Title page

DC-SIGN Polymorphisms Associated with Risk of Hepatitis C Virus Infection Among Men who Have Sex with Men but not Among Injecting Drug Users

Running Title: DC-SIGN SNPs and HCV susceptibility

Gaby S. STEBA[1], Sylvie M. KOEKKOEK[1], Joost W. VANHOMMERIG[1,2], Kees BRINKMAN[3], David KWA[4], Jan T.M. VAN DER MEER[5], Maria PRINS[1,2], Ben BERKHOUT[1], Michael TANCK[6], William A. PAXTON[1,7*], Richard MOLENKAMP[1*], Janke SCHINKEL[1*] on behalf of the MOSAIC (MSM observational Study of Acute infection with Hepatitis C) study group and the ACS (Amsterdam Cohort Studies)

1. Department of Medical Microbiology, Academic Medical Center, Amsterdam
2. Department of Infectious Diseases, Public Health Service of Amsterdam
3. Department of Internal Medicine, Onze Lieve Vrouwe Gasthuis, Amsterdam
4. Department of Microbiology, Onze Lieve Vrouwe Gasthuis, Amsterdam
5. Department of Internal Medicine, Division of Infectious Diseases, Tropical Medicine and AIDS, Center for Infection and Immunity Amsterdam (CINIMA), Amsterdam
6. Department of Clinical Epidemiology, biostatistics and bioinformatics (CEBB), Academic Medical Center, Amsterdam
7. Department of Clinical Infection, Microbiology and Immunology Institute of Infection and Global Health, University of Liverpool

* Contributed equally to this study

Word count text: 1964

Word count abstract: 98

Corresponding author:

William A Paxton, 8 West Derby Street, Liverpool L69 7BE, UK Tel: +44(0)1517959605, E-mail w.a.paxton@liverpool.ac.uk

Footnote page

Funding:

The research leading to these results has received funding from the European Community's Seventh Framework Programme [FP7-2007-2013] under grant agreement n°HEALTH-F3-2012-305578

Conflict of interest:

No

Meetings where data has been presented:

HCV 2015, Strasbourg, France

CROI 2016, Boston, USA

EASL 2017, Amsterdam, the Netherlands

Abstract

We aimed to identify whether genetic polymorphisms within L-SIGN or DC-SIGN correlate with HCV susceptibility. An MSM and an IDU cohort of HCV cases and multiple-exposed uninfected controls were genotyped for numerous L-SIGN and DC-SIGN polymorphisms. DC-SIGN SNPs -139, -871 and -939 correlate with HCV acquisition in the MSM cohort only. When the same SNPs were introduced into a transcription activity assay they demonstrated a reduction in expression with predicted alteration in binding of transcription factors. DC-SIGN promoter SNPs correlate with risk of HCV acquisition via sexual but not IDU exposure, likely through modulation of mRNA expression levels.

Keywords: HCV; HIV-1; Lectins, DC-SIGN; polymorphism, single nucleotide; MSM; sexual transmission

Introduction

Hepatitis C virus (HCV) represents a major global health burden, with 350.000 people dying annually from HCV-related liver disease.[1] Intravenous drug use is now the major transmission route. Nevertheless, since 2000, sexual transmission has been reported frequently among HIV-infected men who have sex with men (MSM) and is associated with high-risk sexual behavior. Interestingly, some individuals remain uninfected despite practicing high-risk behavior(s). Studies have shown that ultimately 10-20% of injecting drug users (IDU) do not seroconvert, suggesting a biological reason why some individuals are less prone to contract HCV. [2]

DC-SIGN (dendritic cell specific ICAM-grabbing non-integrin, CD209) and L-SIGN (DC-SIGN related, CD209L) are c-type lectins, which have been implicated to play a role in HCV transmission and infection.[3] DC-SIGN is a calcium-dependent cell surface lectin on the surface of dendritic cells (DCs).[4] DCs are localized in skin and mucosal tissues and may serve as a replication reservoir for HCV.[4,5] L-SIGN is mainly expressed on liver and lymph node sinusoidal endothelial cells. It shares 77% amino acid identity with DC-SIGN and it has been shown to capture several viruses including HCV.[3] Whereas the DC-SIGN neck region on exon 4 is highly conserved (7 repeats in the majority of individuals) the L-SIGN neck region is very variable.[6] This repeat region has been suggested to affect disease susceptibility and outcome for HIV-1 infection. [7–10]

The objective of this study was to analyze the frequency of previously reported genetic variations in DC/L-SIGN genes in individuals from two well-defined cohorts at risk of HCV infection who either seroconverted or remained uninfected. We identified three DC-SIGN SNPs that were associated with HCV susceptibility through high risk sexual exposure but not with IDU. Furthermore, we assessed whether these SNPs in the DC-SIGN promoter affect its activity.

