```
S TALENs                          (411) TCCTGCTGCTATGCCTCAT             CTTCTGGACTATCAAGGTA
Genotype
A1 (AY233288) (391) TTTTATCATATTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTATTGGTTCTTCTGGATTATCAAGGTATGTTGCCCGTTTGTCCTCTA
A2   (X70185) (391) TTTTATCATATTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGATTATCAAGGTATGTTGCCCGTTTGTCCTCTA
A3 (AB194950) (391) TTTTATCATATTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTATTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
B1 (AB073851) (391) TTTTATCATCTTCCTCTGCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
B2 (AF121245) (391) TTTTATCATCTTCCTCTGCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
B3 (EF473977) (391) TTTTATCATCTTCCTCTGCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
B4 (AB115551) (391) TTTTATCATCTTCCTCTGCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
B5 (AB219428) (391) TTTTATCATATTCCTCTGCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
B6 (AB287316) (391) TTTTATCATCTTCCTTTGCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
C1 (DQ089760) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTGTCAAGGTATGTTGCCCGTTTGTCCTCTA
C2 (AB365451) (391) TTTTATCATATTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTACCAAGGTATGTTGCCCGTTTGTCCTCTA
C3   (X75665) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTACCAAGGTATGTTGCCCGTTTGTCCTCTA
C4 (AB048704) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
C5 (EU410079) (391) TTTTATCATATTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTACCAAGGTATGTTGCCCGTTTGTCCTCTA
C6 (AB493837) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGATTACCAAGGTATGTTGCCCGTTTGTCCTCTA
D1 (AB246348) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
D2 (EU594416) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTCTGTCCTCTA
D3 (EU594435) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
D4 (AB033559) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
 E (AM494708) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
F1 (AY090458) (384) TTTTATCATATTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTCTGTCCTCTA
F2 (AY311369) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
F3 (AB036909) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
F4 (AF223965) (391) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
 G (AB375168) (391) TTTTATCATATTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTG
 H (AY090457) (394) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTGTGTCCTCTA
pCH 9/3091   (1673) TTTTATCATCTTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTATCAAGGTATGTTGCCCGTTTGTCCTCTA
            Surface  F  I  I  F  L  F  I  L  L  L  C  L  I  F  L  L  V  L  L  D  Y  Q  G  M  L  P  V  C  P  L
         Polymerase F  Y  H  L  P  L  H  P  A  A  M  P  H  L  L  V  G  S  S  G  L  S  R  Y  V  A  R  L  S  S
```

**Figure S1.** Comparison of HBV genotype sequences in the region targeted by the S TALEN. On the top row, the two 19 nucleotide sequences in the HBV plus strand that are targeted by each of the left and right subunits comprising the complete TALEN are highlighted in yellow. HBV genome coordinate of the first T residue of the left subunit target is underlined in parentheses. Cognates from HBV isolates, with 20 nt flanking on either side, are shown below. HBV genotypes, subgenotypes, accession numbers, pCH9/3091 and first nucleotide coordinate of each sequence are provided on the left in bold. Sequences within the S TALEN targets that are conserved in each of the genotypes are also highlighted in yellow. Adjacent amino acid sequences of Surface and Polymerase, by the pCH9/3091 plasmid, are shown on the bottom two rows. Individual amino acids are aligned to the first nucleotide of each codon.

```
C TALENs                        (2319) TATCAACACTTCCGGAGAC             CGACG------AGGCAGGTCCCCTA
Genotype
A1 (AY233288) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGAGACCGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
A2   (X70185) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGGGACCGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
A3 (AB194950) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCTGAAACTACTGTTGTTAGACGACGGGACCGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
          Core/HBeAg   P  N  A  P  I  L  S  T  L  P  E  T  T  V  L  R  R  R  D  R  G  R  S  P  R  R  R  T  P  S  P  R  R
Polymerase  (X70185) -  -  M  P  L  S  Y  Q  H  F  R  R  L  L  F  L  D  D  G  T  E  A  G  P  L  E  E  E  L  P  R  L  A

C TALEN                         (2319) TATCAACACTTCCGGAGAC             CGACGAGGCAGGTCCCCTA
B1 (AB073851) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAGACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
B2 (AF121245) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGAAGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
B3 (EF473977) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAGACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
B4 (AB115551) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
B5 (AB219428) (2301) CACCAAATGCCCCTATCTTATCCACACTTCCGGAAACTACTGTTGTTAGACGAAGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
B6 (AB287316) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGAGTCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
C1 (DQ089760) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
C2 (AB365451) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
C3   (X75665) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGAAGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
C4 (AB048704) (2301) CACCAAATGCCCCTATCCTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
C5 (EU410079) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
C6 (AB493837) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAAACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
D1 (AB246348) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAGACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
D2 (EU594416) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAGACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
D3 (EU594435) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAGACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
D4 (AB033559) (2301) CACCAAATGCCCCTATCCTATCAACACTTCCGGAGACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
 E (AM494708) (2301) CACCAAATGCCCCTATCTTATCAACACTTCCGGAGAATACTGTTGTTAGACGAAGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
F1 (AY090458) (2304) CACCAAATGCCCCTATCTTATCCACACTTCCGGAAACTACTGTTGTTAGACGAAGAGGCAGGTCTCCTCGAAGAAGAACTCCCTCGCCTCGCAG
F2 (AY311369) (2301) CACCAAATGCCCCTATCCTATCCACACTTCCGGAAACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
F3 (AB036909) (2301) CACCAAATGCCCCTATCCTATCCACACTTCCGGAAACTACTGTTGTTAGACGAAGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCCG
F4 (AF223965) (2301) CACCAAATGCCCCTATCCTATCCACACTTCCGGAAACTACTGTTGTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
 G (AB375168) (2337) CACCAAATGCCCCTATCCTATCAACACTTCCGGAGACTACTGTTGTTAGACGAAGAGGCAGGTCCCCTCGAAGAAGAACTCCCTCGCCTCGCAG
 H (AY090457) (2304) CACCAAATGCCCCTATCCTATCAACACTTCCGGAGACTACTGTTGTTAGACAACGAGGCAGGGCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
pCH 9/3091     (401) CACCAAATGCCCCTATCCTATCAACACTTCCGGAGACTACTGTTCTTAGACGACGAGGCAGGTCCCCTAGAAGAAGAACTCCCTCGCCTCGCAG
          Core/HBeAg   P  N  A  P  I  L  S  T  L  P  E  T  T  V  L  R  R  R  G  R  S  P  R  R  R  T  P  S  P  R
          Polymerase -  -  M  P  L  S  Y  Q  H  F  R  R  L  L  F  L  D  D  E  A  G  P  L  E  E  E  L  P  R  L  A
```

