Supplementary Figure S1

(a) Annotated AvrBs4 protein sequence

N-Terminal:

▼+287
MDPIRSRTPSPARELLPGPQPDGVQPTADRGVSPPAGGPLDGLPARRTMSRTRLPSPPAPSPAFSAGSFSD
▲Δ1

▼+153
LLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAA
▲Δ135

QPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDM

▼+49
IAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTG
▲Δ239

APLN

Aligned repeat domains:

LTPEQVVAIAS**NI**GGKQALETVQALLPVLCQAHG
LTPDQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LIPQQVVAIAS**NI**GGKQALETVQRLLPVLCQDHG
LTPQQVVAIAS**NS**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NI**GGKQALETVQRLLPVLCQDHG
LTPQQVVAIAS**NS**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NS**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGRPALE

C-Terminal:

▼+17 ▼+47
SIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVADHAQVVRV
▲Δ261 ▲Δ231

LGFFQCHSHPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARSGTLPPASQRWDRILQASGMKRAKPSPT

STQTPDQASLHAFADSLERDLDAPSPMHEGDQTRASSRKRSRSDRAVTGPSAQQSFEVRVPEQRDALHLPL

▼+231
SWRVKRPRTSIGGGLPDPGTPTAADLAASSTVMREQDEDPFAGAADDFPAFNEEELAWLMELLPQ
▲Δ47

(b) AvrBs3 protein sequence (aligned repeat domains only)

LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NS**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NS**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**NS**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGRPALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGRPALE

(c) *CCR5*-specific TALEN, left subunit (repeat domains only)

LTPQQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NI**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQALLPVLCQAHG
LTPQQVVAIAS**HD**GGRPALE

(d) *CCR5*-specific TALEN, right subunit (repeat domains only)

LTPQQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**HD**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG

LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQALLPVLCQAHG
LTPQQVVAIAS**NG**GGRPALE

(e) *IL2RG*-specific TALEN, left subunit (repeat domains only)

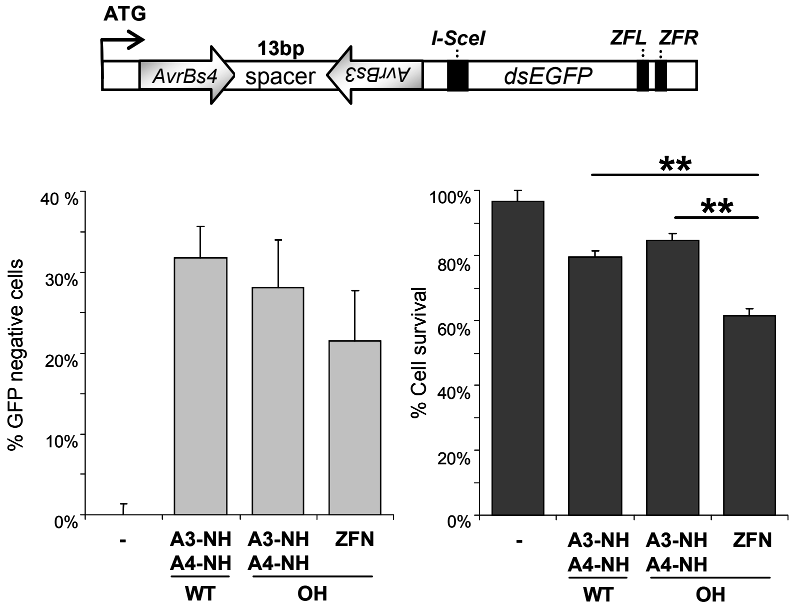
LTPQQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NN**GGKQALETVQAXXPVLCQAHG
LTPQQVVAIAS**NI**GGRPALE

(f) *IL2RG*-specific TALEN, right subunit (repeat domains only)

LTPQQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**HD**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQALLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG
LTPQQVVAIAS**NN**GGKQALETVQALLPVLCQAHG
LTPQQVVAIAS**NI**GGRPALE

Supplementary Figure S1. Amino acid sequences of TALE proteins used in this study. All TALEN are based on the AvrBs4 scaffold. (a) Sequence of AvrBs4 isolated from Xanthomonas. The N-terminal, central repeat domains, and C-terminal sequences are shown. Numbers preceded by a "+" indicate the positions of truncation variants relative to the DNA binding domain, while numbers preceded by a "Δ" indicate the deleted residues (compare to Fig. 1a). Repeat-variable di-residues (RVDs) are highlighted in bold. (b) Sequence of repeat domains of the natural Xanthomonas AvrBs3 protein. (c-f) Sequences of repeat domains of the engineered TALEs, used to target either the human *CCR5* locus (c, d) or human *IL2RG* locus (e, f).