Patients and Methods

Study populations

1. MSM cohort (MOSAIC)

Sixty-two HIV-1 infected, Western European MSM participating in the MSM Observational Study of Acute Infection with Hepatitis C (MOSAIC) cohort were included. Risk behavior data was available from behavioral questionnaires collected at 6 month intervals. Participants were categorized as multiple exposed uninfected (MEU, n = 30) or multiple exposed infected (MEI, n=32) based on reported behavioral risk factors at inclusion or any of the follow up visits, which have been shown to be associated with increased risk of acquiring HCV sexually in the MOSAIC cohort. Distribution of risk factors (i.e. no or inconsistent condom use, anal intercourse with an HCV-infected sex partner, fisting, use of sex toys, rectal bleeding during or after sex, and group sex) is summarized in supplemental Table 1. The MOSAIC study was approved by the Institutional Review Board of the Academic Medical Center under assigned study numbers NL26485.018.09 and NL48572.018.14.

2. IDU cohort (ACS)

Sixty-two Western European participants from the Amsterdam Cohort Studies (ACS) among IDU were selected, who started injecting drugs intravenously before 1990, which was a period with high incidence of HCV among drug users (up to 27.5/100 person years in the 1980s in this cohort).[11] The ACS among IDU was an open prospective cohort study recruiting drug users between 1985 and 2016 investigating the epidemiology, the natural history and pathogenesis of HIV-1 infection and other blood-borne and/or sexually transmitted diseases. Participants who injected more than 2 years and remained HCV seronegative during follow-up (n = 40) were classified as MEU where as 22 MEI seroconverted for HCV during follow up. Total duration of injecting drugs and follow up was similar for MEU and MEI (supplemental Table 1). The ACS study was approved by the Institutional Review Board of the Academic Medical Center under assigned study numbers MEC 07/182 and MEC 09/040.

DNA isolation and genotyping

DNA was isolated from 200 µl participant serum utilising the QIAamp DNA blood mini kit according to the manufacturer's protocol (Qiagen). The number of repeat domains within the L-SIGN repeat region was determined for each subject by PCR. PCR reactions contained 5 µl of template DNA, 400nM forward primer, 400nM reverse primer, 2.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.1 mg/mL Bovine Serum Albumin (BSA), 1.25 units FastStart Taq DNA polymerase in a total volume of 25 µL 1x Faststart PCR buffer.

L-SIGN SNP rs2277998 was assessed using the Ready-to-use hot start reaction mix for High Resolution Melting (HRM) curve analysis using the LightCycler® 480 (Roche). The reaction contained 2.0 µl DNA template, 2.5 mM $MgCl_2$, 8 ng α-casein, 450 nM Fwd primer (Biolegio) and 450 nM Rev primer (Biolegio) in a total volume of 20 µL 1x HRM master mix.

To assess reported DC-SIGN SNPs in the promoter region at positions -939 (rs735240), -871 (rs735239), -336 (rs4804803) and -139 (rs2287886), a DNA fragment covering approximately 1000 bp upstream of the ATG translation start site was amplified with two primer sets. The amplicons were sequenced in both directions with the same primers using Big dye terminator according to manufacturer's instructions (Applied Biosystems, Inc., Norwalk CT). Primers and amplification conditions are summarized in supplemental Table 2.

Cell culture

HEK 293T/17 cells (ATCC number: CRL-11268) were cultured in DMEM (Invitrogen) supplemented with 10% FCS, 1x MEM Non-Essential Amino Acids (Gibco), 100 U/ml penicillin and 100 U/ml streptomycin. Cells were incubated at 37°C in 5% CO_2 and passaged twice a week upon 90% confluence.

Construction DC-SIGN promoter expression construct

The DC-SIGN promoter variants were constructed by amplifying the DC-SIGN promoter region from DNA from one study participant with the -139A, -871A and -939G variants using primers tailed with *XhoI* and *HindIII* restriction sites. The amplicons were cloned into the pGL10.4 vector[luc2] (Promega) at the *XhoI* and *HindIII* sites. Promoter variants (see supplemental Figure 1) were established by site directed mutagenesis. Mutations were made with the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies) with specific mutagenic primers (see Supplemental Table 2).

