

The Etiology of Multiple Sclerosis: Genetic Evidence for the Involvement of the Human Endogenous Retrovirus HERV-Fc1

Bjørn A. Nexø^{1*}, Tove Christensen², Jette Frederiksen³, Anné Møller-Larsen², Annette B. Oturai⁴, Palle Villesen⁵, Bettina Hansen¹, Kari K. Nissen¹, Magdalena J. Laska^{1,6}, Trine S. Petersen¹, Sandra Bonnesen¹, Anne Hedemand¹, Tingting Wu⁷, Xinjie Wang⁷, Xiuqing Zhang⁷, Tomasz Brudek², Romana Maric², Helle B. Søndergaard⁴, Finn Sellebjerg⁴, Klaus Brusgaard⁸, Anders L. Kjeldbjerg⁹, Henrik B. Rasmussen¹⁰, Anders L. Nielsen¹, Mette Nyegaard¹, Thor Petersen⁶, Anders D. Børglum¹, Finn S. Pedersen⁹

1 Department of Human Genetics, University of Aarhus, Aarhus C, Denmark, **2** Department of Medical Microbiology, University of Aarhus, Aarhus C, Denmark, **3** Department of Neurology, Glostrup Hospital, Glostrup, Denmark, **4** Department of Neurology, The Danish Multiple Sclerosis Research Center, Copenhagen University Hospital, Copenhagen, Denmark, **5** Bioinformatics Research Center, University of Aarhus, Aarhus C, Denmark, **6** Department of Neurology, Aarhus University Hospital, Aarhus C, Denmark, **7** BGI-Shenzhen, Shenzhen, China, **8** Department of Clinical Genetics, Odense University Hospital, Odense, Denmark, **9** Department of Molecular Biology, University of Aarhus, Aarhus C, Denmark, **10** Research Institute of Biological Psychiatry, Psychiatry Center Sct. Hans, Roskilde, Denmark

Abstract

We have investigated the role of human endogenous retroviruses in multiple sclerosis by analyzing the DNA of patients and controls in 4 cohorts for associations between multiple sclerosis and polymorphisms near viral restriction genes or near endogenous retroviral loci with one or more intact or almost-intact genes. We found that SNPs in the gene *TRIM5* were inversely correlated with disease. Conversely, SNPs around one retroviral locus, HERV-Fc1, showed a highly significant association with disease. The latter association was limited to a narrow region that contains no other known genes. We conclude that HERV-Fc1 and *TRIM5* play a role in the etiology of multiple sclerosis. If these results are confirmed, they point to new modes of treatment for multiple sclerosis.

Citation: Nexø BA, Christensen T, Frederiksen J, Møller-Larsen A, Oturai AB, et al. (2011) The Etiology of Multiple Sclerosis: Genetic Evidence for the Involvement of the Human Endogenous Retrovirus HERV-Fc1. PLoS ONE 6(2): e16652. doi:10.1371/journal.pone.0016652

Editor: Robert Belshaw, University of Oxford, United Kingdom

Received: September 28, 2010; **Accepted:** December 23, 2010; **Published:** February 2, 2011

Copyright: © 2011 Nexø et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by The Lundbeck Foundation, The Danish Multiple Sclerosis Society, Warwara Larsens Fond, The NovoNordisk Foundation and The Villum Kann Rasmussen Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: BAN, TC, AML, MJL, PV, and FSP have submitted a patent application covering the findings of this paper. This does not alter the authors' adherence to all the PLoS ONE policies on sharing data and materials.

* E-mail: nexo@humgen.au.dk

Introduction

Two views have dominated the discussions on the etiology of multiple sclerosis (MS[MIM 126200]) in recent decades: It could be a genetic disease, or it could be a disease caused by an infectious agent. The genetic view is most clearly expounded by studies of twins, showing that concordance among monozygotic twins is approximately 25 percent, while concordance among dizygotic twins is only 2–3 percent [1]. The alternative view, that multiple sclerosis is caused by an infectious agent, is most clearly backed by animal studies [2–4]. The obvious intersections of a genetic and a viral etiology are the human endogenous retroviruses, which are Mendelian inherited proviral loci in the human genome.

