Supplementary information for:

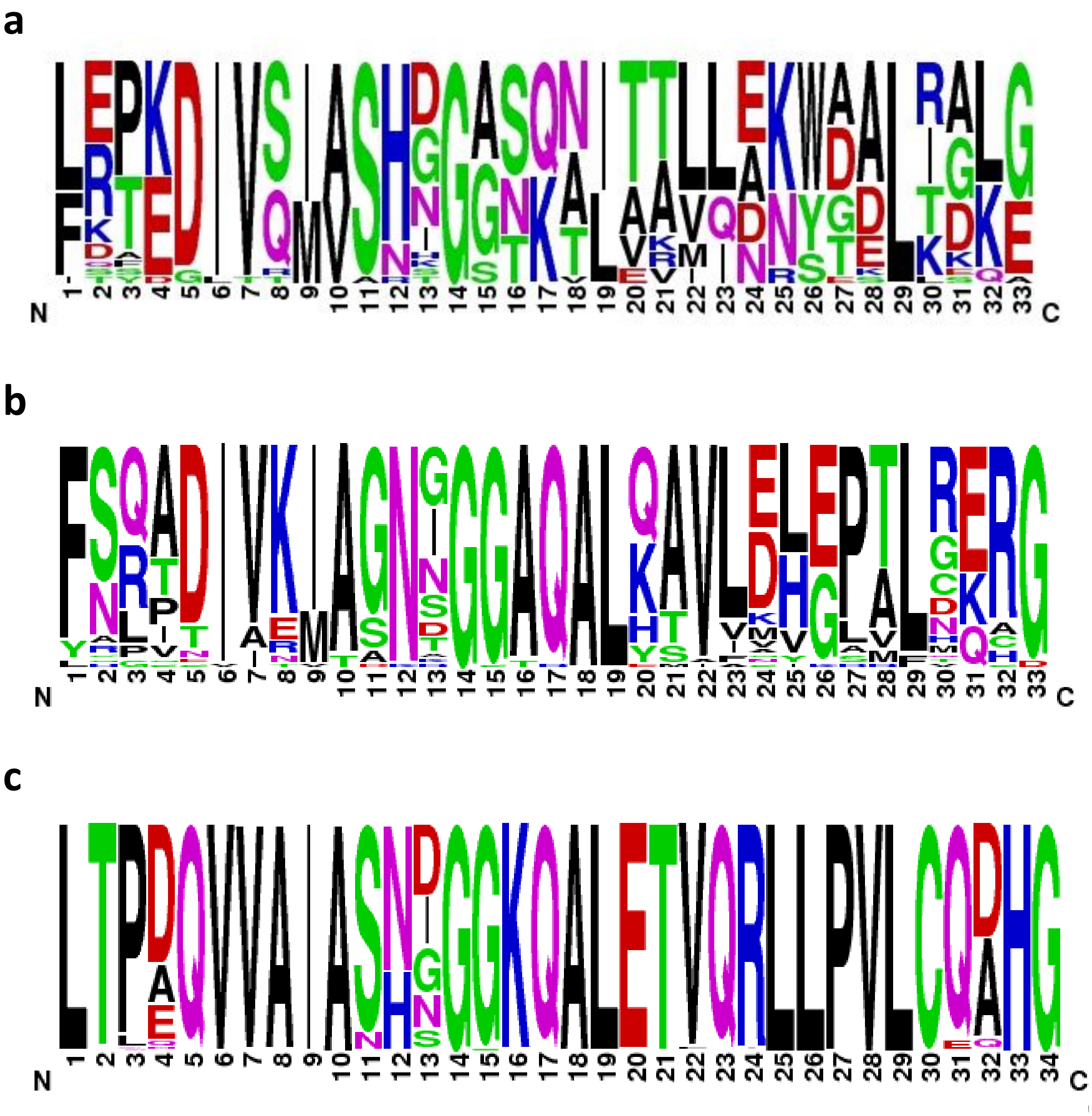
BurrH: a new modular DNA binding protein for genome engineering

Alexandre Juillerat, Claudia Bertonati, Gwendoline Dubois, Valérie Guyot, Séverine Thomas, Julien Valton, Marine Beurdeley, George H. Silva, Fayza Daboussi and Philippe Duchateau

