Host Factors Associated with HPV Infection in Inuit Women of Northern Quebec

Stephanie Metcalfe

Masters of Science, Epidemiology

Department of Epidemiology, Biostatistics and Occupational Health
McGill University
Montreal, Quebec, Canada

August 2012

Thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the degree of Masters of Science

Abstract

Background:
The Human Leukocyte Antigen (HLA) complex, which is a set of genes involved in the identification of foreign antigens and activation of the immune system, has been proposed as a possible co-factor in HPV infections.

Objectives:
To determine the frequencies of HLA alleles, haplotypes and genotypes in a cohort of Inuit women in Nunavik, Quebec. The association between these alleles and haplotypes and HPV incidence, period prevalence, persistence, multiple infections and Squamous Intraepithelial Lesions (SILs) will be investigated.

Methods:
A cohort of 548 Inuit women seeking routine care and living in four different communities of Nunavik was assembled (2002-2010). Cervical specimens were taken at each visit and tested for HPV-DNA. HPV genotypes were classified by oncogenic potential as well as by alpha-papillomavirus species. The HLA alleles were typed using DNA extracted from cervical samples taken at cohort entry. All odds ratios and rate ratios were adjusted for age.

Results:
The most common HLA class I and class II alleles were HLA-E*0103 (66.4%) and HLA-DQB1*03 (94.2%), respectively. Three HLA-G alleles were prevalent in this population: HLA-G*010401 (50.4%), HLA-G*010101 (31.4%), and HLA-G*010102 (15.2%). HLA-G*010102 was associated with a significantly decreased risk of overall (OR=0.64 95% CI=0.42-0.98) and low risk (LR) (OR=0.37 95% CI=0.16-0.83) HPV period prevalence. HLA-G*010101 (and its homozygous and heterozygous genotypes) were associated with an increased risk of LR alpha group 1 and 3. The homozygous HLA-G*010401 genotype was associated with a decreased risk of LR alpha group 3 infection. No HLA alleles, haplotypes or genotypes were significantly associated with HPV persistence. The HLA-DRB1*13 allele was associated with an increased risk of any SIL and LGSIL.

Conclusions:
The distribution of HLA alleles and haplotypes in a population of Nunavik, Quebec is different from that found in other Canadian populations, which included women of varied ethnicities. Several alleles (G*010101, G*010102, and G*010401) were associated with HPV period prevalence, but no alleles were significantly associated with HPV persistence.

Resume

Contexte :
Le complexe d'histocompatibilité (HLA) humain, un ensemble de gènes engagés dans l'identification d'antigènes étrangers et l'activation du système immunitaire, pourrait être un cofacteur des infections au VPH.

Objectifs :
Déterminer les fréquences des allèles, haplotypes et génotypes HLA au sein d'une cohorte de femmes inuites du Nunavik, au Québec. Cette recherche portera sur l'association entre ces allèles et haplotypes et la prévalence, l'incidence, et la persistance du VPH, de même que les infections multiples au VPH et les lésions malpighiennes intraépithéliales(SIL).

Méthodes :
Une cohorte composée de 548 femmes inuites s'étant présentées pour un examen de routine et vivant dans quatre communautés différentes du Nunavik a été constituée (2002-2010). Des échantillons de leur col utérin ont été prélevés à chaque visite et ont fait l'objet d'un test de dépistage d'ADN du VPH. Les génotypes du VPH ont été classés par potentiel oncogène et par espèces alphapapillomavirus. Les allèles HLA ont été typés à l'aide d'ADN provenant des échantillons de cols utérins prélevés lors de l'entré au sein de la cohorte. Tous les rapports de probabilité et taux bruts ont été ajustés en fonction de l'âge.

Résultats :
Les allèles HLA de catégorie I et catégorie II les plus courants sont respectivement les HLA-E*0103 (66,4 %) et HLA-DQB1*03 (94,2 %). Trois allèles HLA-G étaient prévalents au sein de cette population : HLA-G*010401 (50,4 %), HLA-G*010101 (31,4 %) et HLA-G*010102 (15,2 %). Le HLA-G*010102 était associé à un risque significativement réduit de prévalence générale du VPH (RC=0,64 95 % IC=0,42-0,98) et à un faible risque de prévalence de VPH à faible risque (RC=0,37 95 % IC=0,16-0,83). Le HLA-G*010101 et ses génotypes homozygotes et hétérozygotes étaient associés à un risque augmenté de groupes alpha 1 et 3 à faible risque. Le génotype homozygote HLA-G*010401 était associé à un risque réduit d'infection de groupe alpha 3 à faible risque. Aucun allèle, haplotype ni génotype HLA n'était associé de manière significative à la persistance du VPH. L'allèle HLA-DRB1*13 était associé à un risque réduit de toute lésion intraépithéliale malpighienne (SIL) ou de toute lésion intraépithéliale malpighienne de bas grade (LGSIL).

Conclusions :
La distribution des allèles et haplotypes HLA au sein d'une population du Nunavik, au Québec, diffère de celle observée au sein d'autres populations canadiennes, comprenant des femmes d'origines ethniques diverses. Plusieurs allèles (G*010101, G*010102 et G*010401) étaient associés à la prévalence du VPH au cours d'une période donnée, mais aucun n'était associé de façon significative à la persistance du VPH.

Statement of Support

This study was funded by the Canadian Institutes of Health Research (CIHR) through grant to Dr. Paul Brassard and from a clinical research networks/CIHR team grant on HPV infection and associated diseases. I received a Graduate Research Enhancement and Travel (GREAT) Award to present my research at the 15th International Congress on Circumpolar Health (ICCH) Conference in Fairbanks, Alaska.

Acknowledgments

I would like to thank the following people for their contribution towards this thesis:

The Tultattavik Health Center, the Nunavik Regional Board of Health and Social Services, the nurse practitioners who worked on the study and the participating women and communities in Nunavik, without whom this research would not have been possible.

My supervisor, Dr. Paul Brassard, for introducing me to this topic and for his mentorship and encouragement.

The students who previously worked with this data set for their code, in particular Helen Cerigo who first helped me navigate the dataset and Sophie Dellaniello for her help with SAS coding.

Dr. Michel Roger who performed the HLA typing, Dr. Francois Coutlée who performed the HPV typing, and Dr. Eduardo Franco for his knowledge of HPV epidemiology.

Jeff for reading over my thesis and for telling me not to worry, and my friends and family for their love and support.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ASC	Atypical Squamous Cells
BCCR	Biomarkers of Cervical Cancer Risk Study
CIN	Cervical Intraepithelial Neoplasia
DNA	Deoxyribonucleic Acid
HGSIL	High-Grade Squamous Intraepithelial Lesion
HIV	Human Immunodeficiency Virus
HLA	Human Leukocyte Antigen
HPV	Human Papillomavirus
HR	High Risk
IARC	International Agency for Research on Cancer
LGSIL	Low-Grade Squamous Intraepithelial Lesion
LR	Low Risk
MHC	Major Histocompatibility Complex
NK	Natural Killer
OC	Oral Contraceptives
OR	Odds Ratio
Pap	Papanicolaou
PCR	Polymerase Chain Reaction
RR	Rate Ratio
SCC	Squamous Cell Carcinoma
SIL	Squamous Intraepithelial Lesion
STI	Sexually Transmitted Infection
WM	Woman-Months

Literature Review

Human Papillomavirus (HPV)

Human papillomavirus (HPV) infection is the most common sexually transmitted infection (STI) in the world. Estimates of prevalence of HPV infection range from 2% to 44% (prevalence studies vary in age range and HPV detection sensitivity). Over 120 different genotypes of HPV have been found so far and over 40 of them infect the human anogenital tract. Most HPV infections are transient, but it has been established that a persistent infection with a high-oncogenic risk HPV type (HR-HPV) is a necessary, but not sufficient cause of cervical cancer. HPV infection has also been linked to cancers of the anus, vulva, vagina, penis and throat. Infection with low oncogenic risk HPV types (LR-HPV) can cause ano-genital warts. [1] [2]

Classification of HPV

Papillomaviruses with tissue tropism for anogenital mucosa make up the alpha-papillomavirus genus, which is the largest genus of the *Papillomaviridae* family. The alpha genus is further broken down into 15 species. HPV genotypes within each species share similar biological and pathological characteristics. Five phylogenetically related species (alpha-5, 6, 7, 9, and 11) of HPV are considered high-risk and can cause cervical cancer. LR species include a cervical species (alpha-1, 8, 10, 13) and vaginal species (alpha-2, 3, 4, and 15). [3] Figure 1 shows a HPV phylogenetic tree, the alpha papillomavirus can be seen at the top. [4] Classification into genotypes is based on a difference of more than 10% in the L1 gene (that encodes the major capsid protein). [5]

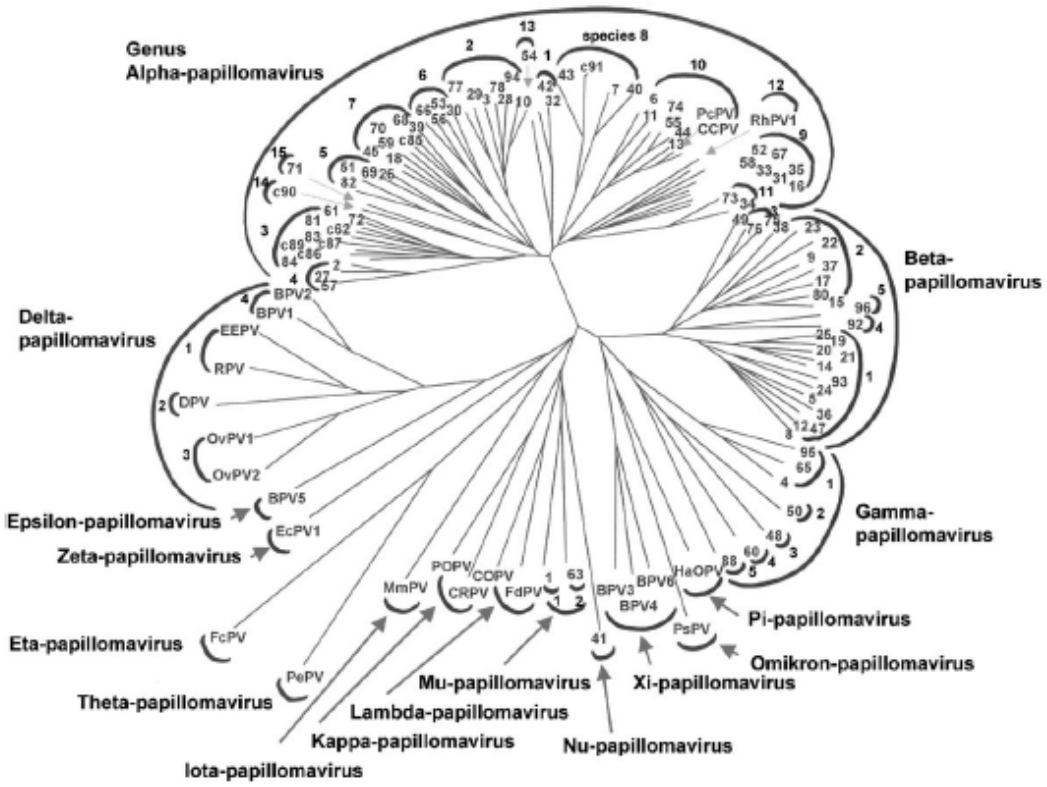

Figure 1: HPV phylogenetic tree [4]

The genotypes (types) are classified as high risk (HR) and low risk (LR) types based on their oncogenic potential for cervical cancer (based on strength of relative risk). [6] Table 1 outlines the most recent classification (2009) of HPV types by the International Agency for Research on Cancer (IARC). [3] HPV 16, 18, 31, and 45 are the most common high risk genotypes. HPV 16 accounts for 50-60% cases of cervical cancer worldwide and HPV type 18 accounts for 10-20%. [7] [8] HPV 6 and 11 are the most common LR genotypes and they account for 90% of all ano-genital warts. [2]

Table 1: Epidemiological Classification of HPV Genotypes [3]

Epidemiological Classification	**HPV Types**
High-Risk	16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59
Probable High-Risk	68
Possible High-Risk	26, 53, 66, 67, 70, 73, 82, 30, 34, 69, 85, 97
Low-Risk	6, 11, 40, 42, 43, 44, 54, 61, 72, 81, 89
Undetermined Risk	34, 57, 83

Each HPV genotype can be further classified into variants. HPV type 16 variants have been the most greatly studied. The phylogenetic tree of HPV type 16 contains 48 variants and has five branches, named after the geographic location from which they originate: Africa (Af1 and Af2), Asia and the Americas (AA), Europe (E), and East Asia (As). Evidence suggests that the evolution of papillomaviruses followed that of humans and their spread around the globe in prehistoric times. [9]

Natural History of HPV Infection

HPV is primarily sexually transmitted (through direct epithelial contact), though there is some evidence of vertical transmission. [8] Prevention of HPV is similar to that of other STIs, but condom use does not appear to provide full protection against transmission. A recent meta-analysis found that there was no consistent evidence that condom use could substantially reduce the risk of HPV transmission. [2]

Papillomaviruses gain access to the epithelial basement membrane through micro-abrasions and infect basal cells. Infected cells then differentiate and migrate up to the upper layers of the epithelium. The virions are assembled in the outermost cell of the epithelium and are released when epithelial cells are sloughed off. Because viral release is not associated with cytolysis or cytopathic death and occurs far from immune surveillance sites, the virus is able to evade the immune system and persist for long periods. Infected epithelial cells appear to have died 'naturally' and there is no inflammation and no activation of the innate immune system. [7 10]

Most HPV infections are transient and clear spontaneously. Carcinogenic development is a result of a productive infection for a prolonged period of time. [1] On average, HPV infections last from 4 to 20 months, with the median duration around a year. [1,2] HR-HPV types usually last longer than LR-HPV types and HPV type 16 tends to last longer than other HR types. [1]

Epidemiology of HPV Infection

HPV infection prevalence is highest among young women (most HPV infections occur soon after initiation of sexual activity) with a decline in young adult women. Some studies report second peak of HPV prevalence in peri- or post-menopausal women. Figure 2 shows the two patterns of age-specific prevalence of HPV infection [1]. The late increase could be due to a cohort effect (older women having been exposed to different co-factors during their lifetime), reactivation of latent infection, or a new virus acquired from new sexual partners. The Society of Obstetricians and Gynaecologists of Canada estimate the prevalence of all HPV infections and HR-HPV infections in Canada to be 10-30% and 11-25%, respectively. They also estimate that 75% of Canadians will be infected with HPV at least once during their lifetime. [11]

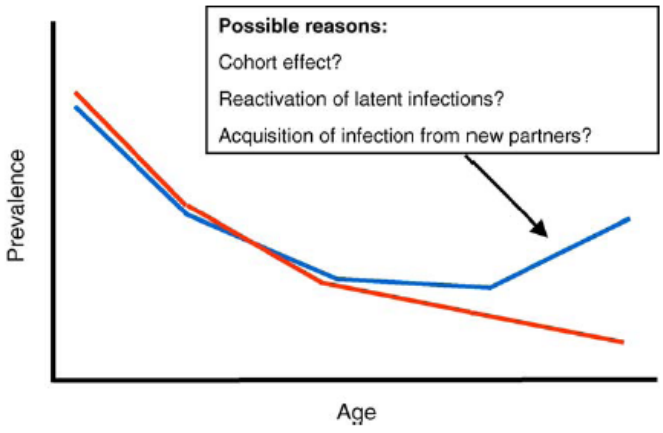

Figure 2: Two patterns of age-specific prevalence of cervical HPV infections [1]

Cohort studies that used polymerase chain reaction (PCR) to detect HPV found that in one month 3% of initially HPV-free women will acquire HPV infections. [1]

Cumulative incidence of HPV has been shown to exceed 40% after 3 years in American and English populations. [1] In a Brazilian population, the cumulative incidence was 24% over 18 months. [1] HR-HPVs have been shown to have a higher incidence rate than LR-HPVs. [1]

Risk Factors for HPV Infection

Various risk factors of HPV infection have been identified, such as number of sexual partners (which includes both lifetime and recent sexual partners), age at first sexual intercourse, cigarette smoking, oral contraceptives (this association may be due to increased surveillance of those who take OCs), high parity, other STIs and immunosuppression. The most consistent determinant of HPV infection is age. [1,2,7,8] Most studies show a decrease of HPV prevalence after the age of 30, independent of sexual activity, which suggests that individuals develop an adaptive immune response to HPV (see figure 2). [1,7]

The current focus of research has been on risk factors for viral persistence which is associated with an increased risk of developing cervical cancer. Studies have reported an association between persistence and older age, HPV variants, HPV DNA viral load, variations of the host immune system (such as genetic polymorphisms of the human leukocyte antigen, HLA), low fruit and vegetable consumption, and low circulating levels of vitamins. [1] HPV DNA viral load is an indicator of the productivity of HPV DNA replication and therefore is thought to be a predictor of a persistent infections and development of cytological abnormalities. HLA molecules present foreign antigens to the immune system and therefore play a critical role in the recognition and clearance of viral infections. [12] Figure 3, is an etiological model of HPV infection (and progression to cervical cancer) including possible co-factors.

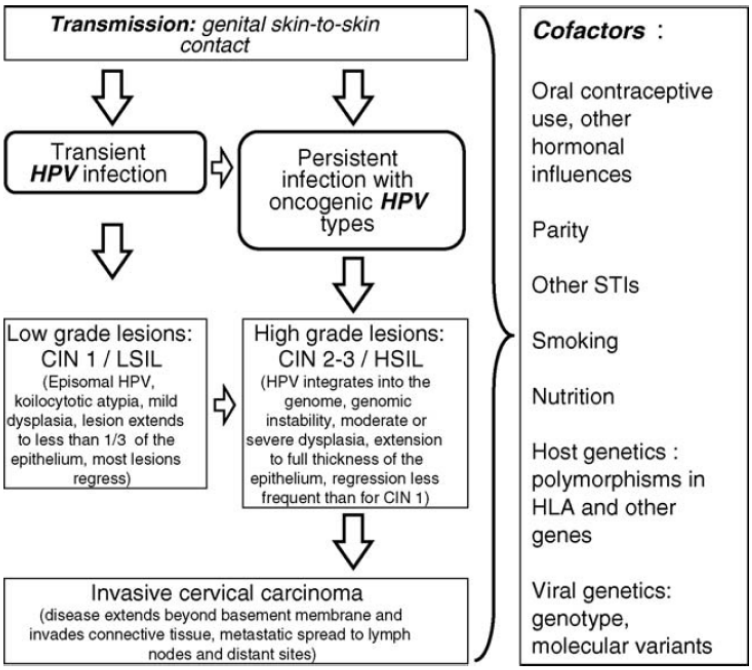

Figure 3: Etiological model of HPV infections [1]

Multiple HPV Infections

Co-infection with more than one HPV type is common, especially in young, sexually active women and the prevalence of multiple infections is significantly higher among women with more lifetime sex partners. [1] In a cohort of Brazilian women, one fifth of the women infected with HPV (at any time over follow-up) had multiple HPV infections. [13] In a cohort of Montreal university students, over half of the women were infected with HPV and 60% of the infected women had multiple infections. [14]

The occurrence of multiple HPV infections is significantly greater than what is expected by chance. [15] This excess is likely caused by the fact that HPV types share common risk factors and a common mode of transmission, rather than interactions between combinations of HPV types. [16] In fact, Chaturvedi *et al*. found that HPV genotypes involved in co-infections occur at random and saw little evidence for enhancement of susceptibility or competitive exclusion/cross-protection. [15] Female sex behaviours alone cannot explain the increase in co-infection prevalence and it is possible that immunological mechanisms may play a role.