The endogenous human retroviruses exist in multiple types, generally named after the amino acid of the tRNA that serves as primer for first strand synthesis. Some endogenous virus types (–H and –W) are old and universal in a large range of species, whereas certain of the –K type integrations are recent and still segregate in the human population. HERV-Fc endogenous viruses can be traced

as far back as the simian forefather, however present copies in our genome are more recent acquisitions and specifically the HERV-Fc1 locus on the X-chromosome, with which this paper is concerned, is shared by man, chimpanzee, and gorilla, but not orangutan. This put the time of integration at about 10 to 15 million years ago [5]. We have confirmed the absence of HERV-Fc1 in orangutan and African Green Monkey by specific PCR (Magdalena J. Laska, unpublished). Homology studies of the *pol* genes have shown that HERV-Fc is closely related to HERV-H, more distantly related to HERV-W and only remotely related to HERV-K [6].

In this study, we have attempted to evaluate the role of retroviruses in the etiology of Multiple Sclerosis, both indirectly by studying genes that restrict retroviruses, and directly by searching for an endogenous retrovirus associated with occurrence of disease.

Results

Inspired by the classical genetic studies of the mouse restriction gene for retroviruses, *Fv-1* [7], we first attempted to evaluate the

importance of retroviruses for MS indirectly. Several human restriction genes are known to influence retroviral replication to various degrees. If MS were caused by a retrovirus, these genes could well influence the risk of disease. Of the restriction genes, *TRIM5* appears to be the closest functional relative of *Fv-1* [8]. Therefore, to substantiate the involvement of a retrovirus in multiple sclerosis we sought for linkage disequilibrium between markers in *TRIM5* and disease. Initially, we investigated a small cohort consisting of 28 verified cases [9] and 24 controls. We tested 12 SNPs, located in *TRIM5* intron 1. The nearby exon 2 encodes the RING domain in the N-terminus of the protein. This domain plays a major role in recognition of retroviral cores. The SNPs and their primers are listed in Table S1. We found that 7 of the SNPs were suggestive of disequilibrium between SNP and disease ($p \leq 0.1$) (Table S2). Table 1 shows the results with the best of these, rs3802981, giving a p-value of 0.004 for association. After Bonferroni correction the p value became 0.05, i.e. the results were just globally significant. The table also shows the results from the analysis of another larger cohort (Cohort2). This cohort consisted of DNA from 350 patients with verified multiple sclerosis [9] living in Western Denmark as well as DNA from 500 persons without sclerosis from the same part of the country. Females constituted 62 percent of cases and 67 percent of the controls. Again, rs3802981 was associated with disease ($p = 0.008$). The combined p-value for the two cohorts using Fisher's formula was 0.0003; after Bonferroni correction the p value was 0.004.

We were similarly inspired by the classical Mendelian studies of the highly active endogenous ecotropic viruses of the AKR mouse [10,11], which showed that genetic mapping of germ line retroviruses is straightforward. Therefore, we attempted to associate multiple sclerosis with specific human endogenous retroviruses by genetic epidemiology. Intact or near-intact human endogenous retroviruses were treated as inheritable loci in the analyses. We speculated that in order to contribute to disease a retroviral sequence should contain at least one functional or near-functional major gene i.e. a gene with at most 2 stop-codons or frame-shifts, and excluded viruses that did not fulfill this criterion using the RetroSearch tool [12]. Moreover, while many such retroviruses are rather similar, their sites of integration are generally unique, so by testing neighboring SNPs (less than 10 kb away from the virus) we should be able to test the viral loci individually. A cartoon of the location of all such endogenous viruses in the human genome can be found in Figure S1. A list of SNPs near these viruses, for which we could perform genotyping, is shown in Table S3, sorted into 13 plexes for the Sequenom and shown together with the primers. Of the 52 retroviruses identified as relevant 48 were covered by this approach, while 4 could not be tested for association.

Initially, we investigated 193 SNPs in the DNA from Cohort 2. There were clear-cut clusters of significant SNPs in a few chromosomal locations. These locations were substantiated by analyzing additional nearby SNPs. A total of 220 SNPs were tested. The most significant results are indicated in Figure S1, and all results are listed in Table S4.

A particularly striking cluster of significant SNPs occurred on chromosome X at approximate chromosome position 97100000 around the HERV-Fc1 proviral locus (NCBI genome build 37.1). Here, 3 close-lying SNPs gave p-values less than 0.001. The SNP rs391745 was lowest with a p-value of 4×10^{-6} (2-sided) for association with disease, when calculating on the basis of C-allele carriers (Table 2, cohort 2). After Bonferroni correction this p-value was 0.001.

