Supplementary Data
Combined activities for control sequences measured with an all-in-one reporter

We created and tested extra non-matched shRNA plasmids including single shRNA vectors plus control combinations (c.c.)). Each were separately transfected with the three all-in-one reporters; the HIV-specific aio sense and anti-sense reporters, and the non-matched control (though it is matched to the 7 control shRNAs shown here). Off-scale values (> 100 %, i.e. no activity) are indicated by open circles and text labels where appropriate. Values shown are representative of 2 or more independently repeated experiments.
Supplementary materials and methods

Detailed conservation profiles for the 10 selected shRNAs

The following table shows the detailed conservation profiles for the 10 selected shRNAs calculated in our previous study. This information is an excerpt from Mcintyre et al. “96 shRNAs designed for maximal coverage of HIV-1 variants” Retrovirology (2009) vol. 6 (1) pp. 55. HIV-1 sequence data was compiled from 2 sources; publicly available sequence from the Los Alamos National Laboratory (LANL; www.hiv.lanl.gov) and proprietary sequence information from Virco (www.virco.com). The LANL data set included all near full-length genome sequences and gene sequence fragments as of December 2006. HIV-2 and SIV sequences were examined and excluded as they were sufficiently divergent to the NL4-3 HIV-1 reference strain [Genbank:AF324493]. The Virco data set was a small, but highly relevant private data set obtained from 105 HIV-1 infected persons from Europe. It contained only gene-specific sequences for the 6 accessory genes; Tat, Rev, Vif, Vpu, Vpr and Nef. Conservations were calculated for LANL and Virco sequences combined, LANL clade B subtypes only, and the Virco sequences only (where applicable). In each group, 6 conservations are given: the conservation for the entire overlapping 23 mer region, and the 5 individual conservations for the 5 overlapping 19 mers that make up the 23 mer profile region.

<table>
<thead>
<tr>
<th>#</th>
<th>Target</th>
<th>LANL + Virco (all)</th>
<th>LANL clade B</th>
<th>Virco (all)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>-2</td>
<td>-1</td>
</tr>
<tr>
<td>0</td>
<td>LTR 510-21</td>
<td>69</td>
<td>69</td>
<td>69</td>
</tr>
<tr>
<td>1</td>
<td>LTR 527-21</td>
<td>90</td>
<td>93</td>
<td>93</td>
</tr>
<tr>
<td>2</td>
<td>Gag 533-20</td>
<td>58</td>
<td>62</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>Pol 248-20</td>
<td>75</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>Pol 2670-21</td>
<td>72</td>
<td>78</td>
<td>78</td>
</tr>
<tr>
<td>5</td>
<td>Pol 2878-20</td>
<td>74</td>
<td>84</td>
<td>76</td>
</tr>
<tr>
<td>6</td>
<td>Vif 9-21</td>
<td>71</td>
<td>71</td>
<td>71</td>
</tr>
<tr>
<td>7</td>
<td>Tat (x1) 140-21</td>
<td>68</td>
<td>81</td>
<td>82</td>
</tr>
<tr>
<td>8</td>
<td>Vpu 143-20</td>
<td>55</td>
<td>56</td>
<td>65</td>
</tr>
<tr>
<td>9</td>
<td>Env 1428-21</td>
<td>72</td>
<td>79</td>
<td>79</td>
</tr>
</tbody>
</table>

Average (all 10) : 70 75 76 74 74 75 82 86 87 85 85 85 67 73 75 73 72 72
Constructing the single shRNA expression plasmids

The inserts for all of the shRNA expression plasmids (except shRNA #1) were built from a single synthetic oligonucleotide ~ 72 - 75 nt. long. Each oligonucleotide template consisted of a partial RE recognition site (Bam HI), the hairpin sense (or upper) sequence, a loop sequence, the hairpin anti-sense sequence, a pol III terminator sequence, and a second partial RE recognition site (Hind III), organized in the following general format: 5’-GCGCGGATCC | core(20 - 21) | NCTCGAGN | core(20 - 21) | [G/A/C] | TTTTTTGAAAGCTT-3’. A short primer sequence (12 nt.) common to the 3’ end of all oligonucleotides was also designed (5’-CGCGAAGCTTCCAAAAA-3’), annealed to each oligonucleotide template and extended with Phi-29 DNA polymerase in a single-step isothermal extension reaction. This step generated double-stranded synthetic inserts that were then digested to create ‘sticky ends’ and cloned as per standard procedures (as previously described). The template for shRNA #1 was created from standard complementary oligonucleotide pairs with offset ends as each shRNA core sequence contained an internal Hind III site making it incompatible with the Phi-29 extension method.