Multiple infections are common among women who are immunosuppressed (HIV-infected women). [15,16] Out of a cohort of 208 HIV-infected women in Brazil, 98% were infected with HPV and multiple HPV types were detected in 78.9% of the cases (with an average of more than three infections per woman). [17]

It is unclear what role multiple infections play on HPV persistence. In general single and multiple infections have similar clearance rates, but some studies have found an association between multiple HPV infections and increased persistence. [1] A recent study found that persistence of an initial HPV infection was a strong risk factor for acquiring a new infection. Therefore it has been hypothesized that rather than co-infections leading to persistence (or persistence leading to co-infections), that both outcomes are an indication of immune dysfunction. [18,19]

Women with multiple infections have increased odds of squamous intraepithelial lesions (SIL). The risk of cervical disease in women co-infected with more than one HPV type is similar to the sum of the risks of the co-infecting types. Therefore, HPV co-infections occur at random and each HPV leads to cervical disease independently. [15,16]

Cervical Cancer

Using PCR amplification techniques, Walboomers *et al.* showed that 99.7% of cervical carcinoma samples, collected from women in 22 countries, contain HPV DNA. [20] The relative risk of developing cervical cancer if infected with a HR-HPV ranges between 50 and 100, which is one of the strongest associations currently identified in cancer epidemiology. Cervical cancer is also the only cancer for which a necessary cause has been identified. [1] Cervical cancer is the second most common cancer in women worldwide and 80% of cervical cancers occur in developing countries. [2,6] According to the Canadian Cancer Society, in 2011 there were 1,300 new cases of cervical cancer and 350 deaths in Canada. [21]

Natural history studies suggest that low-grade squamous intraepithelial lesions (LGSIL) are the result of a productive viral infection where the infected epithelium undergoes maturation and differentiation, and exhibits minor cell abnormalities. High-

grade squamous intraepithelial lesions (HGSIL) are the result of an HPV infection where the infected immature epithelial cells are prevented from undergoing maturation and differentiation. The immature cells continue to replicate which can lead to genetic abnormalities and eventually cancer. [2] LGSILs are caused by both LR-HPV and HR-HPV infections, whereas HGSILs are caused by HR-HPV infections. HGSIL, LGSIL, and ASC (Atypical Squamous Cells) are cytological classifications of squamous cells collected from Papanicolaou (Pap) smears. [2] Cervical biopsies are classified into three grades of cervical intraepithelial neoplasia (CIN 1, 2, 3).

Invasive cancer occurs when the disease extends beyond the basement membrane and invades connective tissue. [1,5] CIN 3 and invasive cancers are associated with the integration of the HPV DNA into the host chromosome. [22] The progression from HPV infection to invasive cervical cancer is depicted in figure 4. [5] The pre-cancerous state is long. The mean age of women with HGSIL is 28 years-old and the mean age of women with cervical cancer is 50 years-old. [2] Only 10% of all HPV infections result in high grade lesions (HGSIL/CIN2 and 3) and less than 1% of all HPV infections result in invasive cervical cancer. [23]

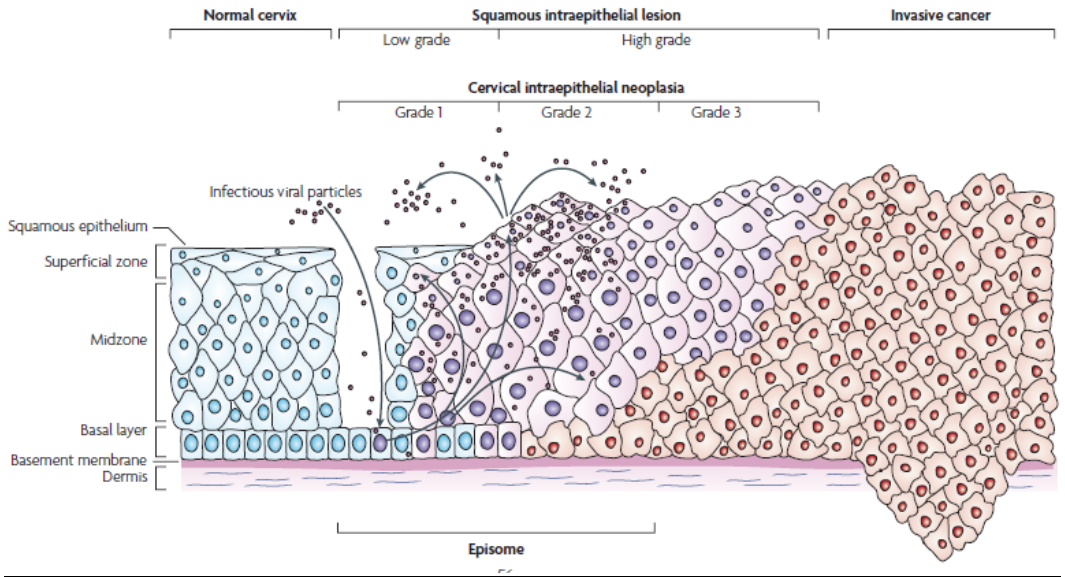

Figure 4: HPV infection mediated progression to cervical cancer [22]

Human Leucocyte Antigen (HLA)

The human leucocyte antigen (HLA), the name of the major histocompatibility complex (MHC) in humans, is a family of genes located on the short arm of chromosome 6. HLA genes code for cell-surface proteins, which present antigens (foreign and self) to immune cells, and other proteins that are involved in internal processing of antigens. The family contains class I and class II genes. Class I proteins are expressed on all nucleated cells and includes major (HLA-A, -B, and -C) and minor proteins (HLA-E, -F, and -G). HLA-A, -B, and –C proteins present antigens to (CD8+) cytotoxic T cells and HLA-E and –G present antigens to natural killer (NK) cells. If immune cells do not recognize the antigens presented by the HLA proteins as 'self' (for example if an infected cell's HLA proteins were presenting a viral protein instead of a host cell protein) they kill the infected cell through activation of the innate immune response. The class II proteins HLA-DP, -DQ, and –DR are expressed on immune cells (lymphocytes and macrophages) and present antigens to (CD4+) helper T cells, which help initiate a cell-mediated immune response. Class II proteins are formed from heterodimeric proteins (αβ). Both HLA class I and class II proteins are important in the regulation of the immune response to viral (and other) infections. [12]

Definitions

It is important to define some genetic terms that will be used. An allele is an alternative form or version of a gene. An individual's genotype is the two alleles inherited for a particular gene (this should not be confused with a virus's genotype which is a classification based on a difference of more than 10% in the L1 gene for HPV). Individuals inherit two alleles of each gene, one from each parent; if both alleles are the same the individual is homozygous for that gene and if the two alleles are different then the individual is heterozygous for that gene. If an individual is heterozygous for a gene the allele that is expressed is called the dominant allele/gene. In order for a recessive allele/gene to be expressed an individual must be homozygous for that allele. [24]

A haplotype is a set of DNA polymorphisms along a chromosome that tend to be inherited together. Two alleles (at two or more loci not necessarily on the same chromosome) are said to be in linkage disequilibrium when their combination occurs in the population less or more often than is expected to occur by chance. [24]

A polymorphism is a place in a DNA sequence where there is variation (the less common variation must be present in at least 1% of the population). The most common type of polymorphism is variation at a single base pair. The wild type allele is the non-mutated version of a gene. [24] HLA alleles are named in this way: HLA–locus name*allele number. The first two numbers after the asterix (*) are the allele number and any numbers after that are the allele subtype. [25] For example HLA-G*010401, 01 is the allele number and 0401 is the allele subtype.

Biological Plausibility of HPV/HLA Association

As noted previously, the majority of HPV infections are transient and only a minority of women in longitudinal studies are still infected (with the same type) on the subsequent visit. Therefore co-factors must be involved in viral acquisition, persistence, and progression to cervical cancer. Host genetics, in particular HLA polymorphisms, have been proposed to influence the natural history of viral infections. [1] Family studies have shown that there is a genetic link to cervical cancer development, in fact a woman's risk of cervical cancer increases two-fold when they have a biological first degree relative who has had a cervical tumor. [26]

It has been shown that when an individual's immune system is compromised, in particular inefficient cell-mediated and innate immunity, that they have an increased risk of HPV infection. [27,28] For example, women infected with human immunodeficiency virus (HIV) have a greater risk of recurrent genital warts than women without HIV (relative risk = 15.93). [29] A study has also shown that women with HIV were at a greater risk of CIN recurrence after treatment than women without HIV, and in the HIV-positive women, a low CD4+ count was a predictor for CIN recurrence (62% recurrence in HIV-

positive women, 18% recurrence in HIV-negative women, and 87% recurrence in HIV-positive women with low CD4+ counts). [30]

HLA genes are highly polymorphic allowing HLA proteins to bind to and present a variety of antigenic peptides. [12] It is hypothesized that individuals with HLA polymorphisms that have a higher binding affinity to HPV antigens are better at clearing the infection, and variation in genes involved in cell-processing might affect which antigens are presented on the surface. Therefore it is likely that some HLA alleles and haplotypes are more (or less) effective at recognition and clearance HPV and other viral infections. Variations in HLA genes have also been associated with Hepatitis B Virus (HBV) persistence and chronic Hepatitis C Virus (HCV). [31,32]

Overall Distribution of HLA Alleles in Caucasian and Black Populations

The distribution of HLA alleles in the Biomarkers of Cervical Cancer Risk (BCCR) case-control study can be seen in Table 2A and 2C below. The Biomarkers of Cervical Cancer Risk (BCCR) study was a hospital-based study conducted in Montréal, Québec (2001-2009) with a total of 1372 women. [33] Another study, The McGill-Concordia Cohort study, described the distribution of HLA class II alleles and haplotypes and class I allele HLA-B*07 (distribution of HLA alleles can be seen in Table 2B) in female university students in Montréal, Québec (1996-1999). This prospective cohort study investigated the natural history of HPV infection and cervical neoplasia. [28]

Table 2: HLA Allele and Haplotype Prevalence in Montreal Women

Study author	HLA allele	Proportion in the population
A: Ferguson *et al.* (BCCR study) [33]	HLA-G*010101 HLA-G*010102 HLA-G*010401 HLA-G*010103 HLA-G*0106 HLA-G*0103 HLA-G*010404 HLA-G*0105N HLA-G*010108	45% 24% 9% 7% 6% 4.5% 3% 1% 0.5%
B: Mahmud *et al.* (McGill-Concordia study) [28]	HLA-B*07 HLA-DQB1*03 HLA-DQB1*0602 HLA-DRB1*13 HLA-DRB1*1501 HLA-DRB1*1501-HLA-DQB1*03 HLA-DRB1*1501-HLA-DQB1*0602	14.7% 63.7% 21% 16.6% 25.2% 13.5% 21%
C : Ades et al. (BCCR study) [34]	HLA-B*07 HLA-DQB1*03 HLA-DQB1*0602 HLA-DRB1*13 HLA-DRB1*1501 HLA-B*07-DRB1*1501 HLA-B*07-DRB1*0602 HLA-DRB1*1501-DQB1*0602 HLA-B*07-DRB1*1501-DQB1*0602	17.7% 57.1% 19.7% 25.3% 21.3% 9.2% 8.5% 19.7% 8.5%

The distribution of HLA alleles in a population depends on the population's ethnicity. In a sub-analysis (which can be seen in figure 5), Mahmud *et al.* determined the prevalence of different HLA alleles and haplotypes according to ethnicity and found that the prevalence of HLA alleles did not differ significantly based on ethnicity, though DRB1*1501 was twice as prevalent in black women as in white and Asian women. The prevalence of the HLA-DRB1*1501-HLA-DQB1*0602 haplotype varied from 6% in Asian women to 25% in white, English students. [28]

Category	White			Asian	Black	Total
	English	French	Other			
Entire cohort	279 (53.2)	49 (9.4)	122 (23.3)	54 (10.3)	20 (3.8)	524 (100)
Alleles						
B*07	50 (17.9)	5 (10.2)	13 (10.7)	4 (7.4)	5 (25.0)	77 (14.7)
DQB1*03	175 (62.7)	33 (67.3)	77 (63.1)	42 (77.8)	7 (35.0)	334 (63.7)
DQB1*0602	71 (25.4)	9 (18.4)	21 (17.2)	3 (5.6)	6 (30.0)	110 (21.0)
DRB1*13	47 (16.8)	13 (26.5)	18 (14.8)	4 (7.4)	5 (25.0)	87 (16.6)
DRB1*1501	81 (29.0)	10 (20.4)	24 (19.7)	9 (16.7)	8 (40.0)	132 (25.2)
Haplotypes						
DRB1*1501-DQB1*03	42 (15.1)	6 (12.2)	14 (11.5)	7 (13.0)	2 (10.0)	71 (13.5)
DRB1*1501-DQB1*0602	71 (25.4)	9 (18.4)	21 (17.2)	3 (5.6)	6 (30.0)	110 (21.0)

NOTE. Data are no. (%) of subjects. Percentages do not sum to 100 because alleles are not mutually exclusive.

Figure 5: Prevalence of HLA Alleles and Haplotypes, According to Ethnicity [28]

Evidence for a HPV/HLA association

The majority of studies that have examined the possible association between HLA polymorphisms and HPV have focused on the effect that HLA polymorphisms (allele or haplotype) have on the development of cervical lesions or cancer. In the majority of these studies the control group included both women who test positive and negative for HPV infections. If an association is seen in these studies it is unclear at which point in the natural history of cervical cancer the particular HLA allele or haplotype has an effect. For example, if a particular allele is associated with an increased or decreased risk of cervical cancer, the allele could be evoking its effect at the point of viral acquisition, viral persistence, or cancer development.

The three groups of class II HLA alleles/haplotypes that have been most frequently associated with cervical cancer are (1) HLA-DQB1*03, (2) HLA-DRB1*1501 and HLA-DQB1*0602 alleles (which are in linkage disequilibrium and often found in the same haplotype), and (3) HLA-DRB1*13 and HLA-DQB1*0603 (also in linkage disequilibrium). HLA-DQB1*03, HLA-DRB1*1501 and HLA-DQB1*0602 alleles are most often associated with an increased risk of cervical cancer, though some studies have found an inverse association between HLA-DRB1*1501 and HLA-DQB1*0602 and cervical cancer. [12,26,27,35] One study found that the HLA-DRB1*1501-HLA-DQB1*0602 haplotype was associated

with an increased risk of HPV infection but a decreased risk of cervical cancer. HLA-DRB1*13 and HLA-DQB1*0603 alleles are associated with a decreased risk of cervical cancer. [34]

HLA class I alleles and haplotypes have been less frequently studied. The HLA-B*07 allele and haplotypes containing HLA-B*07 are often associated with an increased risk of cervical cancer. [12,26,36] Qiu *et al.* found that the HLA-B*07 allele was associated with familial cervical cancer (OR=8.7, CI=1.8-41.1). [36] Three papers have focused on the association between HLA-G polymorphisms and cervical lesions/cancer. A study in Brazilian women found a significant protective association between HLA-G*0103 and SIL and the HLA-G*0101/G*0104 genotype and HGSIL. [37] A Canadian study found a significant protective association between the heterozygous HLA-G*010101 genotype and invasive cancer and the heterozygous HLA-G*010102 genotype and CIN2. They also found that women with the homozygous haplotypes HLA-G*010102 and HLA-G*0106 were at a significantly greater risk of having invasive cervical cancer. [33]

There are three studies that have focused on the association between HLA polymorphisms and HPV infection. One that looked at HLA class I genes [38] and two that looked at HLA class II genes. [28,39] Two other studies (that examined the association between HLA polymorphisms and cervical cancer) did a sub-analysis on the control group to determine if any particular HLA alleles or haplotypes were associated with an increased risk of HPV infection. [33,34] The results of these studies can be seen in tables 3 and 4.

Ferguson *et al.* examined the association between class I HLA-E and –G polymorphisms and HPV cumulative prevalence (infected with HPV at any point over the study period) and persistence (positive for the same HPV type for 2 or more consecutive visits) in a group of 656 female university students. To determine the effect of an allele, they compared subjects who were homozygous or heterozygous for the allele to those who lacked the allele. All significant associations between HLA-G alleles and cumulative prevalence or persistence of HPV infection can be seen in table 3. None of the HLA-E alleles that were investigated were associated with cumulative prevalence or

persistence of HPV. Instead of grouping HPV types into HR and LR groups they grouped HPV types by alpha papillomavirus groups. Group 1 includes LR species α-1, α-8, α-10, and α-13, group 2 includes HR species α-5, α-6, α-7, α-9, and α-11, and Group 3 includes LR species α-3 and α-15. These groups were used for comparison to determine if any HLA alleles recognized epitopes on similar HPV species. [38]

In a second study by Ferguson *et al*., where the main objective was to study the association between HLA-G polymorphisms and CIN2, CIN3 and invasive cervical cancer, they did a sub-analysis within the control group. They found that HLA-G*010102 was significantly associated with a decreased risk of LR-HPV infection (OR=0.57, 95% CI=0.37-0.89) and that HLA-G*0106 was significantly associated with an increased risk of LR-HPV infection (OR=2.08, 95% CI=1.21-3.59). [33]

Table 3: HLA-G Alleles and Their Association with Prevalence and Persistence of HPV Infection in Previous Studies [38]

HLA-G Allele	**Strength of Association OR (95% CI)**	**HPV Cumulative Prevalence/Persistence**
HLA-G*01:01:02	1.90 (1.21-3.01)	Cumulative prevalence of HPV type 16 infection
HLA-G*01:01:03 HLA-G*01:01:05 HLA-G*01:01:08	2.07 (1.03-4.15) 3.87 (1.66-9.02) 2.17 (1.12-4.21)	Cumulative prevalence of HPV species α-1, 8, 10 & 13 infection
HLA-G*01:01:05 HLA-G*01:04:01	2.52 (1.03-6.19) 0.49 (0.25-0.95)	Cumulative prevalence of HPV species α- 2, 3, 4 & 15 infection
HLA-G*01:01:02	2.07 (1.16-3.68)	Persistent HPV type 16
HLA-G*01:03	2.99 (1.12-8.00)	Persistent HPV species α-2, 3, 4 & 15 infection

Mahmud *et al.* examined the association between HLA-B, HLA-DQB1, and HLA-DRB1 alleles and haplotypes and cumulative risk and prevalence of HPV using unconditional logistic regression (age at first intercourse and ethnicity were adjusted for). The haplotypes were chosen a priori because they were in linkage disequilibrium and the comparison was made between individuals who tested positive for both alleles and individuals who tested negative for both alleles. Significant associations from this analysis can be seen in table 4. A subgroup analysis was performed on only those

women who were at high risk of being exposed to a HPV infection (age at first intercourse ≤17 or lifetime sexual partners ≥5). In the subgroup analysis, HLA-DQB1*0602 was associated with an increased risk of LR-HPV infection (OR=2.2, 95% CI=1.0-4.7) and HLA-DRB1*13 was associated with an increased risk of any HPV infection (OR=2.6, 95% CI=1.3-5.2), HR-HPV infection (OR=2.8, 95% CI=1.3-5.9), and HPV type 16 infection (OR=3.3, 95% CI=1.4-8.2). No significant association between any allele/haplotype and HPV persistence was found. [28] Maciag *et al.* performed a similar study on a cohort of 620 women in Brazil. Unconditional logistic regression was used to determine ORs and possible co-factors, age and ethnicity, were controlled for and the results of this study can be found in table 4. [39] Ades *et al.* preformed a sub-analysis on 884 control subjects and found a significant association between the HLA-B7-HLA-DRB1*1501-HLA-DQB1*0602 haplotype and an increased risk of HPV type 16 and 18 infection (see table 4). [34] These studies suggest that there is a link between HLA class II polymorphisms and the risk and persistence of an HPV infection.