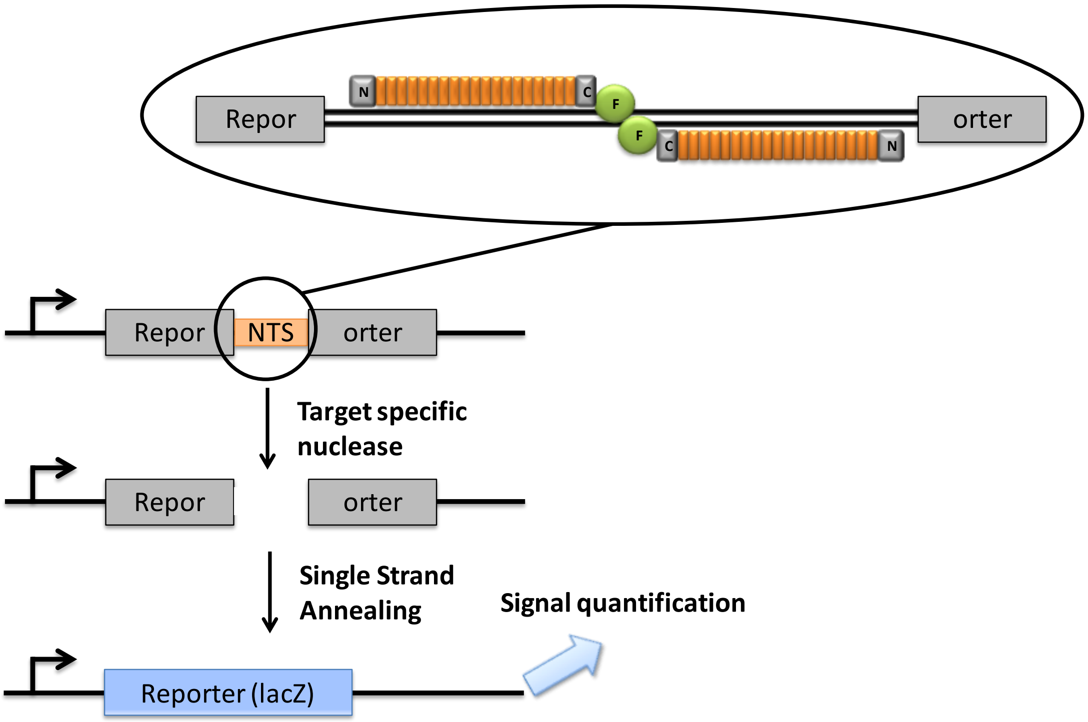
CELLECTIS S.A., 8 Rue de la Croix Jarry, 75013 Paris, France.

Correspondence to: Philippe.Duchateau (philippe.duchateau@cellectis.com).

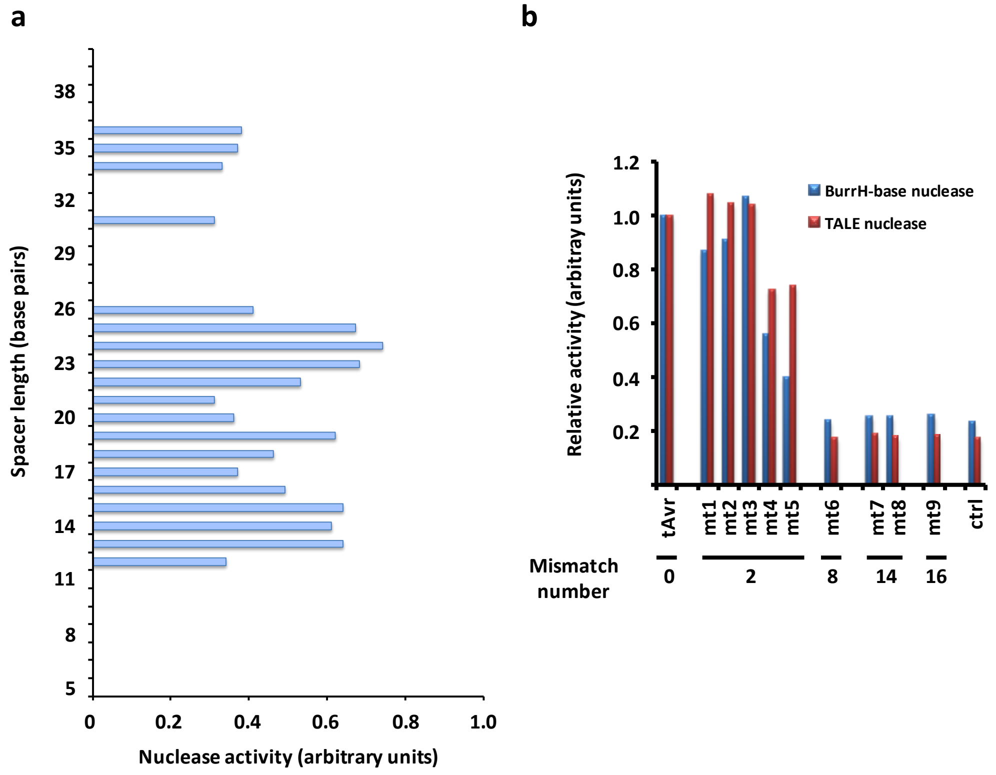
Supplementary Figure 1. Juillerat *et al.*



