1-1-2011

Differential microRNA regulation of *HLA-C* expression and its association with HIV control

Smita Kulkarni
SAIC-Frederick, Inc.

Ram Savan
NCI-Frederick

Ying Qi
Ragon Institute of Massachusetts General Hospital

Xiaojiang Gao
Ragon Institute of Massachusetts General Hospital

Yuko Yuki
Ragon Institute of Massachusetts General Hospital

See next page for additional authors

Authors
Smita Kulkarni, Ram Savan, Ying Qi, Xiaojiang Gao, Yuko Yuki, Sara E. Bass, Maureen P. Martin, Peter Hunt, Steven G. Deeks, Amalio Telenti, Florencia Pereyra, David Goldstein, Steven Wolinsky, Bruce Walker, Howard A. Young, and Mary Carrington

LETTER

doi:10.1038/nature09914

Differential microRNA regulation of *HLA-C* expression and its association with HIV control

Smita Kulkarni[1,2]*, Ram Savan[3]*, Ying Qi[1,2], Xiaojiang Gao[1,2], Yuko Yuki[1,2], Sara E. Bass[1], Maureen P. Martin[1,2], Peter Hunt[4], Steven G. Deeks[4], Amalio Telenti[5], Florencia Pereyra[2], David Goldstein[6], Steven Wolinsky[7], Bruce Walker[2], Howard A. Young[3] & Mary Carrington[1,2]

The *HLA-C* locus is distinct relative to the other classical *HLA* class I loci in that it has relatively limited polymorphism[1], lower expression on the cell surface[2,3], and more extensive ligand–receptor interactions with killer-cell immunoglobulin-like receptors[4]. A single nucleotide polymorphism (SNP) 35 kb upstream of *HLA-C* (rs9264942; termed −35) associates with control of HIV[5–7], and with levels of *HLA-C* messenger RNA transcripts[8] and cell-surface expression[7], but the mechanism underlying its varied expression is unknown. We proposed that the −35 SNP is not the causal variant for differential HLA-C expression, but rather is marking another polymorphism that directly affects levels of HLA-C[7]. Here we show that variation within the 3′ untranslated region (UTR) of *HLA-C* regulates binding of the microRNA hsa-miR-148 to its target site, resulting in relatively low surface expression of alleles that bind this microRNA and high expression of *HLA-C* alleles that escape post-transcriptional regulation. The 3′ UTR variant associates strongly with control of HIV, potentially adding to the effects of genetic variation encoding the peptide-binding region of the *HLA* class I loci. Variation in HLA-C expression adds another layer of diversity to this highly polymorphic locus that must be considered when deciphering the function of these molecules in health and disease.

MicroRNAs (miRNAs) are a class of non-protein-coding RNAs that are estimated to regulate 30% of all genes in animals[9] by binding to specific sites in the 3′ UTR, resulting in post-transcriptional repression, cleavage or destabilization[10–12]. The 3′ UTR of the *HLA-C* gene is predicted to be a target for 26 distinct human miRNAs using three miRNA-target-prediction programs (Supplementary Fig. 1), of which three (miR-148a and miR-148b, which bind the same target site, and miR-657) were shown to have the greatest likelihood of binding. We sequenced the 3′ UTRs of the common *HLA-C* alleles (Supplementary Fig. 2) and show that the two binding sites of these three miRNAs are polymorphic (Supplementary Fig. 3a). The binding site for miR-148a/miR-148b contains a single base pair insertion/deletion at position 263 downstream of the *HLA-C* stop codon (rs67384697G representing the insertion (263ins) and rs67384697− representing the deletion (263del)) along with other precisely linked variants (259C/T, 261T/C, 266C/T). These variants are likely to impose a restriction in miR-148a/miR-148b binding, as prediction algorithms indicate that the binding of these miRNAs to the alleles marked by 263ins (for example, *Cw*0702*, a low-expression allotype) is more stable than to alleles with 263del (for example, *Cw*0602*, a high-expression allotype) (Supplementary Fig. 3b). Similarly, alleles with 307C within the miR-657 target site are predicted to be better targets of miR-657 than those with 307T (Supplementary Fig. 4). Thus, variation in the 3′ UTR of *HLA-C* may influence the interaction between these miRNAs and their putative binding sites in an allele-specific manner, potentially leading to differential levels of HLA-C allotype expression.

To test directly whether the variation in the *HLA-C* 3′ UTR affects levels of protein expression, the full-length 3′ UTRs containing intact miR-148a/miR-148b- and miR-657-binding sites (that is, 263ins and 307C, respectively; *Cw*0702*, *Cw*0303*, *Cw*0401*, *Cw*0701*) and disrupted binding sites (that is, 263del and 307T, respectively; *Cw*0602*, *Cw*0802*, *Cw*1203*, *Cw*1502*) were each cloned downstream of the luciferase gene in a pGL3 reporter construct (Fig. 1a). The constructs were then transfected into HLA class I negative B721.221 cells, and the

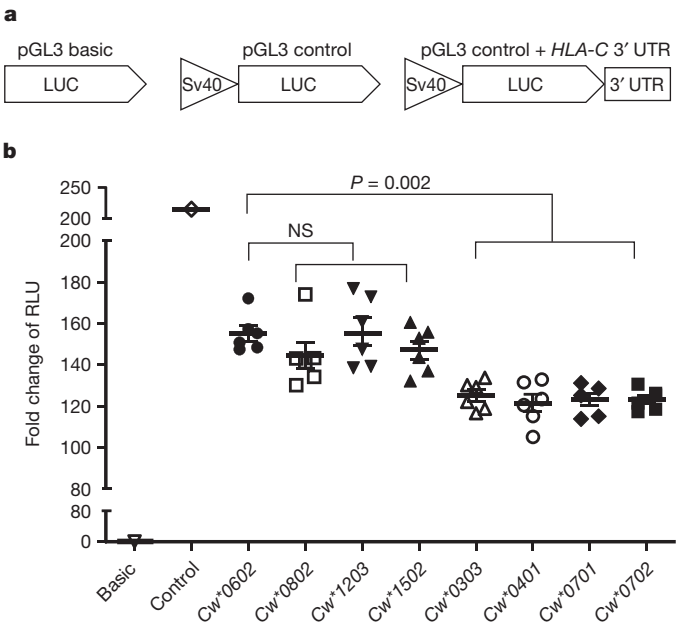


Figure 1 | Variation in the *HLA-C* 3′ UTR differentially affects the expression of a reporter gene. **a**, **b**, Full-length 3′ UTRs of various *HLA-C* alleles cloned into luciferase (LUC) reporter constructs were transfected into B721.221 cell lines and the stability of the mRNA was estimated by dual luciferase reporter assays. The normalized luciferase activity is presented as fold change of relative light units (RLU). The data represent six replicates in each experimental group, the mean ± standard error (s.e.) are depicted as horizontal and vertical bars for each group, respectively, and one of three comparable experiments performed is shown. Non-parametric Wilcoxon–Mann–Whitney tests were used for statistical comparisons and two-tailed *P* values are indicated. NS, not significant. **a**, Schematic representations of the luciferase reporter constructs used in this study. Sv40, Simian virus 40. **b**, Fold change in luciferase activity of 3′ UTRs of *HLA-C* alleles as compared to that of *Cw*0602*.

[1]Cancer and Inflammation Program, Laboratory of Experimental Immunology, SAIC-Frederick, Inc., NCI-Frederick, Frederick, Maryland 21702, USA. [2]Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, Boston, Massachusetts 02114, USA. [3]Cancer and Inflammation Program, Laboratory of Experimental Immunology, NCI-Frederick, Frederick, Maryland, 21702 USA. [4]San Francisco General Hospital AIDS division, University of California, San Francisco, California 94110, USA. [5]Institute of Microbiology, University of Lausanne, Lausanne 1011, Switzerland. [6]Institute for Genome Sciences & Policy, Center for Human Genome Variation, Duke University, Durham, North Carolina 27708, USA. [7]Division of Infectious Diseases, The Feinberg School of Medicine, Northwestern University, Chicago, Illinois 60611, USA.
*These authors contributed equally to this work.

level of luciferase activity was measured (fold increase of relative light units). Although the *Cw*0602* 3′ UTR repressed luciferase activity as compared to the control containing no 3′ UTR, the constructs containing intact miRNA-binding sites (that is, 263ins and 307C; *Cw*0702*, *Cw*0303*, *Cw*0401*, *Cw*0701*) produced significantly lower luciferase activity relative to the construct containing the 3′ UTR of *Cw*0602*, which contains 263del and 307T (Fig. 1b). However, 3′ UTRs from other alleles with the 263del and 307T variants (*Cw*0802*, *Cw*1203*, *Cw*1502*) did not show significant variation in luciferase activity as compared to *Cw*0602* (Fig. 1b). Psicheck2 reporter constructs containing 3′ UTRs of *Cw*0602* also produced significantly higher luciferase activity as compared to those with *Cw*0702* 3′ UTR (Supplementary Fig. 5a), indicating that this effect was reproducible in a distinct reporter construct. Further, pGL3 constructs containing 3′ UTRs of *Cw*0602* and *Cw*0702* in three additional cell lines showed the same pattern as that seen in B721.221 cells, indicating a consistent difference of these 3′ UTRs in the regulation of HLA-C expression that is independent of cell type (Supplementary Fig. 5b–e). Thus, *HLA-C* 3′ UTR alleles characterized by variation at positions 263 and 307 within miRNA-binding regions differentially regulate gene expression.

The expression of endogenous mature miR-148b and miR-152, another miR-148 miRNA family member, was very low as compared to miR-148a, and miR-657 was undetectable in *HLA-C* homozygous B lymphoblastoid cell lines (BLCLs) and B721.221 cells (Supplementary Fig. 6). These data point to the involvement of miR-148a rather than miR-148b, miR-152, or miR-657 in regulation of HLA-C expression. Additionally, disruption of the miR-657-binding site by site-directed mutagenesis had no effect on luciferase activity (Supplementary Fig. 7a, b), indicating that miR-657 does not affect HLA-C expression.

To test whether variants in the miR-148a-binding site account for the differential gene expression patterns, we swapped positions 256–266 of the 3′ UTR of *Cw*0602* to match those of *Cw*0702* and vice versa, thereby providing an intact miR-148a-binding site to the 3′ UTR of *Cw*0602* (*06mut*) and disrupting the binding site for miR-148a in the *Cw*0702* 3′ UTR (*07mut*), but leaving the remainder of the 3′ UTR sequences intact (Fig. 2a). The luciferase activity of *06mut* was significantly lower than that of *07mut* (Fig. 2b), indicating that the polymorphisms between positions 256–266 in the miR-148a-binding region account for the difference in luciferase expression between constructs containing the 3′ UTRs of *Cw*0602* versus *Cw*0702*. Two other polymorphic sites, A256C and A267G, in the miR-148a-binding site (Supplementary Fig. 3a) distinguish different sets of alleles as compared to 263del/ins, but these two variants had no effect on miRNA-mediated suppression (Supplementary Fig. 7c, d).