**Figure S2.** Comparison of HBV genotype sequences in the region targeted by the C TALEN. The sequences are depicted using a similar format to that described in Figure S1. HBV genotype A isolates have an additional 6 nucleotides in the in the sequence targeted by the right TALEN subunit.

```
P1 TALENs                              (1144) TACCCCGTTGCCCGGCAAC             CCAAGTGTTTGCTGACGCA
Genotype
A1 (AY233288) (1119) TCTGAGTAAACAGTATATGAACCTTTACCCCGTTGCCCGGCAACGGCCTGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCT
A2   (X70185) (1119) TCTAAGTAAACAGTACATGAACCTTTACCCCGTTGCTCGGCAACGGCCTGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCC
A3 (AB194950) (1119) TCTAAGTAAACAGTATATGAACCTTTACCCCGTTGCCCGGCAACGGCCTGGTCTATGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCC
B1 (AB073851) (1119) TCTAAGTAAACAGTATATGACCCTTTACCCCGTTGCTCGGCAACGGCCTGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCC
B2 (AF121245) (1119) TTTAAGTAAACAGTATCTGAACCTTTACCCCGTTGCTCGGCAACGGCCTGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCC
B3 (EF473977) (1119) TCTAAGCAAACAATATCTGAACCTTTACCCCGTTGCGCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCC
B4 (AB115551) (1119) TTTAAGTAAACAGTATCTGAACCTTTACCCCGTTGCTCGGCAACGGCCTGGTCTATGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCC
B5 (AB219428) (1119) TCTACACAAACAGTATTTGAACCTTTACCCCGTTGCTCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCC
B6 (AB287316) (1119) TCTAAACAAACAATATCTGAACCTTTACCCCGTTACTCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCT
C1 (DQ089760) (1119) TCTGTGTAAACAATATCAGAACCTTTACCCCGTTGCTCGGCAACGGTCAGGTCTTTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCC
C2 (AB365451) (1119) TCTGTGTCAACAATACCTGCACCTTTACCCCGTTGCCCGGCAACGGTCAGGCCTCTGCCAAGTGTTTGCTGACGCAACCCCCACTGGATGGGGCTTGGCC
C3   (X75665) (1119) TCTGTGTAAACAATATCTGAACCTTTACCCCGTTGCCCGGCAACGGTCTGGTCTTTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCC
C4 (AB048704) (1119) TCTGTGTAAACAATATCTGAACCTTTACCCCGTTGCCCGGCAACGGGCTGGTCTCTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCC
C5 (EU410079) (1119) TCTGTGTAAACAATATTTGACCCTTTATCCCGTTGCTCGGCAACGGCCAGGTCTATGCCAAGTGTTTGCTGATGCAACCCCCACTGGTTGGGGCTTGGCC
C6 (AB493837) (1119) TCTGTGTCAACAATATTTGCACCTTTACCCCGTTGCCCGGCAACGGTCTGGTCTTTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCA
D1 (AB246348) (1119) TCTGTGTAAACAATACATGACCCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCC
D2 (EU594416) (1119) TCTGTGTAAACAATATCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTT
D3 (EU594435) (1119) TCTGTATAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTC
D4 (AB033559) (1119) TCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCAGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGATGCAACCCCCACTGGGTGGGGCTTGGTC
 E (AM494708) (1119) TCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCCGGTCTGTGCCAAGTGTTTGCTGATGCAACCCCCACTGGCTGGGGCTTGGCC
F1 (AY090458) (1122) TCTATGTCAACAATACATGAACCTTTACCCCGTTGCTCGGCAACGGCCAGGCCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCC
F2 (AY311369) (1119) TCTGTGTAAACAATACATGAACCTTTACCCCGTTGCTCGGCAACGGCCAGGCCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCC
F3 (AB036909) (1119) TCTCTGTAAACAATACATGAACCTTTACCCCGTTGCTCGGCAACGGCCAGGCCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCC
F4 (AF223965) (1119) TCTCTGTAAACAGTACATGAACCTTTACCCCGTTGCTCGGCAACGGCCAGGCCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCC
 G (AB375168) (1119) TCTCTGTAAACAATACATGAACCTTTACCCCGTTGCTAGGCAACGGCCCGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCTTGGCC
 H (AY090457) (1122) TCTCTGTAAACAATACATGAACCTTTACCCCGTTGCTCGGCAACGGCCAGGCCTTTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGCG
pCH 9/3091    (2401) TCTGTGTAAACAATACCTGAACCTTTACCCCGTTGCCCGGCAACGGCCAGGTCTGTGCCAAGTGTTTGCTGACGCAACCCCCACTGGCTGGGGCTTGGTC
          Polymerase  L  C  K  Q  Y  L  N  L  Y  P  V  A  R  Q  R  P  G  L  C  Q  V  F  A  D  A  T  P  T  G  W  G  L  V
```

