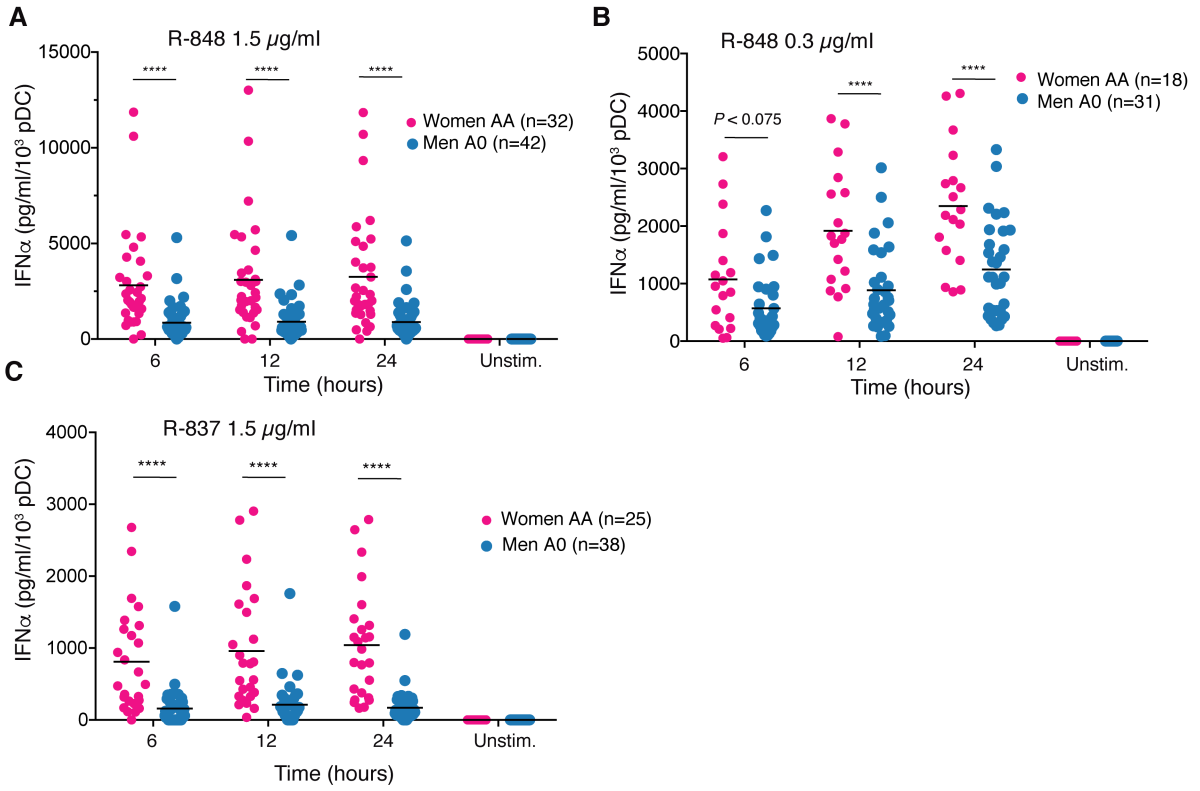
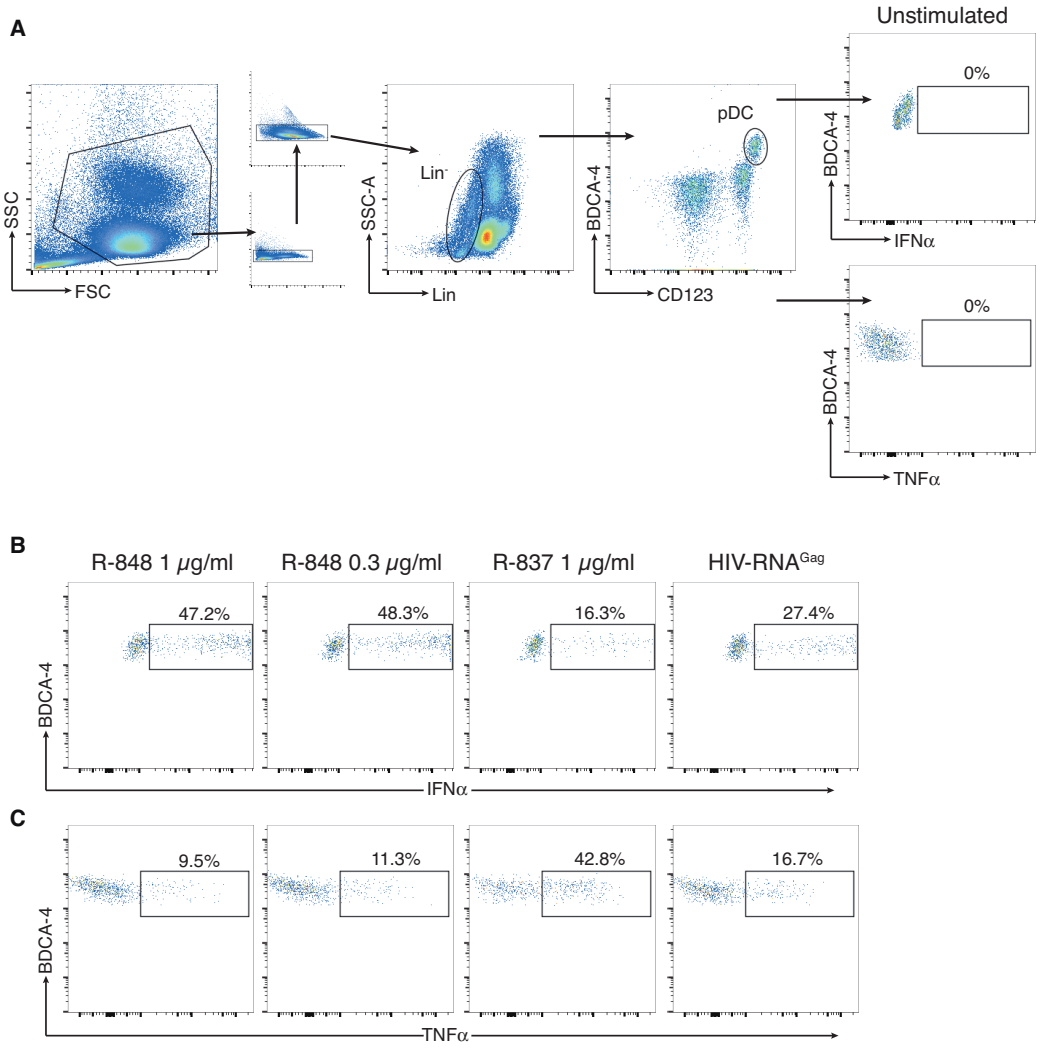
Supplemental Information

