

Procollagen SBP-mGFP-COL1A1 in pLVXPuro (lentiviral backbone, CMV promoter, ampicillin^R)

Construct:

Procollagen-SBP: humanised synthetic signal and N-propeptide sequence of COL1a1 and streptavidin-binding-protein. Signalsequence + N-terminal procollagen protein sequence is identical with AAB94054.3; P02452.5; CAA98968.1

```
GAATTCGAGTTCATGTTCTCCTTTGTGGATCTGCGGCTCCTGCTGCTGCTTGCCGCAACTGCGCTGCT
CACACATGGCCAGGAAGAGGGACAGGTGGAAGGGCAGGATGAGGACATTCTCCCATCACTTGCGTCC
AGAAATGGGCTGAGGTACCACGATCGCGATGTCTGGAAACCCGAACCCTGTCGGATTTGCGTGTGCGAT
AATGGCAAGGTGCTGTGTGACGACGTTATCTGCGACGAAACCAAGAACTGTCCCGGAGCTGAGGTCCC
AGAGGGCGAATGCTGTCCGGTGTGTCCCGATGGCAGCGAGTCTCCACAGACCAGGAAACGACCGGAG
TAGAGGGCCCCGAAAGGGGACACAGGACCACGAGGACCCAGAGGTCCAGCCGGACCACCCGGTTCGTGAT
GGGATACCAGGGCAACCAGGGCTGCCTGGCCCTCCTGGCCACCTGGACCTCCTGGCCCGCCCGGCTT
GGGTGGCAACTTTGCAGCCGCTGCAGCGATGGACGAAAAGACCACCGGATGGAGAGGTGGGCACGTTG
TGGAAGGTCTTGCTGGCGAGTTGGAGCAGCTGCGCGCCAGACTCGAGCATCACCCACAAGGGCAGAG
GAGCCTATGGCAGCTGCCGCC
```

mGFP = humanised monomeric GFP (KP202880.1)

COL1a1 = mature COL1a1 with C-propeptide is identical with NP_000079.2 except for 1 mutation:

A5176T: = Thr -> Ser (corresponds to AS 1434 in NP_000079.2 or base 4426 in NM_000088.3)

Spacers:

1) CACCGGTCGCCACC 2) GCCGCTGCAGCG 3) ATGAAGACACCGGTCGCCACC 4) TGCTGCAGCTGCC

pLVXPuro: lentiviral vector backbone with ampicillin resistance and CMV promoter

Insert sequence including beginning and end of pLVXPuro:

Mutations:

C: T->C no amino acid change; Thr stays

T: A->T Thr -> Ser

```
[...] GACTCTACTAGAGGATCGCTAGCGCTACCGGACTCAGATCTCGACACCGGTCGCCACCATGTTTCAGCTTTGT
GGATCTGCGGCTCCTGCTGCTGCTTGCCGCAACTGCGCTGCTCACACATGGCCAGGAAGAGGGACAGGTGGAAGG
GCAGGATGAGGACATTCTCCCATCACTTGCGTCCAGAATGGGCTGAGGTACCACGATCGCGATGTCTGGAAACC
CGAACCCCTGTCGGATTTGCGTGTGCGATAATGGCAAGGTGCTGTGTGACGACGTTATCTGCGACGAAACCAAGAA
CTGTCCCGGAGCTGAGGTCCAGAGGGCGAATGCTGTCCGGTGTGTCCCGATGGCAGCGAGTCTCCACAGACCA
GGAAACGACCGGAGTAGAGGGCCCCGAAAGGGGACACAGGACCACGAGGACCCAGAGGTCCAGCCGGACCACCCGG
TCGTGATGGGATAACCAGGGCAACCAGGGCTGCCTGGCCCTCCTGGCCACCTGGACCTCCTGGCCCGCCCGGCTT
GGGTGGCAACTTTGCAGCCGCTGCAGCGATGGACGAAAAGACCACCGGATGGAGAGGTGGGCACGTTGTGGAAGG
TCTTGCTGGCGAGTTGGAGCAGCTGCGCGCCAGACTCGAGCATCACCCACAAGGGCAGAGAGAGCCCATGAAGAC
ACCGGTCGCCACCATGGTGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTTCGAGCTGGACGG
CGACGTAACCGGCCACAAGTTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAA
GTTTCATCTGCACCACCGGCAAGCTGCCCGTGCCTGGCCACCCCTCGTGACCACCTGACCTACGGCGTGCAGTG
CTTCAGCCGCTACCCCGACCATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCCCCAAGGCTACGTTCCAGGA
GCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTTCGAGGGCGACACCCCTGGT
GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACCTA
CAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACCTCAAGATCCGCCACAA
CATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCT
GCCCGACAACCACTACCTGAGCACCCAGTCCAAGCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTTCCT
GCTGGAGTTTCGTGACCGCCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTGTGTCAGCTGCCCGCAGCT
```