Further validation of the differential regulation of *HLA-C* alleles by miR-148a was achieved by co-transfection of B221.227 cell lines with either a mimic or an inhibitor of miR-148a along with a luciferase reporter construct that contained the 3′ UTR with 263ins (*Cw*0702* or *06mut)* or with 263del (*Cw*0602* or *07mut*). The normalized luciferase activity in cells transfected with the constructs containing the 263del allele (*Cw*0602* and *07mut* 3′ UTR) was not significantly altered by co-transfection with either the mimic or the inhibitor (Fig. 2c, d). However, the mimic of miR-148a further repressed luciferase activity in cells transfected with the 263ins allele (*Cw*0702* and *06mut* 3′ UTR), whereas co-transfection with inhibitor rescued the suppression significantly (Fig. 2c, d). These data provide further support for allele-specific miR-148a targeting of the *HLA-C* 3′ UTR.

BLCLs from individuals homozygous for either *Cw*0602* (BLCL-Cw*0602Hom) or *Cw*0702* (BLCL-Cw*0702Hom) were used to determine whether the variation in the miR-148a-binding site affected endogenous HLA-C expression on the cell surface. As described previously, overall HLA-C expression on a *Cw*0602* homozygous cell line was higher than that on a *Cw*0702* homozygous cell line. As expected, transfection with mimics or inhibitors of miR-148a (Fig. 3a) and miR-148b (Fig. 3b) had no significant effect on cell-surface expression of Cw*0602, an allele containing 263del in the 3′ UTR that disrupts

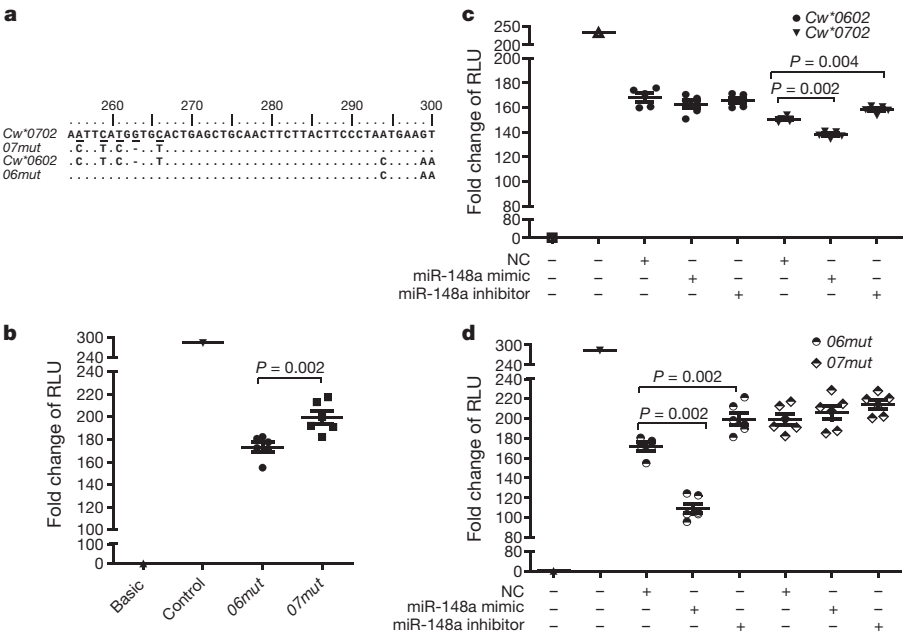


Figure 2 | Disruption of miR-148a target site rescues suppression. **a**, Partial sequence of mutated 3′ UTRs of *Cw*0602* and *Cw*0702* (*06mut* and *07mut*, respectively) are aligned to 3′ UTR sequences of native *Cw*0602* and *Cw*0702*. Identical nucleotides are shown as dots, altered nucleotides are underlined, and deletions are indicated by dashes for optimal alignment. **b**, Fold change in luciferase activity of the modified 3′ UTR (*06mut* and *07mut*). **c**, Fold change in luciferase activity of reporters containing wild-type *Cw*0602* or *Cw*0702* 3′ UTR sequences upon introduction of miR-148a mimic and inhibitor. **d**, Fold change in luciferase activity of reporters containing *06mut* and *07mut* 3′ UTR sequences upon introduction of miR-148a mimic and inhibitor. Presence (+) or absence (−) of each variable, including a negative control (NC) miRNA, a mimic of miR-148a, or an inhibitor of miR-148a is shown. The data represent six replicates in each experimental group, the mean ± s.e. are depicted as horizontal and vertical bars for each group, respectively, and one of three comparable experiments performed is shown. Non-parametric Wilcoxon–Mann–Whitney tests were used for statistical comparisons and two-tailed *P* values are indicated.

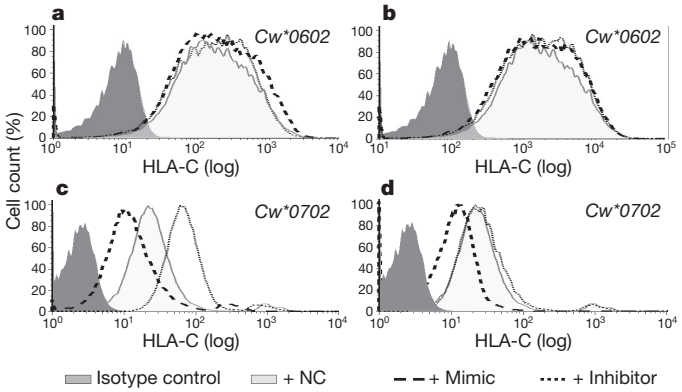


Figure 3 | miR-148a affects cell-surface expression of HLA-C. **a**–**d**, Histograms of HLA-C cell-surface expression on HLA-C homozygous BLCLs using flow cytometry are illustrated. In each plot, a negative control (NC) miRNA that does not bind to the 3′ UTR of *HLA-C* was included. **a**, **b**, *HLA-Cw*0602* homozygous cells (BLCL-Cw*0602Hom) transfected with either a mimic or an inhibitor of miR-148a (**a**) or miR-148b (**b**). **c**, **d**, *HLA-Cw*0702* homozygous cells (BLCL-Cw*0702Hom) transfected with either a mimic or an inhibitor of miR-148a (**c**) or miR-148b (**d**).

miR-148a/miR-148b binding. However, transfection of either miR-148a or miR-148b mimic resulted in decreased expression of Cw*0702 relative to cells transfected with a negative control (Fig. 3c and d, respectively), indicating that increased levels of either of these miRNAs can further downregulate HLA-C expression of alleles that contain an intact binding site for miR-148a/miR-148b (263ins), such as *Cw*0702*. The inhibitor of miR-148a significantly increased the level of endogenous Cw*0702 expression, but the inhibitor of miR-148b had no effect on expression of Cw*0702 (Fig. 3c and d, respectively), confirming the very low levels of miR-148b endogenous expression (Supplementary Fig. 6).

The differential miR-148a regulation of expression across *HLA-C* alleles was precisely reflected in experiments involving: (1) additional *HLA-C* homozygous BLCLs (Supplementary Fig. 8a, b); (2) use of another form of miR-148a inhibitor (peptide nucleic acid inhibitor; Supplementary Fig. 8c); and (3) analysis of total cellular HLA-C protein as determined by western blot (Supplementary Figs 9, 10 and Supplementary Note 1). We conclude that miR-148a regulates the expression of HLA-C in an allele-specific manner that is dependent on variation in the miR-148a-binding site of the *HLA-C* 3′ UTR.

The miR-148a-binding site of the *HLA-C* 3′ UTR is in strong linkage disequilibrium with the −35 SNP that was shown to associate with control of HIV and HLA-C expression levels[5–7] ($D' = 0.75$, $r^2 = 0.74$; $P < 0.0001$, $N = 1{,}760$). Although there is no explanation for a direct causal effect of −35 variation on HLA-C expression[7], the interaction between miR-148a and its polymorphic binding site in the 3′ UTR of *HLA-C* presents a clear rationale for variable levels of HLA-C expression. We determined the frequencies of the 263del/ins genotypes in a cohort of 2,527 HIV-infected European American individuals. Subjects with mean plasma viral loads of <2,000 copies of viral RNA per ml of plasma (controllers) were enriched for 263del, whereas those with viral loads of >10,000 copies of viral RNA per ml of plasma (noncontrollers) had a significantly higher frequency of 263ins (Supplementary Table 1). Because of the strong linkage disequilibrium across the *HLA-C* and *HLA-B* genes (Supplementary Tables 2, 3), we determined whether the 3′ UTR variant has an effect on HIV control that is independent of individual *HLA-A*, *B*, or *C* alleles. A logistic regression approach with stepwise selection of the *HLA-C* 3′ UTR 263 variant along with all *HLA-A*, *B* and *C* alleles that have ≥1% frequency in our cohort (63 alleles) was used. In this analysis, the 263del/del versus 263ins/ins comparison remains significant along with 5 of the 63 *HLA* alleles (*B*5701*, *B*5703*, *B*2705*, *B*5801* and *Cw*1402*; Table 1; for frequencies, see Supplementary Table 1).

Table 1 | Effect of *HLA-C* 3′ UTR 263 on mean viral load

Significant independent variables	*P* value	OR	95% CI
del/del versus ins/ins	2×10^{-14}	0.33	0.25–0.43
*B*2705* versus others	3×10^{-6}	0.34	0.22–0.54
*B*5701* versus others	1×10^{-12}	0.21	0.14–0.32
*B*5703* versus others	3×10^{-5}	0.01	0.002–0.10
*B*5801* versus others	9×10^{-4}	0.27	0.12–0.59
*Cw*1402* versus others	1×10^{-4}	0.26	0.13–0.52

CI, confidence interval; OR, odds ratio. $N = 2{,}527$. A logistic regression analysis with stepwise selection using the *HLA-C* 3′ UTR 263 del/del versus ins/ins comparison and 63 *HLA-A*, *B* and *C* alleles with ≥1% frequency as independent variables in the model was performed using PROC LOGISTIC (SAS 9.1 version, SAS Institute). Significance level for selecting variables shown in the table was $P < 0.05$.

Although these data indicate that the 3′ UTR del/ins variant has an independent effect on HIV control (see Supplementary Note 2 for potential mechanisms that could explain the association), we still cannot completely rule out the possibility that the strong linkage disequilibrium in the region is confounding the results[13] (Supplementary Table 4 and Supplementary Note 3).

The extensive number of disease associations with *HLA* class I and II genes has largely been ascribed to the polymorphic peptide-binding amino acid positions of these molecules. Some reports have speculated that gene expression[14] and/or splicing patterns of the *HLA* genes[15,16] may have a role, but convincing data are missing. Of interest, the *HLA-G* 3′ UTR was shown to encode a polymorphic target site for miR-148a/b[17,18]. Levels of HLA-G have been suggested to alter risk of asthma in children of mothers with asthma[18], although the specificity of assays reporting HLA-G expression beyond implanting placental cells has been questioned[19]. Recently, a variant 35 kb upstream of the HLA-C gene was shown to associate with differential *HLA-C* mRNA expression, cell-surface expression and outcome after HIV infection[5–8]. We have now established a very convincing case that this −35 SNP is marking a functional insertion/deletion variant in the 3′ UTR of *HLA-C* that directly determines expression of the various HLA-C allotypes differentially through miR-148a recognition. These data indicate another tier of diversity to the polymorphic *HLA-C* locus beyond that encoding variation in the peptide-binding region of the gene. We suggest that disease-associated haplotypes may exert their effects through multiple mechanisms, including the type of peptides they bind and their level of expression, and that it is the combination of these that then determines the overall susceptibility status of the haplotype.