Transfection of 293T cells with promoter constructs and analysis of luciferase expression

293T/17 cells were transfected with the various DC-SIGN promoter constructs and a Renilla luciferase expression plasmid (pRL-CMV) (Promega) for normalization in a 50:1 ratio using Xtremegene (Invitrogen) according to manufacturer's protocol. Cells were incubated 24 hours and lysed with Passive lysis buffer (Promega). 5 µl of the lysate was used to measure Firefly and Renilla luciferase activity with Dual-Glo luciferase assay system (Promega) according to manufacturer's protocol.

Prediction of transcription factor (TF) binding sites

TF binding sites were predicted using the PROMO database (http://alggen.lsi.upc.es/) which uses TRANSFAC for prediction. [16]

Statistical analysis

DC/L-SIGN SNP genotype frequencies between MEU and MEI were compared using logistic regression. Initially, an additive/dominance deviation joint 2 degree of freedom test (with two genotype-dependent variables in the regression, one with 0/1/2 coding and the second with 0/1/0 coding) was carried out. Subsequently, in case of dominance deviation ($p<0.1$), a dominant or recessive genetic model was assumed, otherwise an additive genetic model was

assumed in the logistic regression model used to estimate the odds ratio (OR) and corresponding 95% confidence interval. A *p* value <0.05 was considered statistically significant and all analyses were carried out using SPSS software (IBM, version 20).

Results

DC-SIGN -139GG, -871GG and -939AA are associated with reduced HCV susceptibility in MSM

Patient characteristics are summarized in supplemental Table 1. In the MSM cohort, three DC-SIGN SNPs were significantly associated with HCV infection (Table 1). The -139GG was found more frequently in MEU (63.3% in MEU compared to 37.5% in MEI). Additionally, -871GG (36.7% in MEU compared to 12.5% in MEI) and the -939AA (53.3% in MEU compared to 21.9% in MEI) were found more often in MEU, indicating that -139GG, -871GG and -939AA genotypes protect against HCV acquisition (OR: 0.35 p= 0.045, OR: 0.23 p=0.027 and OR: 0.23 p=0.009 respectively). The -336 SNP was not significantly associated with HCV susceptibility. In the ACS IDU cohort, no significant associations were found between SNPs and HCV susceptibility.

No associations between L-SIGN polymorphisms and HCV susceptibility

No association with HCV susceptibility was found for L-SIGN SNP rs2277998. In addition, the L-SIGN repeat distribution between MEI and MEU was similar for both cohorts (supplementary Table 3). No significant difference in zygosity for the L-SIGN repeat was found between MEI and MEU (OR: 0.982 p=0.961) (supplementary Table 4).

DC-SIGN promoter SNPs affect promoter activity

We tested the effect of the promoter variants within the DC-SIGN promoter on transcription activity by using luciferase promoter constructs (Figure 1). The -139G caused a 2.6 fold reduction (p<0.001), the -871G a 3.3 fold reduction (p<0.001) and the -939A a 1.4 fold

reduction (p=0.086). This data suggests that the DC-SIGN promoter variants affect transcription levels and thereby protein and cell surface expression patterns.

Next, we investigated whether the observed decrease in DC-SIGN promoter activity for specific SNPs could be due to alterations in TF binding sites, by a *in silico* comparison of predicted TF binding sites of promoter variants (Figure 1B). The variants at the -139, -871 and -939 sites do affect multiple predicted TF binding sites, with some putative sites lost (GR, C/EBP, Pr-B, Pr-A, HOXD9, HOXD10) and some TF binding sites gained (GR-Alpha, AP-2Alpha). This would indicate that the SNPs identified within the DC-SIGN promoter region can modulate activity through differential binding of transcription factors.

Discussion

Here we investigated whether polymorphisms in DC-SIGN and L-SIGN correlated susceptibility to HCV infection in two well-defined cohorts consisting of individuals at high risk of HCV infection through sexual or intravenous exposure. We selected polymorphisms based on what has been reported within the literature for HCV as well as other infectious agents. In the MSM cohort we identified an association of HCV susceptibility with three DC-SIGN SNPs. These SNPs were not associated with HCV susceptibility in the IDU cohort. No effects were found for the DC-SIGN -336 SNP, the L-SIGN SNP rs2277998 and repeat polymorphism in either cohort.