GTCTTATGGCTATGATGAGAAATCAACCGGAGGAATTTCCGTGCCTGGCCCCATGGGTCCCTCTGGTCCCTCGTGG
TCTCCCTGGCCCCCCTGGTGCACCTGGTCCCCAAGGCTTCCAAGGTCCCCCTGGTGAGCCTGGCGAGCCTGGAGC
TTCAGGTCCCATGGGTCCCCGAGGTCCCCAGGTCCCCCTGGAAAGAATGGAGATGATGGGGAAGCTGGAAAACC
TGGTCGTCCTGGTGAGCGTGGGCCCTCCTGGCCTCAGGGTGTCTGAGGATTGCCCGGAACAGCTGGCCTCCCTGG
AATGAAGGGACACAGAGGTTTCAGTGGTTTGGATGGTGCCAAGGGAGATGCTGGTCCCTGCTGGTCCCTAAGGGTGA
GCCTGGCAGCCCTGGTGAATAATGGAGCTCCTGGTCAGATGGGCCCCCTGGCCTGCCTGGTGAGAGAGGTGCCCC
TGGAGCCCCTGGCCTGCTGGTGTCTGTTGGTGTCTAAGGGTGAAGCTGGTCCCCAAGGGCCCCGAGGCTC
CCCCGCTGGTCCCTGGCTTCCCTGGTGTCTTGGTGTCTAAGGGTGAAGCTGGTCCCCAAGGGCCCCGAGGCTC
TGAAGGTCCCCAGGGTGTGCGTGGTGAGCCTGGCCCCCTGGCCTGCTGGTGTCTGGCCCTGCTGGAAACCC
TGGTGTGATGGACAGCCTGGTGTCTAAGGTGCCAATGGTGTCTCCTGGTATTGCTGGTGTCTCTGGCTTCCCTGG
TGCCCGAGGCCCTCTGGACCCAGGGCCCCGGCGCCCTCCTGGTCCCCAAGGGTAAACAGCGGTGAACCTGGTGTG
TCCTGGCAGCAAAGGAGACACTGGTGTCTAAGGGAGAGCCTGGCCCTGTTGGTGTCTCAAGGACCCCCCTGGCCCTGC
TGGAGAGGAAGGAAAGCGAGGAGCTCGAGGTGAACCCGGACCCACTGGCCTGCCCGGACCCCCCTGGCGAGCGTGG
TGGACCTGGTAGCCGTGGTTTCCCTGGCGCAGATGGTGTCTGGTCCCCAAGGGTCCCGCTGGTGAACGTGGTTT
TCCTGGCCCTGCTGGCCCCAAAGGATCTCCTGGTGAAGCTGGTGTCTCCCGGTGAAGCTGGTCTGCCTGGTGCCAA
GGGTCTGACTGGAAGCCCTGGCAGCCCTGGTCCCTGATGGCAAAACTGGCCCCCTGGTCCC GCCGGTCAAGATGG
TCGCCCCGGACCCCCAGGCCACCTGGTGGCCGTGGTGTGATGGGATTCCCTGGACCTAAAGGTGC
TGCTGGAGAGCCCGCAAGGCTGGAGAGCGAGGTGTTCCCGGACCCCCCTGGCGCTGTCTGGTCCCTGCTGGCAAAGA
