Journal of NeuroVirology, 12: 411–421, 2006
© 2006 Journal of NeuroVirology
ISSN: 1355-0284 print / 1538-2443 online
DOI: 10.1080/13550280601009538

informa
healthcare

Sequence variation in the CC-chemokine ligand 2 promoter of pigtailed macaques is not associated with the incidence or severity of neuropathology in a simian immunodeficiency virus model of human immunodeficiency virus central nervous system disease

Edward K Wright Jr,[1,2] Janice E Clements,[2] and Sheila A. Barber[2]

[1]*McKusick-Nathans Institute of Genetic Medicine and* [2]*Department of Molecular and Comparative Pathobiology, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA*

Increased expression of CC-chemokine ligand 2 (CCL2) in the cerebrospinal fluid (CSF) and brain is consistently observed in human immunodeficiency virus (HIV) and simian immunodeficiency virus (SIV) central nervous system (CNS) disease. The molecular basis for the correlation between increased expression of CCL2 and HIV neuropathogenesis has been linked to a polymorphism at −2578 in the promoter of human CCL2, which was reported to influence the rate of progression to acquired immunodeficiency syndrome (AIDS) and the predisposition of HIV-infected individuals to develop HIV-associated dementia. However, because the rate of neurological deterioration essentially parallels the progression of immunosuppression, it is inherently difficult to uncouple the influence of this polymorphism on increased progression to AIDS from increased propensity to develop CNS complications. To further investigate the correlation between CCL2 and HIV/SIV CNS disease, the authors sequenced the CCL2 promoter of 29 pigtailed macaques examined in their accelerated and consistent SIV model in which all infected macaques develop AIDS but only 69% developed moderate/severe CNS lesions. Sequence analysis identified 39 sites of nucleotide variation in the pigtailed macaque CCL2 promoter/enhancer regions, with the resulting consensus sequence aligning with 94.7% homology to the human CCL2 promoter. After genetic analyses, no variation was found to correlate with the incidence or severity of CNS lesions or with levels of CCL2 in plasma or CSF. These findings suggest that the determinants of neuropathogenesis in this SIV model are distinct from variation in these regions of the CCL2 promoter. *Journal of NeuroVirology* (2006) **12,** 411–421.

Keywords: CCL2; *Macaca nemestrina*; MCP-1; Promoter; SNP

Introduction

Despite the success of highly active antiretroviral therapy in suppressing human immunodeficiency virus (HIV) replication and reconstituting immune function, the prevalence of HIV-associated dementia (HIV-D) has not declined and is actually on the rise (Dore *et al*, 1999; McArthur, 2004; Neuenburg *et al*, 2002; Sacktor *et al*, 2002). Perivascular macrophages and microglia (long-lived resident

Address correspondence to Sheila A. Barber, 1101 Wootton Parkway, 8th Floor, Rockville, MD 20852, USA. E-mail: sbarber1@jhmi.edu
The authors thank Drs. David J. Cutler and Patrick M. Tarwater for their helpful comments and discussions. They also thank the entire Retrovirus Laboratory at Johns Hopkins University School of Medicine for useful discussions. This research was supported by NIH grants to J.E.C. (NS050028, MH70306 and NS47984).
Received 5 July 2006; revised 14 August 2006; accepted 7 September 2006.

macrophages of the brain) are the primary sources of productive HIV replication in the central nervous system (CNS), which is thought to be infected initially by the infiltration of infected monocytes from the peripheral blood (Eugenin *et al*, 2006; McArthur *et al*, 2005). In addition to supporting productive HIV replication, these macrophages also generate many of the toxic products associated with HIV-D. In fact, the abundance of monocytes/macrophages in the brain may provide a better correlate for the severity of HIV-D than CNS viral load (Glass *et al*, 1995; Gonzalez-Scarano and Baltuch, 1999; Hou and Major, 2001; Kaul *et al*, 2001; Zink *et al*, 1999b).

CCL2 (CC-chemokine ligand 2; formerly referred to as monocyte chemoattractant protein-1; MCP-1), which is secreted by macrophages, endothelial cells, and astrocytes in the CNS, is a critical chemoattractant that mediates the infiltration of monocyte/macrophages into the brain (Fuentes *et al*, 1995). CCL2 is present in the cerebrospinal fluid (CSF) of patients with HIV-D and levels of CCL2 correlate with the degree of neurocognitive impairment (Bernasconi *et al*, 1996; Cinque *et al*, 1998; Conant *et al*, 1998; Gonzalez-Scarano and Baltuch, 1999; Hou and Major, 2001; Kaul *et al*, 2001; Kelder *et al*, 1998; Mengozzi *et al*, 1999; Persidsky *et al*, 1999; Weiss *et al*, 1997, 1999; Wu *et al*, 2000; Zink *et al*, 1999b). Moreover, longitudinal analyses of HIV patients indicate that CCL2 levels in the CSF increase prior to the development of clinical neurological disease and tend to associate with the time to onset of HIV-D (Kelder *et al*, 1998). Intriguingly, recent *in vitro* studies have demonstrated that HIV-infected monocytes are preferentially (compared to uninfected monocytes) induced by CCL2 (but not other chemokines) to transmigrate across and disrupt an intact blood brain barrier (Eugenin *et al*, 2006).

A definitive link between increased expression of CCL2 and HIV neuropathogenesis was suggested first by a study in which a polymorphism at −2578 (alternatively designated −2518 relative to the transcriptional start site) in the distal regulatory region of the human CCL2 promoter was found to influence the predisposition of HIV-infected individuals to develop HIV-D (Gonzalez *et al*, 2002). Specifically, adults homozygous for a guanine at −2578 were found to resist initial infection with HIV, but once infected these individuals exhibited accelerated progression to AIDS and increased risk of developing HIV-D. In this study, serum levels of CCL2 were significantly higher in individuals homozygous or heterozygous for the −2578 G allele, demonstrating a link between the polymorphism and peripheral CCL2 expression, but no measurements of CCL2 levels in CSF were provided. In subsequent studies, CCL2 levels in the CSF of HIV-infected individuals supported a gene-dosage effect for the −2578 G allele, with homozygous individuals expressing higher levels of CCL2 in CSF than heterozygotes, which expressed higher levels than individuals homozygous for the ancestral −2578 A allele (Letendre *et al*, 2004). The collective data from these two studies implicated the −2578 G allele as a predisposing determinant for increased expression of CCL2 in the CNS of HIV-infected individuals and the development of HIV-D.

An inherent complexity in the interpretation of these studies in HIV-infected individuals is the consistent observation that HIV-D rarely develops prior to profound immunosuppression (McArthur *et al*, 2005). As a result, it is difficult to unequivocally uncouple the influence of this polymorphism on increased progression to AIDS from increased propensity to develop CNS complications. To this end, we have subverted this complexity by studying our accelerated and consistent simian immunodeficiency virus (SIV) model of HIV CNS disease that closely recapitulates HIV CNS infection: encephalitis with active virus replication in the CNS, characteristic histopathological changes, psychomotor impairment, and neurodegeneration (Murray *et al*, 1992; Sharer *et al*, 1988; Zink and Clements, 2002; Zink *et al*, 1999a). Importantly for our studies, all infected macaques develop acquired immunodeficiency syndrome (AIDS) yet only 69% develop moderate/severe CNS disease by 84 days post inoculation (Mankowski *et al*, 2004). In addition, increased levels of CCL2 in the CSF precede and predict the development of SIV encephalitis: pigtailed macaques (*Macaca nemestrina*) that develop moderate/severe SIV encephalitis express significantly higher CSF CCL2 levels than pigtailed macaques that develop mild/no encephalitis (Mankowski *et al*, 2004; Zink *et al*, 2001). In short, 69% of infected macaques develop AIDS, and exhibit increased CSF CCL2 and moderate/severe CNS lesions, whereas 31% of infected macaques develop AIDS, do not express increased CSF CCL2, and exhibit only mild/no CNS lesions.