TLR7 dosage polymorphism shapes interferogenesis and HIV-1 acute viremia in women

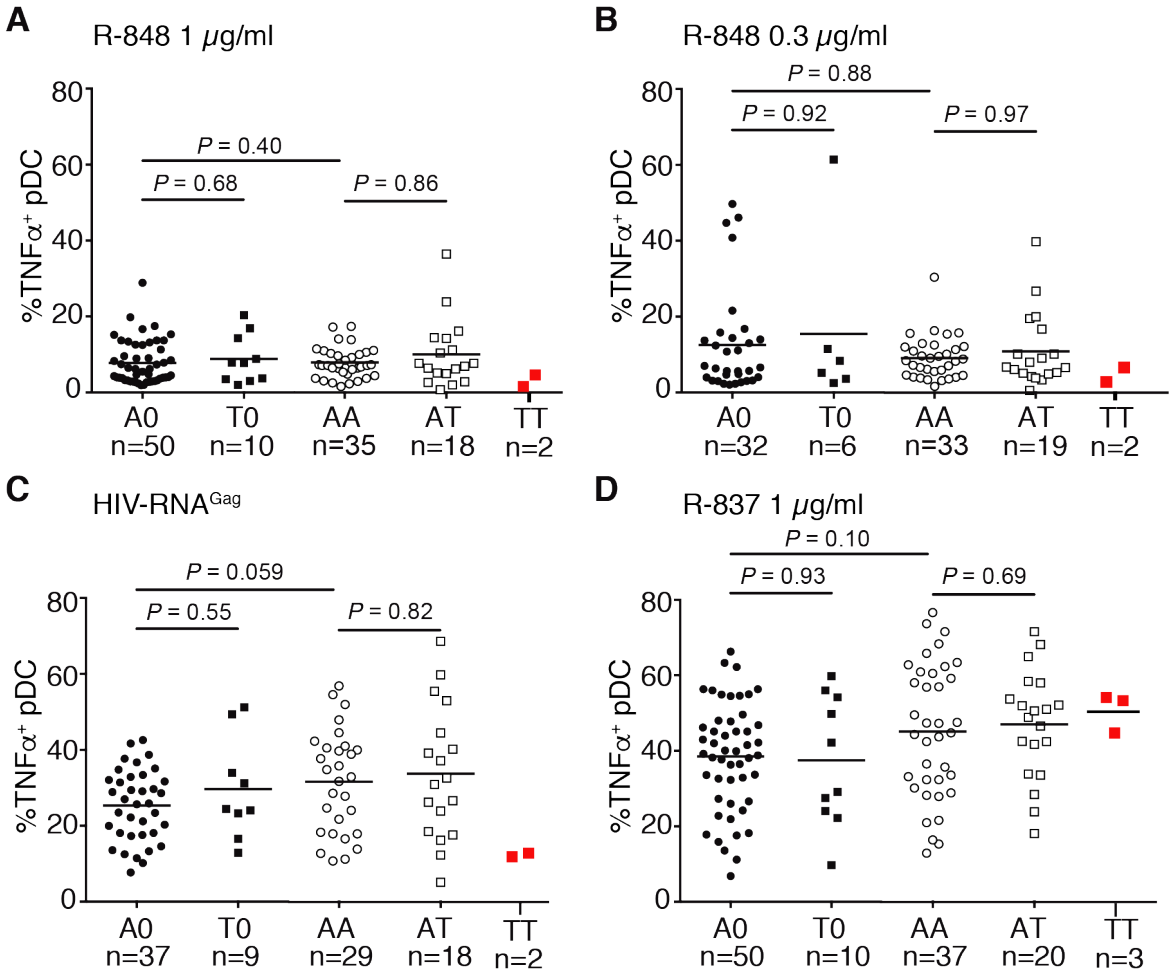
Contents



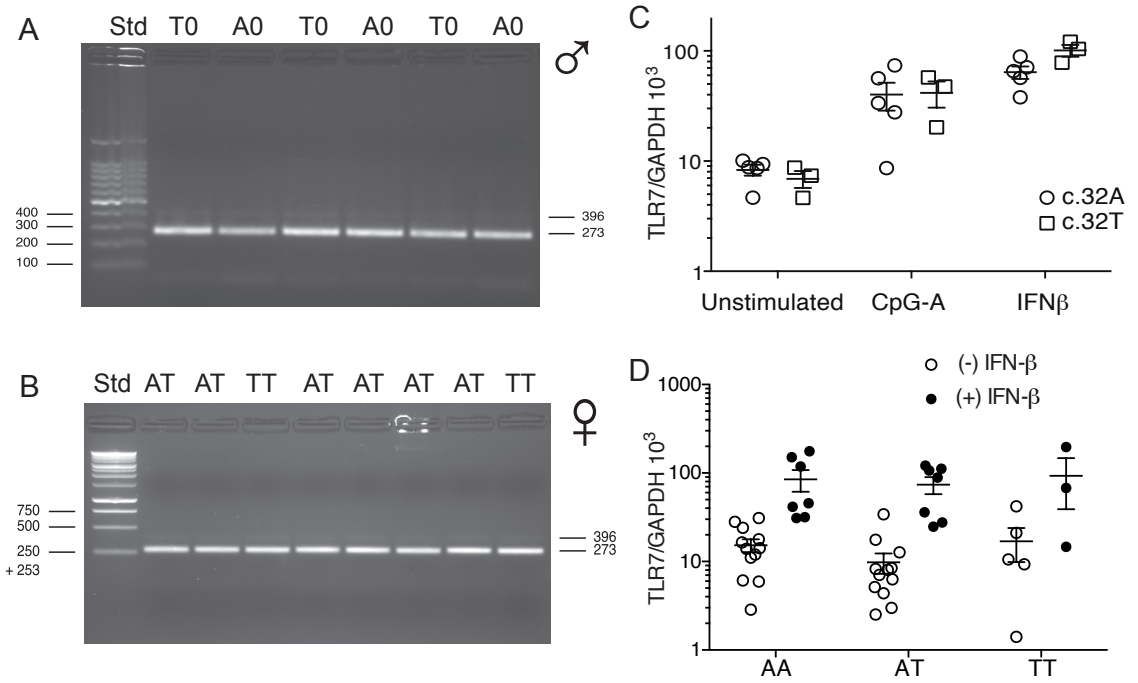
Supplemental Figure 1. Sex-dependent differences in IFN-α production by PBMCs upon TLR7 stimulation. Same data as in Figure 1 but restricted to the homozygous (A/A women) and hemizygous (A/0 men) carriers of the major allele of the *TLR7* SNP, rs179008. Women and men were compared with regard to TLR7-driven production of IFN-α by PBMCs in response to *ex vivo* stimulation with TLR7 ligands as shown. Culture supernatants were sampled 6, 12 and 24 hours after stimulation, or 24 hours for non-stimulated (Unstim.) control cells. IFN-α values (pg/ml) were normalized to flow-cytometrically determined pDC numbers for each donor; see Supplemental Figure 2A for the Lin^{-} $CD123^{+}$ $BDCA\text{-}4^{+}$ gating strategy. Each dot represents one donor; horizontal bars indicate mean values. (**A**) 1.5 μg/ml R-848. (**B**) 0.3 μg/ml R-848. (**C**) 1.5 μg/ml R-837. Groups were compared by a two-way ANOVA test followed by Sidak's multiple comparisons test. ****, $p < 0.0001$; N.D., not detectable, i.e., below the 30 pg/ml detection threshold of the IFN-α ELISA.