**Figure S3.** Comparison of HBV genotype sequences in the region targeted by the P1 TALEN. The sequences are depicted using a similar format to that described in Figure S1.

```
P2 TALENs                      (1299) TCGCAGCAGGTCTGGAGCA              TGATAACTCTGTTGTCCTA
Genotype
A1 (AY233288) (1274) CGGAACTCCTAGCTGCTTGTTTTGCTCGCAGCCGGCCTGGAGCAAAACTCATCGGGACTGATAATTCTGTCGTCCTTTCTCGGAAATATACATC
A2   (X70185) (1274) CGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCCGGTCTGGAGCAAAGCTCATCGGAACTGACAATTCTGTCGTCCTCTCGCGGAAATATACATC
A3 (AB194950) (1274) CGGAACTCCTAGCAGCTTGTTTTGCTCGCAGCCGGTCTGGAGCAAAACTCATCGGGACTGATAATTCTGTCGTCCTTTCTCGGAAATATACATC
B1 (AB073851) (1274) CGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCGAAACTCATCGGGACTGACAATTCTGTCGTGCTCTCCCGCAAGTATACATC
B2 (AF121245) (1274) CGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGGGCAAAACTCATCGGGACTGACAATTCTGTCGTACTCTCCCGCAAGTATACAGC
B3 (EF473977) (1274) CGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCCGGTCTGGAGCGAAACTTATCGGGACTGACAATTCTGTTGTCCTTTCCCGCAAATATACATC
B4 (AB115551) (1274) CGGAACTCCTTGCAGCTTGTTTTGCTCGCAGCCGGTCTGGAGCAAAACTTATCGGGACTGACAATTCTGTCGTGCTATCCCGCAAGTATACATC
B5 (AB219428) (1274) CGGAACTCCTAGCTGCTTGTTTTGCTCGCAGCAGGTCTGGAGCGAAACTTATCGGGACCGATAATTCTGTCGTCTTGTCCCGTAAATATACATC
B6 (AB287316) (1274) CGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGGGCGAACCTCATCGGGACTGACAATTCGGTTGTCCTTTCCCGCAAATATACATC
C1 (DQ089760) (1274) CGGAACTCCTAGCAGCTTGTTTTGCTCGCAGCCGGTCTGGAGCGAAACTTATCGGCACCGACAACTCTGTTGTCCTCTCTCGGAAATACACCTC
C2 (AB365451) (1274) CGGAACTCCTAGCAGCTTGTTTTGCTCGCAGCCGGTCTGGAGCAAAACTTATCGGGACTGACAACTCTGTTGTCCTCTCTCGGAAATACACCTC
C3   (X75665) (1274) CGGAACTCCTAGCGGCTTGTTTTGCTCGCAGCCGGTCTGGAGCAAACATTATCGGAACCGACAACTCTGTCGTCCTCTCTCGGAAATACACATC
C4 (AB048704) (1274) CGGAACTCCTAGCAGCTTGTTTCGCTCGCAGCCGGTCTGGAGCGAACATTCTCGGGACCGACAACTCTGTTGTTCTCTCTCGGAAATACACCTC
C5 (EU410079) (1274) CGGAACTCCTTGCAGCCTGTTTTGCTCGCAGCCGGTCTGGAGCAAACCTTATAGGAACCGACAACTCTGTTGTCCTTTCTCGGAAATACACCTC
C6 (AB493837) (1274) CGGAACTCCTAGCAGCTTGTTTTGCTCGCAGCCGGTCTGGAGCGAAACTTATCGGAACCGACAATTCCGTTGTCCTCTCTCGGAAATACACCTC
D1 (AB246348) (1274) CGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTCTCGGGACGGATAACTCTGTTGTTCTCTCCCGCAAATATACATC
D2 (EU594416) (1274) CGGAACTCCTAGCCGCTTGCTTTGCTCGCAGCAGGTCTGGAGCAAACATTCTCGGGACGGATAACTCTGTTGTTCTCTCCCGCAAATATACATC
D3 (EU594435) (1274) CGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTCTCGGGACTGATAACTCTGTTGTCCTCTCCCGCAAATATACATC
D4 (AB033559) (1274) CGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTCTCGGAACTGACAACTCTGTTGTCCTCTCCCGCAAATATACATC
 E (AM494708) (1274) CGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCGAAACTTATCGGGACGGATAATTCTGTCGTTCTCTCCCGGAAATATACATC
F1 (AY090458) (1277) CAGAACTCCTTGCAGCTTGTTTCGCTCGCAGCCGGTCTGGAGCGAAACTCATCGGCACAGACAACTCTGTTGTCCTCTCTAGGAAGTACACCTC
F2 (AY311369) (1274) CGGAACTCCTAGCAGCTTGTTTCGCTCGCAGCCGGTCTGGAGCGACTCTTAATGGCACGGACAATTCTGTTGTCCTATCTAGGAAGTACACCTC
F3 (AB036909) (1274) CGGAACTCCTTGCAGCCTGTTTCGCTCGCAGCCGGTCTGGAGCGAACATTATCGGCACAGACAACTCTGTTGTCCTCTCTAGGAAGTACACCTC
F4 (AF223965) (1274) CGGAACTCCTTGCAGCTTGTTTCGCTCGCAGCCGGTCTGGAGCGAATATCATCGGGACAGACAACTCTGTTGTACTCTCTAGGAAGTACACCTC
 G (AB375168) (1274) CGGAACTCCTAGCTGCTTGTTTTGCTCGCAGCCGGTCTGGAGCAAAACTCATTGGGACTGACAATTCTGTCGTCCTTTCTCGGAAATATACATC
 H (AY090457) (1277) CGGAACTCCTAGCAGCTTGTTTCGCTCGCAGCCGGTCTGGAGCGGACATTATCGGCACTGACAACTCCGTTGTCCTGTCTCGGAAGTACACCTC
pCH 9/3091    (2556) CGGAACTCCTAGCCGCTTGTTTTGCTCGCAGCAGGTCTGGAGCAAACATTATCGGGACTGATAACTCTGTTGTCCTATCCCGCAAATATACATC
         Polymerase   E  L  L  A  A  C  F  A  R  S  R  S  G  A  N  I  I  G  T  D  N  S  V  V  L  S  R  K  Y  T
```

**Figure S4.** Comparison of HBV genotype sequences in the region targeted by the P2 TALEN. The sequences are depicted using a similar format to that described in Figure S1.

