

An ancient fecundability-associated polymorphism creates a new GATA2 binding site in a distal enhancer of *HLA-F*

Katelyn M. Mika¹, Xilong Li^{2,3}, Francesco J. DeMayo², and Vincent J. Lynch^{1*}

¹Department of Human Genetics, The University of Chicago, 920 E. 58th Street, CLSC 319C, Chicago, IL 60637, USA.

²National Institute of Environmental Health Sciences, Department Reproductive & Developmental Biology Laboratory, 111 TW Alexander Drive, Research Triangle Park, NC, 27709, USA

³Current address: Feed Research Institute, Chinese academy of agricultural sciences. No. 12 Zhongguancun South st. Haidian district, Beijing, China, 100081

*Correspondence: vjlynch@uchicago.edu

Abstract

Variation in female reproductive traits such as fertility, fecundity, and fecundability are heritable in humans, but identifying and functionally characterizing genetic variants associated with these traits has been challenging. Here we explore the functional significance and evolutionary history of a G/A polymorphism of SNP rs2523393, which we have previously shown is an eQTL for the *HLA-F* gene and significantly associated with fecundability (time to pregnancy). We replicated the association between rs2523393 genotype and *HLA-F* expression using GTEx data and demonstrate that *HLA-F* is up-regulated in the endometrium during the window of implantation and by progesterone in decidual stromal cells. Next, we show that the rs2523393 A allele creates a new GATA2 binding site in a progesterone responsive distal enhancer that loops to the *HLA-F* promoter. Remarkably, we found that the A allele is derived in the human lineage, that G/A polymorphism arose before the divergence of modern and archaic humans, and is segregating at intermediate to high frequencies across human populations. Remarkably, the derived A is also has been identified in a GWAS as a risk allele for multiple sclerosis. These data suggests that the polymorphism is maintained by antagonistic pleiotropy and a reproduction-health tradeoff in human evolution.

Introduction

Female reproductive traits such as fertility, fecundity, and fecundability are heritable in humans^{1,2}, however, identifying the genetic bases for these traits has been challenging¹⁻³. We previously performed an integrated expression quantitative trait locus (eQTL) mapping and association study to identify eQTLs in mid-secretory endometrium that influence pregnancy outcomes^{1,2,4-6}. Among the eQTLs we identified was a G/A polymorphism (rs2523393) that was significantly associated with *HLA-F* expression and fecundability. Specifically, the G allele of rs2523393 was associated with longer intervals to pregnancy and lower expression of the *HLA-F* in mid-secretory phase (receptive) endometrium, whereas the A allele was associated with shorter intervals to pregnancy and higher *HLA-F* expression. The median time to pregnancy, for example, was 2.3, 2.6, and 4.9 months among women with the AA, GA, and GG genotypes, respectively^{1-3,6}.

While the functions of HLA-F are enigmatic, it is thought to regulate immune responses and may play an important role in regulating maternal-fetal immunotolerance during pregnancy^{7,8}. HLA-F is highly expressed, for example, in placental villi, on the surface of invasive and migratory extra-villous trophoblasts (EVTs), and decidual stromal cells (DSCs)⁹⁻¹¹. Furthermore HLA-F expression increases during gestation, reaching a peak at term^{9,12}. Like other MHC-I molecules, HLA-F binds natural killer (NK) cell receptors from the LIR and KIR families, including LIR1 and LIR2^{13,14}, KIR2DS4 and KIR3DL2¹⁵, and KIR3DL1 and KIR3DS1^{14,16,17}. Uterine natural killer (uNK) cells, which are essential for the establishment and maintenance of maternal immunotolerance and spiral artery remodeling, express KIR3DL1 and LIR2 suggesting HLA-F expressed by EVT and DSCs mediate interactions with uNK during implantation, trophoblast invasion, and establishment of the uteroplacental circulation. HLA-F expression level is also positively correlated with uNK abundance in mid-luteal endometria and is predictive of achieving pregnancy^{6,18} consistent an important role for HLA-F in the establishment of pregnancy.

Here we explore the functional and evolutionary history of the rs2523393 G/A polymorphism. We first replicate the association between rs2523393 genotype and *HLA-F* expression using GTEx data. We demonstrate that *HLA-F* expression increases during the

menstrual cycle and is up-regulated by progesterone during the differentiation of endometrial stromal fibroblasts into decidual stromal cells. Next, we show that rs2523393 is located within a cAMP/progesterone responsive enhancer that makes long range regulatory interactions with the promoter of *HLA-F*, and that the A allele creates a new binding site for the progesterone receptor (PGR) co-factor GATA2. Remarkably the G/A polymorphism arose before the divergence of modern and archaic humans, is segregating at intermediate to high frequencies across human populations, and associated with a predisposition to several diseases. These data suggest the G/A polymorphism is maintained by antagonistic pleiotropy and that there is a reproduction-health tradeoff in human evolution.

Materials and Methods

rs2523393 is an eQTL for *HLA-F*

We replicated the association between the G/A polymorphism at rs2523393 and *HLA-F* expression levels using GTEx Analysis Release V6 (dbGaP Accession phs000424.v7.p2) data for 35 tissues, including the 101 uterus samples^{19,20}. We also used GTEx data to identify other genes for which rs2523393 was an eQTL. Briefly, we queried the GTEx database using the 'Single tissue eQTLs search form' for SNP rs2523393.

HLA-F is expressed in decidual stromal cells and is differentially regulated by progesterone

Using previously generated microarray expression data, we examined the expression of *HLA-F* in the endometrium across the menstrual cycle (GSE4888²¹) and in endometrial samples from fertile women (n=5), woman had implantation failure following IVF (n=5), and woman with recurrent spontaneous abortion (n=5) (GSE26787²²). These microarray datasets were analyzed with the GEO2R analysis package, which implements the GEOquery²³ and limma R packages^{24,25} from the Bioconductor project to quantify differential gene expression. We also examined *HLA-F* expression in RNA-Seq data previously generated from ESFs treated with control media or differentiated (decidualized) with 100nM medroxyprogesterone acetate, and 1mM 8-bromo-cAMP (all from Sigma-Aldrich Co., St. Louis, MO) into DSCs^{26,27}. We also used previously generated RNA-Seq data to characterize the expression level of *HLA-F* in human endometrial stromal fibroblasts (ESF), decidual stromal cells (DSCs), and DSCs treated with an siRNA to knockdown PGR expression³⁸.

GATA2 ChIP-Seq

ESFs were cultured in 150-mm culture dishes with HESC growth media (DMEM/F12 containing 10% of fetal bovine serum and 1% antibiotic-antimycotic). Cells were grown to each approximately 90% confluency, starting to treat with decidual media (Opti-MEM media containing 2% charcoal-stripped FBS, 1% antibiotic-antimycotic, 1 μ M Medroxyprogesterone acetate, 10 nM E2, and 50 μ M cAMP). Decidual media was changed every 48 h. After 72h of treatment, cells were fixed. Cells were fixed for 15 min with 1/10 volume of freshly prepared formaldehyde solution (11% formaldehyde, 0.1 M NaCl, 1 mM EDTA, and 50 mM HEPES). The fixation was stopped by adding 1/20 volume 2.5 M glycine for 5 min. Fixed DSCs were collected and pelleted at 800 x g for 10 min at 4C. Cell pellets were washed two times with cold PBS-Igepal (0.5% Igepal, 1 mM PMSF). DSCs from six patients were pooled before genomic DNA isolation. GATA2 immunoprecipitation and DNA library generation were performed by Active Motif. ChIP and input DNA were amplified using the Illumina ChIP-Seq DNA Sample Prep Kit. Briefly, DNA ends were polished and 59-phosphorylated using T4 DNA polymerase, Klenow polymerase, and T4 polynucleotide kinase. Addition of 3'-adenine to blunt ends using Klenow fragment (3'-5' exo minus), Illumina genomic adapters were ligated and the sample was size fractionated to 175–225 bp on a 2% agarose gel. After amplification for 18 cycles with Phusion polymerase, the resulting DNA libraries were tested by RT-qPCR at the same specific genomic regions as the original ChIP DNA to assess quality of the amplification reactions. 36-nt sequence reads were identified by the Sequencing Service (using Illumina's Hi-Seq). Reads were mapped to the human genome hg19 using the BWA algorithm. Only tags that map uniquely, have no more than 2 mismatches, and that pass quality control filtering are used in the subsequent analysis. The number of unique alignments without duplicate reads was totaled and MACS analysis was performed at a pvalue 1e-7. Raw and processed data are available in SRA and GEO: GSE108409.