In the present study, we sequenced the distal and proximal regulatory regions of the CCL2 gene obtained from 29 pigtailed macaques examined in our SIV model in an attempt to establish a definitive association between sequence variation in the CCL2 promoter and the incidence and severity of SIV CNS disease. Although all macaques expressed the ancestral −2578 A/A genotype, we identified 39 sites of variation in their CCL2 promoters. After rigorous statistical analysis, however, no single variation was found to associate with the incidence and severity of SIV CNS disease or with plasma or CSF CCL2 levels throughout disease progression. We conclude that any genetic basis for the consistent development of moderate/severe CNS lesions in pigtailed macaques inoculated in this SIV model must be distinct from the −2578 site as well as any of the 39 sites of variation.

Results

High sequence conservation at binding sites in the CCL2 promoter

The nucleotide sequence of the pigtailed macaque CCL2 promoter (distal and proximal regions) is 94.7% homologous to the human sequence (Figure S1). Only one of the known regulatory elements in the human CCL2 promoter, a TRE element in the proximal region, is disrupted by base substitutions in macaques (Figure 1). Single base substitutions were found in the GC box and the TATA box as compared to the human sequence; however, the macaque sequence is identical to the mouse sequence at these bases. Thus, these substitutions appear to have occurred in the chimpanzee/human lineage after separation from pigtailed macaques. The sequence identity of the GC box and the TATA box between mouse and pigtailed macaques implies functional conservation at these binding sites. The two distal nuclear factor (NF)-κB elements, the proximal GAS element, and one of the proximal TRE elements are identical when comparing human, chimpanzee, pigtailed macaque, and mouse sequences. Given the compelling link between CCL2 expression and HIV/SIV neuropathogenesis, the high conservation of transcription factor binding sites revealed in our study strongly suggests that the pigtailed macaque CCL2 promoter is subject to the same regulatory mechanisms during SIV infection as the human CCL2 promoter during HIV infection.

Sequence variation in the CCL2 promoter

All pigtailed macaque distal promoters possessed the ancestral adenine nucleotide at the site of the −2578 guanine polymorphism in the human population (Figure S1). Therefore, sequence variation at −2578 does not influence the development of SIV CNS disease in this macaque model. However, we identified 39 sites of variation in 1456 bp of the macaque CCL2 distal and proximal promoter regions that have not been described to date in the human population. We conclude from this finding that pigtailed macaques are in fact highly variable in these

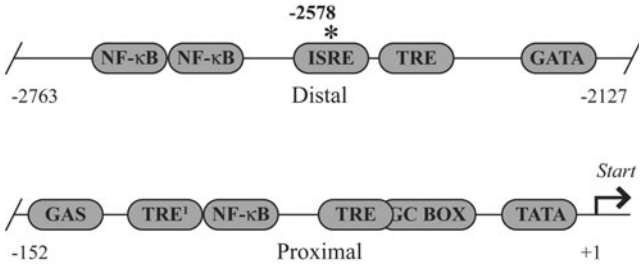

Figure 1 The diagram of the human CCL2 distal and proximal promoter regions is illustrating the relative position of transcription factor binding elements and the translational start site. *Position of the −2578 human polymorphisms. [1]The TRE element that is disrupted by base substitutions in pigtailed macaques.

CCL2 promoter regions, an observation consistent with previous reports suggesting that macaques exhibit higher genomic variation than humans (Tosi *et al*, 2003). Twelve sites of variation were found in only 1 of the 29 pigtailed macaques and another 7 sites were found in fewer than 4 macaques and were not informative in the statistical analyses. The remaining 20 sites of variation were found in greater than four macaques, thus they were informative and statistical analysis was performed. Other than two sites of variation that were found in the proximal NF-κB binding element of only two macaques (noninformative in statistical tests), no variation mapped to any known transcription factor binding site. Neither variation in the proximal NF-κB site could be used in the statistical calculations because neither variable allele was present at a high enough frequency, which argues that variation in these sites does not account for the reproducible differences in severity of SIV CNS disease in this model.

Increased expression of CCL2 correlates with severity of SIV CNS disease

Previous reports from our laboratory have established a correlation between increased expression of CCL2 in the CSF and the severity of CNS lesions in 18 macaques inoculated in our SIV model (Mankowski *et al*, 2004; Zink *et al*, 2001). Here we provide a comprehensive analysis of CCL2 levels and neuropathology encompassing all 29 macaques infected to date in the context of this model. Of the 29 macaques, 9 (31%) exhibited no/mild CNS lesions at necropsy whereas the remaining 20 (69%) exhibited severe/moderate CNS lesions. Although the previous studies (n = 18) reported that CSF CCL2 levels and ratios of CSF:plasma CCL2 levels predict and precede CNS disease severity with significant differences between macaques that develop severe/moderate versus mild/no CNS disease detectable at 28 days post inoculation, increasing the number of macaques to 29 in this study provided more statistical power and revealed that median CSF CCL2 levels are significantly different in macaques that develop severe/moderate (583.9 pg/mL) versus mild/no (213.0 pg/mL) CNS disease as early as 14 days post inoculation ($P = .0081$, after Bonferroni correction) and remain significantly different throughout the course of disease (Figure 2A). The median ratios of CSF:plasma CCL2 levels in this comprehensive analysis are statistically different in macaques that develop moderate/severe versus mild/no CNS lesions at day 56 post inoculation (severe/moderate 7.3, mild/none 2.6, $P = .0072$ after Bonferroni correction; Figure 2C), which leads us to conclude that CSF CCL2 levels are the better predictors of disease with the larger population studied and increasing variation taken into account. Plasma levels of CCL2 were not significantly different in macaques that develop severe/moderate versus mild/no CNS disease, although macaques that developed severe/moderate

lesions tended to have higher median plasma CCL2 levels throughout the course of disease as compared to macaques that developed mild/no CNS lesions, consistent with a more hyper-responsive inflammatory phenotype (Figure 2B).

Sequence variation in the CCL2 promoter does not correlate with severity of SIV CNS disease

Only one site, −2737 (which is not located in any recognized transcription factor binding site), had a significant association with moderate/severe neuropathology by the Fisher's exact test ($P = .0464$; Table 1); however, correction for multiple comparisons (Bonferroni correction, $n = 19$) completely abolished this significance ($P = .882$). Therefore, no allele significantly associated with either the moderate/severe or mild/none groups. Four variable sites (not located in any recognized transcription factor binding site) correlated with CCL2 levels at various stages of disease progression, revealing a total of five significant ($P < .05$) Mann-Whitney test P values (Table 2). Again, correction for multiple comparisons (Bonferroni correction, $n = 228$) eliminated the significance. Therefore, no allele correlated with the level of CCL2 in the CSF or plasma at any stage of disease. Taken together, both tests indicate that no site of variation associates with incidence or severity of SIV CNS disease or with CCL2 levels in plasma or CSF in this model, in keeping with our observations that only 2 of the 39 sites of variation occurred in a known transcription factor binding site.