We retested the association of rs391745 and multiple sclerosis in Cohort 3. This cohort consisted of 542 cases and 1165 controls with 67 percent female cases and 65 percent female controls. rs391745 was again associated with MS ($p = 0.01$) (Table 2, cohort 3). Finally, we tested a third, smaller cohort of verified multiple sclerosis consisting of 235 cases and 407 controls in Eastern Danes (Cohort 4). In this cohort women constituted 70 and 63 percent among cases and controls, respectively. We could not find any association of rs391745 with multiple sclerosis in this cohort ($p = 0.97$) (Table 2, cohort 4), and do not know, why the association was absent. The p-value for all 3 cohorts combined was 0.00002 (2-sided) and the p-value was thus significant after Bonferroni correction ($p = 0.004$). Similar conclusions were reached when the association was calculated on the basis of allele frequencies, taking into account that males are hemizygous for the SNP (Table S5). In this case, Bonferroni-analysis showed that the corrected p-value was 0.0076. The female controls of cohort 3 were in marginal Hardy-Weinberg disequilibrium ($p = 0.067$), the females in the other cohorts were in HW equilibrium ($p > 0.1$).

To substantiate the association of HERV-Fc1 with multiple sclerosis, we performed a scan of the region surrounding the HERV-Fc1 provirus in Cohort 2. Table S3, plex 11–13 includes the extra SNPs tested. Figure 1 shows the association of a number of polymorphisms in and around the locus in relation to their position on the chromosome. It is clear that the association occurred in a very narrow region around the provirus. In contrast, the nearest known genes lie 141 kb upstream, and 57 kb downstream, respectively. No microRNA genes or small nucleolar RNA genes have been reported in the region.

Looking further for an interaction or additive effect between HERV-Fc1 and *TRIM5* we depicted the odds ratio for disease as a function of the SNPs rs391745 and rs3802981, representing the two genes. The results are shown in Figure 2. It is obvious that

Table 1. Association of multiple sclerosis with the polymorphism rs3802981 located in the first intron of *TRIM5*.

Cohort	Status	CC	CT	TT	Fraction of TT genotypes	OR (CI95%) CC + CT vs TT	P-chitest (Pearson)
1	Cases	2	12	14	0.50	0.20 (0.05–0.74)	0.004
	Controls	10	10	4	0.17		
2	Cases	76	154	114	0.49	0.64 (0.47–0.86)	0.008
	Controls	122	292	131	0.31		
1+2	Cases	78	166	128	0.52	0.59 (0.44–0.79)	0.002
	Controls	132	302	135	0.31		

F: p-value as calculated from cohort1 and 2 by Fisher's formula for the combination of p-values.
doi:10.1371/journal.pone.0016652.t001

Table 2. Association of multiple sclerosis with rs391745 located near HERV-Fc1.

Cohort	Status	C-allele carriers	Non-carriers	Fraction C-allele carriers	OR (CI95) for carriers	P (2-sided)
2	Cases	79	246	0.24	2.29 (1.60–3.28)	4×10^{-6}
	Controls	68	485	0.12		
3	Cases	99	407	0.20	1.43 (1.09–1.89)	0.010
	Controls	165	972	0.15		
4	Cases	35	196	0.15	1.01 (0.64–1.59)	0.97
	Controls	59	333	0.15		
2+3+4	Cases	213	849	0.20	1.54 (1.27–1.87)	1.3×10^{-5}
	Controls	292	1790	0.14		

The p-values were calculated from the numbers of carriers and non-carriers using Pearson's Chi-square test. The p-value for all 3 groups marked F was calculated from the individual p-values using Fisher's formula.

doi:10.1371/journal.pone.0016652.t002

disease risk is strongly influenced by the two genes, albeit largely in an additive fashion.

Discussion

By treating the endogenous retroviruses as inheritable loci, we have brought the formidable knowledge of human genetics to bear on their role in disease, and avoided the difficulties of establishing the causality of a ubiquitous agent by conventional means. It is not clear if HERV-Fc1 is infectious, but analyses of the proviral DNA suggest the presence of two stop-codons and a frame-shift in the *pol* gene. The reading frames of the *gag* and *env* genes appear to be complete (NCBI build 37.1). We cannot exclude the contribution of other endogenous viruses to putative infectivity through pseudotyping or recombination. Perhaps the relevant determinant of infectivity in the organism is not the individual locus, but the repertoire of proviral sequences [13]. The p-values obtained are convincing arguments for an association of

HERV-Fc1 with multiple sclerosis. However, the odds ratios are less striking than what could be hoped for. We attribute the latter to the fact that we have not yet found the causative variations, but depend on associations between close SNPs. The latter also applies to the results with *TRIM5*.