TGGAGAGGCTGGAGCTCAGGGACCCCCCTGGCCCTGCTGGTCCCGCTGGCGAGAGAGGTGAACAAGGCCCTGCTGG
CTCCCCCGGATTCCAGGGTCTCCCTGGTCCCTGCTGGTCCCTCCAGGTGAAGCAGGCAAACCTGGTGAACAGGGTGT
TCCTGGAGACCTTGGCGCCCCTGGCCCCCTCTGGAGCAAGAGGCGAGAGAGGTTTCCCTGGCGAGCGTGGTGTGCA
AGGTCCCCCTGGTCCCTGCTGGTCCCCGAGGGGCCAACGGTGTCTCCCGCAACGATGGTGTCTAAGGGTGTGCTGG
TGCCCTGGAGCTCCCGGTAGCCAGGGCGCCCCCTGGCCTTCAGGGAATGCCTGGTGAACGTGGTGTGAGCTGGTCT
TCCAGGGCCTAAGGGTGTACAGAGGTGATGCTGGTCCCCAAGGTGTCTGATGGTCTCCTGGCAAAGATGGCGTCCG
TGGTCTGACCGGCCCATTTGGTCCCTGGCCCTGCTGGTGGCCCTGGTGTGACCAAGGTGAAAGTGGTCCCAGCGG
CCCTGCTGGTCCCCTGGAGCTCGTGGTGGCCCCGGAGACCGTGGTGTGAGCCTGGTCCCCCGGCCCTGCTGGCTT
TGCTGGCCCCCTGGTGTGACGGCCAACCTGGTGTCTAAGGGCGAACCTGGTGTGATGCTGGTGTCTAAGGGCGATGC
TGGTCCCCCTGGCCCTGCCGGACCCGCTGGACCCCTGGCCCCATTGGTAATGTTGGTGTCTCCTGGAGCCAAAGG
TGCTCGCGGCAGCGCTGGTCCCCCTGGTGTCTACTGGTTCCTGGTGTCTGGCCGAGTCCGTCCTCCTGGCCC
CTCTGGAATGCTGGACCCCTGGCCCTCCTGGTCCCTGCTGGCAAAGAAGGCGGCAAAGGTCCCCGTGGTGTGAGAC
TGGCCCTGCTGGACGTCCCTGGTGAAGTTGGTCCCCCTGGTCCCCCTGGCCCTGCTGGCGAGAAAGGATCCCCCTGG
TGCTGATGGTCCCTGCTGGTGTCTCCTGGTACTCCCGGGCCTCAAGGTATTGCTGGACAGCGTGGTGTGGTCCGGCT
GCCTGGTCAGAGAGGAGAGAGAGGCTTCCCTGGTCTTCCCTGGCCCTCTGGTGAACCTGGCAAACAAGGTCCCTC
TGGAGCAAAGTGGTGAACGTGGTCCCCCTGGTCCCATGGGCCCCCTGGATTGGCTGGACCCCTGGTGAATCTGG
ACGTGAGGGGGCTCCTGGTGGCGAAGGTTCCCTGGACGAGACGGTTCCTGGCGCCAAGGGTGACCGTGGTGA