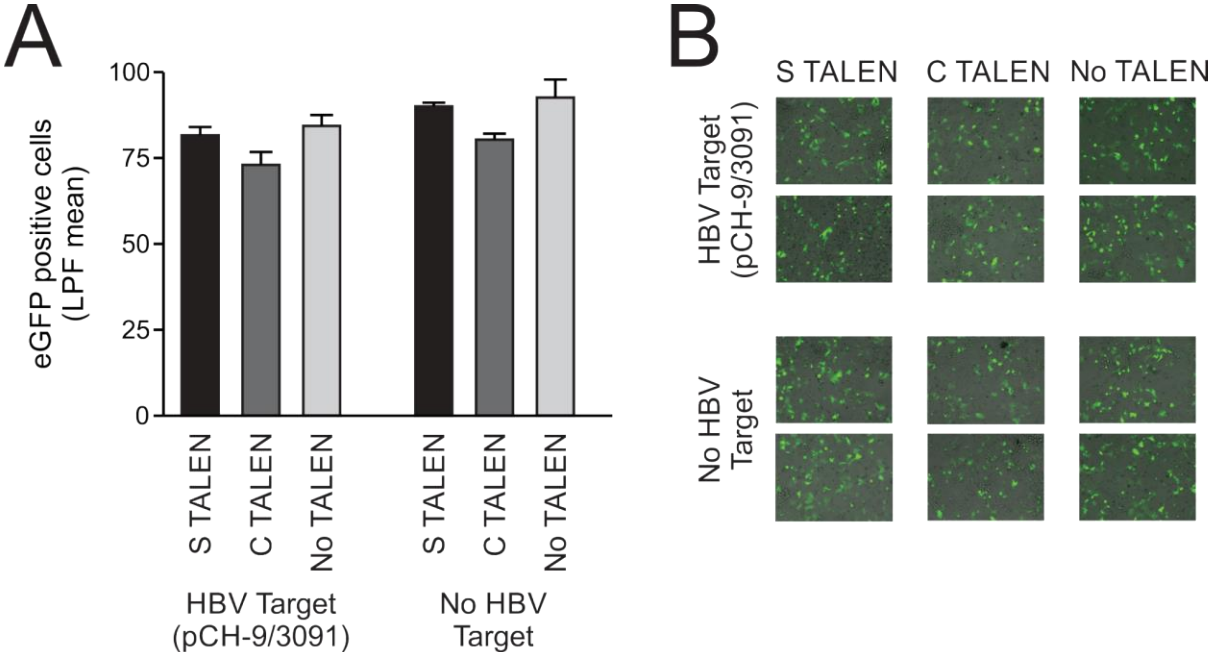

**Figure S5.** Survival assay after co-transfection with GFP- and TALEN-expressing plasmids. **A.** Following transfection of Huh7 cells with plasmids encoding TALENs and GFP in the presence or absence of HBV target DNA, GFP-positive cells were counted in three representative low power fields (LPF, 100 ×) of experiments performed in triplicate. Mean numbers of fluorescent cells observed are shown with standard errors of the mean indicated by error bars. **B.** Representative low power fields of cells transfected with each of the plasmid combinations.

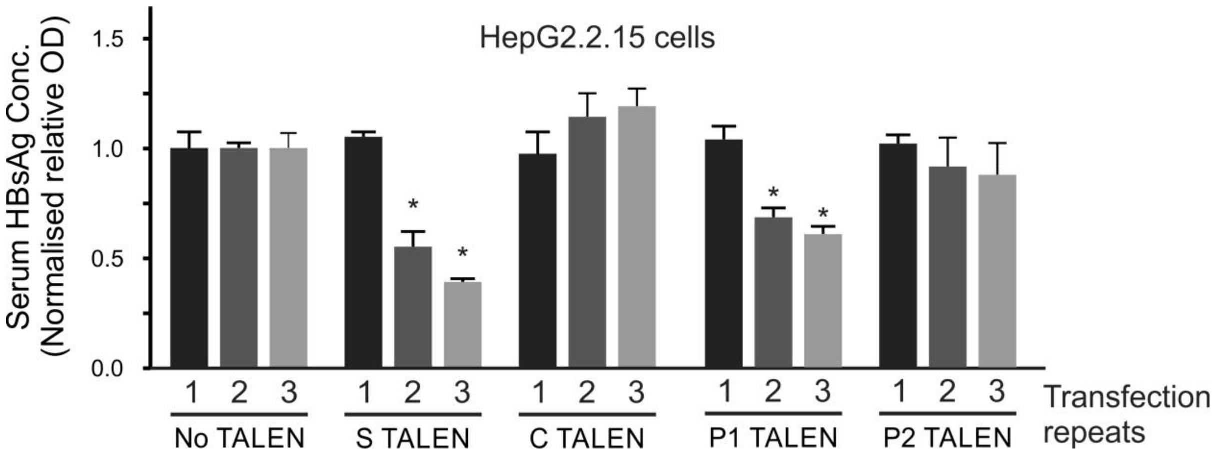

**Figure S6.** HBsAg concentrations in culture supernatants of HepG2.2.15 cells that were subjected to repeat transfections with plasmids encoding each of the panel of 4 HBV-targeting TALENs under hypothermic (30°C) temperatures. Data are represented as the means and the error bars indicate the standard errors of the mean. Statistically significant differences between values obtained after transfection of cells with each of the TALENs and values without TALEN transfection (No TALEN) are indicated by asterisks (*: $p < 0.05$).

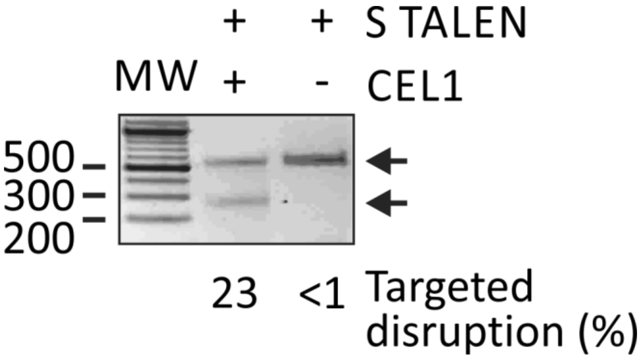

**Figure S7.** Detection of targeted HBV DNA disruption using CEL1 nuclease assay. cccDNA was isolated from HepG2.2.15 cells that had been subjected to three repeat transfections with the S TALEN-encoding plasmids under hypothermic conditions. Amplicons were subjected to digestion with the CEL1 endonuclease. The arrows indicate the larger intact PCR product (520 bp) and smaller digested fragment (~280 bp). The target disruption, measured as the ratio of smaller fragment intensity to total DNA in the lane, is indicated below as a percentage. Sizes of DNA fragments (bp) in the molecular weight marker lane (MW) are shown on the left.

