Prevalence of human endogenous retroviral element associates with Hodgkin’s lymphoma incidence rates

Wee Hong Woo\textsuperscript{a,b}, Liang Shen\textsuperscript{c}, Sai Mun Leong\textsuperscript{d,\#}, Evelyn Siew–Chuan Koay\textsuperscript{b,d}

\textsuperscript{a} School of Chemical & Life Sciences, Singapore Polytechnic, Singapore
\textsuperscript{b} Department of Pathology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
\textsuperscript{c} Dean’s Office, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
\textsuperscript{d} Molecular Diagnosis Centre, Department of Laboratory Medicine, National University Hospital, National University Health System, Singapore

\textbf{1. Introduction}

Leukaemias, lymphomas and multiple myeloma can be the dreadful sequelae of viral infections. RNA oncogenic viruses, which are also known as retroviruses, were initially known to cause lymphoproliferative disorders only in animals, but have also been reported to demonstrate their transforming ability in humans. Retroviruses are capable of integrating their proviral DNA into the host genome, which may in turn, mediate leukemogenesis via insertional mutagenesis. For instance, human T-cell lymphotropic virus type 1 is the causative agent of adult T-cell leukaemia–lymphoma \cite{1}. Akin to exogenous retroviruses, human endogenous retroviruses (HERVs) have integrated into the human genome since at least 30 million years ago and are transmitted vertically via the germ line \cite{2}. Some family members of HERVs (HERV-K and HERV-H) are implicated in leukaemias and lymphomas, although the precise molecular mechanism underlying HERV-mediated carcinogenesis remains to be elucidated \cite{3}. While seroprevalence to exogenous viral infections in endemic regions have been shown to be associated with the incidence rate of lymphoid cancers \cite{1,4}, information about the prevalence of HERV-H in different populations in relation to lymphoid neoplasms remains unknown. Of interest, a recent study reported that endogenous LTR derepression is involved in the pathogenesis of Hodgkin’s lymphoma \cite{5}. This provide the impetus for us to determine the prevalence of HERV-H in the three major ethnic groups (Chinese, Malays and Indians) residing in Singapore, and examine the association between the prevalence rate of each of these ethnic group with the reported incidence rate of Hodgkin's lymphoma retrieved from the Singapore Cancer Registry.

\textbf{2. Materials and methods}

\textbf{2.1. Study design and subjects}

To investigate the prevalence of HERV-H in a cross-section of the Singapore population, 808 apparently healthy subjects, aged 16–60 years, were randomly recruited from the residents in Singapore during April 2010–November 2011. The mean age at recruitment was 23.3 years. Of these, 23 subjects whose genomic materials were classified as indeterminate were excluded from the study, resulting in a total number of 785 subjects.

The study protocol was approved by the Ethics Review Committee of Singapore Polytechnic.
2.2. Genotype analysis

Buccal cells were collected from volunteers using the MasterAmp™ buccal brush (Epicentre, USA) and their genomic DNA was isolated using the ReliaPrep gDNA Tissue Miniprep System (Promega, USA). Each individual’s DNA was tested for the presence of HERV-H element by using polymerase chain reaction (PCR) amplifications as previously described [6]. Briefly, PCR was conducted using the primers HERV-H: forward, 5′-CTT CCC TCG GTG TCT TTA CG-3′ and reverse, 5′-AAG ATT AGA CAC ACT CAG CAA CG-3′. The PCR mix (volume of 50 μl) contained 20 ng of extracted genomic DNA template, 2.0 μM of each primer, 200 μM of each deoxynucleoside triphosphate, 15 mM of MgCl₂, and 2.5 units of GoTaq® DNA polymerase (Promega, USA). PCR amplification was conducted with 2 min of initial denaturation at 95 °C, 35 cycles of 30 s at 95 °C, 30 s at annealing temperature at 60 °C, and 60 s of elongation at 72 °C, followed by a 10 min final extension at 72 °C. To monitor for reagent contamination, a no template control (NTC) was included in every batch of PCR amplification. A housekeeping gene, β-actin, was used to validate the true negatives.

2.3. Statistical methods

Analyses were conducted separately for gender and ethnicity. Prevalence rates and 95% confidence intervals (95% CIs) were calculated using the ProMESA software, version 1.62 (Epicentre, Madison, Wisconsin, USA). Relative risks and odds ratios (ORs) with 95% CIs were calculated using the GraphPad Prism Version 6.0c (GraphPad Software, USA) to estimate the association of HERV-H with the variables. Differences between groups were estimated by chi-square test (statistical significance for P < 0.05). Odds ratios statistical tests were based on 2-sided probability. Correlations between incidence rates and relative risks were calculated using Pearson’s correlation (GraphPad Prism).