Expression levels of different HLA-C allotypes occur as a continuous gradient rather than the bimodal expression pattern that would be expected if miR-148a regulation were the sole mechanism involved. Thus, additional *cis*-acting factors may fine-tune HLA-C expression in an allotype-specific manner. *Trans*-acting factors unlinked to the *HLA-C* locus may also affect expression levels in a manner that is independent of HLA-C allotype, leading to some degree of variation in expression levels of a given HLA-C allotype. Although the system regulating *HLA-C* expression is multifactorial, the significant involvement of miRNA in this process provides new approaches for manipulation of the immune system in the treatment of human disease.

METHODS SUMMARY

DNA from 2,527 HIV$^+$ patients of European descent was used to determine the effect of the *HLA-C* 3′ UTR variation on control of HIV viral load. Viral load measurements were obtained from participants of the Multicentre AIDS cohort study (MACS)[20], Swiss HIV Cohort (http://www.shcs.ch), the SCOPE cohort[21] and the International HIV Controllers Study Cohort (http://www.hivcontrollers.org). Individuals were grouped into those who maintain mean viral load < 2,000 (controllers) and those who have mean viral load > 10,000 (noncontrollers).

Complete *HLA-C* 3′ UTR fragments were amplified, inserted into the XbaI site downstream of the luciferase gene in a pGL3-control vector (Promega), and transfected into B721.221, BLCL and Jurkat cells using AMAXA nucleofector (Lonza) and into 293T cells using Fugene6 (Roche). Luciferase activity was measured using the Dual Luciferase Reporter Assay System (Promega) and presented as fold change of relative light units[22]. For studies of miR-148a/miR-148b mimics and inhibitors (Dharmacon), 20 pmol per well of oligonucleotide mimics or inhibitors of miR-148a and miR-148b were transfected into the cells. Surface expression of

HLA-C on BLCLs was analysed using staining with DT9 antibody (provided by V. Braud)[23].

Total RNA was extracted using the Total RNA Purification Kit (Norgen). Relative quantification of miR-148a and miR-148b was performed using a Taqman real-time PCR assay (Applied Biosystems) and RNU48 served as an endogenous RNA control.

SAS9.1 (SAS Institute) was used for data management and statistical analyses.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 March 2010; accepted 10 February 2011.
Published online 17 April 2011.

1. Zemmour, J. & Parham, P. Distinctive polymorphism at the HLA-C locus: implications for the expression of HLA-C. *J. Exp. Med.* **176,** 937–950 (1992).
2. McCutcheon, J. A., Gumperz, J., Smith, K. D., Lutz, C. T. & Parham, P. Low HLA-C expression at cell surfaces correlates with increased turnover of heavy chain mRNA. *J. Exp. Med.* **181,** 2085–2095 (1995).
3. Snary, D., Barnstable, C. J., Bodmer, W. F. & Crumpton, M. J. Molecular structure of human histocompatibility antigens: the HLA-C series. *Eur. J. Immunol.* **7,** 580–585 (1977).
4. Bashirova, A. A., Martin, M. P., McVicar, D. W. & Carrington, M. The killer immunoglobulin-like receptor gene cluster: tuning the genome for defense. *Annu. Rev. Genomics Hum. Genet.* **7,** 277–300 (2006).
5. Fellay, J. *et al.* A whole-genome association study of major determinants for host control of HIV-1. *Science* **317,** 944–947 (2007).
6. International HIV Controllers Study. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science* **310,** 1551–1557 (2010).
7. Thomas, R. *et al.* HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of *HLA-C*. *Nature Genet.* **41,** 1290–1294 (2009).
8. Stranger, B. E. *et al.* Genome-wide associations of gene expression variation in humans. *PLoS Genet.* **1,** e78 (2005).
9. Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120,** 15–20 (2005).
10. Giraldez, A. J. *et al.* Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312,** 75–79 (2006).
11. Lim, L. P. *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433,** 769–773 (2005).
12. Wu, L., Fan, J. & Belasco, J. G. MicroRNAs direct rapid deadenylation of mRNA. *Proc. Natl Acad. Sci. USA* **103,** 4034–4039 (2006).
13. Corrah, T. W. *et al.* A reappraisal of the relationship between the HIV-1-protective single nucleotide polymorphism 35 kb upstream of the *HLA-C* gene and surface HLA-C expression. *J. Virol.* **85,** 3367–3374 (2011).
14. Schaefer, M. R. *et al.* A novel trafficking signal within the HLA-C cytoplasmic tail allows regulated expression upon differentiation of macrophages. *J. Immunol.* **180,** 7804–7817 (2008).
15. Kralovicova, J. & Vorechovsky, I. Position-dependent repression and promotion of DQB1 intron 3 splicing by GGGG motifs. *J. Immunol.* **176,** 2381–2388 (2006).
16. Krangel, M. S. Secretion of HLA-A and -B antigens via an alternative RNA splicing pathway. *J. Exp. Med.* **163,** 1173–1190 (1986).
17. Castelli, E. C. *et al. In silico* analysis of microRNAs targeting the *HLA-G* 3′ untranslated region alleles and haplotypes. *Hum. Immunol.* **70,** 1020–1025 (2009).
18. Tan, Z. *et al.* Allele-specific targeting of microRNAs to *HLA-G* and risk of asthma. *Am. J. Hum. Genet.* **81,** 829–834 (2007).
19. Apps, R., Gardner, L. & Moffett, A. A critical look at HLA-G. *Trends Immunol.* **29,** 313–321 (2008).
20. Phair, J. *et al.* Acquired immune deficiency syndrome occurring within 5 years of infection with human immunodeficiency virus type-1: the Multicenter AIDS Cohort Study. *J. Acquir. Immune Defic. Syndr.* **5,** 490–496 (1992).
21. Emu, B. *et al.* Phenotypic, functional, and kinetic parameters associated with apparent T-cell control of human immunodeficiency virus replication in individuals with and without antiretroviral treatment. *J. Virol.* **79,** 14169–14178 (2005).
22. Li, H., Wright, P. W. & Anderson, S. K. Identification and analysis of novel transcripts and promoters in the human killer cell immunoglobulin-like receptor (KIR) genes. *Methods Mol. Biol.* **612,** 377–391 (2010).
23. Braud, V. M., Allan, D. S., Wilson, D. & McMichael, A. J. TAP- and tapasin-dependent HLA-E surface expression correlates with the binding of an MHC class I leader peptide. *Curr. Biol.* **8,** 1–10 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This project has been funded in part with federal funds from the National Cancer Institute, National Institutes of Health (NIH), under contracts HHSN261200800001E, N02-CP-55504, R01-DA04334 and R01-DA12568. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government. This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research and the Cancer Inflammation Program Project Award for the year 2009, a grant from the Bill & Melinda Gates Foundation as part of the Collaboration for AIDS Vaccine Discovery, and the Mark and Lisa Schwartz Foundation. We would also like to thank the patients and investigators involved in the Multicenter AIDS Cohort Study (the MACS is funded by the National Institute of Allergy and Infectious Diseases, with supplemental funding from the National Cancer Institute and the National Heart, Lung and Blood Institute (grants UO1-AI-35042, 5-MO1-RR-00722 (GCRC), UO1-AI-35043, UO1-AI-37984, UO1-AI-35039, UO1-AI-35040, UO1-AI-37613 and UO1-AI-35041), the Swiss HIV Cohort Study (see Supplementary Note 4 for the list of members), supported by the Swiss National Science Foundation grant number 33CSCO-108787, and the SCOPE study, which was funded by the UL1 RR024131 (Clinical and Translational Sciences Award) and P30 AI27763 (Center for AIDS Research) grants. We thank S. Anderson, G. O'Connor and R. Thomas for advice, A. Kronfli and K. Ramakrishnan for assistance in plasmid and genomic DNA preparations, A. McFarland for western blots, V. Braud for the DT9 antibody, R. Johnson and G. Nelson for statistical advice and T. Covell for administrative assistance.

Author Contributions S.K. and R.S. performed and evaluated the miRNA experiments. S.K., R.S., H.A.Y. and M.C. designed the study. M.C. directed the study. S.K., R.S. and M.C. wrote the manuscript. X.G., Y.Y., S.B. and M.M. genotyped *HLA*. Statistical analysis was performed by Y.Q. The clinical samples and data were contributed to by P.H., S.G.D., D.D., A.T., D.G., S.W., F.P. and B.W. Intellectual input was provided by all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.C. (carringm@mail.nih.gov)

METHODS

Subjects. DNA from 2,527 HIV$^+$ patients of European descent was used to determine the effect of the *HLA-C* 3′ UTR variation on control of HIV viral load. DNA from 1,760 individuals of European descent was used to determine the patterns of linkage disequilibrium between the *HLA-C* 3′ UTR variant and the *HLA-C* coding regions. Viral load measurements used in categorical analyses were obtained from participants of the Multicentre AIDS cohort study (MACS)[20], the Swiss HIV Cohort (http://www.shcs.ch), the SCOPE cohort[21], and the International HIV Controllers Study Cohort (http://www.hivcontrollers.org). Subjects were classified as HIV controllers if plasma HIV RNA was below 2,000 copies per ml in a minimum of three determinations in the absence of antiretrovirals, spanning at least a 12-month period. Chronically infected individuals were classified as noncontrollers if plasma HIV RNA was above 10,000 copies per ml in the absence of antiretrovirals. The respective institutional review boards approved the study, and all subjects gave written informed consent. All of the patients were of European descent and standard methods for measurements of viral load were used in all three study cohorts.

***HLA-C* 3′ UTR DNA sequencing.** The entire *HLA-C* 3′ UTR was amplified from genomic DNA by PCR using the following pair of primers: forward 5′-gtgag attctggggagctga and reverse 5′-gaacagcaactaggcacagg. The amplicons were sequenced in both directions using the same primers by capillary electrophoresis using an ABI 3730 DNA analyser (Applied Biosystems).

HLA genotyping. DNA samples were genotyped for *HLA-A*, *B* and *C* genes by sequence-based typing of exons 2 and 3 and/or the PCR-sequence-specific oligonucleotide probe typing protocol as recommended by the 13th International Histocompatibility Workshop[24].

Construction of *HLA-C* 3′ UTR luciferase reporters:. The complete 3′ UTR fragments from various *HLA-C* alleles (*Cw*0303*, *Cw*0401*, *Cw*0502*, *Cw*0602*, *Cw*0701*, *Cw*0702*, *Cw*0802*, *Cw*1203*, *Cw*1502*) were amplified from genomic DNA and inserted into the XbaI site downstream of the luciferase gene in the pGL3-control vector (Promega). Mutations at position 307 were introduced using a site-directed mutagenesis kit (Stratagene) and the following sets of primers: for *07-307C>T* (sense, 5′-gcaacttctacttccctaatgaagttaagaatctgaatataaat ttgtgttc; and antisense, 5′-gaacacaaatttatatt cagattcttaacttcattagggaagtaagaagttgc) and for *07-307C>G* (sense, 5′-caacttcttacttccctaatgaa ttaagaagctgaatataaatttgtgtt; and antisense, 5′-aacacaaatttatattcagcttcttaacttcattagggaagtaagaagttg). The intact miR-148-binding site was reconstituted into the *Cw*0602* 3′ UTR by altering only the motif from positions 256–266 and the construct is referred to as *06mut* (5′-CTTTACG-TGT > 5 ′-ATTCATGGTGC; see Fig. 2). For the *07mut* construct, the miR-148-binding site of the *Cw*0702* 3′ UTR was disrupted by altering positions 256–266 (5′-ATTCATGGTGC > 5′-CTTTACG-TGT). Both *06mut* and *07mut* 3′ UTR sequences were synthesized by Genscript.