Table 1 Fisher's exact test: uncorrected P values

	Severe/Moderate No. of Alleles		Mild/None No. of Alleles		
Position	A[1]	a[2]	A	a	2-Tail P value
−2790	33	7	17	1	0.4135
−2768	38	2	16	2	0.5808
−2737	28	12	17	1	0.0464
−2729	32	8	12	6	0.3273
−2702	28	12	12	6	1.0000
−2661	37	3	16	2	0.6410
−2545	38	2	16	2	0.5808
−2425	29	11	14	4	0.7559
−2321	37	3	16	2	0.6410
−2291	38	2	16	2	0.5808
−2251	33	7	18	0	0.0868
−2235	27	13	12	6	1.0000
−2180	32	8	17	1	0.2489
−2125	22	18	10	8	1.0000
−2025	28	12	11	7	0.5543
−1982	21	19	9	9	1.0000
−1939	32	8	15	3	1.0000
−144	35	5	14	4	0.4380
−79	31	9	11	7	0.2189

Note. Variation at sites −2702 and −2653 is linked; therefore, one statistical test is used for both sites, and only −2702 is shown. The Fisher's exact test was calculated for sites that had x > 4 of the least common nucleotide variant. There are no significant p values after correction for multiple comparisons (Bonferroni correction, $n = 19$) [1]A represents the allele at higher frequency in the population. [2]a represents the allele at lower frequency in the population.

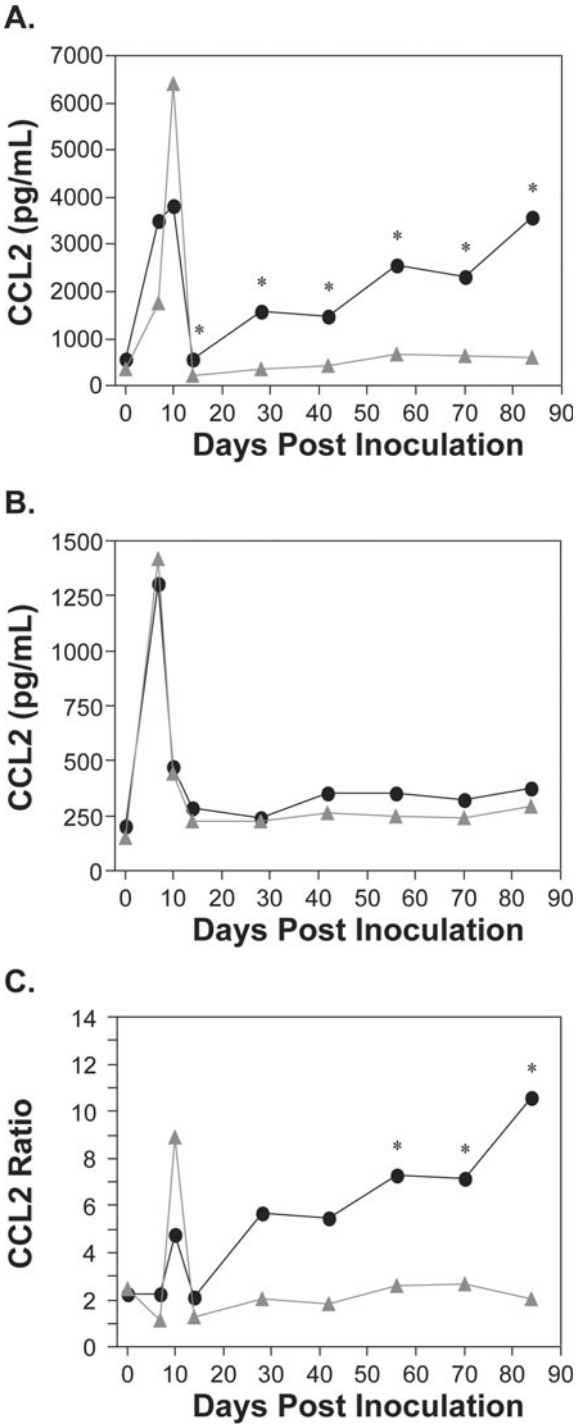

Figure 2 Medians of **(A)** CSF CCL2 levels, **(B)** plasma CCL2 levels, and **(C)** ratios of CSF:plasma CCL2 levels are graphed versus time after inoculation. CSF CCL2 levels are significantly higher in macaques with moderate/severe encephalitis *(black circles)* compared to macaques with mild/no encephalitis (*grey triangles*) starting at day 14 p.i. and continuing throughout disease. There is no significant difference in plasma levels between macaques with moderate/severe encephalitis compared to mild/no encephalitis. The ratio of CSF:plasma CCL2 level is significantly higher in macaques with moderate/severe encphalietis beginning at day 56 p.i. and continuing throughout disease. *Significant difference between macaques that develop severe/moderate and mild/no CNS disease (Mann-Whitney).

Table 2 Pigtailed macaque CCL2 promoter polymorphism–protein level correlation: Mann-Whitney test uncorrected p values

	Preinoculation			Acute			Peak			Necropsy		
Position	CSF	Plasma	CSF/Plasma	CSF	Plasma	CSF/Plasma	CSF	Plasma	CSF/Plasma	CSF	Plasma	CSF/Plasma
−2790	0.421	0.402	0.819	0.317	0.864	0.252	0.788	0.788	0.643	0.691	0.629	0.637
−2768	0.924	0.793	0.694	0.975	0.467	0.635	0.776	0.924	0.393	0.811	0.533	0.337
−2737	0.877	0.888	0.638	0.465	0.073	0.877	0.413	0.298	0.642	0.132	0.099	0.177
−2729	0.364	0.466	0.335	0.947	0.319	0.774	0.493	0.580	0.774	0.864	0.341	0.554
−2702	0.525	0.578	0.889	0.083	0.645	0.283	0.469	0.776	0.844	0.423	0.927	0.878
−2661	0.583	0.322	0.976	0.931	0.260	0.795	0.708	0.977	0.977	0.827	0.337	0.558
−2545	0.924	0.793	0.694	0.975	0.467	0.635	0.776	0.924	0.393	0.811	0.533	0.337
−2425	0.946	0.760	0.466	0.702	0.875	0.291	0.840	0.805	0.875	0.939	0.510	0.914
−2321	0.544	0.675	0.230	0.100	0.624	0.030	0.931	0.795	0.751	0.779	1.000	0.845
−2291	0.591	0.201	0.076	0.728	0.242	0.874	0.874	0.681	0.825	0.433	0.603	0.149
−2251	0.572	0.055	0.230	0.809	0.006	0.319	0.572	0.979	0.809	0.143	0.933	0.119
−2235	0.088	0.290	0.724	0.036	0.550	0.080	0.877	0.413	0.521	0.643	0.594	0.777
−2180	0.109	0.801	0.675	0.373	0.079	0.980	0.939	0.558	0.558	0.398	1.000	0.287
−2125	0.465	0.731	0.099	0.981	0.981	0.981	0.465	0.621	0.465	0.269	0.525	0.055
−2025	0.878	0.765	0.629	0.417	0.196	0.283	0.913	0.776	0.776	0.675	0.197	0.120
−1982	0.164	0.525	0.321	0.010	0.826	0.018	0.196	0.164	0.137	0.193	0.374	0.232
−1939	0.492	0.314	0.168	0.628	0.373	0.702	0.702	0.460	0.524	0.779	0.396	0.696
−144	0.083	0.839	0.072	0.484	0.976	0.879	0.447	0.484	0.447	0.527	0.775	0.172
−79	0.062	0.506	0.360	0.406	0.299	0.836	0.678	0.917	0.795	0.817	0.763	0.689

Note. Variation at sites −2702 and −2653 is linked; therefore, one statistical test is used for both sites, and only −2702 is shown. There are no significant p values after correction for multiple comparisons (Bonferroni correction, $n = 228$).