Phi-29 extension (shRNAs #0, #2 - #9)

Oligos were ordered at the minimal synthesis and purification scales (0.05 μM and desalt, Sigma-Genosys) and were re-suspended in water (1 - 10 μg / μl). Twenty picomoles of each oligo was used in the extension reaction (1x reaction buffer, 2x BSA, 20 μM dNTPs (1 μl of a 10 mM stock), 10 units of Phi29 (New England Biolabs) and water to 20 μl), which was incubated at 30 °C for ~10 min., then 65 °C for 10 min. (to inactivate the polymerase). The extension product was digested (Bam HI plus Hind III), purified using the Nucleotide Removal kit (Qiagen), ligated to the expression plasmid and used to transform electrocompetent GT116 E.coli. Positive clones were confirmed by automated sequencing using our loop digestion method. Each single shRNA construct was digested and sequenced in two reactions, one containing the forward primer, the other containing the reverse. The primers bound to the expression plasmid backbone approximately 100 bases away from the region encoding the base of the hairpin stem. Each reaction contained: 1x RE buffer (NEB-2, New England Biolabs), 1x BSA, ~ 500 ng of template plasmid, 10 pmol of sequencing primer, ~ 10 units XhoI and water to a total volume of 16 μl, and was incubated at 37 °C for 30 – 60 min. prior to shipping, without purification, to an automated sequencing facility (Australian Genome Research Facility, AGRF).

<table>
<thead>
<tr>
<th>Hairpin #</th>
<th>Target (NL4-3)</th>
<th>Oligonucleotide template (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (♂)</td>
<td>LTR 510-21</td>
<td>GCGCGGATCCCCCCTAAGGAGGTTTTTTTTTTTGGAGAGCTT</td>
</tr>
<tr>
<td>1 (♂)</td>
<td>LTR 527-21</td>
<td>Overlapping oligonucleotide pair used instead, see below</td>
</tr>
</tbody>
</table>
Standard complementary oligonucleotide pairs (shRNA #1)

The template for shRNA #1 was made as a synthetic duplexes with overhanging ends identical to those created by restriction enzyme (RE) digestion (BamHI at the 5' and HindIII at the 3'). The coding region for each hairpin was contained within a single oligonucleotide (upper oligo: 5'-GATCC | core(19-29) | NCTCGAGN | core(19-29) | [G/A/C] | TTTTTTGGA-3') and its complementary equivalent (lower oligo: 5'-AGCTTCCAAAAAA | [G/A/C] | core(19-29) | NCTCGAGN | core(19-29) | G-3'). Both oligos were synthesized and re-suspended as previously described. 1 μl from each was added to 98 μl of annealing solution (10 mM Tris-Cl pH 8.0, 50 mM NaCl, 1 mM EDTA), heated to 100 °C for 5 minutes, slowly equilibrated to room temperature and diluted up to 10,000 fold for ligation. The insert and plasmid were ligated, and used to transform electrocompetent GT116 E.coli (Invivogen).
The reporter and control sequences

The fluorescent reporters used in this study were assembled with following target domains derived from the common HIV-1 laboratory strain, NL4-3 [Genbank:AF324493]. n.b. each of the following target domains was fused immediately downstream of GFP with several stop codons placed between the domains.

Target-specific reporters

Nef-LTR (1255 bp): Nef shown in grey, LTR shown in UPPERcase

(matched to shRNA #0 and #1)

tggagggctattccatctccaaaagagaacaagaatctctgtgtgtgatctaccacacacagcattcctcttgtggc
cagaactacacacagggccaggggtcagatctccactggtcttggtggtgctagctctacctgactccaaggctgtggg
cactactacagggccaggtctcagatctccacctgactcttcacactgtaccagttgccagat

tggagggctattccatctccaaaagagaacaagaatctctgtgtgtgatctaccacacacagcattcctcttgtggc
cagaactacacacagggccaggggtcagatctccactggtcttggtggtgctagctctacctgactccaaggctgtggg
cactactacagggccaggtctcagatctccacctgactcttcacactgtaccagttgccagat

tggagggctattccatctccaaaagagaacaagaatctctgtgtgtgatctaccacacacagcattcctcttgtggc
cagaactacacacagggccaggggtcagatctccactggtcttggtggtgctagctctacctgactccaaggctgtggg
cactactacagggccaggtctcagatctccacctgactcttcacactgtaccagttgccagat

tggagggctattccatctccaaaagagaacaagaatctctgtgtgtgatctaccacacacagcattcctcttgtggc
cagaactacacacagggccaggggtcagatctccactggtcttggtggtgctagctctacctgactccaaggctgtggg
cactactacagggccaggtctcagatctccacctgactcttcacactgtaccagttgccagat

tggagggctattccatctccaaaagagaacaagaatctctgtgtgtgatctaccacacacagcattcctcttgtggc
cagaactacacacagggccaggggtcagatctccactggtcttggtggtgctagctctacctgactccaaggctgtggg
cactactacagggccaggtctcagatctccacctgactcttcacactgtaccagttgccagat

Additional file 1 (Mcintyre et. al.)