Supplementary Figure 2. Juillerat *et al.*



Supplementary Figure 3. Juillerat *et al.*



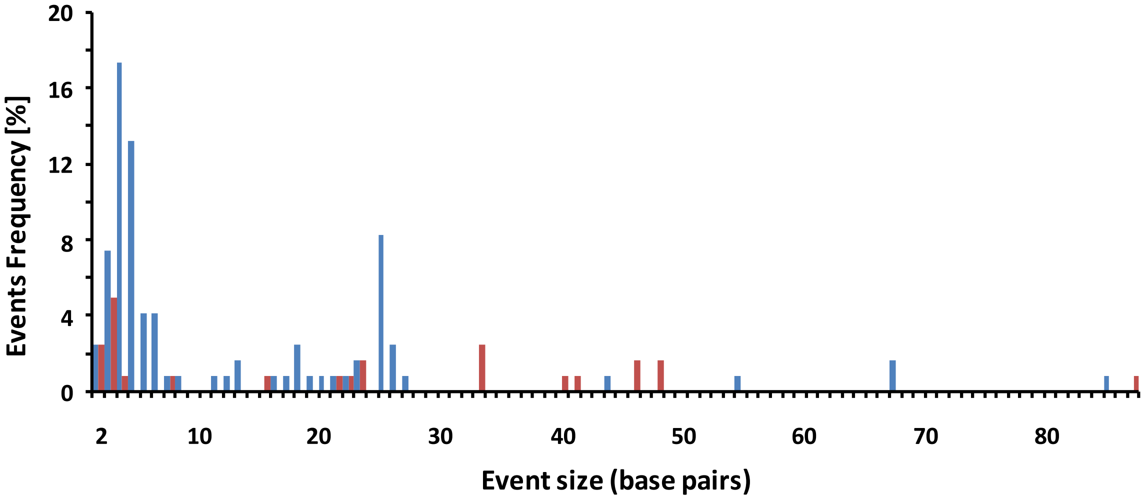
Supplementary Figure 4. Juillerat *et al.*

a

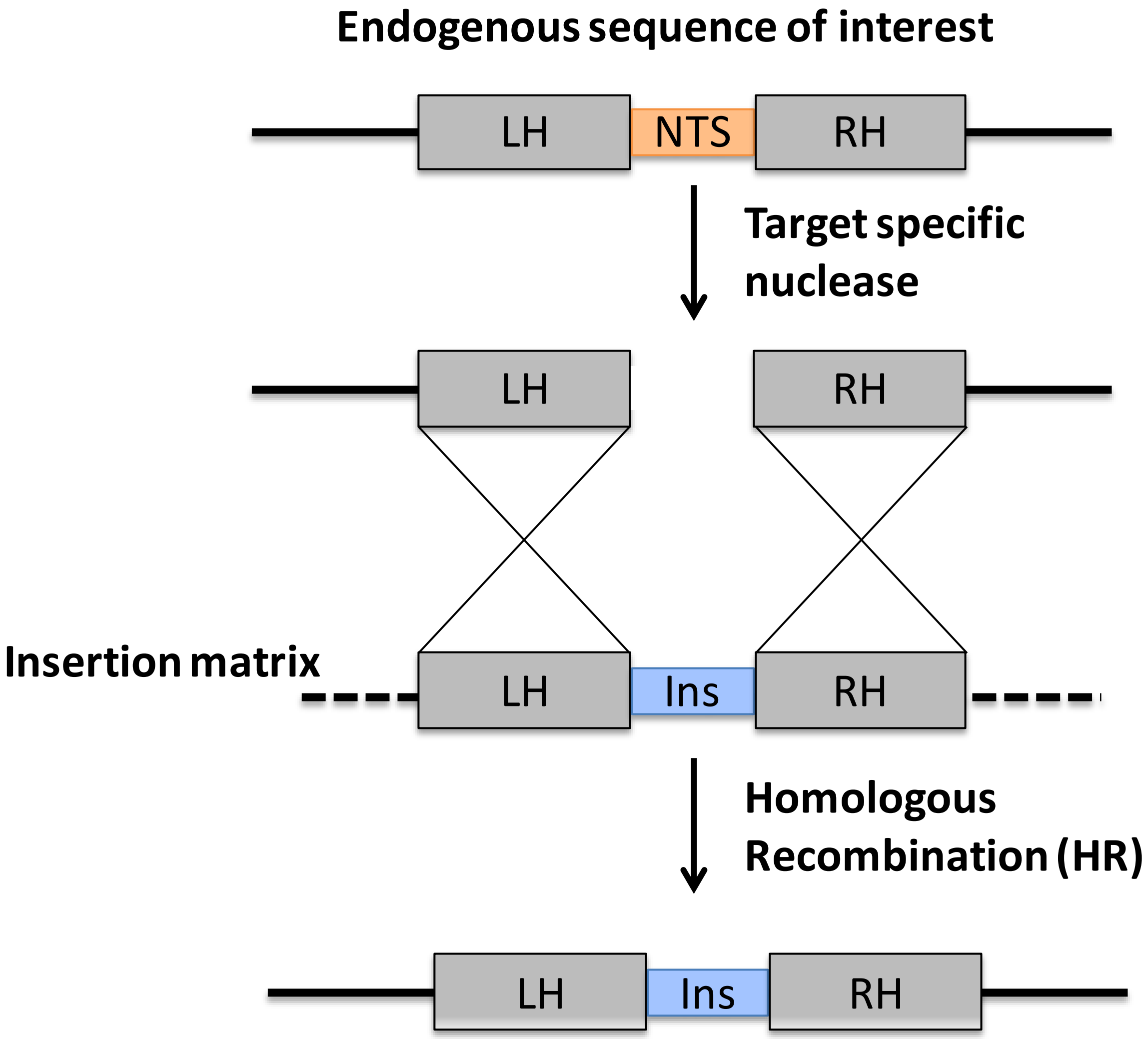
WT: aaccccattgtccgggaacccagagct**cacagccacgatctt**agacccgagcccacagagccagaggtg