We studied four SNPs in the DC-SIGN promoter region, of which three (-139, -871 and -939) were found to correlate with HCV susceptibility in MSM, with -139G showing the strongest effect. Although the same SNPs have previously been associated with other infectious diseases, this is the first time SNPs have been reported to be associated with susceptibility to sexual transmission of HCV. Interestingly, the -139G SNP has also been reported to protect against sexual transmission of HIV-1.[14]

It has been published previously that the combination of -139G and -939A in the DC-SIGN promoter region significantly reduces DC-SIGN expression on immature DCs compared to

-139A and -939G.[15] We now show that the -139G and -871G SNP independently cause a reduction in promoter activity, while the -939A variant failed to reach statistical significance (p = 0.085). The DC-SIGN promoter encodes multiple TF binding sites which are *in silico* predicted to be affected by the -139, -871 and -939 variants. This strongly suggests that the decreased promoter activity observed *in vitro* is (at least partly) caused by a reduction in TF binding, which will require further testing.

As our study was small, our observations clearly need to be confirmed in larger cohorts. However, the functional data supports the associations of the SNPs with protection against HCV acquisition. Collectively, our data suggest that DC-SIGN plays a role in HCV acquisition via sexual and not intravenous exposure. This effect appears to be mediated by reduced DC-SIGN expression, which suggest that DC-SIGN on DCs plays a role in sexual transmission of HCV, similar to its role in HIV infection. [4]. We hypothesize that DCs transfer HCV to the liver through DC-SIGN; individuals with the protective genotypes will have lower DC-SIGN expression, resulting in a reduced susceptibility to sexual acquisition of HCV. Alternatively, DC-SIGN expression on DCs at mucosal surfaces may influence HCV antigen capture and induction of localized immune responses and modulate mucosal protection against HCV acquisition, which does not play a role in intravenous exposure. Further studies into the exact mechanism behind DC-SIGN affecting HCV infection susceptibility are warranted to better understand how DC-SIGN expression levels might influence immune responses, as well as mechanisms of transmission.

Funding

This work was supported by funding from the European Community's Seventh Framework Programme [FP7-2007-2013] under grant agreement n°HEALTH-F3-2012-305578.

Acknowledgements

241 The authors would like to thank participants of the ACS and MOSAIC study, Margreet Bakker
242 for sample storage and handling and Astrid Newsum for data handling. This work was
243 conducted within the framework of the Amsterdam Cohort Studies on HIV infection and AIDS
244 (Web site: www.amsterdamcohortstudies.org), a collaboration between the Public Health
245 Service of Amsterdam, the Academic Medical Center of the University of Amsterdam, Sanquin
246 Blood Supply Foundation, the University Medical Center Utrecht, and the Dutch HIV
247 Monitoring Foundation.

References

1 World Health Organization Data and statistics. 2015.

2 Sutton AJ, Gay NJ, Edmunds WJ, Hope VD, Gill ON, Hickman M. Modelling the force of infection for hepatitis B and hepatitis C in injecting drug users in England and Wales. *BMC Infect Dis* 2006; **6**:93.

3 Cormier EG, Durso RJ, Tsamis F, Boussemart L, Manix C, Olson WC, *et al.* L-SIGN (CD209L) and DC-SIGN (CD209) mediate transinfection of liver cells by hepatitis C virus. *PNAS* 2004; **101**:14067–14072.

4 Geijtenbeek TB, Kwon DS, Torensma R, van Vliet SJ, van Duijnhoven GC, Middel J, *et al.* DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* 2000; **100**:587–97.

5 Goutagny N, Fatmi A, De Ledinghen V, Penin F, Couzigou P, Inchauspé G, *et al.* Evidence of viral replication in circulating dendritic cells during hepatitis C virus infection. *J Infect Dis* 2003; **187**:1951–1958.

6 Barreiro LB, Patin E, Neyrolles O, Cann HM, Gicquel B, Quintana-Murci L. The heritage of pathogen pressures and ancient demography in the human innate-immunity CD209/CD209L region. *Am J Hum Genet* 2005; **77**:869–86.

7 Wichukchinda N, Kitamura Y, Rojanawiwat A, Nakayama EE, Song H, Pathipvanich P, *et al.* The polymorphisms in DC-SIGNR affect susceptibility to HIV type 1 infection. *AIDS Res Hum Retroviruses* 2007; **23**:686–92.

8 Chaudhary O, Bala M, Singh J, Hazarika A, Kumar R, Luthra K. The DC-SIGNR 7/5 genotype is associated with high dendritic cell counts and their subsets in patients infected with HIV-1. *J Clin Immunol* 2013; **33**:788–97.

9 Rathore A, Chatterjee A, Sivarama P, Yamamoto N, Dhole TN. Role of homozygous DC-SIGNR 5/5 tandem repeat polymorphism in HIV-1 exposed seronegative North Indian individuals. *J Clin Immunol* 2008; **28**:50–7.