3. Results

Of 808 subjects recruited, 785 were successfully analysed for HERV-H element. The prevalence rates of HERV-H positivity, based on gender as well as categorised by their ethnic descent, are shown in Table 1. 89.9% (381/427) females and 88.6% (320/361) males were found to possess HERV-H in their genome, resulting in a total of 89.3% HERV-H positivity overall. With respect to gender, there is no statistical difference between males and females (P > 0.05).

The prevalence rates of HERV-H positivity for Chinese, Malays and Indians are 86.1%, 95.4%, and 96.8%, respectively. To examine the significance of ethnic disparities, we performed a Chi-square test. Compared to Chinese, both Malays and Indians were found to possess statistically higher prevalence rates (P < 0.01). In addition, Malays were also found to be 1.11 times likely (95% CI = 1.05–1.17; P < 0.01), and Indians 1.12 times likely (95% CI = 1.07–1.18; P < 0.01) to be HERV-H positive when compared to Chinese.

Fig. 1 depicts the strong association between the Hodgkin’s lymphoma incidence rates and the prevalence rate for HERV-H positivity in three major ethnic groups of Singapore (1A: χ² = 0.9921 for male; 1B: r = 0.9801 for female). The age standardised rate risks data was obtained from the Singapore Cancer Registry Report No. 7 [7].

4. Discussion

HERVs represent the trail of successful ancient retroviral infection of the human germ line. As a consequence to the persistence of vertical transmission, about 8% of the human genome is occupied by HERVs [2]. Their existence, through millions of years of evolution, implies that HERVs have overcome the stringent conditions of evolution, and domesticated and incorporated into the host genomes, although

### Table 1

HERV-H positivity in the study Singapore population. **Source:** Singapore Department of Statistics.

| Race/ethnicity | HERV-H positive | | | HERV-H negative | | | |
|---|---|---|---|---|---|---|---|
| | n | Prevalence rate (%) | 95% CI | Value | n | Prevalence rate (%) | Relative risk | Odds ratio | Value |
| | | | 95% CI | | | | | | |
| Total Gender | 701 | 89.3 | 87.1–91.5 | N.A. | 84 | 10.7 | 8.5–12.9 | N.A. | N.A. |
| Male | 320 | 88.6 | 86.4–90.8 | Reference | 41 | 11.4 | 9.2–13.6 | 1.00 | 1 |
| Female | 381 | 89.9 | 87.8–92.0 | 1.3 (–1.76 to 4.36) | 0.583 | 43 | 11.1 | 7.99–12.2 | 1.01 | 1.14 (0.72–1.79) | 0.583 |
| Race/ethnicity | Chinese | 438 | 86.1 | 83.7–88.5 | Reference | 71 | 13.9 | 11.5–16.3 | Reference | 1 |
| | Malay | 124 | 95.4 | 93.9–96.9 | 9.3 (6.47–12.1) | 0.0035 | 6 | 4.6 | 3.13–6.07 | 1.11 (1.05–1.17) | 3.35 (1.42–7.89) | 0.0037 |
| | Indian | 121 | 96.8 | 95.6–98.0 | 10.7 (7.98–13.42) | 0.0009 | 4 | 3.2 | 1.97–4.43 | 1.12 (1.07–1.18) | 4.90 (1.76–13.7) | 0.0011 |
| | Male | 253 | 87.8 | 85.5–90.1 | Reference | 36 | 16.3 | 13.7–18.9 | Reference | 1 |
| | Female | 253 | 87.8 | 85.5–90.1 | Reference | 35 | 12.2 | 9.91–14.5 | Reference | 1 |
| | Malay | 56 | 94.9 | 93.4–96.4 | Reference | 3 | 3.2 | 2.8–5.6 | 1.14 (1.06–1.23) | 4.41 (1.32–14.8) | 0.0149 |
| | Indian | 121 | 96.8 | 95.6–98.0 | 10.7 (7.98–13.42) | 0.0009 | 4 | 3.2 | 1.97–4.43 | 1.12 (1.07–1.18) | 4.90 (1.76–13.7) | 0.0011 |
| | Male | 56 | 98.2 | 97.1–99.1 | Reference | 1 | 1.8 | 0.87–2.73 | 1.17 (1.10–1.26) | 10.9 (1.46–81.3) | 0.0035 |
| | Female | 65 | 95.6 | 94.2–97.0 | Reference | 3 | 4.4 | 2.97–5.83 | 1.09 (1.02–1.16) | 3.00 (0.89–10.1) | 0.0629 |
| | Other races | 18 | 85.7 | 83.3–88.2 | 0.4 (–3.04–3.84) | 0.9652 | 3 | 14.3 | 11.9–16.8 | 0.996 (0.83–1.19) | 0.97 (0.28–3.39) | 0.9652 |

* Other (Burmesse, Ceylonese, Eurasian, Filipino, Japanese, Javanese, Pakistani, Punjabi, Sikh, Vietnamese),
| * 95% CI: 95% confidence interval; calculated based on Singapore Population in 2011 i.e., 5,183,700.
the accumulative effects of mutations, deletions, frame shifts, premature stop codons and hypermethylation of the promoter sites may have rendered them defective or inactive. While retro-elements such as Alu (in SINE) is implicated in pathogenesis of autoimmune lymphoproliferative syndrome and X-linked agammaglobulinemia, and L1 (in LINE) in haemophilia A and B [8], the role of HERV in carcinogenesis remains poorly understood.

