* coding sequence (Figure 4.4).

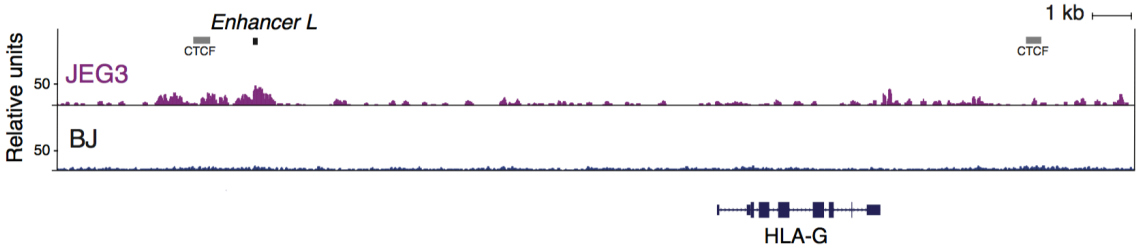


Figure 4.4. *Enhancer L* is part of an open chromatin region specifically in JEG3.

Enhancer L is part of an approximately 1 kb-long DNase Hypersensitive Site (DHS) in JEG3 cells that is absent in control BJ fibroblasts (HLA-G$^-$), as revealed by genome-wide DNase-seq. Binding of the insulator CTCF upstream of *Enhancer L* and downstream of *HLA*-G, according to ENCODE ChIP-seq data, indicates potential boundaries of the *HLA-G* regulatory chromatin domain.

This CTCF binding pattern suggests the existence of an insulated chromatin domain (Dowen et al., 2014) for *HLA-G* transcriptional regulation, corroborated by our observation that *Enhancer L* deletion does not significantly alter the expression of any gene other than *HLA-G* on chromosome 6 (Figures 2.3.2 and 2.3.3). Interestingly, a long-range chromatin interaction mediated by the insulator CTCF has been described in the MHC class II locus (Majumder et al., 2008). Our data suggests that the looping interaction between *Enhancer L* and the promoter of *HLA-G,* is mediated by GATA2/3

(Chen et al., 2012; Deng et al., 2012), possibly in association with CEBPβ (Tong et al., 2005).

In conclusion, we have demonstrated that trophoblast HLA-G expression is contingent upon the activity of a remote enhancer, *Enhancer L*. Our data is consistent with a model where CEBPβ and GATA2/3 associate with *Enhancer L*, are recruited to the core promoter of *HLA-G* via chromatin looping, and upregulate HLA-G expression (Figure 3.3.6). These findings establish chromatin looping mediated by lineage-specific transcription factors as a premier mechanism governing tissue-specific gene expression at the maternal-fetal interface.

Several outstanding questions regarding the trophoblast-restricted expression of *HLA-G* remain to be addressed. Are there trophoblast-restricted transcription factors controlling *HLA-G* expression? Do they act via *Enhancer L*? We observed that the HLA-G negative HEK293T cell line expresses all three factors found by us to associate with *Enhancer L* and to activate *HLA-G* transcription– CEBPβ, GATA2, and GATA3 – and yet HEK293T cells do not express HLA-G. In fact, our DHS and 3C analyses clearly showed that *Enhancer L* is not active and does not loop into the *HLA-G* proximal promoter in this HLA-G negative cell line. The cell type-specific activity of *Enhancer L* therefore suggests that additional transcription factors and/or mechanisms of gene regulation are required to drive HLA-G expression in EVT.

Intriguingly, EVT express neither NLRC5 nor CIITA, and yet constitutively express three MHC class I genes: HLA-C, HLA-E, and HLA-G. How are these three genes expressed

simultaneously in the absence of HLA-A and HLA-B transcription? Our work indicates that HLA-G expression is accomplished by a mechanism fundamentally different from the one previously described for classical MHC class I gene expression: looping of a distant enhancer mediated by CEBP and GATA transcription factors into the proximal promoter.

Yet, this looping mechanism does not exclude the possibility that a transactivator other than NLRC5 and CIITA assembles a transcriptional complex at the *HLA-G* promoter, similar to the well-studied enhanceosome found at the promoters of the classical MHC class I genes. If that is indeed the case, is such alternative enhanceosome also found at the promoters of HLA-C and HLA-E specifically in trophoblasts? Are there co-opted retroviral enhancers analogous to *Enhancer L* upstream of *HLA-C* and *HLA-E*? Functional dissection of the *HLA-C* and *HLA-E* loci using MPRA, analogous to the strategy used in our study for *HLA-G*, is likely to reveal novel cis-regulatory elements required for trophoblast expression of HLA-C and HLA-E in the absence of the classical transcriptional regulators described in other cell types.

We further propose two unbiased experiments aimed at gaining a more complete understanding of HLA-G expression in EVT. The first one would entail constructing a reporter cell line by knocking in a 2A-GFP gene downstream of *HLA-G*. This tool would allow HLA-G transcription levels to be quantitatively measured by GFP expression, a readout that can be exploited for a large scale screen. This reporter cell line would then be utilized in a high-throughput CRISPR/Cas9-based loss-of-function screen for transcriptional regulators of HLA-G expression (Wang et al., 2014a). This concept could be used to create more complex reporter cell lines: a cell line where HLA-G, HLA-C, and

HLA-E are each coupled to a distinct reporter gene would make it possible to identify genes specifically involved in HLA-G expression vs. all three MHC genes expressed by EVT. Of note, the use of reporter genes, as opposed to relying on surface expression, would eliminate gene candidates involved in anything other than gene regulation at the transcriptional level.

The second experiment would focus on the comprehensive identification of transcriptional regulators that bind to *Enhancer L* and to the classical promoter of *HLA-G*. This could be accomplished by carrying out CRISPR/Cas9-based "reverse ChiP" on these regions. In brief, these genomic regions would first be targeted with a catalytically dead Cas9 and sequence-specific gRNAs, and then pulled down by immunoprecipitation of Cas9. In principle, this strategy allows for the pulldown of any genomic region without relying on a known bound transcription factor, leaving all protein-DNA interactions unaffected. All trans-regulatory proteins bound to these specific regions would then be characterized by mass spectrometry analysis (Fujita and Fujii, 2013, 2014). If successful, this approach would allow us to dissect the trophoblast MHC enhanceosome, and further dissect the mechanism by which *Enhancer L* activates trophoblast-restricted HLA-G expression.

Future studies further dissecting the transcriptional regulation of *HLA-G* will not only shed light on immune privilege during pregnancy, but may also enable us to specifically control HLA-G expression in any desired cell type. This knowledge has the potential to translate into novel strategies to address pregnancy disorders and cancer.

APPENDIX 1: Experimental procedures used in CHAPTERS 2 through 4

Cell Culture

JEG3 and HEK293T cells were cultured in RPMI-1640 medium (Gibco) supplemented with 10% FBS, Glutamax, and Penicillin-Streptomycin. Transfections were carried out using FuGENE 6 (Promega) according to the manufacturer's instructions and analyzed 48 h post-transfection.

Flow cytometry

Cells were harvested, blocked in 4% FBS for 30 min, stained with HLA-G PE (clone MEMG/9, Abcam) in 1% FBS for 1 hour, washed thrice and resuspended in 1 % FBS. Cells were acquired using either a FACSCalibur or an LSR-II instrument (BD Biosciences) and analyzed with FlowJo (TreeStar) software.

Quantitative Real Time PCR (qRT-PCR) analysis

Total RNA was isolated using TRIzol (Life Technologies), according to manufacturer's instructions. 1500 ng RNA was used for cDNA synthesis with the qScript cDNA SuperMix (Quanta Biosciences). 30 ng cDNA was used per qRT-PCR reaction, performed using SYBR Green (Life Technologies) on a ViiA7 system real-Time PCR System (Life Technologies). Target gene expression levels were normalized to *GAPDH*.

Molecular Biology. For CRISPR/Cas9 genome editing in JEG3 cells, a human codon-optimized *S. pyogenes* Cas9 gene with a C-terminal nuclear localization signal (Mali et al., 2013) subcloned into a CAG expression plasmid upstream of a 2A-GFP (Ding et al., 2013b) was used. The guide RNAs (gRNAs) were cloned into a separate plasmid

containing the human U6 polymerase III promoter (Mali et al., 2013) using BbsI restriction sites. For lentiviral delivery of CRISPR/Cas9 to primary extravillous trophoblasts (EVT), Cas9 was instead expressed from a human UbC promoter and upstream of a T2A-GFP (Kabadi et al., 2014). gRNAs were subcloned into the lentiGuide-puro vector (Sanjana et al., 2014). Lentiviral production was carried out in HEK293T cells using psPAX2 and VSV-G as packaging plasmids, as described (Kabadi et al., 2014). Tested transcription factor genes were amplified from JEG3 cDNA and directionally cloned into a plasmid containing a CMV promoter upstream of an IRES-GFP.

Massively Parallel Reporter Assay (MPRA)

First, 12,000 oligonucleotides tiling the *HLA-G* locus (27 kb) coupled to distinguishing tags were generated using microarray-based DNA synthesis. The 121 bp-long tiles and tags are separated by two common restriction sites. The oligonucleotides were then PCR amplified from universal primer sites and directionally cloned into a pGL4 plasmid backbone (Promega) using Gibson assembly. An invariant promoter-firefly luciferase segment containing either a minimal TATA box weak (minP) or strong (SV40P) promoter was then inserted between the tiles and tags by double digestion and directional ligation. The resulting reporter plasmid pools were co-transfected into JEG3 cells using FuGENE 6 (Promega). Two biological replicate MPRA experiments were performed. The relative enhancer activities of the different tiles were inferred by sequencing and counting their corresponding tags from the cellular mRNA and the transfected plasmid pool, as described in (Melnikov et al., 2012). Nominal hits were defined as any tile where both enhancer activity measurements were > 1 and with p-values < 0.05 in both replicates.

Those that agreed between SV40P and minP promoter data sets were considered the most confident hits. For the second MPRA experiment, a single-hit scanning mutagenesis (Melnikov et al., 2012), 12,000 *Enhancer L* variants were generated, including all possible single substitutions, multiple series of consecutive substitutions and small insertions at all positions. Each variant was linked to an average of 16 tags each. The remainder of the workflow was as described above.

Luciferase reporter gene assays

Individual candidate regions were amplified from JEG3 genomic DNA and directionally cloned into a pGL4 plasmid (Promega) containing either the minP or the SV40P promoter and firefly luciferase. JEG3 and HEK293T cells were transfected in 24-well plates using FuGENE 6 (Promega) with the individual firefly luciferase constructs and Renilla luciferase at a 10:1 ratio. 48 h post-transfection, firefly luciferase activity was measured using a Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions and normalized to Renilla luciferase to control for cell number and transfection efficiency.

Genome-wide DNase-seq

DNase I digestion followed by sequencing was performed as previously described (Hesselberth et al., 2009; Sherwood et al., 2014). See Extended Experimental Procedures for details. In short, 10 million cells were harvested, washed twice with ice-cold PBS and resuspended in Buffer A containing protease inhibitors and Spermidine (Sigma). Nuclei were extracted using ice-cold 0.05% NP-40 in Buffer A, centrifuged at 800 g for 5 min at 37^{o}C and gently resuspended in ice-cold PBS. An aliquot was taken to estimate nuclei number and integrity of using a cell counter (Bio-Rad). Intact nuclei were

washed twice with ice-cold isotonic buffer and digested with empirically determined limiting concentrations of DNase I (Sigma) for 3 min at 37°C. Digests were stopped with EDTA and the samples were incubated with Proteinase K overnight at 55°C. DNA was phenol/chloroform-extracted and concentrated by ethanol precipitation. Selection of 175-400 bp DNA fragments using the E-gel Agarose System (Invitrogen) was performed to select for regions in which DNase I can cut twice (at both ends), enriching for hypersensitive regions (Sherwood et al., 2014). Library preparation and sequencing were performed at the MIT BioMicroCenter. Prepared libraries were sequenced on a Hiseq2000 sequencing system (Illumina) to a depth of 160M-230M reads per sample using paired-end reads with a length of 40bp. These were aligned to the human genomes (version hg19, canonical chromosomes only) using bwa version 0.6.2 with default parameters. Quality control tests and regions of DNase hypersensitivity were calculated using the tool Hotspot-SPOT (v4) (John et al., 2011) with an FDR of 0.01.

Chromatin Conformation Capture (3C)

3C assays were carried out essentially as described in (Gavrilov et al., 2009; Majumder et al., 2008). 10^7 cells were resuspended in 9 ml of medium and cross-linked using 2% formaldehyde 10 min at room temperature (RT). The cross-linking reaction was quenched with 0.125 M glycine on ice and the cells were washed twice with cold PBS. Cells were lysed in 5 ml cell lysis buffer containing protease inhibitors (Gavrilov et al., 2009) on ice for 10 min. The nuclei were resuspended in NEBuffer DpnII (New England Biolabs) containing 0.3% SDS and incubated in a thermomixer 1 h at 37°C shaking at 1400 rpm. Next, 1.8% Triton X-100 was added to sequester the SDS and the samples were incubated for an additional hour at 37°C shaking at 1400 rpm. The cross-linked DNA was digested with 1,000 units DpnII (New England Biolabs) at 37°C overnight.

DpnII was heat inactivated at 65°C for 20 min. For ligation of DNA ends, T4 DNA ligase was added and the samples were incubated for 4 h at 16°C, followed by 30 min at RT. Cross-links were reversed by incubating with Proteinase K (10 mg/ml) at 65°C overnight. Finally, the DNA was phenol/chloroform-extracted and concentrated by ethanol precipitation. 50 ng DNA were analyzed by PCR.

CRISPR/Cas9 genome editing

JEG3 cells were transfected with Cas9-2A-GFP and guide RNAs targeting *Enhancer L*. GFP^+ cells were sorted 48 hours post-transfection and plated at clonal density in 10 cm dishes. Approximately 10 days after plating, single cell-derived colonies were picked into 96-well plates and cultured for an additional 10 days. For PCR analysis, cells were harvested and genomic DNA extracted using prepGEM Tissue (ZyGEM). Selected WT and KO clones were then expanded and further characterized.

Transcriptome-wide RNA-seq

Total RNA from JEG3 cells was extracted using TRIzol (Life Technologies), according to manufacturer's instructions and then purified by spin column purification (RNeasy mini kit, Qiagen) using a QIAcube system. RNA was quantified using a Nanodrop (Thermo Fisher) and its integrity assessed on a Bioanalyzer (Agilent) using the RNA 6000 RNA chip. 500 ng high-quality total RNA (RNA Integrity Number ≥8) was used as input for Tru-seq library construction using the TruSeq RNA Sample Preparation Kit (Illumina), as described in (Sun et al., 2013; Trapnell et al., 2013) Library purity, correct fragment size, and concentration were assessed using the Bioanalyzer DNA7500 chip. Libraries free of adapter dimers and with a peak region area (220–500 bp) ≥80% of the total area were individually barcoded, pooled, and sequenced on an Illumina HiSeq 2000 platform.

Reads were mapped to the human genome (hg19) using TopHat v2.0.14 (Kim et al., 2013; Trapnell et al., 2009) with the flags: "--no-coverage-search --GTF gencode.v19.annotation.gtf" where gencode.v19.annotation.gtf is the Gencode v19 reference transcriptome available at gencodegenes.org. Cufflinks v2.2.1 (Trapnell et al., 2013) was used to quantify gene expression and assess the significance of differential expression. Briefly, Cuffquant was used to quantify mapped reads against Gencode v19 transcripts of at least 200bp with biotypes: protein_coding, lincRNA, antisense, processed_transript, sense_intronic, sense_overlapping. Cuffdiff was run with default options on the resulting .cxb files. Gene set enrichment analysis was performed with the GSEA v2.1.0 (Mootha et al., 2003; Subramanian et al., 2005) from the Broad Institute. Genes with expression greater than 1 FPKM in WT or KO conditions were sorted in decreasing order by the absolute value of their log2 fold change as determined by Cuffdiff. The resulting ranked list was fed to the GSEA pre-ranked tool with default options and a permutation seed of 42.

First trimester primary extravillous trophoblast (EVT) isolation and transduction

Discarded human placental and decidual material (gestational age 6-12 weeks) was obtained from women undergoing elective pregnancy termination at a local reproductive health clinic. All of the human tissue used for this research was de-identified, discarded clinical material. The Committee on the Use of Human Subjects (the Harvard IRB) determined that this use of all of this human material is exempt from the requirements of IRB review. Extravillous trophoblasts (EVT) were isolated as previously described (Tilburgs et al., 2015a). 50-100k $CD45^{-}HLA\text{-}G^{+}$ EVT were plated in 48-well cell culture plates (Costar) pre-coated with 100 µl of 20ng/ml fibronectin for 45 minutes (BD), in Trophoblast Medium, which consisted of DMEM/F12 medium (Gibco) supplemented with

10% NCS, Glutamax, insulin, transferrin, selenium (100X, Gibco), 5 ng/ml EGF (Peprotech), and 400 units of human gonadotropic hormone (Sigma). Two hours post-plating, EVT were transduced with Cas9-T2A-GFP and *Enhancer L* gRNAs' lentiviral particles pseudotyped with G glycoprotein from vesicular stomatitis virus (VSV-G) in the presence of 8 µg/ml polybrene (Hexadimethrine bromide, Sigma). Lentiviral particles were produced using HEK293T cells and were concentrated 20X using Lenti-X Concentrator (Clontech). Transduction was performed three additional times, 12 h apart. Three days after the first transduction, medium was switched to villous stromal cell (VSC)-conditioned medium (RPMI-1640 medium supplemented with 10% FBS, Glutamax, and Penicillin-Streptomycin) for an additional 2 days. For analysis by flow cytometry, cells were washed with warm PBS, harvested with Trypsin, and resuspended in trophoblast medium for staining with ITGA5 PE (an EVT marker) (Tilburgs et al., 2015a) and HLA-G APC (both Biolegend).

Chromatin immunoprecipitation (ChIP-qPCR)

ChIP-qPCR was performed using Dynabeads Protein G (Life Technologies) according to the manufacturer's instructions. JEG3 cells were harvested, resuspended in PBS and cross-linked using 1% formaldehyde. Glycine was added to stop the cross-linking reaction and cells were washed twice with ice-cold PBS. Nuclei were isolated using ice-cold Cell Lysis Buffer containing protease inhibitors and PMSF (CalBiochem) and then lysed using ice-cold Nuclei Lysis Buffer containing protease inhibitors and PMSF. Cross-linked chromatin was sheared using a Bioruptor Standard Sonication Device UCD-200 (Diagenode) to 200-500 bp fragments, assessed by gel electrophoresis. Samples were then diluted with ChIP Dilution Buffer and incubated with Mock IgG, Pol II, CEBPB, GATA2, or GATA3 antibodies (Santa Cruz Biotechnology) overnight at 4°C. Dynabeads

Protein G were blocked with BSA and glycogen at 4°C during the same period of time. The following day, beads were washed and eluted. The eluates were then reverse cross-linked and DNA was purified using phenol/chloroform extraction. 25 ng DNA were used for qRT-PCR analysis using SYBR Green (Life Technologies).

APPENDIX 2: List of primers used in CHAPTERS 2 through 4

***Enhancer L* CRISPR gRNAs**		
CRISPR 1	Fw:	5'-CACCGGCCAGGCACCTATCGGTTG-3'
	Rev:	5'-AAACCAACCGATAGGTGCCTGGCC-3'
CRISPR 2	Fw:	5'-CACCGAGACAATCACAATGCATTA-3'
	Rev:	5'-AAACTAATGCATTGTGATTGTCTC-3'
***Enhancer L* deletion PCR**		
	Fw:	5'-CATGGTCATAAAGAGATAAAG-3'
	Rev:	5'-CTTACGATCTTCCCGGATGTC-3'
qRT-PCR		
GAPDH	Fw:	5'-GAAGGTGAAGGTCGGAGT-3'
	Rev:	5'-GAAGATGGTGATGGGATTTC-3'
HLA-G	Fw:	5'-GCTGCCCTGTGTGGGACTGAGTG-3'
	Rev:	5'-GACGGAGACATCCCAGCCCCTTT-3'
3C		
Loading control	Fw:	5'-CACAAGAGTAGCGGGGTCAG-3'
	Rev:	5'-GAGCAGCAGGAAGAGGGTTC-3'
Enhancer L-Promoter looping	Fw:	5'-GCATGAAAGGGAAAGCAAG-3'
	Rev:	5'-CACTCCATGAGGTATTTCAG-3'
ChIP-qPCR		
GAPDH 3' UTR	Fw:	5'-TCGACAGTCAGCCGCATCT-3'
	Rev:	5'-CTAGCCTCCCGGGTTTCTCT-3'
HBB Promoter	Fw:	5'-CTGGTGGGGTGAATTCTTTGC-3'
	Rev:	5'-AGTCCAAGCTAGGCCCTTTT-3'
CEBP Positive Control	Fw:	5'-AGACTTTGAAGACGATTCAGCA-3'
	Rev:	5'-ACCCCTGATTGCTCAACACT-3'
GATA Positive Control	Fw:	5'-CTCTGGCCGGTCGATGTTATC-3'
	Rev:	5'-GATGGCGGCTGCGATTAAC-3'
Enhancer L	Fw:	5'-GGCCAGGCACCTATCGGTTG-3'
	Rev:	5'-TAATGCATTGTGATTGTCTC-3'
HLA-G Classical Promoter	Fw:	5'-GTGGCTCTCAGGGTCTCAGG-3'
	Rev:	5'-CGACGCTGATTGGCTTCTCTA-3'

APPENDIX 3: Efficient gene ablation in primary human hematopoietic cells using CRISPR/Cas9

Adapted from: Pankaj K. Mandal, **Leonardo M. R. Ferreira**, Ryan Collins, Torsten B. Meissner, Christian L. Boutwell, Max Friesen, Brian S. Garrison, Alexei Stortchevoi, David Bryder, Kiran Musunuru, Harrison Brand, Todd M. Allen, Michael E. Talkowski, Derrick J. Rossi, and Chad A. Cowan. *Cell Stem Cell* 15 (5), 2014.