10 Liu H, Carrington M, Wang C, Holte S, Lee J, Greene B, *et al.* Repeat-region polymorphisms in the gene for the dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin-related molecule: effects on HIV-1 susceptibility. *J Infect Dis* 2006; **193**:698–702.

11 Van Den Berg CHSB, Smit C, Bakker M, Geskus RB, Berkhout B, Jurriaans S, *et al.* Major decline of hepatitis C virus incidence rate over two decades in a cohort of drug users. *Eur J Epidemiol* 2007; **22**:183–193.

12 Farré D, Roset R, Huerta M, Adsuara JE, Roselló L, Albà MM, *et al.* Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res* 2003; **31**:3651–3653.

13 Messeguer X, Escudero R. Regulatory Elements Using Species-Tailored. 2002; **18**:333–334.

14 Kagoné TS, Bisseye C, Méda N, Testa J, Pietra V, Kania D, *et al.* A variant of DC-SIGN gene promoter associated with resistance to HIV-1 in serodiscordant couples in Burkina Faso. *Asian Pac J Trop Med* 2014; **7**:S93–S96.

15 Mezger M, Steffens M, Semmler C, Arlt EM, Zimmer M, Kristjanson GI, *et al.* Investigation of promoter variations in dendritic cell-specific ICAM3-grabbing non-integrin (DC-SIGN) (CD209) and their relevance for human cytomegalovirus reactivation and disease after allogeneic stem-cell transplantation. *Clin Microbiol Infect* 2008; **14**:228–234.

Table and Figure Legends

Table 1 Distribution of DC/L-SIGN SNPs in MEI and MEI individuals

rs2287886 GG, rs735240 AA and rs735239 GG genotypes are significantly associated with protection against HCV acquisition in the MOSAIC (MSM) cohort. No significant associations within the ACS (IDU) cohort.

Figure 1 Effect SNPs on DC-SIGN promoter activity *A*, The -139 SNP causes a reduction of 2.6 fold (p=0.0005), the -871 SNP of 3.3 fold (p=0.0009) and the -939 SNP a 1.4 fold (not significant). *B*, Protective SNPs affect TF binding sites in the DC-SIGN promoter. Putative binding of TFs to DC-SIGN promoter sequences with and without SNPs. Some TFs do not bind anymore to the sequence containing protective SNPs (orange), some bind both sequences (blue) and some bind exclusively to the SNP containing the protective variant (green).

Supplemental Table legends

Supplemental Table 1 Patients characteristics from the MOSAIC and ACS cohorts

Supplemental Table 2 Primers and PCR conditions used for analysis of the DC/L-SIGN polymorphisms

Supplemental Table 3 Distribution L-SIGN repeat region among MEI and MEU individuals

Supplemental Table 4 Zygosity L-SIGN repeat region compared between MEI and MEU individuals

No difference in L-SIGN zygosity between MEI and MEU individuals.

Supplemental Figure legends

Supplemental Figure 1 Graphical representation of the DC-SIGN promoter and expression plasmid pGL4.10 construct

322 Table 1

MOSAIC							**MEI vs MEU**		
	genotype	MEI (n)	MEI (%)	MEU (n)	MEU (%)	Dominance deviation [a]	OR	95% CI	p value
L-SIGN rs2277998	AA	1	3%	4	13%				
	AG	11	34%	6	20%	0.09[b]	0.83	0.29 to 2.37	0.73
	GG	20	63%	20	67%		(GG vs AG+AA)		
DC-SIGN - 139	AA	3	9%	6	20%				
	AG	17	53%	5	17%	0.01[b]	0.35	0.12 to 0.97	0.04[c]
	GG	12	38%	19	63%		(GG vs AG+AA)		
DC-SIGN - 336	AA	22	69%	25	83%				
	AG	9	28%	5	17%	0.99	2.32	0.74 to 7.32	
	GG	1	3%	0	0%		(per G allele)		
DC-SIGN - 871	AA	16	50%	12	40%				
	AG	12	38%	6	20%	0.08[b]	0.23	0.06 to 0.85	0.03[c]
	GG	4	13%	11	37%		(GG vs AG+AA)		
DC-SIGN - 939	AA	7	22%	16	53%				
	AG	18	56%	5	17%	<0.01[b]	0.23	0.07 to 0.69	0.01[c]
	GG	7	22%	8	27%		(AA vs AG+GG)		
ACS							MEI vs MEU		
	genotype	MEI (n)	MEI (%)	MEU (n)	MEU (%)	Dominance deviation	OR	95% CI	p value
L-SIGN rs2277998	AA	1	4,50%	2	5%				
	AG	10	45,50%	17	43%	0.85	1.06	0.44 to 2.56	0.896
	GG	11	50,00%	21	53%		(per A allele)		
DC-SIGN - 139	AA	5	22,70%	10	25%				
	AG	5	22,70%	16	40%	0.22	0.70	0.36 to 1.38	0.3
	GG	12	54,50%	14	35%		(per A allele)		
DC-SIGN - 336	AA	15	68,20%	28	70%				
	AG	4	18,20%	9	23%	0.50	1.20	0.55 to 2.60	0.65
	GG	3	13,60%	3	8%		(per G allele)		
DC-SIGN - 871	AA	13	59,10%	18	45%				
	AG	6	27,30%	18	45%	0.21	0.79	0.36 to 1.74	0.56
	GG	3	13,60%	4	10%		(per G allele)		
DC-SIGN - 939	AA	4	18,20%	7	18%				
	AG	8	36,40%	18	45%	0.56	0.87	0.42 to 1.79	0.71
	GG	10	45,50%	15	38%		(per A allele)		