GATA2 knockdown and expression profiling

Three HESC subcultures were transfected with siRNA and treated with decidual media (see above) for 3 days. Total RNA was extracted using the Qiagen RNeasy RNA isolation kit (Qiagen). The RNA from 3 replicates (wells) was pooled for each treatment per cell line. The integrity of all RNA samples was tested with the Bioanalyzer 2100 (Agilent Technologies). The concentration of RNA was quantified on the Nanodrop Spectrophotometer (Nanodrop Technologies). The samples with 260/280 greater than 1.8 were used for microarray

hybridization. Microarrays were performed by the Genomic and RNA Profiling Core of Baylor College of Medicine using Affymetrix human genome U133 Plus 2.0 arrays (Affymetrix). Microarray CEL files were analyzed using dChip using the PM-MM model and quantile normalization. Combat was used to normalize differences and for batch correction. Two-side t test and fold changes were used to define differentially expressed genes. Genes with $P=0.05$ and an absolute fold change 1.4 were considered significant. Raw and processed data are available in SRA and GEO: GSE108409.

GATA2 ChIP-Seq

ESFs isolated from 6 patients were cultured separately in 150-mm culture dishes with ESF growth media and allowed to reach approximately 90% confluency before being treated with decidual media. Decidual media was changed every 48 hours. After 72 hours of treatment, cells were fixed for 15 minutes with 1/10 volume of freshly prepared formaldehyde solution (11% formaldehyde, 0.1 M NaCl, 1 mM EDTA, and 50 mM HEPES). The fixation was stopped by adding 1/20 volume 2.5 M glycine for 5 minutes. Fixed HESCs were collected and pelleted at 800 g for 10 minutes at 4°C. Cell pellets were washed 2 times with cold PBS, ESFs from 6 patients were pooled before genomic DNA isolation. GATA2 immunoprecipitation and DNA library generation were performed by Active Motif. DNA libraries were sequenced by Illumina's HiSeq Sequencing Service. 50-nucleotide sequence reads were mapped to the human genome (GRCh Build 37; February 2009) using the Burrows-Wheeler Aligner algorithm with default settings. Alignment information for each read was stored in the Sequence Alignment/Map or Binary version of the Sequence Alignment/Map format. Sequence alignments were extended in silico (using Active Motif software) at their 3'-ends to a length of 150–250 bp and assigned to 32-nucleotide bins along the genome. The resulting histogram of fragment densities was stored in a binary analysis results file. Raw and processed data are available in SRA and GEO: GSE108409.

Luciferase assays

A 1000bp region centered on rs2523393 (hg19, chr6:29705520-29706519) was synthesized with either the A or the G allele (Genscript) was cloned into the pGL3-Basic luciferase vector (Promega). A pGL3-Basic plasmid without the 1kb rs2523393 insert was used as a negative control. GATA2 and PGR expression vectors were also obtained from Genscript. Endometrial stromal fibroblasts (ATCC CRL-4003) immortalized with telomerase were

maintained in phenol red free DMEM (Gibco) supplemented with 10% charcoal stripped fetal bovine serum (CSFBS; Gibco), 1x ITS (Gibco), 1% sodium pyruvate (Gibco), and 1% L-glutamine (Gibco). Confluent ESFs in 96 well plates in 80 μ l of Opti-MEM (Gibco) were transfected with 100ng of the luciferase plasmid, 100ng of GATA2 and/or PGR as needed, and 10ng of pRL-null with 0.1 μ l PLUS reagent (Invitrogen) and 0.25 μ l of Lipofectamine LTX (Invitrogen) in 20 μ l Opti-MEM. The cells incubated in the transfection mixture for 6hrs and the media was replaced with the phenol red free maintenance media overnight. Decidualization was then induced by incubating the cells in the decidualization media: DMEM with phenol red (Gibco), 2% CSFBS (Gibco), 1% sodium pyruvate (Gibco), 0.5mM 8-Br-cAMP (Sigma), and 1 μ M of the progesterone analog medroxyprogesterone acetate (Sigma) for 48hrs. Decidualization controls were incubated in the decidualization control media (phenol red free DMEM (Gibco), 2% CSFBS (Gibco), and 1% sodium pyruvate (Gibco) instead for 48hrs. After decidualization, Dual Luciferase Reporter Assays (Promega) were started by incubating the cells for 15mins in 20 μ l of 1x passive lysis buffer. Luciferase and renilla activity were then measured using the Glomax multi+ detection system (Promega). Luciferase activity values were standardized by the renilla activity values and background activity values as determined by measuring luminescence from the pGL3-Basic plasmid with no insert. Each luciferase experiment was replicated in at least 4 independent experiments.

The rs2523393 polymorphism occurs in region that loops to the *HLA-F* promoter

We also took advantage of the fact that rs2523393 was an eQTL for *HLA-F* in EBV transformed lymphocytes (LCLs) to determine if this locus interacts with the *HLA-F* promoter in capture Hi-C (PCHi-C) data generated from LCLs. We used the CHiCP browser⁸⁵, a web-based tool for the integrative and interactive visualization of promoter capture Hi-C datasets, to identify long range interactions in the high-resolution capture Hi-C dataset of Mifsud et al. from the LCL line GM12878⁸⁶.

The rs2523393 A allele is derived in humans

To reconstruct the evolutionary history of the G/A polymorphism we used a region spanning 50bp upstream and downstream of rs2523393 from hg19 (chr6:32814778-32814878) as a query sequence to BLAT search the chimpanzee (CHIMP2.1.4), gorilla (gorGor3.1), orangutan (PPYG2), gibbon, (Nleu1.0), rhesus monkey (MMUL_1), hamadryas baboon

(Pham_1.0), olive baboon (Panu_2.0), vervet monkey (ChISab1.0), marmoset (C_jacchus3.2.1), Bolivian squirrel monkey (SalBol1.0), tarsier (tarSyr1), mouse lemur (micMur1), and galago (OtoGar3) genomes. For all other non-human species we used the same 101bp region as a query for SRA-BLAST against high-throughput sequencing reads deposited in SRA. The top scoring 100 reads were assembled into contigs using the 'Map to reference' option in Geneious v6.1.2 and the human sequence as a reference. Sequences for the Altai Neanderthal, Denisovan, and Ust-Ishim were obtained from the 'Ancient Genome Browser'. The frequency of the G/A allele across the Human Genome Diversity Project (HGDP) populations was obtained from the 'Geography of Genetic Variants Browser'. We inferred ancestral sequences of the 101bp region using the ancestral sequence reconstruction (ASR) module of Datamonkey²⁹ which implements joint, marginal, and sampled reconstruction methods³⁰, the nucleotide alignment of the 101bp, the best fitting nucleotide substitution model (HKY85), a general discrete model of site-to-site rate variation with 3 rate classes, and the phylogeny shown in **Fig 6A**. All three ASR methods reconstructed the same sequence for the ancestral human sequence at 1.0 support.

Results

rs2523393 is a eQTL for *HLA-F*

We have previously shown that the rs2523393 A/G polymorphism is an eQTL for *HLA-F* in mid-secretory phase endometrium⁶. To replicate this observation in an independent cohort and in additional tissues, we tested if rs2523393 was correlated with *HLA-F* expression using GTEx data. Confirming our previous observation, we found that rs2523393 was an eQTL for *HLA-F* in the uterus ($P=0.028$) as well as 22 other tissues (**Fig. 1**). We also used GTEx data to identify other genes for which rs2523393 was an eQTL and found that it was an eQTL for 15 other genes, including *HLA-A* and *HLA-G* (**Table S1**). The observation that rs2523393 was an eQTL for *HLA-A* and *HLA-G* in GTEx data prompted us to explore whether it was an eQTL for these genes in the mid-secretory endometrium data from our original study⁶. Indeed we found that rs2523393 was also an eQTL for *HLA-G* ($P=0.041$), but not *HLA-A* in mid-secretory endometrium.

The rs2523393 G/A polymorphism occurs within a GATA2 motif

To infer the functional consequences of the G/A polymorphism we used DeepSea³¹, a deep learning-based algorithm that infers the effects of single nucleotide substitutions on

chromatin features such as transcription factors binding, DNase I sensitivity, and histone marks. DeepSea predicted the G allele would have a negative effect on the binding of GATA2 (log2 fold change effect: -2.07, E-value: 0.035) and GATA3 (log2 fold change effect: -1.00, E-value: 0.003) (**Fig 2A**). We next used JASPAR transcription factor binding site (TFBS) motifs³² to identify putative TFBSs in a 36bp window upstream and downstream of rs2523393. We identified a single TFBS in this window, a GATA motif (matrix ID: MA0036.3, score: 8.70) and found that the G/A polymorphism occurs at an invariant A site in the motif (GATAA). Similar to the DeepSea results, substituting the reference A allele with the alternative G allele is predicted to abolish GATA2 binding (**Fig 2B/C**). These results suggest that the G/A polymorphism alters binding of GATA2, a transcription factor that plays an essential role in mediating the transcriptional response to progesterone in decidual stromal cells (DSCs)³³⁻³⁵.