Cell lines and culture conditions. The HLA class I negative B721.221 cell line characterized by complete absence of *HLA-A*, *B* and *C* mRNA transcripts[25] were grown in RPMI 1640 (Gibco) medium with 10% heat-inactivated fetal bovine serum (FBS; Atlanta Biologicals) and Epstein–Barr virus (EBV)-transformed B cell lines derived from peripheral blood lymphocytes of individuals homozygous for *HLA-C* alleles (BLCLs) were grown in RPMI 1640 medium with 15% FBS and 1% penicillin–streptomycin–glutamine (PSG; Gibco) at 37 °C in 5% CO_2. The human embryonic kidney cell line (HEK293T)[26] and a T cell line (Jurkat) were cultured in DMEM (Gibco) supplemented with 10% FBS and 1% PSG.

Cell transfection and luciferase reporter assays. B721.221, BLCLs and Jurkat cells were plated at a density of 1×10^6 cells per well in a 96-well plate. 500 ng per well of the pGL3 and 20 ng per well of *Renilla* reporter constructs were transfected using optimized AMAXA nucleofector. 293T cells were plated at a density of 0.5×10^6 cells per well. 50 ng per well of the pGL3 and 2 ng per well of *Renilla* reporter constructs were transfected using Fugene6 (Roche Applied Bioscience). The transfected B721.221 and BLCLs were incubated for 5 h, Jurkat cells were incubated for 24 h, and 293T cells were incubated for 48 h at 37 °C in a CO_2 incubator. The cells were lysed and the firefly and *Renilla* luciferase activities were measured using the Dual Luciferase Reporter Assay System (Promega) and a multidetection microplate reader (fluoSTAR Omega, BMG LABTECH). Luciferase activity of each reporter construct is calculated as fold change relative to the activity of pGL3-basic construct lacking a promoter as well as *HLA-C* 3′ UTR ((luciferase test / average luciferase basic) × (average *Renilla* basic / *Renilla* test)) as previously described[22]. The luciferase activity is presented as fold change of RLU ± s.e. (RLU, mean ± s.e.). Firefly luciferase activity was normalized relative to the *Renilla* luciferase activity for each transfection. All experiments were performed with six replicates and repeated in at least three independent experiments. Similarly, luciferase assays were carried out by co-transfecting 20 pmol per well of oligonucleotide mimics or inhibitors of miR-148a and miR-148b (Dharmacon) or the negative control along with the reporter plasmids. The negative control represents a universal control for both inhibitors and mimics that are based on the sequences of miRNAs in *C. elegans*. These negative control miRNAs have been confirmed to have minimal sequence identity with miRNAs in human, mouse and rat.

Antibodies and flow cytometry. Before flow cytometry, BLCLs were plated at a density of 1×10^6 cells per well in a 96-well plate and transfected with 20 pmol per well (1 μM final concentration) of either the mimic or inhibitor of miR-148a, miR-148b, or negative control using AMAXA nucleofector (Lonza). The cells were incubated for 24–48 h in a 37 °C CO_2 incubator before determining cell-surface expression for HLA-C or HLA-B.

HLA-C surface expression on BLCLs and peripheral blood lymphocytes was then analysed by staining with the DT9 antibody (provided by V. Braud)[23] or L31 antibody (MediaPharma SRL) followed by a secondary PE-conjugated anti-mouse IgG (Sigma-Aldrich) and collected on an LSRII flow cytometer (BD Biosciences). Isotype controls were obtained from Sigma-Aldrich. HLA-B surface expression on BLCLs was detected using FITC-conjugated anti-Bw4 and anti-Bw6 antibodies (One Lambda) and FITC-conjugated isotype controls. The histograms were plotted using the FlowJo software version 7.5 (TreeStar). L31 binds a linear epitope on HLA class I heavy chain alleles carrying an aromatic residue (Y/F) at position 67 (ref. 27). These include all HLA-C allotypes used in experiments involving L31 in this study and a few crossreacting HLA-B allotypes (HLA B7, B8, B35, B51), which were excluded from samples chosen for our studies involving L31. L31 binds heavy chains that are free of β2m, denatured or unfolded[27].

Lymphocyte separation and miR-148a inhibition. Peripheral blood was obtained from healthy donors and lymphocytes were separated using lymphocyte separation medium as per manufacturer's instructions (Lonza). The lymphocytes were suspended in antibiotic-free RPMI 1640 medium supplemented with 10% FBS and glutamine. PNA inhibitor of miR-148a was added at a final concentration of 0.1 μM for 1×10^6 cells in 200 μl medium and incubated for 48 h. The cells were then analysed for surface expression of HLA-C using the DT9 antibody as described earlier.

Immunoblot analysis. Western blot analysis was performed using protein lysates from 10×10^6 BLCL cell lines transfected with mimic, inhibitor of miR-148a or negative control (1 μM final concentration, transfection conditions as described earlier). The cells were incubated for 48 h before the total cell lysates were prepared. The lysates were subjected to SDS–polyacrylamide (Invitrogen) gel electrophoresis and transferred to Immobilon-P membranes (Millipore). The membranes were blocked at room temperature (20–25 °C) in 5% milk. The membranes were then probed for HLA-C using murine monoclonal antibody L31 (MediaPharma SRL) and β-actin (Santa Cruz Biotechnology). Horseradish peroxidase-conjugated goat anti-mouse IgG (Santa Cruz) was used as the secondary antibody.

Isolation of RNA and miRNA expression analysis by real-time PCR. Total RNA from B721.221 and BLCLs was extracted using the total RNA purification kit (Norgen). Relative quantification of miR-148a, miR-148b, miR-152 and miR-657 was performed using the Taqman real-time PCR assay (Applied Biosystems). RNU48, a small nucleolar RNA that shows abundance and relatively stable expression in both the cell lines used in this study, served as an endogenous control.

Statistical analyses. The effect of polymorphisms in the 3′ UTR of *HLA-C* on HIV viral load control was determined by categorical analyses of the comparison groups: HIV infected controllers versus noncontrollers. Stepwise logistic regression was used to test whether any *HLA* (*HLA-A*, *B* and *C*) allele with a frequency of 1% or greater had an effect in addition to the *HLA-C* 3′ UTR 263 effect. We used SAS procedure PROC LOGISTIC with STEPWISE selection. Sixty-three *HLA* alleles were included in the model. A likelihood ratio test was used to select alleles into the regression model. The significance level for entry of a variable into the model was $P < 0.05$. The *HLA-C* 3′ UTR 263del/del versus 263ins/ins comparison and five *HLA* alleles met the significance level and were selected into the model. SAS9.1 (SAS Institute) was used for data management and statistical analyses. PROC FREQ was used to compute frequencies in each test group. PROC LOGISTIC was used to calculate odds ratios and 95% confidence intervals. A two sided P value of < 0.01 was considered statistically significant. PROC ALLELE was used to compute linkage disequilibrium between the 263del/ins SNP and *HLA-C*/*HLA-B* alleles.

24. Tilanus, M. G. J. in *Immunobiology of the Human MHC. Proceedings of the 13th International Histocompatibility Workshop and Conference* (ed. Hansen, J.) Vol. 1, 304–416 (IHWG Press, 2002).
25. Shimizu, Y. & DeMars, R. Production of human cells expressing individual transferred *HLA-A,-B,-C* genes using an *HLA-A,-B,-C* null human cell line. *J. Immunol.* **142,** 3320–3328 (1989).
26. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **36,** 59–72 (1977).
27. Setini, A. *et al.* Distinctive features of the α_1-domain alpha helix of HLA-C heavy chains free of β_2-microglobulin. *Hum. Immunol.* **46,** 69–81 (1996).

doi:10.1038/nature09914

Supplementary Figure 1. Predicted miRNA targeting in the 3'UTR of *HLA-Cw*0702* and *Cw*0602*. We identified 26 different putative miRNA binding sites in the 3'UTR of *HLA-C* using Miranda[27], PicTar[28] and Sanger miRbase target prediction programs. RNAhybrid[29] was used to select the best miRNA-target interactions. The miR-148a/miR-148b binding site at position 250-272 and the miR-657 binding site at position 290-312 showed the strongest potential interactions as predicted by lower MFE. The predicted miRNA and miRNA-target regions are shown.

```
           -10       0         10        20        30        40        50        60        70        80        90
           .|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|...
Cw*0702    TCTGTAGCCTGAGACAGCTGCCTGTGTGGGACTGAGATGCAGGATTTCTTCACACCTCTCCTTTGTGACTTCAAGAGCCTCTGGCATCTCTTTCTGCAAAGGCACCTGAA
Cw*0701    ...............................................................................................GT.....
Cw*0102    -.............-..............................................................................T.....
Cw*0303    ...............................................................................................T.....
Cw*0304    ...............................................................................................T.....
Cw*0401    ...............................................................................................T.....
Cw*1402    .............-.............................................................................T.....
Cw*0202    ..............................................................................A........T.....
Cw*0501    ...........................................................T...................................T.....
Cw*0602    ...............................................................................................T.....
Cw*0802    ...........................................................T...................................T.....
Cw*1203    ...............................................................................................T.....
Cw*1502    -.............-.............................................................................T.....
Cw*1601    ...............................................................................................T.....

           100       110       120       130       140       150       160       170       180       190       200
           .|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|...
Cw*0702    TGTGTCTGCGTTCCTGTTAGCATAATGTGAGGAGGTGGAGAGACAGCCCACCCCCGTGTCCACCGTGACCCCTGTCCCCACACTGACCTGTGTTCCCTCCCCGATCATCT
Cw*0701    ...............................................................................................
Cw*0102    ...............................................................................................
Cw*0303    ....................................C..........................................................
Cw*0304    ...............................................................................................
Cw*0401    ...............................................................................................
Cw*1402    ...............................................................................................
Cw*0202    ...............................................................................................
Cw*0501    ...............................................................................................
Cw*0602    ...............................................................................................
Cw*0802    ...............................................................................................
Cw*1203    ...............................................................................................
Cw*1502    ...............................................................................................
Cw*1601    ..C............................................................................................

           210       220       230       240       250       260       270       280       290       300       310
           .|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|...
Cw*0702    TTCCTGTTCCAGAGAGGTGGGGCTGGATGTCTCCATCTCTGTCTCAAATTCATGGTGCACTGAGCTGCAACTTCTTACTTCCCTAATGAAGTTAAGAACCTGAATATAAA
Cw*0701    ...............................................................................................
Cw*0102    ...............A.....-.........................C..........G.....................................
Cw*0303    ...............A.....-.........................C..........G...................................G...............
Cw*0304    ...............A.....-.........................C..........G...................................G...............
Cw*0401    ...............A.....-.........................C..........G.....................................
Cw*1402    ...............A.....-.........................C..........G.......................................G..........
Cw*0202    ...............A.....-.........................C..T.C.-..T..................----.....C....AA......T..........
Cw*0501    ...............A.....-.........................C..T.C.-..T...........................C....AA......T..........
Cw*0602    ...............A.....-.........................C..T.C.-..T...........................C....AA......T..........
Cw*0802    ...............A.....-.........................C..T.C.-..T...........................C....AA......T..........
Cw*1203    ...............A.....-.........................C..T.C.-..T...........................C....AA......T..........
Cw*1502    ...............A.....-.........................C..T.C.-..T...........................C....AA......T..........
Cw*1601    ...............A.....-.........................C..T.C.-..T...........................C....AA......T..........

           320       330       340       350       360       370       380       390       400       410       420
           .|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Cw*0702    TTTGTGTTCTCAAATATTTGCTATGAAGCGTTGATGGATTAATTAAATAAGTCAATTCCTAGAAGTTGAGAGAGCAAATAAAGACCTGAGAACCTTCCAGAA
Cw*0701    ...............................................................................................
Cw*0102    .....T......................G...............................G.......................................
Cw*0303    .....T......................G...............................G................................T.......
Cw*0304    .....T......................G...............................G................................T.......
Cw*0401    .....T......................G...............................G.......................................
Cw*1402    .....T......................G...............................G.......................................
Cw*0202    .....T....................GAG...............................G.......................................
Cw*0501    .....T....................GAG...............................G.......................................
Cw*0602    .....T....................GAG...............................G.......................................
Cw*0802    .....T....................GAG...............................G.......................................
Cw*1203    .....T....................GAG...............................G.......................................
Cw*1502    .....T....................GAG...............................G.......................................
Cw*1601    .....T....................GAG.........................C...G.......................................
```