HLA and Multiple HPV infections

HPV co-infections are common, especially in young women. Co-infections occur more often that would be expected by chance, likely because HPV genotypes share the same risk factors. Women with co-infections acquire new infections at an increased rate and they have greater risk of HPV persistence (though not consistently across studies) and occurrence of SIL, which indicates that immunological susceptibility may be involved. [14] In a cohort of Montreal University students, Smith (Master's thesis) found that the alleles HLA-G*010101 (OR=1.66 95% CI=1.02-2.70), HLA-G*010105 (OR=6.44 95% CI=1.66-24.95), and HLA-DQB1*0602 (OR=1.83 95% CI=1.08-3.09) were significantly associated with an increased risk of multiple HPV infections. Smith also found that HLA-G*010103 was significantly associated with a decreased risk of multiple HPV infections (OR=0.34 95% CI=0.14-0.81). [14]

Table 4: HLA Class II Alleles and Haplotypes their Association with Prevalence and Persistence of HPV Infection in Previous Studies

Study Author	**HLA Class II Allele/Haplotype**	**Strength of Association OR (95% CI)**	**HPV Prevalence/Cumulative Risk/Persistence**
Mahmud *et al.* (2007) [28]	**Alleles:**		Cumulative risk of any HPV infection
	HLA-DRB1*13	1.7 (1.0-2.8)	
	Alleles:		Cumulative risk of HPV type 16 infection
	HLA-DRB1*1501	2.1 (1.1-4.1)	
	Haplotypes:		
	HLA-DRB1*1501-HLA-DQB1*0602	2.0 (1.0-4.1)	
Maciag *et al.* (2002) [39]	**Alleles:**		Cumulative risk of HPV infection
	HLA-DQB1*0201	0.64 (0.5-0.9)	
	HLA-DRB1*0301	0.46 (0.3-0.7)	
	HLA-DRB1*1601	3.99 (1.3-12.7)	
	Haplotypes:		
	HLA-DRB1*1601–HLA-DQB1-0502	5.27 (1.5-19.1)	
	HLA-DRB1*0301-HLA-DQB1*0201	0.45 (0.3-0.7)	
	HLA-DRB1*04–HLA-DQB1-0301	6.90 (1.5-32.2)	
	Alleles:		HPV persistence
	HLA-DQB1*0201	0.62 (0.4-0.9)	
	HLA-DQB1*0502	2.63 (1.2-5.8)	
	HLA-DRB1*0301	0.45 (0.2-0.8)	
	HLA-DRB1*0807	3.15 (1.1-9.5)	
	HLA-DRB1*1601	5.93 (1.8-20.1)	
	Haplotypes:		
	HLA-DRB1*0807–HLA-DQB1-0402	3.19 (1.1-9.6)	
	HLA-DRB1*1601–HLA-DQB1-0502	7.79 (2.0-29.8)	
	HLA-DRB1*0301-HLA-DQB1*0201	0.44 (0.2-0.8)	
	HLA-DRB1*1102-HLA-DQB1*0301	0.11 (0.0-0.8)	
Ades *et al.* (2008) [34]	**Haplotypes:**		Prevalence of HPV type 16 or 18 infection
	HLA-B7-HLA-DRB1*1501	2.70 (1.30-5.60)	
	HLA-B7-HLA-DQB1*0602	3.22 (1.58-6.58)	
	HLA-B7-HLA-DRB1*1501-HLA-DQB1*0602	2.87 (1.38-5.97)	

Study Setting and Rational

Study Setting

The data for this study was collected in Nunavik, the arctic and sub-arctic region of Northern Quebec. Nunavik covers an area of 507,000 km, but has a population of only 11,000 people. The population lives in 14 communities located along the coast of the Hudson Bay, Hudson Straight, and Ungava Bay. The location of Nunavik and a map of its communities can be seen in figure 6. The population is not only geographically isolated but has a unique culture and sociodemographic makeup. The population of Nunavik is young and growing: around half of the residents are under the age of 20 and the growth rate in the early 2000's was 12.0% (the growth rate in Quebec was 4.3%). 90% of residents of Nunavik self-identify as Inuit. The Inuit population has a high rate of infectious diseases, including STIs, and a rising rate of chronic disease. [40,41]

Figure 6: Location of Nunavik and its communities [42]

Canadian Aboriginal women have a higher risk of developing and dying from cervical cancer than the general Canadian population. [43] Previous analysis of the population of Nunavik suggests that the prevalence of HPV in Nunavik was two-fold higher than the reported rates among the Canadian general population and in women younger than 20 the difference was almost three-fold. [43] The prevalence was similar to that found in other high risk populations, such as Montreal University students. The overall prevalence of HPV infection in Nunavik women was 29.8%, and in women 15-19 years old the prevalence was 58%. HR-HPV types are also more prevalent in Nunavik women than in the general Canadian population: the prevalence of HR-HPV was 20.4% and in women 15-19 years old the prevalence was 47%. The population showed a U-shaped HPV prevalence curve. [43] In women 40 or more years old, HR-HPV infection was associated with the number of sexual partners. In women younger than 40, HR-HPV infection was associated with the number of sexual partners, the age at first sexual intercourse, and marital status. [41]

Rational

Previous analysis of a population of Inuit women in Nunavik, Quebec suggests this population is at a greater risk of HPV infection than the rest of Canada. This increased risk could be due to behavioural differences between the two populations, increased burden of infection with high risk HPV types or variants, and/or increased host susceptibility. Analysis of behavioural risk factors for HPV infection and the prevalence of different HPV types have been completed on this population, and therefore this study will focus on susceptibility of Inuit women to HPV infection. [41,43] It is possible that host susceptibility, in particular HLA polymorphisms, may play a part in the increased risk of HPV prevalence in the Inuit population.

In the present study, subjects were followed for two to seven years and therefore period prevalence, incidence of new infections, and duration of infection can be calculated. Previous studies have investigated the association between HLA polymorphisms and HPV cumulative prevalence (also called cumulative risk) and

persistence/duration. This study will also determine the association between HLA polymorphisms and HPV incidence and concurrent multiple HPV infections.

It is important to investigate why the population of Nunavik, Quebec (and other Canadian Inuit populations) are at an increased risk of acquiring HPV and developing cervical cancer. Host susceptibility may be one of the factors that places this population at an increased risk of HPV infection.

Methodology

Hypotheses

- In a population, HLA allele prevalence depends on the ethnicity of the population. All participants in this study self-identified as Inuit and therefore the prevalence of HLA alleles, haplotypes and genotypes in this study population will likely vary from other populations in Canada.
- According to previous literature, HLA-G alleles may be associated with an increased (or decreased) risk and/or duration of HPV infection, but HLA-E alleles will likely have no association with HPV infection.
- HLA-DQB1*03, HLA-DRB1*1501, and HLA-DQB1*0602 alleles will likely be associated with an increased risk and/or duration of HPV infection, and the HLA-DRB1*13 alleles will be associated with a decreased risk and/or duration of HPV infection.

Objectives

The objectives of this study are:

1. To describe the prevalence of HLA class I (HLA-B*07, HLA-G, and HLA-E) and HLA class II alleles (HLA-DQB1*03, HLA-DQB1*0602, HLA-DRB1*1501, and HLA-DRB1*13), haplotypes and HLA-G genotypes in a sample of Inuit women in Nunavik, Quebec.
2. Determine if there is an association between HLA alleles, haplotypes and/or genotypes and the prevalence, incidence, and duration of HPV infection in this population.
3. The association between HLA alleles and/or haplotypes and multiple HPV infections and squamous intraepithelial lesions will also be determined.

Study Design

Overview

The data used for this analysis was taken from a prospective cohort of Inuit women living in communities along the Ungava and Hudson Bays in Nunavik, Quebec.

Target population

The target population for this study was all Inuit women aged 15-69 living in Nunavik between January 2002 and December 2007. In 2001, the population of Nunavik was 9600, which included approximately 4320 female Inuit residents (90% of this population self-identify as Inuit and approximately half of Inuit residents are female). In 2006, the population grew to 10570 which included approximately 4760 female Inuit residents.

Eligibility Criteria

Women were eligible for this study if they:

- Self-identified as Inuit
- Were born in Nunavik, Quebec
- Were between 15 and 69 years of age
- Had an intact uterus and no current referral for hysterectomy
- Had not used vaginal medication in the two days prior to cohort entry
- Had not received treatment for cervical disease in the six months prior to cohort entry
- Were no more than 12 weeks pregnant

Subject Recruitment

The sampling frame was all women (aged 15-69) presenting for a regularly scheduled Pap test at a clinic in one of the four participating communities in Ungava Bay (Kuujjuak, Kangiqsualujjuaq, Kangiqsujuaq, and Kangirsuk) between January 2002 and December 2007. Subjects were followed until July 2010. A small number of women were recruited from a mobile mammography program (between August and October 2004) along the

coasts of Ungava and Hudson Bay. Nurse practitioners systematically asked women (who were not already enrolled) to participate in the study and interested women were assessed for eligibility. All eligible women provided written informed consent. (See appendix 1) Any women who wanted to withdraw from the study could do so at any time.

Subject Follow-up

This study did not use a pre-set testing interval and women were not followed up in a systematic manner. Cervical specimens were collected at any subsequent, regularly scheduled clinic visits for a gynecological exam. Therefore, the follow up reflected the women's normal Pap smear history. Every four months, members of the research team visited the participating communities. During their visits they made announcements on local FM radio stations encouraging women to get yearly Pap tests.

Ethical Considerations

Ethics approval for this study was obtained from the McGill Institutional Review Board and the Tulattavik Health Center (who provided services to study participants). Written informed consent was obtained from all study participants.

Data Collection

Questionnaire

At enrolment, nurse practitioners collected sociodemographic and behavioural characteristics from a standardized questionnaire. The questionnaire was provided in English, French, and Inuktitut (the Inuktitut version was back-translated into English to ensure accurate translation). The questionnaire was validated by a steering committee, which included Nunavik community members, the Tulattavik Health Center, and the Nunavik Regional Board of Health and Social Services. It was also piloted with a group of ten Inuit women in order to ensure its comprehensibility and ease of use. (See appendix 2)

Medical Chart Review

Research team members reviewed the medical charts of the enrolled subjects and extracted additional information on the subject's medical history (using a standardized data retrieval form). The extracted information included: reproductive history, STI diagnosis, major surgeries, organ transplants, immunosuppression, the use of steroid medication, and results of Pap tests performed throughout the study period. (See appendix 3)

Cervical Specimen Testing and DNA Extraction

Cervical specimens were collected at each visit for a Pap test, including at the entry to the study. Ectocervical and endocervical cells were collected using a Dacron swab for both a Pap smear and HPV-DNA testing. Once used for the Pap smear, the swab was immersed in a tube of 1.5mL of PreservCyt (Cytyc Corporation, Boxborough, MA) and agitated to release the cells. PreservCyt is a methanol-based liquid preservative and was used as a transport medium to preserve the integrity of the exfoliated cells. The cells were kept at 4^{o}C before being placed on wet ice and transported to the Montréal laboratory of François Coutlée for HPV typing (CHUM – Hospital Notre Dame).
The cervical smear slides were sent to Québec City to be read by an experienced cytopathologist who was blind to the subject's HPV status. The Bethesda classification system for cytological diagnoses was used as the basis for cytopathology reports. [44] Cytology results were sent to the treating physician in the patient's respective community and recorded in their medical chart.

HPV-DNA Testing and Typing

The cervical cells were separated from the preservative medium using centrifugation at 13000xg for 15 minutes at 22^{o}C. The supernatant was discarded and the cell pellet left to dry. Once dry, the cells were re-suspended in 300µL of 20mM Tris-EDTA buffer (pH 8.3). The DNA was purified with Master pure[74] (Epicenter, Madison, WI).

HPV DNA was detected through PCR amplification with PGMY09-PGMY11 primers and quality-controlled reverse line blot assay (Roche Diagnostics). Twenty-seven genital HPV genotypes (6, 11, 16, 18, 26, 31, 33, 35, 39, 40, 42, 44, 45, 51, 52, 53, 54, 56, 57, 58, 59, 66, 68, 73, 82 (IS39 and MM4 subtypes), 83, and 84) were identified using oligonucleotide hybridization. In April 2004 an extended line blot was introduced and 10 additional HPV genotypes (34, 61, 62, 67, 69, 70, 71, 72, 81, and 84 (CP6108)) were added.

A 268-bp region of the β-globin gene was amplified using GH20 and PC04 primers in order to assess the quality of DNA samples. Samples with a negative β-globin test were classified as having a baseline HPV result below acceptable quality levels.

Each specimen was assigned a unique identifier and laboratory personnel were blinded to information on study subjects. Standard procedures were followed at all times to prevent contamination.

HLA Typing

Enrolment cervical specimens were used for HLA typing. Low resolution typing was used to type all HLA class II alleles and HLA-B*07. Once the DNA was purified, HLA alleles HLA-B*07, HLA-DQB1*03, HLA-DQB1*0602, HLA-DRB1*1501, and HLA-DRB1*13 were typed using a PCR technique that used sequence-specific primers (PCR-SSP), according to a well-established procedure. [45] In this procedure, pairs of oligonucleotides were used as primers for amplification of the HLA alleles. As a control, a 796-basepair (bp) fragment from the third intron of HLA-DRB1 was also amplified in order to verify that the amplification was successful. The PCR product was viewed using electrophoresis. The DNA was stained with ethidium bromide, run on a 1% agarose gel and visualized using UV illumination. The investigators were blinded to HPV and questionnaire results.[28]

Class I HLA-E and G alleles were typed using high resolution DNA typing and therefore all of the alleles were captured. HLA-G and HLA-E alleles were determined through direct DNA sequencing of the nucleotide regions encompassing the HLA-E exon 3 (218bp) and

HLA-G exons 2-4 (1718bp) using purified DNA from cervical samples as described previously by Ferguson *et al*. [38]

Data Management

Each subject was assigned a unique identifier at recruitment so that their information from the questionnaire and medical chart review could be linked with HPV-DNA and HLA test results. All identifying information apart from the unique ID was excluded from the analysis dataset in order to ensure confidentiality for all subjects. Members of the research team were the only people allowed to access the data collection sheets and consent forms.

Statistical Analysis

Inclusion in Dataset

Subjects were included in the dataset if they met the eligibility criteria, completed a baseline questionnaire, had a HPV-DNA and HLA test of acceptable quality, and did not withdraw at any time during the study.

Study Variables

HPV Status

HPV genotypes can be classified into HR and LR groups based on their oncogenic potential. This study classified HPV genotypes into HR and LR based on the most recent classification used by the International Agency for Research on Cancer (IARC). [3] Those HPV genotypes that were classified as probable or possible risk where grouped with the HR genotypes and those that were of indeterminate risk were grouped with LR genotypes. The HR group included HPV type 16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 67, 68, 69, 70, 73 and 82. The LR group included HPV type 6, 11, 34, 40, 42, 44, 54, 61, 62, 64, 71, 72, 81, 83, 84 and 89.

The HPV genotypes were also grouped according to alpha-papillomavirus species. The three alpha groups were formed based on phylogenetic analyses done by de Villiers *et*

al. [4] Alpha group 1 included α-1, α-8, α-10, and α-13, alpha group 2 included α-5, α-6, α-7, α-9, and α-11, and alpha group 3 included α-3 and α-15. Group 2 includes HR HPV types and groups 1 and 3 include LR HPV types. Groups 1 and 2 include cervical HPV types and group 3 includes vaginal HPV (that can cause vaginal warts). [4] These groups were used for comparison to determine if any HLA alleles recognized epitopes on similar HPV species.

HLA Polymorphisms

To estimate the effect of each allele, subjects who tested positive for an allele (homozygous or heterozygous) were compared to those who tested negative for that allele. Comparing those who are homozygous or heterozygous for an allele to those in which the allele is absent is called the dominant allele model. To estimate the effect of a haplotype or genotype, subjects who tested positive for both alleles in the haplotype or genotype were compared to those who tested negative for both alleles. The alleles HLA-B*7, HLA-DQB1*03, HLA-DQB1*0602, HLA-DRB1*1501, and HLA-DRB1*13 were chosen a priori, because in previous literature they have been associated with both HPV infection and cervical cancer. HLA-E and HLA-G alleles have been less frequently studied and therefore high resolution DNA testing was used to determine all the HLA-G and HLA-E alleles present in the population.

Independent Variables

The questionnaire that study participants completed at study entry provided information on sociodemographic, lifestyle, and sexual characteristics as well as medical history. The medical chart review provided additional information. The main independent variables used in this study (as well as the categorization scheme for the non-continuous variables) are listed below.

<u>Sociodemographic Variables:</u>

- Age

- Marital Status (0=married or living with parents, 1=single, divorced, or widowed)
- Education (0=less than grade 9, 1=grade 9 or more)
- Employment Status (0=unemployed, 1=employed)

Lifestyle variables:

- Current Smoker (0=no, 1=yes)
- Current Alcohol (0=no, 1=yes)
- Current Birth Control (0=no, 1=yes)

Sexual History:

- Age at first sexual intercourse
- Number of lifetime sexual partners (0=fewer than 10, 1=10 or more)
- Number of sexual partners in the past year (0=0, 1=one or two, 3=three or more)

Medical history:

- Lifetime number of live births
- Self-reported STI history (0=no, 1=yes)
- History of Pap test in the past 3 years (0=no, 1=yes)

Coverage of the Target Population and Selection Bias

In order to evaluate the coverage of the target population, 2006 Canadian census data was used. In order not to overburden the nurse practitioners in the communities, who were recruiting participants as well as their regular clinic duties, no information on the total number of women asked to participate or the sociodemographic data of those women who declined to participate was recorded. The study was designed to be as unintrusive and undemanding as possible. The simple research protocol also enhanced the likelihood of sustainability of the study in an area where there is a high turnover rate of the health care staff. A partial evaluation of selection bias was conducted where the characteristics of the study population was compared to published statistics regarding Inuit female residents of Nunavik. [46]

Outcomes

Odds ratios (and their corresponding 95% confidence intervals) were used to determine the association between HLA alleles/haplotypes/genotypes and the following binary outcomes: HPV period prevalence, HPV duration (or persistence), multiple HPV infections, and SIL. Odds ratios were calculated using unconditional logistic regression. Rate ratios (and their corresponding 95% confidence intervals) were used to determine the association between HLA alleles/haplotypes/genotypes and HPV incidence. Rate ratios were calculated using Poisson regression. Crude and age-adjusted odds and rate ratios were determined for all comparisons. It was decided to adjust for age because age is the strongest determinant of an HPV infection and age had no missing values in our dataset (no multiple imputations necessary). The alleles, haplotypes, and genotypes were only used in the comparisons if they were present in more than 5% of the population or they were associated with HPV infection in previous literature. SAS version 9.2 and R version 2.11.1 statistical programs were used for data analysis.

Period Prevalence

Women who tested positive for any HPV at any time over the follow-up period (including those who were HPV positive at baseline) were compared to women who tested negative for HPV at baseline and remained negative for the rest of the follow-up period.

Period prevalence was also calculated for LR HPV infections only, HR HPV infections only, high-risk HPV 16 and 18, and alpha papillomavirus groups 1, 2 and 3. In each of the comparisons, the comparison group included women who tested negative for HPV at baseline and remained negative for the rest of the follow-up period.

Both crude and age-adjusted ORs will be calculated. Age-adjusted rates were calculated by adding age at baseline to the regression equation.

Period Prevalence Sensitivity Analysis

Three period prevalence sensitivity analyses were performed where:

1. The cohort was restricted to women with three or more visits
2. The cohort was restricted to those women who had a high likelihood of exposure to HPV (women with 10 or more lifetime partners)
3. The cohort was restricted both to women with three or more visits and women with 10 or more lifetime partners

The sensitivity analyses were performed for overall HPV infections and the three alpha papillomavirus groups.

HPV Duration/Persistence

HPV persistence was defined as a woman who tested positive for the same HPV type on two or more visits over the time period. A transient HPV infection was defined as a woman who tested positive once for an HPV infection at any time over the study period. Three comparisons were made:

1. Women with persistent infections were compared to women with transient infections and HPV-negative women
2. Women with persistent infections were compared to women with transient infections
3. Women with persistent infections were compared to HPV-negative women

The second is the most valuable comparison, because an association here indicates that the allele has an effect on whether an infection is cleared or if it persists. These three comparisons were also performed for HR alpha papillomavirus group 2 infection persistence. The cohort was restricted to only those women with three or more visits, to avoid detection opportunity bias. Again, both crude and age-adjusted ORs will be calculated.

Persistence Sensitivity Analysis

A sensitivity analysis for persistence (of any HPV and for HR alpha papillomavirus group 2) will be performed in which the definition for persistence will be more stringent. In the

sensitivity analysis, persistence will be defined as a woman who tests positive for the same HPV type on two or more consecutive visits.

Multiple HPV infections

A woman was considered to have a co-infection if she tested positive for two different HPV types on the same visit (this is often called concurrent co-infections). Women who had multiple HPV types over the follow-up period but not on the same visit were not considered to have a co-infection.

Three comparisons were made:

1. Women with co-infections were compared to women who had single HPV infections and HPV-negative women
2. Women with co-infections were compared to women who had single HPV infections
3. Women with co-infections were compared to HPV-negative women

The second is the most valuable comparison, because an association here indicates that the allele has an effect on whether a woman has one or more infections.

Crude and age-adjusted ORs will be calculated.

Abnormal Cytology

Women with any SIL (HGSIL and LGSIL) cytology result over the follow-up period were compared with women who had normal (no SIL) cytology results for the entire study period. Women with a LGSIL cytology result at any point over follow-up were compared with those who had no SIL cytology result. This comparison was also completed for women with HGSIL.

All cytology comparisons were limited to only those women who were HPV positive, because HPV-positivity is a prerequisite for a SIL. Crude and age-adjusted ORs will be calculated.

HPV Incidence

HPV incidence was calculated for any HPV infection and for the three alpha-papillomavirus species groups. The incidence rate was calculated as the number of new HPV infections per 1000 woman-months (WM) of follow-up. Women did not contribute to the person-time for an infection if they tested positive for an infection of that type at enrolment.