A.1. ABSTRACT

Genome editing via CRISPR/Cas9 has rapidly become the tool of choice by virtue of its efficacy and ease of use. However, CRISPR/Cas9 mediated genome editing in clinically relevant primary human somatic cells remains untested. Here, we report the CRISPR/Cas9 targeting of two clinically relevant genes, *B2M* and *CCR5*, in primary human $CD4^+$ T cells and $CD34^+$ hematopoietic stem and progenitor cells (HSPCs). Use of single RNA guides led to highly efficient mutagenesis in $CD34^+$ HSPCs but not in $CD4^+$ T cells. A dual guide approach improved gene deletion efficacy in both cell types. HSPCs that had undergone genome editing with CRISPR/Cas9 retained multi-lineage potential. We examined predicted on- and off-target mutations via target capture sequencing in HSPCs and observed low levels of off-target mutagenesis at one only site. These results demonstrate that CRISPR/Cas9 can efficiently ablate genes in HSPCs with minimal off-target mutagenesis, which could have broad applicability for hematopoietic cell-based therapy.

A.2. INTRODUCTION

The hematopoietic system is at the forefront of regenerative medicine and cell-based gene therapies due to the fact that diverse progenitor and terminally differentiated cells can be readily obtained, manipulated, and reintroduced into patients (Weissman, 2000). The development of genome editing methodologies such as meganucleases, zinc-finger nucleases (ZFNs) (Urnov et al., 2010), and transcription activator-like effector nucleases (TALENs) (Joung and Sander, 2013; Scharenberg et al., 2013), have enabled site-specific gene repair or ablation and raised the possibility of using these technologies to treat a broad range of diseases at the genetic level (Pan et al., 2013), with targeted therapies in hematopoietic cells at the fore of such efforts (Aiuti et al., 2013; Biffi et al., 2013; Genovese et al., 2014; Tebas et al., 2014). Despite much promise, limitations associated with first generation gene editing technologies, which include low targeting efficacy and laborious *de novo* engineering of proteins for each target have precluded wide-spread adoption of these technologies for therapeutic use (Silva et al., 2011). The recent emergence of the clustered, regularly interspaced, palindromic repeats (CRISPR) system for gene editing has the potential to overcome these limitations (Jinek et al., 2012). The CRISPR technology utilizes a fixed nuclease, in most cases the CRISPR-associated protein 9 (Cas9) from *Streptococcus pyogenes,* in combination with a short guide RNA (gRNA) to target the nuclease to a specific DNA sequence (Cong et al., 2013; Jinek et al., 2012; Jinek et al., 2013; Mali et al., 2013). CRISPR/Cas9 relies on simple base-pairing rules between the target DNA and the engineered gRNA rather than protein-DNA interactions required by ZFNs and TALENs (Gaj et al., 2013; Wei et al., 2013). As a result, the CRISPR/Cas9 system has proven extremely simple and flexible (Sander and Joung, 2014). Perhaps most important, this system has achieved highly

efficacious alteration of the genome in a number of cell types and organisms (Ding et al., 2013b; Hwang et al., 2013; Niu et al., 2014; Wang et al., 2013; Wei et al., 2013).

Given the importance of the hematopoietic system in cell-based gene therapies, we tested the CRISPR/Cas9 system in primary human $CD4^+$ T cells and $CD34^+$ hematopoietic stem and progenitor cells (HSPCs) targeting two clinically relevant genes, beta-2 microglobulin (*B2M*) and chemokine receptor 5 (*CCR5*). *B2M* encodes the accessory chain of major histocompatibility complex (MHC) class I molecules and is required for their surface expression (Bjorkman et al., 1987b; Zijlstra et al., 1990). Deletion of *B2M* is a well-established strategy to ablate MHC class I surface expression (Riolobos et al., 2013), and could be used to generate hypoimmunogenic cells for transplantation and adoptive immunotherapy. *CCR5* is the main co-receptor used by CCR5-tropic strains of HIV-1 (Trkola et al., 1996) and a validated target for gene ablation, as mutations resulting in loss of protein expression or haploinsufficiency protect against HIV infection(Catano et al., 2011; Hutter et al., 2009; Martinson et al., 1997; Samson et al., 1996). Moreover, transplantation of CCR5 homozygous mutant (delta32) hematopoietic stem cells provides long-term protection against HIV rebound even after discontinuation of antiretroviral therapy (Allers et al., 2011; Hutter et al., 2009). Several attempts have been made to target *CCR5* in $CD4^+$ T cells (Perez et al., 2008; Tebas et al., 2014) and $CD34^+$ HSPCs (Holt et al., 2010) using ZFNs. However, the efficiency of gene targeting reported was not sufficient to protect against viral recrudescence (Tebas et al., 2014). Recently, *CCR5* has been targeted using the CRISPR/Cas9 system in cell lines (Cho et al., 2013) and iPS cells (Ye et al., 2014). However, CRISPR/Cas9 gene editing in primary human hematopoietic cells remains untested. Here we report that use of CRISPR/Cas9 with single gRNAs led to highly efficient *CCR5* ablation in $CD34^+$

HSPCs but not *B2M* in $CD4^+$ T cells. Employing a dual gRNA approach identified gRNA pairs that improved gene deletion efficacy in both cell types with biallelic inactivation frequencies reaching 34% for *B2M* in primary $CD4^+$ T cells, and 42% for *CCR5* in primary $CD34^+$ HSPCs. Importantly, CRISPR/Cas9 *CCR5*-edited $CD34^+$ HSPCs retained multi-lineage potential *in vitro* and *in vivo* upon xenotransplantation.

Deep target capture sequencing of predicted on- and off-target sites in $CD34^+$ HSPCs targeted with multiple single or dual gRNA combinations revealed highly efficacious on-target mutagenesis, and exceedingly low off-target mutagenesis.

A.3. EXPERIMENTAL PROCEDURES

Molecular Biology

We subcloned a human-codon-optimized *Cas9* gene with a C-terminal nuclear localization signal (Mali et al., 2013) into a CAG expression plasmid with GFP (Ding et al., 2013a). The guide RNAs (gRNAs) were separately expressed from a plasmid with the human *U6* polymerase III promoter (Mali et al., 2013). Each gRNA sequence was introduced in this plasmid using BbsI restriction sites.

Primary blood cell isolation

Primary $CD4^+$ T cells were isolated from peripheral blood (Leukopacs, MGH) using RosetteSep $CD4^+$ T cell enrichment cocktail (STEMCELL Technologies). $CD34^+$ cells from G-CSF mobilized peripheral blood were purchased from AllCells.

Cell culture

HEK293T cells were grown as adherent monolayers in RPMI-1640 medium supplemented with 10% FBS. K562 and T cells were cultured in RPMI-1640 medium supplemented with 10% FBS. $CD34^+$ HSPCs were cultured in DMEM/F12 medium supplemented with 10% FBS, β-mercaptoethanol, GlutaMax, Pencillin-Streptomycin, minimum non-essential amino acid and human cytokine cocktails (GM-CSF, SCF, TPO, Flt3 ligand, IL3, IL6). Cell lines were passaged every 3-4 days.

Transfection of Cells

Human primary $CD4^+$ T cells and $CD34^+$ HSPCs were transfected with Cas9-2A-GFP and gRNA encoding plasmids using respective Amaxa Nucelofection kits and cell-specific Nucleofection program with an Amaxa Nucleofection II device as per manufacturers instructions with minor modifications. HEK293T cells were seeded in 6-well plates the day before transfection and transfected using Fugene 6 (Promega).

Cell sorting

For the CCR5 targeting experiments in $CD34^+$ HSPCs, cells were plated in antibiotic free medium following transfection. 24 hours post-transfection, cells were harvested in sample medium (2% FBS and 2 mM EDTA in PBS without Ca^{2+} and Mg^{2+}) and HSPCs were stained with anti-CD34-PE/Cy7 (clone: 581, Biolegend, 1:100) for 20 min on ice. Live, GFP^+ $CD34^+$ HSPCs were sorted using an Aria II sorter (BD Bioscience) and plated in complete DMEM/F12 medium supplemented with human cytokine cocktail and culture for 72 hours prior to analysis. For the B2M experiments, cells were stained with mouse monoclonal anti-B2M-APC antibody (clone: 2M2, Biolegend) 48 or 72 hours post-transfection to estimate loss of B2M expression. FACS data were analyzed using FlowJo software.

Colony forming cell (CFC) assay

1500 $CD34^+$ sorted cells were plated in 1.5 ml of methylcellulose (MethoCult™ H4034 Optimum, Stem Cell Technologies) on a 35 mm cell culture dish and cultured for two weeks at 37 °C in a 5% CO2 incubator. Colonies were then counted and scored.

Surveyor/CEL assay

Amplicons spanning the different targeted regions were PCR amplified using the Phusion polymerase and HF Buffer (New England Biolabs) and CEL assay was carried out using the Surveyor Mutation detection kit (Transgenomic) according to the manufacturer's instructions, with minor modifications.

Clonal analysis

Colonies grown in MethoCult™ H4034 Optimum were individually picked and lysed in 50 µl of lysis buffer containing detergent and Proteinase K buffer (van der Burg et al., 2011). Samples were digested at 56 °C for 1 h followed by Proteinase K inactivation at 95 °C for 15 min. 50 µl of water with RNase A were added to the samples. 2 µl of samples were use for PCR. A 436 bp amplicon spanning the targeted region was PCR amplified using GoTaq® Green Master Mix (Promega) as per manufacturer's instructions. For single gRNA experiments, PCR products were analyzed by Sanger sequencing (Macrogen). For dual gRNA experiments, PCR products were analyzed by agarose gel electrophoresis.

In vivo transplantation of $CD34^+$ HSPCs

NOD/SCID/IL2R$\gamma^{-/-}$ (NSG) mice (The Jackson Laboratory) were housed in a pathogen-free facility, maintained in microisolator cages, and fed autoclaved food and water. Adult

(6-8 weeks of age) NSG mice were conditioned with sub-lethal (2 Gy) whole-body irradiation. The conditioned recipients were transplanted with 75,000-sorted $CD34^+$ HSPCs expressing Cas9 alone (control group, n=2) or Cas9 with crCCR5_D+Q gRNAs (experimental group, n=5). At 12 weeks post-transplantation, all mice were euthanized and blood, bone marrow, and spleen samples were taken for characterization of human hematopoietic cell chimerism. Human $CD45^+$ cells were sorted for DNA isolation and analysis of *CCR5* deletion.

Single Cell PCR assay

48 h after electroporation with Cas9 and different gRNA combinations, GFP^+ primary $CD4^+$ T cells were sorted into 384-well plates (Twin tec skirted PCR plate, Eppendorf) containing 4 µl of prepGEM Tissue (ZyGEM) per well. Cells were lysed and digested following the manufacturer's instructions to release the genomic DNA. A multiplexed nested PCR was then carried out in the same plate with the primer combinations represented in Supplemental Figures S2C and S2F. The resulting DNA was then used in two subsequent PCR reactions, one amplifying a positive control region, to determine successful genomic DNA isolation from a single cell, and another one amplifying a region lying between the two gRNA binding sites, allowing us to quantify the percentage of cells homozygous for the dual gRNA induced deletion (Supplemental Figures S2D and S2G). Cells were scored based on the melting curves of the PCR amplicons. PCR reactions were performed in a Applied Biosystems ViiA 7 real-Time PCR System (Life Technologies).

Off-Target Prediction and Capture Sequencing

Degenerate gRNA off-target sequences were predicted for each gRNA targeting *CCR5* using the CRISPR Design off-target prediction tool (Hsu et al., 2013). Off-target sequences were further supplemented by alignment of each gRNA to the human genome using BOWTIE of which all results up to and including 3 mismatches were added to the total off-target list (Langmead et al., 2009). All instances of each predicted off-target sequence existent in the human genome reference build GRCh37v71 were recorded (Supplementary Table T1). Each guide RNA target site (n=6) and predicted off-target site (n=172) was selected for capture sequencing using the Agilent SureSelectXT Target Enrichment System. Capture intervals were expanded by approximately 500 bp in both the 5' and 3' directions to ensure exhaustive capture of the targeted region and detection of any genetic lesion occurring at or near a predicted gRNA on- or off-target site, as we have previously shown accurate capability to detect translocations and inversions using targeted capture of probes in proximity to a rearrangement breakpoint using a CapBP procedure as described (Talkowski et al., 2011). Probes were tiled with 60-fold greater density over each predicted 23bp on- or off-target gRNA binding site than the flanking kilobase of sequence. Isogenic $CD34^+$ HSPCs-mPB were transfected with CRISPR/Cas9 plasmids (one Cas9 only-treated control group, three treatment groups transfected with a single gRNA, and three treatment groups transfected with dual gRNAs). Sorted $CD34^+$ genome edited HSPCs were cultured for two weeks prior to DNA isolation. Capture libraries were prepared from DNA extracted from seven treatment groups. Capture libraries were sequenced as 101 bp paired-end reads on an Illumina HiSeq2000 platform.

NGS Data Processing and Computational Analysis

Read pairs were aligned to GRCh37v71 with Bwa-MEM v0.7.10-r789 (Li, arXiv 2013). Alignments were processed using PicardTools and SAMBLASTER (Faust and Hall, 2014). The Genome Analysis Toolkit (GATK) v3.1-1-g07a4bf8 was applied for base quality score recalibration, insertion/deletion (InDel) realignment, duplicate removal, and single nucleotide variant (SNV) and InDel discovery and genotyping per published best-practice protocols (McKenna et al, Genome Res 2010; DePristo et al, Nat Genet 2011; Van der Auwera et al, 2013). SNVs and InDels were annotated using ANNOVAR (Wang et al., 2010). Structural variants (SVs) were detected with LUMPY v0.2.5 considering both anomalous pair and split read evidence at a minimum call weight threshold of 7 and an evidence set score ≤ 0.05 (Layer et al., 2014). Candidate copy number variants (CNVs) were further statistically assessed by Student's t-test for a concomitant change in depth of coverage across the putative CNV. As a final exhaustive measure, each on- and off-target site was manually scrutinized in each capture library for evidence supporting predictable mutagenesis that is not detectable by the computational algorithms due to low levels of mosaicism in the sequenced population.

Evaluation of Off-Target Mutation Frequency

A statistical framework was developed to assess off-target mutational burden for each gRNA. For each off-target site (n=172), all reads with at least one nucleotide of overlap with that 23bp off-target site were collected and their CIGAR information was tabulated into categories as follows: reads representing small InDels (CIGAR contains at least one "I" or "D"), reads potentially representative of other rearrangements (CIGAR contains at least one "S" or "H"), and reads reflecting reference sequence (CIGAR did not match either of the two former categories). Such counts were gathered at all 172 sites in all

seven libraries and were further pooled to form comparison groups of “treatment” libraries (transfected gRNA matches corresponding off-target site gRNA) and “control” libraries (transfected gRNA does not match corresponding off-target site gRNA). Next, at each off-target site, relative n-fold enrichment of each read classification between treatment and control libraries was evaluated. Finally, a one-tailed Fisher’s Exact Test was performed to assess the statistical significance of enrichment of variant reads in treatments versus controls at each off-target site, followed by Bonferroni correction to retain an experiment-wide significance threshold of $\alpha = 0.05$.

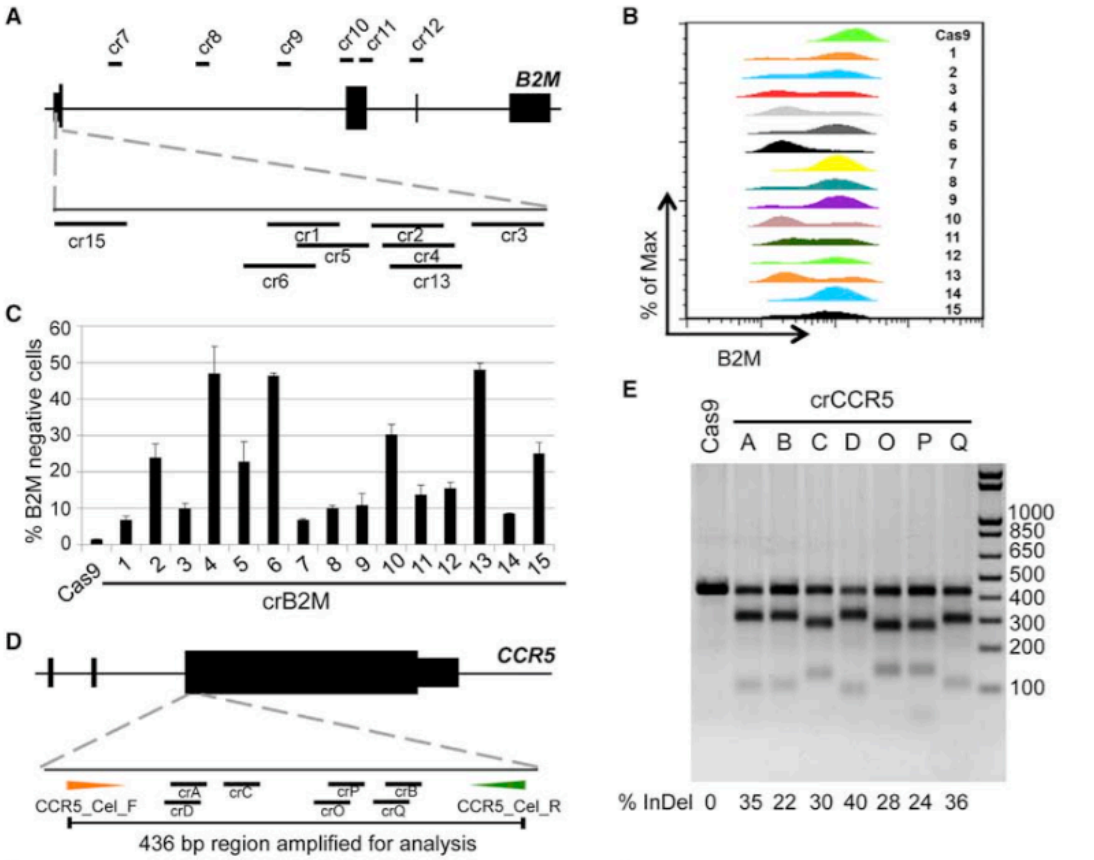
A.4. RESULTS

Considerable variation in targeting of *B2M* and *CCR5* in primary blood cells via CRISPR/Cas9

We designed several gRNAs to target Cas9 to the *B2M* gene (Figure A.4.1A). Each guide was then tested for the ability to direct site-specific mutations in HEK293T cells. As B2M is a surface antigen, we utilized flow cytometry to measure the efficiency of each guide to direct Cas9-mediated ablation of B2M expression 72 hours post-transfection (Figure A.4.1B). We observed that B2M surface expression was abrogated in 6.93% (±1.02 SEM, n=3) to 48% (±1.80 SEM, n=3) of HEK293T cells depending upon the gRNA utilized (Figures A.4.1C and A.4.2A). Similar results were observed using the CEL surveyor assay, with guide-specific mutation frequencies of 0% to 26% in HEK293T cells (Figure A.4.2B). In parallel, we also designed multiple gRNAs to target Cas9 to *CCR5* (Figure A.4.1D). Upon introducing these into K562 cells, we measured targeting efficacy using the CEL Surveyor assay and observed mutation frequencies ranging from 22-40% (Figure A.4.1E). Considerable variation in the efficiency with which a specific gRNA

directed Cas9-mediated ablation was observed, even between gRNAs targeting the same exon or nearly overlapping DNA sites at both the *B2M* and *CCR5* loci (Figures A.4.1A through A.4.1E) indicating that on-target efficiency of site directed mutation with the CRISPR system is highly gRNA dependent, as previously noted (Hsu et al., 2013).