Δ5: aaccccattgtccgggaacccagagct**caca-----gatctt**agacccgagcccacagagccagaggtg (x10)
Δ4a: aaccccattgtccgggaacccagagct**cacagc----atctt**agacccgagcccacagagccagaggtg (x9)
Δ4b: aaccccattgtccgggaacccagagct**cacagcc----tctt**agacccgagcccacagagccagaggtg (x9)
Δ26: aaccccattgtccgggaacccagagct**--------------------------**cacagagccagaggtg (x6)
Δ3: aaccccattgtccgggaacccagagct**cacag---cgatctt**agacccgagcccacagagccagaggtg (x6)

b



Supplementary Figure 5. Juillerat *et al.*



Supplementary figure Legends

Figure S1. (a) Sequence logo representation of the modules of the four fragments belonging to metagenomic sample. **(b)** Sequence logo representation of the modules of three proteins found in the proteome of *Burkholderia rhizoxinica*. **(c)** Sequence logo representation of the repeat sequences of the *Xanthomonas* TALE family. WEBLOGO software was used to create the logos.

Figure S2. A reporter plasmid containing a Nuclease Target Sequence (NTS) is flanked by overlapping truncated LacZ genes. Upon target cleavage in yeast, the restoration of the LacZ marker through the single strand annealing (SSA) pathway of recombination restores a functional LacZ gene, which can be quantified and related to the nuclease efficiency

Figure S3. (a) A collection of 36 targets comprising two AvrBs3 sequences facing each other on both DNA strand with spacer length (between the two targets) ranging from 5 to 40 bp were designed and assayed in the yeast SSA assay using the BurrH-based nuclease bn18473 to determine the optimal cleavage conditions. **(b)** To evaluate the impact of mismatches on the nuclease activity, a collection of 9 targets containing 2, 8, 14 or 16 mutations with respect to the AvrBs3 target (tAvr) were assayed in the yeast SSA assay using the BurrH-based nuclease bn18473 and an engineered AvrBs3 TALEN. Activity on the reference target (0 mutations) was normalized to 1 for each of the two nucleases. The control is a point without nuclease. N≥6.

Figure S4. (**a**) Alignment of the WT genomic sequence and the most predominant events (deletions are indicated by dashes) induced by the BurrH nuclease (18 modules) at the CAPNS1 locus. Binding sites of the two half BurrH-based nuclease are underlined. (**b**) Size repartition of the deletions (blue bars) and insertions (red bars) events induced at the CAPNS1 locus by the BurrH nuclease composed of 18 modules.

Figure S5. Schematic of the *in vivo* targeted gene insertion strategy. The donor DNA consists of two homologous regions (LH, left homology; RH, right homology) flanking a central insert sequence. Double strand breaks resulting from the nuclease activity at the NTS can be repaired by homologous recombination with the donor DNA, resulting in specific insertion events.

Supplementary table 1. Juillerat *et al.*

Amino acid sequences of N-terminal (including a NLS and a HA tag) domain, central DNA binding core and C-terminal (including FokI) domains of the WT BurrH-based nuclease (bn17421) with the targeted sequence (upper case, spacer in lower case)

Nter Sequence	MGDPKKKRKVIDYPYDVPDYAIDIASTAFVDQDKQMANRLNLSPLERSKIEKQYGGATTLAFISNKQNELAQILSRADILKIASYDCAAHALQAVLDCGPMLGKRG
DNA binding core	FSQSDIVKIAGNIGGAQALQAVLDLESMLGKRGFSRDDIAKMAGNIGGAQTLQAVLDLESAFRERGFSQADIVKIAGNNGGAQALYSVLDVEPTLGKRGFSRADIVKIAGNTGGAQALHTVLDLEPALGKRGFSRIDIVKIAANNGGAQALHAVLDLGPTLRECGFSQATIAKIAGNIGGAQALQMVLDLGPALGKRGFSQATIAKIAGNIGGAQALQTVLDLEPALCERGFSQATIAKMAGNNGGAQALQTVLDLEPALRKRDFRQADIIKIAGNDGGAQALQAVIEHGPTLRQHGFNLADIVKMAGNIGGAQALQAVLDLKPVLDEHGFSQPDIVKMAGNIGGAQALQAVLSLGPALRERGFSQPDIVKIAGNTGGAQALQAVLDLELTLVEHGFSQPDIVRITGNRGGAQALQAVLALELTLRERGFSQPDIVKIAGNSGGAQALQAVLDLELTFRERGFSQADIVKIAGNDGGTQALHAVLDLERMLGERGFSRADIVNVAGNNGGAQALKAVLEHEATLNERGFSRADIVKIAGNGGGAQALKAVLEHEATLDERGFSRADIVRIAGNGGGAQALKAVLEHGPTLNERGFNLTDIVEMAANSGGAQALKAVLEHGPTLRQRGLSLIDIVEIASNGGAQALKAVLKYGPVLMQAG
Cter Sequence	RSNEEIVHVAARRGGAGRIRKMVAPLLERQGRSGSDPISRSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFAAD
Target sequence	AAGAGAAGCAAAGACGTTACtagcatgaaggtaccGTAACGTCTTTGCTTCTCTT

Supplementary table 2. Juillerat *et al.*

Amino acid sequences of N-terminal (including a NLS and a HA tag) domain, central DNA binding core and C-terminal (including FokI) domains of the BurrH-based nuclease (bn18473) targeting AvrBs3 sequence.