Women months were calculated for each woman based on the time from enrolment to the first positive test (of the infection of interest) or until the end of follow-up (the last HPV-DNA negative test) if they did not test positive over the study period. The date of positivity was estimated to be halfway in between the positive HPV-DNA test and the preceding negative HPV-DNA test. Crude and age-adjusted rate ratios, as well as their corresponding confidence intervals, were calculated using SAS version 9.2 Poisson regression.

Missing Data

Approximately one-fourth of HPV-DNA tests were missing a date. Missing baseline HPV-DNA test dates were given the date that the questionnaire was filled out. Using the medical chart review data, if there was a cytology test for a corresponding missing HPV-DNA test then the date of the HPV-DNA test was given the date of the cytology test. If there was no corresponding cytology test, then the date was estimated based on the order in which the tests were received at the lab for analysis. It is reasonable to assume that the date of a missing test occurred between the preceding test and the subsequent test and therefore missing dates were assigned to the midpoint between the dates of the preceding and following tests.

Results

Cohort Recruitment and Study Eligibility

667 women were recruited into the cohort between January 2002 and December 2007 and followed until July 2010. 548 women met the eligibility criteria (completed a baseline questionnaire and had a HPV-DNA and HLA test of acceptable quality).

Coverage of Target Population

This study captured 40.9% of the target population. The coverage was 42.5% in Kuujjuak, 54.3% in Kangiqsualujjuaq, 23.9% in Kangiqsujuaq, and 35.2% in Kangirsuk. The study captured 30.6% of women aged 15 to 19, 48.0% of 20 to 24 year olds, 50.8% of 25 to 44 year olds, and 24.8% of women 45 years old and older. The age distribution of the study population roughly mirrors the age distribution of the target population; this was also true for each village individually. Figure 7 shows the age distribution of the study population and the target population, which was determined using the 2006 Canadian census. [46]

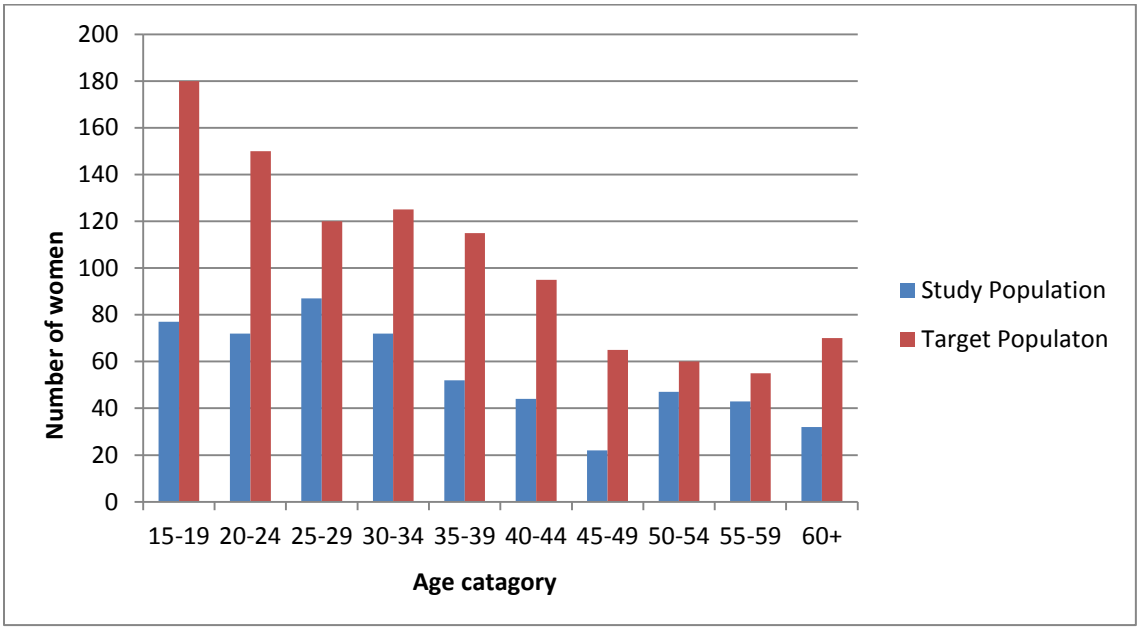

Figure 7: Age Distribution of Study and Target Population

Characteristics of Study Population

Sociodemographic Characteristics

Just over half of the study population was married or living with their partner. Thirty percent of the population was unemployed and forty percent had less than a grade 9 education. (Table 5)

Table 5: Distribution of Sociodemographic Characteristics

Independent variable	n	N	%	% missing
Marital status:				
Married or living with partner	299	528	56.63	3.65
Single	229	528	43.37	
Employed:				
Yes	367	524	70.04	4.38
No	157	524	29.96	
Education:				
Less than grade 9	211	517	40.81	5.66
Grade 9 or more	306	517	59.19	

Smoking and Alcohol

Three quarters of the study population were current smokers and 66 percent of the population reported current alcohol use. (Table 6)

Table 6: Distribution of Smoking and Alcohol Characteristics

Independent variable	n	N	%	% missing
Smoker:				
Yes	386	528	73.11	3.65
No	142	528	26.89	
Alcohol use:				
Yes	353	530	66.60	3.28
No	177	530	33.40	

Reproductive Characteristics

The percentage of women with none, one or two, and three or more lifetime deliveries was 17%, 29% and 54%, respectively. 36% of the population reported using birth control and 67% of the population self-reported having had a sexually transmitted infection (STI). (Table 7)

Table 7: Distribution of Reproductive Characteristics

Independent variable	n	N	%	% missing
Lifetime deliveries:				
None	94	543	17.31	0.91
One or two	158	543	29.10	
Three or more	291	543	53.59	
Birth control:				
Yes	187	517	36.17	5.66
No	330	517	63.83	
Self-reported history of STI:				
Yes	350	519	67.44	5.29
No	169	519	32.56	

Sexual Behaviour Characteristics

The percentage of women whose age at first intercourse was less than 15, 15 to 20, and over 20 years old was 38%, 59%, and 3% respectively. 35% of the study participants had 10 or more lifetime sexual partners. The percentage with none, one or two, and three or more sexual partners in the past year was 11%, 62%, and 28% respectively. (Table 8)

Table 8: Distribution of Sexual Behaviour Characteristics

Independent variable	n	N	%	% missing
Age at first sexual intercourse:				
<15	192	506	37.94	7.66
15-20	299	506	59.09	
>20	15	506	2.96	
Number of lifetime sexual partners:				
Less than 10	320	495	64.65	9.67
10 or more	175	495	35.35	
Number of sexual partners in the past year:				
None	54	505	10.69	7.85
1 or 2	311	505	61.58	
3 or more	140	505	27.72	

Follow-up of Participants

548 eligible women attended their first visit and had an adequate HPV-DNA and HLA test. 67% of these women returned for a second visit and 50% returned for a third visit. Less than 20% of the original 548 women returned for the sixth visit. The percentage of women per visit post-enrollment is shown in figure 8. The average length of time between follow-up visits was 1.36 years (or 16.3 months).

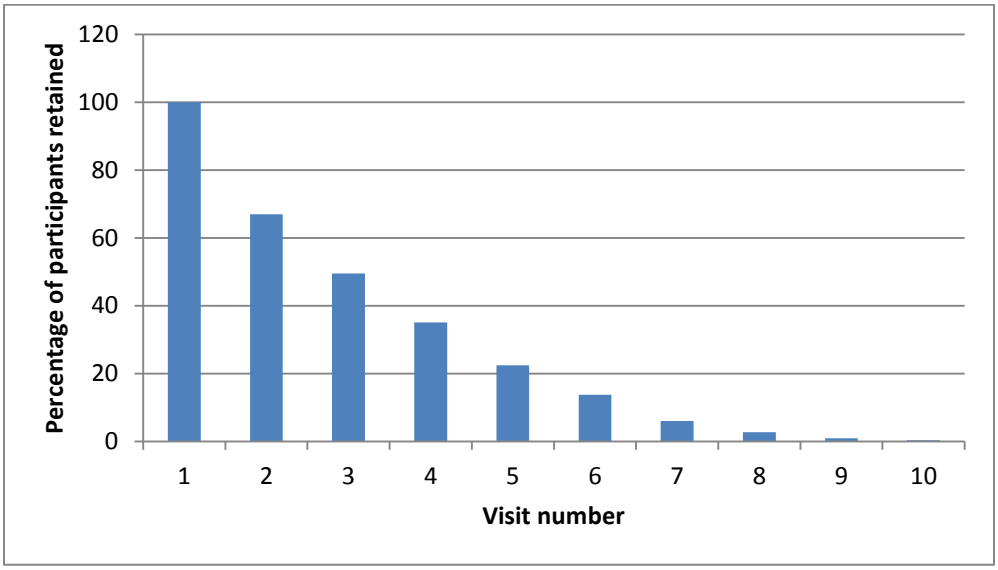

Figure 8: Percentage of Women Per Visits Post Enrollment

Distribution of HLA Alleles and Haplotypes

Alleles

Table 9 shows the distribution of HLA-B*07 and class II HLA alleles. This table shows the number of women with each allele. The highlighting indicates those alleles that will be used in future comparisons. Just less than 5% of the population tested positive for the HLA-B*07 allele. The most common class II HLA allele in this population was HLA-DQB1*03. It was present in 94% of women in this population. The other three class II HLA alleles (that were tested for) were present in less than 5% of the population: HLA-DQB1*0602 (3.6%), HLA-DRB1*13 (4.7%) and HLA-DRB1*1501 (3.8%). (Table 9)

Table 9: The distribution of HLA-B*07, DQB1*03, DQB1*0602, DRB1*13, and DRB1*1501 alleles

Allele	n	N	%
HLA-B*07	23	514	4.47
HLA-DQB1*03	486	516	94.19
HLA-DQB1*0602	19	523	3.63
HLA-DRB1*13	25	532	4.70
HLA-DRB1*1501	20	527	3.80

Table 10 shows the distribution of HLA-E and HLA-G alleles in this population; for each woman both HLA-E and HLA-G alleles were determined and therefore the denominator is twice the number of the women in the cohort. Two HLA-E alleles were found in this population. HLA-E*0103 was the most common, found in 66% of the women, and HLA-E*0101 was found in 33% of the women. There were eight different HLA-G alleles that were found in this population. The most common HLA-G alleles were HLA-G*010401 (50%), HLA-G*010101 (31%), and HLA-G*010102 (15%). The HLA-G wild type is HLA-G*0101. 48% of the alleles in this population were the wild type. (Table 10)

Table 10: Distribution of HLA-E and HLA-G alleles

Allele	n	N	%
HLA-E*0101	343	1044	32.85
HLA-E*0103	693	1044	66.38
HLA-G*010101	328	1044	31.42
HLA-G*010102	159	1044	15.23
HLA-G*010103	11	1044	1.05
HLA-G*010108	4	1044	0.38
HLA-G*0103	4	1044	0.38
HLA-G*010401	526	1044	50.38
HLA-G*0106	14	1044	1.34
HLA-G*0107	1	1044	0.10

Haplotypes

Table 11 shows the HLA haplotypes of the 5 HLA alleles which were chosen *a priori* because they were often associated with HPV and cervical cancer. None of the HLA haplotypes in table 11 were present in more than 5% of the population. HLA-DRB1*1501-DQB1*0602 and B*07-DRB1*1501-DQB1*0602 have been linked to cervical cancer in previous literature and therefore will be used in further analysis.

Table 11: HLA-B, DQB1, and DRB1 Haplotype Distribution

HLA Haplotype	n	N	%
B*07-DQB1*03	18	494	3.64
B*07-DRB1*13	1	510	0.20
B*07-DRB1*1501	8	502	1.59
B*07-DRB1*0602	8	479	1.67
DQB1*03-DQB1*0602	13	505	2.57
DQB1*03-DRB1*13	15	511	2.94
DQB1*03-DRB1*1501	14	507	2.76
DQB1*0602-DRB1*13	1	517	0.19
DRB1*1501-DQB1*0602	19	511	3.72
DRB1*1501-DRB1*13	1	520	0.19
B*07-DRB1*1501-DQB1*0602	8	487	1.64

Table 12 shows the HLA-G genotypes of the most common three HLA-G alleles in this population (HLA*G-010101, HLA-G*010102, and HLA-G*010401).

Table 12: HLA-G Genotype Distribution

HLA-Genotype	n	N	%
G*010101			
Homozygous	51	522	9.77
Heterozygous	227	522	43.49
Absent	244	522	46.74
G*010102			
Homozygous	8	522	1.53
Heterozygous	138	522	26.44
Absent	376	522	72.03
G*010401			
Homozygous	131	522	25.10
Heterozygous	268	522	51.34
Absent	123	522	23.56

HPV Period Prevalence

HPV period prevalence was determined for any HPV, by oncogenic grouping and by alpha-papillomavirus species groups. 548 women were included in this analysis. 287 of these women were HPV-negative at baseline and remained HPV-negative over the follow up period. Of the 261 women who were HPV-positive at any point over the study period: 88 women had only LR infections; 171 women had only HR infections; and 72 women had HPV type 16 or 18 infections.

Only one HLA allele, HLA-G*010102, was significantly associated with a decreased risk of an overall HPV infection (OR=0.64 95% CI=0.42-0.98). The homozygous HLA-G*010102 genotype was also associated with a decreased risk of overall HPV infection (OR=0.25 95% CI=0.05-1.32) but the association was not significant. The HLA-DQB1*0602 and DRB1*1501 alleles were associated with a decreased risk of overall HPV infection, but the associations were not significant. (Table 13)

Table 13: Any HPV Type Period Prevalence by HLA Alleles

Allele	HPV Negative n	HPV Positive n	Any HPV type Age-Adjusted OR (95% CI)
Cohort (n=548)	287	261	
HLA-B*07	12	11	0.68 (0.25-1.84)
HLA-E*0101	151	142	1.00 (0.68-1.47)
HLA-E*0103	246	222	0.93 (0.49-1.76)
HLA-G*010101	144	133	1.02 (0.70-1.49)
HLA-G*010102	84	67	**0.64 (0.42-0.98)**
HLA-G*010103	4	7	1.38 (0.36-5.30)
HLA-G*010401	206	186	1.38 (0.88-2.17)
HLA-G*0106	6	7	0.77 (0.24-2.46)
HLA-DQB1*03	250	236	1.66 (0.732-3.83)
HLA-DQB1*0602	13	6	0.37 (0.13-1.06)
HLA-DRB1*13	15	10	0.47 (0.19-1.20)
HLA-DRB1*1501	13	7	0.37 (0.13-1.04)

Period Prevalence of Only LR, Only HR and Type 16/18 Infections

None of the women who had the B*07 allele were positive for only LR infection. The HLA-G*010102 allele was significantly associated with a decreased risk of being infected with only LR HPV infection (age-adjusted OR=0.37 95% CI= 0.16-0.83). None of the associations between HLA alleles and haplotypes and only HR infections and HR HPV type 16/18 infections were significant. (See appendix 4)

Period Prevalence of Alpha Papillomavirus Groups 1, 2 and 3

42 women tested positive for an alpha group 1 (includes α-1, α-8, α-10, and α-13) HPV infections, 209 women tested positive for an alpha group 2 (includes α-5, α-6, α-7, α-9, and α-11) HPV infections, and 126 women tested positive for an alpha group 3 (includes α-3 and α-15) HPV infection. Group 1 and 3 includes LR HPV species and group 2 includes HR HPV species.

The HLA-G*010101 allele was associated with a significantly increased risk of LR alpha group 1 (OR=2.23 95% CI=1.08-4.59). Both the homozygous (OR=2.72 95% CI=0.86-8.60) and heterozygous (OR=2.14 95% CI=1.03-4.52) genotypes of this allele both appeared to be associated with an increased risk of alpha group 1 infection. The latter association was significant. None of the HR HPV alpha group 2 associations were significant. The HLA-G*010101 allele was associated with a significantly increased risk of LR alpha group 3 (OR=1.70 95% CI=1.09-2.65). Both the homozygous (OR=1.81 95% CI=0.84-3.90) and heterozygous (OR=1.69 95% CI=1.07-2.67) genotypes of this allele appeared to be associated with an increased risk of alpha group 3 infection, the later association was significant. The homozygous genotype of HLA-G*010401 (OR=0.50 95% CI=0.25-0.97) was significantly associated with a decreased risk of HPV alpha group 3 infections. (Table 14)

Table 14: Alpha Papillomavirus Group Period Prevalence by HLA Alleles, Haplotypes and Genotypes

Allele / Haplotype / Genotype	α Group 1 (α-1, α-8, α-10, α-13) Age-Adjusted OR (95% CI)	α Group 2 (α-5, α-6, α-7, α-9, α-11) Age-Adjusted OR (95% CI)	α Group 3 (α-3, α-15) Age-Adjusted OR (95% CI)
HLA-B*07	n/a	0.84 (0.31-2.25)	0.22 (0.03-1.69)
HLA-E*0101	0.96 (0.49-1.89)	1.04 (0.71-1.52)	0.95 (0.61-1.46)
HLA-E*0103	2.02 (0.46-8.92)	1.13 (0.59-2.14)	1.29 (0.60-2.77)
HLA-G*010101	**2.23 (1.08-4.59)**	1.15 (0.79-1.68)	**1.70 (1.09-2.65)**
HLA-G*010102	0.45 (0.18-1.11)	1.13 (0.74-1.71)	0.90 (0.55-1.47)
HLA-G*010103	1.14 (0.13-10.06)	0.67 (0.16-2.87)	0.39 (0.05-3.43)
HLA-G*010401	0.90 (0.40-2.01)	0.71 (0.46-1.10)	0.74 (0.45-1.21)
HLA-G*0106	n/a	1.54 (0.42-5.69)	2.72 (0.74-9.99)
HLA-DQB1*03	0.99 (0.21-4.64)	0.99 (0.43-2.30)	0.60 (0.25-1.43)
HLA-DQB1*0602	n/a	0.54 (0.18-1.61)	0.95 (0.32-2.84)
HLA-DRB1*13	0.43 (0.05-3.50)	0.99 (0.42-2.34)	1.15 (0.44-3.05)
HLA-DRB1*1501	0.48 (0.06-3.95)	0.64 (0.23-1.80)	1.09 (0.39-3.06)
DRB1*1501-DQB1*0602	n/a	0.54 (0.18-1.62)	0.93 (0.31-2.77)
B*07-DRB1*1501-DQB1*0602	n/a	0.50 (0.10-2.54)	0.41 (0.05-3.50)
G010101:			
Homozygous	2.72 (0.86-8.60)	1.71 (0.89-3.28)	1.81 (0.84-3.90)
Heterozygous	**2.14 (1.03-4.52)**	1.05 (0.71-1.57)	**1.69 (1.07-2.67)**
Absent*	1.00	1.00	1.00
G010102:			
Homozygous	1.03 (0.11-9.32)	0.49 (0.09-2.62)	0.39 (0.05-3.46)
Heterozygous	0.40 (0.15-1.07)	1.18 (0.78-1.80)	0.94 (0.57-1.53)
Absent*	1.00	1.00	1.00
G010401:			
Homozygous	0.68 (0.24-1.89)	0.62 (0.36-1.07)	**0.50 (0.25-0.97)**
Heterozygous	1.02 (0.44-2.37)	0.77 (0.49-1.22)	0.89 (0.53-1.49)
Absent*	1.00	1.00	1.00

*Reference group

Period Prevalence Sensitivity Analysis

Three sensitivity analyses were performed for period prevalence of any HPV and alpha groups 1, 2 and 3: (1) where the cohort was limited to those women with three or more visits, (2) where the cohort was restricted to women with 10 or more lifetime partners, and (3) where both restrictions were put on the cohort.

In the first analysis, the direction of the associations mentioned above remained the same but the majority of the associations moved toward the null, including that between HLA-G*010102 and period prevalence of any HPV (OR=0.75 95% CI=0.42-1.35). In this analysis, the homozygous HLA-G*010101 genotype was significantly associated with an increased risk of any HPV (OR= 2.55 95% CI=1.06-6.13) and alpha papillomavirus groups 1 (OR=4.41 95% CI=1.12-17.44), 2 (OR=2.83 95% CI=1.14-7.06) and 3 (OR=3.29 95% CI=1.06-10.15).

In the second analysis, the direction of the alpha group associations remained the same. The homozygous HLA-G*010401 genotype was significantly associated with a decreased risk of any HPV and LR alpha group 3. The direction of the associations between the HLA-G*010102, DQB1*0602 and DRB1*1501 alleles and any HPV all changed.