The gene *TRIM5* encodes an ubiquitin ligase that restricts the replication of many retroviruses to varying degrees [9]. The gene product accelerates unpacking of the viral cores and leads to abortive infections. In many ways, TRIM5 action resembles the effect of the mouse restriction gene *Fv-1*. The association of *TRIM5* with multiple sclerosis is independent confirmation that a retrovirus plays a role in the etiology of MS.

Our results should be seen in combination with previous findings in genomic scans that demonstrate that the major histocompatibility complex on chromosome 6 is important for this disease [14], as seems the receptor for interleukin 7 [15]. Retroviral infections often develop into running battles between

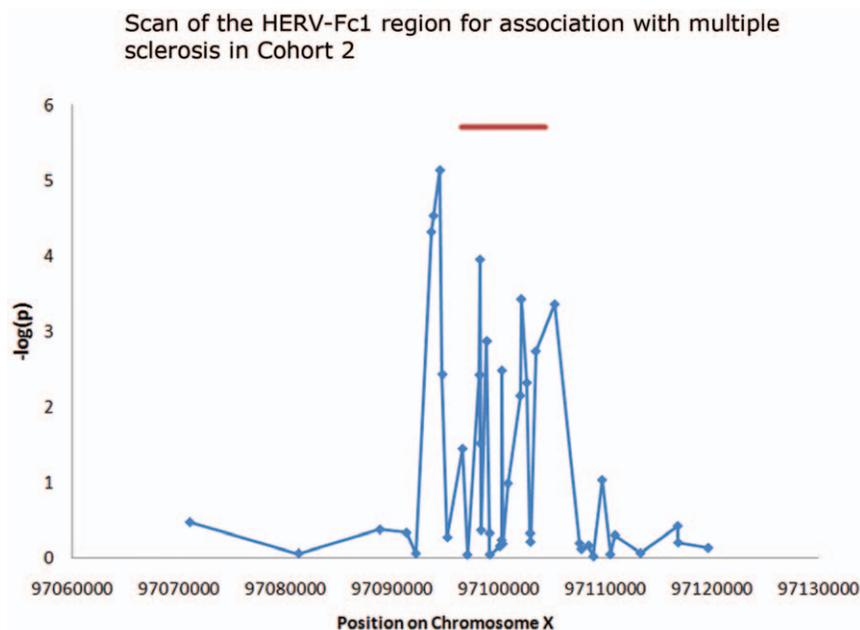


Figure 1. A scan of polymorphisms in the region around HERV-Fc1 for association with multiple sclerosis. The provirus stretches from 97096500 to 97104400 and is marked in red.