***HLA-F* is up-regulated by progesterone and GATA2 in decidual stromal cells**

Previous studies have shown that *HLA-F* is expressed in placental villi, extra-villous trophoblasts (EVT), and DSCs^{9-11,36}. To determine if the expression of *HLA-F* varies in the endometrium across the menstrual cycle, we used previously generated microarray expression data²¹. We found that *HLA-F* expression reached peak expression in the late secretory phase of the menstrual cycle, and was significantly differentially expressed throughout the menstrual cycle (**Fig 3A**). To test if *HLA-F* is up-regulated by progesterone, we examined its expression in RNA-Seq data from ESFs treated with control media or differentiated with cAMP/progesterone into DSCs treated with either a PGR specific siRNA or a scrambled control siRNA^{26,27,37}. We found that *HLA-F* was up-regulated 2.11-fold ($P=3.59\times 10^{-5}$; **Fig 3B**) in DSCs and that knockdown of PGR down-regulated *HLA-F* 1.67-fold ($P=6.15\times 10^{-8}$; **Fig 3B**). Next we used an siRNA to knockdown the expression of GATA2 in DSCs and assayed global gene expression using an Agilent Human Gene Expression 8x60K microarray. Consistent with our previous results, we observed a 2.20-fold down-regulation of *HLA-F* ($P=0.02$) upon GATA2 knockdown (**Fig 3C**). We also observed that *HLA-F* expression was lower in the endometria of women with implantation failure (1.38-fold, $P=1.54\times 10^{-3}$) or recurrent spontaneous abortion (1.35-fold, 0.011) compared to fertile controls (**Fig 3D**). Thus we conclude that *HLA-F* expression increases as the menstrual cycle progresses, is up-regulated as by progesterone, PGR, and GATA2 during the differentiation of ESFs into DSCs, and dysregulation is associated with implantation failure and recurrent spontaneous abortion.

The rs2523393 A allele creates a new enhancer that loops to the *HLA-F*

Our observations that rs2523393 is an eQTL for *HLA-F* and that progesterone and GATA2 up-regulate *HLA-F* in DSCs suggests rs2523393 may be located within or linked to a progesterone responsive enhancer. To identify such a regulatory element we generated new GATA2 ChIP-Seq data from DSCs and used previously published ChIP-Seq data from DSCs for the transcription factors PGR^{27,38,39}, FOXO1³⁸, FOSL2³⁸, and NR2F2 (COUP-TFII)⁴⁰, H3K27ac which marks active enhancers²⁷, H3K4me3 which marks active promoters²⁷, and DNase-Seq and FAIRE-Seq to identify regions of open chromatin²⁷. We also took advantage of the fact that rs2523393 was an eQTL for *HLA-F* in multiple tissues, including EBV transformed lymphocytes (LCLs), to identify additional transcription factor binding sites using ChIP-Seq data generated by ENCODE and promoter capture Hi-C (PCHi-C) data generated from LCLs. We found that rs2523393 was located in a region of open chromatin and within a GATA2 ChIP-Seq peak in DSCs, 170bp upstream of a FOSL2 ChIP-Seq peak in DSCs, and nearby a cluster of transcription factor binding sites in ENCODE data including a binding site for GATA2 (**Fig 4**). Finally, we found that this region loops to the *HLA-F* promoter in LCLs located ~14.5kb upstream of rs2523393, suggesting this region is an enhancer for *HLA-F*.

To test if this locus has regulatory potential we synthesized a 1000bp region spanning 500bp upstream and downstream of rs2523393 (**Fig 4**) with either the reference A allele or the alternate G allele and cloned them into the pGL3-Basic[*minP*] luciferase reporter vector, which lacks an enhancer but encodes a minimal promoter. Next we transiently transfected either the pGL3-Basic[*minP*]-rs2523393A or pGL3-Basic[*minP*]-rs2523393G luciferase reporter along with the pRL-null internal control vector into DSCs and quantified luciferase and renilla activity using a dual luciferase assay. We found that luciferase activity was significantly lower in DSCs transfected either pGL3-Basic[*minP*]-rs2523393A (Wilcoxon test $P=3.80 \times 10^{-4}$) or pGL3-Basic[*minP*]-rs2523393G (Wilcoxon test $P=3.82 \times 10^{-5}$) compared to empty vector pGL3Basic[*minP*] controls (**Fig 5**). Next, we co-transfected pGL3-Basic[*minP*]-rs2523393A or pGL3-Basic[*minP*]-rs2523393G along either a GATA2 expression vector or a PGR expression vector. While co-transfection of neither the GATA2 expression vector nor the PGR expression vector enhanced luciferase expression above empty vector controls, there were significant differences between alleles (**Fig 5**). Finally we co-transfected pGL3-Basic[*minP*]-rs2523393A or pGL3-Basic[*minP*]-rs2523393G with both GATA2 and PGR expression vectors. Consistent with cooperative interaction between GATA2 and PGR, we observed a significant increase in luciferase

expression from both the pGL3-Basic[*minP*]-rs2523393A (Wilcox test $P=3.00\times 10^{-10}$) and pGL3-Basic[*minP*]-rs2523393G (Wilcox test $P=3.61\times 10^{-7}$) vectors compared to empty vector controls. We also observed that the pGL3-Basic[*minP*]-rs2523393A vector drove significantly higher luciferase expression than the pGL3-Basic[*minP*]-rs2523393G vector (Wilcox test $P=3.65\times 10^{-3}$). Thus we conclude that rs2523393 is located within a GATA2 binding site in a progesterone responsive enhancer, and that the G/A polymorphism affects enhancer function.

The rs2523393 A allele is derived in humans

To reconstruct the evolutionary history of the G/A polymorphism we identified a region spanning 50bp upstream and downstream of rs2523393 from 37 primates, including species from each of the major primate lineages, multiple sub-species of African apes (Homininae), as well as modern and archaic (Altai Neanderthal and Denisova) humans. Next we used maximum likelihood methods to reconstruct ancestral sequences for this 101bp region. We found that the G allele was ancestral in primates and that the A variant was only found in Neanderthal, Denisovan, and modern human populations (**Fig 6A**). Next we examined the frequency of these alleles across the Human Genome Diversity Project (HGDP) populations and found that the derived and ancestral alleles were segregating at intermediate to high frequencies in nearly all human populations (**Fig 6B**). These results indicate that the derived A variant at rs2523393, which creates a new GATA2 binding site in an enhancer of HLA-F expression, arose before the population split between archaic and modern humans between 550,000 and 765,000 years ago⁴¹.

The rs2523393 G/A polymorphism is associated with disease

Our observation that the rs2523393 G/A polymorphism is relatively old and segregating at intermediate to high frequencies implies the polymorphism is being maintained by balancing selection, frequency dependent selection, or antagonistic pleiotropy⁴². Intriguingly, a previous GWAS found the rs2523393 A allele was associated with multiple sclerosis (MIM: 126200) with an odds ratio of 1.28 [1.18-1.39] ($P=1.04\times 10^{-17}$)⁴³ suggesting antagonistic pleiotropy maintains both alleles. To explore if other disease were associated with rs2523393, we took advantage of the UK Biobank GWAS results of ~2,000 phenotypes (**Fig 7**). We found that the rs2523393 G allele was most significantly associated with malabsorption/coeliac disease ($P=8.8\times 10^{-24}$, beta =-0.0016), multiple sclerosis ($P=2.7\times 10^{-10}$, beta =-0.0009), and hyperthyroidism/thyrotoxicosis ($P=1.2\times 10^{-8}$, beta =-0.0012).