Next, we tested selected individual guides in primary $CD4^+$ T cells from peripheral blood and $CD34^+$ hematopoietic stem and progenitor cells isolated from mobilized peripheral blood (HSPCs-mPB). To our surprise, gRNAs that had proven to be highly efficacious at targeting *B2M* in HEK293T cells invariably exhibited much lower targeting efficiencies in primary $CD4^+$ T cells ranging from 1.4% (±0.2 SEM, n=6) to 4.7% (±0.9 SEM, n=6) when measuring *B2M* surface expression (Figures A.4.1F, A.4.2C, and A.4.2D) or from 3% to 11% using the CEL surveyor assay (Figures A.4.2B and A.4.2E). For instance, gRNA crB2M_13 targeting *B2M* exhibited more than 10-fold reduced efficacy in $CD4^+$ T cells (4.7%±0.9) as compared to HEK293T cells (48.0%±1.8) (Figures A.4.1F and A.4.2C). Interestingly, individual gRNAs targeting *CCR5* showed comparably high mutation frequencies in $CD34^+$ HSPCs from mobilized peripheral blood (mPB) as had been observed in K562 cells, as determined by the CEL Surveyor assay (Figures A.4.1E and A.4.1G). To explore this further, we performed direct Sanger sequencing of several hundred colonies derived from individual HSPCs clones targeted with gRNAs crCCR5_A or crCCR5_B from two donors and observed very high monoallelic and biallelic mutation frequency in all cases (Figure A.4.1H). As only FACS sorted cells expressing Cas9 were analyzed in these experiments, it is unlikely that the observed differences in on-target mutation efficiency were due to differences in transfection efficiencies but rather may reflect intrinsic differences amongst primary hematopoietic cell types.



H

	Donor 1			Donor 2		
crCCR5	# clones	+/- (%)	-/- (%)	# clones	+/- (%)	-/- (%)
A	66	24.2	25.8	90	13.3	30
B	70	21.4	30	68	26.4	29.4

Figure A.4.1 Targeting clinically relevant loci in human cells.

A) Schematic of gRNAs targeting *B2M* locus. B) Histogram of B2M surface expression in 293T cells 72 h after CRISPR/Cas9 transfection. C) B2M deletion efficiency with various gRNAs in HEK293T cells. Pooled data from three independent experiments are shown

(Continued) (mean±SEM). D) Schematic of gRNAs targeting the *CCR5* gene. Orange and green arrows represent primer pairs used to amplify the targeted region for analysis. E) CEL surveyor assay of each gRNA targeting *CCR5* in K562 cells. % InDels is indicated under each guide. F) B2M deletion efficiency of selected gRNAs in primary $CD4^+$ T cells in comparison to 293T cells. Pooled data from 6 independent T-cell donors is shown (mean±SEM). G) CEL Surveyor assay of gRNA crCCR5_A and crCCR5_B targeting *CCR5* in K562 cells and HSPC-mPB. H) Clonal deletion efficiency of gRNA crCCR5_A and crCCR5_B targeting *CCR5* in HSPC-mPB from two donors as determined by Sanger sequencing. (Note: crB2M_14 is not depicted in panel A schematic, as it is located 20 Kb downstream of coding sequence.)

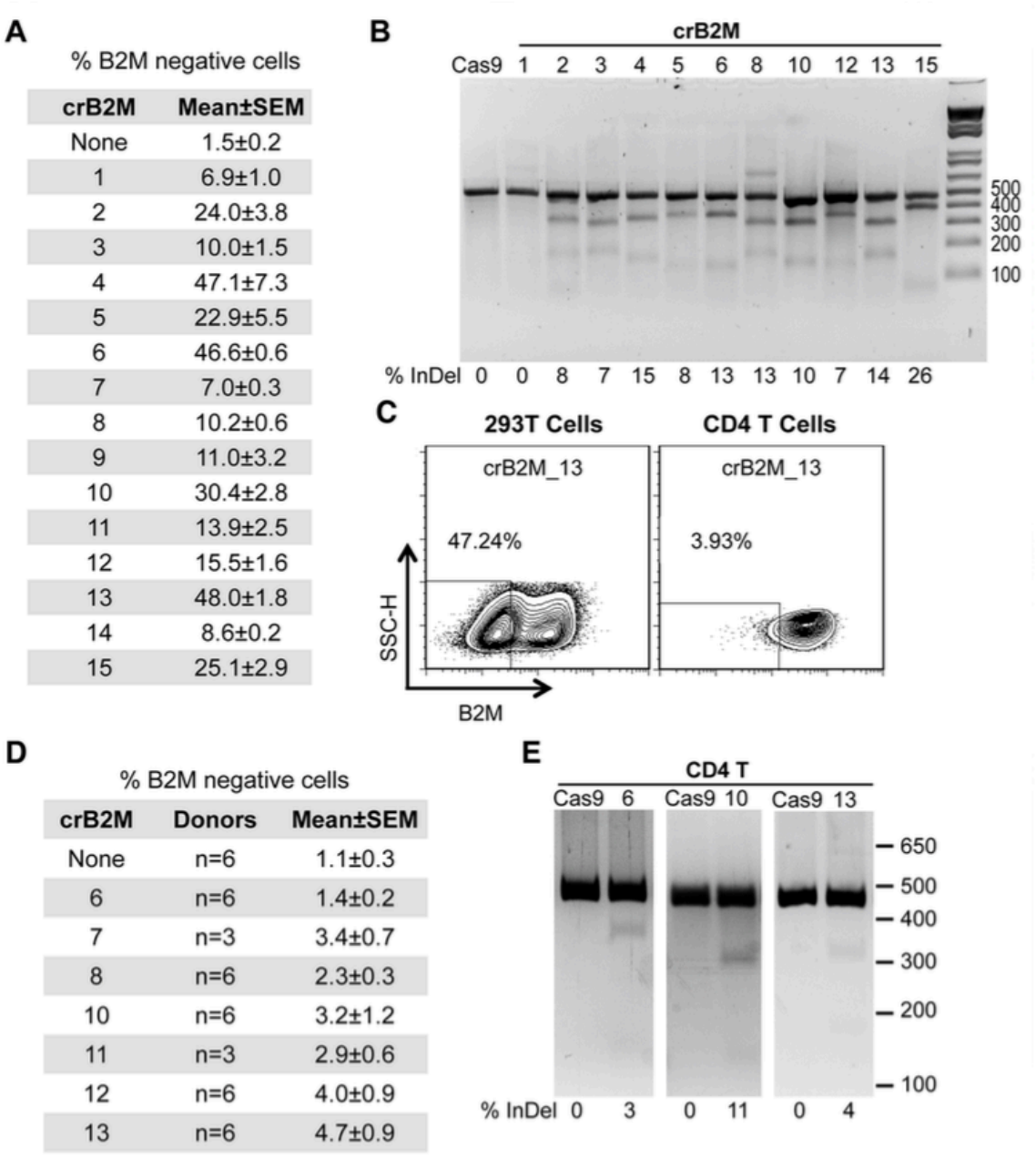


Figure A.4.2. Evaluation of on target mutational efficiencies of various gRNAs targeting *B2M*.

A) B2M deletion efficiency for all gRNAs targeting *B2M* locus in HEK293T cells as measured by flow cytometry. Pooled data from 3 independent experiments shown as mean±SEM. B) B2M deletion efficiencies of selected guides in HEK293T cells, measured as % InDels by CEL Surveyor assay. C) Comparison of B2M surface

(Continued) expression in HEK293T cells and primary $CD4^+$ T cells when transfected with Cas9 and guide crB2M_13. D) D) B2M deletion efficiency for selected guides targeting the *B2M* locus in primary $CD4^+$ T cells, as measured by flow cytometry. E) B2M deletion efficiencies of selected guides in primary CD4+ T cells, measured as % InDels by CEL Surveyor assay.

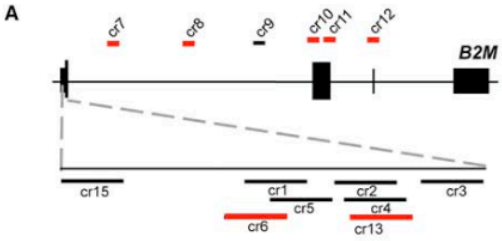
A dual guide strategy allows for predictable gene ablation in primary human hematopoietic cells

Therapeutic translation utilizing CRISPR/Cas9 gene editing in the hematopoietic system would be best enabled if predictable gene editing could be achieved. We reasoned that using two gRNAs directed against the same locus might generate predictable mutations (deletions) over that achieved by reliance on the error-prone non-homologous end joining (NHEJ) pathway, which represents the predominant DNA double strand break (DSB) repair pathway in HSPCs (Beerman et al., 2014). Indeed, this approach has previously been utilized for ZFNs, TALENs and the CRISPR/Cas9 system to achieve predictable deletions (Bauer et al., 2013; Canver et al., 2014; Gupta et al., 2013; Lee et al., 2010; Wang et al., 2014b; Zhou et al., 2014). Six dual gRNA combinations targeting *B2M* (crB2M_10+11, crB2M_10+12, crB2M_6+7, crB2M_8+10, crB2M_13+8, and crB2M_6+8) with DNA sequence lengths between their predicted Cas9 cleavage sites ranging from 81 to 2261 nucleotides were introduced in $CD4^+$ T cells together with Cas9 (Figure A.4.3A). We observed a trend of improved targeting efficacy for most of the tested gRNA pairs and greatly improved efficacy for one gRNA pair (crB2M_13+8), which resulted in 18.0% (±8.35 SEM, n=3) ablation of B2M surface expression (Figures A.4.3B, A.4.3C, and A.4.4A). B2M ablation led to a reduction of MHC class I cell surface

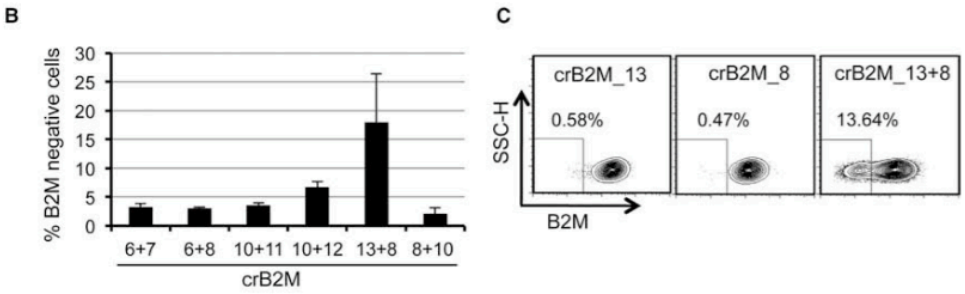
expression consistent with the role of B2M in stabilizing MHC class I (Figure A.4.4B). We further interrogated mutation frequency at a clonal level via a single-cell quantitative PCR (qPCR) approach that revealed that 28.2% (n=301) of $CD4^+$ T cells were homozygous null for *B2M* (Figure A.4.4C). Upon Sanger sequencing across the predicted Cas9 cutting sites, we observed deletion of the intervening sequence as predicted (Figure A.4.4D).

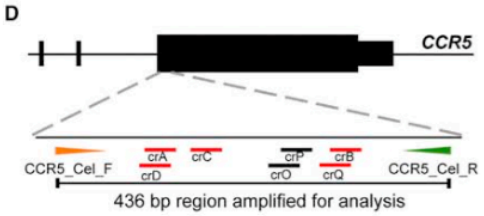
We next applied the dual guide strategy to primary HSPCs by introducing three gRNA pairs (crCCR5_A+B, crCCR5_C+D and crCCR5_D+Q) along with Cas9 into $CD34^+$ HSPCs-mPB (Figure A.4.3D). Sorted $CD34^+$ HSPCs expressing Cas9 were plated into methylcellulose and emergent clonal colonies were individually picked two weeks post-plating for analysis. As the CEL Surveyor assay cannot be utilized to effectively quantify deletion efficacy, individual colonies were analyzed by PCR to quantify the deletion efficacy at one or both *CCR5* alleles (Figures A.4.3D and A.4.3E). Remarkably, although variation in *CCR5* ablation was noted among different donors and gRNA pairs, we consistently observed high monoallelic and biallelic inactivation of *CCR5* in all cases (Figures A.4.3E and A.4.4E). For example, one dual gRNA combination (crCCR5_D+Q) generated biallelic *CCR5* deletion in $CD34^+$ HSPCs-mPB at a rate of 26.8% (±7.1 SEM) across 4 donors (Figures A.4.3E A.4.4E). It should be noted however that the mutation rates determined by this PCR strategy underestimate actual mutation frequency, since small insertions or deletions (InDels) are not detected by this approach. A similar dual gRNA approach targeting *CCR5* (crCCR5_A+B) in $CD4^+$ T cells resulted in a biallelic inactivation rate of 8.8% at the single cell level (n=363 cells analyzed) (Figure A.4.4F). Again, after Sanger sequencing, we noted predictable excision of the DNA between the Cas9 cleavage sites (Figure A.4.4G).

Taken together, these data demonstrate that highly efficacious ablation of clinically relevant genes can be achieved in primary hematopoietic $CD4^+$ T cells and $CD34^+$ HSPCs using a dual guide strategy.

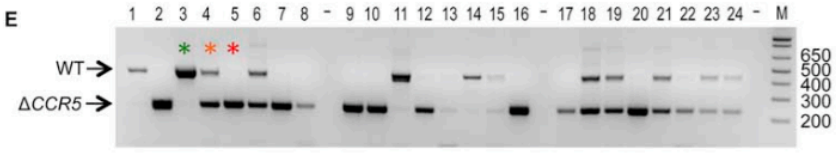


crB2M combination	Offset in bp
10+11	81
10+12	907
6+7	974
8+10	1616
13+8	2222
6+8	2261





crCCR5 combination	Offset in bp
C+D	34
A+B	205
D+Q	205



crCCR5 combination	Donors	#clones	+/- (%mean±SEM)	-/- (%mean±SEM)
A+B	n=4	460	13.5±4.7	19.3±6.8
D+Q	n=4	387	15.0±5.2	26.8±7.1
C+D	n=3	372	19.3±6.5	22.0±8.7

Figure A.4.3. A dual gRNA approach for CRISPR/Cas9 genome editing in primary human hematopoietic stem and effector cells.

A) Schematic of dual gRNA approach for targeting the *B2M* locus. gRNA pairs are in red. The offset in base pairs between actual cutting sites for each gRNA combinations are

(Continued) shown in the right panel. B) B2M deletion efficiency in $CD4^+$ T cells for 6 dual gRNA combinations from three independent donors showing mean ± SEM. C) FACS plots showing loss of B2M expression 72 h after transfection of either single gRNA alone (crB2M_13 or crB2M_8) or in combination (crB2M_13+8) in primary $CD4^+$ T cells. D) Schematic of dual gRNA approach for targeting *CCR5*. gRNA pairs are shown in red. Orange and green arrowheads represent primer pairs used to amplify the targeted region. The offset between the actual cutting sites of each gRNA pair is shown in the right panel. E) Agarose gel electrophoresis image of 24 $CD34^+$ HSPCs-mPB derived clones targeted with gRNAs crCCR5_D+Q analyzed by PCR. Note the deletion of the 205 bp region between the two gRNA cutting sites (top panel; WT: wild type band; Δ*CCR5*: deleted band; green asterisk denotes a WT clone; orange asterisk denotes a heterozygous clone; and red asterisk denotes a homozygous deleted clone). Clonal deletion efficiency for 3 dual gRNA combinations targeting *CCR5* in $CD34^+$ HSPC-mPB (n=4; % mean±SEM; bottom panel).

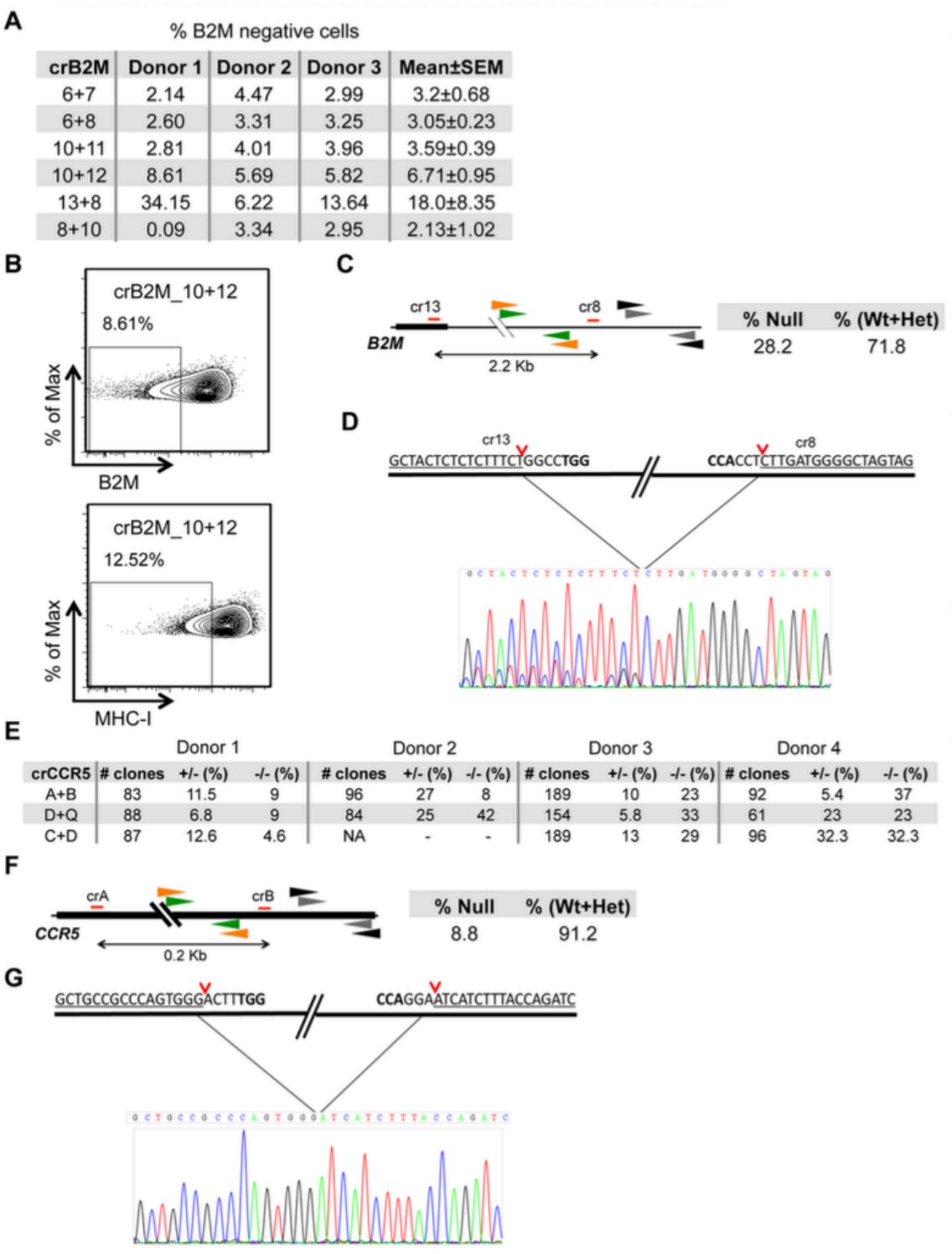


Figure A.4.4. Targeting efficiency of dual gRNA combinations.

A) B2M deletion efficiency for 6 dual gRNA combinations from three independent donors as measured by flow cytometry. B) FACS plots showing loss of MHC class I surface

(Continued) expression (bottom panel) following *B2M* deletion (top panel). C) Schematic of the single cell nested PCR strategy for the *B2M* locus (left panel), black and gray arrowheads: control primer pairs, orange and green arrowheads: primer pairs flanking targeting region. % B2M null single cells is shown (right panel, n=301). D) Sanger sequencing chromatogram showing predicted deletion of targeted region at *B2M* locus. E) Clonal *CCR5* deletion efficiency for three dual gRNA combinations in $CD34^+$ HSPC-mPB obtained from multiple donors. DNA isolated from individual colony was analyzed by PCR and gel electrophoresis. F) Schematic of the single cell nested PCR strategy (left panel) for determining deletion of CCR5 in primary $CD4^+$ T cells. % CCR5 null single cells is shown (right panel, n=363). G) Sanger sequencing chromatogram shows predicted deletion at targeted region.