Supplementary Figure 2. *HLA-C* 3'UTR sequence alignments. Alignment of full length 3'UTR sequences of the common *HLA-C* alleles is shown. The 3' UTR was amplified by PCR and sequenced using genomic DNA from at least 10 individuals homozygous for each of the *HLA-C* alleles, except for *Cw*1402*, for which 3 homozygous individuals were genotyped. Identical nucleotides are shown as dots and deletions are indicated by introducing hyphens (-).

doi:10.1038/nature09914

a

```
           240       250       260       270       280       290       300       310
           .|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Cw*0702    CTCCATCTCTGTCTCAAATTCATGGTGCACTGAGCTGCAACTTCTTACTTCCCTAATGAAGTTAAGAACCTGAATAT
Cw*0701    ..........................................................................
Cw*0102    ................C.........G.................................................
Cw*0303    ................C.........G.................................G............
Cw*0304    ................C.........G.................................G............
Cw*0401    ................C.........G.................................................
Cw*1402    ................C.........G.......................................G........
Cw*0202    ................C..T.C.-..T.................----.....C....AA......T........
Cw*0501    ................C..T.C.-..T.........................C....AA......T........
Cw*0602    ................C..T.C.-..T.........................C....AA......T........
Cw*0802    ................C..T.C.-..T.........................C....AA......T........
Cw*1203    ................C..T.C.-..T.........................C....AA......T........
Cw*1502    ................C..T.C.-..T.........................C....AA......T........
Cw*1601    ...............C..T.C.-..T.....................C....AA......T........
```

miR-148a
miR-148b

miR-657

b

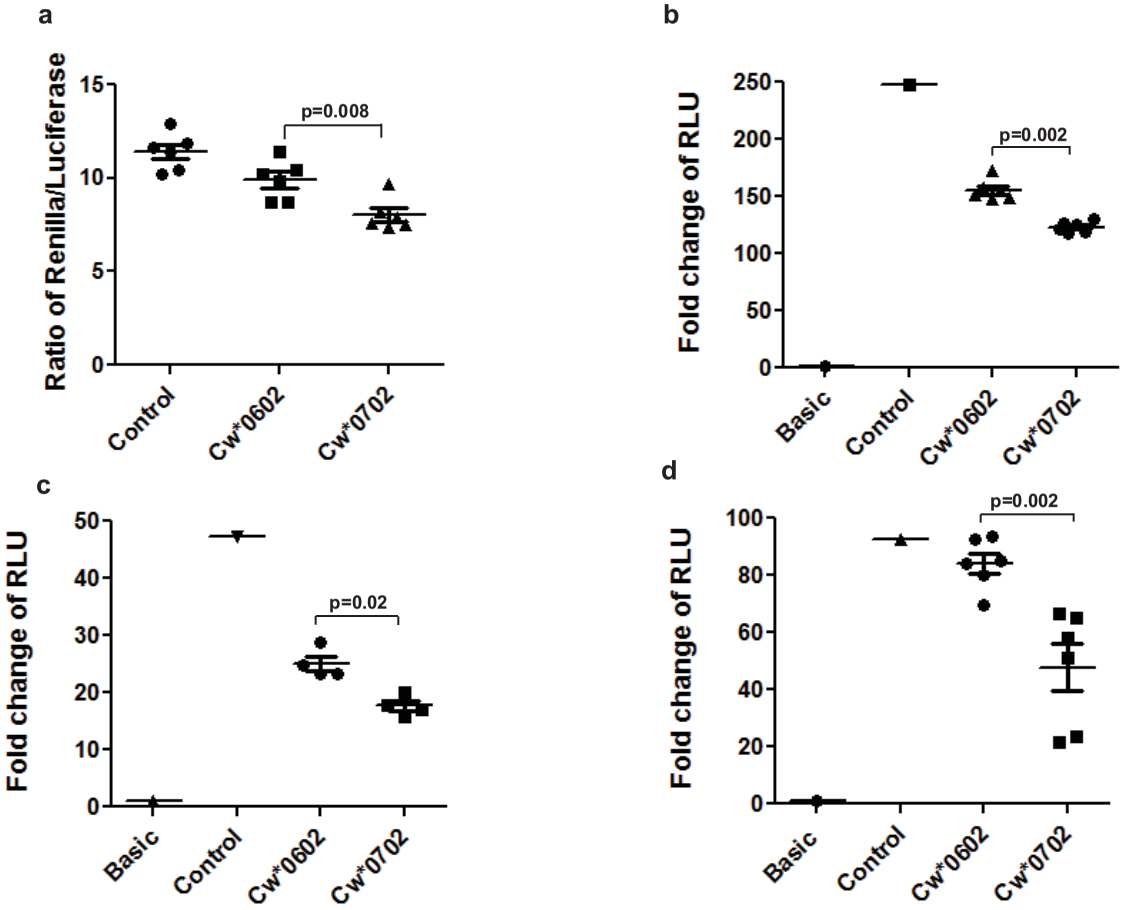
```
Cw*0702      5' U    A   G            3'
                 CAAA UUC  UGGUGCACUGA          MFE -24.7
                 GUUU AAG  AUCACGUGACU
miR-148a     3' U    C   AC           5`

Cw*0303      5'      C   A            G 3'
                     UUC  UGGUGCGCUGA           MFE -23.0
                     AAG  AUCACGUGACU
miR-148a     3' UGUUUC   AC             5'

Cw*0602      5'       A  U  C         G 3'
                       CU UA GUGUACUGA          MFE -19.0
                       GA AU CACGUGACU
miR-148a     3'UGUUUCAA  C              5`

Cw*0702      5'U        A            G  3'
                CAAA UUC  UGGUGCACUGA           MFE -24.4
                GUUU AAG  ACUACGUGACU
miR-148b     3'U    C   AC              5`

Cw*0303      5'       C   A            G 3'
                       UUC  UGGUGCGCUGA         MFE -22.7
                       AAG  ACUACGUGACU
miR-148b     3' UGUUUC    AC             5`

Cw*0602      5'             CG        G   3'
                              UGUACUGA          MFE -17.0
                              ACGUGACU
miR-148b     3' UGUUUCAAGACACU            5'
```

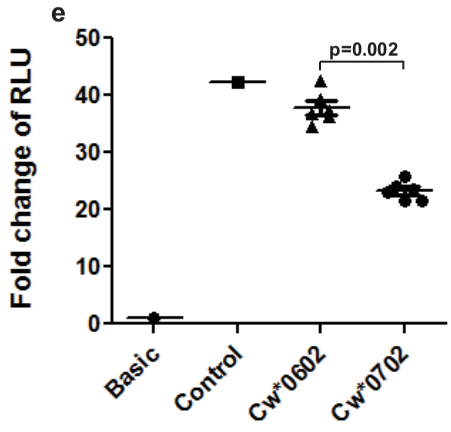
Supplementary Figure 3. Polymorphisms in the putative miRNA binding sites in the 3'UTR of *HLA-C*. **a**. Partial 3'UTR sequences of *HLA-C* alleles are aligned relative to *Cw*0702* showing a pattern of linked polymorphic sites in the putative binding regions of miR-148a/miR-148b and miR-657. Identical nucleotides are shown as dots and deletions are indicated by hyphens (-) for optimal alignment. Variations that disrupt miRNA pairing with the 3'UTR of the high expression *HLA-C* alleles are shown in bold. DNAs from at least 10 individuals homozygous for each allele were sequenced, except for *Cw*1402,* where only 3 individuals homozygous for this allele were sequenced. Sequences across individuals carrying the same allele were identical. **b.** Sequences of RNA hybrids between mature miR-148a/miR-148b and *HLA-C* 3'UTR sequences are shown. The genotype marked by a deletion at position 263 in the 3'UTR of *Cw*0602* and other linked variations (259T, 261C, 266T) within the miRNA binding region (underlined) disrupt pairing with miR-148a/miR-148b. The minimum free energy (MFE) of the RNA duplex shown on the right was determined by RNA hybrid software.

doi:10.1038/nature09914


```
Cw*0702   5'   AU  A UUA        A     3'
           CCUA  GA G   AGAACCUG            MFE: -19.7 kcal/mol
           GGAU  CU C   UCUUGGAC
miR-657   3'   CU  C CAC        GG    5'

Cw*0602   5'     CU  AAAUA     U   A  3'
             CCUA  GA     AGAA  CUG        MFE: -17.4 kcal/mol
             GGAU  CU     UCUU  GAC
miR-657   3'     CU  CCCAC     G   GG 5'

307C>T    5'   AU  A UUA    U   A     3'
           CCUA  GA G   AGAA CUG            MFE: -17.1 kcal/mol
           GGAU  CU C   UCUU GAC
miR-657   3'   CU  C CAC    G   GG    5'

307C>G    5'     AU  A UUA    G   A   3'
             CCUA  GA G   AGAA CUG          MFE: -15.2 kcal/mol
             GGAU  CU C   UCUU GAC
miR-657   3'     CU  C CAC    G   GG  5'
```

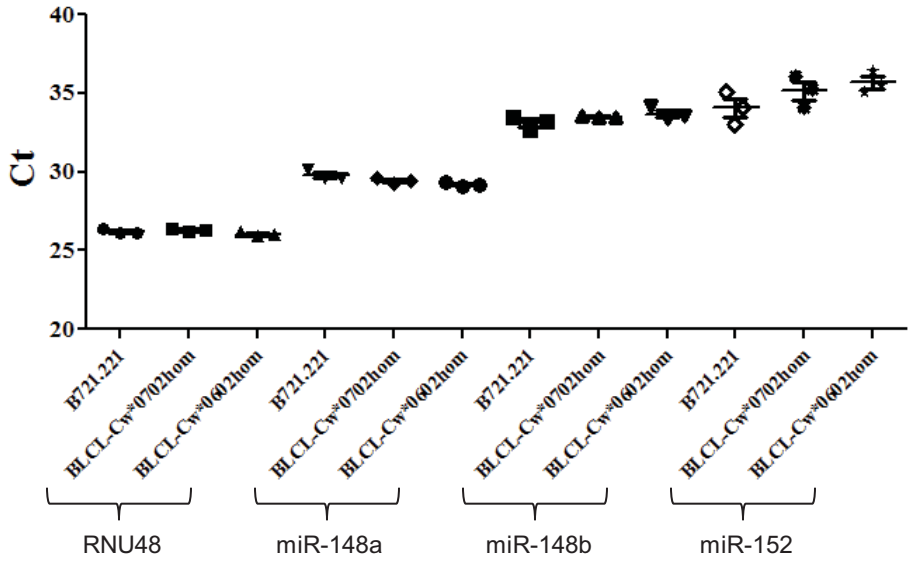
Supplementary Figure 4. Predicted binding of miR-657 to the *HLA-C* 3'UTR. A schematic representation of the RNA hybrid structure of the mature miR-657 binding to the 3'UTR of *Cw*0702* and *Cw*0602* is shown. The 307C/T/G polymorphism is highlighted. This polymorphism is predicted to disrupt binding of miR-657 to *Cw*0602*. The minimum free energy (MFE) of the RNA duplex was analyzed by RNA hybrid.