In the third analysis, again the direction of the alpha group associations remained the same but the direction of the associations between the HLA-G*010102, DQB1*0602 and DRB1*1501 alleles and any HPV changed. (See appendix 4)

HPV Persistence

No HLA alleles, haplotypes or genotypes were significantly associated with HPV persistence. HLA-B*07 (OR=7.30 95% CI=0.86-61.96) and HLA-G*010103 (OR=2.75 95% CI=0.28-27.39) were associated with an increased risk of HPV persistence versus transient infections and HLA-G*0106 (OR=0.42 95% CI=0.04-4.82) was associated with a decreased risk of HPV persistence versus transient infections. (See appendix 4)

HR HPV Alpha Group 2 Persistence

Again no HLA alleles, haplotypes or genotypes were significantly associated with HR HPV alpha group 2 persistence. HLA-B*07 (OR=3.29 95% CI=0.62-17.45) and G*010103 (OR=1.86 95% CI=0.16-21.61) were associated with an increased risk of alpha group 2 persistence, but the associations were not significant. Also, the HLA-DQB1*03 allele was associated with a decreased, but non-significant, risk of alpha group 2 persistence (OR=0.41 95% CI=0.10-1.72). (See appendix 4)

HPV Persistence Sensitivity Analysis

A sensitivity analysis was performed for any HPV persistence and for HR alpha group 2 persistence, in which the definition of persistence was limited to only those women who tested positive for the same HPV infection on two or more consecutive visits. In these analyses, many of the associations moved towards the null or changed directions. The association between HLA-G*010103 (OR=2.24 95% CI=0.13-37.26) remained positively associated with any HPV persistence, but the association was still non-significant. (See appendix 4)

Multiple HPV Infections

548 women were included in the analysis of multiple HPV infections. Out of the 261 women who tested positive for an HPV infection over the follow-up period, 117 women had concurrent co-infections (when a women tests positive for two different HPV types on the same visit). The remaining 144 HPV positive women had mono-infections. HLA-G*010101 allele (OR=1.54 95% CI=0.90-2.66) and its heterozygous genotype (OR=1.83 95% CI=1.03-3.21) were associated with an increased risk of multiple infections versus single infections. The latter was significant. The HLA-G*010102 allele was associated with a significantly decreased risk of multiple infections versus HPV-negative (OR=0.45 95% CI=0.24-0.82). The homozygous genotype of HLA-G*010401 was associated with a significantly decreased risk of multiple infections versus HPV-negative (OR=0.44 95% CI=0.21-0.92). The HLA-DQB1*0602 and DRB1*1501 alleles and their

respective haplotypes were associated with an increased, but not significant, risk of multiple infections. (Table 15)

Table 15: Multiple HPV Infections by HLA Alleles, Haplotypes and Genotypes

Allele / Haplotype / Genotype	Co-infection vs. Mono-infection Age-Adjusted OR (95% CI)	Co-infection vs. HPV-negative Age-Adjusted OR (95% CI)
HLA-B*07	3.91 (0.66-23.15)	1.20 (0.38-3.74)
HLA-E*0101	1.04 (0.60-1.79)	0.92 (0.54-1.57)
HLA-E*0103	0.84 (0.34-2.08)	0.90 (0.38-2.16)
HLA-G*010101	1.54 (0.90-2.66)	1.36 (0.80-2.32)
HLA-G*010102	0.72 (0.38-1.35)	**0.45 (0.24-0.82)**
HLA-G*010103	1.08 (0.21-5.67)	1.71 (0.29-10.02)
HLA-G*010401	0.65 (0.34-1.26)	1.09 (0.60-1.98)
HLA-G*0106	n/a	1.73 (0.50-6.04)
HLA-DQB1*03	0.91 (0.28-2.96)	1.59 (0.54-4.71)
HLA-DQB1*0602	2.96 (0.46-19.21)	0.52 (0.15-1.79)
HLA-DRB1*13	1.37 (0.33-5.66)	0.50 (0.16-1.58)
HLA-DRB1*1501	2.81 (0.42-18.63)	0.49 (0.14-1.71)
DRB1*1501-DQB1*0602	2.94 (0.46-18.91)	0.52 (0.15-1.78)
B*07-DRB1*1501-DQB1*0602	3.00 (0.12-72.29)	0.47 (0.05-4.22)
G010101:		
Homozygous	1.16 (0.50-2.70)	1.80 (0.76-4.30)
Heterozygous	**1.83 (1.03-3.21)**	1.44 (0.84-2.48)
Absent*	1.00	1.00
G010102:		
Homozygous	1.86 (0.15-22.86)	1.21 (0.16-9.25)
Heterozygous	0.71 (0.39-1.27)	1.14 (0.63-2.07)
Absent*	1.00	1.00
G010401:		
Homozygous	0.70 (0.33-1.49)	**0.44 (0.21-0.92)**
Heterozygous	1.00 (0.53-1.85)	0.87 (0.46-1.60)
Absent*	1.00	1.00

*Reference category

High- and Low-Grade Squamous Intraepithelial Lesion Period Prevalence

A total of 257 women were included in the analysis of Squamous Intraepithelial Lesions. 59 women had an SIL, 37 of those women had a LGSIL and 22 had a HGSIL.

None of the women with the B*07 allele had a SIL. The HLA-G*010102 allele was associated with a decreased risk of HGSIL (OR= 0.54 95% CI=0.17-1.67), but the association was not significant. HLA-G*0106 was associated with an increased risk of any SIL, LGSIL, and HGSIL, but none of the associations were significant.

The HLA-DQB1*03 allele was associated (OR= 0.46 95% CI=0.13-1.69), non-significantly, with a decreased risk of LGSIL. The HLA-DRB1*13 allele was associated with an increased risk of any SIL (OR= 4.34 95% CI=1.27-14.78), LGSIL (OR= 5.56 95% CI=1.43-21.63), and HGSIL (OR= 3.25 95% CI=0.58-18.30). The first two associations were significant. The homozygous G*010401 genotype was associated with an increased risk of both LGSIL and HGIL. (Table 16)

Table 16: Any SIL, LGSIL, and HGSIL by HLA Alleles

Allele	Any SIL vs. Normal Cytology Age-Adjusted OR (95% CI)	LGSIL vs. Normal Cytology Age-Adjusted OR (95% CI)	HGSIL vs. vs. Normal Cytology Age-Adjusted OR (95% CI)
HLA-E*0101	1.81 (0.94-3.47)	2.11 (0.93-4.74)	1.44 (0.56-3.63)
HLA-E*0103	0.59 (0.23-1.53)	0.47 (0.16-1.39)	0.95 (0.20-4.52)
HLA-G*010101	0.86 (0.46-1.59)	1.08 (0.51-2.29)	0.63 (0.26-1.55)
HLA-G*010102	0.84 (0.42-1.67)	1.07 (0.48-2.41)	0.54 (0.17-1.67)
HLA-G*010103	n/a	n/a	n/a
HLA-G*010401	0.94 (0.47-1.85)	0.88 (0.39-1.98)	0.99 (0.36-2.72)
HLA-G*0106	2.40 (0.55-10.50)	1.52 (0.25-9.31)	3.74 (0.63-22.33)
HLA-DQB1*03	0.66 (0.20-2.16)	0.46 (0.13-1.69)	1.28 (0.15-10.84)
HLA-DQB1*0602	1.42 (0.25-7.86)	1.12 (0.12-10.39)	1.92 (0.21-17.57)
HLA-DRB1*13	**4.34 (1.27-14.78)**	**5.56 (1.43-21.63)**	3.25 (0.58-18.30)
HLA-DRB1*1501	1.02 (0.19-5.48)	0.79 (0.09-7.15)	1.35 (0.15-12.00)

HPV Incidence

HPV incidence was determined for any incident HPV infection and by alpha groups. None of the HLA alleles, haplotypes or genotypes were significantly associated with any incident HPV infections.

Incident Alpha Groups Infections

Two alleles were associated with a significantly increased rate of LR alpha group 1 infection: HLA-G*010101 (RR=2.57 95% CI=1.01-6.53) and HLA-DRB1*13 (RR=3.53 95% CI=1.17-10.70). The heterozygous HLA-G*010101 genotype was also associated with a significantly increased rate of LR alpha group 1 infection. No alleles, haplotypes or genotypes were significantly associated with incident HR alpha group 2 infections. HLA-G*010101 and its homozygous and heterozygous genotypes were all significantly associated with an increased rate of LR alpha group 3 infections. The HLA-G*010401 homozygous genotype was associated with a significantly decreased rate of alpha group 3 infections (RR=0.39 95% CI=0.16-0.95). (Table 17)

Table 17: HPV Incidence by HLA Alleles, Haplotypes and Genotypes

Allele / Haplotype / Genotype	Any HPV infection	α Group 1 (α-1, α-8, α-10, α-13)	α Group 2 (α-5, α-6, α-7, α-9, α-11)	α Group 3 (α-3, α-15)
	Age-Adjusted Rate Ratio (95% CI)	Age-Adjusted Rate Ratio (95% CI)	Age-Adjusted Rate Ratio (95% CI)	Age-Adjusted Rate Ratio (95% CI)
HLA-B*07	0.77 (0.28-2.09)	1.56 (0.36-6.70)	0.76 (0.18-3.12)	0.92 (0.29-2.95)
HLA-E*0101	1.02 (0.68-1.52)	1.25 (0.54-2.90)	1.06 (0.65-1.74)	1.25 (0.73-2.14)
HLA-E*0103	0.90 (0.47-1.74)	1.07 (0.25-4.58)	0.80 (0.38-1.67)	0.64 (0.30-1.35)
HLA-G*010101	1.27 (0.85-1.90)	**2.57 (1.01-6.53)**	1.25 (0.76-2.04)	**2.05 (1.16-3.62)**
HLA-G*010102	0.64 (0.39-1.05)	0.95 (0.38-2.42)	0.95 (0.55-1.64)	0.64 (0.34-1.22)
HLA-G*010103	0.90 (0.22-3.63)	n/a	n/a	0.99 (0.14-7.16)
HLA-G*010401	0.69 (0.45-1.07)	0.63 (0.26-1.54)	0.83 (0.47-1.46)	0.67 (0.38-1.16)
HLA-G*0106	1.47 (0.60-3.63)	n/a	0.28 (0.04-2.06)	0.95 (0.23-3.93)
HLA-DQB1*03	1.84 (0.68-5.00)	0.46 (0.14-1.55)	5.56 (0.77-40.08)	1.25 (0.39-4.02)
HLA-DQB1*0602	0.70 (0.26-1.91)	n/a	0.35 (0.09-1.43)	0.54 (0.13-2.23)
HLA-DRB1*13	1.02 (0.44-2.36)	**3.53 (1.17-10.70)**	0.42 (0.10-1.74)	n/a
HLA-DRB1*1501	0.66 (0.24-1.80)	n/a	0.35 (0.08-1.41)	0.56 (0.14-2.31)
DRB1*1501-DQB1*0602	0.70 (0.26-1.90)	n/a	0.35 (0.09-1.45)	0.55 (0.13-2.27)
B*07-DRB1*1501-DQB1*0602	0.47 (0.07-3.39)	n/a	0.59 (0.08-2.31)	0.82 (0.11-6.04)
G010101:				
Homozygous	1.42 (0.67-3.00)	0.77 (0.09-6.48)	1.61 (0.73-3.57)	**2.86 (1.28-6.38)**
Heterozygous	1.20 (0.74-1.97)	**2.60 (1.00-6.78)**	1.02 (0.60-1.76)	**2.03 (1.08-3.80)**
Absent	1.00	1.00	1.00	1.00
G010102:				
Homozygous	n/a	1.95 (0.26-14.70)	n/a	0.88 (0.12-6.37)
Heterozygous	0.84 (0.49-1.45)	0.78 (0.26-2.34)	1.05 (0.60-1.86)	0.56 (0.27-1.14)
Absent	1.00	1.00	1.00	1.00
G010401:				
Homozygous	0.92 (0.47-1.80)	0.66 (0.18-2.33)	1.06 (0.52-2.18)	**0.39 (0.16-0.95)**
Heterozygous	0.83 (0.47-1.48)	0.63 (0.23-1.72)	0.83 (0.44-1.56)	0.75 (0.41-1.36)
Absent	1.00	1.00	1.00	1.00

*Reference category

Discussion

The data presented here represents a cohort study of HPV Infection in Inuit women form Nunavik, Quebec who were followed from January 2002 to July 2010. The primary results presented here are: (1) HLA allele, haplotype and genotype prevalence, (2) the association between HLA alleles, haplotypes and genotypes and HPV period prevalence, (3) the association between HLA alleles, haplotypes and genotypes and HPV persistence, (4) the association between HLA alleles, haplotypes and genotypes and multiple concurrent HPV infections, (5) the association between HLA alleles, haplotypes and genotypes and LGSIL and HGSIL period prevalence, and (6) the association between HLA alleles, haplotypes and genotypes and HPV Incidence.

HLA Allele and Haplotype Prevalence

Eight HLA-G alleles were found in this population, the three most common alleles were HLA-G*010101 (31%), HLA-G*010102 (15%), and HLA-G*010401 (50%). The five remaining alleles were each found in less than 5% of the population. HLA-DQB1*03 was the most common HLA class II allele (94%), the other three class II alleles investigated were each present in less than 5% of the population.

There are few studies which have outlined the distribution of HLA alleles and haplotypes in Canadian populations and therefore the possibility of comparison is limited. The distribution of HLA alleles and haplotypes in the Inuit population is unique when compared to that found in Montreal (includes women from varied ethnic backgrounds). In Montreal women, the HLA-G*010101 and HLA-G*010102 alleles were also quite prevalent, but HLA-G*010401 (which was the most prevalent HLA-G allele found in the Inuit population) was only found in 9% of the population. [33] The HLA-B*07 allele was more prevalent in the Montreal populations (15-18%) than the Nunavik population (<5%). The HLA-DQB1*0602, DRB1*13, and DRB1*1501 alleles were also found in a greater proportion of the Montreal populations (20%, 17-25%, and 21-25%, respectively) than in the Nunavik one. HLA-DQB1*03 was found in 57-64% of the

Montreal student population, whereas it was found in almost all of the Inuit women (94%). [28,34] (For comparison table see appendix 4)

Two studies investigated the distribution of HLA alleles in native populations of Alaska and Greenland, but they are not exactly comparable to this study because of the difference in allele nomenclature. In the Greenland population, HLA-B*07, DRB1*15 and DQB1*06 were all found in less than 2% of the population. [47] In the Alaskan population HLA-B*0702, DRB1*1501, DRB1*1301 and DQB1*0602 were again found in less than 2% of the population. [48] DQB1*0301 was found in 61.7% of the population of Alaskan Natives. [48] In the Nunavik population HLA-B*07, DRB1*1501, DRB1*03 and DQB1*0602 were all found in less than 5% of the population and HLA-DQB1*03 was found in the majority of the population (94%). These three populations have somewhat similar HLA allele distributions, unfortunately the Greenland and Alaskan studies did not report on HLA typing for HLA-E or G thus precluding comparisons. (See appendix 4)

A study by Matte *et al.* compiled the frequencies of HLA-G in various populations around the world (which can be seen in figure 8). In the majority of the populations the most common alleles were HLA-G*010101, G*010102, and G*010401. In all populations but one (German/Croatian) the most frequent allele was G*010101. In the Inuit population of Nunavik, the most common allele was HLA-G*010401 (50%) which makes it unique among these populations. [49]

Allele	Japanese (n = 82)	Hutterite (n = 160)	Finnish (n = 194)	Danish (n = 144)	German/ Croatian (n = 344)	Portuguese (n = 117)	Spanish (n = 228)	African American (n = 84)	African Ghanaian (n = 84)	African Shona (n = 216)
G*01011	43	46	58	62	32	37	38	70	83	39.3
G*01012	14	21	38	27	36	31	22	6	2.4	14.4
G*01013	5	2	5	5	7	17	7	2.4	0	0
G*01014	—	4	—	—	—	—	—	—	—	0
G*01015	—	—	—	—	—	—	—	—	—	0
G*01016	—	—	—	—	—	—	—	—	—	0
G*01017	—	—	—	—	—	—	—	—	—	0
G*01018	—	—	—	—	9.1	—	—	—	—	14.4
G*0102	—	—	—	—	—	—	—	—	—	0
G*0103	—	4	—	—	2.3	2	0	—	—	0
G*01041	38	20	—	5	6	13	11	13	9.5	20.4
G*01043	—	—	—	—	—	—	—	—	—	0.4
G*0105N	0	—	—	0.6	2.3	0	3	8.3	4.8	11.1
References	[23]	[13]	[12]	[15]	[16]	[17]	[14]	[23]	[23]	

Abbreviation: — = not determined.

Figure 8: HLA-G allele frequencies (%) in different populations around the world [49]

HPV Period Prevalence

Only one HLA allele or haplotype was associated with a significantly increased or decreased risk of any HPV. The HLA-G*010102 allele was associated with a significantly decreased risk of any HPV and only LR HPV period prevalence (OR=0.64 95%CI=0.42-0.98 and OR=0.37 95% CI=0.16-0.83, respectively). Ferguson *et al.* also found that the HLA-G*010102 allele was associated with a significantly decreased risk of LR HPV infection (OR=0.57, 95% CI=0.37-0.89). [33] The direction of the association between HLA-G*010102 and any HPV changed when the cohort was limited to those women at high risk of acquiring HPV, therefore this may not be a true association.

The HLA-G*010101 allele and its respective homozygous and heterozygous genotypes were associated with an increased risk of LR alpha groups 1 and 3. In fact, in the first sensitivity analysis (in which the cohort was restricted to those women who had three or more visits) the homozygous HLA-G*010101 genotype was significantly associated with an increased risk of any HPV and all three alpha papillomavirus groups. Both the homozygous and heterozygous genotypes were associated with an increased risk of HPV infection and therefore the effect of the allele is most likely dominant. In the paper by Ferguson et al, they found that the HLA-G*010101 allele was associated with a slightly increased risk of alpha papillomavirus groups 1 and 3, but the association was not significant. [38]

The homozygous HLA-G*010401 genotype was associated with a significantly decreased risk of alpha group 3 infection. This association was also significant in the sensitivity analysis where only women at high risk of exposure were included in the analysis. Ferguson *et al.* also found that HLA-G*010401 was significantly associated with an increased risk of alpha group 3 infection. [38] The heterozygous HLA-G*010401 genotype had a null association with HPV infection and therefore it is likely that the effect of this allele is recessive.

In this study, HLA-G*010101 and G*010102 were associated with an increased and a decreased risk of HPV infection, respectively. The HLA-G*010102 allele behaves differently from the wild type (HLA-G*010101), even though the G*010102 mutation

does not change the protein, but the mechanism is unknown. The G*010102 mutation does lie near Glu-63, which interacts with the P2 position of the loaded peptide and could affect the loading of antigens. [33] This could explain why these two alleles have different effects on HPV infection.

In the sensitivity analyses, the associations between the HLA alleles, haplotypes, and genotypes and any HPV either went towards the null or changed direction. As for the alpha group associations, the directions remained in the same but the associations were closer to the null. This could indicate that using alpha groups rather than overall HPV period prevalence as an outcome may be more useful for determining true associations. The period prevalence analysis above suggests that HLA alleles interact differently with different alpha groups. The three alpha papillomavirus groups were used because the alpha species that are included in each group are phylogenetically related and therefore would have similar epitopes to which HLA proteins bind.

HPV Persistence

No alleles, haplotypes or genotypes were significantly associated with a persistent infection. Though three HLA alleles were associated with HPV persistence, when the definition of persistence was limited (to women who had the same HPV type on two or more consecutive visits) only one alleles (HLA-G*010103) retained its association. Three alleles were associated with HR alpha group 2 persistence, but none of them retained their association in the sensitivity analysis. HLA-G*010103 was positively associated with HPV persistence in three of the comparisons, but the associations were not significant (likely because HLA-G*010103 was present in such a small percentage of the population). Ferguson *et al.* also found that HLA-G*010103 was associated with an increased risk of HR alpha group 2 persistence (OR=1.39 95% CI=0.75-2.61). [38]

Measuring HPV persistence in this population is imprecise because the majority of women were already sexually active at enrolment; therefore it is unknown when the women who were HPV positive first became infected.