Nter Sequence	MGDPKKKRKVIDYPYDVPDYAIDIASTAFVDQDKQMANRLNLSPLERSKIEKQYGGATTLAFISNKQNELAQILSRADILKIASYDCAAHALQAVLDCGPMLGKRG
DNA binding core	FSQSDIVKIAGHDGGAQALQAVLDLESMLGKRGFSRDDIAKMAGNGGGAQTLQAVLDLESAFRERGFSQADIVKIAGNIGGAQALYSVLDVEPTLGKRGFSRADIVKIAGNGGGAQALHTVLDLEPALGKRGFSRIDIVKIAANIGGAQALHAVLDLGPTLRECGFSQATIAKIAGNIGGAQALQMVLDLGPALGKRGFSQATIAKIAGNIGGAQALQTVLDLEPALCERGFSQATIAKMAGHDGGAQALQTVLDLEPALRKRDFRQADIIKIAGHDGGAQALQAVIEHGPTLRQHGFNLADIVKMAGNGGGAQALQAVLDLKPVLDEHGFSQPDIVKMAGNIGGAQALQAVLSLGPALRERGFSQPDIVKIAGNIGGAQALQAVLDLELTLVEHGFSQADIVKIAGHDGGTQALHAVLDLERMLGERGFSRADIVNVAGHDGGAQALKAVLEHEATLNERGFSRADIVKIAGHDGGAQALKAVLEHEATLDERGFSRADIVNVAGNGGGAQALKAVLEHEATLNERGFNLTDIVEMAAHDGGAQALKAVLEHGPTLRQRGLSLIDIVEIAGNGGGAQALKAVLKYGPVLMQAG
Cter Sequence	RSNEEIVHVAARRGGAGRIRKMVAPLLERQGRSGSDPISRSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFAAD

Supplementary table 3. Juillerat *et al.*

Amino acid sequences of N-terminal (including a NLS and a HA tag) domain, central DNA binding core and C-terminal (including FokI) domains of the control TALEN targeting AvrBs3 sequence.

Nter Sequence	MADPIRSRTPSPARELLPGPQPDGVQPTADRGVSPPAGGPLDGLPARRTMSRTRLPSPPAPSPAFSAGSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
DNA binding core	LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGRPALE
Cter Sequence	SIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARSGTLPPASQRWDRILQASGMKRAKPSPTSTQTPDQASLHAFADSLERDLDAPSPMHEGDQTRASSRKRSRSDRAVTGPSAQQSFEVRVPEQRDALHLPLSWRVKRPRTSIGGGLPDPISRSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFAAD

Amino acid sequences of N-terminal (including a NLS and a HA tag) domain, central DNA binding core and C-terminal (including FokI) domains of the engineered TALEN targeting AvrBs3 sequence.

Nter Sequence	MADPIRSRTPSPARELLPGPQPDGVQPTADRGVSPPAGGPLDGLPARRTMSRTRLPSPPAPSPAFSAGSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
DNA binding core	LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGRPALE
Cter Sequence	SIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLGDPISRSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFAAD

Supplementary table 4. Juillerat *et al.*

AvrBs3 target sequences (upper case) with spacer (lower case) length ranging from 5 to 40 base pairs.