Discussion

Because of their close evolutionary relationship to the *Homo sapiens* species, macaques are utilized to study many human diseases (Sibal and Samson, 2001). Because SIV has proven to be the ancestor of HIV, it comes as no surprise that SIV-infected macaques are considered the best animal models to study HIV infection in humans, capitalizing on the close evolutionary relationships between both the viruses and hosts. As with any animal model, a major goal is to closely recapitulate human disease to a degree where the likelihood and confidence that prevention or treatment discoveries will translate effectively to humans. In the present work, we have studied our accelerated and consistent SIV/macaque model of HIV CNS disease. The pigtailed macaques used in this model exhibit the same compelling link between CCL2 and SIV CNS disease as described for CCL2 and HIV CNS disease. Like HIV, increased levels of CCL2 in CSF precede and predict the severity of CNS disease, and only a subset of SIV-infected pigtailed macaques exhibit this increase in CSF CCL2 and develop marked CNS lesions. However, despite these strong similarities, the −2578 polymorphism linked to HIV CNS disease does not even exist in the 29 pigtailed macaques that initially established the link between CCL2 and SIV CNS disease. Furthermore, no other variation in the CCL2 enhancer and promoter regions provided a molecular basis for the CCL2 connection. Nevertheless, several questions have emerged from these observations that warrant discussion.

First, is there a genetic basis for the propensity to develop SIV CNS disease? It is conceivable that there is; however, the molecular mechanism (e.g., SNP, single nucleotide polymorphism) must be distinct from the −2578 site in pigtailed macaques. Interestingly, a study of over 3000 individuals enrolled in five natural cohorts in the United States failed to confirm the association between −2578 and HIV transmission that was reported in the initial study linking CCL2 to HIV neuropathogenesis, but did associate one of seven identified haplotypes encompassing the MCP-1–MCP-3 and eotaxin gene cluster with susceptibility to HIV infection (Modi *et al*, 2003). Although no assessment was made with regard to HIV-D in this population, the influential SNPs in the implicated haplotype (which coincidentally contained the ancestral "A" at −2578) were located in the CCL2 promoter, intron 1 of CCL2 and the eotaxin promoter. Furthermore, although a single haplotype was found to correlate with HIV transmission, some SNPs in that haplotype were also found in other haplotypes. Thus, it is entirely possible that such a multigenic haplotype exists in pigtailed macaques involving sequence variations in the CCL2 gene encompassing or distinct from the enhancer/promoter regions examined in our study. Candidate distinct sites within the CCL2 gene would include the poorly characterized far upstream region (−3.6 to −2.7 kb) of the CCL2 enhancer, CCL2 introns, UTR sequences, or other yet unidentified regulatory regions (Sekine *et al*, 2002; Wagner *et al*, 2001). Because increased expression of CCL2 correlates with the severity of HIV/SIV CNS disease, such variation would undoubtedly affect the expression of CCL2 rather than its functionality. It is also conceivable that independent sequence variation exist in the coding regions of transcription factors that regulate expression of CCL2 and/or components of the signaling pathways that activate those factors. Additional rigorous sequencing efforts will

be required to address these possibilities, which are beyond the scope of this study.

Second, how might the results from this study impact future experimentation in our accelerated SIV model? Perhaps the most extraordinary observation in this study is the ability to predict with statistically significant accuracy which macaques will develop CNS disease by 84 days post infection (p.i.) based on CSF CCL2 levels at 14 days p.i. Therefore, when evaluating the efficacy of intervention therapies for CNS disease in SIV-infected macaques following acute infection, CSF CCL2 levels at 14 days p.i. could be quantitated to ensure that all macaques studied would ostensibly develop CNS disease in our accelerated SIV model. In this way, the risk that by sheer random selection the macaques studied would not have developed SIV CNS disease in the first place (i.e., in the absence of treatment) would be minimized. The benefit of extrapolation of such an early biomarker to predict HIV CNS disease becomes immediately apparent; unfortunately, although CCL2 has been shown to be a good biomarker for HIV CNS disease at end stage, it would be doubtful that CCL2 could be used as a biomarker as early in HIV infection as in our accelerated model due to the extreme differences between the normal course of HIV infection and the course of infection in our accelerated SIV model, which has only a small decrease of peripheral viral load (no set point) and macaques rapidly progress to terminal disease essentially skipping the asymptomatic phase of AIDS. Moreover, the variably chronic course of disease and insidious onset of neurological symptoms in addition to confounding factors not present in our SIV model (e.g., nonspecific pathogen–free environment, secondary infections, and intravenous drug use) likely preclude the development of a temporal template for early predictive CSF CCL2 levels that models most HIV patients. However, monitoring CCL2 levels of an HIV-infected individual may provide an early indicator of ensuing CNS complications, akin to monitoring peripheral blood CD4 counts to detect the onset of AIDS, because in this accelerated SIV model elevated CCL2 levels persist from post acute levels through end-stage disease.

Third, what is the significance of the number of variable sites in the CCL2 promoter/enhancer of pigtailed macaques, the evolution of the G polymorphism in humans and the −2578 site itself in the regulation of CCL2 in humans? An abundance of sequence variation in pigtailed macaques relative to the human population suggests that pigtailed macaques have a larger effective population size (Nei, 1987). The lack of the −2578 G polymorphism in the CCL2 promoter of pigtailed macaques and chimpanzees (Gonzalez *et al*, 2002) suggests that this specific mutation occurred after the human-chimpanzee divergence in evolution. The maintenance of this regulatory polymorphism in the human population that reaches an allelic frequency of up to 60.2% in some populations (Hong *et al*, 2006) suggests either that a G at −2578 exerts a selective advantage compared to the A allele or that the G is passive yet serendipitously survived neutral evolution. A quick search of the literature reveals no fewer than 15 human diseases currently associated with −2578 G: asthma, coronary artery disease, rheumatoid arthritis, type 1 diabetes, type 2 diabetes, pulmonary tuberculosis, HLA-B27–associated acute anterior uveitis, breast cancer, systemic sclerosis, lupus, Alzheimer's disease, hepatitis C virus (HCV)-related liver disease, and HIV-associated dementia (Aguilar *et al*, 2001; Flores-Villanueva *et al*, 2005; Ghilardi *et al*, 2005; Gonzalez *et al*, 2002; Gonzalez-Escribano *et al*, 2003; Karrer *et al*, 2005; Kim *et al*, 2002; Muhlbauer *et al*, 2003; Pola *et al*, 2004; Simeoni *et al*, 2004; Szalai *et al*, 2001a, 2001b; Wegscheider *et al*, 2005; Yang *et al*, 2004). In contrast, only a single study has associated the −2578 A polymorphism with a disease (Mori *et al*, 2005). Therefore, on balance, the G polymorphism is apparently more unfavorable, and in theory should have been selected against. However, a polymorphism that contributes to disease pathogenesis without inducting prepubescent mortality or impairing reproductive potential may not be selected against.