CRISPR/Cas9 *CCR5*-edited $CD34^+$ HSPCs retain multi-lineage potential

In order to determine whether $CD34^+$ HSPCs that had undergone genome editing with CRISPR/Cas9 retained their potential to differentiate into myeloid and lymphoid effector cells, we performed *in vitro* and *in vivo* differentiation assays. Towards this, *CCR5*-edited $CD34^+$ HSPCs-mPB were plated in methylcellulose and clonal colonies that emerged two weeks post-plating were counted and scored for contribution to granulocyte, macrophage, erythrocyte and megakaryocyte lineages (Bernstein et al., 1991). Comparable colony numbers and colony types were observed regardless of whether single, dual or no gRNAs were used demonstrating that $CD34^+$ HSPC colony forming potential was not impacted by CRISPR/Cas9 (Figure A.4.5A) despite the high monoallelic and biallelic *CCR5* inactivation frequencies observed in these experiments (Figures A.4.1H and A.4.3E).

We next tested the *in vivo* reconstitution potential of HSPCs following CRISPR/Cas9 targeting of *CCR5* by xenotransplantation of control (Cas9-only), and *CCR5* edited (Cas9 + crCCR5 D+Q) CD34$^+$ HSPC-mPB into NOD-*Prkdc*^{Scid}-*IL2rγ*^{null} (NSG) recipients. Recipients were then sacrificed at 12 weeks post-transplant and human hematopoietic cell engraftment (hCD45$^+$) was examined in the bone marrow revealing contribution to CD19$^+$ lymphoid cells and CD11b$^+$ myeloid cells (Figure A.4.5B). Human CD45$^+$ hematopoietic cells were also found in the spleens of transplanted mice (Figure A.4.5C). PCR analysis on DNA isolated from sorted human CD45$^+$ cells from the bone marrow and spleen of reconstituted mice demonstrated that *CCR5* edited cells (ΔCCR5) were robustly contributing to the observed human hematopoietic cell chimerism (Figure A.4.5D). Taken together, these results demonstrate that CRISPR/Cas9 *CCR5*-edited CD34$^+$ HSPCs retained multi-lineage potential *in vitro* and *in vivo*.

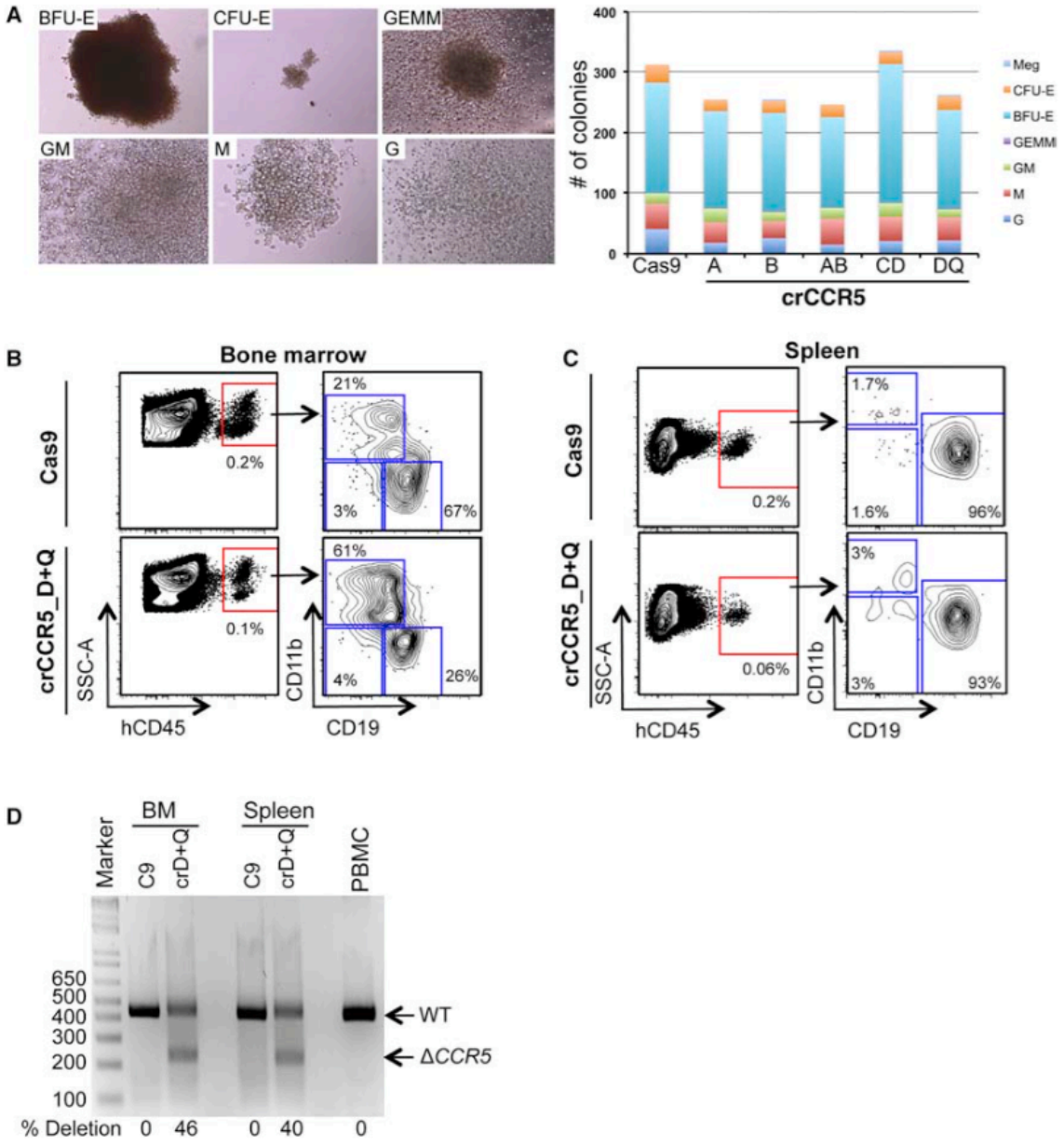


Figure A.4.5. *CCR5*-edited CD34+ HSPCs retain multi-lineage potential.

A) CFC assay results showing colony forming and differentiation potential of $CD34^+$ HSPC-mPB cells after genome editing. Representative pictures of colonies formed in methylcellulose CFC assay (left panel) with quantified data on colony number and types are presented (right panel). Representative FACS plot showing human cell engraftment and multi-lineage reconstitution at 12 weeks post-transplantation in the bone marrow (B)

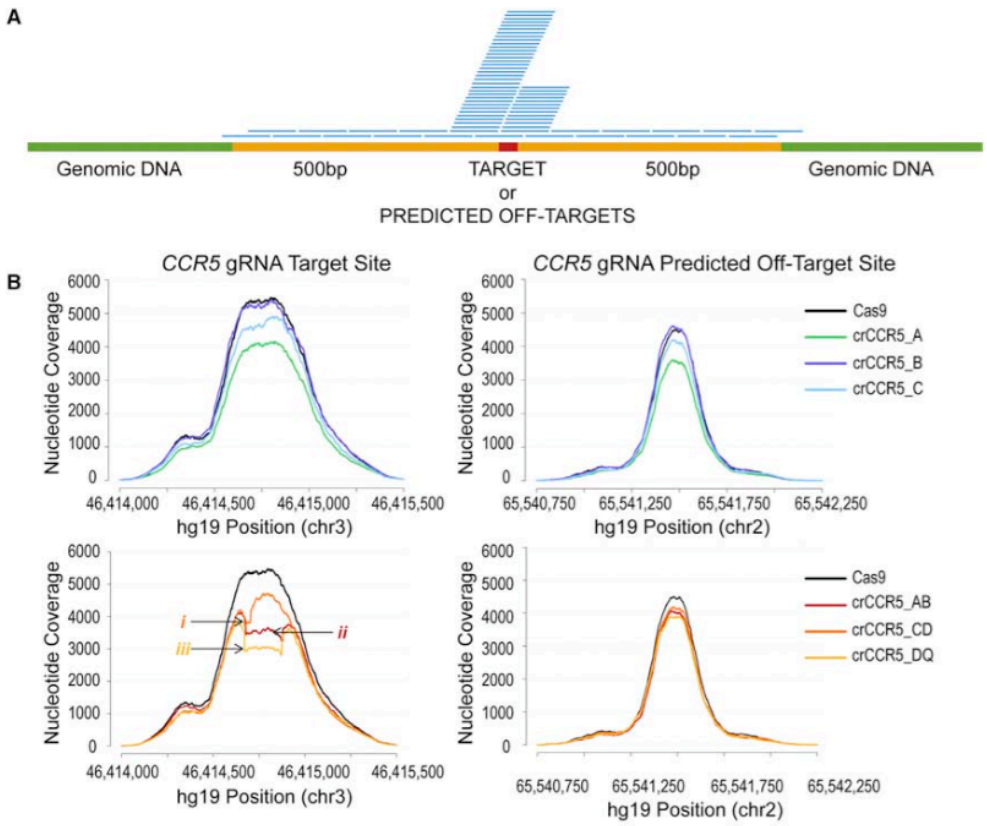
(Continued) and spleen (C) of NSG recipient mice. D) PCR results confirmed predicted deletion of targeted region at *CCR5* locus in human cells sorted from bone marrow and spleen of NSG mice transplanted with CRISPR/Cas9-treated HSPCs. PBMC (human peripheral blood mononuclear cells) from healthy donor taken as control. WT, wild type; Δ*CCR5*, deleted band; BFU-E, Burst Forming Unit-Erythrocyte; CFU-E, Colony Forming Unit-Erythroid; GEMM, Granulocyte/Erythrocyte/Macrophage/Megakaryocyte; GM, Granulocyte/Macrophage

Mutational analysis of genome edited HSPCs by target capture deep sequencing

CRISPR/Cas9 has previously been shown to generate off-target mutations to varying degrees depending upon experimental setting and cell type (Cho et al., 2014; Cradick et al., 2013; Fu et al., 2013; Fu et al., 2014; Hruscha et al., 2013; Lin et al., 2014). To examine this in primary $CD34^+$ HSPCs, we performed target capture sequencing of $CD34^+$ HSPCs-mPB subjected to CRISPR/Cas9 *CCR5*-editing. Our experimental design included capture of each gRNA target site (n=6) and predicted off-target sites (n=172) with expanded capture intervals of 500 base pairs flanking each site to ensure accurate detection of any genetic lesion occurring at or near the selected sites (Figure A.4.6A). We have previously shown that this approach can also capture structural variation breakpoints, such as translocations and inversions, in proximity to the capture site (Talkowski et al., 2011). Sorted $CD34^+$ HSPCs treated with Cas9 alone or in combination with multiple single gRNA (crCCR5_A, crCCR5_B, or crCCR5_C) or dual gRNA combinations (crCCR5_A+B, crCCR5_C+D, or crCCR5_D+Q) were cultured for two weeks and then sequenced to a mean target coverage of 3,390X across each 23 bp gRNA sequence (range 379.6X - 7,969.5X)(Figure A.4.6B). Analysis of the resulting data

revealed highly efficacious on-target mutagenesis with a diverse array of mutated sequence variants observed in both single-gRNA and dual-gRNA treatments (Figure A.4.6C). As expected, we detected small InDels of up to 10bp in addition to varying single nucleotide substitutions at the predicted target sites in the single gRNA libraries. Predicted deletions (i.e. deletions spanning between the two gRNA target sites) were the most common mutations observed (crCCR5_A+B: 19.95%; crCCR5_C+D: 20.45%; crCCR5_D+Q: 42.13%), while small InDels (crCCR5_A+B: 3.06%; crCCR5_C+D: 0.50%; crCCR5_D+Q: 2.95%) were also frequent (Figure A.4.6C). Interestingly, for two dual gRNA combinations (crCCR5_A+B and crCCR5_D+Q) we also observed inversions between the two predicted Cas9 cleavage sites (crCCR5_A+B: 3.06%; crCCR5_D+Q: 2.48%). The most efficacious dual gRNA combination crCCR5_D+Q led to mutations in approximately 48% of the captured sequence reads (Figure A.4.6C).

We next examined the capture sequence reads at predicted off-target sites in the genome. An N-fold enrichment analysis was performed, wherein we compared the total number of non-reference sequencing reads at each predicted off-target site in gRNA treated and control (Cas9 only) samples. This analysis generated a ratio where 1.0 indicates an equivalent number of non-reference sequence reads in both treated and control samples, values less than 1.0 indicate fewer non-reference reads in treated samples, and values greater than 1.0 indicate a greater number of non-reference reads in treated samples (Figure A.4.6D).



C

	crCCR5					
Mutation Class	**A**	**B**	**C**	**AB**	**CD**	**DQ**
Predicted deletion	-	-	-	19.95%	20.45%	42.13%
Small InDel	5.10%	13.54%	13.30%	3.06%	0.50%	2.95%
Inversion	-	-	-	3.12%	-	2.48%
Single Nucleotide Substitution	0.03%	0.02%	0.00%	0.05%	0.15%	0.04%
Others	0.66%	3.09%	3.72%	0.05%	0.08%	0.39%
Total On-Target	**5.79%**	**16.65%**	**17.02%**	**26.23%**	**21.19%**	**47.99%**

D

crCCR5	Off-Target Sites		N-Fold Enrichment of Reads with Variants (Treatment compared to Control)			
	Total Sites	*Significant Sites**	*Off-Target InDel*	*Off-Target Other*	*Off-Target Combined*	***On-Target** Combined***
A	19	0	0.83	1.16	1.06	**39.87**
B	30	1*	1.14*	1.28	1.22	**63**
C	23	0	1.07	0.94	0.96	**23.78**
D	18	0	1.08	0.84	0.88	**40.35**
Q	36	0	0.98	0.78	0.85	**57.66**

Figure A.4.6. Targeted capture and extremely deep sequencing of on-target and predicted off-target sites in CD34^{+} HSPCs.

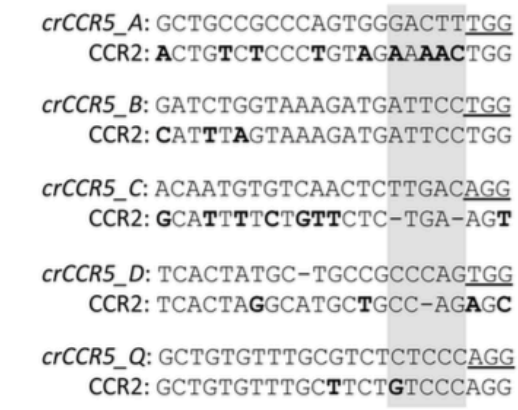
A) Schematic overview of targeted capture and deep sequencing of on-target and predicted off-target sites (red bar). 500 bp flanking cutting site (in yellow) were included in sequence analysis for detection of structural rearrangements, including translocations.

(Continued) Probe sets are indicated in blue. B) Plots showing consistent sequencing depth coverage at both on-target (left panel) and off-target (right panel) sites, achieving a coverage exceeding 3,000x for all on-target sites. Decrease in sequencing depth at the on-target sites in dual-gRNA libraries is marked by arrow, supporting predicted deletions (bottom left; i=35 bp, ii=206 bp, iii=205 bp). C) Precise estimation of on-target mutation allele frequencies by capture sequencing. Notably, the observed rate of effective null mutation exceeds previous estimates by PCR validation of predictable deletions, as smaller InDels and inversions also occur at appreciable frequencies. D) Estimation of mutation frequencies at predicted off-target sites (*One off-target site was statistically different from controls following correction for multiple comparisons; $p \leq 7.6 \times 10^{-11}$). N-fold enrichment is determined based on the ratio of non-reference reads in treated libraries compared to untreated library. Each value represents the average of all off-target sites for a given single gRNA or dual gRNA experiment. Enrichment of 1 is equivalent to baseline (untreated control). **For reference to on-target enrichments, on-target combined represents the proportion of non-reference reads (including single and dual gRNA treatments using a given gRNA) to total reads at on-target sites in treatment compared to control.

Strikingly, our analysis showed that the mean enrichment of mutations at off-target sites in all the gRNA-treated samples compared to control closely conformed to the null hypothesis (i.e., 0.99-fold enrichment compared to controls) indicating that off-target mutation events were extremely rare. Indeed, statistical evaluation of all captured off-target sites yielded a single site (1/172; 0.6%) in the sample treated with gRNA crCCR5_B alone that passed multiple test correction for a statistically significant

enrichment for off-target InDels in the gRNA crCCR5_B treated libraries versus control (gRNA crCCR5_B; $p \leq 7.6x10^{-11}$). When we scrutinized the sequencing reads from the only statistically significant off-target site, which was located in the highly homologous *CCR2* gene (Figure A.4.7A), we found that all sequence variants (36 out of 5,963 total reads) were one or two base InDels, (Figure A.4.7B). Of note, in the other sample in which gRNA crCCR5_B was used (in combination with gRNA crCCR5_A) only 13 out of 5,339 reads supported mutation. However these events did not meet statistical significance above control or samples treated with other gRNAs (Figure A.4.7). Thus, off-target mutagenesis was exceedingly rare and, moreover, the use of two gRNAs in combination did not increase the very low incidence of off-target mutagenesis. We also performed targeted analyses for structural variation at all sites and, though we could easily detect on-target inversions in dual gRNA combination crCCR5_A+B and crCCR5_D+Q, there was no evidence for inversion or translocation at any off-target sites in any of the treatments. These data indicate that on-target mutagenesis efficiency was very high, while off-target mutagenesis was extremely infrequent for both single- and dual gRNA treatments.

A



B

Mutation	crCCR5 treatment								
	B			A+B			A		
	Reads Supporting Mutation	*Total Reads at Site*	*Frequency*	*Reads Supporting Mutation*	*Total Reads at Site*	*Frequency*	*Reads Supporting Mutation*	*Total Reads at Site*	*Frequency*
One Base Insertion	30	5,963	0.50%	2	5,339	0.04%	0	4,678	0.00%
Two Base Insertion	0	5963	0.00%	1	5,339	0.02%	0	4,678	0.00%
One Base Deletion	5	5,963	0.08%	9	5,339	0.17%	4	4,678	0.09%
Two Base Deletion	1	5,963	0.02%	1	5,339	0.02%	0	4,678	0.00%
Total	**36**	5,963	**0.60%**	**13**	**5,339**	**0.24%**	**4**	**4,678**	**0.09%**

Figure A.4.7. Potential off-target sites identified in *CCR5* homologue *CCR2* and analysis of events detected at the single off-target site in which mutagenesis was significantly detected above background.

A) Sequence alignment of *CCR5* gRNAs utilized in this study in relation to the closest homologous sequence in *CCR2* showing mismatched nucleotides in bold. Noteworthy is the fact that guide crCCR5_B, which yielded the sole significantly detected off-target mutagenesis in *CCR2* (detailed in panel B), has 3 nucleotide mismatches, which are distal to the PAM (underlined) and seed (grey box) sequences. **B)** In-depth analyses of all sequence reads at the single off-target site in which mutagenesis was significantly detected above background in both capture libraries treated with the associated gRNA (B; libraries treated with single gRNA crCCR5_B & dual-gRNA crCCR5_A+B), as well as the library treated with gRNA crCCR5_A as a comparison. Total off-target mutation frequency at this site was 0.6% in the single gRNA treatment (crCCR5_B) and notably

(Continued) decreased to 0.24% in the dual gRNA treatment (crCCR5_A+B) in which gRNA plasmid concentration of each gRNA was half of that utilized in single gRNA treatments.

A.5. DISCUSSION

In this study, we utilized the CRISPR/Cas9 system in human primary $CD4^+$ T cells and $CD34^+$ HSPCs to target two clinically relevant genes *B2M* and *CCR5*. To our surprise, the activity of the CRISPR/Cas9 system was remarkably variable in different human cell types, with the same gRNA exhibiting highly efficacious on target mutagenic activity in HEK293T cells but little activity in $CD4^+$ T cells. In contrast, the targeting efficacy in K562 cells and $CD34^+$ HSPCs was comparable.

Moreover, consistent with previous reports (Hsu et al., 2013) we observed that the efficiency of the CRISPR/Cas9 system was gRNA specific, as even gRNAs with partially overlapping sequences within the same exon displayed significantly different targeting efficiencies. Further, a dual gRNA approach yielded increased gene ablation efficacy with certain gRNA pairs in both $CD4^+$ T cells and $CD34^+$ HSPCs leading to predicted deletions at the targeted loci.