Page 5
atagtaagaatgtatagcctaccagcatctggcacataagacaaagagacaccctttagagactagactagttagacccgat
tctataaaacctaaagagccagcagcagcctcacaagaggttaaaaaattgatgacagaaacctgttgtggtcacaataagcgaaa
caccagttgaagactatttttaaaagcttggaagcagcagcagcactagaagaatgtacagacagcatgtcagggagtggg
ggacccaggccataaagcagaggttttgctgaagcataaatggcgaagtaaaatccagctacattataatgacagaaag
gcaatatt

**Pol-1 (1 - 436) (436 bp)**
(matched to shRNA #3)
ttttttagggaagatctggccttcaccaagggcagccaggaatttttcctagacgacagcagacagcactagaagaatgtacagacaggttttgttggtcacaataagcgaacaccagttgaagactatttttaaaagcttggaagcagcagcagcactagaagaatgtacagacagcatgtcagggagtggg

gagAGCAGATGATACAGTATTagaagaaatgaattttgcaggaagagatgggaacacccaaataggtgggggaatttgggaggttttatcaaagtaagacagtatgatcagatactcatagaaatctgcggacataaagctataggtacagtattagga
cctacacctgtcacaataattggaagaaatctgt

**Pol-2670 (41 bp)**
(matched to shRNA #4)
ttaagacagcAGTACAAATGGCAGTATTCatccacaatatt

**Pol-2878 (41 bp)**
(matched to shRNA #5)
cctctggaaaGGTGAAGGGGCAGTAGTAAatcaagataata

**Vif (579 bp)**
(matched to shRNA #6)
AtggaaaaaCAGATGGCAGGTGTGATTgtgtggcaagtagcaggatagagatttaacatagggaaagatttag

taaacacatgtagtatatttccaggaagaaactgaagctgttgtttagagcactacttagaagagaatcactaattcccaaaataggtcc
gagagctacacatcccatcagggagtgtctaaattgtagaataaacaacatattgggtgtctgctacagaggaaagagactgctcat
tgggtcagggaggtctctacaataagggaaacagatagtcacacacaaactgtgaccctgctagccgacaccactaatatt
catctgcaactattttgtagtttttcagaataaggctctagattaataggtgataggtgtagttctcattggtgaatatacagc
gagacataaacaaggtctaggactctcagatgtggtcagcactagcaggacattataaaaaccaaaaaagatazagacgacccactttgcct
taggttaggaaaactgacagagggacagggacaggggcaacagcagagaggggcccagagacaggggaggccacaatcatgga
gacacagt

**Tat x12 (261 bp)**
(matched to shRNA #7)
atggagccagttagctctagactagagccctggaagacatccaggaagctgctcaaaacactgtgtcataaatgtctattgtaaaagttgtgtctctagttgtttcatgcacaaagacctgggcatcctctctATGGCAGGAAGAAAGCGGAGagc

cagcgacagagagctcagacagaagagtctacactcaagaggtttctctctatcaagagacaccctcaccacactccgaggggacccgcagcgaagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagagaga
Vpu (246 bp)
(matched to shRNA #8)
atgcaacctataatagtagcattagtagcattagcataataatatgcaaatagttgtgtggtccatagtaatcatagaata
tagaaaatatagacaagaagaataatagacaggtagttaattgatagactaatagaaagaacaacagcagcctgtctttaggatattgatgatctgtgagt