GACCGGCCCCCGCTGGACCCCTGGTGTCTCCTGGTGTCTCCTGGTGGCCCTGGCCCCGTTGGCCCTGCTGGCAAAG
TGGTGTGATCGTGGTGTGAGACTGGTCCCTGCTGGTCCC GCCGGTCTGTGCGCCCTGTTGGCGCCCGTGGCCCCG
ACCCCAAGGCCCCCGTGGTGTGACAAGGGTGTGACAGGGCGAACAGGGCGACAGAGGCATAAAGGGTACCCTGGCTT
CTCTGGCCTCCAGGGTCCCCCTGGCCCTCCTGGTCTCCTGGTGAACAAGGTCCCTCTGGAGCCTCTGGTCCCTGC
TGGTCCCCGAGGTCCCCCTGGCTCTGCTGGTGTCTCCTGGCAAAGATGGACTCAACGGTCTCCCTGGCCCCATTGG
GCCCCCTGGTCCCTCGCGGTGCGACTGGTGTGCTGGTCCCTGTTGGTCCCCCGGCCCTCCTGGACCTCCTGGTCC
CCCTGGTCCCTCCAGCGCTGGTTTCGACTTCAGCTTCTGCCCCAGCCACCTCAAGAGAAGGCTCACGATGGTGG
CCGCTACTACCGGGCTGATGATGCCAATGTGGTTCGTGACCGTGACCTCGAGGTGGACACCACCCTCAAGAGCCT
GAGCCAGCAGATCGAGAACATCCGGAGCCCAGAGGGCAGCCGCAAGAACCCCGCCGCACCTGCCGTGACCTCAA
GATGTGCCACTCTGACTGGAAGAGTGGAGAGTACTGGATTGACCCCAACCAAGGCTGCAACCTGGATGCCATCAA
AGTCTTCTGCAACATGGAGACTGGTGTGAGACCTGCGTGTACCCCACTCAGCCCAGTGTGGCCAGAGAAGCTGGTA
CATCAGCAAGAACCCCAAGGACAAGAGGCATGTCTGGTTCGGCGAGAGCATGACCGATGGATTCCAGTTCGAGTA
TGGCGGCCAGGGCTCCGACCCTGCCGATGTGGCCATCCAGCTGACCTTCTGCGCCTGATGTCCACCGAGGCCCTC
CCAGAACATCACCTACCCTGCAAGAACAGCGTGGCCTACATGGACCAGCAGACTGGCAACCTCAAGAAGGCCCT
GCTCCTCCAGGGCTCCAACGAGATCGAGATCCGCGCCGAGGGCAACAGCCGCTTACCTACAGCGTCACTGTGCA
TGGCTGCACGAGTACACCCGGAGCCTGGGGCAAGACAGTGTGATGAATACAAAACCACCAAGTCCCTCCCGCCTGCC
CATCATCGATGTGGCCCCCTGGACGTTGGTGGCCCCAGACCAGGAATTCGGCTTCGACGTTGGCCCTGTCTGCTT
CCTGTAACTGCAGTCGACGGTACCGCGGGCCCCGGATCCCGGACTCTAGATAATTCTACCGGGTAGGGGAGGCC
CTT [...]