Other than the controversial association between the −2578 polymorphism and HIV pathogenesis, no molecular mechanism has emerged to link the −2578 G to CCL2 expression. The −2578 site lies within an interferon stimulated response element (ISRE), which other than the −2578 G polymorphism, is 100% conserved between macaques and humans. In the initial report (Gonzalez *et al*, 2002), the transcriptional activator IRF-1 was found to bind less efficiently to the G allele *in vitro* than the A allele in the context of an Electrophoretic Mobility Shift Assay (EMSA) (although neither the G nor the A make a better consensus for an Interferon Stimulated Response Element (ISRE)). However, this finding appears counterintuitive to the association of the G allele with increased expression of CCL2. One other study reported a protein complex preferentially bound in EMSA to the G allele not the A allele, which, if the complex activates transcription, might account for the hyperresponsive G phenotype (Muhlbauer *et al*, 2003). To date, no definitive molecular mechanism has been proposed.

Finally, should the results from this study be interpreted to conclude that intervention treatments found to be effective in macaques at suppressing CCL2 and CNS disease will not in all likelihood translate to HIV patients? As far as transcriptional regulation is concerned (one potential target of intervention therapy), this study has demonstrated impressive conservation between macaques and humans at the CCL2 enhancer and promoter regions. The single difference in consensus binding sites localizes to a TRE element in the proximal promoter that is disrupted in macaques, but constitutively occupied in humans (Zhou *et al*, 1998). Disruption of the human TRE element has no affect on basal transcription, and

when stimulated with interferon (IFN)-γ results in only a nominal reduction (Zhou *et al*, 1998). Thus, at best, macaques would be expected to express lower levels of CCL2 than humans, yet both HIV and SIV CNS diseases are associated with high levels of CCL2. Therefore, there is no reason *a priori* to conclude that effective treatments in macaques will not translate to humans. In this regard, we have recently reported that the antibiotic minocycline suppresses the incidence and severity of SIV CNS disease in our rigorous macaque model, which correlated with suppression of many markers of CNS disease including suppressed expression of CCL2 in the CNS (Zink *et al*, 2005). Researchers at Johns Hopkins University School of Medicine are now recruiting HIV-infected patients on highly active antiretroviral therapy (HAART) with neurological symptoms for clinical trials to test the efficacy of minocycline against HIV CNS disease.

Materials and methods

Animal tissues
For sequencing: DNA was extracted from small punches of brain tissue that were snap-frozen at necropsy and stored at −80°C.

For quantitation of CCL2: Blood (plasma) and CSF from 29 pigtailed macaques were sampled prior to and 7, 10, 14, 28, 42, 56, and 70 days after inoculation with SIV/DeltaB670 (50 animal infectious dose $[AID]_{50}$) and SIV/17E-Fr (10,000 AID_{50}) as well as at necropsy as described previously (Zink *et al*, 2001).

Neuropathological assessment: Brain sections were examined microscopically and scored as no, mild, moderate, or severe pathology as described (Zink *et al*, 1999a).

DNA extraction
Tissue samples (50 mg) were transferred to FastDNA tubes (Bio101) containing 300 μL ATL buffer from the Qiagen Dneasy Tissue Kit and disrupted in the FastPrep machine twice for 15 s. Samples were centrifuged briefly at 14,000 rpm at room temperature, and DNA was extracted from the supernatant using the Qiagen Dneasy tissue kit according to the manufacturer's instructions.

Polymerase chain reaction (PCR) and sequencing
CCL2 distal and proximal promoter regions were PCR-amplified from genomic DNA. CCL2 regions were amplified in 25-μL reactions with Epicentre's Fail-Safe PCR system. The 927-bp CCL2 enhancer region was PCR-amplified using the forward primer 5′-CCGAGATGTTCCCAGCACAG-3′ and the reverse primer 5′-CTGCTTTGCTTGTGCCTCTT-3′ (Rovin, 1999), with the following thermal profile: 35 cycles of 94°C (30 s.), 57°C (30 s.), and 72°C (1 min and 30 s.), with an initial denaturation step of 94°C (5 min). The 529-bp CCL2 promoter region was PCR-amplified using the above thermal profile, with forward primer 5′-TAATGCATTGTCAGGGAGCC-3′ and reverse primer 5′-GCAGAGACTTTCATGCTGGA-3′, but with an extension time of 30 s. PCR reactions were resolved on a 1.2% agarose gel, and bands were gel-purified (Invitrogen) and sequenced on an Applied Biosystems 3730 DNA analyzer. Sequencing was performed using the primers listed above and the following internal primers for complete double coverage of the CCL2 distal promoter region: 5′-GGGTGTGAATCAGAAAAGAAAGTC-3′ and 5′-CTCTGTGACCACGGCCTAAT-3′. Individual pigtailed macaque promoter regions were sequenced two times from independent PCR reactions.

Alignment of sequences
Consensus sequences were established for each animal using Sequencher, version 4.6. Alignments of all pigtailed consensus CCL2 regions were completed, and variation was determined from these alignments.

Enzyme-linked immunosorbent assay (ELISA)
CCL2 protein in plasma and CSF was quantitated by ELISA (R&D Systems, Minneapolis, MN).

Data analysis
Comparison of CCL2 levels and severity of neurological disease: Animals were separated into two groups based on the degree of neuropathology at the time of sacrifice: mild/no CNS disease ($n = 9$) and severe/moderate CNS disease ($n = 20$). Median CSF and plasma CCL2 levels were determined for each group at every time point relative to SIV inoculation and compared using the Mann-Whitney test.

Comparison of sequence variation in the CCL2 promoter and severity of neurological disease: Alleles were summed at variable sites, categorized by the severity of neuropathology and subjected to the Fisher's two-tailed exact test.

Comparison of sequence variation in the CCL2 promoter and CCL2 levels: Differences in CSF and plasma CCL2 protein levels, as well as ratios of CSF:plasma CCL2 levels, between pigtailed macaques grouped according to nucleotide variant were evaluated using the Mann-Whitney test. The most common homozygous variants were tested against the heterozygous and least common homozygous variants at four stages of disease (preinoculation, acute infection, the peak protein level achieved after acute infection, and at necropsy). The Mann-Whitney test could not be calculated for nucleotides that had an insufficient number of variable alleles at that position ($x < 4$ alleles). All P values were adjusted using the Bonferroni correction for multiple comparison testing.

Accession numbers

Individual macaque consensus sequences were deposited in GenBank, accession numbers: DQ514923–DQ514939, DQ514941–DQ514943, DQ514950, DQ514951, DQ514957–DQ514959, DQ514992–DQ515007, DQ515009–DQ515011, DQ515018, DQ515019, DQ515026–DQ515029, and DQ826454–DQ826461.

References

Aguilar F, Gonzalez-Escribano MF, Sanchez-Roman J, Nunez-Roldan A (2001). MCP-1 promoterpolymorphism in Spanish patients with systemic lupus erythematosus. *Tissue Antigens* **58:** 335–338.