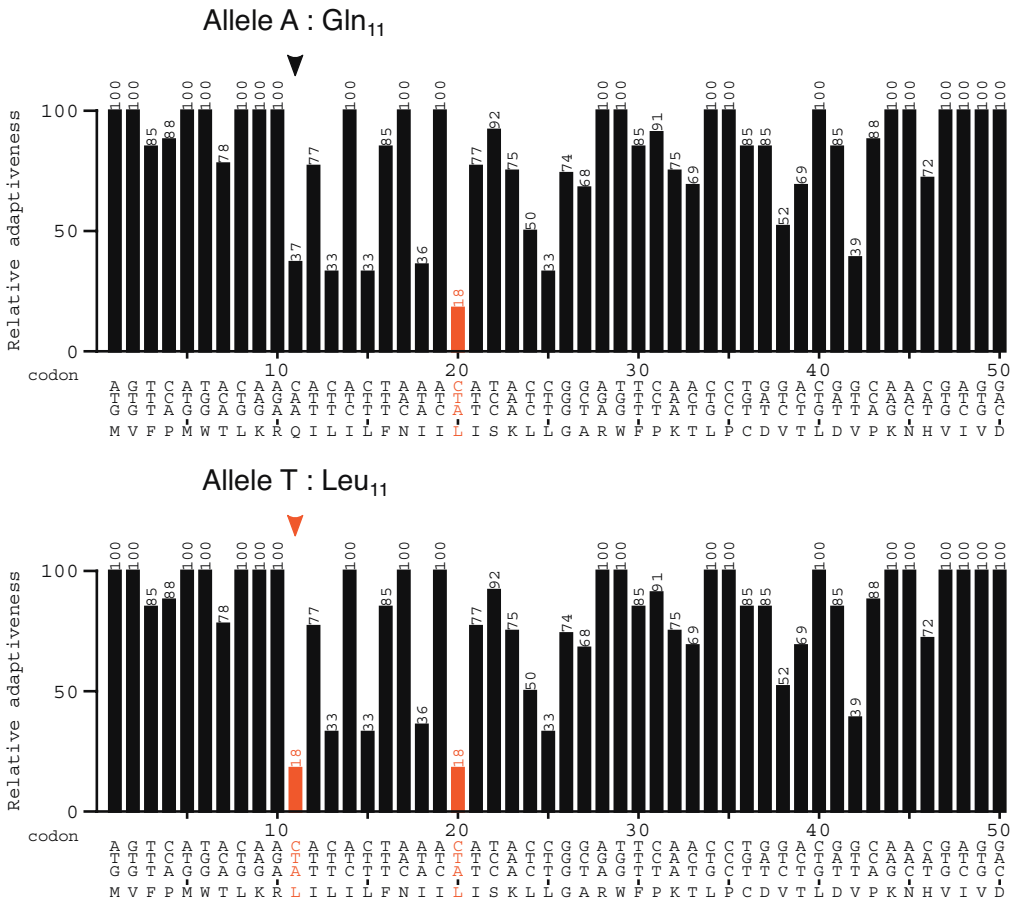
Supplemental Figure 2. Flow cytometric analysis of cytokine-producing pDCs. Freshly isolated PBMCs were stimulated for 5 hours with TLR7 ligands R-848 (1 µg/ml and 0.3 µg/ml), R-837 (1 µg/ml) or HIV-1-derived $Gag_{RNA1166}$, in the presence of brefeldin A for the final 3 hours. The cells were then harvested, surface stained, fixed, permeabilized, and intracellularly stained for IFN-α and TNF-α. (**A**) Gating strategy to analyze the population of cytokine-producing human Lin^- $CD123^+$ $BDCA\text{-}4^+$ pDCs (non-stimulated cells). (**B**) Frequencies of IFN-α-producing pDCs. (**C**) Frequencies of TNF-α-producing pDCs.