The lack of CRISPR/Cas9 activity observed in T cells, especially with single gRNAs, may be due to a number of factors including, inefficient plasmid DNA delivery, the innate immune response of T cells to foreign nucleic acid (Monroe et al., 2014), and/or active DNA repair machinery. Given the efficacy of the CRISPR/Cas9 system in a wide variety of cell types and species both *in vitro* and *in vivo* (Sander and Joung, 2014), the lack of

activity we observed in T cells is likely the exception and not the rule. Nonetheless, our results highlight the fact that CRISPR/Cas9 targeting efficacy can greatly differ between cell lines and primary cells. Ultimately, further studies will be necessary to determine how variable the activity of the CRISPR/Cas9 system is in different primary human cell types.

Our mutational analysis revealed highly efficacious mutagenesis of on-target sites in $CD34^+$ HSPCs. Single gRNAs generated a range of mutations with the vast majority comprised of small InDels. In contrast, dual gRNA combinations largely led to predicted deletions, though a diverse array of mutations including InDels and even inversions were detected. Importantly, we only identified one statistically significant off-target site in the highly homologous *CCR2* gene, which occurred in one out of six experimental settings (gRNA crCCR5_B alone). Sequence analysis of gRNA crCCR5_B in comparison to the identified off-target site in *CCR2* indicated that it perfectly matched in the seed region and contained 3 sequence mismatches at the 5' end of the gRNA sequence (positions 1, 4 and 6). This data is consistent with previous studies showing that mismatches in the 5' proximal end of the gRNA are well tolerated by Cas9 (Lin et al., 2014; Wu et al., 2014). Our data therefore supports the idea that judicious guide design is critical for minimizing off-target mutations. Of note, our very deep sequencing analysis enabled detection of the sole off-target event we describe, whereas sequence analysis performed at lower sequencing depth -- such as 50X coverage that has been used in previous off-target analyses (Smith et al., 2014; Suzuki et al., 2014; Veres et al., 2014) -- would have been unable to detect this event. Overall, our analysis of CRISPR/Cas9 mutational activity in $CD34^+$ HSPCs revealed very high on-target mutation rates and extremely low incidence of off-target mutagenesis.

Utilizing a dual gRNA strategy is uniquely suited for facile deletion of long (multi-kb) portions of the genome with unprecedented precision and efficiency. In theory, gRNAs can be designed to target sites as distant from each other on a chromosome as desired, allowing the deletion of one or even multiple entire genes. Remarkably, the whole 350 kb long human *GBP* locus has been successfully deleted in a human cell line using a dual CRISPR approach (Ohshima et al., 2014). Subsequent studies systematically characterized dual CRISPR-mediated genomic deletions in human cell lines focusing on either the *HPRT* locus or various other loci. In both cases, deletions of 1 Mb were achieved with a frequency of almost 1% (Canver et al., 2014; He et al., 2014). In fact, more recent reports have used CRISPR/Cas9 to edit even larger genomic regions, including an 11 Mb fragment inversion, mimicking a cancer chromosomal translocation (Maddalo et al., 2014), and a 30 Mb deletion on chromosome 19, yielding the first ever fully haploid human cell line (Essletzbichler et al., 2014).

The ability to direct efficient and predictable deletions using dual gRNAs also opens the possibility of using this strategy to target non-coding regions in the genome such as enhancers and silencers that control the expression of disease relevant genes. Recently, a dual CRISPR approach has been used to ablate expression of the oncogene *TAL1* in acute leukemia by disrupting a leukemia-specific enhancer (Mansour et al., 2014). This study suggests that enhancers hold the potential to be highly specific and sensitive targets for perturbation by CRISPR/Cas9 in cancer therapy. A major challenge in the field of cancer therapy is specificity, as the main approaches used, radiotherapy and chemotherapy, affect not only cancer cells but also healthy tissues (Begg et al., 2011; Bouwman and Jonkers, 2012). Utilizing CRISPR/Cas9 could help solve this problem, by

specifically targeting noncoding regions used exclusively by tumor cells (Mansour et al., 2014).

Still, blood disorders will most likely be the first to benefit from clinical translation of a dual gRNA strategy. Recent studies identified regulatory regions (both enhancers and silencers) that control expression of fetal hemoglobin (Bauer et al., 2013), which if deleted increase fetal globin expression in cells otherwise restricted to expressing adult β-globin (Bauer et al., 2013; Xu et al., 2011). Targeted deletion of such regions in $CD34^+$ HSPCs followed by transplantation into patients may provide a durable therapy for the treatment of β-hemoglobinopathies such as sickle cell anemia and β-thalassemia (Xu et al., 2011).

Overall, our data demonstrate that the CRISPR/Cas9 system can be used to ablate genes of clinical significance in primary human $CD4^+$ T cells and $CD34^+$ HSPCs with an efficiency that is therapeutically meaningful for a number of clinical settings, such as the treatment of HIV. Our demonstration that CRISPR/Cas9 targeted $CD34^+$ HSPCs retain multi-lineage potential *in vitro* and *in vivo*, combined with very high on-target and minimal off target mutation rates suggests that CRISPR/Cas9 could have broad applicability enabling novel gene and cell-based therapies of the blood.

REFERENCES

Adams, E.J., and Parham, P. (2001). Species-specific evolution of MHC class I genes in the higher primates. Immunol Rev *183*, 41-64.

Ait-Azzouzene, D., Gendron, M.C., Houdayer, M., Langkopf, A., Burki, K., Nemazee, D., and Kanellopoulos-Langevin, C. (1998). Maternal B lymphocytes specific for paternal histocompatibility antigens are partially deleted during pregnancy. J Immunol *161*, 2677-2683.

Aiuti, A., Biasco, L., Scaramuzza, S., Ferrua, F., Cicalese, M.P., Baricordi, C., Dionisio, F., Calabria, A., Giannelli, S., Castiello, M.C.*, et al.* (2013). Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. Science *341*, 1233151.

al-Lamki, R.S., Skepper, J.N., and Burton, G.J. (1999). Are human placental bed giant cells merely aggregates of small mononuclear trophoblast cells? An ultrastructural and immunocytochemical study. Hum Reprod *14*, 496-504.

Allers, K., Hutter, G., Hofmann, J., Loddenkemper, C., Rieger, K., Thiel, E., and Schneider, T. (2011). Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. Blood *117*, 2791-2799.

Amiot, L., Onno, M., Drenou, B., Monvoisin, C., and Fauchet, R. (1998). HLA-G class I gene expression in normal and malignant hematopoietic cells. Hum Immunol *59*, 524-528.

Amiot, L., Onno, M., Renard, I., Drenou, B., Guillaudeux, T., Le Bouteiller, P., and Fauchet, R. (1996). HLA-G transcription studies during the different stages of normal and malignant hematopoiesis. Tissue Antigens *48*, 609-614.

Andrews, D.M., Sullivan, L.C., Baschuk, N., Chan, C.J., Berry, R., Cotterell, C.L., Lin, J., Halse, H., Watt, S.V., Poursine-Laurent, J.*, et al.* (2012). Recognition of the nonclassical MHC

class I molecule H2-M3 by the receptor Ly49A regulates the licensing and activation of NK cells. Nat Immunol *13*, 1171-1177.

Apps, R., Murphy, S.P., Fernando, R., Gardner, L., Ahad, T., and Moffett, A. (2009). Human leucocyte antigen (HLA) expression of primary trophoblast cells and placental cell lines, determined using single antigen beads to characterize allotype specificities of anti-HLA antibodies. Immunology *127*, 26-39.

Arck, P.C., and Hecher, K. (2013). Fetomaternal immune cross-talk and its consequences for maternal and offspring's health. Nat Med *19*, 548-556.

Arenas-Hernandez, M., Sanchez-Rodriguez, E.N., Mial, T.N., Robertson, S.A., and Gomez-Lopez, N. (2015). Isolation of Leukocytes from the Murine Tissues at the Maternal-Fetal Interface. J Vis Exp, e52866.

Arnaiz-Villena, A., Morales, P., Gomez-Casado, E., Castro, M.J., Varela, P., Rojo-Amigo, R., and Martinez-Laso, J. (1999). Evolution of MHC-G in primates: a different kind of molecule for each group of species. J Reprod Immunol *43*, 111-125.

Arnould, S., Delenda, C., Grizot, S., Desseaux, C., Paques, F., Silva, G.H., and Smith, J. (2011). The I-CreI meganuclease and its engineered derivatives: applications from cell modification to gene therapy. Protein Eng Des Sel *24*, 27-31.

Arruvito, L., Billordo, A., Capucchio, M., Prada, M.E., and Fainboim, L. (2009). IL-6 trans-signaling and the frequency of CD4+FOXP3+ cells in women with reproductive failure. J Reprod Immunol *82*, 158-165.

Avril, T., Iochmann, S., Brand, D., Bardos, P., Watier, H., and Thibault, G. (2003). Human choriocarcinoma cell resistance to natural killer lysis due to defective triggering of natural killer cells. Biol Reprod *69*, 627-633.

Avril, T., Jarousseau, A.C., Watier, H., Boucraut, J., Le Bouteiller, P., Bardos, P., and Thibault, G. (1999). Trophoblast cell line resistance to NK lysis mainly involves an HLA class I-independent mechanism. J Immunol *162*, 5902-5909.

Bainbridge, D., Ellis, S., Le Bouteiller, P., and Sargent, I. (2001). HLA-G remains a mystery. Trends Immunol *22*, 548-552.

Bainbridge, D.R., Ellis, S.A., and Sargent, I.L. (2000). HLA-G suppresses proliferation of CD4(+) T-lymphocytes. J Reprod Immunol *48*, 17-26.

Balkundi, D.R., Hanna, N., Hileb, M., Dougherty, J., and Sharma, S. (2000). Labor-associated changes in Fas ligand expression and function in human placenta. Pediatr Res *47*, 301-308.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299-308.

Barakonyi, A., Kovacs, K.T., Miko, E., Szereday, L., Varga, P., and Szekeres-Bartho, J. (2002). Recognition of nonclassical HLA class I antigens by gamma delta T cells during pregnancy. J Immunol *168*, 2683-2688.

Bauer, D.E., Kamran, S.C., Lessard, S., Xu, J., Fujiwara, Y., Lin, C., Shao, Z., Canver, M.C., Smith, E.C., Pinello, L.*, et al.* (2013). An erythroid enhancer of BCL11A subject to genetic variation determines fetal hemoglobin level. Science *342*, 253-257.

Beerman, I., Seita, J., Inlay, M.A., Weissman, I.L., and Rossi, D.J. (2014). Quiescent Hematopoietic Stem Cells Accumulate DNA Damage during Aging that Is Repaired upon Entry into Cell Cycle. Cell stem cell.

Begay, V., Smink, J., and Leutz, A. (2004). Essential requirement of CCAAT/enhancer binding proteins in embryogenesis. Mol Cell Biol *24*, 9744-9751.

Begg, A.C., Stewart, F.A., and Vens, C. (2011). Strategies to improve radiotherapy with targeted drugs. Nat Rev Cancer *11*, 239-253.

Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell *98*, 387-396.

Bennett, C.L., Christie, J., Ramsdell, F., Brunkow, M.E., Ferguson, P.J., Whitesell, L., Kelly, T.E., Saulsbury, F.T., Chance, P.F., and Ochs, H.D. (2001). The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. Nat Genet *27*, 20-21.

Bernstein, I.D., Andrews, R.G., and Zsebo, K.M. (1991). Recombinant human stem cell factor enhances the formation of colonies by CD34+ and CD34+lin- cells, and the generation of colony-forming cell progeny from CD34+lin- cells cultured with interleukin-3, granulocyte colony-stimulating factor, or granulocyte-macrophage colony-stimulating factor. Blood *77*, 2316-2321.

Biffi, A., Montini, E., Lorioli, L., Cesani, M., Fumagalli, F., Plati, T., Baldoli, C., Martino, S., Calabria, A., Canale, S.*, et al.* (2013). Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. Science *341*, 1233158.

Billingham, R.E., and Beer, A.E. (1984). Reproductive immunology: past, present, and future. Perspect Biol Med *27*, 259-275.

Billington, W.D. (2015). Origins and evolution of reproductive immunology: a personal perspective. J Reprod Immunol *108*, 2-5.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987a). The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. Nature *329*, 512-518.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. (1987b). Structure of the human class I histocompatibility antigen, HLA-A2. Nature *329*, 506-512.

Blaschitz, A., Hutter, H., Leitner, V., Pilz, S., Wintersteiger, R., Dohr, G., and Sedlmayr, P. (2000). Reaction patterns of monoclonal antibodies to HLA-G in human tissues and on cell lines: a comparative study. Hum Immunol *61*, 1074-1085.

Bodmer, J.G., Marsh, S.G., and Albert, E. (1990). Nomenclature for factors of the HLA system, 1989. Immunol Today *11*, 3-10.

Bouwman, P., and Jonkers, J. (2012). The effects of deregulated DNA damage signalling on cancer chemotherapy response and resistance. Nat Rev Cancer *12*, 587-598.

Boyington, J.C., Motyka, S.A., Schuck, P., Brooks, A.G., and Sun, P.D. (2000). Crystal structure of an NK cell immunoglobulin-like receptor in complex with its class I MHC ligand. Nature *405*, 537-543.

Braud, V.M., Allan, D.S., O'Callaghan, C.A., Soderstrom, K., D'Andrea, A., Ogg, G.S., Lazetic, S., Young, N.T., Bell, J.I., Phillips, J.H.*, et al.* (1998). HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. Nature *391*, 795-799.

Brown, J.A., Dorfman, D.M., Ma, F.R., Sullivan, E.L., Munoz, O., Wood, C.R., Greenfield, E.A., and Freeman, G.J. (2003). Blockade of programmed death-1 ligands on dendritic cells enhances T cell activation and cytokine production. J Immunol *170*, 1257-1266.

Brown, R., Kabani, K., Favaloro, J., Yang, S.H., Ho, P.J., Gibson, J., Fromm, P., Suen, H., Woodland, N., Nassif, N.*, et al.* (2012). CD86(+) or HLA-G(+) can be transferred via trogocytosis from myeloma cells to T cells and are associated with poor prognosis. Blood *120*, 2055-2063.

Brunkow, M.E., Jeffery, E.W., Hjerrild, K.A., Paeper, B., Clark, L.B., Yasayko, S.A., Wilkinson, J.E., Galas, D., Ziegler, S.F., and Ramsdell, F. (2001). Disruption of a new forkhead/winged-helix protein, scurfin, results in the fatal lymphoproliferative disorder of the scurfy mouse. Nat Genet *27*, 68-73.

Cantoni, C., Verdiani, S., Falco, M., Pessino, A., Cilli, M., Conte, R., Pende, D., Ponte, M., Mikaelsson, M.S., Moretta, L.*, et al.* (1998). p49, a putative HLA class I-specific inhibitory NK receptor belonging to the immunoglobulin superfamily. Eur J Immunol *28*, 1980-1990.

Canver, M.C., Bauer, D.E., Dass, A., Yien, Y.Y., Chung, J., Masuda, T., Maeda, T., Paw, B.H., and Orkin, S.H. (2014). Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. J Biol Chem *289*, 21312-21324.

Carroll, D. (2011). Genome engineering with zinc-finger nucleases. Genetics *188*, 773-782.

Catano, G., Chykarenko, Z.A., Mangano, A., Anaya, J.M., He, W., Smith, A., Bologna, R., Sen, L., Clark, R.A., Lloyd, A.*, et al.* (2011). Concordance of CCR5 genotypes that influence cell-mediated immunity and HIV-1 disease progression rates. The Journal of infectious diseases *203*, 263-272.

Caumartin, J., Favier, B., Daouya, M., Guillard, C., Moreau, P., Carosella, E.D., and LeMaoult, J. (2007). Trogocytosis-based generation of suppressive NK cells. EMBO J *26*, 1423-1433.

Chen, L., and Flies, D.B. (2013). Molecular mechanisms of T cell co-stimulation and co-inhibition. Nat Rev Immunol *13*, 227-242.

Chen, T., Darrasse-Jeze, G., Bergot, A.S., Courau, T., Churlaud, G., Valdivia, K., Strominger, J.L., Ruocco, M.G., Chaouat, G., and Klatzmann, D. (2013). Self-specific memory regulatory T cells protect embryos at implantation in mice. J Immunol *191*, 2273-2281.

Chen, Y., Bates, D.L., Dey, R., Chen, P.H., Machado, A.C., Laird-Offringa, I.A., Rohs, R., and Chen, L. (2012). DNA binding by GATA transcription factor suggests mechanisms of DNA looping and long-range gene regulation. Cell Rep *2*, 1197-1206.

Cheng, Y.H., and Handwerger, S. (2005). A placenta-specific enhancer of the human syncytin gene. Biol Reprod *73*, 500-509.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nature biotechnology *31*, 230-232.

Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome research *24*, 132-141.

Chu, W., Gao, J., Murphy, W.J., and Hunt, J.S. (1999). A candidate interferon-gamma activated site (GAS element) in the HLA-G promoter does not bind nuclear proteins. Hum Immunol *60*, 1113-1118.

Chumbley, G., King, A., Robertson, K., Holmes, N., and Loke, Y.W. (1994). Resistance of HLA-G and HLA-A2 transfectants to lysis by decidual NK cells. Cell Immunol *155*, 312-322.

Chuong, E.B., Rumi, M.A., Soares, M.J., and Baker, J.C. (2013). Endogenous retroviruses function as species-specific enhancer elements in the placenta. Nat Genet *45*, 325-329.

Ciejek, E.M., Tsai, M.J., and O'Malley, B.W. (1983). Actively transcribed genes are associated with the nuclear matrix. Nature *306*, 607-609.

Clements, C.S., Kjer-Nielsen, L., Kostenko, L., Hoare, H.L., Dunstone, M.A., Moses, E., Freed, K., Brooks, A.G., Rossjohn, J., and McCluskey, J. (2005). Crystal structure of HLA-G: a nonclassical MHC class I molecule expressed at the fetal-maternal interface. Proc Natl Acad Sci U S A *102*, 3360-3365.

Cockerill, P.N., and Garrard, W.T. (1986). Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. Cell *44*, 273-282.

Colonna, M., Brooks, E.G., Falco, M., Ferrara, G.B., and Strominger, J.L. (1993). Generation of allospecific natural killer cells by stimulation across a polymorphism of HLA-C. Science *260*, 1121-1124.

Comiskey, M., Goldstein, C.Y., De Fazio, S.R., Mammolenti, M., Newmark, J.A., and Warner, C.M. (2003). Evidence that HLA-G is the functional homolog of mouse Qa-2, the Ped gene product. Hum Immunol *64*, 999-1004.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A.*, et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.

Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.

Cradick, T.J., Fine, E.J., Antico, C.J., and Bao, G. (2013). CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. Nucleic acids research *41*, 9584-9592.

Croy, B.A., Chen, Z., Hofmann, A.P., Lord, E.M., Sedlacek, A.L., and Gerber, S.A. (2012). Imaging of vascular development in early mouse decidua and its association with leukocytes and trophoblasts. Biol Reprod *87*, 125.

Davis, D.M., Reyburn, H.T., Pazmany, L., Chiu, I., Mandelboim, O., and Strominger, J.L. (1997). Impaired spontaneous endocytosis of HLA-G. Eur J Immunol *27*, 2714-2719.

de Kruijf, E.M., Sajet, A., van Nes, J.G., Natanov, R., Putter, H., Smit, V.T., Liefers, G.J., van den Elsen, P.J., van de Velde, C.J., and Kuppen, P.J. (2010). HLA-E and HLA-G expression in classical HLA class I-negative tumors is of prognostic value for clinical outcome of early breast cancer patients. J Immunol *185*, 7452-7459.

De Plaen, E., Naerhuyzen, B., De Smet, C., Szikora, J.P., and Boon, T. (1997). Alternative promoters of gene MAGE4a. Genomics *40*, 305-313.

de Wet, J.R., Wood, K.V., DeLuca, M., Helinski, D.R., and Subramani, S. (1987). Firefly luciferase gene: structure and expression in mammalian cells. Mol Cell Biol *7*, 725-737.

de Wit, E., and de Laat, W. (2012). A decade of 3C technologies: insights into nuclear organization. Genes Dev *26*, 11-24.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell *149*, 1233-1244.

Deng, W., Rupon, J.W., Krivega, I., Breda, L., Motta, I., Jahn, K.S., Reik, A., Gregory, P.D., Rivella, S., Dean, A.*, et al.* (2014). Reactivation of developmentally silenced globin genes by forced chromatin looping. Cell *158*, 849-860.

Deniz, G., Christmas, S.E., Brew, R., and Johnson, P.M. (1994). Phenotypic and functional cellular differences between human CD3- decidual and peripheral blood leukocytes. J Immunol *152*, 4255-4261.

Diehl, M., Munz, C., Keilholz, W., Stevanovic, S., Holmes, N., Loke, Y.W., and Rammensee, H.G. (1996). Nonclassical HLA-G molecules are classical peptide presenters. Curr Biol *6*, 305-314.

