SUPPLEMENTARY INFORMATION

MATERIALS AND METHODS

The design, engineering and characterization of 3- and 4-fingers that target hCCR5 and mTYR genes are as previously described (1). The ZFN designs and their cognate sites are shown in **Supplementary Table S1**. The 4-finger ZFPs that target hCCR5 and the obligate heterodimer FokI nuclease domain variants were generated by using overlapping oligonucleotides (2). Analysis of the sequence-specificity of designer CCR5-specific 3- and 4-finger ZFNs using cell-free assays are shown in **Supplementary Figure 1S**. The substrate, pc.CCR5 plasmid, is from NIH AIDS Research and Reference Reagent Program. HEK293 Flp-In cells were purchased from Invitrogen and generation of HEK293 cell lines with an integrated mutant eGFP was performed according to the Invitrogen manual. The locus-specific primers and PCR conditions are indicated in **Supplementary Table S2**. The media and growth conditions for mouse albino melanocyte cell culture are as described elsewhere (3).

Redesign of the dimer interface of the FokI nuclease domains of the ZFNs to reduce cytotoxicity:

The dimer interface of the FokI nuclease domains was re-engineered to reduce the cytotoxicity of the designed 3- and 4-finger ZFNs. We re-designed the dimer interface of the FokI nuclease domains as follows: The original ZFN architecture requires dimerization of the cleavage domain to induce a DSB, which necessitates the design and engineering of two different ZFNs for binding to adjacent half-sites [**Supplementary Figure 2S (A)**]. Since the protein-protein interaction for the cleavage domain dimer formation is weak, the ZFNs, like FokI, are likely to remain monomeric at concentrations less than 10 μM; and dimerize only when bound to their specific targets. Although this requirement for dimerization of the cleavage domain restricts cleavage by a pair of ZFNs to long sequences, it could also introduce a potential problem: since this dimer interaction does not select for the heterodimer species, homodimer species which form could lead to unwanted off-target cleavage elsewhere in the genome of cells. Two different homodimers could result from the two individual ZFNs with different sequence specificities, and

more often than not, they do occur [**Supplementary Figure 2S (B)**]. These homodimers affect the safety and efficacy of ZFN-mediated gene targeting. We re-designed the FokI nuclease domains as described elsewhere (4, 5).

Construction of ZFNs and the donor plasmid substrate:

Zinc finger proteins (ZFPs) containing 3 or 4 fingers that recognize the targeted CCR5 sequence were designed as previously described (1). The 4-finger ZFPs that target human CCR5 reported in literature were generated using overlapping oligonucleotides (2). The DNA sequences encoding the desired ZFPs were synthesized from 16 complete overlapping oligonucleotides. The oligonucleotides were first phosphorylated using polynucleotide kinase and ATP; and then ligated by using Taq ligase in a thermocycler and amplified by PCR. The ZFP constructs were digested with NdeI/SpeI and then ligated to NdeI/SpeI cut pET15b vector carrying the FokI nuclease domain. The FokI obligate heterodimer nuclease domain variants were also constructed from 30 completely overlapping oligonucleotides as described above. The amplified DNA was cloned into the BamHI site of pUC19 plasmid. The clones were sequenced and then subcloned into the pET15b vector at the SpeI/BamHI sites. The DNA and amino acid sequences for the FokI obligate heterodimer nuclease domain variants generated at PBPL will be reported elsewhere (manuscript in preparation). The ZFN constructs were then subcloned into pIRES (Clonetech Laboratories Inc.) at the NheI/MluI sites under the CMV promoter for expression in mammalian cells. Nuclear localization signal (PKKKRKV) was added to N terminal of the ZFN constructs in order to direct the proteins to the nucleus from the cytoplasm.