Supplementary Figure S2



Supplementary Figure S2. Impact of the *Fok*I cleavage domain on TALEN activity. HEK293 cells harboring an integrated *dsEGFP* gene that contains a heterodimeric TALEN recognition site separated by 13-bp spacer were transfected with expression vectors encoding TALEN pairs with either wild-type (WT) or obligate heterodimeric (OH) *Fok*I variants, or a ZFN pair with OH. The graphs display the percentage of EGFP-negative cells 5 days after transfection (left panel) and relative cell survival as compared to cells expressing a nonfunctional nuclease (indicated by "–"; right panel). Statistically significant differences are indicated by ** ($P<0.01$).

Supplementary Table S1

Target	Sequence (5' to 3') §	Spacer (bp)
EBE_AvrBs4	TATAATTAATAATCCACTT	
EBE_AvrBs3	TCTATAAACCTAACCCTCT	
EBE_AvrBs4-6-AvrBs4	TATAATTAATAATCCACTTgtcagtAAGTGGATTATTAATTATA	6
EBE_AvrBs4-8-AvrBs4	TATAATTAATAATCCACTTaagaattcAAGTGGATTATTAATTATA	8
EBE_AvrBs4-9-AvrBs4	TATAATTAATAATCCACTTgaattctttAAGTGGATTATTAATTATA	9
EBE_AvrBs4-10-AvrBs4	TATAATTAATAATCCACTTgaattctttaAAGTGGATTATTAATTATA	10
EBE_AvrBs4-12-AvrBs4	TATAATTAATAATCCACTTaagaattctttaAAGTGGATTATTAATTATA	12
EBE_AvrBs4-15-AvrBs4	TATAATTAATAATCCACTTagtactgaattctttAAGTGGATTATTAATTATA	15
EBE_AvrBs4-16-AvrBs4	TATAATTAATAATCCACTTagtactgaattctttaAAGTGGATTATTAATTATA	16
EBE_AvrBs4-18-AvrBs4	TATAATTAATAATCCACTTagtgacactgaattctttAAGTGGATTATTAATTATA	18
EBE_AvrBs4-21-AvrBs4	TATAATTAATAATCCACTTagtgacactgaattcgcttttAAGTGGATTATTAATTATA	21
EBE_AvrBs4-27-AvrBs4	TATAATTAATAATCCACTTagtgacactgaattcacagtagcatttAAGTGGATTATTAATTATA	27
EBE_AvrBs4-13-AvrBs3	TATAATTAATAATCCACTTaggctcgagttacAGAGGGTTAGGTTTATAGA	13
EBE_hCCR5	TGTGGGCAACATGCTGGTCatcctcatcctgataAACTGCAAAAGGCTGAAGA	15
EBE_hCCR2	TGTGGGCAACATGCTGGTCgtcctcatcttaataAACTGCAAAAAGCTGAAGT	15
EBE_hIL2RG	TACACGTTTCGTGTTCGGAgccgctttaacccacTCTGTGGAAGTGCTCAGCA	15

§ Small letters represent the spacers

Supplementary Table S2

Number	Sequence (5' to 3')	Target
#78	tgcaccgctggataacgacat	*in vitro* cleavage
#77	aagtcgtgctgcttcatgtggt	*in vitro* cleavage
#13	tcacggggatttccaagtctc	dsGFP-F
#559	ctcggcgcgggtcttgtag	dsGFP-R
#985	aagatggattatcaagtgtcaagtcc	hCCR5-F
#986	caaagtcccactgggcg	hCCR5-R
#1093	tcagtgaagggagcagtgtg	hIL2RG-F
#1094	aacaacacgctaacccaacc	hIL2RG-R
#1474	ggcaaagactgggaagttgc	hCCR2-F
#1475	gggacagaagcaaacacagc	hCCR2-R

Nucleotide sequences of primers used for genotyping.