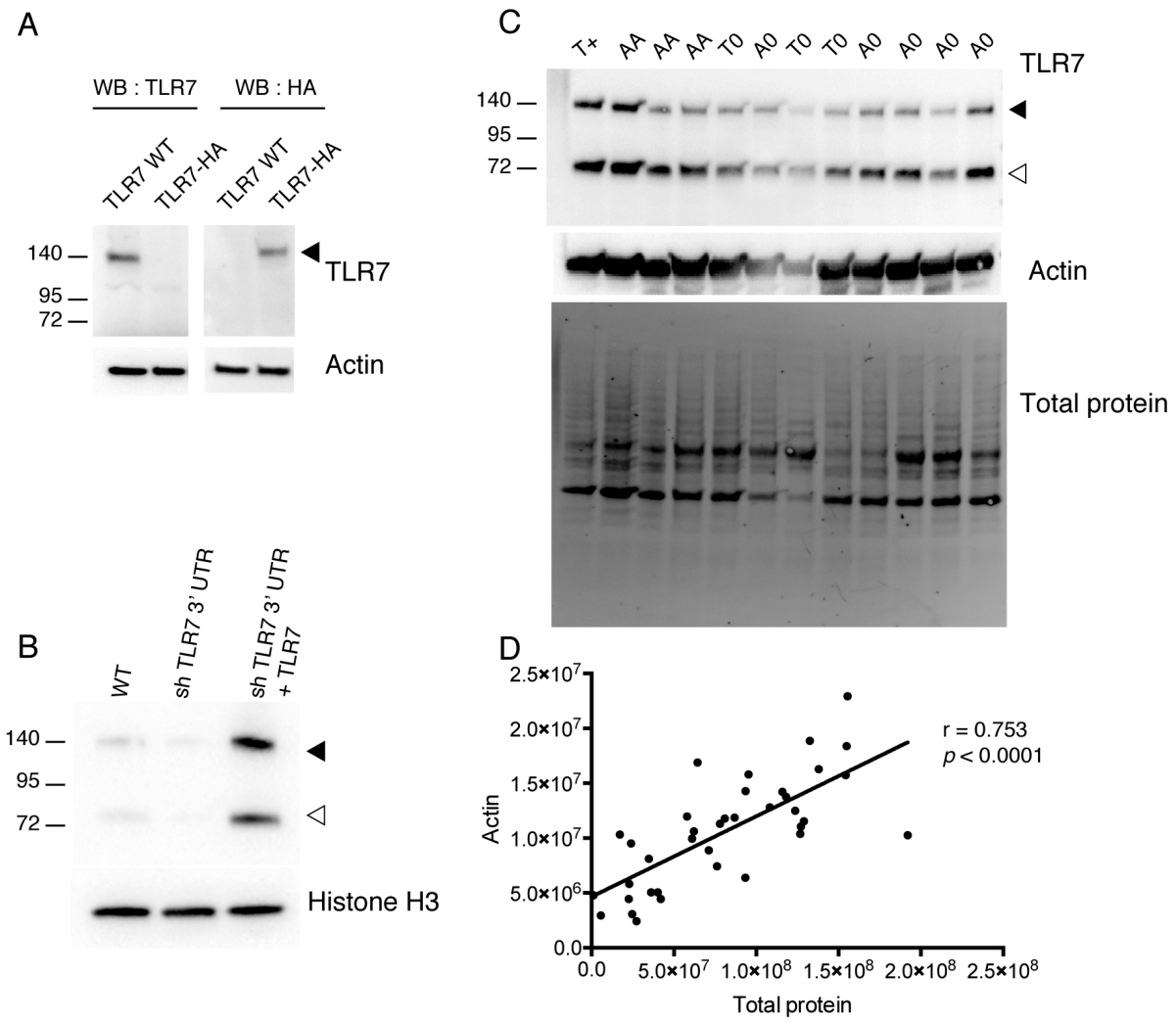
Supplemental Figure 3. Carriage of rs179008 allele T does not modify TLR7-driven TNF-α production in pDCs. Freshly isolated PBMCs from men and women were stimulated for 5 hours with a TLR ligand, in the presence of brefeldin A for the final 3 hours, and the frequency of TNF-α-producing Lin⁻ CD123⁺ BDCA-4⁺ pDCs determined by flow cytometry. (**A**) 1 μg/ml R-848, a ligand of TLR7 and TLR8. (**B**) 0.3 μg/ml R-848. (**C**) $Gag_{RNA1166}$, a HIV-1-derived RNA ligand of TLR7 and TLR8. (**D**) 1 μg/ml R-837, a ligand of TLR7. Each dot represents one donor. Horizontal bars indicate mean values. Mann-Whitney tests.



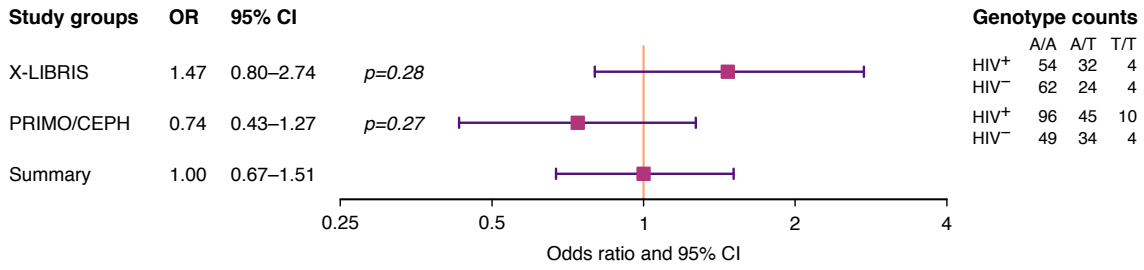
Supplemental Figure 4. SNP rs179008 affects neither *TLR7* mRNA levels nor splicing. (**A**, **B**) Analysis of the *TLR7* exon 2-exon 3 splice junction by PCR amplification and agarose gel electrophoresis of a cDNA fragment extending from the 3′ end of exon 1 to the 5′ region of exon 3. RNA was isolated from male (**A**) and female (**B**) PBMCs. Amplimer sizes were as expected from correct splicing. The faint minor band at 396 bp corresponds to a known alternative splicing form. Std, DNA size standards. **(C, D**) Quantitation of *TLR7* mRNA by real-time PCR. (**C**) PBMCs from men (i.e., obligate hemizygotes for *TLR7*) of rs179008 A/0 (n = 5) or T/0 (n = 3) genotype were cultured in medium supplemented or not with 1 μg/ml class A CpG oligodeoxynucleotides (CpG-A) as TLR9 ligands or with IFN-β for 24 h. The cells were then harvested, and B lymphocytes isolated for *TLR7* mRNA quantitation by real-time PCR. Data were normalized to the *GAPDH* internal control (× 10^3). (**D**) PBMCs from women were analyzed for rs179008 genotype-dependent *TLR7* mRNA expression as in (**C**), either at rest or after 24-h stimulation with IFN-β. No significant differences between genotypes were observed in resting cells or after incubation with IFN-β (Kruskal-Wallis test with Dunn's post hoc test to correct for multiple comparisons).