The donor substrate for eGFP gene correction was constructed as follows: The eGFP gene was amplified, deleting the first 12 bp including the start codon with primers 5'-GGCGAGGAGCTGTTCAC-3' and 5'- CCGACTCGAGTTACTT GTACAG CTCGTC-3'. Another fragment of DNA with overlapping bps to the 3' end of the eGFP gene was generated by amplifying a 700 bp DNA segment of pCDNA/FRT/TO vector with primers 5'-CCGCTCGAGTCTGAGGGCCCGTTTAAAC-3' and 5'-CCCGCAGGACATATCC ACG-3'. The two overlapping fragments were then used to generate a 1.5 kb DNA fragment by PCR,

which was subsequently cloned into pGEMT vector (Promega). The GFP donor template plasmid had been made non-functional for GFP expression by removing the first 12 amino acids of the GFP protein, including the start codon, and by using a promoter-less GFP.

Mutant eGFP construct encoding CCR5 ZFN binding sites:

The mutant eGFP gene was constructed by amplifying the eGFP gene in two fragments and inserting the ZFN binding site between two BstXI sites as follows: The 5' fragment of eGFP gene was amplified using primers 5'-CGCGGATCCGTTATGGTGAGCAAGGGCGAGGAGC-3' and 5'-ACTGACT GTA**U**GGAAAATGAGAGAGCCCATTATTATGGGTAGCGGCTGAAGCAC-3'. The 3' fragment of the eGFP gene was amplified using the primers 5'-ATACAGTCAG**U**ATCAATCCATGATGATGGAAGCAGC ACGACTTCTTC-3' and 5'-CCGACTCGAGTTACTTGTACAGCTCGTC -3'. The amplified fragments were digested with USER enzyme (New England Biolabs) and ligated by T4 DNA ligase. The ligated fragment was digested with BamHI and XhoI to be cloned in pCDNA5/FRT/TO vector (Invitrogen). The binding sites for various ZFN pairs were constructed by replacing the BstXI fragment as follows: Binding site oligonucleotides with BstXI overhangs are annealed and ligated to BstXI cut pcDNA5/FRT/TO vector containing the mutant eGFP gene.

Generating HEK293 cell lines with an integrated mutant eGFP:

HEK293 Flp-In cells were purchased from Invitrogen. HEK293 cell lines with an integrated mutant GFP encoding the CCR5 ZFN target site was generated using a similar procedure as discussed above. The cells were grown in DMEM supplemented with 10% fetal bovine serum. Cells were transfected with pOG44 (Invitrogen) vector and pCDNA5/FRT/TO vector carrying various CCR5 ZFN binding sites in the ratio 9:1 with lipofectamine 2000 (Invitrogen) transfection agent and hygromycin resistant single cell colonies were identified as above. Genomic DNA from the single cell colonies was then isolated to amplify the mutant eGFP gene with primers 5'-CGCAAATGGGCGGTAGGCGTG-3' and 5'-CCGACTCG AGTTACTTGTACAGCTCGTC-3'. The mutant eGFP gene locus was amplified, cloned and sequenced to confirm the presence of the ZFN binding sites. Similarly, we developed another HEK293 Flp-In cell line,

where the GFP gene is mutated by the insertion of an in-frame stop codon, followed by the recognition sites for the designed 3-finger ZFNs that target the mouse tyrosinase gene.

For ZFN-mediated gene targeting experiments, one million cells were then transfected with 400 ng of each pIRES vector carrying the corresponding CCR5 ZFNs and 1 ug of the donor template plasmid by lipofectamine 2000 as per manufacturer's protocol. eGFP expression was induced by adding 1ug/ml doxycycline 5-6 hours post-transfection. Gene targeting occurs when the transfected ZFNs create a DSB at their cognate recognition site in the mutant GFP gene and the DSB is repaired using the GFP donor plasmid as a template for HR. With effective gene targeting and HR, the cell becomes GFP positive, as detected by flow cytometry (6).