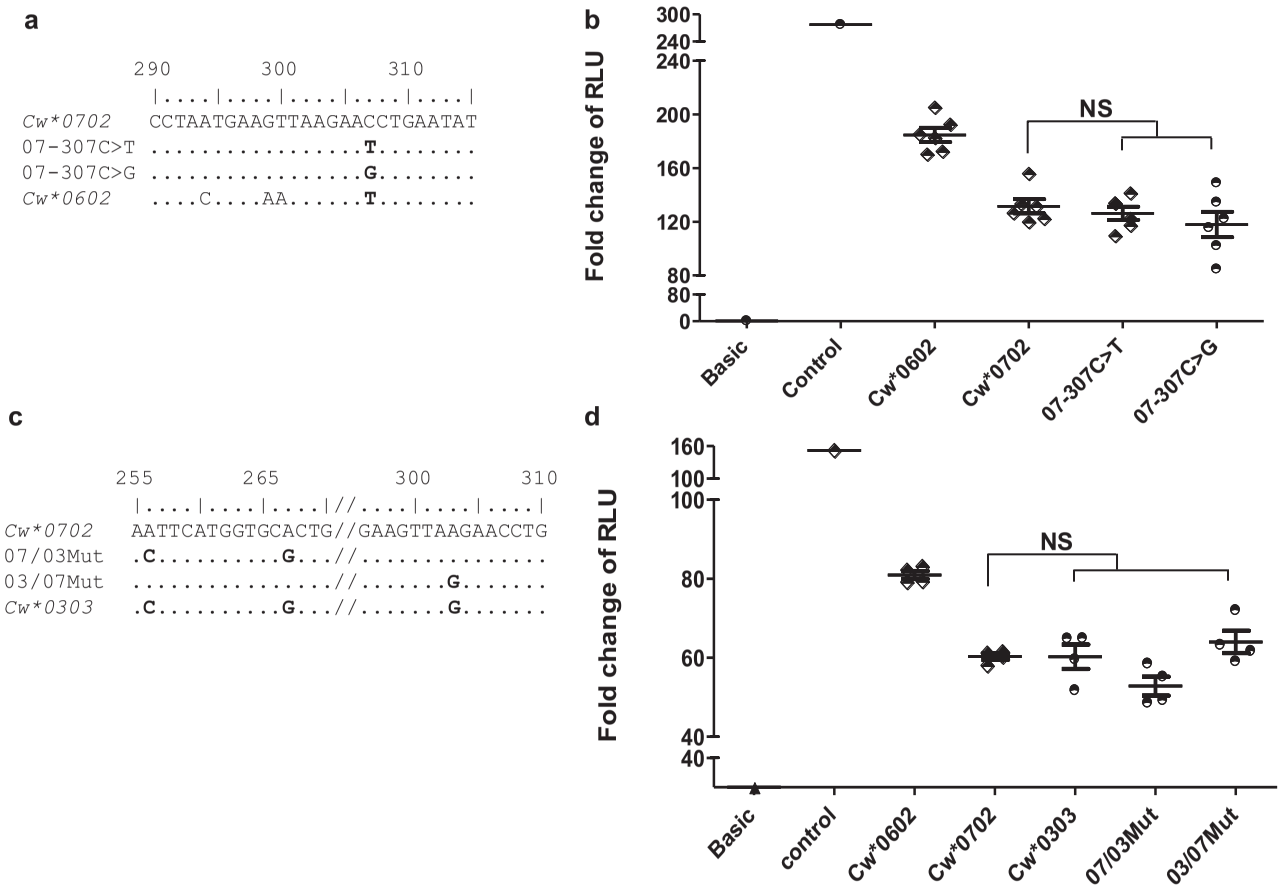
Supplementary Figure 5. 3'UTR mediated gene regulation is consistent across different vector systems and cell lines. a. 3'UTRs of *Cw*0602* and *Cw*0702* were cloned downstream of Renilla in a psicheck2 dual reporter vector. The cloned UTRs were then transfected into the B721.221 cell line and luciferase activity was determined after 5 hours. The luciferase activity is presented as the ratio of relative light units (RLU) of Renilla and firefly luciferase, which are encoded in the same vector to normalize variation in transfection efficiency. **b-e**. The plots show the fold change in luciferase activity in B721.221 (**b**), BLCL-Cw*0702Hom (EBV transformed B cell line derived from an *HLA-Cw*0702* homozygous individual) (**c**), HEK293T (human embryonic kidney cell line) (**d**), and Jurkat (T cell line) (**e**) when transfected with the basic or control pGL3 vectors, or with the pGL3 vectors containing the 3'UTR of *Cw*0602* or *Cw*0702*. The normalized luciferase activity of reporters is plotted as fold change of RLU. Non-parametric Wilcoxon-Mann-Whitney tests were used for statistical comparisons and two tailed p values are indicated. The data represent 4-6 replicates in each experimental group as shown. The mean ±SE are depicted as horizontal and vertical bars, respectively, for each group and one of three comparable experiments performed is shown.

doi:10.1038/nature09914

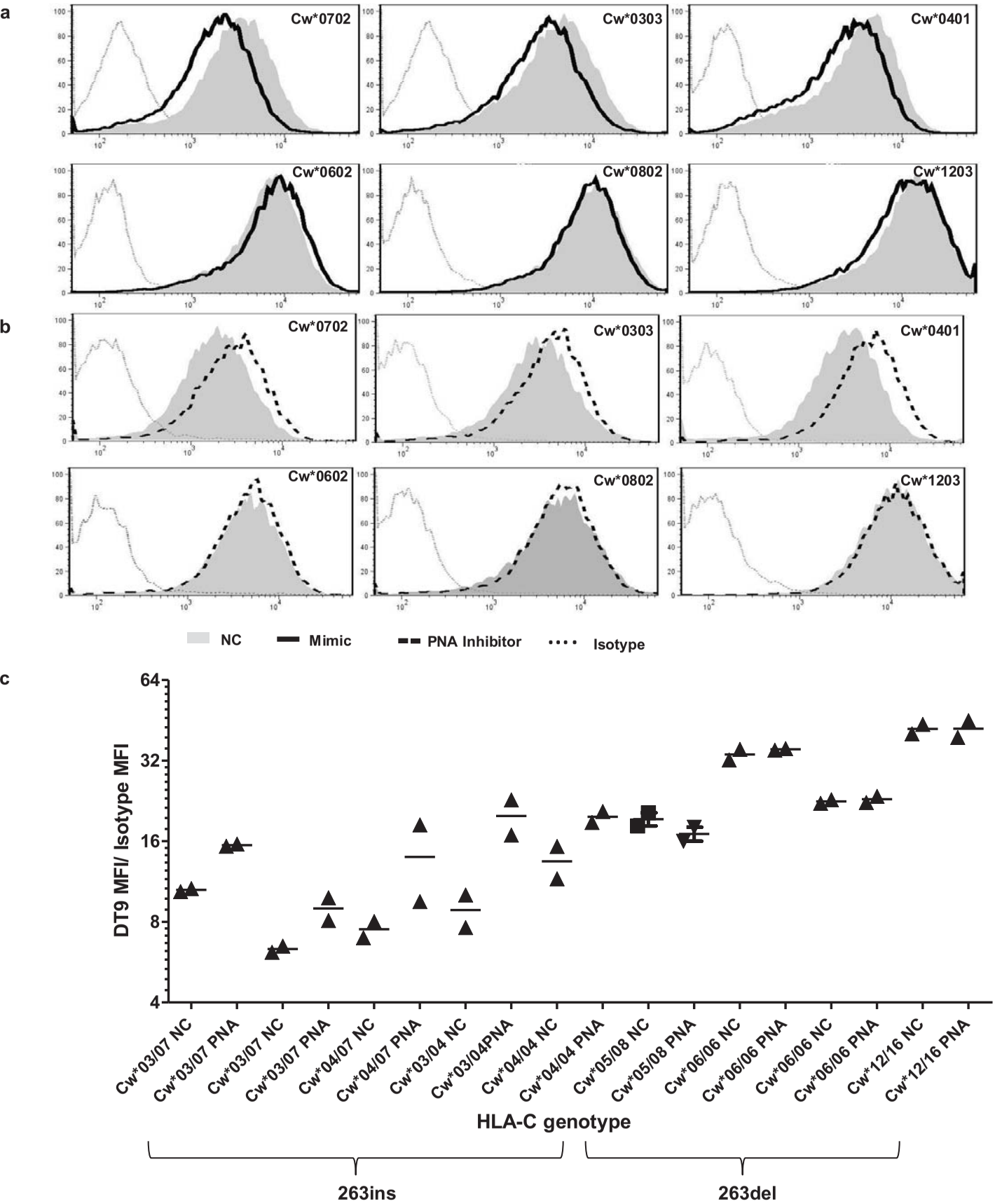


Supplementary Figure 6. miR-148b and miR-152 expression levels are lower than that of miR-148a. Endogenous miR-148a, miR-148b and miR-152 expression levels were estimated in the total RNA pool of three cell lines (B721.221, BLCL -Cw*0702Hom, BLCL-Cw*0602Hom) using a *Taq*man real time RT-PCR assay. Threshold cycles of amplification (C_t) are plotted. RNU48, which is expressed at constitutively high levels, served as an endogenous RNA control. The data represent three replicates in each group. The Mean ±SE are depicted as horizontal and vertical bars, respectively, for each group.