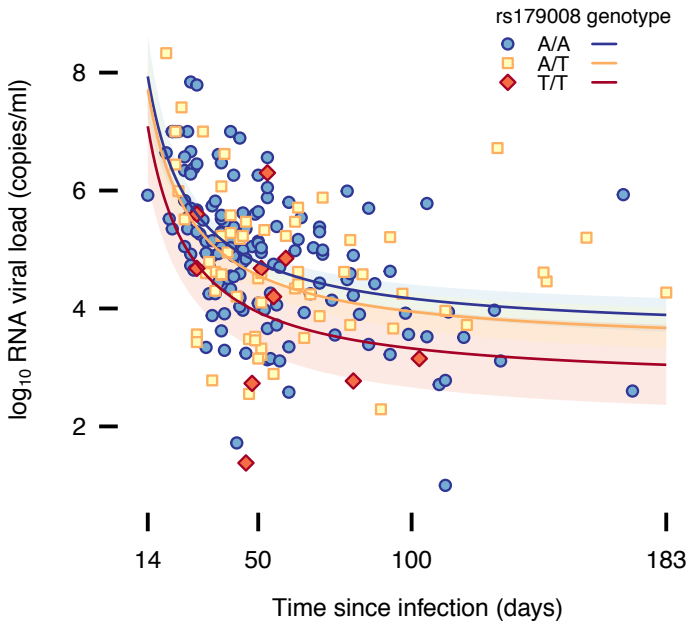
Supplemental Figure 5. Adaptation of codon usage at the 5′ end of the TLR7 coding sequence. Analyses with the *gcua* software show the adaptiveness score of each codon to the human codon usage bias in the A (top) and T alleles of rs179008. Rare, low-adaptiveness (< 20%) codons are shown in red. The amino acid translation is shown at the bottom, and arrows point to the 11th codon and amino acid. The CAA codon for the 11th residue in allele A encodes 27% of glutamine residues across the human genome. A CTA codon is found at this position in allele T, which is used for only 7% of leucine residues. 5′-to-3′ directional scanning of the mRNA by the translating ribosome, from left to right on these images, predicts the Leu_{11} codon in allele T to be the first low-adaptiveness codon in the path of the ribosome. Unlike the CTA triplet encoding Leu_{20} (also in red), the Leu_{11} codon is close enough to the initial ATG codon to suggest interference with translation initiation.



Supplemental Figure 6. Highly specific immunodetection of TLR7. (**A**) Western blot analysis of TLR7 protein from HEK 293T cells transfected with TLR7 expression vectors encoding either native TLR7 or a modified form carrying a C-terminal HA epitope tag. Detection was performed with a monoclonal antibody to the C-terminus of TLR7 (left) or to the HA epitope (right). (**B**) GEN2.2 pDC leukemic cells expressing an shRNA targeting the 3′ UTR of *TLR7* (1) were further transduced with a lentiviral vector of the full-length *TLR7* cDNA as previously described (1) and assessed for TLR7 protein expression. (**C**) Western blot detection of TLR7, β-actin, and total protein in PBMCs from women of different rs179008 genotypes using the antibody to the TLR7 C-terminus, from a representative subset of the experiments shown in Figure 4E. (T+) denotes the control cell lysate from an A/A woman that was used as an internal standard for inter-gel normalization of TLR7 expression data throughout the experiments. (**D**) Correlation between the densitometric signal of β-actin and total protein staining. Spearman's correlation coefficient.



Supplemental Figure 7. Association analysis between rs179008 T and HIV-1 infection status. The forest plot shows the measure of association (odds ratio and 95% CI) determined from the proportions of A/T and T/T (carrier) versus A/A (non-carrier) genotypes in two different case-control settings, and a meta-analytical summary of the two studies. Genotype counts are shown to the right of the plot; p-values from exact Cochran-Armitage trend tests, two-sided.