FACS and Microscopy analysis:

Three, five or seven days after transfection with ZFNs and donor plasmid, eGFP gene correction was measured by FACS using a BD FACS Canto II. GFP fluorescence (GFP) was measured using BP 530/30 filter. BD FACS Diva™ Software, v6.1.1 was used for analyses. A similar analysis was performed one, two and three weeks post-transfection with ZFNs and donor plasmid. GFP positive cells were sorted and examined by microscopy to confirm GFP expression (6).

RFLP and Nucleotide Sequence analysis:

Five days post-transfection with ZFNs and the donor plasmids, GFP positive cells were sorted, serially diluted to get individual clones and grown. Single cell colonies were isolated from sorted GFP^+ cells by seeding into 96-well plates at a density of less than one cell/well and grown for 30 days. The genomic DNA was then isolated from these GFP^+ clones as well as from GFP^- control clones. The eGFP at the FRT site was PCR amplified and then digested with BstXI. The PCR and digested products were resolved using a 1.8% agarose gel for RFLP analysis. The mutant eGFP gene has two BstXI sites, where the ZFN binding sites are inserted. Correction of the eGFP gene by homology-directed repair results in the loss of the BstXI sites. The PCR product size of the corrected eGFP gene is 930 bp as

compared to 990 bp for the mutant gene. BstXI digestion of the mutant eGFP PCR product generates two bands: 450 bp and 540 bp, respectively. All GFP positive cells are resistant to BstXI digestion, confirming ZFN-mediated eGFP gene correction in these cells. Furthermore, the PCR products were also cloned into pGEMT (Promega) and then several of the recombinant plasmids were sequenced. Similarly, endogenous CCR5 and CCR2 regions were PCR amplified from GFP^+ clones using locus-specific primers, the PCR products were cloned into pGEMT and five recombinant clones were sequenced for each of the above loci for each of the GFP^+ clones, namely GFP^+1, GFP^+2 and GFP^+3, respectively, and six recombinant clones in the case of the GFP^+4 clone. The locus-specific primers and PCR conditions are indicated in **Supplementary Table S2**. The DNA sequence information from the three different loci of each GFP^+ clone is shown in Table 1.

TYR-specific ZFN constructs and mouse albino melanocyte cell culture experiments:

The media and growth conditions for mouse albino melanocyte cell culture are described elsewhere (4). Mouse albino mutant melanocyte Melan-c cells were a gift from Dr. DC. Bennett (London, UK). A plasmid encoding a full-length wild type mouse tyrosinase (Tyrs-J) was obtained from Dr. Y. Yamamoto (Tohoku University, Japan). The plasmid has a pUC18 backbone and inserted full-length wild-type mouse tyrosinase. The individual ZFs, which had been previously shown in other studies to bind to component triplets of the target site, were tethered together to form the 3-finger ZFPs. They were then converted into ZFN and characterized using cell-free assays as described elsewhere (1). The ZFN constructs were then cloned into pIRES under the control of CMV promoters to form pIRES:ZFN for expression of ZFN in mammalian cells.

Amaxa-based nucleofection protocols were used to transfect albino mouse melanocytes with the efficiency of ~60% (4). Melan-c cells ($5x10^5$ cells) were nucleofected with a mixture of plasmids containing indicated amounts of pIRES:ZFNs (also known as ZFN-encoding plasmids) and exogenous Tyrs-J (also known as “correcting” template plasmid). Conditions were optimized for the transfection by fixing the amount of ZFN-encoding plasmid and varying the correcting plasmid, and vice versa. Optimal

results for nucleofection experiments were obtained at concentrations of 100-200 ng of the ZFN-encoding plasmid and 4 μg of the correcting plasmid. Cells were cultured (37^0C, 10% CO_2) for 18 hr after nucleofection without disturbance. Then, culture media was removed by aspiration and the cells were cultured in fresh medium for another 6-12 days, with culture medium changes every 3 days. Pigmented cells were observed as early as 4 days after nucleofection.