Discussion

The mechanisms that underlie maternal tolerance of the antigenically distinct fetus are complex⁴⁴ and have been the subject of intense study since Medawar proposed tolerance was achieved by maternal immunosuppression and immaturity of fetal antigens⁴⁵. It is now clear that rather than being suppressed at the maternal-fetal interface, the immune system plays an active role in establishing a permissive environment for implantation, placentation, and gestation. Among the diverse immune cells that contribute to the successful establishment and maintenance of pregnancy are maternal regulatory T cells (Tregs) and $\gamma\delta$ T cells⁴⁶⁻⁴⁸, uterine natural killer cells^{49,50}, uterine dendritic cells⁵¹, uterine macrophage⁵², as well as DSCs themselves⁵³⁻⁵⁵. DSCs, for example, recruit uterine natural killer cells to the endometrium via IL-15 expression⁵⁶⁻⁶⁴, stimulate the migration of macrophage into the endometrium via CSF1 expression^{56-58,65,66}, and suppress cytokine secretion by allogenic CD4+ T cells via PD-1 ligand interactions⁵⁴, which are all essential for establishing maternal-fetal immunotolerance. These data indicate that DSCs directly regulate local immune responses through multiple mechanisms, potentially also through HLA-F.

Major histocompatibility complex (MHC) genes play an important role in the rejection of non-self tissues, but also contribute to maternal tolerance of the fetus^{1-3,5,67-69,4,70-81}. While the precise functions of HLA-F are unclear, our observations that its expression level in the endometrium during the window of implantation is associated with fecundability⁶, and that it is up-regulated by progesterone implicates it in the establishment of pregnancy. Intriguingly HLA-F binds LIR and KIR natural killer cell receptors¹³⁻¹⁷, suggesting HLA-F expressed by DSCs directly signal to uterine natural killer (uNK) cells which are essential for the establishment of maternal immunotolerance and spiral artery remodeling. Consistent with this hypothesis HLA-F expression level is positively correlated with uNK abundance in mid-luteal endometria and is predictive of achieving pregnancy^{6,18}.

Our observation that the G/A polymorphism is shared between modern humans, Neanderthals, and Denisovans indicates it is relatively ancient and may be maintained by balancing selection, frequency dependent selection, or antagonistic pleiotropy. We found that while the derived A allele creates a new GATA2 binding site and augments the function of a progesterone responsive enhancer of HLA-F, a previous GWAS found it was also associated

with multiple sclerosis⁴³ and the G polymorphism is also associated with malabsorption/coeliac disease, multiple sclerosis, and hyperthyroidism/thyrotoxicosis in the UK Biobank. These data are consistent with maintenance through antagonistic pleiotropy. We have previously shown that another fecundability-associated variant, which switches a repressor into an enhancer of endometrial *TAP2* expression, was also shared between modern humans, Neanderthals, and Denisovans⁸². Remarkably, while the ancestral T allele in the *TAP2* *cis*-regulatory element was associated with shorter time to pregnancy and ulcerative colitis⁸³, the derived C allele was associated with longer time to pregnancy and Crohn's disease (MIM: 266600)⁸⁴ suggesting these alleles are also maintained by antagonistic pleiotropy. Collectively these data suggest there is a reproduction-health tradeoff in human evolution.

Acknowledgments

This work was funded by a Burroughs Wellcome Preterm Birth Initiative grant (1013760), an NIH National Institute of General Medical Sciences Graduate Training Grant (T32GM007197), and by the March of Dimes Transdisciplinary Center (TDC) at UChicago-Northwestern-Duke.

Web Resources

GTEEx database (<http://www.gtexportal.org/home/eqtls/bySnp>), GEO2R analysis package (<http://www.ncbi.nlm.nih.gov/geo/geo2r/>), Geography of Genetic Variants Browser (<http://popgen.uchicago.edu/ggv/>), Datamonkey web-server (<http://www.datamonkey.org>), Online Mendelian Inheritance in Man (<http://www.omim.org>), Oxford Brain Imaging Genetics (BIG) Server (<http://big.stats.ox.ac.uk>), CHiCP, a web-based tool for the integrative and interactive visualization of promoter capture Hi-C datasets (<https://www.chicp.org>).

References

1. Kosova, G., Abney, M., and Ober, C. (2010). Heritability of reproductive fitness traits in a human population. *Proc. Natl. Acad. Sci. U.S.A.* *107*, 1772–1778.
2. Christensen, K., Kohler, H.P., Basso, O., Olsen, J., Vaupel, J.W., and Rodgers, J.L. (2003). The correlation of fecundability among twins: evidence of a genetic effect on fertility?. *Epidemiology.* *14*, 60-64.
3. Tropf, F.C., Verweij, R.M., van der Most, P.J., Stulp, G., Bakshi, A., Briley, D.A., Robinson, M., Numan, A., Esko, T., Metspalu, A., Medland, S.E., Martin, N.G., Sneider, H., Lee, S.H., and Mills, M.C. (2016). Mega-analysis of 31,396 individuals from 6 countries uncovers strong gene-environment interaction for human fertility. *bioRxiv*. doi: <https://doi.org/10.1101/049163>
4. Ober, C., Elias, S., Kostyu, D.D., and Hauck, W.W. (1992). Decreased fecundability in Hutterite couples sharing HLA-DR. *Am. J. Hum. Genet.* *50*, 6–14.
5. Ober, C., Hyslop, T., Elias, S., Weitkamp, L.R., and Hauck, W.W. (1998). Human leukocyte antigen matching and fetal loss: results of a 10 year prospective study. *Hum. Reprod.* *13*, 33–38.
6. Burrows, C.K., Kosova, G., Herman, C., Patterson, K., Hartmann, K.E., Valez Edwards, D.R., Stephenson, M.D., Lynch, V.J., and Ober, C. (2016). Expression Quantitative Trait Locus Mapping Studies in Mid-Secretory Phase Endometrial Cells Identifies HLA-F and TAP2 as Fecundability-Associated Genes. *PLoS Genet.* *12*, e1005858.
7. Ishitani, A., Sageshima, N., and Hatake, K. (2006). The involvement of HLA-E and -F in pregnancy. *J. Reprod. Immunol.* *69*, 101–113.
8. Persson, G., Melsted, W.N., Nilsson, L.L., and Hviid, T.V.F. (2017). HLA class Ib in pregnancy and pregnancy-related disorders. *Immunogenetics* *69*, 581–595.
9. Hackmon, R., Pinnaduwege, L., Zhang, J., Lye, S.J., Geraghty, D.E., and Dunk, C.E. (2017). Definitive class I human leukocyte antigen expression in gestational placentation: HLA-F, HLA-E, HLA-C, and HLA-G in extravillous trophoblast invasion on placentation, pregnancy, and parturition. *Am. J. Reprod. Immunol.* *77*, e12643.
10. Ishitani, A., Sageshima, N., Lee, N., Dorofeeva, N., Hatake, K., Marquardt, H., and Geraghty, D.E. (2003). Protein Expression and Peptide Binding Suggest Unique and Interacting Functional Roles for HLA-E, F, and G in Maternal-Placental Immune Recognition. *J. Immunol.* *171*, 1376–1384.
11. Nagamatsu, T., Fujii, T., Matsumoto, J., Yamashita, T., Kozuma, S., and Taketani, Y. (2006). Human leukocyte antigen F protein is expressed in the extra-villous trophoblasts but not on the cell surface of them. *Am. J. Reprod. Immunol.* *56*, 172–177.
12. Shobu, T., Sageshima, N., Tokui, H., Omura, M., Saito, K., Nagatsuka, Y., Nakanishi, M., Hayashi, Y., Hatake, K., and Ishitani, A. (2006). The surface expression of HLA-F on decidual trophoblasts increases from mid to term gestation. *J. Reprod. Immunol.* *72*, 18–32.