Bernasconi S, Cinque P, Peri G, Sozzani S, Crociati A, Torri W, Vicenzi E, Vago L, Lazzarin A, Poli G, Mantovani A (1996). Selective elevation of monocyte chemotactic protein-1 in the cerebrospinal fluid of AIDS patients with cytomegalovirus encephalitis. *J Infect Dis* **174:** 1098–1101.

Cinque P, Vago L, Mengozzi M, Torri V, Ceresa D, Vicenzi E, Transidico P, Vagani A, Sozzani S, Mantovani A, Lazzarin A, Poli G (1998). Elevated cerebrospinal fluid levels of monocyte chemotactic protein-1 correlate with HIV-1 encephalitis and local viral replication. *AIDS* **12:** 1327–1332.

Conant K, Garzino-Demo A, Nath A, McArthur JC, Halliday W, Power C, Gallo RC, Major EO (1998). Induction of monocyte chemoattractant protein-1 in HIV-1 Tat-stimulated astrocytes and elevation in AIDS dementia. *Proc Natl Acad Sci U S A* **95:** 3117–3121.

Dore GJ, Correll PK, Li Y, Kaldor JM, Cooper DA, Brew BJ (1999). Changes to AIDS dementia complex in the era of highly active antiretroviral therapy. *AIDS* **13:** 1249–1253.

Eugenin EA, Osiecki K, Lopez L, Goldstein H, Calderon TM, Berman JW (2006). CCL2/monocyte chemoattractant protein-1 mediates enhanced transmigration of human immunodeficiency virus (HIV)-infected leukocytes across the blood-brain barrier: a potential mechanism of HIV-CNS invasion and NeuroAIDS. *J Neurosci* **26:** 1098–1106.

Flores-Villanueva PO, Ruiz-Morales JA, Song CH, Flores LM, Jo EK, Montano M, Barnes PF, Selman M, Granados J (2005). A functional promoter polymorphism in monocyte chemoattractant protein-1 is associated with increased susceptibility to pulmonary tuberculosis. *J Exp Med* **202:** 1649–1658.

Fuentes ME, Durham SK, Swerdel MR, Lewin AC, Barton DS, Megill JR, Bravo R, Lira SA (1995). Controlled recruitment of monocytes and macrophages to specific organs through transgenic expression of monocyte chemoattractant protein-1. *J Immunol* **155:** 5769–5776.

Ghilardi G, Biondi ML, La Torre A, Battaglioli L, Scorza R (2005). Breast cancer progression and host polymorphisms in the chemokine system: role of the macrophage chemoattractant protein-1 (MCP-1) −2518 G allele. *Clin Chem* **51:** 452–455.

Glass JD, Fedor H, Wesselingh SL, McArthur JC (1995). Immunocytochemical quantitation of human immunodeficiency virus in the brain: correlations with dementia. *Ann Neurol* **38:** 755–762.

Gonzalez E, Rovin BH, Sen L, Cooke G, Dhanda R, Mummidi S, Kulkarni H, Bamshad MJ, Telles V, Anderson SA, Walter EA, Stephan KT, Deucher M, Mangano A, Bologna R, Ahuja SS, Dolan MJ, Ahuja SK (2002). HIV-1 infection and AIDS dementia are influenced by a mutant MCP-1 allele linked to increased monocyte infiltration of tissues and MCP-1 levels. *Proc Natl Acad Sci U S A* **99:** 13795–13800.

Gonzalez-Escribano MF, Torres B, Aguilar F, Rodriguez R, Garcia A, Valenzuela A, Nunez-Roldan A (2003). MCP-1 promoter polymorphism in Spanish patients with rheumatoid arthritis. *Hum Immunol* **64:** 741–744.

Gonzalez-Scarano F, Baltuch G (1999). Microglia as mediators of inflammatory and degenerative diseases. *Annu Rev Neurosci* **22:** 219–240.

Hong SB, Jin SY, Park HJ, Jung JH, Sim WY (2006). Analysis of the monocyte chemoattractant protein 1 −2518 promoter polymorphism in Korean patients with alopecia areata. *J Korean Med Sci* **21:** 90–94.

Hou J, Major EO (2001). Direct and indirect mechanisms of HIV-1 neuropathogenesis in the human central nervous system. *Adv Exp Med Biol* **493:** 29–34.

Karrer S, Bosserhoff AK, Weiderer P, Distler O, Landthaler M, Szeimies RM, Muller-Ladner U, Scholmerich J, Hellerbrand C (2005). The −2518 promotor polymorphism in the MCP-1 gene is associated with systemic sclerosis. *J Invest Dermatol* **124:** 92–98.

Kaul M, Garden GA, Lipton SA (2001). Pathways to neuronal injury and apoptosis in HIV-associated dementia. *Nature* **410:** 988–994.

Kelder W, McArthur JC, Nance-Sproson T, McClernon D, Griffin DE (1998). Beta-chemokines MCP-1 and RANTES are selectively increased in cerebrospinal fluid of patients with human immunodeficiency virus-associated dementia. *Ann Neurol* **44:** 831–835.

Kim HL, Lee DS, Yang SH, Lim CS, Chung JH, Kim S, Lee JS, Kim YS (2002). The polymorphism of monocyte chemoattractant protein-1 is associated with the renal disease of SLE. *Am J Kidney Dis* **40:** 1146–1152.

Letendre S, Marquie-Beck J, Singh KK, de Almeida S, Zimmerman J, Spector SA, Grant I, Ellis R (2004). The monocyte chemotactic protein-1 −2578G allele is associated with elevated MCP-1 concentrations in cerebrospinal fluid. *J Neuroimmunol* **157:** 193–196.

Mankowski JL, Queen SE, Clements JE, Zink MC (2004). Cerebrospinal fluid markers that predict SIV CNS disease. *J Neuroimmunol* **157:** 66–70.

McArthur JC (2004). HIV dementia: an evolving disease. *J Neuroimmunol* **157:** 3–10.

McArthur JC, Brew BJ, Nath A (2005). Neurological complications of HIV infection. *Lancet Neurol* **4:** 543–555.

Mengozzi M, De Filippi C, Transidico P, Biswas P, Cota M, Ghezzi S, Vicenzi E, Mantovani A, Sozzani S, Poli G (1999). Human immunodeficiency virus replication induces monocyte chemotactic protein-1 in human macrophages and U937 promonocytic cells. *Blood* **93:** 1851–1857.

Modi WS, Goedert JJ, Strathdee S, Buchbinder S, Detels R, Donfield S, O'Brien SJ, Winkler C (2003). MCP-1-MCP-3-Eotaxin gene cluster influences HIV-1 transmission. *AIDS* **17:** 2357–2365.

Mori H, Kaneko Y, Narita I, Goto S, Saito N, Kondo D, Sato F, Ajiro J, Saga D, Ogawa A, Sakatsume M, Ueno M, Tabei K, Gejyo F (2005). Monocyte chemoattractant protein-1 A-2518G gene polymorphism and renal survival of Japanese patients with immunoglobulin A nephropathy. *Clin Exp Nephrol* **9:** 297–303.

Muhlbauer M, Bosserhoff AK, Hartmann A, Thasler WE, Weiss TS, Herfarth H, Lock G, Scholmerich J, Hellerbrand C (2003). A novel MCP-1 gene polymorphism is associated with hepatic MCP-1 expression and severity of HCV-related liver disease. *Gastroenterology* **125:** 1085–1093.