Genotypic analysis of pigmented Melan-c cells

Genomic DNA was recovered from 3 of the gene-corrected clones (that had originated from individual pigmented Melan-c cells), and the DNA was subjected to RFLP analysis to determine the genotype of the gene-corrected cells. For the analysis, a 352 bp fragment of the mouse TYR gene was amplified from genomic DNA by PCR; the amplified fragment was then digested by DdeI restriction endonuclease, and the products were analyzed using 12% PAGE (4). Presence of both 179 bp (which is a marker for the wild-type allele) and 144 bp (which is a marker for the mutant allele) fragments suggest targeted correction of one of the mutant tyrosinase alleles in the pigmented mouse albino melanocytes.

References:

(1) M. Mani, K. Kandavelou, J.F.Dy, S. Durai, S. Chandrasegaran, Design, engineering and characterization of zinc finger nucleases. Biochem Biophys Res Commun 335 (2005) 447- 457.

(2) E.E. Perez, J. Wang, J.C. Miller, Y. Jouvenot et al, Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol 26 (2008) 808–816.

(3) V. Alexeev, K. Yoon, Stable and inheritable changes in genotype and phenotype of albino melanocytes induced by an RNA-DNA oligonucleotide. Nat Biotechnol 16 (1998)1343-1347.

(4) J.C. Miller, M.C. Holmes, J. Wang, D.Y. Guschin, Y.L. Lee, et a, An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25 (2007) 778-785.

(5) M. Szczepek, V. Brondani, J. Buchel, L. Serrano, D.J. Segal, T. Cathomen, Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol 25 (2007) 786-793.

(6) F.D. Urnov, J.C. Miller, Y.L. Lee, C.M. Beausejour, J.M. Rock, et al, Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435 (2005) 646-51.

(A)

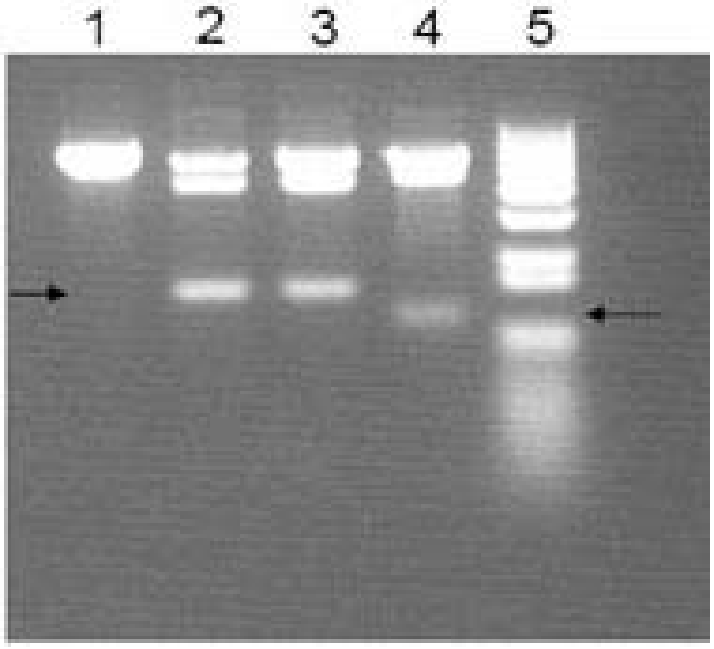

Figure S1: Analysis of the specificity of designer ZFNs. **(A)** Characterization of 3- and 4-finger ZFNs using the cell-free assays (8). The plasmid, pc.CCR5, encoding the human CCR5 cDNA, was liniearized at the unique HpaI site and used as the substrate. The expected cleavage products from the 3-finger ZFNs and the 4-finger ZFNs that overlap 3-finger sites are the same, namely 1.5 kb and 4.6 kb, respectively. The 4-finger ZFN sites that overlap the Δ32 deletion site of the hCCR5 gene are located about 300 bp downstream from the 3-finger sites; the expected products from cleavage at this site are 1.2 kb and 4.9 kb respectively. The agarose gel profile shows that the products of cleavage are consistent with the sizes for the expected products. Lanes: 1, control (substrate treated with IVTT mixture containing no ZFN plasmids); 2, +3-finger ZFNs; +4-finger ZFNs that overlap the 3-finger ZFN sites; 4, +4 finger ZFNs that overlap the Δ32 deletion site; and 5, 1 kb ladder. The expected 1.5 kb (lanes 2 & 3) and 1.2 kb (lane 4) products from the cleavage reactions are indicated by arrows.