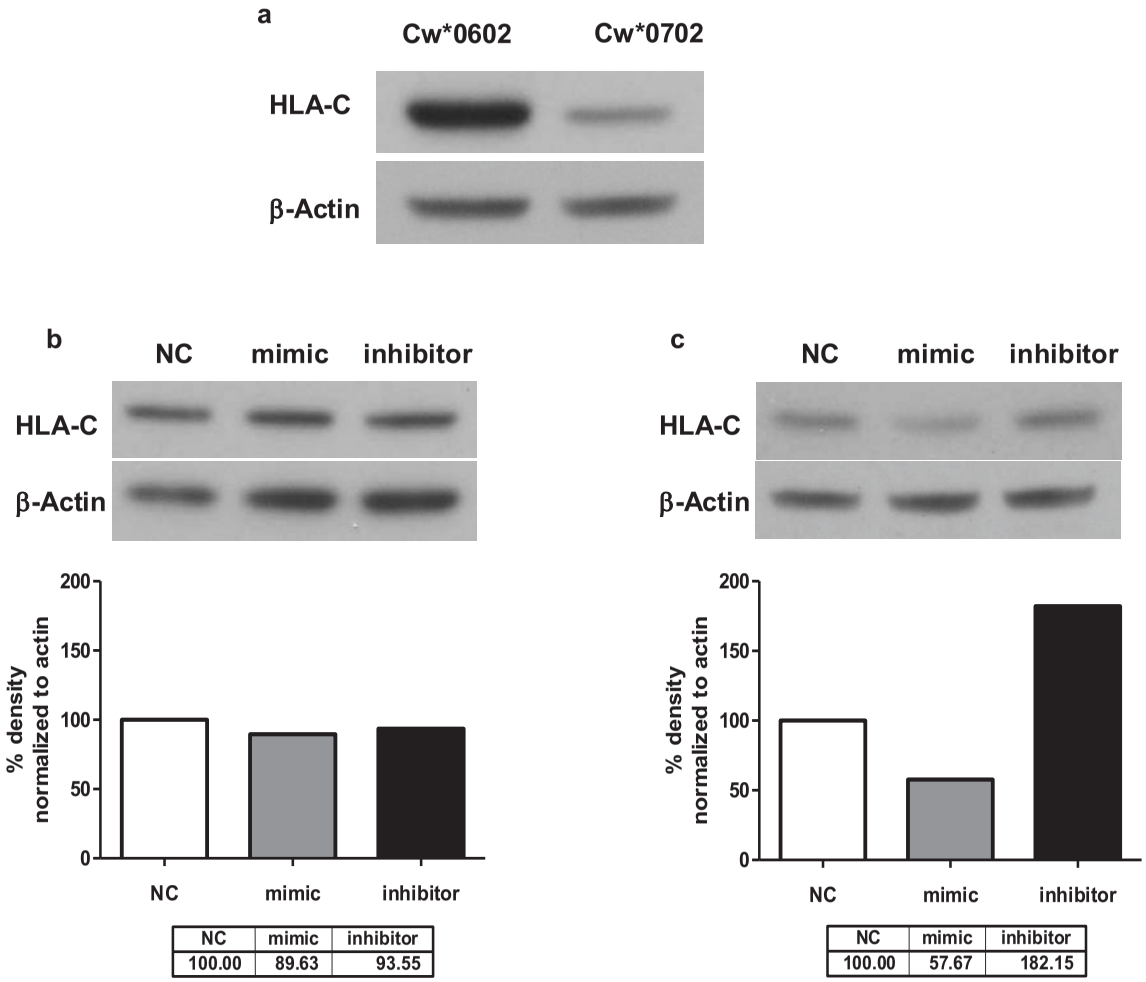
doi:10.1038/nature09914



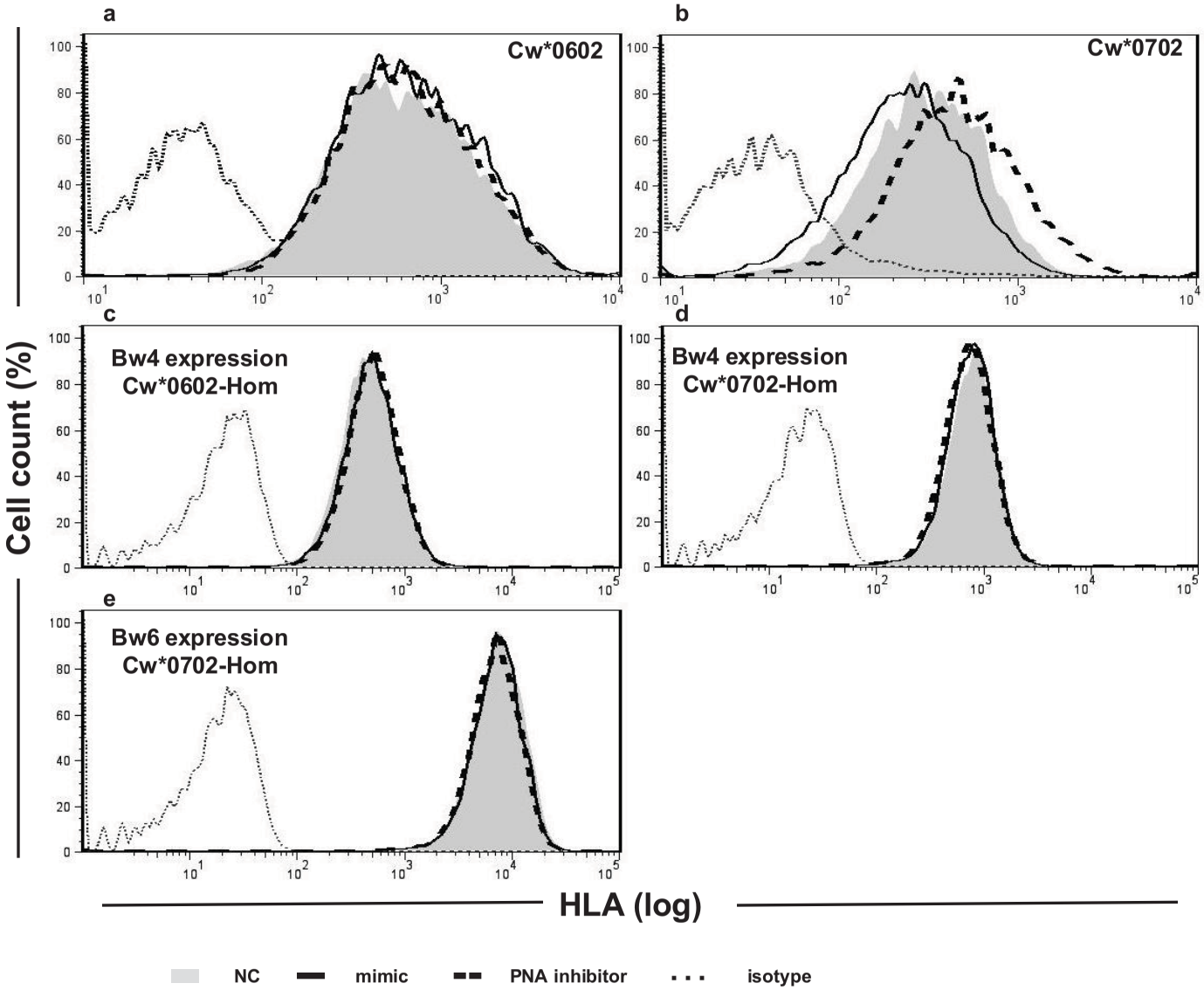
Supplementary Figure 7. Polymorphisms at positions at 256, 267, and 307 in the *HLA-C* 3'UTR do not affect post-transcriptional gene regulation. **a.** Alignment of partial 3'UTR sequences showing mutations that were introduced at position 307 (C>T and C>G) into the 3'UTR sequence of *HLA-Cw*0702* and cloned into a luciferase reporter. **b.** The graph shows the fold change of luciferase activity (Mean ± SE) of the 307C>T (*07-307C>T*) and 307C>G (*07-307C>G*) 3'UTR reporters relative to the wild type *Cw*0702* 3'UTR reporter. **c**. Alignment of partial 3'UTR sequences showing mutations introduced at positions 256 and 267 into the 3'UTR sequence of *HLA-Cw*0702* and *HLA-Cw*0303* and cloned into a luciferase reporter. 07/03Mut is identical to *Cw*0702* except at positions 256 and 267 where it represents the variants characteristic of *Cw*0303*. 03/07Mut is identical to *Cw*0303* except at positions 256 and 267 where it represents the variants characteristic of *Cw*0702*. **d**. The plot shows the fold change of luciferase activity (Mean ± SE) of the A256C/A267G mutant of *Cw*0702* (*07/03 Mut*), C256A/G267A mutant of *Cw*0303 (03/07Mut)*, wild type *Cw*0602*, wild type *Cw*0702* and wild type *Cw*0303* 3'UTR reporters. The data represent six replicates in each experimental group. The Mean ±SE are depicted as horizontal and vertical bars, respectively, for each group, and one of three comparable experiments performed is shown. Non-parametric Wilcoxon-Mann-Whitney tests were used for statistical comparisons. ns = not significant.



Supplementary Figure 8. **Effect of mimic and inhibitor of miR-148a on a panel of primary lymphocytes and BLCL. a-b.** Histograms showing the effect of HLA-C cell surface expression on various *HLA-C* homozygous BLCL after transfection with mimic (**a**) or PNA inhibitor of miR-148a (**b**). A panel of 6 BLCL homozygous for either 263I (*Cw*0702, Cw*0303*, *Cw*0401*) or 263D alleles (*Cw*0602, Cw*0802*, *Cw*1203*) were each tested for the effects of mimic and inhibitor. **c.** Primary peripheral blood lymphocytes homozygous for 263I or 263D were treated with either PNA inhibitor of miR-148a or NC PNA. HLA-C expression was analyzed using the DT9 antibody and is expressed as the ratio of MFI of DT9/MFI of isotype control. NC miRNA, which does not bind to the 3'UTR of *HLA-C*, was included in each experiment.



Supplementary Figure 9. Effect of mimic and inhibitor of miR-148a on total protein levels of HLA-C. a. Total protein levels of HLA-C as determined by L31 mAb western blot of the cell lysates from BLCL homozygous for *Cw*0602* and *Cw*0702* are shown. L31 antibody recognizes heavy chain conformers of HLA-C[30] (along with a limited number of HLA-B allotypes that were selected against to ensure specificity of L31 binding). **b-c.** Changes in the total protein as determined by L31 western blot of HLA-C in the cell lysates of BLCL homozygous for *Cw*0602* (**b**) or *Cw*0702* (**c**) after treatment with NC, mimic or inhibitor of miR-148a. The band intensities are measured by densitometry, and changes in the expression are normalized to actin and plotted as percent change relative to treatment with NC.



Supplementary Figure 10. Effect of mimic and inhibitor of miR-148a on surface expression of free heavy chains of HLA-C and HLA-B. The histograms show cell surface expression of HLA-C or -B on *HLA-C* homozygous BLCL. In each plot, NC miRNA that does not bind to the 3'UTR of *HLA-C* or *-B* was included. **a-b**. HLA-C expression was measured by L31 binding of Cw*0602 on BLCL (Cw*0602-Hom) (**a)** or Cw*0702 (Cw*0702Hom) **(b)** transfected with either a mimic or inhibitor of miR-148a. **c-e**. Effects of mimic or inhibitor of miR-148a on surface HLA-B expression was measured by mAbs that recognize HLA-Bw4 allotypes, including B*5701 and B*5802 (present on Cw*0602-Hom) **(c)** and B*4901 (present on BLCL-Cw*0702Hom) (**d**), or Bw-6 allotypes, including B*1801 (present in Cw*0702-Hom) (**e**).

Supplementary Table S1. Frequencies of the independent variables in the logistic regression model by stepwise selection of *HLA-C* 3'UTR del/ins and *HLA-A*, *-B*, *-C* alleles.

GENOTYPE	VL<2000		VL>10000		GENOTYPE	VL<2000		VL>10000	
N=2527	N	%	N	%	N=1301	N	%	N	%
263del/del	238	28.5	237	14.0	263del/del	238	57.2	237	26.7
263del/ins	421	50.3	805	47.6	263ins/ins	178	42.7	648	73.2
263ins/ins	178	21.3	648	38.3	B*2705	52	12.5	46	5.2
					B*5701	101	24.3	37	4.2
					B*5703	19	4.6	1	0.11
					B*5801	14	3.4	16	1.8
					Cw*1402	19	4.6	20	2.3

VL = viral load; OR = odds ratio; CI = confidence interval; N = 2527. A logistic regression analysis with stepwise selection using the *HLA-C* 3'UTR 263 del/del vs. ins/ins comparison and 63 *HLA-A, -B* and *-C* alleles with ≥1% frequency as independent variables in the model was performed using PROC LOGISTIC (SAS 9.1 version, SAS Institute). Significance level for selecting variables shown in the table was $p < 0.05$. Frequencies of *HLA-C* 3'UTR 263 del/ins and the *HLA* alleles that remained significant in the model are shown.