Spacer size	Target sequence
5	TATATAAACCTAACCCTCTaggtaAGAGGGTTAGGTTTATATA
6	TATATAAACCTAACCCTCTaaggtaAGAGGGTTAGGTTTATATA
7	TATATAAACCTAACCCTCTaaggtacAGAGGGTTAGGTTTATATA
8	TATATAAACCTAACCCTCTgaaggtacAGAGGGTTAGGTTTATATA
9	TATATAAACCTAACCCTCTgaaggtaccAGAGGGTTAGGTTTATATA
10	TATATAAACCTAACCCTCTtgaaggtaccAGAGGGTTAGGTTTATATA
11	TATATAAACCTAACCCTCTtgaaggtacctAGAGGGTTAGGTTTATATA
12	TATATAAACCTAACCCTCTatgaaggtacctAGAGGGTTAGGTTTATATA
13	TATATAAACCTAACCCTCTatgaaggtaccttAGAGGGTTAGGTTTATATA
14	TATATAAACCTAACCCTCTcatgaaggtaccttAGAGGGTTAGGTTTATATA
15	TATATAAACCTAACCCTCTtagcatgaaggtaccAGAGGGTTAGGTTTATATA
16	TATATAAACCTAACCCTCTgcatgaaggtaccttgAGAGGGTTAGGTTTATATA
17	TATATAAACCTAACCCTCTgcatgaaggtaccttgtAGAGGGTTAGGTTTATATA
18	TATATAAACCTAACCCTCTagcatgaaggtaccttgtAGAGGGTTAGGTTTATATA
19	TATATAAACCTAACCCTCTagcatgaaggtaccttgtcAGAGGGTTAGGTTTATATA
20	TATATAAACCTAACCCTCTtagcatgaaggtaccttgtcAGAGGGTTAGGTTTATATA
21	TATATAAACCTAACCCTCTtagcatgaaggtaccttgtcgAGAGGGTTAGGTTTATATA
22	TATATAAACCTAACCCTCTtagcatgaaggtaccttgtcgtAGAGGGTTAGGTTTATATA
23	TATATAAACCTAACCCTCTctagcatgaaggtaccttgtcgtAGAGGGTTAGGTTTATATA
24	TATATAAACCTAACCCTCTctagcatgaaggtaccttgtcgttAGAGGGTTAGGTTTATATA
25	TATATAAACCTAACCCTCTactagcatgaaggtaccttgtcgttAGAGGGTTAGGTTTATATA
26	TATATAAACCTAACCCTCTactagcatgaaggtaccttgtcgttgAGAGGGTTAGGTTTATATA
27	TATATAAACCTAACCCTCTcactagcatgaaggtaccttgtcgttgAGAGGGTTAGGTTTATATA
28	TATATAAACCTAACCCTCTcactagcatgaaggtaccttgtcgttgaAGAGGGTTAGGTTTATATA
29	TATATAAACCTAACCCTCTccactagcatgaaggtaccttgtcgttgaAGAGGGTTAGGTTTATATA
30	TATATAAACCTAACCCTCTccactagcatgaaggtaccttgtcgttgatAGAGGGTTAGGTTTATATA
31	TATATAAACCTAACCCTCTaccactagcatgaaggtaccttgtcgttgatAGAGGGTTAGGTTTATATA
32	TATATAAACCTAACCCTCTaccactagcatgaaggtaccttgtcgttgattAGAGGGTTAGGTTTATATA
33	TATATAAACCTAACCCTCTgaccactagcatgaaggtaccttgtcgttgattAGAGGGTTAGGTTTATATA
34	TATATAAACCTAACCCTCTgaccactagcatgaaggtaccttgtcgttgattcAGAGGGTTAGGTTTATATA
35	TATATAAACCTAACCCTCTtgaccactagcatgaaggtaccttgtcgttgattcAGAGGGTTAGGTTTATATA
36	TATATAAACCTAACCCTCTtgaccactagcatgaaggtaccttgtcgttgattcaAGAGGGTTAGGTTTATATA
37	TATATAAACCTAACCCTCTctgaccactagcatgaaggtaccttgtcgttgattcaAGAGGGTTAGGTTTATATA
38	TATATAAACCTAACCCTCTctgaccactagcatgaaggtaccttgtcgttgattcagAGAGGGTTAGGTTTATATA
39	TATATAAACCTAACCCTCTtctgaccactagcatgaaggtaccttgtcgttgattcagAGAGGGTTAGGTTTATATA
40	TATATAAACCTAACCCTCTtctgaccactagcatgaaggtaccttgtcgttgattcagtAGAGGGTTAGGTTTATATA

Supplementary table 5. Juillerat *et al.*

Sequences of the targets that contained mismatches. Binding sites are represented in upper case and spacer in lower case. Mismatches are highlighted in grey boxes

name	Target sequence
tAvr	ATATAAACCTAACCCTCTtagcatgaaggtaccAGAGGGTTAGGTTTATAT
mt1	ATATAAACCTAACCCTC**A**tagcatgaaggtacc**T**GAGGGTTAGGTTTATAT
mt2	ATATAAACCTAACCCTC**C**tagcatgaaggtacc**G**GAGGGTTAGGTTTATAT
mt3	ATATAAACCTAACCCTC**G**tagcatgaaggtacc**C**GAGGGTTAGGTTTATAT
mt4	ATATAAACCT**G**ACCCT**T**Ttagcatgaaggtacc AGAGGGTTAGGTTTATAT
mt5	ATATAAACCTAACC**A**TC**C**tagcatgaaggtaccAGAGGGTTAGGTTTATAT
mt6	**C**T**C**TA**TCT**C**A**A**C**CCCT**T**Ttagcatgaaggtacc AGAGGGTTAGGTTTATAT
mt7	**T**TATA**G**A**G**CTAA**AGT**T**A**TcatcagttaacattataaatgA**T**A**ACT**TTAG**C**T**C**TATA**C**
mt8	**T**T**T**TA**C**ACCT**GCAG**CTCTcattttccatacagtcagtAGAG**CTGC**AGGT**G**TA**C**A**A**
mt9	**C**T**C**TA**TCT**C**A**A**C**CCCT**T**TtagcatgaaggtaccttgtcgA**A**AGGG**G**T**T**G**AGA**TA**G**A**G**

Supplementary table 6. Juillerat *et al.*

Amino acid sequences of N-terminal (including a NLS and a HA tag) domain, central DNA binding core and C-terminal (including FokI) domains of the 20 modules BurrH-based nuclease targeting CAPNS1 (left arm) with the targeted sequence.