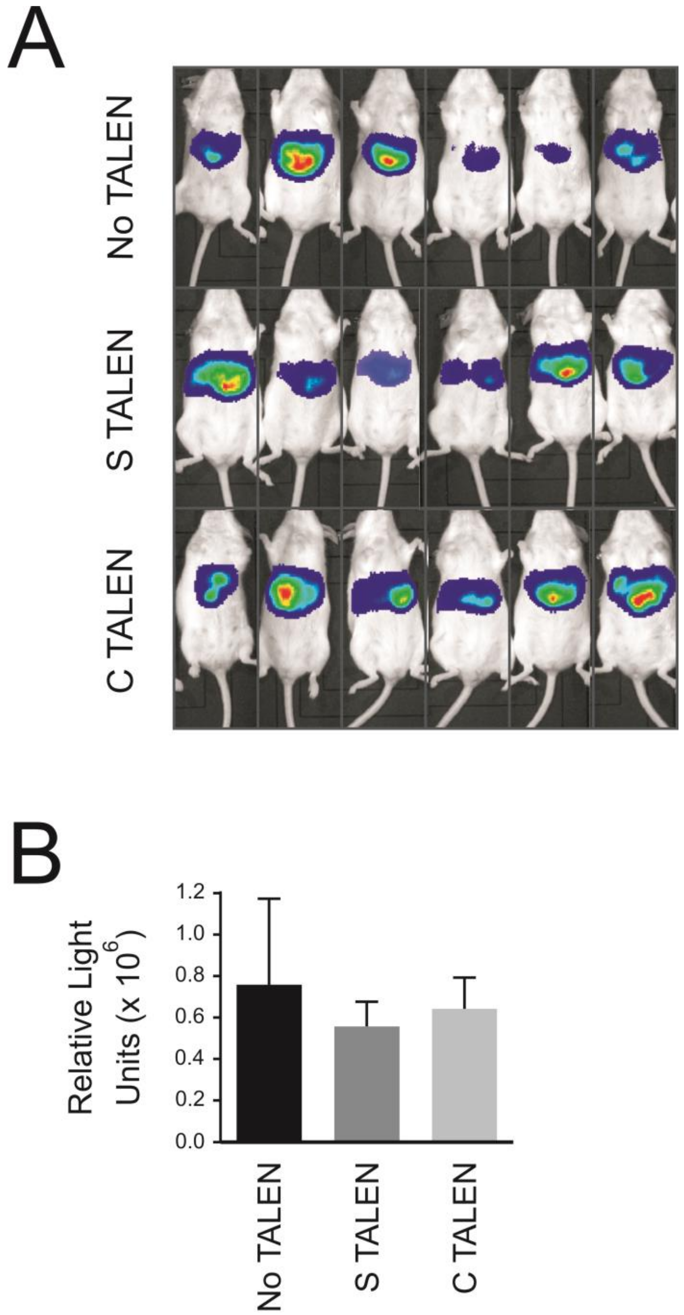

**Figure S8.** Equivalent reporter gene expression in groups of mice that were subjected to HDI. **A.** Each animal from the three groups of mice was subjected to bioluminescence imaging following injection of the TALENs, HBV sequences and a Firefly luciferase-encoding plasmid. **B.** Mean relative light units emitted from the hepatic region was determined for each of the groups of animals. Values are presented as the means with error bars indicating standard deviations.

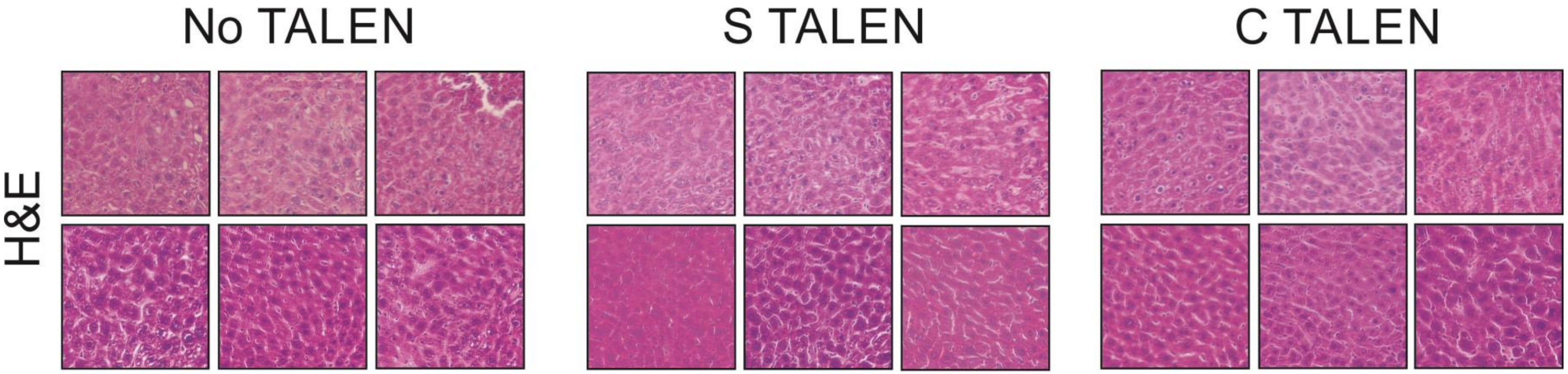

**Figure S9.** Representative liver histology (hematoxylin and eosin staining) obtained from liver sections of each animal that had been subjected to HDI with or without TALEN-expressing plasmids. Low power fields (100 ×) are shown.

**Table S1.** BLAST analysis of the *Mus musculus* genome to identify potential binding sites of combined and individual left and right subunits (SL and SR) of the S TALEN.

| | **Chrom. No.** | **Nucleotide Range** | **S TALEN Sites** | | **Identities**$^{\#}$ | **Sequence ID** |
|---|---|---|---|---|---|---|
| | | | **SL** | **SR** | | |
| | | | ***TCCTGCTGCTATGCCTCAT***<br>AGGACGACGATACGGAGTA | ***CTTCTGGACTATCAAGGTA***<br>GAAGACCTGATAGTTCCAT | | |
| **Potential combined SL and SR sites** | 6 | 6830661 to 6830677 | ●CCTGCTGCTATGCCTCA●* | | 17/19 | NW_001030820.1 |
| | | 11962065 to 11962080 | TCCTGCTGCTATGCCT●●● | | 16/19 | |
| | | 41232469 to 41232483 | | CTTCTGGACTATCAA●●●● | 15/19 | |
| | 4 | 44188029 to 44188045 | ●●CTGCTGCTATGCCTCAT | | 17/19 | NT_187032.1 |
| | | 31532806 to 31532820 | ●CCTGCTGCTATGCCT●●● | | 15/19 | |
| | | 8318087 to 8318104 | ●CCTGCTGCTA●GCCTCAT | | 15/19 | |
| | | 36732721 to 36732735 | | ●TTCTGGACTATCAAG●●● | 17/19 | |
| | 4 | 22544951 to 22544966 | TCCTGCTGCTATGCCT●●● | | 16/19 | NT_187033.1 |
| | | 10954765 to 10954779 | | ●●●●TGGACTATCAAGGTA | 15/19 | |
| | 8 | 9218161 to 9218176 | TCCTGCTGCTATGCCT●●● | | 16/19 | NT_078575.7 |
| | | 26838227 to 26838244 | | ●TTCTGGA●TATCAAGGTA | 17/19 | |
| | 9 | 3697632 to 3697647 | TCCTGCTGCTATGCCT●●● | | 16/19 | NT_039474.8 |
| | | 44495841 to 44495855 | | CTTCTGGACTATCAA●●●● | 15/19 | |
| | 13 | 1175041 to 1175059 | TCCTGCTGCT●TGCCTCAT | | 18/19 | NW_001030529.1 |
| | | 4966540 to 4966554 | | ●●●●TGGACTATCAAGGTA | 15/19 | |
| | 9 | 43318590 to 43318605 | TCCTGCTGCTATGCCT●●● | | 16/19 | NW_001030907.1 |
| | | 44102972 to 44102987 | TCCTGCTGCTATGCCT●●● | | 16/19 | |
| | | 10709989 to 10710005 | | CTTCTGGACTA●CAAGG●● | 16/19 | |
| | 7 | 1111381 to 1111396 | TCCTGCTGCTATGCCT●●● | | 16/19 | NT_039433.8 |
| | | 38299233 to 38299248 | TCCTGCTGCTATGCCT●●● | | 16/19 | |
| | | 53113291 to 53113306 | TCCTGCTGCTATGCCT●●● | | 16/19 | |
| | | 50396556 to 50396574 | TCCTGCTGCT●TGCCTCAT | | 18/19 | |
| | | 2339469 to 2339486 | TCCTGCTGCTATG●CTCA● | | 17/19 | |
| | | 8016639 to 8016656 | TCCTGCTGCT●TGCCTCA● | | 17/19 | |
| | | 2033241 to 2033255 | | CTTCTGGACTATCAA●●●● | 15/19 | |
| | 3 | 41300341 to 41300356 | TCCTGCTGCTATGCCT●●● | | 16/19 | NT_039240.8 |
| | | 27971181 to 27971195 | | ●TTCTGGACTATCAAG●●● | 15/19 | |
| | | 66053798 to 66053816 | | CTTCTGG●CTATCAAGGTA | 18/19 | |
| | 17 | 69593674 to 69593692 | TCCTGCT●CTATGCCTCAT | | 18/19 | NT_039649.8 |
| | | 30114485 to 30114500 | | ●●TCTGGACTATCAAGGT● | 16/19 | |