Complete insert sequence:

LOCUS ORIGIN	EMBOSS_001	5271 bp	DNA	linear	UNC 06-MAR-2017	
1	ATGTTTCAGCT	TTGTGGATCT	GCGGCTCCTG	CTGCTGCTTG	CCGCAACTGC	GCTGCTCACA
61	CATGGCCAGG	AAGAGGGACA	GGTGAAGGG	CAGGATGAGG	ACATTCCTCC	CATCACTTGC
121	GTCCAGAAATG	GGCTGAGGTA	CCACGATCGC	GATGTCTGGA	AACCCGAACC	CTGTCCGATT
181	TGCGTGTGCG	ATAATGGCAA	GGTGTCTGTG	GACGACGTTA	TCTGCGACGA	AACCAAGAAC
241	TGTCCCGGAG	CTGAGGTCCC	AGAGGGCGAA	TGCTGTCCGG	TGTGTCCCGA	TGGCAGCGAG
301	TCTCCCACAG	ACCAGGAAAC	GACCGGAGTA	GAGGGCCCGA	AAGGGGACAC	AGGACCACGA
361	GGACCCAGAG	GTCCAGCCGG	ACCACCCGGT	CGTGATGGGA	TACCAGGGCA	ACCAGGGCTG
421	CCTGGCCCTC	CTGGCCACC	TGGACCTCCT	GGCCCGCCCG	GCTTGGGTGG	CAACTTGTCA
481	BCCGCTGCAG	CGATGGACGA	AAAGACCACC	GGATGGAGAG	GTGGGCACGT	TGTGGAAGGT
541	CTTGCTGGCG	AGTTGGAGCA	GCTGCGCGCC	AGACTCGAGC	ATCACCACA	AGGGCAGAGA
601	GAGCCCATGA	AGACACCGGT	CGCCACCATG	GTGAGCAAGG	GCGAGGAGCT	GTTCCACGGG
661	GTGGTGCCCA	TCCTGGTCTGA	GCTGGACGGC	GACGTAACG	GCCACAAGTT	CAGCGTGTCC
721	GGCAGGGCG	AGGGCGATGC	CACCTACGGC	AAGCTGACCC	TGAAGTTCAT	CTGCACCACC
781	GGCAAGCTGC	CCGTGCCCTG	GCCCACCTC	GTGACCACCC	TGACCTACGG	CGTGCAGTGC
841	TTCAGCCGCT	ACCCCGACCA	CATGAAGCAG	CACGACTTCT	TCAAGTCCGC	CATGCCGAA
901	GGCTACGTCC	AGGAGCGCAC	CATCTTCTTC	AAGGACGACG	GCAACTACAA	GACCCGCGCC
961	GAGGTGAAGT	TCGAGGGCGA	CACCCTGGTG	AACCGCATCG	AGCTGAAGGG	CATCGACTTC
1021	AAGGAGGACG	GCAACATCCT	GGGGCACAAG	CTGGAGTACA	ACTACAACAG	CCACAACGTC
1081	TATATCATGG	CCGACAAGCA	GAAGAACGGC	ATCAAGGTGA	ACTTCAAGAT	CCGCCACAAC
1141	ATCGAGGACG	GACGCGTGCA	GCTCGCCGAC	CACTACCAGC	AGAACACCCC	CATCGGCGAC
1201	GGCCCCGTGC	TGCTGCCCGA	CAACCACTAC	CTGAGCACCC	AGTCCAAGCT	GAGCAAAGAC
1261	CCCAACGAGA	AGCGCGATCA	CATGGTCTCTG	CTGGAGTTCG	TGACCCGCCG	CGGGATCACT
1321	CTCGGCATGG	ACGAGCTGTA	CAAGTGCTGC	AGCTGCCCCC	AGCTGTCTTA	TGGCTATGAT
1381	GAGAAATCAA	CCGGAGGAAT	TTCCGTGCTT	GGCCCCATGG	GTCCCTCTGG	TCCTCGTGGT
1441	CTCCCTGGCC	CCCTTGGTGC	ACCTGGTCCC	CAAGGCTTCC	AAGGTCCCCC	TGGTGAGCCT
1501	GGCGAGCCTG	GAGCTTCAGG	TCCCATGGGT	CCCCGAGGTC	CCCCAGGTCC	CCCTGGAAAG