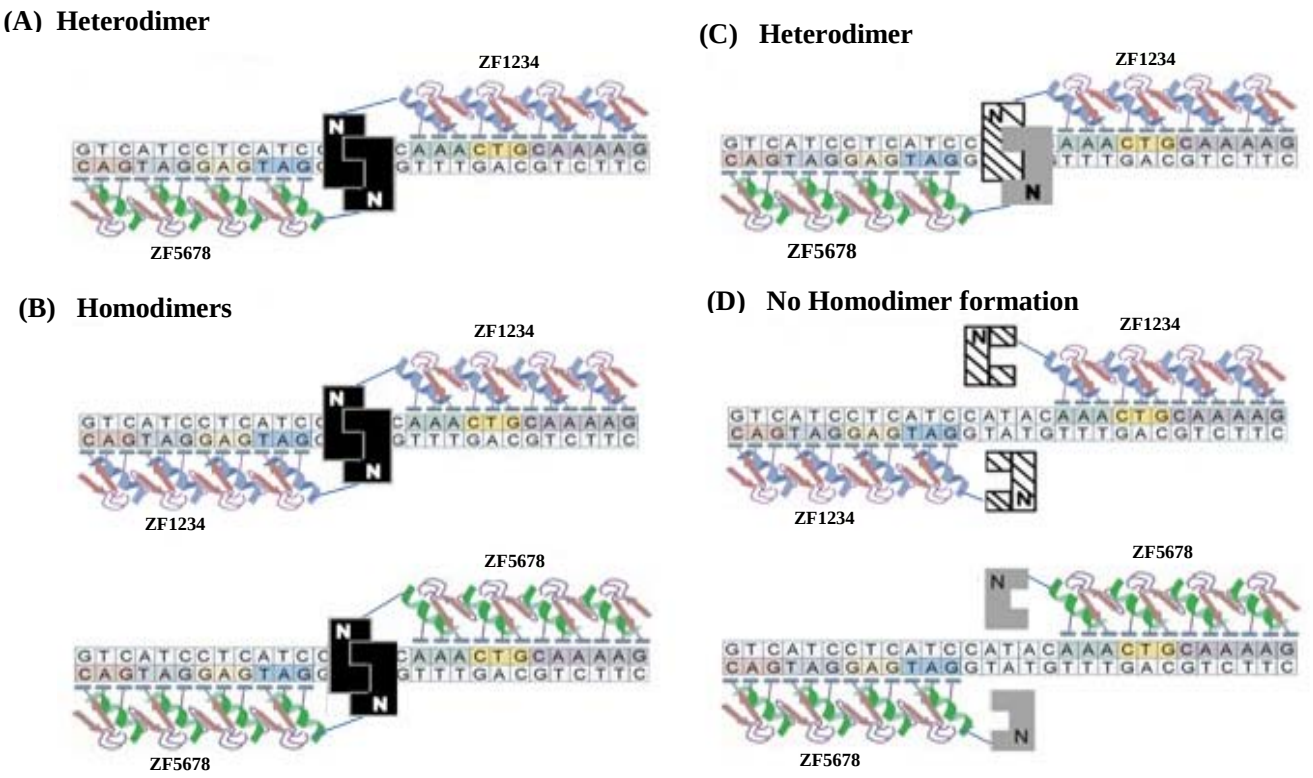

Figure S2: Re-design of the dimer interface of FokI nuclease domains to promote obligate heterodimer formation and reduce off-target cleavage. **(A)** Schematic representation of a pair of 4-finger ZFNs bound to their cognate heterodimer sites in the CCR5 gene. **(B)** Potential off-target cleavage resulting from the homodimer formation between ZFNs fused to the wild-type FokI cleavage domains. Two different homodimers could result from the two individual ZFNs with different sequence specificities. **(C & D)** The architecture of the ZFN dimer interface was re-designed to select for obligate heterodimer formation **(C)** and inhibit homodimer formation **(D)** between a pair of ZFNs to greatly reduce the off-target cleavage, and hence toxicity of ZFNs.

Table S1: ZFN designs and their target sequences within hCCR5 and mTYR genes

Number of ZFs	ZFN Designs*	ZFN Target sequences	α-helix amino acid residues at positions -1 +1 +2 +3 +4 +5 +6	Reference
3	TYR-2772-R TYR-2749-L	5'- GTG GAT GAC - 3' 5'- GAA GGG GAA- 3'	RSDALTR TTSNLAR DRSNLTR QSSNLAR RSDHLTK QSSNLAR	1
3	CCR5-276-R CCR5-247-L	5'- GCT GCC GCC - 3' 5'- GAA GGG GAC- 3‘	QSSDLTR DRSDLTR ERGTLAR QSSNLAR RSDHLTK DRSNLTR	1
4	CCR5-279-R CCR5-244-L	5'- GCT GCC GCC CAG - 3' 5'- GAA GGG GAC AGT - 3'	QSSDLTR DRSDLTR ERGTLAR RSDNLRE QSSNLAR RSDHLTK DRSNLTR HRTTLLN	2