[a] P-value of dominance deviation test

[b] Dominance deviation p-value < 0.1

[c] Statistically significant (<0.05)

323

324

Supplementary Table 1

	MOSAIC		ACS	
Characteristics	**MEI**	**MEU**	**MEI**	**MEU**
n (total=124)	32	30	22	40
Mean age ± SD	43.0 ±6.9	48.5 ±7.9	52.0 ±6.9	52.8 ±7.2
% Male gender	100%	100%	50%	72.5%
% Dutch Nationality	87.5%	96.7%	86.4%	92.5%
% HIV positive at entry	100%	100%	0%	0%
HIV seroconversion during follow-up	n.a	n.a	13.6%	0%
Median start date of Follow-up (IQR)	22/2/2011 (4/2/2010-2/8/2011)	14/2/2011 (19/5/2010-20/12/2011)	23/2/1988 (15/1/1987-08/02/1992)	20/10/1992 (12/09/1988-22/04/1998)
Median time of follow-up ± SD	4.01 ± 1.80	3.78 ± 1.30	14.96 ± 5.65	14.31 ±5.62
Mean duration IDU in years	4 IDU in last 6 months (no duration)	n.a	7.21 ±3.42	8.45±4.83
% Reported sharing of needles[§]	0%	0%	75%	55%
Having an HCV-infected sex partner*	7	1	n.a	n.a
Fisting[§]				
With steady partner	9	5	n/a	n/a
With casual partner(s)	10	8	n/a	n/a
Use of sex toys[§]				
With steady partner	13	12	n/a	n/a
With casual partner(s)	15	4	n/a	n/a
Rectal bleeding during or after sex[§]				
With steady partner	`	10	n/a	n/a
With casual partner(s)	15	8	n/a	n/a
Groupsex[§]	24	23	n/a	n/a
Rectal bleeding during or after sex*	17	5		
CD4 count last negative moment(cases)/last visit (controls)*	523±138	621±222	n.a	n.a
CD4 count nadir	277±160	269±179	n.a	n.a
Baseline Mosaic Risk score (medium)*[#]	**2,9**	**1,1**		

n.a. = not applicable,
* = $p < 0,05$
[§] reported at least once
[#] [10]

344 **Supplementary table 2**[1]

Name	Orientation	Primer sequence 5'-> 3'	fragment length
L-SIGN repeat	Fwd	CCTAAGTCAGGAACAATCCGA	284bp, 353bp, 422bp, 491bp, 560bp, 629bp, 698bp (3/4/5/6/7/8/9 repeats, respectively)
	Rev	GAACTCACCAAATGCAGTCTTCAAATC	
L-SIGN SNP rs2277998	Fwd	GTCTAACTCCCAGCGGA	45bp
	Rev	TGGCAGGCGGTGACG	
DC-SIGN promotor PCR 1	Fwd	GCAGTCTTGGTTCCTTGGAG	630bp
	Rev	ACTTGCAGTGCCTCCTCAGT	
DC-SIGN promotor PCR 2	Fwd	TGCTGCTGTCCTCATTTTTG	638bp
	Rev	AGCATACAGAAACCCCGTTG	
Mutagenesis primer -139	Fwd	TAGGGATCTGTCATCCAAAAGGCTAGTGGAAAGCATCAGAGCA	
	Rev	TGCTCTGATGCTTTCCACTAGCCTTTTGGATGACAGATCCCTA	
Mutagenesis primer -871	Fwd	AGTACTAGTACATTTAATAACGTAGATAAATCTCACAAAACAG	
	Rev	CTGTTTTGTGAGATTTATCTACGTTATTAAATGTACTAGTACT	
Mutagenesis primer -939	Fwd	CACACTGTAAGATTTGATTTTATGTGAATTTTGAGAACAGGCA	
	Rev	TGCCTGTTCTCAAAATTCACATAAAATCAAATCTTACAGTGTG	