Env-1300 (1300 - 1740) (440 bp)
(matched to shRNA #9)
ccccctcccatcagtggacaaattagatgttcatcaaatattactggcctgctattaacaagagatggtggtaataacaacaatg
ggtcgcgagatctttcacaacctgaggggcatatgaggggacaaTTGGAGAAGTGGAATTATATAaatataaagtagttaattggaaggatggtgtaatatataaatagagagatggtgtgaatatagaacacaagttagaatgtaggtggtcgcgagagaaaaaagacagttggaatatagggcgtttttctggggcttgggttcttgggagccagcaggaagaactattggtcagctacgctacacccatcgtattgtgggttaataaggccggtttaatgagggctatttgagcagcagcttacttgggtataatataaagtaaagatggtggtaataacaacaatg
ggtccggatcttcagacctggaggaggcgatatgaggggacaaTTGGAGAAGTGGAATTATATAaatataaagtagttaattggaaggatggtgtaatatataaatagagagatggtgtgaatatagaacacaagttagaatgtaggtggtcgcgagagaaaaaagacagttggaatatagggcgtttttctggggcttgggttcttgggagccagcaggaagaactattggtcagctacgctacacccatcgtattgtgggttaataaggccggtttaatgagggctatttgagcagcagcttacttgggtataatataaagtaaagatggtggtaataacaacaatg
ggtccggatcttcagacctggaggaggcgatatgaggggacaaTTGGAGAAGTGGAATTATATAaatataaagtagttaattggaaggatggtgtaatatataaatagagagatggtgtgaatatagaacacaagttagaatgtaggtggtcgcgagagaaaaaagacagttggaatatagggcgtttttctggggcttgggttcttgggagccagcaggaagaactattggtcagctacgctacacccatcgtattgtgggttaataaggccggtttaatgagggctatttgagcagcagcttacttgggtataatataaagtaaagatggtggtaataacaacaatg
ggtccggatcttcagacctggaggaggcgatatgaggggacaaTTGGAGAAGTGGAATTATATAaatataaagtagttaattggaaggatggtgtaatatataaatagagagatggtgtgaatatagaacacaagttagaatgtaggtggtcgcgagagaaaaaagacagttggaatatagggcgtttttctggggcttgggttcttgggagccagcaggaagaactattggtcagctacgctacacccatcgtattgtgggttaataaggccggtttaatgagggctatttgagcagcagcttacttgggtataatataaagtaaagatggtggtaataacaacaatg
ggtccggatcttcagacctggaggaggcgatatgaggggacaaTTGGAGAAGTGGAATTATATAaatataaagtagttaattggaaggatggtgtaatatataaatagagagatggtgtgaatatagaacacaagttagaatgtaggtggtcgcgagagaaaaaagacagttggaatatagggcgtttttctggggcttgggttcttgggagccagcaggaagaactattggtcagctacgctacacccatcgtattgtgggttaataaggccggtttaatgagggctatttgagcagcagcttacttgggtataatataaagtaaagatggtggtaataacaacaatg
ggtccggatcttcagacctggaggaggcgatatgaggggacaaTTGGAGAAGTGGAATTATATAaatataaagtagttaattggaaggatggtgtaatatataaatagagagatggtgtgaatatagaacacaagttagaatgtaggtggtcgcgagagaaaaaagacagttggaatatagggcgtttttctggggcttgggttcttgggagccagcaggaagaactattggtcagctacgctacacccatcgtattgtgggttaataaggccggtttaatgagggctatttgagcagcagcttacttgggtataatataaagtaaagatggtggtaataacaacaatg
ggtccggatcttcagacctggaggaggcgatatgaggggacaaTTGGAGAAGTGGAATTATATAaatataaagtagttaattggaaggatggtgtaatatataaatagagagatggtgtgaatatagaacacaagttagaatgtaggtggtcgcgagagaaaaaagacagttggaatatagggcgtttttctggggcttgggttcttgggagccagcaggaagaactattggtcagctacgctacacccatcgtattgtgggttaataaggccggtttaatgagggctatttgagcagcagcttacttgggtataatataaagtaaagatggtggtaataacaacaatg
ggtccggatcttcagacctggaggaggcgatatgaggggacaaTTGGAGAAGTGGAATTATATAaatataaagtagttaattggaaggatggtgtaatatataaatagagagatggtgtgaatatagaacacaagttagaatgtaggtggtcgcgagagaaaaaagacagttggaatatagggcgtttttctggggcttgggttcttgggagccagcaggaagaactattggtcagctacgctacacccatcgtattgtgggttaataaggccggtttaatgagggctatttgagcagcagcttacttgggtataatataaagtaaagatggtggtaataacaacaatg
ggtccggatcttcagacctggaggaggcgatatgaggggacaaTTGGAGAAGTGGAATTATATAaatataaagtagttaattggaaggatggtgtaatatataaatagagagatggtgtgaatatagaacacaagttagaatgtaggtggtcgcgagagaaaaaagacagttggaatatagggcgtttttctggggcttgggttcttgggagccagcaggaagaactattggtcagctacgctacacccatcgtattgtgggttaataaggccggtttaatgagggctatttgagcagcagcttacttgggtataatataaagtaaagatggtggtaataacaacaatg