Though persistence of HPV infection has been linked to cervical cancer, there is no consensus on what defines a persistent infection verses a transient infection. Natural history studies suggest that more than 50% of women infected with HPV will be HPV-negative within a year of first becoming infected and 90% will be HPV-negative within two years. Most longitudinal studies define viral persistence as the repeated detection of a specific HPV type at two consecutive follow-up visits. In these studies the time interval between consecutive follow-up visits is often inconsistent, and therefore the duration of a persistent and transient infection is inconsistent. [22,50]

It is also unclear whether undetected HPV infections have truly been cleared or if there is a period of latency where the viral load is too low to be detected. In this study, if a woman tested negative between two positive visits it was still defined as a persistent infection. It was assumed that on the middle visit that the virus load was too low to be detected. [22,50]

Multiple HPV Infections

HLA-G*010101 (OR=1.54 95% CI=0.90-2.66) was associated with an increased risk of co-infection versus mono-infection, but the association was non-significant. Both the homozygous and heterozygous genotypes of HLA-G*010101 appeared to be associated with an increased risk of multiple infections (the second was significant); therefore the effect of the allele is most likely dominant. Smith, in her Master's thesis, found that HLA-G*010101 was significantly associated with an increased risk of co-infection versus mono-infection (OR=1.66 95% CI=1.02-2.70). [14]

The HLA-G*010102 allele and the homozygous HLA-G*010401 genotype were both associated with a significantly decreased risk of co-infection versus HPV-negative (neither of these allele were included in Smith's analysis). The association between the heterozygous HLA-G*010401 genotype and multiple infections was close to the null. Therefore it is likely that the effect of the allele is recessive.

In this study, the HLA-DRB1*1501 and DQB1*0602 alleles were associated with an increased risk of co-infection versus mono-infection, but the associations were not

significant. Smith also found these two alleles to be associated with an increased risk of co-infection, but the associations were not significant either. [14]

It has been shown that HPV types that are involved in co-infections occur at random, the risk of being co-infected with more than one HPV type was similar to the sum of the risks of being infected with each individual HPV type. Thus, HPV co-infections occur at random and contribute to cervical disease independently. This is likely because all HPV types have a common mode of transmission and have common risk factors. Since female sex behaviours cannot alone explain the increased prevalence, immunological mechanisms may also play a role. As previously stated, the prevalence of multiple infections is high among immunosuppressed HIV-infected women. HPV co-infections may indicate immunological susceptibility to HPV infections. [13,15]

High and Low Grade Squamous Intraepithelial Lesion Period Prevalence

Low-grade squamous intraepithelial lesions (LGSIL) are the result of a productive viral infection and can be caused by both LR and HR HPV types. High-grade squamous intraepithelial lesions (HGSIL) occur when infected immature epithelial cells are prevented from undergoing maturation and differentiation and are caused by HR HPV types. HGSIL is equivalent to CIN 2 and 3. [2,22] Only 45 women in this study had a HGSIL and therefore there was little power to find any association in this analysis. HGSIL is a more important outcome than LGSIL because HGSIL is a precursor lesion whereas LGSIL is indicative of a productive viral infection. Once a HGSIL is found it is treated before it can further develop into invasive cancer and therefore HGSIL is a surrogate measure for cervical cancer. [2,22]

In the majority of previous literature, HLA-DQB1*03, DQB1*0602, and DQB1*1501 were associated with an increased risk of cervical lesions. [12,27] This study found that all three of these alleles were associated with an increased risk of HGSIL, but the associations were not significant. In this study HLA-DRB1*13 was associated with a significantly increased risk of any SIL and LGSIL. It was also associated with an increased risk of HGSIL but the association was not significant. In the majority of the literature this

allele is usually associated with a decreased risk of cervical lesions. [12,27] This difference could be because of a different distribution of HLA-DRB1*13 allele subtypes in the Inuit population. Maciag *et al*., Wu *et al.* and Yang *et al.* found that different DRB1*13 allele subtypes have different directions of association with cervical cancer. [39,51,52] For example, Yang *et al.* did a meta-analysis of the associations between HLA-DRB1 alleles and cervical squamous cell carcinomas (SCC) and found that HLA-DRB1*1301 and DRB1*1302 were associated with a significantly decreased risk of SCC and DRB1*1305 was associated (non-significantly) with an increased risk of SCC (OR=5.51 95% CI=0.61-49.99). [52]

Three HLA-G alleles were found to be associated with a SIL in this study: G*010102, G*010401, and G*0106. The HLA-G*010102 allele (OR=0.54 95% CI=0.17-1.67) and its heterozygous genotype (OR=0.41 95% CI=0.11-1.45) were associated with a decreased risk of HGSIL. Ferguson *et al.* found that the heterozygous HLA-G*010102 genotype was significantly associated with a decreased risk of CIN2 (OR=0.63 95% CI=0.40-0.98). [33] In this study, the homozygous HLA-G*010401 genotype was associated with an increased risk of any SIL, LGSIL and HGSIL, but the associations were not significant. Ferguson *et al.* found that the homozygous HLA-G*010401 genotype was associated with an increased risk of CIN3 (OR=3.1 95% CI=0.14-75.8) and Simões et al. found that the HLA-G*0104/14bp genotype was associated with an increased risk of HGSIL. [33,37] The heterozygous genotype association was close to the null; this again suggests that the effect of the allele is recessive. In the population of Nunavik, HLA-G*0106 was associated with an increased risk of any SIL, LGSIL, and HGSIL. Ferguson *et al.* found that the homozygous genotype of this allele was associated with an increased risk of both CIN2 and CIN3, but neither of the associations were significant. [33] Cancer cells display tumor-associated antigens that are presented by HLA class I proteins, HLA-G proteins in particular play an important role in the elimination of tumor cells. A common tumor-escape mechanism involves the downregulation of HLA-G proteins. It is thought that certain HLA-G polymorphisms influence HLA-G expression, and therefore would be associated with cervical cancer and its precursors. [53]

HPV Incidence

No HLA alleles, haplotypes or genotypes were associated with significantly increased or decreased rate of any HPV infection. The HLA-G*010101 allele and its homozygous and heterozygous genotypes were associated with an increased rate of both LR alpha group 1 and 3. Ferguson *et al.* found that the HLA-G*010101 allele was associated with an increased risk of alpha group 1 (OR=1.43 95% CI=0.80-2.57) and 3 (OR=1.42 95% CI=0.82-3.46), but the associations were not significant. [38] HLA-DRB1*13 was associated with a significantly increased rate of LR alpha group 1 incidence. This allele has been associated with an increased risk of LR HPV infection in studies by both Maciag *et al.* (OR=2.2 95% CI=1.1-4.3) and Mahmud *et al.* (OR=1.5 95% CI=0.7-3.3). [28,39] The homozygous HLA-G*010401 genotype was associated with a significantly decreased rate of LR alpha group 3. As previously mentioned, Ferguson *et al.* found that HLA-G*010401 was associated with a decreased risk of alpha group 3. [38]

Adjustment

In this study, all odds ratios and rate ratios were adjusted for age at baseline (entry into the cohort). In order to avoid over-adjustment it was decided to adjust for just one variable. Age was chosen because it is the most consistent determinate of HPV infection and because there was no missing values for age. [1,2] Previous studies which have investigated the association between HLA polymorphisms and HPV infection have adjusted for ethnicity and age or age at first intercourse. The majority of studies adjusted for age at study entry rather than age at first intercourse. [34,38,39]

Ethnicity is considered to be a possible confounder because it is associated with both an individual's genetics and lifestyle behaviour factors. In this study, there was no need to adjust for ethnicity because study participants all self-identified as Inuit. Other than ethnicity, there are no other known confounders of this association because no other factors affect both an individual's genetics as well as their likelihood to be infected with HPV.

Limitations

While reviewing the results of this study it is important to also consider its limitations. In all of the analyses the 95% confidence intervals were quite wide because the study population was fairly small (and many HLA alleles were found in only a small percentage of the population), therefore this study had little power to find significant associations. In the case of rare alleles or a rare outcome, the study may lack sufficient power to detect an association. For any HPV period prevalence, there was less than 50% power to detect an association (at the 0.05 level) for any of the alleles present in less than 5% of the population. There was 55-70% power to detect an association for HLA-E*0101, G*010101, and G*010102 and >90% power to detect an association for HLA-E*0103, G*010401, and DQB1*03.

As mentioned above, the frequency of HLA polymorphisms in a population depends greatly on the ethnicity of the population. This population was ethnically homogenous, but there may be some regional differences in HLA frequency within Nunavik (for example between Hudson and Ungava bay communities) but our study sample was not able to address this specifically.

This was an exploratory study of a small population and therefore the data was not adjusted for multiple testing. Since the results were not adjusted for multiple testing, some of the associations found could be due to chance. All associations must therefore be looked at with caution and either compared with previous studies or repeated in future studies to be verified. Mahmud *et al.* completed a small analysis to determine if any associations were likely to be occurred by chance, and found that most of the 72 observed associations in their study were unlikely to have occurred by chance.[28]

The recruitment was non-random (an analysis of selection and non-participation bias can be found below) and some dates were missing from the incidence analysis (see methods for how this was dealt with).

Selection and Non-participation Bias

Women who presented for regularly scheduled Pap tests at clinics in one of the four participating communities were asked to participate in the study. Nurse practitioners did not collect any information from the women who declined to participate in the study, so no evaluation of the difference between participants and non-participants was possible. The nurse practitioners estimated that the participation rate was over 80%, but the total number of women approached was not recorded.

A partial evaluation of selection bias was performed by comparing sociodemographic and lifestyle characteristics of the study participants to data on the general population of Nunavik collected by Statistics Canada and Santé-Québec.

Statistics Canada reported in their 2006 Census Aboriginal Profiles on Nunavik that only 12% of Inuit females over the age of 15 years old had a high school certificate (or its equivalent). In women aged 25-34, 20% of the census population and 19.7% of the study population had at least a high school education. In women aged 35-64, 6.6% of the census population and 9.2% of the study population had at least a high school education. [46] The 2006 census reported that 64.6% of Inuit females in Nunavik were single, 27.1% were married, 2.4% were separated, 1.6% were divorced and 4.3% were widowed. [46] In the study population, 42.2% of women were single, 27.5% were living with a partner, 26.9% were married, 1.9% were divorced or separated, and 1.6% were widowed. The distribution of education and marital status found in the study population mirrors that found in the 2006 Canadian census. The study population did differ from the census on employment: 70% of the study population was employed whereas only 53.2% of the census population was employed. [46]

According to a 2004 Nunavik Inuit Health Survey (completed by the Institut National de Santé Publique du Québec and the Nunavik Regional Board of Health and Social Services), three-fourths of the Nunavik population were smokers and were occasional or regular drinkers (men and women). [54] In this study, 73% of the women were smokers and 66.6% reported alcohol consumption. These two measures were similar to what was reported in the Nunavik Inuit Health Survey. The section of the

survey on women's health and sexual health reported that 30.3% of the women had no sexual partners in the past year, 51.2% had one sexual partner and 18.5% had two or more. [54] This study found that 10.7% of the women had no sexual partners in the last year, 61.6% had one or two sexual partners and 27% had three or more. The Nunavik Inuit Health Survey also reported that 28.0% of women reported the use of birth control in the past year. [54] In this study, 36.2% of women reported using birth control.aore sexually active than the general population of Nunavik and that the proportion of women in the study using birth control is also greater than the proportion in the general population.

Using information from the Canadian Census and Nunavik Inuit Health Survey it appears that the study population is representative of the general population of Nunavik in terms of marital status, education, smoking and alcohol consumption. The study population did differ from the general population in that more women in the study were employed, more of them reported using birth control and they were more sexually active, which suggests that the study population is at greater risk of acquiring HPV.

The differences outlined above between the study population and the target population should not have affected the distribution of alleles, haplotypes, and genotypes. It is therefore reasonable to assume that the distribution of HLA alleles, haplotypes, and genotypes in the study population is similar to that of the target population, the Inuit population of Nunavik.

Strengths

Even though this study does have some limitations, it also has many strengths. This study captured 41% of the target population and the age distribution of the women in the study was similar to that of the target population.

To the authors best knowledge this is the first time that HLA polymorphisms have been investigated in this unique and isolated population. There have only been a few published analyses reporting the association between HLA polymorphisms and HPV

period prevalence, persistence, and multiple infections and no published analyses (to the best of the author's knowledge) have reported the association between HLA polymorphisms and HPV incidence.

This study attempted to report the distribution of HLA alleles, haplotypes, and genotypes in Nunavik and to determine their influence on HPV infection. Through interpretation of the results, this study may contribute to a better understanding of the relationship between HPV infections and the host immune response.

Conclusions

This is the first analysis of the distribution of HLA alleles and haplotypes in the Inuit population of Nunavik, Quebec. The association between these alleles, haplotypes and genotypes and HPV incidence, period prevalence, persistence, multiple infections and SILs was investigated. This data, along with other previous studies, will hopefully provide insight into the role of the host immune system, in particular HLA polymorphisms, in HPV infections. This study, along with the previous analysis performed on this population, will further our understanding of HPV infection the Nunavik population.

The distribution of HLA alleles, haplotypes, and genotypes in Nunavik, Quebec is different from that found in other Canadian populations and other populations around the world. Two alleles in particular, HLA-G*010401 and HLA-DQB1*03 (found in 76% and 94% of individuals, respectively) were found more frequently in the Nunavik population than in other populations which have been studied.

The results of this study show an association between the HLA-G*010101 allele and an increased risk of LR alpha group 1 and 3 period prevalence and incidence. Both the homozygous and heterozygous HLA-G*010101 genotypes were also associated with an increased risk of HPV infections, which suggests that this allele's effect is dominant. The HLA-G*010102 allele was associated with a significantly decreased risk of any HPV infection and LR HPV infections. The homozygous HLA-G*010401 genotype was associated with a significantly decreased risk of LR alpha group 3 infection period prevalence and incidence. The association between the heterozygous HLA-G*010401 genotype and HPV infection was close to the null, which suggests that this allele's effect is recessive. These results are consistent with the results of Ferguson *et al.* and Smith.[14,33,38] In the period prevalence sensitivity analyses, the associations between HLA and overall HPV infection changed according to the different sensitivity analysis whereas the alpha papillomavirus group associations remained the same. This suggests that the alpha groups are a more indicative outcome than any HPV type and that HLA alleles

interact differently with different alpha groups. It could be that certain HLA polymorphism recognize (or fail to recognize) epitopes on similar HPV species.

No HLA alleles, haplotypes or genotypes were found to be significantly associated with HPV persistence (duration of infection). This suggests that HLA polymorphisms may play a greater role in the acquisition of an HPV infection rather than virus clearance. The HLA complex likely acts early in the etiological pathway at stage of host immune recognition rather than the stage of clearance of a viral infection. [28]

The HLA-G*010101 allele and the homozygous HLA-G*010401 genotype were found to be associated with an increased and decreased risk of concurrent multiple infections, respectively. It has been shown that co-infections occur at random and co-infections may indicate immunological susceptibility to HPV infections.

The HLA class II allele HLA-DRB1*13 was associated with a significantly increased rate of LR alpha groups 1 infections, any SIL and LGSIL (a LGSIL is indicative of a productive viral infection). HLA-DRB1*13 is usually associated with a decreased risk of cervical cancer, but few studies have investigated the association between HLA-DRB1*13 and LR HPV infections.

This study's data supports the theory that HLA polymorphisms play a role in the natural history of HPV infections. Regardless of the unique distribution of HLA alleles, haplotypes, and genotypes in Nunavik, this and other studies suggest that HLA polymorphisms act at the stage of viral acquisition and that HLA polymorphisms interact differently with different alpha papillomavirus groups. [28,38] Knowledge about an individual's (or family's) genetic susceptibility to HPV infection or cervical cancer may allow for elaboration of cervical cancer screening and HPV vaccination.

References

1. Trottier H, Franco EL. The epidemiology of genital human papillomavirus infection. Vaccine 2006;24 Suppl 1:S1-15.
2. Baseman JG, Koutsky LA. The epidemiology of human papillomavirus infections. J Clin Virol 2005;32 Suppl 1:S16-24.
3. Bouvard V, Baan R, Straif K, et al. A review of human carcinogens--Part B: biological agents. Lancet Oncol 2009;10:321-2.
4. de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. Virology 2004;324:17-27.
5. Wheeler CM. Natural history of human papillomavirus infections, cytologic and histologic abnormalities, and cancer. Obstet Gynecol Clin North Am 2008;35:519-36; vii.
6. Munoz N, Bosch FX, de Sanjose S, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med 2003;348:518-27.
7. Scheurer ME, Tortolero-Luna G, Adler-Storthz K. Human papillomavirus infection: biology, epidemiology, and prevention. Int J Gynecol Cancer 2005;15:727-46.
8. Bosch FX, de Sanjose S. The epidemiology of human papillomavirus infection and cervical cancer. Dis Markers 2007;23:213-27.
9. Ho L, Chan SY, Burk RD, et al. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. J Virol 1993;67:6413-23.
10. Stanley M. Immune responses to human papillomavirus. Vaccine 2006;24 Suppl 1:S16-22.
11. Incidence and prevalence of HPV in Canada. (Accessed at http://www.hpvinfo.ca/health-care-professionals/what-is-hpv/incidence-and-prevalence-of-hpv-in-canada/.)
12. Hildesheim A, Wang SS. Host and viral genetics and risk of cervical cancer: a review. Virus Res 2002;89:229-40.
13. Rousseau MC, Pereira JS, Prado JC, Villa LL, Rohan TE, Franco EL. Cervical coinfection with human papillomavirus (HPV) types as a predictor of acquisition and persistence of HPV infection. J Infect Dis 2001;184:1508-17.
14. Smith MA. Occurrence, Determinants and Dynamics of HPV Coinfections in a Cohort of Montreal University Students: Queen's University; 2011
15. Chaturvedi AK, Katki HA, Hildesheim A, et al. Human papillomavirus infection with multiple types: pattern of coinfection and risk of cervical disease. J Infect Dis 2011;203:910-20.
16. Cervantes JL. Multiple human papillomavirus infection: don't forget the genetic background. J Infect Dis 2011;204:1816; author reply -7.
17. Levi JE, Kleter B, Quint WG, et al. High prevalence of human papillomavirus (HPV) infections and high frequency of multiple HPV genotypes in human immunodeficiency virus-infected women in Brazil. J Clin Microbiol 2002;40:3341-5.
18. Trottier H, Mahmud S, Prado JC, et al. Type-specific duration of human papillomavirus infection: implications for human papillomavirus screening and vaccination. J Infect Dis 2008;197:1436-47.

19. Moscicki AB, Ma Y, Jonte J, et al. The role of sexual behavior and human papillomavirus persistence in predicting repeated infections with new human papillomavirus types. Cancer Epidemiol Biomarkers Prev 2010;19:2055-65.
20. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999;189:12-9.
21. Statistics CCSsSCoC. Canadian Cancer Statistics 2011. Toronto: Canadian Cancer Society; 2011.
22. Woodman CB, Collins SI, Young LS. The natural history of cervical HPV infection: unresolved issues. Nat Rev Cancer 2007;7:11-22.
23. Franco EL, Villa LL, Sobrinho JP, et al. Epidemiology of acquisition and clearance of cervical human papillomavirus infection in women from a high-risk area for cervical cancer. J Infect Dis 1999;180:1415-23.
24. Talking Glossary of Genetic Terms. (Accessed at http://www.genome.gov/glossary/index.cfm.)
25. HLA nomenclature. 2012. (Accessed at http://hla.alleles.org/announcement.html.)
26. Chattopadhyay K. A comprehensive review on host genetic susceptibility to human papillomavirus infection and progression to cervical cancer. Indian J Hum Genet 2011;17:132-44.
27. de Araujo Souza PS, Sichero L, Maciag PC. HPV variants and HLA polymorphisms: the role of variability on the risk of cervical cancer. Future Oncol 2009;5:359-70.
28. Mahmud SM, Robinson K, Richardson H, et al. HLA polymorphisms and cervical human Papillomavirus infection in a cohort of Montreal University students. J Infect Dis 2007;196:82-90.
29. Fennema JS, van Ameijden EJ, Coutinho RA, van den Hoek AA. HIV, sexually transmitted diseases and gynaecologic disorders in women: increased risk for genital herpes and warts among HIV-infected prostitutes in Amsterdam. Aids 1995;9:1071-8.
30. Fruchter RG, Maiman M, Sedlis A, Bartley L, Camilien L, Arrastia CD. Multiple recurrences of cervical intraepithelial neoplasia in women with the human immunodeficiency virus. Obstet Gynecol 1996;87:338-44.
31. Zhang SY, Gu HX, Li D, et al. Association of human leukocyte antigen polymorphism with hepatitis B virus infection and genotypes. Jpn J Infect Dis 2006;59:353-7.
32. Cangussu LO, Teixeira R, Campos EF, et al. HLA class II alleles and chronic hepatitis C virus infection. Scand J Immunol 2011;74:282-7.
33. Ferguson R, Ramanakumar AV, Koushik A, Coutlee F, Franco E, Roger M. Human leukocyte antigen G polymorphism is associated with an increased risk of invasive cancer of the uterine cervix. Int J Cancer 2011.
34. Ades S, Koushik A, Duarte-Franco E, et al. Selected class I and class II HLA alleles and haplotypes and risk of high-grade cervical intraepithelial neoplasia. Int J Cancer 2008;122:2820-6.
35. de Araujo Souza PS, Villa LL. Genetic susceptibility to infection with human papillomavirus and development of cervical cancer in women in Brazil. Mutat Res 2003;544:375-83.