Interestingly, our analysis reveals strong positive correlation between age standardized incidence rate of Hodgkin’s lymphoma and prevalence rate for HERV-H positivity. Based on genetic similarity in the pol region, HERV-H belongs to Class I of the Gammaretrovirus-like family. Apart from possessing long terminal repeats (LTR)s, HERV-H is also known to utilise histidine (H) transfer RNA (tRNA) at its primer binding site to initiate reverse transcription [2]. A previous study demonstrated that derepression of endogenous LTR due to hypomethylation results in up-regulation of CSF1R proto-oncogene expression, which in turn contributes to the pathogenesis of Hodgkin’s lymphoma [5]. In addition, endogenous retroviral Np9 and Rec were identified as oncogenic players in leukemia [9]. Perturbation to HERV-H-mediated pluripotency regulation may lead to deregulation in self-renewal control and emergence of cancer stem cell, though this hypothesis remains to be proven.

In contrast to the low insertion frequency in Asian population for HERV-K113 (13%) and -K115 (4%) [10], we demonstrate herein high insertion frequency (89.3%) of HERV-H in the Singapore population, and observe interesting correlation between prevalence and age-standardized incidence rate of Hodgkin’s lymphoma for the three races residing in Singapore. Together, current and other study [10] provides the differential prevalence rate of various HERV elements in different human race, and can serve to demonstrate these “jumping genes” retrotransposons (e.g., HERV-H vs. HERV-K) have vastly different distribution frequency. Further prospective or retrospective HERV-H gene expression studies involving cohort of Hodgkin’s lymphoma patients are required to directly implicated endogenous retroviral elements in lymphomagenesis and investigate the molecular mechanisms underlying the process.

Authors’ contributions

W.H.W. performed experiments. W.H.W., L.S., S.M.L. and E.S-C. K. interpreted the data. W.H.W. designed the study and drafted the manuscript. S.M.L. and E.S-C.K. revised the manuscript critically for important intellectual content. All authors approved the final version of the manuscript.

Acknowledgements

This work was supported by Grants from Singapore Polytechnic [Grant numbers 11–27801-45–2214, 11–27801-45–2554, 11–27801–45–2762 and 11–27801–45–3286].

References

[1] Proietti FA, Carneiro-Proietti AB, Catalán-Soares BC, Murphy EL. Global epidemiology of HTLV-I infection and associated diseases. Oncogene 2005;24:6058–68.
[2] Bannert N, Kurth R. Retroelements and the human genome: new perspectives on an old relation. Proc Natl Acad Sci USA 2004;101(Suppl. 2):14572–9.
[3] Romanish MT, Cohen CJ, Mager DL. Potential mechanisms of endogenous retroviral-mediated genomic instability in human cancer. Semin Cancer Biol 2010;20:246–53.
[4] Magnath I. Epidemiology: clues to the pathogenesis of Burkitt lymphoma. Br J Haematol 2012;156:744–56.
[5] Lamprecht B, Walter K, Kreher S, Kumar R, Hummel M, Lenze D, et al. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. Nat Med 2010;16:571–8.
[6] Alves PM, Levy N, Stevenson BJ, Bouzourene H, Theiler G, Bricard G, et al. Identification of tumor-associated antigens by large-scale analysis of genes expressed in human colorectal cancer. Cancer Immun 2009;2:10.
[7] National Registry of Diseases Office. Trends in cancer incidence in Singapore 1968–2007. Singapore Cancer Registry Report No. 7. Singapore Cancer Registry, Singapore; 2010.
[8] Solyom S, Kazazian Jr HH. Mobile elements in the human genome: implications for disease. Genome Med 2012;4:12.
[9] Santoni FA, Guerra J, Luban J. HERV-H RNA is abundant in human embryonic stem cells and a precise marker for pluripotency. Retrovirology 2012;9:111.
[10] Jha AR, Pillai SK, York VA, Sharp EF, Storm EC, Wachter DJ, et al. Cross-sectional dating of novel haplotypes of HERV-K 113 and HERV-K 115 indicate these proviruses originated in Africa before Homo sapiens. Mol Biol Evol 2009;26:2617–26.