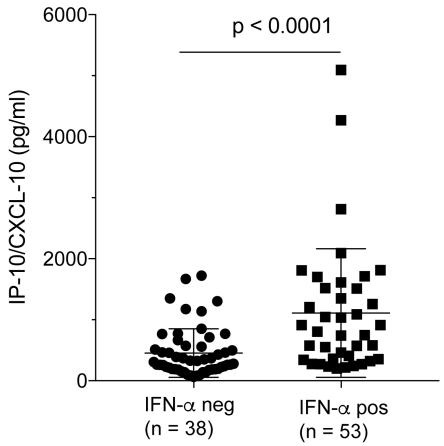
	Estimate	95% CI	p-value
(Intercept)	3.56	3.21, 3.91	<2E–16
Time^{-1}	61.31	47.48, 75.14	6.75E–16
rs179008 A/T	–0.22	–0.53, 0.09	0.16
rs179008 T/T	–0.85	–1.51, –0.18	0.013

Supplemental Figure 8. Regression model for viral load among PRIMO female patients. The variation in viral load among female patients was modelled with reference to the time elapsed since infection and to the rs179008 genotype. The model coefficients are shown under the plot. The coefficients for the genotypes are relative to the baseline genotype, rs179008 A/A.

Supplemental Table 1. Signs and symptoms in female PRIMO patients homozygous for rs179008

	A/A genotype N=148		T/T genotype N=10	
Signs and symptoms	**n**	**Percentage (95% CI)**	**n**	**Percentage (95% CI)**
Symptomatic in general	116	78% (71–84)	5	50% (24–76)
Fatigue	90	61% (53–68)	5	50% (24–76)
Fever	89	60% (52–68)	5	50% (24–76)
Arthralgia, myalgia	60	41% (33–49)	3	30% (11–60)
Tonsillitis, pharyngitis [a]	55	37% (30–45)	0	0% (0–28)
Weight loss	48	32% (25%–40%)	3	30% (11–60)
Rash	47	32% (25%–40%)	3	30% (11–60)
Headache	42	28% (22–36)	1	10% (2–40)
Cervical lymphadenopathy	37	25% (19–33)	2	20% (6–51)
Nausea, vomiting	37	25% (19–33)	2	20% (6–51)
Diarrhea	35	24% (18–31)	2	20% (6–51)
Oral ulcer	29	20% (14–27)	1	10% (2–40)
Hepatosplenomegaly	6	4% (2–9)	1	10% (2–40)
Other neurological signs	6	4% (2–9)	0	0% (0–28)
Meningitis	3	2% (1–6)	0	0% (0–28)
Mononeuritis	1	1% (0–4)	0	0% (0–28)
Other signs and symptoms	33	22% (16–30)	1	10% (2–40)

[a] p = 0.011, Boschloo's exact test, two-sided.



Supplemental Figure 9. Association of IP-10 values with detectable IFN-α. The plot shows the relationship between measured values for IP-10 and the detection (IFN-α pos) or non-detection of IFN-α in a representative sub-cohort of female Caucasian PRIMO patients.

Supplemental Table 2. Relationship between symptomatic presentation of primary HIV-1 infection and measurements for IFN-α, IP-10 and HIV-1 nucleic acids in the n=92 PRIMO subset

	Symptomatic presentation		
	NO, n=18	**YES, n=74**	**p-value**
HIV-1 RNA (log_{10} copies/ml)	3.8 (0.3)	4.9 (0.1)	<0.001[a]
Cell-associated HIV-1 DNA (log_{10} copies/10^6 cells)	2.8 (0.2) [n=16]	3.3 (0.1) [n=67]	0.03[a]
IP-10 (pg/ml)	385.2 (100.7)	819.6 (100.2)	0.04[a]
IFN-α			
– Detectable	4 (22.2%)	34 (46.6%)	0.068[b]
– Non-detectable	14 (77.8%)	39 (53.4%)	

[a] Student's t-test. Values expressed as mean (SD).
[b] Fisher's exact test. Values expressed as n (%).

Supplemental Table 3. Association between rs179008 genotypes, parameters of primary HIV-1 infection, and plasma IP-10 in a subset of PRIMO female patients

	A/A (n=41)	**A/T or T/T (n=51)**	**T/T (n=10)**	**p-value A/A vs A/T + T/T**	**p-value A/A vs T/T**	**Notes**
Symptomatic primary infection	35 (85%)	39 (76%)	5 (50%)	0.31	**0.03**	2
Viral load (log_{10} HIV-1 RNA copies/ml)	5.0 (4.3–5.8)	4.6 (3.7–5.2)	4.4 (2.8–4.8)	**0.03**	0.05	1
Cell-associated HIV-1 DNA (log_{10} copies/10^6 cells)	3.5 (3.0–3.8)	3.0 (2.6–3.5)	3.0 (2.7–3.5)	**0.003**	0.09	1
Detectable IFN-α	18 (44%)	20 (40%)	3 (30%)	0.83	0.49	1
IP-10 (pg/ml)	575.8 (275.7–1209.6)	364.8 (216.4–769.0)	398.3 (190.8–808.5)	**0.04**	0.26	1

[1] Data for continuous variables represented by the median and IQR. Wilcoxon's test.
[2] Categorical data expressed as case counts and percentages with 95% CI. χ^2 or Fisher's exact test.
n.a., not applicable.