13. Lepin, E.J., Bastin, J.M., Allan, D.S., Roncador, G., Braud, V.M., Mason, D.Y., van der Merwe, P.A., McMichael, A.J., Bell, J.I., Powis, S.H., and O'Callaghan, C.A. (2000). Functional characterization of HLA-F and binding of HLA-F tetramers to ILT2 and ILT4 receptors. *Eur. J. Immunol.* *30*, 3552–3561.
14. Dulberger, C.L., McMurtrey, C.P., Hölzemer, A., Neu, K.E., Liu, V., Steinbach, A.M., Garcia-Beltran, W.F., Sulak, M., Jabri, B., Lynch, V.J., and Altfeld, M. (2017). Human Leukocyte Antigen F Presents Peptides and Regulates Immunity through Interactions with NK Cell Receptors. *Immunity* *46*, 1018–1029.
15. Goodridge, J.P., Burian, A., Lee, N., and Geraghty, D.E. (2013). HLA-F and MHC class I open conformers are ligands for NK cell Ig-like receptors. *J. Immunol.* *191*, 3553–3562.
16. Garcia-Beltran, W.F., Hölzemer, A., Martrus, G., Chung, A.W., Pacheco, Y., Simoneau, C.R., Rucevic, M., Lamothe-Molina, P.A., Pertel, T., Kim, T.-E., and Dugan, H. (2016). Open conformers of HLA-F are high-affinity ligands of the activating NK-cell receptor KIR3DS1. *Nat. Immunol.* *17*, 1067–1074.
17. Burian, A., Wang, K.L., Finton, K.A.K., Lee, N., Ishitani, A., Strong, R.K., and Geraghty, D.E. (2016). HLA-F and MHC-I Open Conformers Bind Natural Killer Cell Ig-Like Receptor KIR3DS1. *PLoS ONE* *11*, e0163297.
18. Kofod, L., Lindhard, A., Bzorek, M., Eriksen, J.O., Larsen, L.G., and Hviid, T.V.F. (2017). Endometrial immune markers are potential predictors of normal fertility and pregnancy after in vitro fertilization. *Am. J. Reprod. Immunol.* *78*. e12684.
19. GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* *348*, 648–660.
20. Melé, M., Ferreira, P.G., Reverter, F., DeLuca, D.S., Monlong, J., Sammeth, M., Young, T.R., Goldmann, J.M., Pervouchine, D.D., Sullivan, T.J., and Johnson, R. (2015). The human transcriptome across tissues and individuals. *Science* *348*, 660–665.
21. Talbi, S., Hamilton, A.E., Vo, K.C., Tulac, S., Overgaard, M.T., Dosiou, C., Le Shay, N., Nezhat, C.N., Kempson, R., Lessey, B.A., and Nayak, N.R. (2006). Molecular phenotyping of human endometrium distinguishes menstrual cycle phases and underlying biological processes in normo-ovulatory women. *Endocrinology* *147*, 1097–1121.
22. Lédée, N., Munaut, C., Aubert, J., Sérazin, V., Rahmati, M., Chaouat, G., Sandra, O., and Foidart, J.M. (2011). Specific and extensive endometrial deregulation is present before conception in IVF/ICSI repeated implantation failures (IF) or recurrent miscarriages. *J. Pathol.* *225*, 554-564.
23. Davis, S., and Meltzer, P.S. (2007). GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* *23*, 1846–1847.
24. Smyth, G.K. (2005). Limma: linear models for microarray data. *Bioinformatics and computational biology solutions using R and Bioconductor*. Springer, New York, NY. 397-420.
25. Gentleman, R., Carey, V., Huber, W., Irizarry, R., and Dudoit, S. (2006). *Bioinformatics and*

computational biology solutions using R and Bioconductor. Springer, New York, NY.

26. Tamura, I., Ohkawa, Y., Sato, T., Suyama, M., Jozaki, K., Okada, M., Lee, L., Maekawa, R., Asada, H., Sato, S., and Yamagata, Y. (2014). Genome-wide analysis of histone modifications in human endometrial stromal cells. *Mol. Endocrinol.* *28*, 1656–1669.
27. Lynch, V.J., Nnamani, M.C., Kapusta, A., Brayer, K., Plaza, S.L., Mazur, E.C., Emera, D., Sheikh, S.Z., Grutzner, F., Bauersachs, S., and Graf, A.. (2015). Ancient Transposable Elements Transformed the Uterine Regulatory Landscape and Transcriptome during the Evolution of Mammalian Pregnancy. *Cell Rep.* *10*, 551-561.
28. Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* *8*, 118–127.
29. Delport, W., Poon, A.F.Y., Frost, S.D.W., and Kosakovsky Pond, S.L. (2010). Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics.* *26*, 2455–2457.
30. Kosakovsky Pond, S.L. and Frost, S.D. (2005). Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* *22*, 1208-1222.
31. Zhou, J. and Troyanskaya, O.G. (2015). Predicting effects of noncoding variants with deep learning-based sequence model. *Nat. Methods* *12*, 931–934.
32. Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R., and Zhang, A.W. (2015). JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* *44*, D110-D115.
33. He, B., Lanz, R.B., Fiskus, W., Geng, C., Yi, P., Hartig, S.M., Rajapakshe, K., Shou, J., Wei, L., Shah, S.S., and Foley, C. (2014). GATA2 facilitates steroid receptor coactivator recruitment to the androgen receptor complex. *Proc. Natl. Acad. Sci. U.S.A.* *111*, 18261–18266.
34. Rubel, C.A., Franco, H.L., Jeong, J.-W., Lydon, J.P., and DeMayo, F.J. (2012). GATA2 is expressed at critical times in the mouse uterus during pregnancy. *Gene Expr. Patterns.* *12*, 196–203.
35. Rubel, C.A., Franco, H.L., Camper, S.A., Lanz, R.B., Jeong, J.W., Lydon, J.P., and DeMayo, F.J. (2011). Gata2 Is a Master Regulator of Endometrial Function and Progesterone Signaling. *Biol. Reprod.* *85*, 179.
36. Blanco, O., Tirado, I., Munoz-Fernandez, R., Abadia-Molina, A.C., Garcia-Pacheco, J.M., Pena, J., and Olivares, E.G. (2008). Human decidual stromal cells express HLA-G: Effects of cytokines and decidualization. *Hum. Reprod.* *23*, 144–152.
37. Pabona, J.M.P., Simmen, F.A., Nikiforov, M.A., Zhuang, D., Shankar, K., Velarde, M.C., Zelenko, Z., Giudice, L.C., and Simmen, R.C.M. (2012). Krüppel-like factor 9 and progesterone receptor coregulation of decidualizing endometrial stromal cells: implications for the pathogenesis of endometriosis. *J. Clin. Endocrinol. Metab.* *97*, E376–E392.

38. Mazur, E.C., Vasquez, Y.M., Li, X., Kommagani, R., Jiang, L., Chen, R., Lanz, R.B., Kovanci, E., Gibbons, W.E., and DeMayo, F.J. (2015). Progesterone receptor transcriptome and cistrome in decidualized human endometrial stromal cells. *Endocrinology* *156*, 2239–2253.
39. Kaya, H.S., Hantak, A.M., Stubbs, L.J., Taylor, R.N., Bagchi, I.C., and Bagchi, M.K. (2015). Roles of progesterone receptor A and B isoforms during human endometrial decidualization. *Mol. Endocrinol.* *29*, 882–895.
40. Li, X., Large, M.J., Creighton, C.J., Lanz, R.B., Jeong, J.-W., Young, S.L., Lessey, B.A., Palomino, W.A., Tsai, S.Y., and DeMayo, F.J. (2013). COUP-TFII regulates human endometrial stromal genes involved in inflammation. *Mol. Endocrinol.* *27*, 2041–2054.
41. Meyer, M., Arsuaga, J.-L., de Filippo, C., Nagel, S., Aximu-Petri, A., Nickel, B., Martínez, I., Gracia, A., Bermúdez de Castro, J.M., Carbonell, E., and Viola, B. (2016). Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature*. *531*, 504–507.
42. Carter, A.J.R. and Nguyen, A.Q. (2011). Antagonistic pleiotropy as a widespread mechanism for the maintenance of polymorphic disease alleles. *BMC Med. Genet.* *12*, 160.
43. De Jager, P.L., Jia, X., Wang, J., de Bakker, P.I.W., Ottoboni, L., Aggarwal, N.T., Piccio, L., Raychaudhuri, S., Tran, D., Aubin, C., and Briskin, R. (2009). Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat. Genet.* *41*, 776–782.
44. Erlebacher, A. (2001). Why isn't the fetus rejected?. *Curr. Opin. Immunol.* *13*, 590–593.
45. Medawar, P.B. (1953). Some immunological and endocrinological problems raised by the evolution of viviparity in vertebrates. *Symp. Soc. Exp. Biol.* *7*, 320-337.
46. Erlebacher, A. (2013). Mechanisms of T cell tolerance towards the allogeneic fetus. *Nat. Rev. Immunol.* *13*, 23–33.
47. Guerin, L.R., Prins, J.R., and Robertson, S.A. (2009). Regulatory T-cells and immune tolerance in pregnancy: a new target for infertility treatment? *Hum. Reprod. Update.* *15*, 517-535.
48. Moffett, A., and Loke, Y.W. (2004). The immunological paradox of pregnancy: a reappraisal. *Placenta* *25*, 1–8.
49. González, I.T., Barrientos, G., Freitag, N., Otto, T., Thijssen, V.L., Moschansky, P., von Kwaitkowski, P., Klapp, B.F., Winterhager, E., Bauersachs, S., and Blois, S.M. (2012). Uterine NK cells are critical in shaping DC immunogenic functions compatible with pregnancy progression. *PLoS ONE.* *7*, e46755.
50. Felker, A.M. and Croy, B.A. (2016). Uterine natural killer cell partnerships in early mouse decidua basalis. *J. Leukoc. Biol.* *100*, 645-655.
51. Tagliani, E. and Erlebacher, A. (2011). Dendritic cell function at the maternal-fetal interface. *Expert Rev. Clin. Immunol.* *7*, 593–602.