|  |  |  |  |  |  |  |
|---|---|---|---|---|---|---|
| **Potential individual SL sites** | 13 | 17430998 to 17431014 | TCCTGCTGCTATGCCTC●● |  | 17/19 | NW_001030517.1 |
|  |  | 2227826 to 2227841 | TCCTGCTGCTATGCCT●●● |  | 16/19 |  |
|  | 6 | 845523 to 845539 | ●CCTGCTGCTATGCCTCA● |  | 17/19 | NT_039360.8 |
|  |  | 5969096 to 5969111 | TCCTGCTGCTATGCCT●●● |  | 16/19 |  |
|  | 4 | 12854748 to 12854763 | TCCTGCTGCTATGCCT●●● |  | 16/19 | NT_039264.7 |
|  |  | 25385538 to 25385553 | TCCTGCTGCTATGCCT●●● |  | 16/19 |  |
|  | 9 | 36145333 to 36145348 | TCCTGCTGCTATGCCT●●● |  | 16/19 | NT_039472.8 |
|  |  | 36938186 to 36938201 | TCCTGCTGCTATGCCT●●● |  | 16/19 |  |
|  | 14 | 28926064 to 28926081 | ●CCTGCTGCTATGCCTCAT |  | 18/19 | NT_039606.8 |
|  |  | 60263242 to 60263256 | ●CCTGCTGCTATGCCT●●● |  | 15/19 |  |
|  | 19 | 49262317 to 49262333 | TCCTGCTGCTATGCCTC●● |  | 17/19 | NT_039687.8 |
|  |  | 2286312 to 2286326 | ●●CTGCTGCTATGCCTC●● |  | 15/19 |  |
|  |  | 15839462 to 15839476 | TCCTGCTGCTATGCC●●●● |  | 15/19 |  |
|  | 7 | 22820518 to 22820533 | TCCTGCTGCTATGCCT●●● |  | 16/19 | NW_001030877.1 |
|  |  | 37540231 to 37540246 | TCCTGCTGCTATGCCT●●● |  | 16/19 |  |
|  |  | 34806807 to 34806825 | TCCTGCTGCT●TGCCTCAT |  | 18/19 |  |
|  | 11 | 55118539 to 55118554 | TCCTGCTGCTATGCCT●●● |  | 16/19 | NT_096135.6 |
|  |  | 60207885 to 60207900 | TCCTGCTGCTATGCCT●●● |  | 16/19 |  |
| **Potential individual SR sites** | 3 | 27971181 to 27971195 |  | ●TTCTGGACTATCAAG●●● | 15/19 | NT_039240.8 |
|  |  | 66053798 to 66053816 |  | CTTCTGG●CTATCAAGGTA | 18/19 |  |
|  | 11 | 8445188 to 8445204 |  | CTTCTGG●CTATCAAGG●● | 16/19 | NW_001030434.1 |
|  |  | 12712726 to 12712742 |  | ●TTCTG●ACTATCAAGGT● | 16/19 |  |
|  | 2 | 16581880 to 16581894 |  | CTTCTGGACTATCAA●●●● | 15/19 | NT_039207.8 |
|  |  | 25991220 to 25991236 |  | ●TTCTG●ACTATCAAGGT● | 16/19 |  |
|  |  | 68974756 to 68974772 |  | CTTCTGGACT●TCAAGG●● | 16/19 |  |
|  | 16 | 61477073 to 61477087 |  | ●TTCTGGACTATCAAG●●● | 15/19 | NW_001030584.1 |
|  |  | 36662006 to 36662022 |  |  | 16/19 |  |
|  |  | 71304901 to 71304917 |  | CTTCTGGACTA●CAAGG●● | 16/19 |  |
|  |  |  |  | ●TTCTGGA●TATCAAGGT● |  |  |
|  | 19 | 26586727 to 26586741 |  | CTTCTGGACTATCAA●●●● | 15/19 | NT_039687.8 |
|  |  | 26777773 to 26777787 |  | CTTCTGGACTATCAA●●●● | 15/19 |  |
|  |  | 26955911 to 26955925 |  | CTTCTGGACTATCAA●●●● | 15/19 |  |

#A minimum sequence identity between TALEN subunit cognates and sites within the *Mus musculus* genome of 15 out of 19 sense or antisense nucleotides was used as a cut off.

*Mismatches between the TALEN subunit cognates and the aligned sequence in the *Mus musculus* genome are indicated by '●'.

**Table S2.** BLAST analysis of the human genome to identify potential binding sites of combined and individual left and right subunits (SL and SR) of the S TALEN.