Supplemental Note 1. Genotype and haplotype notation. Female genotypes for a single *TLR7* SNP are written A/B, where allele A is on one X chromosome and allele B on the other X chromosome. In male *TLR7* genotypes, obligate hemizygosity is denoted by zero in lieu of the second allele, e.g., A/0. In writing haplotypes formed by two or three SNPs, the SNPs are always ordered according to their telomere-to-centromere position along the X chromosome: rs179012–rs179008-rs3853839. In females, as above, the alleles carried on one X chromosome are separated by the "/" symbol from the alleles carried on the homologous chromosome. For instance, the GA/GA genotype for rs179012-rs179008 is parsed as follows: alleles rs179012 G and rs179008 A are carried on one X chromosome, and the same alleles are carried on the homologous chromosome; the individual is thus homozygous for rs179012 G and also for rs179008 A.

Supplemental Note 2. Statistical analyses with R. Specialized analyses in the R 3.6.1 computing environment (2) were carried out as follows.

- In Supplemental Figure 7, two-sided Cochran-Armitage trend tests (CATT) (3) under a dominant model were used to test genotype counts from case and control groups for genetic association with HIV-1 acquisition, using package *MaXact* (4). The meta-analysis of odds ratios was performed using package *gmeta* (5) with method = exact1.
- Estimation analyses of HIV-1 acute infection parameters as a function of genotype (Figure 6A and Figure 7C) were performed with package *dabestr* 0.2.2 (6). The computation of 95% confidence intervals was done within *dabestr* by a bias-corrected and accelerated boostrap procedure with 5000 replications.
- 95% confidence ellipses in bivariate plots of HIV-1 infection parameters (Figure 5B) were drawn with function dataEllipse() from package *car* 3.0.3 (7), based on a bootstrap distribution of the bivariate mean with 5000 replications. Hotelling's T^2 tests were performed with package *rrcov* 1.4.9 (8).
- The coefficients of pairwise linkage disequilibrium between *TLR7* SNPs in Figure 6B were computed and plotted as a heatmap with package *LDheatmap* 0.99.7 (9).
- The linear model for viral load in Supplemental Figure 8 was computed on the basis of two predictive variables, the time elapsed since infection and the rs179008 genotype. A fractional polynomial function of degree −1 was used for the time since infection. The distributions for the time since infection for the rs179008 genotype groups, A/A and (A/T + T/T), were first quantile-normalized as the average empirical distribution of the two distributions; this was done to take into account the slight difference in the median time since infection between the two genotype groups (Table 1). Quantile normalization was performed with the normalizeQuantileRank() function from package *aroma.light* 3.14.0 (10).

Supplemental References

1. Souyris M, Cenac C, Azar P, Daviaud D, Canivet A, Grunenwald S, et al. TLR7 escapes X chromosome inactivation in immune cells. *Sci Immunol*. 2018;3(19):eaap8855.
2. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2019. https://www.R-project.org/
3. Balding DJ. A tutorial on statistical methods for population association studies. *Nat Rev Genet.* 2006;7(10):781–791.
4. Tian J, Xu C. MaXact: Exact max-type Cochran-Armitage trend test (CATT). R package version 0.2.1; 2013. https://CRAN.R-project.org/package=MaXact
5. Yang G, Cheng JQ, Xie M, Qian W. gmeta: Meta-Analysis via a unified framework of confidence distribution. R package version 2.3-0; 2017. https://CRAN.R-project.org/package=gmeta
6. Ho J, Tumkaya T, Aryal S, Choi H, Claridge-Chang A. Moving beyond P values: data analysis with estimation graphics. *Nat Methods.* 2019;16(7):565–566.
7. Fox J, Weisberg S. An R Companion to Applied Regression, 3rd Edition. Thousand Oaks, California: SAGE Publications; 2019.
8. Todorov V, Filzmoser P. An object-oriented framework for robust multivariate analysis. *J Stat Soft.* 2009;32(3):1–47. http://www.jstatsoft.org/v32/i03/
9. Shin J-H, Blay S, McNeney B, Graham J. LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *J Stat Soft.* 2006;16(Code Snippet 3):1–9. http://stat.sfu.ca/statgen/research/ldheatmap.html
10. Bengtsson H, Neuvial P, Speed TP. aroma.light: light-weight methods for normalization and visualization of microarray data using only basic R data types; 2019. https://bioconductor.org/packages/release/bioc/html/aroma.light.html