[1] **Amplification conditions:**

L-SIGN repeat: denaturation at 95°C for 5 min, followed by 45 cycles at 95°C for 30s, 60°C for 30s and 72°C for 1 min and a final extension step at 72°C for 10 min.

L-SIGN SNP rs2277998: 50°C for 2 min, denaturation at 95°C for 10 min, followed by 45 cycles at 95°C for 15s and 60°C for 15s, 72°C for 20s, followed by an HRM protocol of 95°C for 1 min, 40°C for 1 min and a fluorescence acquisition step at 60°C for 45s.

DC-SIGN SNPs: denaturation at 95°C for 5 min, followed by 5 cycles at 94°C for 30s, 61°C for 30s (-0.5°C every cycle) and 72°C for 45s followed by 32 cycles at 94°C for 30s, 60°C for 30s and 72°C for 45s and a final extension step at 72°C for 10min.

Supplementary Table 3

MOSAIC						MEI vs MEU		
	genotype	MEI (n)	MEI (%)	MEU (n)	MEU (%)	OR	95% CI	p value
DC-SIGN -139	AA	3	11%	2	25%	0.25 (GG vs AG+AA)	0.105 - 1.30	0.09
	AG	16	59%	1	13%			
	GG	8	30%	5	63%			
DC-SIGN -871	AA	15	56%	4	50%	0.08 (GG vs AG+AA)	0.01 - 0.59	< 0.01
	AG	10	37%	0	0%			
	GG	2	7%	4	50%			
DC-SIGN -939	AA	5	19%	5	63%	0.14 (AA vs AG+GG)	0.02 - 0.77	0.02
	AG	16	59%	0	0%			
	GG	6	22%	3	50%			