Supplementary Table S2. Linkage disequilibrium between *HLA-C* alleles and 263 del/ins from the genotypes of European American individuals (N=1760).

HLA-C alleles	3'UTR 263	Frequency	D'	p	-35 allele in strong LD
0102	ins	0.042	1	<0.0001	C
0202	del	0.049	1	<0.0001	C
0302	ins	0.002	1	NS	C
0303	ins	0.044	1.	<0.0001	T
0304	ins	0.067	1	<0.0001	T
0401	ins	0.111	1	<0.0001	T
0501	del	0.079	1	<0.0001	C
0602	del	0.117	1	<0.0001	C
0701	ins	0.144	1	<0.0001	T
0702	ins	0.119	1	<0.0001	T
0704	ins	0.012	1	<0.0001	T
0801	del	0.002	1	NS	C
0802	del	0.047	1	<0.0001	C
0804	del	0.0004	1	NS	C
1202	del	0.015	1	<0.0001	C
1203	del	0.06	1	<0.0001	C
1402	ins	0.014	1	<0.0001	C
1502	del	0.022	1	<0.0001	T
1505	del	0.003	1	NS	T
1506	del	0.0004	1	NS	T
1601	del	0.028	1	<0.0001	T
1602	del	0.002	1	NS	T
1604	del	0.002	1	NS	T
1701	ins	0.005	1	NS	T

D' = linkage disequilibrium; NS = not significant

doi:10.1038/nature09914

Supplementary Table S3. Linkage disequilibrium between *HLA-B* alleles and 263 del/ins from the genotypes of European American individuals (N=2521).

HLA-B alleles	3'UTR 263	Frequency	D'	p
0702	ins	0.11	0.96	<0.0001
0802	ins	0.09	0.99	<0.0001
1302	del	0.03	1	<0.0001
1401	del	0.02	1	<0.0001
1402	del	0.03	1	<0.0001
1501	ins	0.06	0.91	<0.0001
1801	del	0.03	0.37	<0.0001
2705	del	0.02	0.15	0.05
3501	ins	0.05	0.97	<0.0001
3502	ins	0.01	0.91	0.0007
3503	ins	0.01	0.54	0.005
3701	del	0.01	0.96	<0.0001
3801	del	0.02	0.93	<0.0001
3901	del	0.01	0.48	0.001
4001	ins	0.04	0.97	<0.0001
4002	del	0.01	0.60	<0.0001
4402	del	0.07	0.71	<0.0001
4403	del	0.02	0.3	0.0003
4901	ins	0.02	1	<0.0001
5001	del	0.01	0.9	<0.0001
5101	del	0.03	0.21	0.006
5201	del	0.02	0.96	<0.0001
5501	ins	0.01	0.97	0.0002
5701	del	0.06	0.93	0.0002
5801	ins	0.01	1	0.0002

D' = linkage disequilibrium; NS = not significant

Supplementary Table S4. Effect of *HLA-C* 3'UTR 263 on HIV viral load

a. Single point associations for *HLA-A*, *-B*, *-C* loci and *HLA-C* 3'UTR 263 with HIV viral load control

Locus	VL<2000 (N)	VL>10000 (N)	df	p
263del/ins	837	1690	1	2×10^{-26}
HLA-A	834	1690	70	2×10^{-20}
HLA-B	837	1690	98	2×10^{-62}
HLA-C	837	1690	41	1×10^{-47}

df= degree of freedom; SAS genetics procedure PROC CASECONTROL was used to test the single point associations for the *HLA-A*, *-B*, *-C* loci and *HLA-C* 3'UTR 263 variant with HIV viral load control. If each *HLA* class I locus and position 263 in the *HLA-C* 3'UTR are considered individually, each appear to be very significantly associated with HIV control.

b. Comparison of the nested logistic regression models of *HLA-C* 3'UTR and *HLA-A*, *-B*, *-C*

Model	Model terms	Deviance	DF	Model comparison	ΔDeviance	Δ DF	p-value
0		3209.50	2526				
1	263 del/ins[a]	3096.70	2524	Model 1 vs 0	112.80	2	3×10^{-25}
2	*HLA-A*[b]	3111.32	2507	Model 2 vs 0	98.18	19	1×10^{-12}
3	*HLA-B*[c]	2754.71	2499	Model 3 vs 0	454.79	27	3×10^{-79}
4	*HLA-C*[d]	2940.66	2509	Model 4 vs 0	268.84	17	3×10^{-47}
5	*HLA-A,-B,-C*[e]	2633.70	2463	Model 5 vs 0	575.80	63	7×10^{-84}
6	*HLA-A*[b] + del/ins	3015.49	2505	Model 6 vs 2	92.82	2	2×10^{-21}
7	*HLA-B*[c] + del/ins	2748.76	2497	Model 7 vs 3	5.95	2	0.05
8	*HLA-C*[d] + del/ins	2939.13	2507	Model 8 vs 4	1.53	2	0.5
9	*HLA-A,-B,-C*[e] + del/ins	2631.32	2461	Model 9 vs 5	2.38	2	0.3

Df is Df is the Df is degrees of the freedom in the model; N=2527
The Δ Deviance is the difference of deviance between the compared models. A large value (small p-value) indicates the smaller model does not predict the outcome well, thus the smaller model is rejected
The Δ DF is the difference in degrees of the freedom between the compared models
[a] Includes the main effect of *HLA-C* 3'UTR
[b] Includes the main effect of all alleles at *HLA-A*
[c] Includes the main effect of all alleles at *HLA-B*
[d] Includes the main effect of all alleles at *HLA-C*
[e] Includes the main effect of all alleles at *HLA-A*, *HLA-B* and *HLA-C*

Table 4b considers each locus as a whole and attempts to determine whether position 263 has a significant independent effect. The validity of this approach is hampered by the excessively strong LD between position 263 and *HLA-C*/*HLA-B*. Any effect of *HLA-C* as a whole will include the effect of 3'UTR because there is perfect one-way LD between these loci (that is, there has been no recombination between the *HLA-C* coding region and the 3'UTR of this gene; if the *HLA-C* allele is known, then the 3'UTR variant is known). This is a problem with *HLA-B*, as well; the LD makes it difficult to know whether the effect is due to *HLA-B* completely or *HLA-B* with contribution from *HLA-C* 3'UTR. Thus, lack of significance in model 7, 8 and 9 is not unexpected.

Supplementary notes

1. EBV transformed B lymphocytes are known to express free heavy chains[30] and expression of free heavy chains is presumably independent of other accessory interactions in HLA surface presentation. Thus, free heavy chains on the cell surface likely reflect changes in HLA protein expression directly. Western blot analysis was performed on protein lysates from cell lines transfected with mimic, inhibitor of miR-148a or NC using the L31 mAb, which specifically recognizes heavy chain open conformers of HLA-C[30] (along with a limited number of HLA-B allotypes that we selected against for this analysis to ensure HLA-C specificity in L31 binding). While there is no effect of the miR-148a mimic or inhibitor on expression of Cw*0602 free heavy chains (Supplementary Fig. 9a), there is an effect on Cw*0702 free heavy chains on the surface of EBV transformed B cell lines that is consistent with data using the DT9 antibody (Supplementary Fig. 9b). On the other hand, HLA-B surface expression as detected by antibodies specific for the Bw4 or Bw6 motif remained unchanged upon transfection with mimic or inhibitor (Supplementary Fig. 9c-e).
2. The association between the 3'UTR 263del/ins variant and HIV control suggests that high HLA-C expression may provide some protection against HIV. HLA-C allotypes are ligands for inhibitory KIR and their expression levels are known to affect KIR signalling on natural killer cells[31], so differential expression of HLA molecules may alter potentiation of KIR-expressing effector cells. Differential HLA-C expression patterns may also affect the efficiency of antigen presentation to CTL. Unlike HLA-A and HLA-B, HLA-C is not downregulated by Nef and it has been suggested that preservation of HLA-C expression prevents NK cell lysis of infected cells[32] as mediated through

inhibitory receptors. While the virus may have evolved to become stealth to NK cell responses, the expression of HLA-C does leave the infected cell vulnerable to HLA-C restricted CTL responses. Recent data show that high levels of HLA-C expression does indeed exert selection pressure on the virus, as HIV-1 Nef variants from -35CC (which marks 263del/del) individuals have an increased ability to impair both MHC-II antigen presentation and helper T-cell function possibly to counteract an HLA-C mediated immune response [33].

3. The effect of the 263 3'UTR variant appears to be primarily recessive when comparing controllers to noncontrollers, since in a second analysis where the 263del/ins vs. 263ins/ins comparison is considered along with the 63 individual *HLA* class I alleles, 263del/ins vs. 263ins/ins does not remain in the model (data not shown).

4. The members of the Swiss HIV Cohort Study (SHCS) are M. Battegay, E. Bernasconi, J. Böni, H.C. Bucher, P. Bürgisser, A. Calmy, S. Cattacin, M. Cavassini, R. Dubs, M. Egger, L. Elzi, M. Fischer, M. Flepp, A. Fontana, P. Francioli (President of the SHCS), H. Furrer (Chairman of the Clinical and Laboratory Committee), C.A. Fux, M. Gorgievski, H.F. Günthard (Chairman of the Scientific Board), H.H. Hirsch, B. Hirschel, I. Hösli, C. Kahlert, L. Kaiser, U. Karrer, C. Kind, T. Klimkait, B. Ledergerber, G. Martinetti, N. Müller, D. Nadal, F. Paccaud, G. Pantaleo, A. Rauch, S. Regenass, M. Rickenbach (Head of Data Center), C. Rudin (Chairman of the Mother & Child Substudy), P. Schmid, D. Schultze, J. Schüpbach, R. Speck, B.M. de Tejada, P. Taffé, A. Telenti, A. Trkola, P. Vernazza, R. Weber, and S. Yerly.

Supplementary References

27. Betel, D., Wilson, M., Gabow, A., Marks, D.S., & Sander, C. The microRNA.org resource: targets and expression. *Nucleic Acids Res* 36, D 149-153 (2008).

28. Krek, A. *et al*. Combinatorial microRNA target predictions. *Nat Genet* 37, 495-500 (2005).

29. Rehmsmeir, M., Steffen, P., Hochsmann, M. & Giegerich, R. Fast and effective prediction of micron/target duplexes. *RNA* 10, 1507-1517 (2004).

30. Setini, A. *et al.* Distinctive features of the alpha 1-domain alpha helix of HLA-C heavy chains free of beta 2-microglobulin. *Hum Immunol* **46**, 69-81 (1996).

31. Almeida, C.R. & Davis, D.M. Segregation of HLA-C from ICAM-1 at NK cell immune synapses is controlled by its cell surface density. *J Immunol* **177**, 6904-6910 (2006).

32. Cohen, G.B. *et al.* The selective downregulation of class I major histocompatibility complex proteins by HIV-1 protects HIV-infected cells from NK cells. *Immunity* **10**, 661-671 (1999).

33. Specht, A. *et al.* Counteraction of HLA-C-mediated immune control of HIV-1 by Nef. *J Virol* **84**, 7300-7311 (2010).