52. Nagamatsu, T. and Schust, D.J. (2010). Review: The Immunomodulatory Roles of Macrophages at the Maternal—Fetal Interface. *Reprod. Sci.* *17*, 209-218.
53. 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.
54. Nagamatsu, T., Schust, D.J., Sugimoto, J., and Barrier, B.F. (2009). Human decidual stromal cells suppress cytokine secretion by allogenic CD4+ T cells via PD-1 ligand interactions. *Hum. Reprod.* *24*, 3160–3171.
55. Komatsu, T., Konishi, I., Mandai, M., Mori, T., Hiai, H., and Fukumoto, M. (1998). Expression of class I human leukocyte antigen (HLA) and beta2-microglobulin is associated with decidualization of human endometrial stromal cells. *Hum. Reprod.* *13*, 2246–2251.
56. Dai, X.-M., Zong, X.-H., Sylvestre, V., and Stanley, E.R. (2004). Incomplete restoration of colony-stimulating factor 1 (CSF-1) function in CSF-1-deficient *Csf1^{op}/Csf1^{op}* mice by transgenic expression of cell surface CSF-1. *Blood* *103*, 1114–1123.
57. Ryan, G.R., Dai, X.M., Dominguez, M.G., Tong, W., Chuan, F., Chisholm, O., Russell, R.G., Pollard, J.W., and Stanley, E.R. (2001). Rescue of the colony-stimulating factor 1 (CSF-1)-nullizygous mouse (*Csf1^{op}/Csf1^{op}*) phenotype with a CSF-1 transgene and identification of sites of local CSF-1 synthesis. *Blood* *98*, 74–84.
58. Ovadia, S., Insogna, K., and Yao, G.-Q. (2006). The cell-surface isoform of colony stimulating factor 1 (CSF1) restores but does not completely normalize fecundity in CSF1-deficient mice. *Biol. Reprod.* *74*, 331–336.
59. Ashkar, A.A., Black, G.P., Wei, Q., He, H., Liang, L., Head, J.R., and Croy, B.A. (2003). Assessment of requirements for IL-15 and IFN regulatory factors in uterine NK cell differentiation and function during pregnancy. *J. Immunol.* *171*, 2937–2944.
60. Allavena, P., Giardino, G., Bianchi, G., and Mantovani, A. (1997). IL-15 is chemotactic for natural killer cells and stimulates their adhesion to vascular endothelium. *J. Leukoc. Biol.* *61*, 729–735.
61. Laskarin, G., Strbo, N., Bogovic Crncic, T., Juretic, K., Ledee Bataille, N.L., Chaouat, G., and Rukavina, D. (2005). Physiological role of IL-15 and IL-18 at the maternal-fetal interface. *Chem. Immunol. Allergy.* *89*, 10–25.
62. Barber, E.M. and Pollard, J.W. (2003). The uterine NK cell population requires IL-15 but these cells are not required for pregnancy nor the resolution of a *Listeria monocytogenes* infection. *J. Immunol.* *171*, 37–46.
63. Kitaya, K., Yamaguchi, T., and Honjo, H. (2005). Central Role of Interleukin-15 in Postovulatory Recruitment of Peripheral Blood CD16(-) Natural Killer Cells into Human Endometrium. *J. Clin. Endocrinol. Metab.* *90*, 2932–2940.
64. Gu, L., Tseng, S., Horner, R.M., Tam, C., Loda, M., and Rollins, B.J. (2000). Control of TH2 polarization by the chemokine monocyte chemoattractant protein-1. *Nature* *404*, 407–411.

65. Erlebacher, A. (2013). Immunology of the maternal-fetal interface. *Annu. Rev. Immunol.* *31*, 387–411.
66. Svensson, J., Jenmalm, M.C., Matussek, A., Geffers, R., Berg, G., and Ernerudh, J. (2011). Macrophages at the fetal-maternal interface express markers of alternative activation and are induced by M-CSF and IL-10. *J. Immunol.* *187*, 3671–3682.
67. Fernandez, N., Cooper, J., Sprinks, M., AbdElrahman, M., Fiszer, D., Kurpisz, M., and Dealtry, G. (1999). A critical review of the role of the major histocompatibility complex in fertilization, preimplantation development and feto-maternal interactions. *Hum. Reprod. Update.* *5*, 234-248.
68. Ober, C., Elias, S., O'Brien, E., Kostyu, D.D., Hauck, W.W., and Bombard, A. (1988). HLA sharing and fertility in Hutterite couples: evidence for prenatal selection against compatible fetuses. *Am. J. Reprod. Immunol.* *18*, 111-115.
69. Ober, C.L., Martin, A.O., Simpson, J.L., Hauck, W.W., Amos, D.B., Kostyu, D.D., Fotino, M., and Allen, F.H. (1983). Shared HLA antigens and reproductive performance among Hutterites. *Am. J. Hum. Genet.* *35*, 994–1004.
70. Ober, C., Aldrich, C.L., Chervoneva, I., Billstrand, C., Rahimov, F., Gray, H.L., and Hyslop, T. (2003). Variation in the HLA-G promoter region influences miscarriage rates. *Am. J. Hum. Genet.* *72*, 1425-1435.
71. Tan, C.Y., Ho, J., Chong, Y.S., Loganath, A., Chan, Y.H., Ravichandran, J., Lee, C.G., and Chong, S.S. (2008). Paternal contribution of HLA-G* 0106 significantly increases risk for pre-eclampsia in multigravid pregnancies. *Mol. Hum. Reprod.* *14*, 317-324.
72. Loisel, D.A., Billstrand, C., Murray, K., Patterson, K., Chaiworapongsa, T., Romero, R., and Ober, C. (2012). The maternal HLA-G 1597ΔC null mutation is associated with increased risk of pre-eclampsia and reduced HLA-G expression during pregnancy in African-American women. *Mol. Hum. Reprod.* *19*, 144-152.
73. O'Brien, M., McCarthy, T., Jenkins, D., Paul, P., Dausset, E. D. C., 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.
74. Larsen, M.H., Hylenius, S., Andersen, A.M.N., and Hviid, T.V.F. (2010). The 3'-untranslated region of the HLA-G gene in relation to pre-eclampsia: revisited. *HLA.* *75*, 253-261.
75. Moreau, P., Contu, L., Alba, F., Lai, S., Simoes, R., Orru, S., Carcassi, C., Roger, M., Rabreau, M., and Carosella, E.D. (2008). HLA-G gene polymorphism in human placentas: possible association of G* 0106 allele with preeclampsia and miscarriage. *Biol. Reprod.* *79*, 459-467.
76. Yie, S.M., Li, L.H., Xiao, R., and Librach, C.L. (2008). A single base-pair mutation in the 3'-untranslated region of HLA-G mRNA is associated with pre-eclampsia. *Mol. Hum. Reprod.* *14*, 649-653.

77. Aldrich, C.L., Stephenson, M.D., Karrison, T., Odem, R.R., Branch, D.W., Scott, J.R., Schreiber, J.R., and Ober, C. (2001). HLA-G genotypes and pregnancy outcome in couples with unexplained recurrent miscarriage. *Mol. Hum. Reprod.* 7, 1167-1172.
78. Pfeiffer, K.A., Fimmers, R., Engels, G., van der Ven, H., and van der Ven, K. (2001). The HLA-G genotype is potentially associated with idiopathic recurrent spontaneous abortion. *Mol. Hum. Reprod.* 7, 373-378.
79. Suryanarayana, V., Rao, L., Kanakavalli, M., Padmalatha, V., Raseswari, T., Deenadayal, M., and Singh, L. (2008). Association between novel HLA-G genotypes and risk of recurrent miscarriages: a case-control study in a South Indian population. *Reprod. Sci.* 15, 817-824.
80. Fan, W., Li, S., Huang, Z., and Chen, Q. (2014). Relationship between HLA-G polymorphism and susceptibility to recurrent miscarriage: a meta-analysis of non-family-based studies. *J. Assist. Reprod. Gen.* 31, 173-184.
81. Hylenius, S., Nybo Andersen, A., Melbye, M., and Hviid, T.V.F. (2004). Association between HLA-G genotype and risk of pre-eclampsia: a case-control study using family triads. *Mol. Hum. Reprod.* 10, 237-246.
82. Mika, K.M. and Lynch, V.J. (2016). An Ancient Fecundability-Associated Polymorphism Switches a Repressor into an Enhancer of Endometrial TAP2 Expression. *Am. J. Hum. Genet.* 99, 1059–1071.
83. Anderson, C.A., Boucher, G., Lees, C.W., Franke, A., D'Amato, M., Taylor, K.D., Lee, J.C., Goyette, P., Imielinski, M., Latiano, A., and Lagacé, C. (2011). Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat. Genet.* 43, 246–252.
84. Franke, A., McGovern, D.P.B., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., Lees, C.W., Balschun, T., Lee, J., Roberts, R., and Anderson, C.A. (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat. Genet.* 42, 1118–1125.
85. E. C. Schofield, T. Carver, P. Achuthan, P. Freire-Pritchett, M. Spivakov, J. A. Todd, O. S. Burren; CHiCP: a web-based tool for the integrative and interactive visualization of promoter capture Hi-C datasets, *Bioinformatics*, Volume 32, Issue 16, 15 August 2016, Pages 2511–2513
86. Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L, Wingett SW, Andrews S, Grey W, Ewels PA, Herman B, Happe S, Higgs A, LeProust E, Follows GA, Fraser P, Luscombe NM, Osborne CS. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat Genet.* 2015 Jun;47(6):598-606.