Ding, Q., Lee, Y.K., Schaefer, E.A., Peters, D.T., Veres, A., Kim, K., Kuperwasser, N., Motola, D.L., Meissner, T.B., Hendriks, W.T.*, et al.* (2013a). A TALEN genome-editing system for generating human stem cell-based disease models. Cell Stem Cell *12*, 238-251.

Ding, Q., Regan, S.N., Xia, Y., Oostrom, L.A., Cowan, C.A., and Musunuru, K. (2013b). Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. Cell Stem Cell *12*, 393-394.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Doisne, J.M., Balmas, E., Boulenouar, S., Gaynor, L.M., Kieckbusch, J., Gardner, L., Hawkes, D.A., Barbara, C.F., Sharkey, A.M., Brady, H.J.*, et al.* (2015). Composition, Development, and Function of Uterine Innate Lymphoid Cells. J Immunol *195*, 3937-3945.

Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K.*, et al.* (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell *159*, 374-387.

Drier, Y., Cotton, M.J., Williamson, K.E., Gillespie, S.M., Ryan, R.J., Kluk, M.J., Carey, C.D., Rodig, S.J., Sholl, L.M., Afrogheh, A.H., *et al.* (2016). An oncogenic MYB feedback loop drives alternate cell fates in adenoid cystic carcinoma. Nat Genet *48*, 265-272.

Ellis, S.A., Palmer, M.S., and McMichael, A.J. (1990). Human trophoblast and the choriocarcinoma cell line BeWo express a truncated HLA Class I molecule. J Immunol *144*, 731-735.

Ellis, S.A., Sargent, I.L., Redman, C.W., and McMichael, A.J. (1986). Evidence for a novel HLA antigen found on human extravillous trophoblast and a choriocarcinoma cell line. Immunology *59*, 595-601.

Engleman, E.G., McMichael, A.J., Batey, M.E., and McDevitt, H.O. (1978). A suppressor T cell of the mixed lymphocyte reaction in man specific for the stimulating alloantigen. Evidence that identity at HLA-D between suppressor and responder is required for suppression. J Exp Med *147*, 137-146.

Erlebacher, A. (2013). Immunology of the maternal-fetal interface. Annu Rev Immunol *31*, 387-411.

Essletzbichler, P., Konopka, T., Santoro, F., Chen, D., Gapp, B.V., Kralovics, R., Brummelkamp, T.R., Nijman, S.M., and Burckstummer, T. (2014). Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. Genome Res *24*, 2059-2065.

Fallarino, F., Grohmann, U., Vacca, C., Bianchi, R., Orabona, C., Spreca, A., Fioretti, M.C., and Puccetti, P. (2002). T cell apoptosis by tryptophan catabolism. Cell Death Differ *9*, 1069-1077.

Faust, G.G., and Hall, I.M. (2014). SAMBLASTER: fast duplicate marking and structural variant read extraction. Bioinformatics *30*, 2503-2505.

Ferreira, L.M., Meissner, T.B., Mikkelsen, T.S., Mallard, W., O'Donnell, C.W., Tilburgs, T., Gomes, H.A., Camahort, R., Sherwood, R.I., Gifford, D.K.*, et al.* (2016). A distant trophoblast-specific enhancer controls HLA-G expression at the maternal-fetal interface. Proc Natl Acad Sci U S A.

Flajollet, S., Poras, I., Carosella, E.D., and Moreau, P. (2009). RREB-1 is a transcriptional repressor of HLA-G. J Immunol *183*, 6948-6959.

Fraser, J., Ferrai, C., Chiariello, A.M., Schueler, M., Rito, T., Laudanno, G., Barbieri, M., Moore, B.L., Kraemer, D.C., Aitken, S.*, et al.* (2015). Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. Mol Syst Biol *11*, 852.

Frumento, G., Franchello, S., Palmisano, G.L., Nicotra, M.R., Giacomini, P., Loke, Y.W., Geraghty, D.E., Maio, M., Manzo, C., Natali, P.G.*, et al.* (2000). Melanomas and melanoma cell lines do not express HLA-G, and the expression cannot be induced by gammaIFN treatment. Tissue Antigens *56*, 30-37.

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nature biotechnology *31*, 822-826.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol *32*, 279-284.

Fujita, T., and Fujii, H. (2013). Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin

immunoprecipitation (enChIP) using CRISPR. Biochem Biophys Res Commun *439*, 132-136.

Fujita, T., and Fujii, H. (2014). Identification of proteins associated with an IFNgamma-responsive promoter by a retroviral expression system for enChIP using CRISPR. PLoS One *9*, e103084.

Gaj, T., Gersbach, C.A., and Barbas, C.F., 3rd (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends in biotechnology *31*, 397-405.

Gardner, L., and Moffett, A. (2003). Dendritic cells in the human decidua. Biol Reprod *69*, 1438-1446.

Gasser, S., Orsulic, S., Brown, E.J., and Raulet, D.H. (2005). The DNA damage pathway regulates innate immune system ligands of the NKG2D receptor. Nature *436*, 1186-1190.

Gavrilov, A., Eivazova, E., Priozhkova, I., Lipinski, M., Razin, S., and Vassetzky, Y. (2009). Chromosome conformation capture (from 3C to 5C) and its ChIP-based modification. Methods Mol Biol *567*, 171-188.

Gays, F., Fraser, K.P., Toomey, J.A., Diamond, A.G., Millrain, M.M., Dyson, P.J., and Brooks, C.G. (2001). Functional analysis of the molecular factors controlling Qa1-mediated protection of target cells from NK lysis. Journal of Immunology *166*, 1601-1610.

Genovese, P., Schiroli, G., Escobar, G., Di Tomaso, T., Firrito, C., Calabria, A., Moi, D., Mazzieri, R., Bonini, C., Holmes, M.C.*, et al.* (2014). Targeted genome editing in human repopulating haematopoietic stem cells. Nature.

Geraghty, D.E., Koller, B.H., and Orr, H.T. (1987). A human major histocompatibility complex class I gene that encodes a protein with a shortened cytoplasmic segment. Proc Natl Acad Sci U S A *84*, 9145-9149.

Gershon, R.K., and Kondo, K. (1970). Cell interactions in the induction of tolerance: the role of thymic lymphocytes. Immunology *18*, 723-737.

Gobin, S.J., Biesta, P., de Steenwinkel, J.E., Datema, G., and van den Elsen, P.J. (2002). HLA-G transactivation by cAMP-response element-binding protein (CREB). An alternative transactivation pathway to the conserved major histocompatibility complex (MHC) class I regulatory routes. J Biol Chem *277*, 39525-39531.

Gobin, S.J., Keijsers, V., van Zutphen, M., and van den Elsen, P.J. (1998). The role of enhancer A in the locus-specific transactivation of classical and nonclassical HLA class I genes by nuclear factor kappa B. J Immunol *161*, 2276-2283.

Gobin, S.J., and van den Elsen, P.J. (1999). The regulation of HLA class I expression: is HLA-G the odd one out? Semin Cancer Biol *9*, 55-59.

Gobin, S.J., and van den Elsen, P.J. (2000). Transcriptional regulation of the MHC class Ib genes HLA-E, HLA-F, and HLA-G. Hum Immunol *61*, 1102-1107.

Gobin, S.J., van Zutphen, M., Woltman, A.M., and van den Elsen, P.J. (1999). Transactivation of classical and nonclassical HLA class I genes through the IFN-stimulated response element. J Immunol *163*, 1428-1434.

Golos, T.G., Bondarenko, G.I., Dambaeva, S.V., Breburda, E.E., and Durning, M. (2010). On the role of placental Major Histocompatibility Complex and decidual leukocytes in implantation and pregnancy success using non-human primate models. Int J Dev Biol *54*, 431-443.

Gorer, P.A. (1937). The genetic and antigenic basis of tumor transplantation,. The Journal of Pathology and Bacteriology *44*, 691–697.

Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982). Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol Cell Biol *2*, 1044-1051.

Greger, W.P., and Steele, M.R. (1957). Human fetomaternal passage of erythrocytes. N Engl J Med *256*, 158-161.

Groschel, S., Sanders, M.A., Hoogenboezem, R., de Wit, E., Bouwman, B.A., Erpelinck, C., van der Velden, V.H., Havermans, M., Avellino, R., van Lom, K.*, et al.* (2014). A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. Cell *157*, 369-381.

Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W.*, et al.* (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature *453*, 948-951.

Guidry, P.A., and Stroynowski, I. (2005). The murine family of gut-restricted class Ib MHC includes alternatively spliced isoforms of the proposed HLA-G homolog, "blastocyst MHC". J Immunol *175*, 5248-5259.

Guleria, I., Khosroshahi, A., Ansari, M.J., Habicht, A., Azuma, M., Yagita, H., Noelle, R.J., Coyle, A., Mellor, A.L., Khoury, S.J.*, et al.* (2005). A critical role for the programmed death ligand 1 in fetomaternal tolerance. J Exp Med *202*, 231-237.

Guleria, I., and Sayegh, M.H. (2007). Maternal acceptance of the fetus: true human tolerance. J Immunol *178*, 3345-3351.

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y.*, et al.* (2015). CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. Cell *162*, 900-910.

Gupta, A., Hall, V.L., Kok, F.O., Shin, M., McNulty, J.C., Lawson, N.D., and Wolfe, S.A. (2013). Targeted chromosomal deletions and inversions in zebrafish. Genome research *23*, 1008-1017.

Gustafsson, C., Mjosberg, J., Matussek, A., Geffers, R., Matthiesen, L., Berg, G., Sharma, S., Buer, J., and Ernerudh, J. (2008). Gene expression profiling of human decidual macrophages: evidence for immunosuppressive phenotype. PLoS One *3*, e2078.

Hanna, J., Goldman-Wohl, D., Hamani, Y., Avraham, I., Greenfield, C., Natanson-Yaron, S., Prus, D., Cohen-Daniel, L., Arnon, T.I., Manaster, I.*, et al.* (2006). Decidual NK cells regulate key developmental processes at the human fetal-maternal interface. Nat Med *12*, 1065-1074.

Harrison, G.A., Humphrey, K.E., Jakobsen, I.B., and Cooper, D.W. (1993). A 14 bp deletion polymorphism in the HLA-G gene. Hum Mol Genet *2*, 2200.

He, Z., Proudfoot, C., Mileham, A.J., McLaren, D.G., Whitelaw, C.B., and Lillico, S.G. (2014). Highly efficient targeted chromosome deletions using CRISPR/Cas9. Biotechnol Bioeng.

Heinz, S., Romanoski, C.E., Benner, C., and Glass, C.K. (2015). The selection and function of cell type-specific enhancers. Nat Rev Mol Cell Biol *16*, 144-154.

Heng, H.H., Goetze, S., Ye, C.J., Liu, G., Stevens, J.B., Bremer, S.W., Wykes, S.M., Bode, J., and Krawetz, S.A. (2004). Chromatin loops are selectively anchored using scaffold/matrix-attachment regions. J Cell Sci *117*, 999-1008.

Herzenberg, L.A., and Gonzales, B. (1962). Appearance of H-2 agglutinins in outcrossed female mice. Proc Natl Acad Sci U S A *48*, 570-573.

Hesselberth, J.R., Chen, X., Zhang, Z., Sabo, P.J., Sandstrom, R., Reynolds, A.P., Thurman, R.E., Neph, S., Kuehn, M.S., Noble, W.S.*, et al.* (2009). Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. Nat Methods *6*, 283-289.

Heuchel, R., Matthias, P., and Schaffner, W. (1989). Two closely spaced promoters are equally activated by a remote enhancer: evidence against a scanning model for enhancer action. Nucleic Acids Res *17*, 8931-8947.

Heyborne, K., Fu, Y.X., Nelson, A., Farr, A., O'Brien, R., and Born, W. (1994). Recognition of trophoblasts by gamma delta T cells. J Immunol *153*, 2918-2926.

Heyborne, K.D., Cranfill, R.L., Carding, S.R., Born, W.K., and O'Brien, R.L. (1992). Characterization of gamma delta T lymphocytes at the maternal-fetal interface. J Immunol *149*, 2872-2878.

Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A.*, et al.* (2016). Activation of proto-oncogenes by disruption of chromosome neighborhoods. Science *351*, 1454-1458.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C.*, et al.* (2011). Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol *29*, 731-734.

Holets, L.M., Hunt, J.S., and Petroff, M.G. (2006). Trophoblast CD274 (B7-H1) is differentially expressed across gestation: influence of oxygen concentration. Biol Reprod *74*, 352-358.

Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G.M., Kohn, D.B., Gregory, P.D., Holmes, M.C.*, et al.* (2010). Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nature biotechnology *28*, 839-847.

Holwerda, S., and de Laat, W. (2012). Chromatin loops, gene positioning, and gene expression. Front Genet *3*, 217.

Honig, A., Rieger, L., Kapp, M., Sutterlin, M., Dietl, J., and Kammerer, U. (2004). Indoleamine 2,3-dioxygenase (IDO) expression in invasive extravillous trophoblast supports role of the enzyme for materno-fetal tolerance. J Reprod Immunol *61*, 79-86.

Houser, B.L., Tilburgs, T., Hill, J., Nicotra, M.L., and Strominger, J.L. (2011). Two unique human decidual macrophage populations. J Immunol *186*, 2633-2642.

Hruscha, A., Krawitz, P., Rechenberg, A., Heinrich, V., Hecht, J., Haass, C., and Schmid, B. (2013). Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. Development *140*, 4982-4987.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262-1278.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O*., et al.* (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol *31*, 827-832.

Huang, J.F., Yang, Y., Sepulveda, H., Shi, W., Hwang, I., Peterson, P.A., Jackson, M.R., Sprent, J., and Cai, Z. (1999). TCR-Mediated internalization of peptide-MHC complexes acquired by T cells. Science *286*, 952-954.

Hunt, J.S., and Petroff, M.G. (2013). IFPA Senior Award Lecture: Reproductive immunology in perspective--reprogramming at the maternal-fetal interface. Placenta *34 Suppl*, S52-55.

Hunt, J.S., Vassmer, D., Ferguson, T.A., and Miller, L. (1997). Fas ligand is positioned in mouse uterus and placenta to prevent trafficking of activated leukocytes between the mother and the conceptus. J Immunol *158*, 4122-4128.

Hutter, G., Nowak, D., Mossner, M., Ganepola, S., Mussig, A., Allers, K., Schneider, T., Hofmann, J., Kucherer, C., Blau, O.*, et al.* (2009). Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. The New England journal of medicine *360*, 692-698.

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R., and Joung, J.K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. Nature biotechnology *31*, 227-229.

Ibrahim, E.C., Morange, M., Dausset, J., Carosella, E.D., and Paul, P. (2000). Heat shock and arsenite induce expression of the nonclassical class I histocompatibility HLA-G gene in tumor cell lines. Cell Stress Chaperones *5*, 207-218.

Ikeno, M., Suzuki, N., Kamiya, M., Takahashi, Y., Kudoh, J., and Okazaki, T. (2012). LINE1 family member is negative regulator of HLA-G expression. Nucleic Acids Res *40*, 10742-10752.

Jeremias, I., Herr, I., Boehler, T., and Debatin, K.M. (1998). TRAIL/Apo-2-ligand-induced apoptosis in human T cells. Eur J Immunol *28*, 143-152.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816-821.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. Elife *2*, e00471.

John, S., Sabo, P.J., Thurman, R.E., Sung, M.H., Biddie, S.C., Johnson, T.A., Hager, G.L., and Stamatoyannopoulos, J.A. (2011). Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. Nat Genet *43*, 264-268.

Jokhi, P.P., King, A., Sharkey, A.M., Smith, S.K., and Loke, Y.W. (1994). Screening for cytokine messenger ribonucleic acids in purified human decidual lymphocyte populations by the reverse-transcriptase polymerase chain reaction. J Immunol *153*, 4427-4435.

Joung, J.K., and Sander, J.D. (2013). TALENs: a widely applicable technology for targeted genome editing. Nature reviews Molecular cell biology *14*, 49-55.

Kabadi, A.M., Ousterout, D.G., Hilton, I.B., and Gersbach, C.A. (2014). Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. Nucleic Acids Res *42*, e147.

Kam, E.P., Gardner, L., Loke, Y.W., and King, A. (1999). The role of trophoblast in the physiological change in decidual spiral arteries. Hum Reprod *14*, 2131-2138.

Kelley, D., and Rinn, J. (2012). Transposable elements reveal a stem cell-specific class of long noncoding RNAs. Genome Biol *13*, R107.

Kellum, R., and Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. Cell *64*, 941-950.

Keskin, D.B., Allan, D.S., Rybalov, B., Andzelm, M.M., Stern, J.N., Kopcow, H.D., Koopman, L.A., and Strominger, J.L. (2007). TGFbeta promotes conversion of CD16+ peripheral blood NK cells into CD16- NK cells with similarities to decidual NK cells. Proc Natl Acad Sci U S A *104*, 3378-3383.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Kind, J., Pagie, L., de Vries, S.S., Nahidiazar, L., Dey, S.S., Bienko, M., Zhan, Y., Lajoie, B., de Graaf, C.A., Amendola, M., *et al.* (2015). Genome-wide maps of nuclear lamina interactions in single human cells. Cell *163*, 134-147.

King, A., Allan, D.S., Bowen, M., Powis, S.J., Joseph, S., Verma, S., Hiby, S.E., McMichael, A.J., Loke, Y.W., and Braud, V.M. (2000). HLA-E is expressed on trophoblast and interacts with CD94/NKG2 receptors on decidual NK cells. Eur J Immunol *30*, 1623-1631.

King, A., and Loke, Y.W. (1990). Human trophoblast and JEG choriocarcinoma cells are sensitive to lysis by IL-2-stimulated decidual NK cells. Cell Immunol *129*, 435-448.

Kitaya, K., Yasuda, J., Yagi, I., Tada, Y., Fushiki, S., and Honjo, H. (2000). IL-15 expression at human endometrium and decidua. Biol Reprod *63*, 683-687.

Kobayashi, K.S., and van den Elsen, P.J. (2012). NLRC5: a key regulator of MHC class I-dependent immune responses. Nat Rev Immunol *12*, 813-820.

Koller, B.H., Geraghty, D.E., DeMars, R., Duvick, L., Rich, S.S., and Orr, H.T. (1989). Chromosomal organization of the human major histocompatibility complex class I gene family. J Exp Med *169*, 469-480.

Koopman, L.A., Kopcow, H.D., Rybalov, B., Boyson, J.E., Orange, J.S., Schatz, F., Masch, R., Lockwood, C.J., Schachter, A.D., Park, P.J., *et al.* (2003). Human decidual natural killer cells are a unique NK cell subset with immunomodulatory potential. J Exp Med *198*, 1201-1212.

Kopcow, H.D., Allan, D.S., Chen, X., Rybalov, B., Andzelm, M.M., Ge, B., and Strominger, J.L. (2005). Human decidual NK cells form immature activating synapses and are not cytotoxic. Proc Natl Acad Sci U S A *102*, 15563-15568.

Kosak, S.T., Skok, J.A., Medina, K.L., Riblet, R., Le Beau, M.M., Fisher, A.G., and Singh, H. (2002). Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science *296*, 158-162.

Kourtis, A.P., Read, J.S., and Jamieson, D.J. (2014). Pregnancy and infection. N Engl J Med *370*, 2211-2218.

Kovats, S., Main, E.K., Librach, C., Stubblebine, M., Fisher, S.J., and DeMars, R. (1990). A class I antigen, HLA-G, expressed in human trophoblasts. Science *248*, 220-223.

Krasnow, J.S., Tollerud, D.J., Naus, G., and DeLoia, J.A. (1996). Endometrial Th2 cytokine expression throughout the menstrual cycle and early pregnancy. Hum Reprod *11*, 1747-1754.

Kudo, Y., Boyd, C.A., Sargent, I.L., and Redman, C.W. (2003). Decreased tryptophan catabolism by placental indoleamine 2,3-dioxygenase in preeclampsia. Am J Obstet Gynecol *188*, 719-726.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology *10*, R25.

Layer, R.M., Chiang, C., Quinlan, A.R., and Hall, I.M. (2014). LUMPY: a probabilistic framework for structural variant discovery. Genome biology *15*, R84.

Le Discorde, M., Le Danff, C., Moreau, P., Rouas-Freiss, N., and Carosella, E.D. (2005). HLA-G*0105N null allele encodes functional HLA-G isoforms. Biol Reprod *73*, 280-288.

Le Gal, F.A., Riteau, B., Sedlik, C., Khalil-Daher, I., Menier, C., Dausset, J., Guillet, J.G., Carosella, E.D., and Rouas-Freiss, N. (1999). HLA-G-mediated inhibition of antigen-specific cytotoxic T lymphocytes. Int Immunol *11*, 1351-1356.