36. Qiu X, Zhang F, Chen D, et al. HLA-B*07 is a High Risk Allele for Familial Cervical Cancer. Asian Pac J Cancer Prev 2011;12:2597-600.
37. Simoes RT, Goncalves MA, Castelli EC, et al. HLA-G polymorphisms in women with squamous intraepithelial lesions harboring human papillomavirus. Mod Pathol 2009;22:1075-82.
38. Ferguson R, Ramanakumar AV, Richardson H, et al. Human leukocyte antigen (HLA)-E and HLA-G polymorphisms in human papillomavirus infection susceptibility and persistence. Hum Immunol 2011;72:337-41.
39. Maciag PC, Schlecht NF, Souza PS, Rohan TE, Franco EL, Villa LL. Polymorphisms of the human leukocyte antigen DRB1 and DQB1 genes and the natural history of human papillomavirus infection. J Infect Dis 2002;186:164-72.
40. Cerigo H, Macdonald ME, Franco EL, Brassard P. Awareness and knowledge about human papillomavirus among Inuit women in Nunavik, Quebec. J Community Health 2011;36:56-62.
41. Hamlin-Douglas LK, Coutlee F, Roger M, Hanley J, Franco EL, Brassard P. Determinants of human papillomavirus infection among Inuit women of northern Quebec, Canada. Sex Transm Dis 2010;37:377-81.
42. Image Gallery of Nunavik. (Accessed at http://www.parcoursnunavik.com/carte_nunavik.asp?langue=2.)
43. Hamlin-Douglas LK, Coutlee F, Roger M, Franco EL, Brassard P. Prevalence and age distribution of human papillomavirus infection in a population of Inuit women in Nunavik, Quebec. Cancer Epidemiol Biomarkers Prev 2008;17:3141-9.
44. Crum CP. Genital papillomaviruses and related neoplasms: causation, diagnosis and classification (Bethesda). Mod Pathol 1994;7:138-45.
45. Bunce M, O'Neill CM, Barnardo MC, et al. Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP). Tissue Antigens 1995;46:355-67.
46. Aboriginal Population Profile Region du Nunavik. 2006. (Accessed at http://www12.statcan.ca/census-recensement/2006/dp-pd/prof/92-594/details/page.cfm?Lang=E&Geo1=HR&Code1=2417&Geo2=PR&Code2=24&Data=Count&SearchText=Region%20du%20Nunavik&SearchType=Begins&SearchPR=01&B1=All&GeoLevel=PR&GeoCode=2417.)
47. Welinder L, Graugaard B, Madsen M. HLA antigen and gene frequencies in Eskimos of East Greenland. Eur J Immunogenet 2000;27:93-7.
48. Leffell MS, Fallin MD, Erlich HA, et al. HLA antigens, alleles and haplotypes among the Yup'ik Alaska natives: report of the ASHI Minority Workshops, Part II. Hum Immunol 2002;63:614-25.
49. Matte C, Lacaille J, Zijenah L, Ward B, Roger M. HLA-G and HLA-E polymorphisms in an indigenous African population. The ZVITAMBO Study Group. Hum Immunol 2000;61:1150-6.
50. Gravitt PE. The known unknowns of HPV natural history. J Clin Invest 2011;121:4593-9.
51. Wu Y, Liu B, Lin W, et al. Human leukocyte antigen class II alleles and risk of cervical cancer in China. Hum Immunol 2007;68:192-200.

52. Yang YC, Chang TY, Lee YJ, et al. HLA-DRB1 alleles and cervical squamous cell carcinoma: experimental study and meta-analysis. Hum Immunol 2006;67:331-40.
53. Donadi EA, Castelli EC, Arnaiz-Villena A, Roger M, Rey D, Moreau P. Implications of the polymorphism of HLA-G on its function, regulation, evolution and disease association. Cell Mol Life Sci 2011;68:369-95.
54. Rochette MAaL. Nunavik Inuit Health Survey 2004. Qanuippitaa? How are we?: Institut national de santé publique du Québec (INSPQ) & Nunavik Regional Board of Health and Social Services (NRBHSS); 2008.

Appendix 1: Consent Form

McGill University Health Center
Division of Clinical Epidemiology of the Royal Victoria Hospital

Study on Human papillomavirus and cervical cancer among Inuit Women in Northern Quebec

Information Document

Researchers: Dr Paul Brassard, Division of Clinical Epidemiology, Royal Victoria Hospital, **Dr Eduardo L. Franco**, Department of Oncology, McGill University, **Dr François Coutlée** and **Dr Michel Roger**, Department of Microbiology, Notre-Dame Hospital, University of Montreal Hospital Center.

A) Purpose of this study

Human papillomavirus (HPV) is a virus that causes genital warts and is normally sexually transmitted. HPV infection is detected by collecting samples of cells from the cervix that can be obtained during a Pap test. The sample is then examined to determine the presence of HPV. If HPV is detected, further analysis is conducted to classify the type of HPV. Inuit Women seem to be more at risk for HPV infection. When HPV infection is present for a long time, it can cause in some women cervical cancer, a disease that can be prevented by early detection with the Pap test and treatment. We are doing this study to investigate how many women will acquire and stay with the infection over time. We also want to understand if there are human factors that predispose women to develop cervical disease. We call these host susceptibility factors (including HLA) for cancer and infection. The information found can then be used by Nunavik public health officials for developing effective cancer-screening program and prevention program for women in the area.

B) Procedure

If you agree to participate, you will be asked to complete a self-administered questionnaire after your Pap test at the Hygiene clinic. A nurse will accompany you in the process. If you feel uncomfortable to complete the questionnaire at once, you can always make arrangement with the nurse for another visit. Specimens of your Pap test will be collected for the study. It will be sent to the laboratory at Notre-Dame Hospital in Montreal to be tested for HPV infection. The Pap specimens will be taken and sent to Montreal each time you come back at the Hygiene clinic for your regular Pap test or for any other health condition that requires a cervical exam, over a 5 year period (60 months). Those specimens (including the host susceptibility factors) will be kept for the length of this study and the length of other studies that could come from this global project for a total of 10 years. We will also need to review your medical file to collect further information concerning your health status.

C) Risks and Benefits

There is no additional risk related to this study as the Pap smear is a safe examination. You will not have more visits to the clinic; you will only spend a little more time to fill out the questionnaire. Your participation will help in developing prevention for HPV infection. As well, if any lesions are detected, you may benefit directly by getting the proper treatment.

Information Document

D) Participation

Your participation in this study is **of your own free will**. You can stop participating in any part of the project at any time and this will not affect your health care or treatment in any way. You will also get a copy of this consent form. You have the right to ask any question you want to the nurse about the study before accepting to participate.

E) Confidentiality

In order to ensure your privacy and confidentiality, your name will not appear on any study record or results presented by the research team. Instead a patient identification number will be assigned to you and will appear in all your records. Only the nurse and the researchers in Montreal will have access to the study number. You understand that all information about you and you Pap smear results and host susceptibility factors will be treated in the same confidential manner as other medical records and you will not be identified in any subsequent reporting of results.

F) Questions

If you have any specific questions, now or at any time about this study, please do not hesitate to contact the Director of Professional services of the Ungava Tulattavik Health center, Dr Nathalie Boulanger at (819) 964 2905 *or* make a collect call to the chief investigator, Dr Paul Brassard at (514) 842 1231 ext 36910.

McGill University Health Care Center
Division of Clinical Epidemiology of the Royal Victoria Hospital

Study on human papillomavirus and cervical cancer among Inuit Women in Northern Quebec

Voluntary consent

By signing this form, I acknowledge having received and read a copy of the information paper concerning this study. I have had the opportunity to ask any questions I may have about this study, and they have been answered to my satisfaction. I agree to participate in this study and I understand that I may withdraw this agreement at any time. I understand that my decision whether or not to participate will not change any health care I might receive or my legal rights. I also understand that all information will be kept strictly confidential. My file will be coded and kept in place where only the research team will have access.

1) I agree to complete the questionnaire on risk factors for HPV infection

Yes: ___; No: ___

2) I agree for my Pap specimens collected in the next 5 years be sent to Notre-Dame Hospital in Montreal for detection of HPV infection and host susceptibility factors for cancer and infection including HLA. Those specimens will be kept for the length of this study and the length of other studies that could come from this global project for a total of 10 years.

Yes: ___; No: ___

3) I agree that my medical chart can be reviewed up to 5 years following my agreement to participate.

Yes: ___; No: ___

Signature: _____

Write your name in block letters: _____

Date: _____

Telephone number: _____

Nurse section

I recognize having offered to the participant a copy of this consent form and a copy of the information document.

Participant ID number: IN- _____
Signature of the nurse: _____
Name of the nurse in block letters: _____
Date: _____

Appendix 2: Questionnaire

HPV-INUIT

Research Coordinator Section

Date	Chart number	Date of birth
dd/mm/yyyy		dd/mm/yyyy

A. Participant Identification

2- What is your current marital status ?

Single (Not married and not living with partner) Married Divorced/Separated Widowed Living with partner

B. Socio economic status

3- Are you employed ? Yes No

4- What is your highest level of schooling ? Less than grade 9 Grade 9 to 13 More than grade 13

5- What is the current employement status of your husband or living partner ? Employed Unemployed

If employed, what type of work does he do ?

C. Life habits

6- Are you a current smoker ? Yes No

If yes, a) how many cigarettes do you smoke a day ? :

b) how long (in years) have you been smoking ? :

If no, a) have you ever smokec ? : Yes No

If yes, how long (in years) has it been since you stopped smoking ? :

7- Do you drink alcohol ? Yes No

If yes, a) how long (in years) have you been drinking ? :

b) how often do you drink :

Beer :	Never	Occasionally	Once a week	More than once a week	Every day
Wine :	Never	Occasionally	Once a week	More than once a week	Every day
Wisky/Gin/Vodka or any hard liquor	Never	Occasionally	Once a week	More than once a week	Every day

8- Are you currently using any birth control method ? Yes No

If yes, a) what kind of method do you currently use ? (you can put down more than one)

For how long in years	Birth control methods
	Birth control pills
	Latex safe (condom)
	Spermicides (gel)
	I.U.D (coil)
	Diaphragm
	Depo-Provera (injections)
	Rythm, calendar, natural method
	Other :

D. Sexual behavior

9- Have you ever had sex ? Yes No **If no,** go to question 17.

10- How old were you when you first had sex ? :

11- Throughout your life, what is the number of partners with whom you have had sex ? (approximately)

0-4 5-9 10 and more

12- How many sexual partners have you had in the last year ?

13- How many sexual partners have you had in the last month ?

14- Does your partner(s) have other sexual partner(s) currently ? Yes No Unknown

If yes, how many partners does he currently have (approximately) ?

E. Gynecological and obstetric events

15- Are you pregnart ? Yes No

16- Up to now :

How many times did you deliver a living baby ?

How many times did you have an abortion ?

How many times did you have a miscariage ?

17- Have you ever had a gynecological exam in the past (excluding the current one) ? Yes No Unknown

Year

If yes, what year did you have the first one (approximately) :

what year did you have the last one (approximately) :

18- Have you ever experienced sexually transmitted disease (STD's or infection with herpes, chlamydia, gonorrhea, syphilis) in the past ? Yes No Unknown

If yes, Once 2-4 times 5 times or more

F. Health conditions

19- Are you experiencing one or more of the following health problems ?

HIV infection

Had an organ transplant

Use of cortisone (injection or pills) for more than 1 month

Other health problems:

G. Comments (please, write down any comment you want about a specific item or about the study in general) :

Appendix 3: Medical Chart Review Form

CHARACTERIZATION OF THE HUMAN PAPILLOMAVIRUS INFECTION AMONG A POPULATION OF INUIT WOMEN IN QUEBEC

RÉVISION DE DOSSIER MÉDICAL

No d'identification de l'étude : ☐

No de dossier médical : ☐ **Date de naissance :** ☐

Historique médicale de la patiente :

Maladies chroniques : ☐

Chirurgies : ☐

Immunosupression : ☐

Parité : G ☐ P ☐ A ☐

A. Identification du participant

1- Quel est votre langue principale ? ☐ Inuktitut ☐ Anglais ☐ Francais ☐ Autre _____

2- Quel est votre état civil ? ☐ Célibataire ☐ Mariée ☐ Divorcée/Séparée ☐ Veuve ☐ Habitant avec un conjoint

B. Statut socio-économique

3- Avez-vous un emploi ? ☐ Oui ☐ Non

4- Quel est le plus haut degré de scolarisation que vous avez obtenu ? ☐ Moins que la 9e année ☐ Entre la 9e et la 13e année ☐ Plus que la 13e année

5- Quel est le statut de votre mari ou de votre conjoint ? ☐ Employé ☐ Sans emploi

Si employé, quel type de travail fait-il ? _____

C. Habitudes de vie

6- Fumez-vous présentem ent ? ☐ Oui ☐ Non

Si oui, a) Combien de cigarettes fumez-vous chaque jour ? : _____

b) Depuis combien de temps fumez-vous ? : _____

Si non, a) Avez-vous déjà fumé ? : ☐ Oui ☐ Non

Si oui, depuis combien de temps avez-vous arrêté de fumer ? : _____

7- Est-ce que vous buvez de l'alcool ? ☐ Oui ☐ Non

Si oui, a) Depuis combien de temps bu vez-vous ? : _____

b) À quel fréquence buvez-vous de l'alcool :

Bière : ☐ Jamais ☐ Occasionellement ☐ Une fois par semaine ☐ Plus d'une fois par semaine ☐ Tous les jours

Vin : ☐ Jamais ☐ Occasionellement ☐ Une fois par semaine ☐ Plus d'une fois par semaine ☐ Tous les jours

Wisky/Gin/Vodka ou tout autre liqueur forte : ☐ Jamais ☐ Occasionellement ☐ Une fois par semaine ☐ Plus d'une fois par semaine ☐ Tous les jours

8- Est-ce que vous prenez de la drogue ? ☐ Oui ☐ Non

Si oui, a) Depuis combien de temps vous droguez-vous? _____

b) Quelle sorte de drogue prenez -vous ? _____

LISTE DES RÉSULTATS DE CYTOLOGIE OBTENUS

No.Étude:_____

Date	Normal	Inflammation	Atypie	LGSIL	HGSIL	Inadéquat	Autre

LISTE DES MTS No.Étude:_____

Date	Source		MTS							
	Résultat de lab.	Note du médecin	Gonorrhée	Chlamydia	Syphylis	Lymphogranulome vénérien	Herpes	Condylome	Chancre mou	Autre

LISTE DES MÉTHODES CONTRACEPTIVES UTILISÉES

No.Étude:_____

	Date de début	Date de fin
Contraceptif oraux		
Condom		
Stérilet		
Diaphragme		
Injection Dépo-Provera		
Mousse-gel contraceptif		
Symptothermique		
Autre (spécifiez):		

Appendix 4: Tables

Frequency of HLA alleles in Nunavik, Montreal and Alaska

	Inuit Population of Nunavik	McGill-Concordia Cohort study* [28]	Biomarkers of Cervical Cancer Risk (BCCR) study* [34]	Yup'ik Alaskan Natives** [48]
HLA Allele:				
B*07	4.47%	14.7%	17.7%	1.6%
DQB1*03	94.19%	63.7%	57.1%	61.7%
DQB1*0602	3.63%	21%	19.7%	0.8%
DRB1*13	4.70%	16.6%	25.3%	0.4%
DRB1*1501	3.80%	25.2%	21.3%	0.8%
HLA Haplotype:				
B*07-DQB1*03	3.64%	-	-	-
B*07-DRB1*13	0.20%	-	-	-
B*07-DRB1*1501	1.59%	-	9.2%	-
B*07-DRB1*0602	1.67%	-	8.5%	-
DQB1*03-DQB1*0602	2.57%	-	-	-
DQB1*03-DRB1*13	2.94%	-	-	-
DQB1*03-DRB1*1501	2.76%	13.5%	-	-
DQB1*0602-DRB1*13	0.19%	-	-	-
DRB1*1501-DQB1*0602	3.72%	21%	19.7%	0.8%
DRB1*1501-DRB1*13	0.19%	-	-	-
B*07-DRB1*1501-DQB1*0602	1.64%	-	8.5%	-

*Montreal, Canada

**Difference in allele nomenclature so not exactly comparable

HLA Allele:	Inuit Population of Nunavik	BCCR study* [33]
E*0101	32.85%	-
E*0103	66.38%	-
G*010101	31.42%	45%
G*010102	15.23%	24%
G*010103	1.05%	7%
G*010108	0.38%	0.5%
G*0103	0.38%	-
G*010401	50.38%	9%
G*0106	1.34%	6%
G*0107	0.10%	-

*Montreal, Canada

LR and HR HPV Period Prevalence

Allele / Haplotype / Genotypes	LR-HPV only Age-Adjusted OR (95% CI)	HR-HPV only Age-Adjusted OR (95% CI)	HPV type 16/18 Age-Adjusted OR (95% CI)
Cohort (n=548)			
HLA-B*07	n/a	0.90 (0.29-2.85)	1.11 (0.26-4.68)
HLA-E*0101	0.79 (0.43-1.44)	1.18 (0.72-1.93)	1.03 (0.53-2.02)
HLA-E*0103	1.06 (0.39-2.93)	0.89 (0.41-1.92)	1.20 (0.39-3.68)
HLA-G*010101	0.93 (0.51-1.70)	0.74 (0.45-1.20)	0.86 (0.45-1.65)
HLA-G*010102	**0.37 (0.16-0.83)**	1.05 (0.63-1.75)	0.70 (0.35-1.40)
HLA-G*010103	1.28 (0.14-11.85)	1.96 (0.39-9.81)	2.13 (0.21-22.03)
HLA-G*010401	1.45 (0.68-3.09)	1.28 (0.72-2.26)	1.61 (0.73-3.51)
HLA-G*0106	1.44 (0.28-7.53)	0.30 (0.03-2.78)	1.18 (0.23-6.12)
HLA-DQB1*03	2.07 (0.46-9.41)	2.03 (0.68-6.05)	1.56 (0.44-5.57)
HLA-DQB1*0602	0.34 (0.04-2.69)	0.29 (0.06-1.39)	0.22 (0.03-1.87)
HLA-DRB1*13	0.60 (0.13-2.76)	0.27 (0.07-1.02)	0.46 (0.11-1.87)
HLA-DRB1*1501	0.34 (0.04-2.66)	0.29 (0.06-1.41)	0.22 (0.02-1.86)
DRB1*1501-DQB1*0602	0.34 (0.04-2.67)	0.29 (0.06-1.39)	0.22 (0.03-1.89)
B*07-DRB1*1501-DQB1*0602	n/a	0.54 (0.06-4.90)	n/a
G010101:			
Homozygous	1.23 (0.53-2.85)	1.27 (0.60-2.70)	0.82 (0.25-2.71)
Heterozygous	0.95 (0.56-1.61)	0.80 (0.50-1.29)	0.84 (0.43-1.64)
Absent*	1.00	1.00	1.00
G010102:			
Homozygous	0.59 (0.11-3.14)	0.58 (0.11-3.10)	0.70 (0.08-6.61)
Heterozygous	**0.35 (0.18-0.70)**	1.00 (0.62-1.63)	0.81 (0.40-1.65)
Absent*	1.00	1.00	1.00
G010401:			
Homozygous	1.72 (0.86-3.45)	1.26 (0.65-2.43)	1.70 (0.65-4.45)
Heterozygous	1.06 (0.57-2.01)	1.25 (0.73-2.14)	1.41 (0.63-3.15)
Absent*	1.00	1.00	1.00