4	CCR5-566-R CCR5-536-L	5'- GTC AGT ATC AAT -3' 5'- GAA AAT GAG AGC -3‘	DRSALAR HRTTLLN RRSACRR TTGNLTV QSGNLAR TTGNLTV RSDNLAR DASHLHT	2
4	CCR5-179-R CCR5-151-L	5'- AAA CTG CAA AAG -3' 5'- GAT GAG GAT GAC -3'	RSDNLSV QKINLQV RSDVLSE QRNHRTT DRSNLSR ISSNLNS RSDNLAR TSGNLTR	3

***R, Right; L, Left; m, mouse; h, human**

References:

(1) M. Mani, K. Kandavelou, J.F. Dy, S. Durai, S. Chandrasegaran, Design, engineering and characterization of zinc finger nucleases, Biochem Biophys Res Commun. 335 (2005) 447-457.

(2) Current manuscript.

(3) E.E. Perez, J. Wang, J.C. Miller, Y. Jouvenot Y, et al, Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases, Nat Biotechnol.26 (2008) 808–816.

Table S2: Primer sequences and amplification conditions for eGFP, CCR5 and CCR2 loci

Gene locus	Primer Name	Primer sequence	Amplification conditions
eGFP	CMV-F eGFP-R	5'- CGCAAATGGGCGGTAGGCGTGT - 3' 5'- CCGACTCGAGTTACTTGTACAGCTCGTC - 3'	95°C for 5 min, then 30 cycles of 95°C for 30sec, 55°C for 1min, followed by extension at 72°C for 5 min
CCR5	CCR5-F CCR5-R	5'- ATGGATTATCAAGTG TCAAGTCCA - 3' 5'- TCACAAGCCCACAGATATTTCC - 3'	95°C for 5 min, then 30 cycles of 95°C for 30sec, 58°C for 1min, followed by extension at 72°C for 5 min
CCR2	CCR2-F CCR2-R	5'- CAAGTGTCAAGTCCAATCTATG - 3' 5'- CCAGAATTGATACTGACTGTATG - 3'	95°C for 5 min, then 30 cycles of 95°C for 30sec, 55°C for 1min, followed by extension at 72°C for 5 min