| **Potential combined SL and SR sites** | **Chrom. No.** | **Nucleotide Range** | **S TALEN Sites** | | **Identities** | **Sequence ID** |
|---|---|---|---|---|---|---|
| | | | **SL** | **SR** | | |
| | | | ***TCCTGCTGCTATGCCTCAT***<br>AGGACGACGATACGGAGTA | ***CTTCTGGACTATCAAGGTA***<br>GAAGACCTGATAGTTCCAT | | |
| | 5 | 5316824 to 5316838 | ●●CTGCTGCTATGCCTC●● | | 15/19 | NW_004078023.1 |
| | | 1348622 to 1348638 | | ●TTCTGGACTATC●AGGT● | 16/19 | |
| | 6 | 16508260 to 16508274 | ●●CTGCTGCTATGCCTC●● | | 15/19 | NW_004078028.1 |
| | | 36973928 to 36973944 | ●●CTGCTGCTAT●CCTCAT | | 16/19 | |
| | | 28906310 to 28906326 | | ●TTCTGGACTAT●AAGGT● | 16/19 | |
| | | 66068358 to 66068374 | | ●TTCTGGACTATC●AGGT● | 16/19 | |
| | 3 | 52993974 to 52993988 | TCCTGCTGCTATGCC●●●● | | 15/19 | NT_005612.16 |
| | | 81226837 to 81226851 | ●●●TGCTGCTATGCCTCA● | | 15/19 | |
| | | 9589025 to 9589042 | | ●TTCTGGA●TATCAAGGTA | 17/19 | |
| | 17 | 28496793 to 28496807 | ●●●●GCTGCTATGCCTCAT | | 15/19 | NT_010783.15 |
| | | 35558057 to 35558074 | ●CCTGCTGCT●TGCCTCAT | | 17/19 | |
| | | 13845926 to 13845940 | | ●●●●TGGACTATCAAGGTA | 15/19 | |
| | 16 | 9515596 to 9515610 | ●●●TGCTGCTATGCCTCA● | | 15/19 | NT_010393.16 |
| | | 10180724 to 10180741 | | CTTCTGGACTATCAAGGT● | 18/19 | |
| | | 20197304 to 20197318 | | ●●●●TGGACTATCAAGGTA | 15/19 | |
| | 4 | 28481514 to 28481528 | TCCTGCTGCTATGCC●●●● | | 15/19 | NT_016354.19 |
| | | 86675613 to 86675630 | ●CCTGCTGCT●TGCCTCAT | | 17/19 | |
| | | 5401736 to 5401752 | TCCTGCTGC●ATGCCTC●● | | 16/17 | |
| | | 26630796 to 26630810 | | ●●●●TGGACTATCAAGGTA | 15/19 | |
| | | 91203543 to 91203560 | | ●TTCTGGACTAT●AAGGTA | 17/19 | |
| | | 42836405 to 42836421 | | CTTCTG●ACTATCAAGG●● | 16/19 | |
| | 19 | 8260049 to 8260063 | TCCTGCTGCTATGCC●●●● | | 15/19 | NT_011109.16 |
| | | 5398156 to 5398173 | | CTTCT●GACTATCAAGGT● | 17/19 | |
| | 9 | 32244124 to 32244141 | ●CCTGCTGCTA●GCCTCAT | | 17/19 | NT_008413.18 |
| | | 26866298 to 26866315 | | CTTCTGGA●TATCAAGGT● | 17/19 | |
| | 1 | 61198693 to 61198710 | TCCTGCTGCT●TGCCTCA● | | 17/19 | NT_032977.9 |
| | | 11611961 to 11611975 | | ●●TCTGGACTATCAAGG●● | 15/19 | |
| | | 47930542 to 47930556 | | ●●●CTGGACTATCAAGGT● | 15/19 | |
| | | 66083096 to 66083110 | | ●●●CTGGACTATCAAGGT● | 15/19 | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 6 | 10893712 to 10893726 | ●CCTGCTGCTATGCCT●●● | | 15/19 | NT_007299.13 |
| | | 9894805 to 9894819 | | ●TTCTGGACTATCAAG●●● | 15/19 | |
| | | 14658716 to 14658733 | | ●TTCTGGACT●TCAAGGTA | 17/19 | |
| **Potential individual SL sites** | 22 | 13395721 to 13395736 | ●CCTGCTGCTATGCCTC●● | | 16/19 | NW_004078112.1 |
| | | 23229742 to 23229758 | TCCTGCTGCTAT●CCTC●● | | 16/19 | |
| | | 26356393 to 26356409 | TCCTGCTGCT●TGCCTC●● | | 16/19 | |
| | | 30107958 to 30107974 | ●CCTGCTGCT●TGCCTCA | | 16/19 | |
| | 15 | 69598949 to 69598963 | ●CCTGCTGCTATGCCT●●● | | 15/19 | NW_004078084.1 |
| | | 7011373 to 7011389 | ●●CTGCTGCT●TGCCTCAT | | 16/19 | |
| | | 42322452 to 42322468 | ●●CTGCTGCTAT●CCTCAT● | | 16/19 | |
| | 17 | 12292150 to 12292164 | ●●●●GCTGCTATGCCTCAT | | 15/19 | NW_001838454.2 |
| | | 5230569 to 5230586 | ●CCTGCTGCT●TGCCTCAT | | 17/19 | |
| | 3 | 11943588 to 11943602 | ●CCTGCTGCTATGCCT●●● | | 15/19 | NW_001838877.2 |
| | | 20444778 to 20444795 | TCCTGCTGCTAT●CCTCA● | | 17/19 | |
| | | 1113755 to 1113771 | ●●CTGCTGCTATGC●TCAT | | 16/19 | |
| | 11 | 3402649 to 3402663 | ●●CTGCTGCTATGCCTC●● | | 15/19 | NT_033899.8 |
| | | 21709223 to 21709239 | TCCTGCTGCT●TGCCTC●● | | 16/19 | |
| | 6 | 16509724 to 16509738 | ●●CTGCTGCTATGCCTC●● | | 15/19 | NT_025741.15 |
| | | 36971717 to 36971733 | ●●CTGCTGCTAT●CCTCAT | | 16/19 | |
| **Potential individual SR sites** | 13 | 67246550 to 67246566 | | CTTCTGGACTATCAAGG●● | 17/19 | NT_024524.14 |
| | | 36698126 to 36698140 | | CTTCTGGACTATCAA●●●● | 15/19 | |
| | | 37005816 to 37005830 | | ●●●CTGGACTATCAAGGT● | 15/19 | |
| | | 32823986 to 32824003 | | ●TTCTGGA●TATCAAGGTA | 17/19 | |
| | 21 | 8582147 to 8582162 | | ●●●CTGGACTATCAAGGTA | 16/19 | NT_011512.11 |
| | | 26035956 to 26035973 | | CTTCTGGACTATC●AGGT● | 17/19 | |
| | 12 | 2866821 to 2866835 | | ●●●CTGGACTATCAAGGT● | 15/19 | NW_001838060.2 |
| | | 485378 to 485394 | | ●●TTCTGGACTATCCAGGT● | 16/19 | |
| | 1 | 5704171 to 5704185 | | ●●●CTGGACTATCAAGGT● | 15/19 | NW_001838579.2 |
| | | 23791575 to 23791591 | | CTTCTGGACTA●CAAGG●● | 16/19 | |
| | 9 | 49639304 to 49639318 | | ●TTCTGGACTATCAAG●●● | 15/19 | NT_008470.19 |
| | | 30386318 to 30386335 | | ●TTCTGGACTA●CAAGGTA | 17/19 | |

#A minimum sequence identity between TALEN subunit cognates and sites within the human genome of 15 out of 19 sense or antisense nucleotides was used as a cut off.