*Reference group

HPV Period Prevalence Sensitivity Analysis 1*

Allele / Haplotype / Genotypes	Any HPV type	α Group 1 (α-1, α-8, α-10, α-13)	α Group 2 (α-5, α-6, α-7, α-9, α-11)	α Group 3 (α-3, α-15)
	Age-Adjusted OR (95% CI)			
HLA-B*07	0.70 (0.21-2.34)	n/a	0.84 (0.25-2.87)	0.40 (0.05-3.41)
HLA-E*0101	1.07 (0.63-1.82)	0.84 (0.35-2.03)	1.04 (0.59-1.81)	1.04 (0.54-2.01)
HLA-E*0103	0.88 (0.37-2.07)	2.40 (0.29-19.57)	0.87 (0.35-2.16)	0.82 (0.28-2.36)
HLA-G*010101	1.20 (0.71-2.02)	2.23 (0.86-5.78)	1.27 (0.73-2.20)	1.82 (0.93-3.58)
HLA-G*010102	0.75 (0.42-1.35)	0.32 (0.09-1.14)	0.82 (0.44-1.50)	0.66 (0.32-1.39)
HLA-G*010103	0.33 (0.03-3.38)	n/a	0.41 (0.04-4.22)	0.73 (0.07-7.66)
HLA-G*010401	0.66 (0.37-1.19)	0.76 (0.28-2.02)	0.58 (0.31-1.07)	0.55 (0.27-1.11)
HLA-G*0106	0.89 (0.16-4.86)	n/a	0.64 (0.09-4.37)	1.25 (0.17-9.03)
HLA-DQB1*03	0.81 (0.27-2.41)	1.47 (0.17-12.85)	0.91 (0.29-2.90)	0.45 (0.14-1.47)
HLA-DQB1*0602	0.54 (0.16-1.76)	n/a	0.35 (0.09-1.44)	0.94 (0.25-3.51)
HLA-DRB1*13	1.49 (0.42-5.25)	n/a	1.75 (0.49-6.22)	1.96 (0.47-8.14)
HLA-DRB1*1501	0.64 (0.21-1.98)	0.49 (0.06-4.19)	0.47 (0.13-1.71)	1.10 (0.32-3.80)
DRB1*1501-DQB1*0602	0.56 (0.17-1.82)	n/a	0.37 (0.09-1.51)	0.94 (0.25-3.49)
B*07-DRB1*1501-DQB1*0602	0.61 (0.11-3.58)	n/a	0.76 (0.13-4.51)	0.82 (0.08-8.07)
G010101:				
Homozygous	**2.55 (1.06-6.13)**	**4.41 (1.12-17.44)**	**2.83 (1.14-7.06)**	**3.29 (1.06-10.15)**
Heterozygous	0.95 (0.54-1.67)	1.88 (0.69-5.11)	0.97 (0.53-1.76)	1.59 (0.78-3.24)
Absent**	1.00	1.00	1.00	1.00
G010102:				
Homozygous	0.52 (0.08-3.22)	1.10 (0.11-11.18)	0.34 (0.03-3.36)	0.49 (0.05-4.89)
Heterozygous	0.78 (0.43-1.41)	0.24 (0.05-1.06)	0.86 (0.46-1.61)	0.68 (0.32-1.47)
Absent**	1.00	1.00	1.00	1.00
G010401:				
Homozygous	0.70 (0.34-1.46)	0.69 (0.20-2.44)	0.62 (0.29-1.33)	0.51 (0.20-1.30)
Heterozygous	0.64 (0.34-1.20)	0.80 (0.28-2.24)	0.54 (0.28-1.06)	0.57 (0.27-1.20)
Absent**	1.00	1.00	1.00	1.00

*cohort restricted to women with three or more visits

**Reference group

HPV Period Prevalence Sensitivity Analysis 2*

Allele / Haplotype / Genotypes	Any HPV type	α Group 1 (α-1, α-8, α-10, α-13)	α Group 2 (α-5, α-6, α-7, α-9, α-11)	α Group 3 (α-3, α-15)
	Age-Adjusted OR (95% CI)			
HLA-B*07	0.35 (0.07-1.82)	n/a	0.43 (0.08-2.29)	0.40 (0.05-3.53)
HLA-E*0101	0.94 (0.50-1.77)	0.48 (0.14-1.67)	0.91 (0.46-1.80)	1.10 (0.48-2.49)
HLA-E*0103	0.79 (0.23-2.65)	n/a	0.76 (0.21-2.72)	0.51 (0.13-1.95)
HLA-G*010101	1.03 (0.55-1.95)	1.75 (0.47-6.50)	0.88 (0.44-1.74)	2.08 (0.88-4.88)
HLA-G*010102	1.24 (0.58-2.64)	n/a	1.39 (0.63-3.06)	0.58 (0.19-1.78)
HLA-G*010103	0.45 (0.04-5.18)	n/a	0.56 (0.05-6.47)	0.97 (0.08-11.41)
HLA-G*010401	0.51 (0.24-1.07)	1.20 (0.23-6.32)	0.53 (0.24-1.18)	0.41 (0.17-1.01)
HLA-G*0106	2.04 (0.37-11.29)	n/a	2.17 (0.37-12.84)	2.27 (0.34-15.36)
HLA-DQB1*03	0.86 (0.23-3.18)	0.26 (0.04-1.77)	1.18 (0.26-5.47)	0.51 (0.12-2.11)
HLA-DQB1*0602	2.22 (0.51-9.63)	n/a	2.44 (0.52-11.42)	3.33 (0.66-17.03)
HLA-DRB1*13	3.25 (0.34-31.07)	n/a	2.54 (0.24-27.43)	7.06 (0.72-69.75)
HLA-DRB1*1501	2.27 (0.53-9.77)	n/a	2.51 (0.54-11.66)	3.32 (0.65-16.86)
DRB1*1501-DQB1*0602	2.25 (0.52-9.72)	n/a	2.48 (0.53-11.57)	3.32 (0.65-16.83)
B*07-DRB1*1501-DQB1*0602	0.83 (0.13-5.40)	n/a	1.07 (0.16-7.14)	1.05 (0.10-11.10)
G010101:				
Homozygous	1.86 (0.64-5.35)	3.15 (0.47-21.24)	1.74 (0.57-5.37)	3.32 (0.93-11.91)
Heterozygous	0.87 (0.44-1.73)	1.44 (0.35-5.85)	0.72 (0.34-1.51)	1.89 (0.77-4.62)
Absent**	1.00	1.00	1.00	1.00
G010102:				
Homozygous	1.37 (0.12-16.01)	n/a	1.67 (0.14-19.64)	n/a
Heterozygous	1.23 (0.56-2.68)	n/a	1.36 (0.60-3.10)	0.64 (0.21-1.97)
Absent**	1.00	1.00	1.00	1.00
G010401:				
Homozygous	0.48 (0.20-1.18)	1.17 (0.19-7.34)	0.57 (0.22-1.46)	**0.27 (0.08-0.89)**
Heterozygous	0.53 (0.24-1.18)	1.08 (0.18-6.43)	0.52 (0.22-1.22)	0.46 (0.17-1.23)
Absent**	1.00	1.00	1.00	1.00

*cohort restricted to women with 10 or more lifetime partners

**Reference group

HPV Period Prevalence Sensitivity Analysis 3*

Allele / Haplotype / Genotypes	Any HPV type	α Group 1 (α-1, α-8, α-10, α-13)	α Group 2 (α-5, α-6, α-7, α-9, α-11)	α Group 3 (α-3, α-15)
	Age-Adjusted OR (95% CI)			
HLA-B*07	0.53 (0.09-3.48)	n/a	0.64 (0.10-4.25)	0.63 (0.06-6.88)
HLA-E*0101	1.03 (0.39-2.70)	0.94 (0.19-4.61)	1.05 (0.38-2.92)	1.31 (0.40-4.31)
HLA-E*0103	0.60 (0.09-3.87)	n/a	0.49 (0.07-3.26)	0.34 (0.04-2.98)
HLA-G*010101	1.22 (0.46-3.22)	2.53 (0.42-15.02)	0.88 (0.32-2.41)	3.36 (0.86-13.04)
HLA-G*010102	1.59 (0.48-5.20)	n/a	2.07 (0.61-7.01)	0.52 (0.09-3.23)
HLA-G*010103	1.07 (0.06-19.27)	n/a	1.36 (0.07-25.20)	2.62 (0.13-53.79)
HLA-G*010401	0.38 (0.13-1.17)	0.67 (0.10-4.39)	0.33 (0.10-1.10)	0.40 (0.10-1.49)
HLA-G*0106	1.68 (0.13-22.54)	n/a	2.06 (0.14-29.35)	1.94 (0.07-56.97)
HLA-DQB1*03	0.50 (0.05-5.27)	0.40 (0.02-7.53)	0.86 (0.07-10.56)	0.27 (0.02-2.96)
HLA-DQB1*0602	3.16 (0.32-31.42)	n/a	2.94 (0.26-33.22)	5.97 (0.48-73.71)
HLA-DRB1*13	n/a	n/a	n/a	n/a
HLA-DRB1*1501	3.40 (0.35-33.28)	n/a	3.18 (0.29-34.91)	5.96 (0.50-71.15)
DRB1*1501-DQB1*0602	3.26 (0.33-32.27)	n/a	3.06 (0.27-34.28)	5.84 (0.48-71.57)
B*07-DRB1*1501-DQB1*0602	1.88 (0.15-22.96)	n/a	2.36 (0.19-29.80)	2.94 (0.15-58.37)
G010101:				
Homozygous	2.88 (0.64-13.01)	6.99 (0.46-105.31)	2.31 (0.49-10.96)	7.46 (0.88-63.11)
Heterozygous	0.83 (0.28-2.46)	1.95 (0.28-13.79)	0.48 (0.14-1.61)	2.82 (0.68-11.73)
Absent**	1.00	1.00	1.00	1.00
G010102:				
Homozygous	0.76 (0.04-12.96)	n/a	0.94 (0.05-16.06)	n/a
Heterozygous	1.81 (0.50-6.47)	n/a	2.39 (0.64-8.94)	0.72 (0.11-4.78)
Absent**	1.00	1.00	1.00	1.00
G010401:				
Homozygous	0.29 (0.07-1.18)	0.52 (0.05-5.63)	0.33 (0.08-1.37)	0.18 (0.03-1.28)
Heterozygous	0.45 (0.14-1.45)	0.79 (0.11-5.44)	0.36 (0.10-1.28)	0.53 (0.13-2.06)
Absent**	1.00	1.00	1.00	1.00

*cohort restricted to women with three or more visits AND to women with 10 or more lifetime partners

**Reference group

HPV Persistence

Allele / Haplotype / Genotype	Transient Infections n	Persistent Infections n	Persistent Infections vs. Transient Infections Age-Adjusted OR (95% CI)	Persistent Infections vs. HPV Negative Age-Adjusted OR (95% CI)
Cohort (n=282)	87	80		
HLA-B*07	1	7	7.30 (0.86-61.96)	0.94 (0.29-3.06)
HLA-E*0101	47	49	1.30 (0.68-2.49)	1.40 (0.71-2.75)
HLA-E*0103	74	66	0.47 (0.16-1.35)	0.70 (0.25-1.93)
HLA-G*010101	43	48	1.32 (0.69-2.52)	1.14 (0.58-2.23)
HLA-G*010102	24	20	0.81 (0.40-1.63)	0.67 (0.32-1.40)
HLA-G*010103	1	3	2.75 (0.28-27.39)	2.00 (0.24-16.44)
HLA-G*010401	59	52	0.77 (0.38-1.54)	0.77 (0.38-1.59)
HLA-G*0106	2	1	0.42 (0.04-4.82)	0.80 (0.07-8.91)
HLA-DQB1*03	81	70	0.79 (0.23-2.72)	1.13 (0.32-4.01)
HLA-DQB1*0602	2	3	1.46 (0.23-9.15)	0.53 (0.12-2.29)
HLA-DRB1*13	4	4	0.90 (0.21-3.81)	0.69 (0.16-3.02)
HLA-DRB1*1501	2	4	1.88 (0.33-10.78)	0.61 (0.16-2.36)
DRB1*1501-DQB1*0602	2	3	1.46 (0.23-9.14)	0.52 (0.12-2.24)
B*07-DRB1*1501-DQB1*0602	0	1	n/a	0.34 (0.04-3.27)
G010101:				
Homozygous	11	9	1.07 (0.39-2.95)	1.51 (0.53-4.29)
Heterozygous	30	33	1.37 (0.66-2.84)	0.99 (0.49-2.04)
Absent*	35	27	1.00	1.00
G010102:				
Homozygous	0	1	n/a	0.27 (0.03-2.57)
Heterozygous	22	15	0.70 (0.33-1.50)	0.69 (0.31-1.53)
Absent*	54	53	1.00	1.00
G010401:				
Homozygous	17	15	0.90 (0.35-2.31)	1.09 (0.43-2.81)
Heterozygous	39	34	0.84 (0.39-1.85)	0.71 (0.31-1.63)
Absent*	20	20	1.00	1.00

*Reference group

HPV Alpha Group 2 Persistence (α-5, α-6, α-7, α-9, α-11)

Allele / Haplotype / Genotype	Transient Infections n	Persistent Infections n	Persistent Infections vs. Transient Infections Age-Adjusted OR (95% CI)	Persistent Infections vs. HPV Negative Age-Adjusted OR (95% CI)
Cohort (n=282)	75	63		
HLA-B*07	2	6	3.29 (0.62-17.45)	1.50 (0.44-5.15)
HLA-E*0101	42	38	1.27 (0.61-2.64)	1.29 (0.65-2.55)
HLA-E*0103	64	51	0.39 (0.12-1.28)	0.68 (0.25-1.83)
HLA-G*010101	38	38	1.29 (0.63-2.65)	1.24 (0.62-2.45)
HLA-G*010102	23	18	0.86 (0.40-1.85)	1.14 (0.54-2.38)
HLA-G*010103	1	2	1.86 (0.16-21.61)	1.51 (0.20-11.54)
HLA-G*010401	49	39	0.76 (0.36-1.63)	0.61 (0.29-1.26)
HLA-G*0106	1	0	n/a	n/a
HLA-DQB1*03	71	53	0.41 (0.10-1.72)	0.86 (0.27-2.77)
HLA-DQB1*0602	2	2	1.10 (0.15-8.20)	0.46 (0.09-2.47)
HLA-DRB1*13	3	4	1.34 (0.28-6.33)	0.99 (0.25-4.02)
HLA-DRB1*1501	2	3	1.51 (0.24-9.51)	0.61 (0.14-2.69)
DRB1*1501-DQB1*0602	2	2	1.10 (0.15-8.19)	0.46 (0.09-2.45)
B*07-DRB1*1501-DQB1*0602	1	0	n/a	n/a
G010101:				
Homozygous	10	8	1.17 (0.39-3.52)	1.92 (0.66-5.57)
Heterozygous	26	24	1.14 (0.50-2.57)	0.93 (0.44-1.96)
Absent*	29	21	1.00	1.00
G010102:				
Homozygous	0	1	n/a	0.62 (0.06-5.93)
Heterozygous	21	14	0.80 (0.35-1.80)	1.24 (0.56-2.77)
Absent*	44	38	1.00	1.00
G010401:				
Homozygous	14	10	0.78 (0.27-2.23)	0.75 (0.28-2.01)
Heterozygous	33	26	0.78 (0.33-1.84)	0.64 (0.28-1.44)
Absent*	18	17	1.00	1.00

*Reference group

HPV Persistence Sensitivity Analysis*

Allele / Haplotype / Genotype	Transient Infections n	Persistent Infections n	Persistent Infections vs. Transient Infections Age-Adjusted OR (95% CI)	Persistent Infections vs. HPV Negative Age-Adjusted OR (95% CI)
Cohort (n=266)	105	46		
HLA-B*07	4	2	1.06 (0.18-6.21)	1.06 (0.20-5.69)
HLA-E*0101	64	25	0.83 (0.40-1.74)	1.10 (0.51-2.37)
HLA-E*0103	92	38	0.96 (0.28-3.35)	1.52 (0.44-5.24)
HLA-G*010101	61	26	1.07 (0.51-2.25)	1.59 (0.74-3.41)
HLA-G*010102	28	9	0.72 (0.30-1.69)	0.60 (0.25-1.45)
HLA-G*010103	1	1	2.24 (0.13-37.26)	1.77 (0.15-21.39)
HLA-G*010401	73	28	0.79 (0.36-1.72)	0.60 (0.27-1.38)
HLA-G*0106	2	1	1.34 (0.12-15.37)	1.12 (0.11-11.91)
HLA-DQB1*03	96	40	0.71 (0.20-2.57)	0.42 (0.10-1.84)
HLA-DQB1*0602	4	2	1.12 (0.20-6.46)	0.79 (0.15-4.08)
HLA-DRB1*13	5	3	1.56 (0.35-6.94)	1.36 (0.27-6.91)
HLA-DRB1*1501	5	2	1.00 (0.18-5.44)	0.82 (0.16-4.23)
DRB1*1501-DQB1*0602	4	2	1.21 (0.21-6.97)	0.82 (0.16-4.24)
B*07-DRB1*1501-DQB1*0602	1	2	3.93 (0.32-47.59)	2.44 (0.27-16.08)
G010101:				
Homozygous	15	6	0.97 (0.31-3.01)	2.36 (0.69-8.09)
Heterozygous	46	20	1.12 (0.51-2.48)	1.38 (0.61-3.12)
Absent**	40	16	1.00	1.00
G010102:				
Homozygous	1	1	2.25 (0.14-37.20)	0.62 (0.06-6.36)
Heterozygous	27	8	0.66 (0.27-1.61)	0.60 (0.24-1.50)
Absent**	73	33	1.00	1.00
G010401:				
Homozygous	19	9	0.94 (0.34-2.62)	0.63 (0.23-1.74)
Heterozygous	54	19	0.73 (0.32-1.69)	0.61 (0.25-1.49)
Absent**	28	14	1.00	1.00

*persistence is defined as a woman who tests positive for the same HPV type on two or more consecutive visits

**Reference group

HPV Alpha Group 2 Persistence Sensitivity Analysis (α-5, α-6, α-7, α-9, α-11)

Allele / Haplotype / Genotype	Transient Infections	Persistent Infections	Persistent Infections vs. Transient Infections	Persistent Infections vs. HPV Negative
	n	n	Age-Adjusted OR (95% CI)	Age-Adjusted OR (95% CI)
Cohort (n=266)	88	39		
HLA-B*07	4	1	0.59 (0.06-5.38)	0.7 (0.08-6.63)
HLA-E*0101	56	21	0.68 (0.30-1.54)	1.10 (0.50-2.41)
HLA-E*0103	76	32	0.72 (0.20-2.66)	1.12 (0.33-2.85)
HLA-G*010101	51	22	0.98 (0.44-2.19)	1.45 (0.66-3.21)
HLA-G*010102	27	7	0.50 (0.20-1.29)	0.65 (0.25-1.68)
HLA-G*010103	0	1	n/a	1.46 (0.13-16.00)
HLA-G*010401	57	26	1.17 (0.49-2.79)	0.82 (0.34-1.98)
HLA-G*0106	1	1	2.30 (0.14-37.88)	0.82 (0.07-9.16)
HLA-DQB1*03	78	35	1.80 (0.36-8.90)	0.77 (0.14-4.34)
HLA-DQB1*0602	3	1	0.77 (0.08-7.65)	0.37 (0.04-3.20)
HLA-DRB1*13	5	2	0.92 (0.17-5.03)	0.96 (0.16-5.73)
HLA-DRB1*1501	4	1	0.62 (0.07-5.76)	0.39 (0.04-3.44)
DRB1*1501-DQB1*0602	3	1	0.84 (0.08-8.36)	0.39 (0.04-3.41)
B*07-DRB1*1501-DQB1*0602	1	1	2.55 (0.15-42.97)	1.51 (0.15-15.58)
G010101:				
Homozygous	13	4	0.77 (0.21-2.84)	1.47 (0.39-5.60)
Heterozygous	38	18	1.09 (0.47-2.56)	1.42 (0.62-3.27)
Absent**	32	14	1.00	1.00
G010102:				
Homozygous	1	0	n/a	n/a
Heterozygous	26	7	0.52 (0.20-1.35)	0.75 (0.28-1.96)
Absent**	56	29	1.00	1.00
G010401:				
Homozygous	14	8	1.48 (0.48-4.62)	0.81 (0.28-2.39)
Heterozygous	43	18	1.08 (0.42-2.73)	0.83 (0.33-2.11)
Absent**	26	10	1.00	1.00

*persistence is defined as a woman who tests positive for the same HPV type on two or more consecutive visits

**Reference group