Lee, H.J., Kim, E., and Kim, J.S. (2010). Targeted chromosomal deletions in human cells using zinc finger nucleases. Genome Res *20*, 81-89.

Lee, N., Goodlett, D.R., Ishitani, A., Marquardt, H., and Geraghty, D.E. (1998). HLA-E surface expression depends on binding of TAP-dependent peptides derived from certain HLA class I signal sequences. J Immunol *160*, 4951-4960.

Lefebvre, S., Berrih-Aknin, S., Adrian, F., Moreau, P., Poea, S., Gourand, L., Dausset, J., Carosella, E.D., and Paul, P. (2001). A specific interferon (IFN)-stimulated response element of the distal HLA-G promoter binds IFN-regulatory factor 1 and mediates enhancement of this nonclassical class I gene by IFN-beta. J Biol Chem *276*, 6133-6139.

LeMaoult, J., Caumartin, J., Daouya, M., Favier, B., Le Rond, S., Gonzalez, A., and Carosella, E.D. (2007). Immune regulation by pretenders: cell-to-cell transfers of HLA-G make effector T cells act as regulatory cells. Blood *109*, 2040-2048.

LeMaoult, J., Zafaranloo, K., Le Danff, C., and Carosella, E.D. (2005). HLA-G up-regulates ILT2, ILT3, ILT4, and KIR2DL4 in antigen presenting cells, NK cells, and T cells. FASEB J *19*, 662-664.

Li, C., Houser, B.L., Nicotra, M.L., and Strominger, J.L. (2009). HLA-G homodimer-induced cytokine secretion through HLA-G receptors on human decidual macrophages and natural killer cells. Proc Natl Acad Sci U S A *106*, 5767-5772.

Li, X.J., Zhang, X., Lin, A., Ruan, Y.Y., and Yan, W.H. (2012). Human leukocyte antigen-G (HLA-G) expression in cervical cancer lesions is associated with disease progression. Hum Immunol *73*, 946-949.

Lin, H., Mosmann, T.R., Guilbert, L., Tuntipopipat, S., and Wegmann, T.G. (1993). Synthesis of T helper 2-type cytokines at the maternal-fetal interface. J Immunol *151*, 4562-4573.

Lin, Y., Cradick, T.J., Brown, M.T., Deshmukh, H., Ranjan, P., Sarode, N., Wile, B.M., Vertino, P.M., Stewart, F.J., and Bao, G. (2014). CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic acids research *42*, 7473-7485.

Link, N., Kurtz, P., O'Neal, M., Garcia-Hughes, G., and Abrams, J.M. (2013). A p53 enhancer region regulates target genes through chromatin conformations in cis and in trans. Genes Dev *27*, 2433-2438.

Lissauer, D., Piper, K., Goodyear, O., Kilby, M.D., and Moss, P.A. (2012). Fetal-specific CD8+ cytotoxic T cell responses develop during normal human pregnancy and exhibit broad functional capacity. J Immunol *189*, 1072-1080.

Little, C.C. (1924). The genetics of tissue transplantation in mammals. J Cancer Res *8*, 75-95.

Llano, M., Lee, N., Navarro, F., Garcia, P., Albar, J.P., Geraghty, D.E., and Lopez-Botet, M. (1998). HLA-E-bound peptides influence recognition by inhibitory and triggering CD94/NKG2 receptors: preferential response to an HLA-G-derived nonamer. Eur J Immunol *28*, 2854-2863.

Lonfat, N., and Duboule, D. (2015). Structure, function and evolution of topologically associating domains (TADs) at HOX loci. FEBS Lett *589*, 2869-2876.

Lonfat, N., Montavon, T., Darbellay, F., Gitto, S., and Duboule, D. (2014). Convergent evolution of complex regulatory landscapes and pleiotropy at Hox loci. Science *346*, 1004-1006.

Lynch, V.J., Leclerc, R.D., May, G., and Wagner, G.P. (2011). Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. Nat Genet *43*, 1154-1159.

Ma, G.T., and Linzer, D.I. (2000). GATA-2 restricts prolactin-like protein A expression to secondary trophoblast giant cells in the mouse. Biol Reprod *63*, 570-574.

Maddalo, D., Manchado, E., Concepcion, C.P., Bonetti, C., Vidigal, J.A., Han, Y.C., Ogrodowski, P., Crippa, A., Rekhtman, N., de Stanchina, E.*, et al.* (2014). In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. Nature *516*, 423-427.

Majumder, P., Gomez, J.A., and Boss, J.M. (2006). The human major histocompatibility complex class II HLA-DRB1 and HLA-DQA1 genes are separated by a CTCF-binding enhancer-blocking element. J Biol Chem *281*, 18435-18443.

Majumder, P., Gomez, J.A., Chadwick, B.P., and Boss, J.M. (2008). The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. J Exp Med *205*, 785-798.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.

Manaster, I., Goldman-Wohl, D., Greenfield, C., Nachmani, D., Tsukerman, P., Hamani, Y., Yagel, S., and Mandelboim, O. (2012). MiRNA-mediated control of HLA-G expression and function. PLoS One *7*, e33395.

Mandal, P.K., Ferreira, L.M., Collins, R., Meissner, T.B., Boutwell, C.L., Friesen, M., Vrbanac, V., Garrison, B.S., Stortchevoi, A., Bryder, D.*, et al.* (2014). Efficient Ablation of Genes in Human Hematopoietic Stem and Effector Cells using CRISPR/Cas9. Cell Stem Cell *15*, 643-652.

Mandelboim, O., Reyburn, H.T., Vales-Gomez, M., Pazmany, L., Colonna, M., Borsellino, G., and Strominger, J.L. (1996). Protection from lysis by natural killer cells of group 1 and 2

specificity is mediated by residue 80 in human histocompatibility leukocyte antigen C alleles and also occurs with empty major histocompatibility complex molecules. J Exp Med *184*, 913-922.

Manley, J.L., Fire, A., Cano, A., Sharp, P.A., and Gefter, M.L. (1980). DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. Proc Natl Acad Sci U S A *77*, 3855-3859.

Mansour, M.R., Abraham, B.J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A.D., Etchin, J., Lawton, L., Sallan, S.E., Silverman, L.B.*, et al.* (2014). Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. Science *346*, 1373-1377.

Martinson, J.J., Chapman, N.H., Rees, D.C., Liu, Y.T., and Clegg, J.B. (1997). Global distribution of the CCR5 gene 32-basepair deletion. Nature genetics *16*, 100-103.

Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K.*, et al.* (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res *34*, D108-110.

Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J.*, et al.* (2012). Systematic localization of common disease-associated variation in regulatory DNA. Science *337*, 1190-1195.

Medawar, P.B. (1953). Some immunological and endocrinological problems raised by the evolution of viviparity in vertebrates. Symp Soc Exp Biol *44*, 320-338.

Meissner, T.B., Li, A., Biswas, A., Lee, K.H., Liu, Y.J., Bayir, E., Iliopoulos, D., van den Elsen, P.J., and Kobayashi, K.S. (2010). NLR family member NLRC5 is a transcriptional regulator of MHC class I genes. Proc Natl Acad Sci U S A *107*, 13794-13799.

Meissner, T.B., Li, A., and Kobayashi, K.S. (2012a). NLRC5: a newly discovered MHC class I transactivator (CITA). Microbes Infect *14*, 477-484.

Meissner, T.B., Liu, Y.J., Lee, K.H., Li, A., Biswas, A., van Eggermond, M.C., van den Elsen, P.J., and Kobayashi, K.S. (2012b). NLRC5 cooperates with the RFX transcription factor complex to induce MHC class I gene expression. J Immunol *188*, 4951-4958.

Meissner, T.B., Mandal, P.K., Ferreira, L.M., Rossi, D.J., and Cowan, C.A. (2014). Genome editing for human gene therapy. Methods Enzymol *546*, 273-295.

Melnikov, A., Murugan, A., Zhang, X., Tesileanu, T., Wang, L., Rogov, P., Feizi, S., Gnirke, A., Callan, C.G., Jr., Kinney, J.B.*, et al.* (2012). Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. Nat Biotechnol *30*, 271-277.

Melo, C.A., Drost, J., Wijchers, P.J., van de Werken, H., de Wit, E., Oude Vrielink, J.A., Elkon, R., Melo, S.A., Leveille, N., Kalluri, R.*, et al.* (2013). eRNAs are required for p53-dependent enhancer activity and gene transcription. Mol Cell *49*, 524-535.

Miller, J.D., Weber, D.A., Ibegbu, C., Pohl, J., Altman, J.D., and Jensen, P.E. (2003). Analysis of HLA-E peptide-binding specificity and contact residues in bound peptide required for recognition by CD94/NKG2. J Immunol *171*, 1369-1375.

Mincheva-Nilsson, L., Baranov, V., Yeung, M.M., Hammarstrom, S., and Hammarstrom, M.L. (1994). Immunomorphologic studies of human decidua-associated lymphoid cells in normal early pregnancy. J Immunol *152*, 2020-2032.

Mirkovitch, J., Mirault, M.E., and Laemmli, U.K. (1984). Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. Cell *39*, 223-232.

Moffett, A., and Loke, C. (2006). Immunology of placentation in eutherian mammals. Nat Rev Immunol *6*, 584-594.

Moffett-King, A. (2002). Natural killer cells and pregnancy. Nat Rev Immunol *2*, 656-663.

Monroe, K.M., Yang, Z., Johnson, J.R., Geng, X., Doitsh, G., Krogan, N.J., and Greene, W.C. (2014). IFI16 DNA sensor is required for death of lymphoid CD4 T cells abortively infected with HIV. Science *343*, 428-432.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E.*, et al.* (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet *34*, 267-273.

Moreau, P., Flajollet, S., and Carosella, E.D. (2009). Non-classical transcriptional regulation of HLA-G: an update. J Cell Mol Med *13*, 2973-2989.

Moreau, P., Lefebvre, S., Gourand, L., Dausset, J., Carosella, E.D., and Paul, P. (1998). Specific binding of nuclear factors to the HLA-G gene promoter correlates with a lack of HLA-G transcripts in first trimester human fetal liver. Hum Immunol *59*, 751-757.

Moreau, P., Paul, P., Gourand, L., Prost, S., Dausset, J., Carosella, E., and Kirszenbaum, M. (1997). HLA-G gene transcriptional regulation in trophoblasts and blood cells: differential binding of nuclear factors to a regulatory element located 1.1 kb from exon 1. Hum Immunol *52*, 41-46.

Mosser, D.M., and Edwards, J.P. (2008). Exploring the full spectrum of macrophage activation. Nat Rev Immunol *8*, 958-969.

Mueller-Storm, H.P., Sogo, J.M., and Schaffner, W. (1989). An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. Cell *58*, 767-777.

Munn, D.H., Zhou, M., Attwood, J.T., Bondarev, I., Conway, S.J., Marshall, B., Brown, C., and Mellor, A.L. (1998). Prevention of allogeneic fetal rejection by tryptophan catabolism. Science *281*, 1191-1193.

Naji, A., Menier, C., Morandi, F., Agaugue, S., Maki, G., Ferretti, E., Bruel, S., Pistoia, V., Carosella, E.D., and Rouas-Freiss, N. (2014). Binding of HLA-G to ITIM-bearing Ig-like transcript 2 receptor suppresses B cell responses. J Immunol *192*, 1536-1546.

Nancy, P., and Erlebacher, A. (2014). T cell behavior at the maternal-fetal interface. Int J Dev Biol *58*, 189-198.

Nancy, P., Tagliani, E., Tay, C.S., Asp, P., Levy, D.E., and Erlebacher, A. (2012). Chemokine gene silencing in decidual stromal cells limits T cell access to the maternal-fetal interface. Science *336*, 1317-1321.

Natarajan, K., Dimasi, N., Wang, J., Mariuzza, R.A., and Margulies, D.H. (2002). Structure and function of natural killer cell receptors: multiple molecular solutions to self, nonself discrimination. Annu Rev Immunol *20*, 853-885.

Navarro, F., Llano, M., Bellon, T., Colonna, M., Geraghty, D.E., and Lopez-Botet, M. (1999). The ILT2(LIR1) and CD94/NKG2A NK cell receptors respectively recognize HLA-G1 and HLA-E molecules co-expressed on target cells. Eur J Immunol *29*, 277-283.

Nehar-Belaid, D., Courau, T., Derian, N., Florez, L., Ruocco, M.G., and Klatzmann, D. (2016). Regulatory T Cells Orchestrate Similar Immune Evasion of Fetuses and Tumors in Mice. J Immunol *196*, 678-690.

Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., Kang, Y., Zhao, X., Si, W., Li, W., *et al.* (2014). Generation of Gene-Modified Cynomolgus Monkey via Cas9/RNA-Mediated Gene Targeting in One-Cell Embryos. Cell *156*, 836-843.

Noordermeer, D., de Wit, E., Klous, P., van de Werken, H., Simonis, M., Lopez-Jones, M., Eussen, B., de Klein, A., Singer, R.H., and de Laat, W. (2011). Variegated gene expression caused by cell-specific long-range DNA interactions. Nat Cell Biol *13*, 944-951.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., *et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.

Northcott, P.A., Lee, C., Zichner, T., Stutz, A.M., Erkek, S., Kawauchi, D., Shih, D.J., Hovestadt, V., Zapatka, M., Sturm, D., *et al.* (2014). Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. Nature *511*, 428-434.

O'Brien, M., McCarthy, T., Jenkins, D., Paul, P., Dausset, J., Carosella, E.D., and Moreau, P. (2001). Altered HLA-G transcription in pre-eclampsia is associated with allele specific inheritance: possible role of the HLA-G gene in susceptibility to the disease. Cell Mol Life Sci *58*, 1943-1949.

O'Garra, A., and Arai, N. (2000). The molecular basis of T helper 1 and T helper 2 cell differentiation. Trends Cell Biol *10*, 542-550.

Ohshima, J., Lee, Y., Sasai, M., Saitoh, T., Su Ma, J., Kamiyama, N., Matsuura, Y., Pann-Ghill, S., Hayashi, M., Ebisu, S., *et al.* (2014). Role of mouse and human autophagy proteins in IFN-gamma-induced cell-autonomous responses against Toxoplasma gondii. J Immunol *192*, 3328-3335.

Orr, H.T., Bach, F.H., Ploegh, H.L., Strominger, J.L., Kavathas, P., and DeMars, R. (1982). Use of HLA loss mutants to analyse the structure of the human major histocompatibility complex. Nature *296*, 454-456.

Pan, Y., Xiao, L., Li, A.S., Zhang, X., Sirois, P., Zhang, J., and Li, K. (2013). Biological and biomedical applications of engineered nucleases. Molecular biotechnology *55*, 54-62.

Parham, P. (1996). Immunology: keeping mother at bay. Curr Biol *6*, 638-641.

Patel, D.M., Arnold, P.Y., White, G.A., Nardella, J.P., and Mannie, M.D. (1999). Class II MHC/peptide complexes are released from APC and are acquired by T cell responders during specific antigen recognition. J Immunol *163*, 5201-5210.

Pattillo, R.A., and Gey, G.O. (1968). The establishment of a cell line of human hormone-synthesizing trophoblastic cells in vitro. Cancer Res *28*, 1231-1236.

Pattillo, R.A., Hussa, R.O., Huang, W.Y., Delfs, E., and Mattingly, R.F. (1972). Estrogen production by trophoblastic tumors in tissue culture. J Clin Endocrinol *34*, 59-61.

Paul, P., Cabestre, F.A., Le Gal, F.A., Khalil-Daher, I., Le Danff, C., Schmid, M., Mercier, S., Avril, M.F., Dausset, J., Guillet, J.G.*, et al.* (1999). Heterogeneity of HLA-G gene transcription and protein expression in malignant melanoma biopsies. Cancer Res *59*, 1954-1960.

Paul, P., Rouas-Freiss, N., Khalil-Daher, I., Moreau, P., Riteau, B., Le Gal, F.A., Avril, M.F., Dausset, J., Guillet, J.G., and Carosella, E.D. (1998). HLA-G expression in melanoma: a way for tumor cells to escape from immunosurveillance. Proc Natl Acad Sci U S A *95*, 4510-4515.

Pazmany, L., Mandelboim, O., Vales-Gomez, M., Davis, D.M., Reyburn, H.T., and Strominger, J.L. (1996). Protection from natural killer cell-mediated lysis by HLA-G expression on target cells. Science *274*, 792-795.

Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K.A., Liu, O., Wang, N., Lee, G., Bartsevich, V.V., Lee, Y.L*., et al.* (2008). Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nature biotechnology *26*, 808-816.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. Cell *137*, 1194-1211.

Phillips, T.A., Ni, J., Pan, G., Ruben, S.M., Wei, Y.F., Pace, J.L., and Hunt, J.S. (1999). TRAIL (Apo-2L) and TRAIL receptors in human placentas: implications for immune privilege. J Immunol *162*, 6053-6059.

Piccinni, M.P., Beloni, L., Livi, C., Maggi, E., Scarselli, G., and Romagnani, S. (1998). Defective production of both leukemia inhibitory factor and type 2 T-helper cytokines by decidual T cells in unexplained recurrent abortions. Nat Med *4*, 1020-1024.

Ponte, M., Cantoni, C., Biassoni, R., Tradori-Cappai, A., Bentivoglio, G., Vitale, C., Bertone, S., Moretta, A., Moretta, L., and Mingari, M.C. (1999). Inhibitory receptors sensing HLA-G1 molecules in pregnancy: decidua-associated natural killer cells express LIR-1 and CD94/NKG2A and acquire p49, an HLA-G1-specific receptor. Proc Natl Acad Sci U S A *96*, 5674-5679.

Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. Nature *322*, 697-701.

Quach, K., Grover, S.A., Kenigsberg, S., and Librach, C.L. (2014). A combination of single nucleotide polymorphisms in the 3'untranslated region of HLA-G is associated with preeclampsia. Hum Immunol *75*, 1163-1170.

Rajagopalan, S., and Long, E.O. (1999). A human histocompatibility leukocyte antigen (HLA)-G-specific receptor expressed on all natural killer cells. J Exp Med *189*, 1093-1100.

Rajagopalan, S., and Long, E.O. (2012). Cellular senescence induced by CD158d reprograms natural killer cells to promote vascular remodeling. Proc Natl Acad Sci U S A *109*, 20596-20601.

Rajagopalan, S., Moyle, M.W., Joosten, I., and Long, E.O. (2010). DNA-PKcs controls an endosomal signaling pathway for a proinflammatory response by natural killer cells. Sci Signal *3*, ra14.

Real, L.M., Cabrera, T., Collado, A., Jimenez, P., Garcia, A., Ruiz-Cabello, F., and Garrido, F. (1999). Expression of HLA G in human tumors is not a frequent event. Int J Cancer *81*, 512-518.

Reddy, K.L., Zullo, J.M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. Nature *452*, 243-247.

Redman, C.W., McMichael, A.J., Stirrat, G.M., Sunderland, C.A., and Ting, A. (1984). Class 1 major histocompatibility complex antigens on human extra-villous trophoblast. Immunology *52*, 457-468.

Riolobos, L., Hirata, R.K., Turtle, C.J., Wang, P.R., Gornalusse, G.G., Zavajlevski, M., Riddell, S.R., and Russell, D.W. (2013). HLA engineering of human pluripotent stem cells. Molecular therapy : the journal of the American Society of Gene Therapy *21*, 1232-1241.

Rouas-Freiss, N., Goncalves, R.M., Menier, C., Dausset, J., and Carosella, E.D. (1997a). Direct evidence to support the role of HLA-G in protecting the fetus from maternal uterine natural killer cytolysis. Proc Natl Acad Sci U S A *94*, 11520-11525.

Rouas-Freiss, N., Marchal, R.E., Kirszenbaum, M., Dausset, J., and Carosella, E.D. (1997b). The alpha1 domain of HLA-G1 and HLA-G2 inhibits cytotoxicity induced by natural killer cells: is HLA-G the public ligand for natural killer cell inhibitory receptors? Proc Natl Acad Sci U S A *94*, 5249-5254.

Rouas-Freiss, N., Moreau, P., Ferrone, S., and Carosella, E.D. (2005). HLA-G proteins in cancer: do they provide tumor cells with an escape mechanism? Cancer Res *65*, 10139-10144.

Rouet, P., Smih, F., and Jasin, M. (1994). Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci U S A *91*, 6064-6068.

Rousseau, P., Le Discorde, M., Mouillot, G., Marcou, C., Carosella, E.D., and Moreau, P. (2003). The 14 bp deletion-insertion polymorphism in the 3' UT region of the HLA-G gene influences HLA-G mRNA stability. Hum Immunol *64*, 1005-1010.