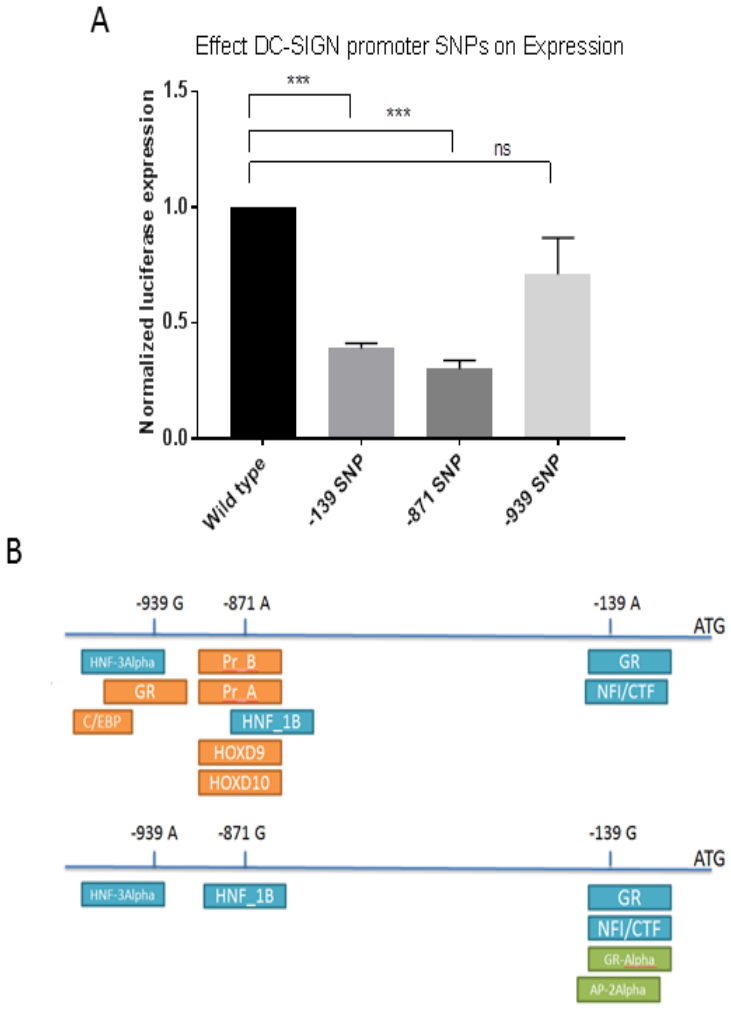
Supplementary Table 4

	genotype n(%)											
	4/5	4/6	4/7	5	5/6	5/7	5/9	6	6/7	6/9	7	7/9
MEI	0 (0.0)	0 (0.0)	5 (9.6)	2 (3.8)	2 (3.8)	9 (17.3)	1 (1.9)	1 (1.9)	5 (9.6)	1 (1.9)	22 (42.3)	4 (7.7)
MEU	1 (1.5)	1 (1.5)	1 (1.5)	6 (8.8)	5 (7.4)	13 (19.1)	1 (1.5)	3 (4.4)	13 (19.1)	0 (0.0)	24 (35.3)	0 (0.0)

Supplementary Table 5

	homozygous	heterozygous	OR	p value
MEI	25 (48.1%)	27 (51.9%)	0.9820	0.9608
MEU	33 (48.5%)	35 (51.5%)		
Total	58 (48.3%)	62 (51.7%)		

`

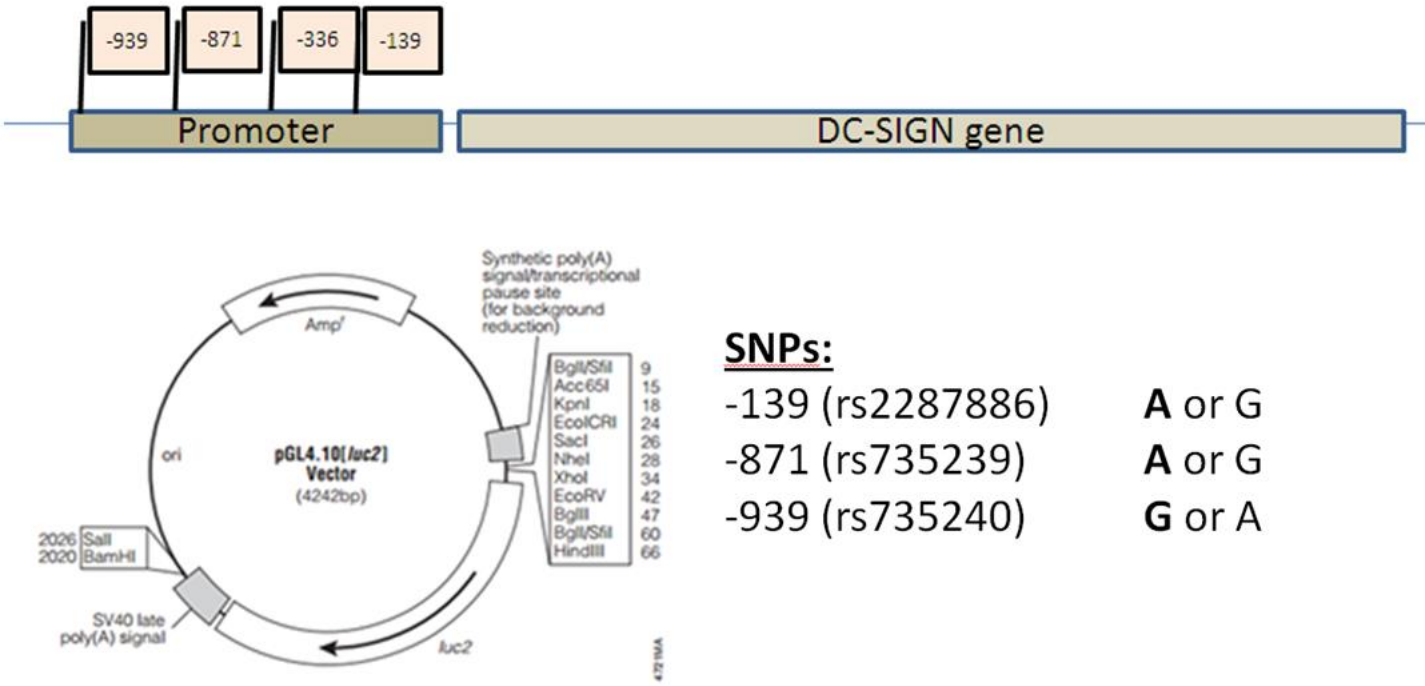


361

362 **Fig 1**

363

364



365

366 Fig S1