doi:10.1371/journal.pone.0016652.g001

Odds ratio for risk of MS as a function of two SNPs

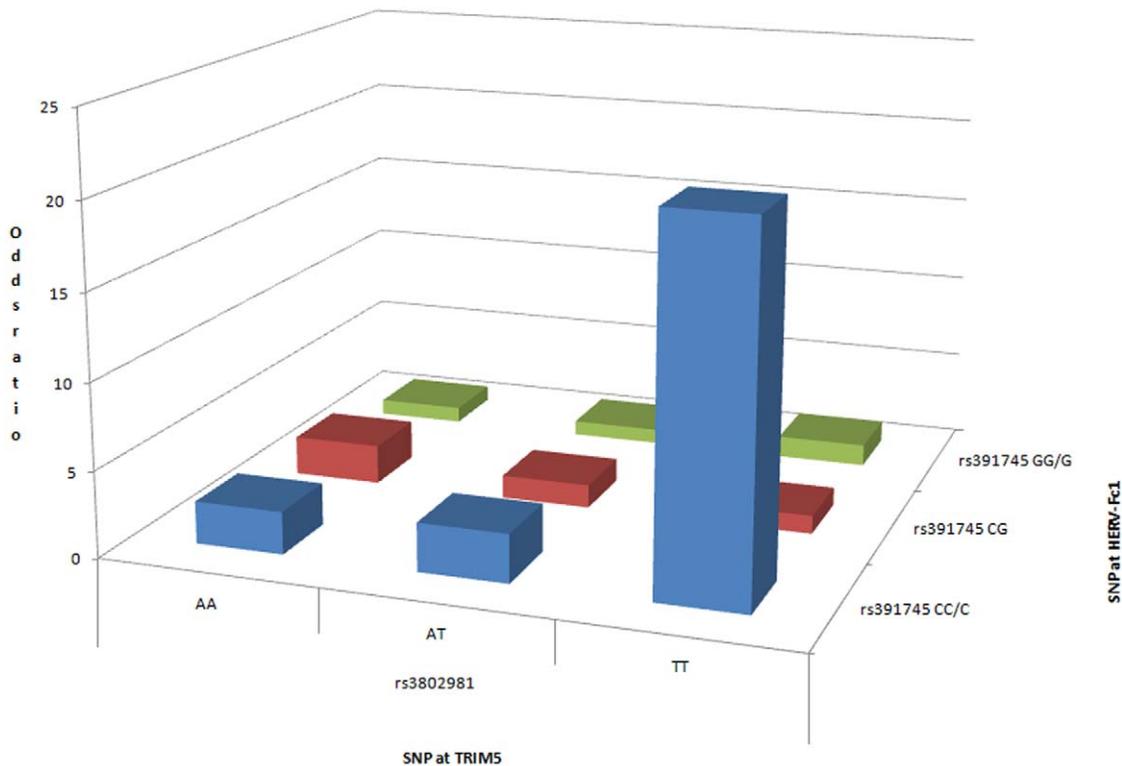


Figure 2. Odds ratio for MS in relation to 2 SNPs at HERV-Fc1 and *TRIM5*, respectively.
doi:10.1371/journal.pone.0016652.g002

the immune system and virus, with the virus mutating repeatedly to avoid the immune system, and the immune system repeatedly catching up [16]. One can see the episodic nature of multiple sclerosis as such a running battle. We also note that HERV-Fc1 is closely related to the HERV-H viruses previously described as activated in patients with multiple sclerosis [17].

Our suggestive evidence that a gene for multiple sclerosis is located on chromosome X could provide a natural explanation for the old observation that more women than men suffer from this disease. This is the normal pattern for X-dominant diseases. For Cohort 3 we obtained information about the clinical subtypes of the patients, but the cohort was too small for rigorous analyses of the subtypes. However, there were no signs of association with rs391745 among the patients with a primary progressive disease, and the association of the remaining (relapsing-remitting and secondarily progressing) was slightly stronger ($p = 0.008$) than that of the total cohort. This may indicate a different etiology of primary progressive sclerosis and is in accordance with the observation that the primary progressive form of multiple sclerosis has a sex ratio around 1.

In general, the finding that a disease is caused by an infectious agent is an encouraging one. These are the diseases, which we know best how to treat. If our results are confirmed, they point to new modes of treatment of Multiple Sclerosis.

Materials and Methods

The studies were approved by the Science Ethical Committee of Mid-Jutland and were performed with written and oral consent from all persons. Four cohorts of ethnically Danish patients, with verified multiple sclerosis [9], were identified together with

controls of similar ethnicity and gender from the same part of the country. A blood sample was drawn from each person and DNA was purified by conventional means. SNPs were tested using a Sequenom[®] platform (San Diego, California) and the iPLEX gold reaction. The SNPs, and the primers used, are listed in Table S1. Each PCR reaction contained 10 ng of genomic DNA, 0.5 U HotStart Taq from Qiagen (Hilden, Germany), 1.25×Enzyme Buffer (Qiagen), 3.5 mM MgCl₂, and 1 mM of each deoxynucleotide. The primers were added to a final concentration of 500 nM each. The total reaction volume was 4 μl. PCR reactions with primers of each plex were performed in 384 well plates with V-shape by performing a 15 min preheat to 94 C followed by cycling the samples 45 times, 94 C for 20 sec, 56 C for 30 sec and 72 C for 1 min, followed by 3 min at 72 C. The plates were stored at -20 C. Treatment with Shrimp Alkaline Phosphatase and extension with molecular weight-modified nucleotides was performed according to the manufacturer's recommendations. The reactions were cleaned with resin and spotted on SpectroChip arrays according to the manufacturer's recommendations. The reactions were analyzed by MALDITOF mass spectrometry on the Sequenom equipment, and the results were analyzed using the MassARRAY Typer 4.0 (Sequenom). Significance was tested by a χ^2 test (SPSS, IBM, Armonk, NY).