Rousseau, P., Paul, P., O'Brien, M., Dausset, J., Carosella, E.D., and Moreau, P. (2000). The X1 box of HLA-G promoter is a target site for RFX and Sp1 factors. Hum Immunol *61*, 1132-1137.

Saito, S., and Sakai, M. (2003). Th1/Th2 balance in preeclampsia. J Reprod Immunol *59*, 161-173.

Samson, M., Libert, F., Doranz, B.J., Rucker, J., Liesnard, C., Farber, C.M., Saragosti, S., Lapoumeroulie, C., Cognaux, J., Forceille, C*., et al.* (1996). Resistance to HIV-1

infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. Nature *382*, 722-725.

Samstein, R.M., Josefowicz, S.Z., Arvey, A., Treuting, P.M., and Rudensky, A.Y. (2012). Extrathymic generation of regulatory T cells in placental mammals mitigates maternal-fetal conflict. Cell *150*, 29-38.

Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol *32*, 347-355.

Sanjana, N.E., Shalem, O., and Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods *11*, 783-784.

Santillan, M.K., Pelham, C.J., Ketsawatsomkron, P., Santillan, D.A., Davis, D.R., Devor, E.J., Gibson-Corley, K.N., Scroggins, S.M., Grobe, J.L., Yang, B.*, et al.* (2015). Pregnant mice lacking indoleamine 2,3-dioxygenase exhibit preeclampsia phenotypes. Physiol Rep *3*.

Scharenberg, A.M., Duchateau, P., and Smith, J. (2013). Genome engineering with TAL-effector nucleases and alternative modular nuclease technologies. Current gene therapy *13*, 291-303.

Schlesinger, M. (1962). Uterus of rodents as site for manifestation of transplantation immunity against transplantable tumors. J Natl Cancer Inst *28*, 927-945.

Schmidt, C.M., Ehlenfeldt, R.G., Athanasiou, M.C., Duvick, L.A., Heinrichs, H., David, C.S., and Orr, H.T. (1993). Extraembryonic expression of the human MHC class I gene HLA-G in transgenic mice. Evidence for a positive regulatory region located 1 kilobase 5' to the start site of transcription. J Immunol *151*, 2633-2645.

Schreiber, R.D., Old, L.J., and Smyth, M.J. (2011). Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. Science *331*, 1565-1570.

Schust, D.J., Hill, A.B., and Ploegh, H.L. (1996). Herpes simplex virus blocks intracellular transport of HLA-G in placentally derived human cells. J Immunol *157*, 3375-3380.

Searle, R.F., and Wren, A.M. (1992). Antigen presenting capacity of human decidua: no evidence for human decidual antigen presenting cell mediated immunoregulation. J Reprod Immunol *21*, 211-221.

Seitz, C., Uchanska-Ziegler, B., Zank, A., and Ziegler, A. (1998). The monoclonal antibody HCA2 recognises a broadly shared epitope on selected classical as well as several non-classical HLA class I molecules. Mol Immunol *35*, 819-827.

Seligman, L.M., Chisholm, K.M., Chevalier, B.S., Chadsey, M.S., Edwards, S.T., Savage, J.H., and Veillet, A.L. (2002). Mutations altering the cleavage specificity of a homing endonuclease. Nucleic Acids Res *30*, 3870-3879.

Sexton, T., and Cavalli, G. (2015). The role of chromosome domains in shaping the functional genome. Cell *160*, 1049-1059.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458-472.

Sherwood, R.I., Hashimoto, T., O'Donnell, C.W., Lewis, S., Barkal, A.A., van Hoff, J.P., Karun, V., Jaakkola, T., and Gifford, D.K. (2014). Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. Nat Biotechnol *32*, 171-178.

Shih, I.M., and Kurman, R.J. (1996). Expression of melanoma cell adhesion molecule in intermediate trophoblast. Lab Invest *75*, 377-388.

Shiroishi, M., Tsumoto, K., Amano, K., Shirakihara, Y., Colonna, M., Braud, V.M., Allan, D.S., Makadzange, A., Rowland-Jones, S., Willcox, B.*, et al.* (2003). Human inhibitory receptors Ig-like transcript 2 (ILT2) and ILT4 compete with CD8 for MHC class I binding and bind preferentially to HLA-G. Proc Natl Acad Sci U S A *100*, 8856-8861.

Silva, G., Poirot, L., Galetto, R., Smith, J., Montoya, G., Duchateau, P., and Paques, F. (2011). Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. Current gene therapy *11*, 11-27.

Simmons, R.L., and Russell, P.S. (1962). The antigenicity of mouse trophoblast. Ann N Y Acad Sci *99*, 717-732.

Simpson, E., Chandler, P., and Pole, D. (1981). A model of T-cell unresponsiveness using the male-specific antigen, H-Y. Cell Immunol *62*, 251-257.

Sindram-Trujillo, A.P., Scherjon, S.A., van Hulst-van Miert, P.P., Kanhai, H.H., Roelen, D.L., and Claas, F.H. (2004). Comparison of decidual leukocytes following spontaneous vaginal delivery and elective cesarean section in uncomplicated human term pregnancy. J Reprod Immunol *62*, 125-137.

Slukvin, II, Lunn, D.P., Watkins, D.I., and Golos, T.G. (2000). Placental expression of the nonclassical MHC class I molecule Mamu-AG at implantation in the rhesus monkey. Proc Natl Acad Sci U S A *97*, 9104-9109.

Smith, C., Gore, A., Yan, W., Abalde-Atristain, L., Li, Z., He, C., Wang, Y., Brodsky, R.A., Zhang, K., Cheng, L.*, et al.* (2014). Whole-genome sequencing analysis reveals high

specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. Cell Stem Cell *15*, 12-13.

Smith, E., and Shilatifard, A. (2014). Enhancer biology and enhanceropathies. Nat Struct Mol Biol *21*, 210-219.

Smith, R.N., and Powell, A.E. (1977). The adoptive transfer of pregnancy-induced unresponsiveness to male skin grafts with thymus-dependent cells. J Exp Med *146*, 899-904.

Smith, S.D., Dunk, C.E., Aplin, J.D., Harris, L.K., and Jones, R.L. (2009). Evidence for immune cell involvement in decidual spiral arteriole remodeling in early human pregnancy. Am J Pathol *174*, 1959-1971.

Snell, G.D. (1948). Methods for the study of histocompatibility genes. J Genet *49*, 87-108.

Solier, C., Mallet, V., Lenfant, F., Bertrand, A., Huchenq, A., and Le Bouteiller, P. (2001). HLA-G unique promoter region: functional implications. Immunogenetics *53*, 617-625.

Somerset, D.A., Zheng, Y., Kilby, M.D., Sansom, D.M., and Drayson, M.T. (2004). Normal human pregnancy is associated with an elevation in the immune suppressive CD25+ CD4+ regulatory T-cell subset. Immunology *112*, 38-43.

Starzl, T.E., and Zinkernagel, R.M. (2001). Transplantation tolerance from a historical perspective. Nat Rev Immunol *1*, 233-239.

Steimle, V., Siegrist, C.A., Mottet, A., Lisowska-Grospierre, B., and Mach, B. (1994). Regulation of MHC class II expression by interferon-gamma mediated by the transactivator gene CIITA. Science *265*, 106-109.

Stenqvist, A.C., Nagaeva, O., Baranov, V., and Mincheva-Nilsson, L. (2013). Exosomes secreted by human placenta carry functional Fas ligand and TRAIL molecules and convey

apoptosis in activated immune cells, suggesting exosome-mediated immune privilege of the fetus. J Immunol *191*, 5515-5523.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Sun, L., Goff, L.A., Trapnell, C., Alexander, R., Lo, K.A., Hacisuleyman, E., Sauvageau, M., Tazon-Vega, B., Kelley, D.R., Hendrickson, D.G.*, et al.* (2013). Long noncoding RNAs regulate adipogenesis. Proc Natl Acad Sci U S A *110*, 3387-3392.

Suzuki, K., Yu, C., Qu, J., Li, M., Yao, X., Yuan, T., Goebl, A., Tang, S., Ren, R., Aizawa, E.*, et al.* (2014). Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. Cell stem cell *15*, 31-36.

Swets, M., Konig, M.H., Zaalberg, A., Dekker-Ensink, N.G., Gelderblom, H., van de Velde, C.J., van den Elsen, P.J., and Kuppen, P.J. (2016). HLA-G and classical HLA class I expression in primary colorectal cancer and associated liver metastases. Hum Immunol.

Talkowski, M.E., Ernst, C., Heilbut, A., Chiang, C., Hanscom, C., Lindgren, A., Kirby, A., Liu, S., Muddukrishna, B., Ohsumi, T.K.*, et al.* (2011). Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research. American journal of human genetics *88*, 469-481.

Tanabe, M., Sekimata, M., Ferrone, S., and Takiguchi, M. (1992). Structural and functional analysis of monomorphic determinants recognized by monoclonal antibodies reacting with the HLA class I alpha 3 domain. J Immunol *148*, 3202-3209.

Tebas, P., Stein, D., Tang, W.W., Frank, I., Wang, S.Q., Lee, G., Spratt, S.K., Surosky, R.T., Giedlin, M.A., Nichol, G.*, et al.* (2014). Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med *370*, 901-910.

Tilburgs, T., Claas, F.H., and Scherjon, S.A. (2010). Elsevier Trophoblast Research Award Lecture: Unique properties of decidual T cells and their role in immune regulation during human pregnancy. Placenta *31 Suppl*, S82-86.

Tilburgs, T., Crespo, A.C., van der Zwan, A., Rybalov, B., Raj, T., Stranger, B., Gardner, L., Moffett, A., and Strominger, J.L. (2015a). Human HLA-G+ extravillous trophoblasts: Immune-activating cells that interact with decidual leukocytes. Proc Natl Acad Sci U S A *112*, 7219-7224.

Tilburgs, T., Evans, J.H., Crespo, A.C., and Strominger, J.L. (2015b). The HLA-G cycle provides for both NK tolerance and immunity at the maternal-fetal interface. Proc Natl Acad Sci U S A *112*, 13312-13317.

Tilburgs, T., Roelen, D.L., van der Mast, B.J., de Groot-Swings, G.M., Kleijburg, C., Scherjon, S.A., and Claas, F.H. (2008). Evidence for a selective migration of fetus-specific CD4+CD25bright regulatory T cells from the peripheral blood to the decidua in human pregnancy. J Immunol *180*, 5737-5745.

Tilburgs, T., van der Mast, B.J., Nagtzaam, N.M., Roelen, D.L., Scherjon, S.A., and Claas, F.H. (2009). Expression of NK cell receptors on decidual T cells in human pregnancy. J Reprod Immunol *80*, 22-32.

Tong, Q., Tsai, J., Tan, G., Dalgin, G., and Hotamisligil, G.S. (2005). Interaction between GATA and the C/EBP family of transcription factors is critical in GATA-mediated suppression of adipocyte differentiation. Mol Cell Biol *25*, 706-715.

Topalian, S.L., Drake, C.G., and Pardoll, D.M. (2015). Immune checkpoint blockade: a common denominator approach to cancer therapy. Cancer Cell *27*, 450-461.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol *31*, 46-53.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Trkola, A., Dragic, T., Arthos, J., Binley, J.M., Olson, W.C., Allaway, G.P., Cheng-Mayer, C., Robinson, J., Maddon, P.J., and Moore, J.P. (1996). CD4-dependent, antibody-sensitive interactions between HIV-1 and its co-receptor CCR-5. Nature *384*, 184-187.

Uckan, D., Steele, A., Cherry, Wang, B.Y., Chamizo, W., Koutsonikolis, A., Gilbert-Barness, E., and Good, R.A. (1997). Trophoblasts express Fas ligand: a proposed mechanism for immune privilege in placenta and maternal invasion. Mol Hum Reprod *3*, 655-662.

Urnov, F.D., Miller, J.C., Lee, Y.L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D., and Holmes, M.C. (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature *435*, 646-651.

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., and Gregory, P.D. (2010). Genome editing with engineered zinc finger nucleases. Nature reviews Genetics *11*, 636-646.

Vacca, P., Montaldo, E., Croxatto, D., Loiacono, F., Canegallo, F., Venturini, P.L., Moretta, L., and Mingari, M.C. (2015). Identification of diverse innate lymphoid cells in human decidua. Mucosal Immunol *8*, 254-264.

Vacca, P., Vitale, C., Montaldo, E., Conte, R., Cantoni, C., Fulcheri, E., Darretta, V., Moretta, L., and Mingari, M.C. (2011). CD34+ hematopoietic precursors are present in human decidua and differentiate into natural killer cells upon interaction with stromal cells. Proc Natl Acad Sci U S A *108*, 2402-2407.

Vacchio, M.S., and Hodes, R.J. (2005). Fetal expression of Fas ligand is necessary and sufficient for induction of CD8 T cell tolerance to the fetal antigen H-Y during pregnancy. J Immunol *174*, 4657-4661.

van de Lagemaat, L.N., Landry, J.R., Mager, D.L., and Medstrand, P. (2003). Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. Trends Genet *19*, 530-536.

van der Burg, M., Kreyenberg, H., Willasch, A., Barendregt, B.H., Preuner, S., Watzinger, F., Lion, T., Roosnek, E., Harvey, J., Alcoceba, M.*, et al.* (2011). Standardization of DNA isolation from low cell numbers for chimerism analysis by PCR of short tandem repeats. Leukemia *25*, 1467-1470.

Vantourout, P., and Hayday, A. (2013). Six-of-the-best: unique contributions of gammadelta T cells to immunology. Nat Rev Immunol *13*, 88-100.

Veres, A., Gosis, B.S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., Erdin, S., Talkowski, M.E., and Musunuru, K. (2014). Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. Cell stem cell *15*, 27-30.

Verma, S., Hiby, S.E., Loke, Y.W., and King, A. (2000). Human decidual natural killer cells express the receptor for and respond to the cytokine interleukin 15. Biol Reprod *62*, 959-968.

Vietri Rudan, M., and Hadjur, S. (2015). Genetic Tailors: CTCF and Cohesin Shape the Genome During Evolution. Trends Genet *31*, 651-660.

Vignali, D.A., Collison, L.W., and Workman, C.J. (2008). How regulatory T cells work. Nat Rev Immunol *8*, 523-532.

Vince, G.S., Starkey, P.M., Jackson, M.C., Sargent, I.L., and Redman, C.W. (1990). Flow cytometric characterisation of cell populations in human pregnancy decidua and isolation of decidual macrophages. J Immunol Methods *132*, 181-189.

Walker, J.A., Barlow, J.L., and McKenzie, A.N. (2013). Innate lymphoid cells--how did we miss them? Nat Rev Immunol *13*, 75-87.

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell *153*, 910-918.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research *38*, e164.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014a). Genetic screens in human cells using the CRISPR-Cas9 system. Science *343*, 80-84.

Wang, X., Wang, Y., Huang, H., Chen, B., Chen, X., Hu, J., Chang, T., Lin, R.J., and Yee, J.K. (2014b). Precise gene modification mediated by TALEN and single-stranded oligodeoxynucleotides in human cells. PloS one *9*, e93575.

Wegmann, T.G., Lin, H., Guilbert, L., and Mosmann, T.R. (1993). Bidirectional cytokine interactions in the maternal-fetal relationship: is successful pregnancy a TH2 phenomenon? Immunol Today *14*, 353-356.

Wei, C., Liu, J., Yu, Z., Zhang, B., Gao, G., and Jiao, R. (2013). TALEN or Cas9 - rapid, efficient and specific choices for genome modifications. Journal of genetics and genomics = Yi chuan xue bao *40*, 281-289.

Weil, P.A., Luse, D.S., Segall, J., and Roeder, R.G. (1979). Selective and accurate initiation of transcription at the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. Cell *18*, 469-484.

Weissman, I.L. (2000). Translating stem and progenitor cell biology to the clinic: barriers and opportunities. Science *287*, 1442-1446.

Wiendl, H., Mitsdoerffer, M., Hofmeister, V., Wischhusen, J., Bornemann, A., Meyermann, R., Weiss, E.H., Melms, A., and Weller, M. (2002). A functional role of HLA-G expression in human gliomas: an alternative strategy of immune escape. J Immunol *168*, 4772-4780.

Winter, C.C., and Long, E.O. (1997). A single amino acid in the p58 killer cell inhibitory receptor controls the ability of natural killer cells to discriminate between the two groups of HLA-C allotypes. J Immunol *158*, 4026-4028.

Wu, X., Scott, D.A., Kriz, A.J., Chiu, A.C., Hsu, P.D., Dadon, D.B., Cheng, A.W., Trevino, A.E., Konermann, S., Chen, S.*, et al.* (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol *32*, 670-676.

Xu, J., Peng, C., Sankaran, V.G., Shao, Z., Esrick, E.B., Chong, B.G., Ippolito, G.C., Fujiwara, Y., Ebert, B.L., Tucker, P.W.*, et al.* (2011). Correction of sickle cell disease in adult mice by interference with fetal hemoglobin silencing. Science *334*, 993-996.

Xu, Y., Plazyo, O., Romero, R., Hassan, S.S., and Gomez-Lopez, N. (2015). Isolation of Leukocytes from the Human Maternal-fetal Interface. J Vis Exp, e52863.

Yan, W.H., and Fan, L.A. (2005). Residues Met76 and Gln79 in HLA-G alpha1 domain involve in KIR2DL4 recognition. Cell Res *15*, 176-182.

Ye, L., Wang, J., Beyer, A.I., Teque, F., Cradick, T.J., Qi, Z., Chang, J.C., Bao, G., Muench, M.O., Yu, J.*, et al.* (2014). Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Delta32 mutation confers resistance to HIV infection. Proceedings of the National Academy of Sciences of the United States of America *111*, 9591-9596.

Yelavarthi, K.K., Schmidt, C.M., Ehlenfeldt, R.G., Orr, H.T., and Hunt, J.S. (1993). Cellular distribution of HLA-G mRNA in transgenic mouse placentas. J Immunol *151*, 3638-3645.

Yie, S.M., Li, L.H., Li, G.M., Xiao, R., and Librach, C.L. (2006a). Progesterone enhances HLA-G gene expression in JEG-3 choriocarcinoma cells and human cytotrophoblasts in vitro. Hum Reprod *21*, 46-51.

Yie, S.M., Xiao, R., and Librach, C.L. (2006b). Progesterone regulates HLA-G gene expression through a novel progesterone response element. Hum Reprod *21*, 2538-2544.

Yokoyama, W.M. (1997). The mother-child union: the case of missing-self and protection of the fetus. Proc Natl Acad Sci U S A *94*, 5998-6000.

Yuhki, N., Beck, T., Stephens, R.M., Nishigaki, Y., Newmann, K., and O'Brien, S.J. (2003). Comparative genome organization of human, murine, and feline MHC class II region. Genome Res *13*, 1169-1179.

Zhang, J., Chen, Z., Smith, G.N., and Croy, B.A. (2011). Natural killer cell-triggered vascular transformation: maternal care before birth? Cell Mol Immunol *8*, 1-11.

Zhang, Y.H., Tian, M., Tang, M.X., Liu, Z.Z., and Liao, A.H. (2015). Recent Insight into the Role of the PD-1/PD-L1 Pathway in Feto-Maternal Tolerance and Pregnancy. Am J Reprod Immunol *74*, 201-208.

Zhao, L., Teklemariam, T., and Hantash, B.M. (2012). Reassessment of HLA-G isoform specificity of MEM-G/9 and 4H84 monoclonal antibodies. Tissue Antigens *80*, 231-238.

Zheng, Y., Josefowicz, S., Chaudhry, A., Peng, X.P., Forbush, K., and Rudensky, A.Y. (2010). Role of conserved non-coding DNA elements in the Foxp3 gene in regulatory T-cell fate. Nature *463*, 808-812.

Zhou, J., Wang, J., Shen, B., Chen, L., Su, Y., Yang, J., Zhang, W., Tian, X., and Huang, X. (2014). Dual sgRNAs facilitate CRISPR/Cas9-mediated mouse genome targeting. The FEBS journal *281*, 1717-1725.

Zhu, J., Yamane, H., and Paul, W.E. (2010). Differentiation of effector CD4 T cell populations (*). Annu Rev Immunol *28*, 445-489.

Zijlstra, M., Bix, M., Simister, N.E., Loring, J.M., Raulet, D.H., and Jaenisch, R. (1990). Beta 2-microglobulin deficient mice lack CD4-8+ cytolytic T cells. Nature *344*, 742-746.

Zinkernagel, R.M., and Doherty, P.C. (1974). Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. Nature *248*, 701-702.

Zipursky, A., Pollock, J., Neelands, P., Chown, B., and Israels, L.G. (1963). The Transplacental Passage of Foetal Red Blood-Cells and the Pathogenesis of Rh Immunisation during Pregnancy. Lancet *2*, 489-493.