Murray EA, Rausch DM, Lendvay J, Sharer LR, Eiden LE (1992). Cognitive and motor impairments associated with SIV infection in rhesus monkeys. *Science* **255:** 1246–1249.

Nei M (1987). *Molecular evolutionary genetics*. New York: Columbia University Press.

Neuenburg JK, Brodt HR, Herndier BG, Bickel M, Bacchetti P, Price RW, Grant RM, Schlote W (2002). HIV-related neuropathology, 1985 to 1999: rising prevalence of HIV encephalopathy in the era of highly active antiretroviral therapy. *J Acquir Immune Defic Syndr* **31:** 171–177.

Persidsky Y, Ghorpade A, Rasmussen J, Limoges J, Liu XJ, Stins M, Fiala M, Way D, Kim KS, Witte MH, Weinand M, Carhart L, Gendelman HE (1999). Microglial and astrocyte chemokines regulate monocyte migration through the blood-brain barrier in human immunodeficiency virus-1 encephalitis. *Am J Pathol* **155:** 1599–1611.

Pola R, Flex A, Gaetani E, Proia AS, Papaleo P, Giorgio AD, Straface G, Pecorini G, Serricchio M, Pola P (2004). Monocyte chemoattractant protein-1 (MCP-1) gene polymorphism and risk of Alzheimer's disease in Italians. *Exp Gerontol* **39:** 1249–1252.

Rovin BH, Lu L, Saxena R (1999). A novel polymorphism in the MCP-1 gene regulatory region that influences MCP-1 expression. *Biochem Biophys Res Comm* **259:** 344–348.

Sacktor N, McDermott MP, Marder K, Schifitto G, Selnes OA, McArthur JC, Stern Y, Albert S, Palumbo D, Kieburtz K, De Marcaida JA, Cohen B, Epstein L (2002). HIV-associated cognitive impairment before and after the advent of combination therapy. *J NeuroVirol* **8:** 136–142.

Sekine O, Nishio Y, Egawa K, Nakamura T, Maegawa H, Kashiwagi A (2002). Insulin activates CCAAT/enhancer binding proteins and proinflammatory gene expression through the phosphatidylinositol 3-kinase pathway in vascular smooth muscle cells. *J Biol Chem* **277:** 36631–36639.

Sharer LR, Baskin GB, Cho ES, Murphey-Corb M, Blumberg BM, Epstein LG (1988). Comparison of simian immunodeficiency virus and human immunodeficiency virus encephalitides in the immature host. *Ann Neurol* **23 (Suppl):** S108–S112.

Sibal LR, Samson KJ (2001). Nonhuman primates: a critical role in current disease research. *Ilar J* **42:** 74–84.

Simeoni E, Hoffmann MM, Winkelmann BR, Ruiz J, Fleury S, Boehm BO, Marz W, Vassalli G (2004). Association between the A-2518G polymorphism in the monocyte chemoattractant protein-1 gene and insulin resistance and type 2 diabetes mellitus. *Diabetologia* **47:** 1574–1580.

Szalai C, Duba J, Prohaszka Z, Kalina A, Szabo T, Nagy B, Horvath L, Csaszar A (2001a). Involvement of polymorphisms in the chemokine system in the susceptibility for coronary artery disease (CAD). Coincidence of elevated Lp(a) and MCP-1 −2518 G/G genotype in CAD patients. *Atherosclerosis* **158:** 233–239.

Szalai C, Kozma GT, Nagy A, Bojszko A, Krikovszky D, Szabo T, Falus A (2001b). Polymorphism in the gene regulatory region of MCP-1 is associated with asthma susceptibility and severity. *J Allergy Clin Immunol* **108:** 375–381.

Tosi AJ, Morales JC, Melnick DJ (2003). Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. *Evol Int J Org Evol* **57:** 1419–1435.

Wagner K, Dendorfer U, Chilla S, Schlondorff D, Luckow B (2001). Identification of new regulatory sequences far upstream of the mouse monocyte chemoattractant protein-1 gene. *Genomics* **78:** 113–123.

Wegscheider BJ, Weger M, Renner W, Posch U, Ulrich S, Hermann J, Ardjomand N, Haller-Schober EM, El-Shabrawi Y (2005). Role of the CCL2/MCP-1 −2518A>G gene polymorphism in HLA-B27 associated uveitis. *Mol Vis* **11:** 896–900.

Weiss JM, Nath A, Major EO, Berman JW (1999). HIV-1 Tat induces monocyte chemoattractant protein-1-mediated monocyte transmigration across a model of the human blood-brain barrier and up-regulates CCR5 expression on human monocytes. *J Immunol* **163:** 2953–2959.

Weiss L, Si-Mohamed A, Giral P, Castiel P, Ledur A, Blondin C, Kazatchkine MD, Haeffner-Cavaillon N (1997). Plasma levels of monocyte chemoattractant protein-1 but not those of macrophage inhibitory protein-1alpha and RANTES correlate with virus load in human immunodeficiency virus infection. *J Infect Dis* **176:** 1621–1624.

Wu DT, Woodman SE, Weiss JM, McManus CM, D'Aversa TG, Hesselgesser J, Major EO, Nath A, Berman JW (2000). Mechanisms of leukocyte trafficking into the CNS. *J NeuroVirol* **6 (Suppl) 1:** S82–S85.

Yang B, Houlberg K, Millward A, Demaine A (2004). Polymorphisms of chemokine and chemokine receptor genes in type 1 diabetes mellitus and its complications. *Cytokine* **26:** 114–121.

Zhou ZH, Chaturvedi P, Han YL, Aras S, Li YS, Kolattukudy PE, Ping D, Boss JM, Ransohoff RM (1998). IFN-gamma induction of the human monocyte chemoattractant protein (hMCP)-1 gene in astrocytoma cells: functional interaction between an IFN-gamma-activated site and a GC-rich element. *J Immunol* **160:** 3908–3916.

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.

Zink MC, Coleman GD, Mankowski JL, Adams RJ, Tarwater PM, Fox K, Clements JE (2001). Increased macrophage chemoattractant protein-1 in cerebrospinal fluid precedes and predicts simian immunodeficiency virus encephalitis. *J Infect Dis* **184:** 1015–1021.

Zink MC, Suryanarayana K, Mankowski JL, Shen A, Piatak M Jr, Spelman JP, Carter DL, Adams RJ, Lifson JD, Clements JE (1999a). High viral load in the cerebrospinal fluid and brain correlates with severity of simian immunodeficiency virus encephalitis. *J Virol* **73:** 10480–10488.

Zink MC, Uhrlaub J, DeWitt J, Voelker T, Bullock B, Mankowski J, Tarwater P, Clements J, Barber S (2005). Neuroprotective and anti-human immunodeficiency virus activity of minocycline. *JAMA* **293:** 2003–2011.

Zink WE, Zheng J, Persidsky Y, Poluektova L, Gendelman HE (1999b). The neuropathogenesis of HIV-1 infection. *FEMS Immunol Med Microbiol* **26:** 233–241.