Fig 1. GTEx multi-tissue eQTL plot for rs2523393. Left, Tissue type, sample size, beta, p-value, and posterior probability that rs2523393 in each tissue tested is also an eQTL across other tissues. Center, single-tissue effect size and 95% confidence interval. Right, single-tissue eQTL p-value versus multi-tissue posterior probability.

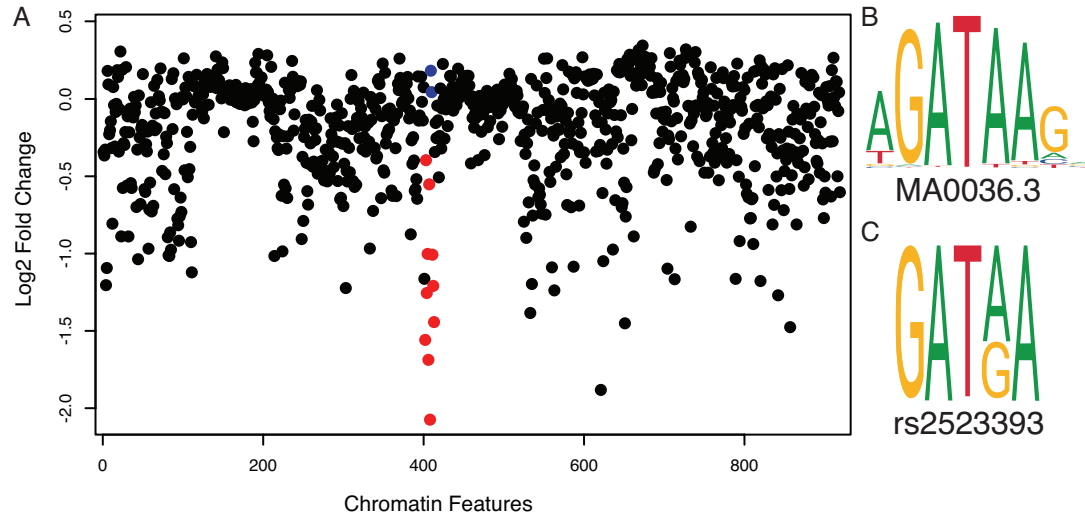


Fig 2. The rs2523393 G allele is predicted to abolish a GATA2 binding site. (A) DeepSea plot showing predicted effects of the G allele on chromatin features. GATA1 binding, blue. GATA2 and GATA3 binding, red. **(B)** JASPAR GATA2 motif. **(C)** GATA2 motif at rs2523393.

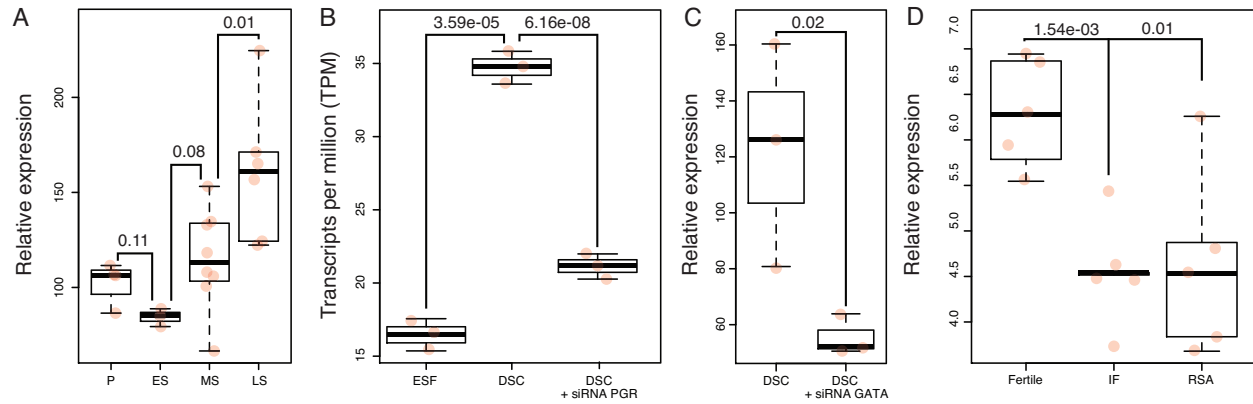


Fig 3. HLA-F is regulated by upregulated progesterone and GATA2 during the menstrual cycle and in human decidual stromal cells. (A) Relative expression of HLA-F throughout the menstrual cycle. P, proliferative phase. ES, early secretory phase. MS, mid-secretory phase. LS, late secretory phase. N=3-8, T-test. (B) Expression level of HLA-F in endometrial stromal fibroblasts (ESF), decidual stromal cells (DSCs), and DSCs treated with an siRNA to knockdown PGR expression. (C) Relative expression level of HLA-F in DSCs and DSCs treated with an siRNA to knockdown GATA2 expression. N=3, T-test. (D) Relative expression level of HLA-F in the endometrial fertile women, and women with implantation failure (IF) or recurrent spontaneous abortion (RSA) N=5, T-test.

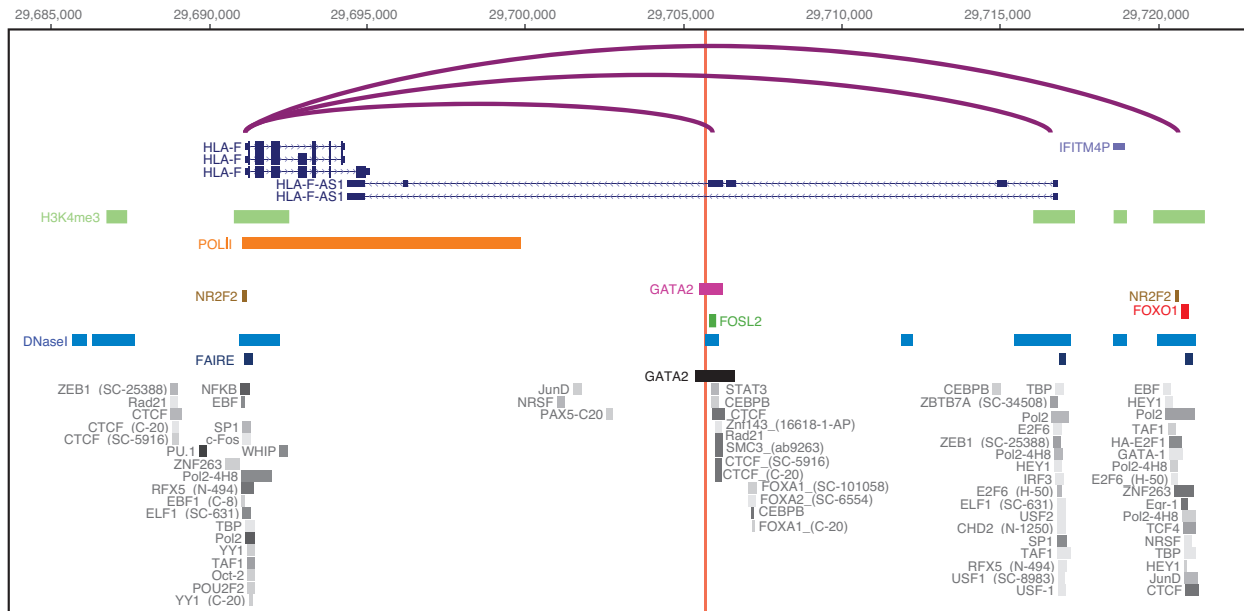


Fig 4. The rs2523393 G/A polymorphism is located in a distal enhancer of HLA-F. Location of rs2523393 polymorphism (vertical red line) with respect to histone modifications that characterize promoters (H3K4me4 ChIP-Seq), enhancers (H3K27ac ChIP-Seq), open chromatin (DNaseI-Seq, FAIRE-Seq), as well as GATA2 and NR2F2 ChIP-Seq binding sites in human DSCs. ENCODE transcription factor binding sites are shown in grey and intrachromosomal loop (PCHiC) data from LCLs is shown in purple.