Nter Sequence	MGDPKKKRKVIDYPYDVPDYAIDIASTAFVDQDKQMANRLNLSPLERSKIEKQYGGATTLAFISNKQNELAQILSRADILKIASYDCAAHALQAVLDCGPMLGKRG
DNA binding core	FSQSDIVKIAGNGGGAQALQAVLDLESMLGKRGFSRDDIAKMAGNGGGAQTLQAVLDLESAFRERGFSQADIVKIAGNNGGAQALYSVLDVEPTLGKRGFSRADIVKIAGNGGGAQALHTVLDLEPALGKRGFSRIDIVKIAAHDGGAQALHAVLDLGPTLRECGFSQATIAKIAGHDGGAQALQMVLDLGPALGKRGFSQATIAKIAGNNGGAQALQTVLDLEPALCERGFSQATIAKMAGNNGGAQALQTVLDLEPALRKRDFRQADIIKIAGNNGGAQALQAVIEHGPTLRQHGFNLADIVKMAGNIGGAQALQAVLDLKPVLDEHGFSQPDIVKMAGNIGGAQALQAVLSLGPALRERGFSQPDIVKIAGHDGGAQALQAVLDLELTLVEHGFSQPDIVRITGHDGGAQALQAVLALELTLRERGFSQPDIVKIAGHDGGAQALQAVLDLELTFRERGFSQADIVKIAGNIGGTQALHAVLDLERMLGERGFSRADIVNVAGNNGGAQALKAVLEHEATLNERGFSRADIVKIAGNIGGAQALKAVLEHEATLDERGFSRADIVRIAGNNGGAQALKAVLEHGPTLNERGFNLTDIVEMAAHDGGAQALKAVLEHGPTLRQRGLSLIDIVEIASNGGGAQALKAVLKYGPVLMQAG
Cter Sequence	RSNEEIVHVAARRGGAGRIRKMVAPLLERQGRSGSDPISRSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFAAD
Target sequence	TTGTCCGGGAACCCAGAGCT

Amino acid sequences of N-terminal (including a NLS and a S tag) domain, central DNA binding core and C-terminal (including FokI) domains of the 20 modules BurrH-based nuclease targeting CAPNS1 (right arm) with the targeted sequence.

Nter Sequence	MGDPKKKRKVIDKETAAAKFERQHMDSIDIASTAFVDQDKQMANRLNLSPLERSKIEKQYGGATTLAFISNKQNELAQILSRADILKIASYDCAAHALQAVLDCGPMLGKRG
DNA binding core	FSQSDIVKIAGNNGGAQALQAVLDLESMLGKRGFSRDDIAKMAGNNGGAQTLQAVLDLESAFRERGFSQADIVKIAGHDGGAQALYSVLDVEPTLGKRGFSRADIVKIAGNGGGAQALHTVLDLEPALGKRGFSRIDIVKIAAHDGGAQALHAVLDLGPTLRECGFSQATIAKIAGNGGGAQALQMVLDLGPALGKRGFSQATIAKIAGNNGGAQALQTVLDLEPALCERGFSQATIAKMAGNGGGAQALQTVLDLEPALRKRDFRQADIIKIAGNNGGAQALQAVIEHGPTLRQHGFNLADIVKMAGNNGGAQALQAVLDLKPVLDEHGFSQPDIVKMAGNNGGAQALQAVLSLGPALRERGFSQPDIVKIAGHDGGAQALQAVLDLELTLVEHGFSQPDIVRITGNGGGAQALQAVLALELTLRERGFSQPDIVKIAGHDGGAQALQAVLDLELTFRERGFSQADIVKIAGNNGGTQALHAVLDLERMLGERGFSRADIVNVAGNNGGAQALKAVLEHEATLNERGFSRADIVKIAGNNGGAQALKAVLEHEATLDERGFSRADIVRIAGNGGGAQALKAVLEHGPTLNERGFNLTDIVEMAAHDGGAQALKAVLEHGPTLRQRGLSLIDIVEIASNGGGAQALKAVLKYGPVLMQAG
Cter Sequence	RSNEEIVHVAARRGGAGRIRKMVAPLLERQGRSGSDPISRSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFAAD
Target sequence	GGCTCTGTGGGCTCGGGTCT

Supplementary table 7. Juillerat *et al.*

Amino acid sequences of N-terminal (including a NLS and a HA tag) domain, central DNA binding core and C-terminal (including FokI) domains of the 18 modules BurrH-based nuclease targeting CAPNS1 (left arm) with the targeted sequence.

Nter Sequence	MGDPKKKRKVIDYPYDVPDYAIDIASTAFVDQDKQMANRLNLSPLERSKIEKQYGGATTLAFISNKQNELAQILSRADILKIASYDCAAHALQAVLDCGPMLGKRG
DNA binding core	FSQSDIVKIAGNNGGAQALQAVLDLESMLGKRGFSRDDIAKMAGNGGGAQTLQAVLDLESAFRERGFSQADIVKIAGHDGGAQALYSVLDVEPTLGKRGFSRADIVKIAGHDGGAQALHTVLDLEPALGKRGFSRIDIVKIAANNGGAQALHAVLDLGPTLRECGFSQATIAKIAGNNGGAQALQMVLDLGPALGKRGFSQATIAKIAGNNGGAQALQTVLDLEPALCERGFSQATIAKMAGNIGGAQALQTVLDLEPALRKRDFRQADIIKIAGNIGGAQALQAVIEHGPTLRQHGFNLADIVKMAGHDGGAQALQAVLDLKPVLDEHGFSQPDIVKMAGHDGGAQALQAVLSLGPALRERGFSQPDIVKIAGHDGGAQALQAVLDLELTLVEHGFSQADIVKIAGNIGGTQALHAVLDLERMLGERGFSRADIVNVAGNNGGAQALKAVLEHEATLNERGFSRADIVKIAGNIGGAQALKAVLEHEATLDERGFSRADIVNVAGNNGGAQALKAVLEHEATLNERGFNLTDIVEMAAHDGGAQALKAVLEHGPTLRQRGLSLIDIVEIAGNGGGAQALKAVLKYGPVLMQAG
Cter Sequence	RSNEEIVHVAARRGGAGRIRKMVAPLLERQGRSGSDPISRSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFAAD
Target sequence	GTCCGGGAACCCAGAGCT