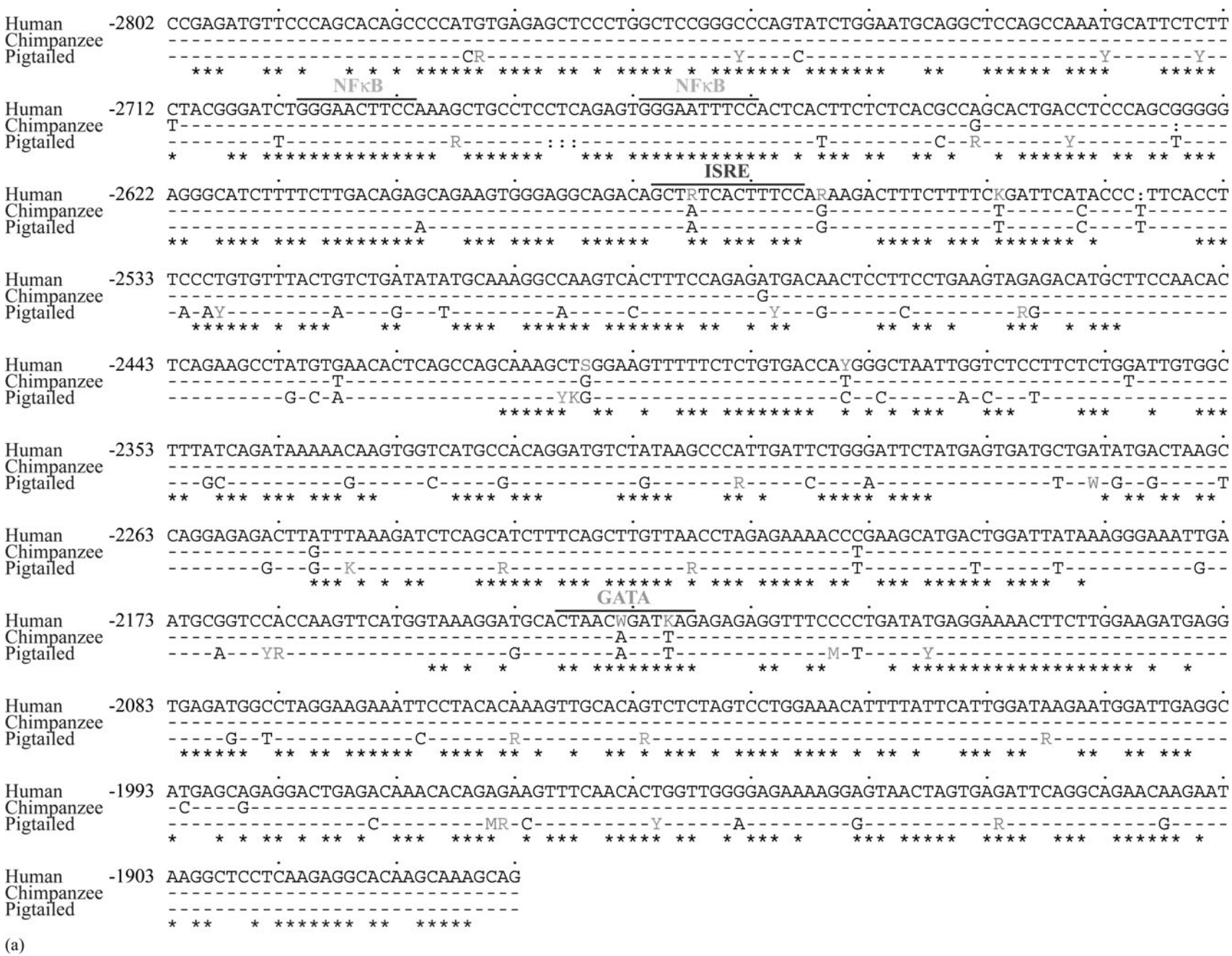

Figure S1 Pigtailed macaque sequence alignment with human and chimpanzee at the distal and proximal CCL2 promoter Regions. Sequence for the (**A**) distal and (**B**) proximal CCL2 promoter regions of pigtailed macaques used in an SIV model of HIV CNS disease. The consensus for the pigtailed macaques was aligned to human and chimpanzee sequences from NCBI. Asterisks (*) at the bottom of the aligned sequence indicate nucleotides that are conserved between human and mouse (determined by alignment of UCSC genome browser). Red variable bases indicate positions that had sufficient quantities of each allele so they could be used in the Fisher's exact test and Mann-Whitney test, where as green positions did not contain enough macaques with the minor allele to perform the statistical tests.

```
Human      -513 TAATGCATTGTCAGGGAGCCGGCCAAAGCTTGAGAGCTCCTTCCTGGCTGGGAGGCCCCTTGGAATGTGGCCTGAAGGTAAGCYGGCAGC
Chimpanzee      ------------------------------------------------------------A----------------------------T------
Pigtailed       --------------------R-----------------------------------K---------C------A-----------T------
                * **       ** ***  * *** *    ** * *  * ****  ****  ***  ******   * ***   *** ** ****

Human      -423 GAGCCTGACATGCTTTCATCTAGTWTCCTCGCTTCCTTCCTTTTCTGCAGTTTTCGCTTCASAGAAAGCAGAATCCTTAAAAATAACCCT
Chimpanzee      -------------------------T-----------------------------------G----------------------------
Pigtailed       R------------------------T------------------------------A-----G-------G---------------GT---
                   **   ***  * *  * *** ***** *    *****           ***   ***** *********  *  *

Human      -333 CTTAGTTCACATCTGTGGTCAGTCTGGGCTTAATGGCACCCCATCCTCCCCATTTGCTCATTTGGTCTCAGCAGTGAATGGAAAAAGTGT
Chimpanzee      ------------------------------------------------------------------------------------------
Pigtailed       --------------------------T-----------------------------------------------Y-----------------
                     ******  ********   * *  *** ** *  ***    *****   ********** ** *** ****  * ****** *
                                                   GAS                                                  TRE
Human      -243 CTCGTCCTGACCCCCTGCTTCCCTTTCCTACTTCCTGGAAATCCACAGGATGCTGCATTTGCTCAGCAGATTTAACAGCCCACTTATCAC
Chimpanzee      -------------A----------------------------------------------------------------------G----------
Pigtailed       --T----------A-------------------------------C---R-C-----------------------------------T-T
                *   * *  **** * * *** ** * * ************ **  ***   ***** ** ****** ****  *    ***********
                   NFκB                 TRE GC BOX                          TATA Box
Human      -153 TCATGGAAGATCCCTCCTCCTGCTTGACTCCKCCYTCTCT::CCCTCTGCCCGCTTTCAATAAGAGGCAGAGACAGCAGCCAGAGGAACC
Chimpanzee      ------G-----------------------G--C-----::-------------------------------------------------A
Pigtailed       --------K--S-------------------A--C-----CT---------T-----------A------------Y------------A
                * **  * * *** **  ***** ******* ***         * ***   **** ****** **** *   *********** *  *
                                                                                  +1
Human       -65 GAGAGGCTGAGACTAACCCAGAAACATCCAATTCTCAAACTGAAGCTCGCACTCTCGCCTCCAGCATGAAAGTCTCTGC
Chimpanzee      ------------------------------------------------------------------------------------
Pigtailed       ------------------------R------G----------------T-S-YC---A----------------------
                ***** * *** * * ***** * ** **** *****  *******  ** ***** ****** **** * *** ***
```

(b)

Figure S1 (Continued).