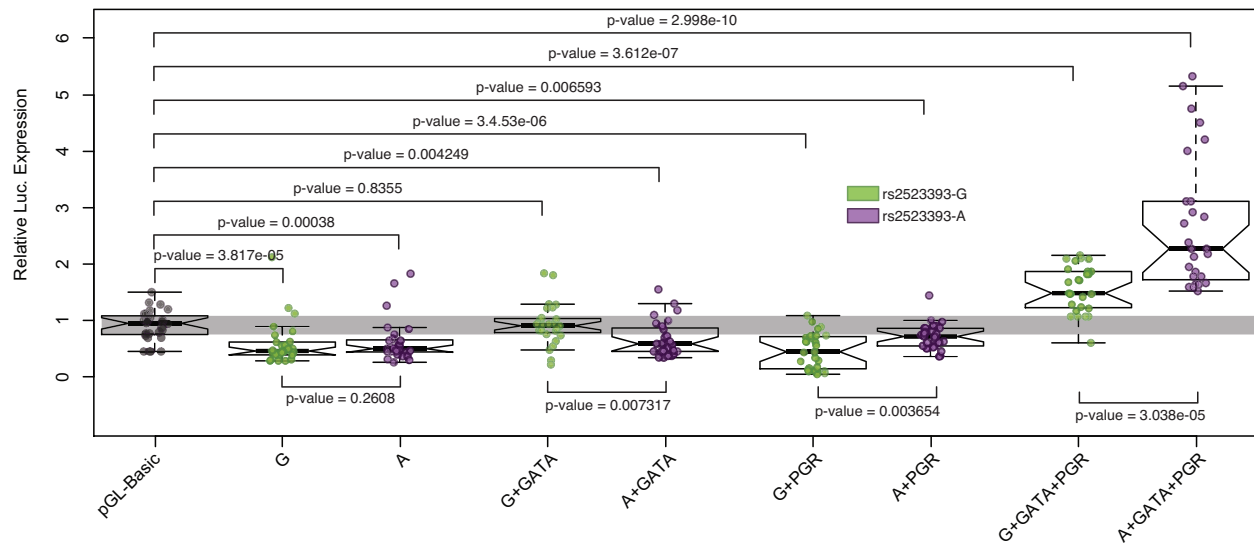


Fig 5. The rs2523393 G/A polymorphism alters the function of a progesterone response enhancer. Luciferase assay results testing the regulatory potential of the G (pGL3Basic-rs2523393G) and A (pGL3Basic-rs2523393A) alleles in DSCs treated with cAMP/MPA for 48 hours. Data are shown as luciferase activity from the pGL3Basic-rs2523393G or pGL3Basic-rs2523393A reporter relative to renilla activity (pRL-null and empty vector (pGL3Basic) controls. +GATA, Relative luciferase activity from the pGL3Basic-rs2523393G or pGL3Basic-rs2523393A in DSCs co-transfected with GATA. +PGR, Relative luciferase activity from the pGL3Basic-rs2523393G or pGL3Basic-rs2523393A in DSCs co-transfected with PGR. +GATA & PGR, Relative luciferase activity from the pGL3Basic-rs2523393G or pGL3Basic-rs2523393A in DSCs co-transfected with GATA and PGR. n=14, p=Wilcox test.

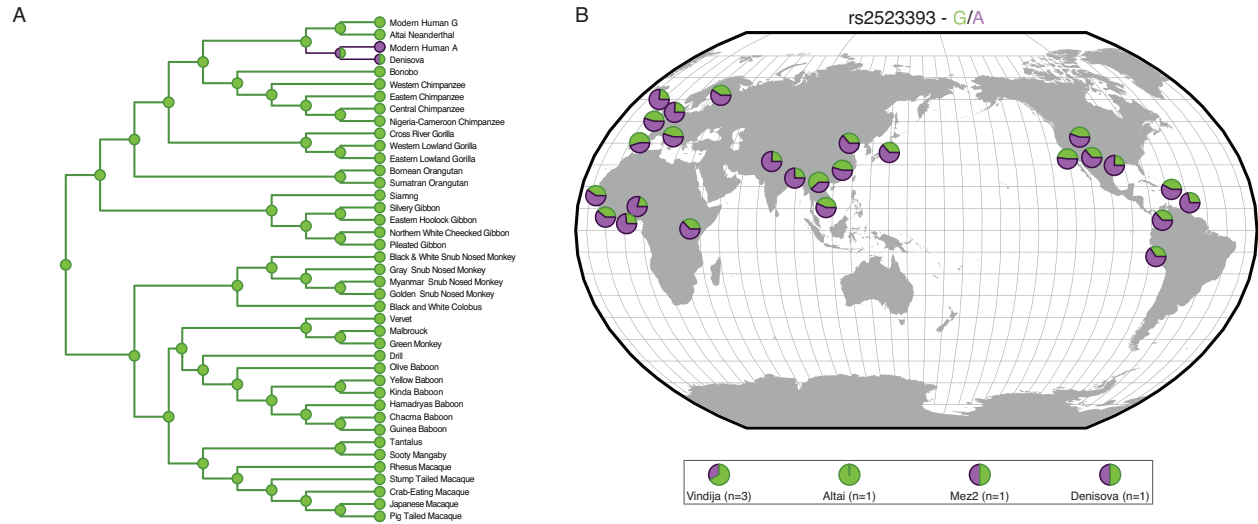


Fig 6. Evolutionary history G/A polymorphism at rs2523393. (A) rs2523393 genotype across primates and ancestral reconstructions. Genotype in extant species is shown as circles next to species names at terminal branches. Genotype at internal nodes based on ancestral reconstruction is shown as circles. Purple, A. Green, G. **(B)** Distribution of the G and A alleles of rs2523393 across HGDP populations.

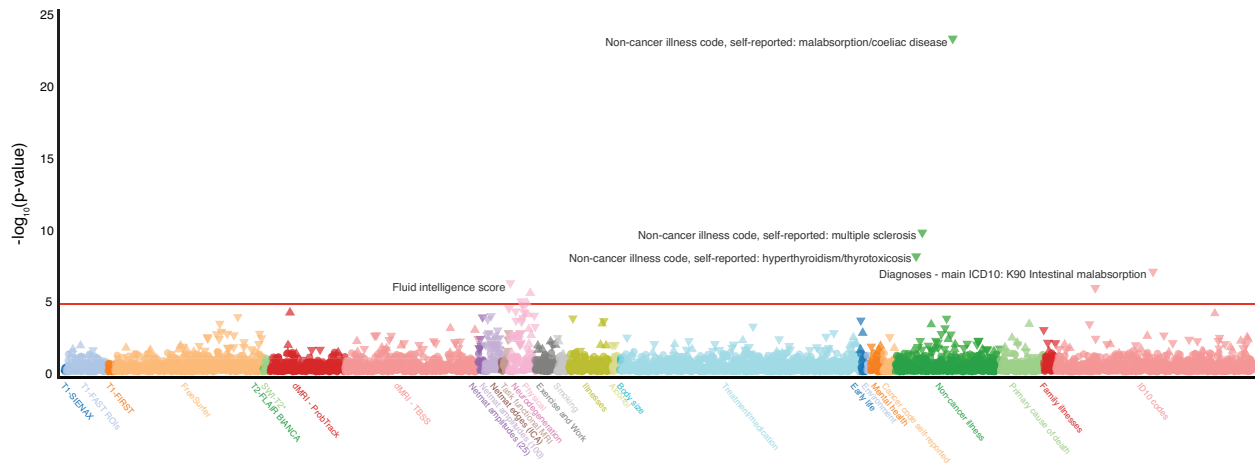


Fig 7. PheWAS plot of phenotypes associated with rs2523393 in the UK Biobank. Associations are shown for 3870 ICD10 codes, associations with $P < 0.05$ are shown as triangles and those with $P \geq 0.05$ are shown as circles. The most statistically significant phenotype associations are labeled.