Supporting Information

Figure S1 A cartoon of the human genome with locations of the relevant endogenous viruses and indication of significant markers nearby.
(TIF)

Table S1 SNPs used for the analysis of *TRIM5* and their primers.
(DOC)

Table S2 SNPs used for the analysis of *TRIM5* and their association with MS.
(DOC)

Table S3 SNPs used for the analysis of selected endogenous retroviruses and the primers used.
(DOC)

Table S4 SNPs used for the analysis of selected endogenous retroviruses and their association with MS.
(DOC)

Table S5 Association of the alleles of rs391745 with MS in the sex-stratified cohorts.
(DOC)

Acknowledgments

BAN is indebted to Drs. Janice Clements, Opendra Narayan, Dianne Griffin, and Richard T. Johnson, Department of Neurology, Johns Hopkins Hospital, who taught him about Visna and other slow virus infections in the spring of 1979. Web resources: Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim/>.

Author Contributions

Conceived and designed the experiments: BAN FSP TC AML TP. Performed the experiments: BH KKN MJL TSP SB AH TW XW XZ. Analyzed the data: BAN PV TW XW XZ ALK ALN MN ADB. Contributed reagents/materials/analysis tools: TC JF AML ABO TB RM HBS FS KB HBR TP. Wrote the paper: BAN.

References

- Hansen T, Skytthe A, Stenager E, Petersen HC, Brønnum-Hansen H, et al. (2005) Concordance for multiple sclerosis in Danish twins: an update of a nationwide study. *Mult Scler* 11: 497–9.
- Gardner MB (1993) Genetic control of retroviral disease in aging wild mice. *Genetica* 91: 199–209.
- Pepin M, Vitu C, Russo P, Mormes JF, Peterhans E (1998) Maedi-visna virus infection in sheep: a review. *Vet Res* 29: 341–67.
- Zink MC, Clements JE (2002) A novel simian immunodeficiency virus model that provides insight into mechanisms of human immunodeficiency virus central nervous system disease. *J Neurovirol* 8 Suppl 2: 42–48.
- Bénit L, Calteau A, Heidmann T (2003) Characterization of the low-copy HERV-Fc family: evidence for recent integrations in primates of elements with coding envelope genes. *Virology* 312: 159–168.
- Tristem M (2000) Identification and characterization of novel human endogenous retroviruses families by phylogenetic screening of the human genome mapping project database. *J Virol* 74: 3715–3730.
- Pincus T, Rowe WP, Lilly F (1971) A major locus affecting resistance to infection with murine leukemia viruses. II. Apparent identity to a major locus described for resistance to friend murine leukemia virus. *J Exp Med* 133: 1234–41.
- Stremiau M, Owens CW, Perron MJ, Kiessling M, Autissier P, et al. (2004) The cytoplasmic body component TRIM5- α restricts HIV-1 infection in Old World monkeys. *Nature* 427: 848–853.
- McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, et al. (2001) Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 50: 121–7.
- Rowe WP (1972) Studies of genetic transmission of murine leukemia virus by AKR mice. I. crosses with FV-1 n strains of mice. *J Exp Med* 136: 1272–85.
- Chattopadhyay SK, Rowe WP, Teich NM, Lowy DR (1975) Definitive evidence that the murine C-type virus inducing locus Akv-1 is viral genetic material. *Proc Natl Acad Sci U S A* 72: 906–10.
- Villesen P, Aagard L, Wiuf C, Pedersen FS (2004) Identification of endogenous retroviral reading frames in the human genome. *Retrovirology* 1: 32.
- Nexø BA (1985) The biology of endogenous leukemia viruses: A study of the DBA/2 mouse in Gallo RC, Stehelin D, Varnier OE, eds. *Retroviruses and human pathology*. Clifton New Jersey, USA: Humana Press. pp 63–76.
- Lincoln MR, Montpetit A, Cader MZ, Saarela J, Dymont DA, et al. (2005) A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. *Nat Genet* 37: 1108–12.
- Gregory SG, Schmidt S, Seth P, Oksenberg JR, Hart J, et al. (2007) Interleukin 7 receptor α chain (IL7R) shows allelic and functional association with multiple sclerosis. *Nat Genet* 39: 1083–1091.
- Clements JE, Gdovin SL, Montelaro RC, Narayan O (1988) Antigenic variation in Lentiviral Diseases. *Ann Rev Immunol* 6: 139–159.
- Christensen T (2005) Association of human endogenous retroviruses with multiple sclerosis and possible interaction with herpes viruses. *Rev Med Virol* 15: 179–211.