*Mismatches between the TALEN subunit cognates and the aligned sequence in the human genome are indicated by ‘●’.

**Table S3.** BLAST analysis of the *Mus musculus* genome to identify potential binding sites of combined and individual left and right subunits (CL and CR) of the C TALEN.

| | **Chrom. No.** | **Nucleotide Range** | **C TALEN Sites** | | **Identities** | **Sequence ID** |
|---|---|---|---|---|---|---|
| | | | **CL** | **CR** | | |
| | | | ***TATCAACACTTCCGGAGAC***<br>ATAGTTGTGAAGGCCTCTG | ***CGACGAGGCAGGTCCCCTA***<br>GCTGCTCCGTCCAGGGGAT | | |
| **Potential combined CL and CR sites** | 5 | 18010816 to 18010830 | ●●TCAACACTTCCGGAG●● | | 15/19 | NW_001030796.1 |
| | | 19827052 to 19827066 | ●●TCAACACTTCCGGAG●● | | 15/19 | |
| | | 16218494 to 16218508 | | ●●●●GAGGCAGGTCCCCTA | 15/19 | |
| | 9 | 47261339 to 47261354 | TATCAA●ACTTCCGGA●●● | | 15/19 | NT_039474.8 |
| | | 7863713 to 7863727 | | ●●●●GAGGCAGGTCCCCTA | 15/19 | |
| | 17 | 38384398 to 38384414 | ●ATCAACACTTCC●GAGA● | | 16/19 | NT_039649.8 |
| | | 43566525 to 43566540 | ●●●CAACACTTCCG●AGAC | | 15/19 | |
| | | 70416673 to 70416687 | | ●●●●GAGGCAGGTCCCCTA | 15/19 | |
| **Potential CL sites** | 3 | 38866300 to 38866314 | ●ATCAACACTTCCGGA●●● | | 15/19 | NT_039240.8 |
| | | 53841380 to 53841397 | TATCA●CACTTCCGGAGA● | | 17/19 | |
| | | 34012433 to 34012449 | ●●TCAA●ACTTCCGGAGAC | | 16/19 | |
| | | 59437720 to 59437735 | TATCA●CACTTCCGGA●●● | | 15/19 | |
| | 5 | 531836 to 531850 | ●●TCAACACTTCCGGAG●● | | 15/19 | NT_078458.7 |
| | | 2348932 to 2348946 | ●●TCAACACTTCCGGAG●● | | 15/19 | |
| | 14 | 6519500 to 6519514 | ●ATCAACACTTCCGGA●●● | | 15/19 | NT_039606.8 |
| | | 6659119 to 6659133 | ●ATCAACACTTCCGGA●●● | | 15/19 | |
| | | 6798768 to 6798782 | ●ATCAACACTTCCGGA●●● | | 15/19 | |
| | | 26304835 to 26304850 | ●●●CAACACTTC●GGAGAC | | 15/19 | |
| | 2 | 93924002 to 93924019 | ●ATCAACACTTCC●GAGAC | | 17/19 | NT_039207.8 |
| | | 9619835 to 9619851 | TATCAACACTTC●GGAG●● | | 16/19 | |
| | | 87013076 to 87013092 | ●●TCAACAC●TCCGGAGAC | | 16/19 | |
| **Potential CR sites** | Sequences in the *Mus musculus* genome with 15 or more nucleotide identities to the CR cognate were not found. | | | | | |

#A minimum sequence identity between TALEN subunit cognates and sites within the *Mus musculus* genome of 15 out of 19 sense or antisense nucleotides was used as a cut off.

*Mismatches between the TALEN subunit cognates and the aligned sequence in the *Mus musculus* genome are indicated by '●'.

**Table S4.** BLAST analysis of the human genome to identify potential binding sites of combined and individual left and right subunits (CL and CR) of the C TALEN.

| | Chrom. No. | Nucleotide Range | C TALEN Sites | | Identities | Sequence ID |
|---|---|---|---|---|---|---|
| | | | CL | CR | | |
| | | | ***TATCAACACTTCCGGAGAC***<br>ATAGTTGTGAAGGCCTCTG | ***CGACGAGGCAGGTCCCCTA***<br>GCTGCTCCGTCCAGGGGAT | | |
| **Potential combined CL and CR sites** | 3 | 51696616 to 51696632 | TATCAACACTTC●GGAG●● | | 16/17 | NT_022517.18 |
| | | 65720310 to 65720324 | | ●●●●GAGGCAGGTCCCCTA | 15/19 | |
| | 2 | 56850349 to 56850365 | ●●TCAACACTTC●GGAGAC | | 16/19 | NW_004078008.1 |
| | | 15604047 to 15604063 | | ●GACGAGGCAGG●CCCCT● | 16/19 | |
| | 21 | 4560786 to 4560802 | ●●TCAACACTTC●GGAGAC | | 16/19 | NW_004078109.1 |
| | | 30289322 to 30289338 | | CGAC●AGGCAGGTCCCC●● | 16/19 | |
| **Potential CL sites** | 8 | 4869274 to 4869289 | ●●●CAACACTTCCGGAGAC | | 16/19 | NT_167187.1 |
| | | 12795942 to 12795958 | ●●TCAACACTTC●GGAGAC | | 16/19 | |
| | 7 | 30035998 to 30036012 | ●●TCAACACTTCCGGAG●● | | 15/19 | NT_007819.17 |
| | | 36477879 to 36477895 | TATCAACACTTC●GGAG●● | | 16/19 | |
| | | 42858728 to 42858744 | ●●TCAACACTTC●GGAGAC | | 16/19 | |
| | 6 | 5016435 to 5016449 | ●●TCAACACTTCCGGAG●● | | 15/19 | NT_007592.15 |
| | | 31733101 to 31733119 | TATCAACACTTCC●GAGAC | | 18/19 | |
| | 3 | 30147609 to 30147625 | TATCAACA●TTCCGGAG●● | | 16/19 | NT_005612.16 |
| | | 55554086 to 55554102 | ●●TCAACACTTC●GGAGAC | | 16/19 | |
| **Potential CR sites** | 2 | 553285 to 553301 | | ●GACGAGGCAGG●CCCCT● | 16/17 | NT_022135.16 |
| | | 830629 to 830645 | | ●GACGAGGCAGG●CCCCT● | 16/17 | |

#A minimum sequence identity between TALEN subunit cognates and sites within the human genome of 15 out of 19 sense or antisense nucleotides was used as a cut off.

*Mismatches between the TALEN subunit cognates and the aligned sequence in the human genome are indicated by '●'.