Amino acid sequences of N-terminal (including a NLS and a S tag) domain, central DNA binding core and C-terminal (including FokI) domains of the 18 modules BurrH-based nuclease targeting CAPNS1 (right arm) with the targeted sequence.

Nter Sequence	MGDPKKKRKVIDKETAAAKFERQHMDSIDIASTAFVDQDKQMANRLNLSPLERSKIEKQYGGATTLAFISNKQNELAQILSRADILKIASYDCAAHALQAVLDCGPMLGKRG
DNA binding core	FSQSDIVKIAGHDGGAQALQAVLDLESMLGKRGFSRDDIAKMAGNGGGAQTLQAVLDLESAFRERGFSQADIVKIAGHDGGAQALYSVLDVEPTLGKRGFSRADIVKIAGNGGGAQALHTVLDLEPALGKRGFSRIDIVKIAANNGGAQALHAVLDLGPTLRECGFSQATIAKIAGNGGGAQALQMVLDLGPALGKRGFSQATIAKIAGNNGGAQALQTVLDLEPALCERGFSQATIAKMAGNNGGAQALQTVLDLEPALRKRDFRQADIIKIAGNNGGAQALQAVIEHGPTLRQHGFNLADIVKMAGHDGGAQALQAVLDLKPVLDEHGFSQPDIVKMAGNGGGAQALQAVLSLGPALRERGFSQPDIVKIAGHDGGAQALQAVLDLELTLVEHGFSQADIVKIAGNNGGTQALHAVLDLERMLGERGFSRADIVNVAGNNGGAQALKAVLEHEATLNERGFSRADIVKIAGNNGGAQALKAVLEHEATLDERGFSRADIVNVAGNGGGAQALKAVLEHEATLNERGFNLTDIVEMAAHDGGAQALKAVLEHGPTLRQRGLSLIDIVEIAGNGGGAQALKAVLKYGPVLMQAG
Cter Sequence	RSNEEIVHVAARRGGAGRIRKMVAPLLERQGRSGSDPISRSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFAAD
Target sequence	CTCTGTGGGCTCGGGTCT

Supplementary table 8. Juillerat *et al.*

Amino acid sequences of N-terminal (including a NLS and a HA tag) domain, central DNA binding core and C-terminal (including FokI) domains of the TALEN targeting CAPNS1 (left arm) with the targeted sequence. T0 is indicated in brackets.

Nter Sequence	MGDPKKKRKVIDYPYDVPDYAIDIADPIRSRTPSPARELLPGPQPDGVQPTADRGVSPPAGGPLDGLPARRTMSRTRLPSPPAPSPAFSAGSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
DNA binding core	LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQALLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGRPALE
Cter Sequence	SIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLGDPISRSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFAAD
Target sequence	(T)CCGGGAACCCAGAGCT

Amino acid sequences of N-terminal (including a NLS and a S tag) domain, central DNA binding core and C-terminal (including FokI) domains of the TALEN targeting CAPNS1 (right arm) with the targeted sequence.

Nter Sequence	MGDPKKKRKVIDKETAAAKFERQHMDSIDIADPIRSRTPSPARELLPGPQPDGVQPTADRGVSPPAGGPLDGLPARRTMSRTRLPSPPAPSPAFSAGSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
DNA binding core	LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPQQVVAIASNGGGRPALE
Cter Sequence	SIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLGDPISRSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVEENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFAAD
Target sequence	(T)CTGTGGGCTCGGGTCT

Supplementary table 9. Juillerat *et al.*

Oligonucleotides used to amplify the CAPNS1 locus in targeted mutagenesis experiments and to identify Knock-in positive clones in Targeted Gene Insertion experiments.

Targeted mutagenesis	**Forward -1**	CCATCTCATCCCTGCGTGTCTCCGACTCAGTACTGAGCTAcgaaccctcagacctgttcgatc
	Forward -2	CCATCTCATCCCTGCGTGTCTCCGACTCAGCATAGTAGTGcgaaccctcagacctgttcgatc
	Reverse	CCTATCCCCTGTGTGCCTTGGCAGTCTCAGgtgagatccagagcccagcctg
Tageted Gene Insertion	**R1**	AATTGCGGCCGCGGTCCGGCGC
	F1	AAAAAGGCCGGTAGCCCATACC
	R2	GCCGCCGCCGCCCTTCAAGAACGAGTTAACC
	F2	TTAAGGCGCGCCGGACCGCGGC