All-in-one reporters
The **aio** reporters were created using custom generated target domains (GenScript; www.genscript.com) comprised of 9 target domains matched to our 10 shRNAs, or 7 unmatched control domains (for the control reporter) which were transferred via PCR into our EGFP reporter base plasmids. The sequence of the 406 bp combined target domain synthesized for the aio sense reporter was (**Xho I - Bgl II** [for cloning]) - #9 - #2 - #0|1 - #3 - #4 - #5 - #7 - #6 - #8 - **Sal I - Bam HI** [for cloning]) (5’-3’): ctcgag (**Xho I**) agatct (**Bgl II**) tgaggga-
caatTTGGAGAAGTGGAATTATATAaatataaagta (**#9, 21**) ttatcagaaggagcaacccaccaacctgcttaagccaaatataaagcttgccttgagtgcttcaagtagtttagatacaggagca

cgtcgggctcaacacgctccagggcaagaatctgtgct

Additional file 1 (Mcintyre et. al.)
ggcagtagataatacaagataatggcactctcctatggcaggaagaagcggagacagcgacgaagccatggaaaacagatgcagagtgaatggcatctcctatggcaggaagaagcggagacagcgacgaagccatggaaaacagatg
gcaggtgatgattgtggaagactctctcatctctcattgccactgtcttctgctctttctattagtacttgccacacaatcatcactctgccatctgttttc
catggtgcttcgtcgtctcgctctcctgccataaggagagtgcctttccacttccttccacctccctcagagga
aatttggagtgaatatgcctatttgctgctttctattatactctctctcttacactctactctgactctgctgtctagctctgtcttattactatcatctcattttctactcttgaagcactcaaggcaagtttattagggcttttaagcaggtgggtccttaggtgcatgttatattaaattgtggtggtggtctctgtgatata
tattttatatataattcatctttctgacttctcagagatct

non-matched control (308 bp)
ctcgagagatctatgaaatataaatatattggttaagaggttaacagggagtcgtaccataaatttagaagacccaccgaggaaagactttccaaaaa
AAGACAGTCCAACACACGCCACCTGTCTC
GAGA
GAGACAGGTGGCGTGTGTTGGACTGTCTT
G
acatagattgaacgcgcaaacgggaagagcgtactctcgaggtgaacggtgtttagttaggtgagcgttagacaattttcttgcagtttactcattgcctcaagatct

The 7 non-matched shRNA controls used to make the 2 (3.8), 3 (3.8.9), 4 (3.8.9.2), 5 (3.8.9.2.7), 6 (3.8.9.2.7.6) and 7 (3.8.9.2.7.6.0) cassette control plasmids were derived from the backwards complement of shRNAs #3, #8, #9, #2, #7, #6 and #0. In this way they were unmatched to the aio reporters yet had identical nucleotide compositions (but in reverse order) to retain similar thermodynamic profiles. The following table lists the sense (upper) strand core sequences for the 7 non-matched controls.
<table>
<thead>
<tr>
<th>shRNA #</th>
<th>Matched shRNA p0 core (5' to 3')</th>
<th>control #</th>
<th>Backwards controls (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>CCCACTGCTTAAGCCTCAA</td>
<td>c0</td>
<td>ATAACTCCGAATTCGTCACCC</td>
</tr>
<tr>
<td>2</td>
<td>GAGCCACCCCACAAGATTT</td>
<td>c2</td>
<td>ATTTGAACACCCCACCGAG</td>
</tr>
<tr>
<td>3</td>
<td>GAGCAGATGATACAGTATT</td>
<td>c3</td>
<td>ATTATGACATAGTAGACGAG</td>
</tr>
<tr>
<td>6</td>
<td>CAGATGGCAGGTGATGATT</td>
<td>c6</td>
<td>TGTTAGTAGGGACGGTAGAC</td>
</tr>
<tr>
<td>7</td>
<td>ATGGCAGGAAGAAGCGGAG</td>
<td>c7</td>
<td>CAGAGCCGAAGAAGGAAGGTA</td>
</tr>
<tr>
<td>8</td>
<td>GAGCAGAAGACAGTGCCAA</td>
<td>c8</td>
<td>TAACGTTGACAGAAGACGAG</td>
</tr>
<tr>
<td>9</td>
<td>TTGGAGAAGTGAATTATAT</td>
<td>c9</td>
<td>AATATATAGTGAAGGTT</td>
</tr>
</tbody>
</table>