1561	AATGGAGATG	ATGGGGAAGC	TGGAACAACCT	GGTCGTCTCTG	GTGAGCGTGG	GCCTCCTGGG
1621	CCTCAGGGTG	CTCGAGGATT	CCCCGGAACA	GCTGGCCTCC	CTGGAATGAA	GGGACACAGA
1681	GGTTTCAGTG	GTTTGGATGG	TGCCAAGGGA	GATGCTGGTC	CTGCTGGTCC	TAAGGGTGAG
1741	CCTGGCAGCC	CTGGTGAAAA	TGGAGCTCCT	GGTCAGATGG	GCCCCCGTGG	CCTGCCTGGT
1801	GAGAGAGGTC	GCCCTGGAGC	CCCTGGCCCT	GCTGGTCTCT	GTGGAATGAA	TGGTGTACT
1861	GGTGTGCCG	GGCCCCCTGG	TCCCACCGGC	CCCCTGGTTC	CTCCTGGCTT	CCCTGGTGCT
1921	GTTGGTGTCTA	AGGGTGAAGC	TGGTCCCCAA	GGGCCCGGAG	GCTCTGAAGG	TCCCAGGGT
1981	GTGCGTGGTG	AGCCTGGCCC	CCCTGGCCCT	GCTGGTCTCTG	CTGGCCCTGC	TGAAAACCCCT
2041	GGTGTCTGATG	GACAGCCTGG	TGCTAAAGGT	GCCAATGGTG	CTCCTGGTAT	TGCTGGTGCT
2101	CCTGGCTTCC	CTGGTGCCCG	AGGCCCTCT	GGACCCAGG	GCCCCGGCGG	CCCTCCTGGT
2161	CCCAAGGTA	ACAGCGGTGA	ACCTGGTCT	CCTGGCAGCA	AAGGAGACAC	TGGTGTAAAG
2221	GGAGAGCCTG	GCCCTGTGG	TGTTCAAGGA	CCCCCTGGCC	CTGCTGGAGA	GGAAGGAAAG
2281	CGAGGAGCTC	GAGGTGAACC	CGGACCCACT	GGCTGCCCCG	GACCCCTGAG	CGAGCGTGGT
2341	GGACCTGGTA	GCCGTGGTGT	CCCTGGCGCA	GATGGTGTGT	CTGGTCCCAA	GGGTCCCCTG
2401	GGTGAACGTG	GTTCTCCTGG	CCCTGTCTGG	CCCAAAGGAT	CTCCTGGTGA	AGCTGTCTGT
2461	CCCGGTGAAG	CTGGTCTGCC	TGGTGCCAAG	GGTCTGACTG	GAAGCCCTGG	CAGCCCTGGT
2521	CCTGATGGCA	AAACTGGCCC	CCCTGGTCCC	GCCGGTCAAG	ATGGTCCGCC	CGGACCCCA
2581	GGCCACCTG	GTGCCCTGG	TCAGGTGGT	GTGATGGGAT	TCCCTGGACC	TAAAGGTGCT
2641	GCTGGAGAGC	CCGGCAAGGC	TGGAGAGCGA	GGTGTTCCTG	GACCCCTGG	CGCTGTCCGT
2701	CCTGCTGGCA	AAGATGGAGA	GGCTGGAGCT	CAGGGACCCC	CTGGCCCTGC	TGGTCCCCTG
2761	GGCGAGAGAG	GTGAACAAGG	CCCTGCTGGC	TCCCCCGGAT	TCCAGGGTCT	CCCTGGTCTCT
2821	GCTGGTCTCTC	CAGGTGAAGC	AGGCAAACCT	GGTGAACAGG	GTGTTCTCTG	AGACCTTGGC
2881	GCCCTGGCC	CCTCTGGAGC	AAGAGGCGAG	AGAGGTTTCC	CTGGCGAGCG	TGGTGTGCAA
2941	GGTCCCCTG	GTCCCTGCTGG	TCCCCGAGGG	GCCAACGGTG	CTCCCGCAA	CGATGGTCT
3001	AAGGGTGATG	CTGGTGCCCC	TGGAGCTCCC	GGTAGCCAGG	GCGCCCTGG	CCTTACGGGA
3061	ATGCCTGGTG	AACGTGGTGC	AGCTGGTCTT	CCAGGGCCTA	AGGGTGACAC	AGGTGATGCT
3121	GTGCCCAAAG	GTGCTGATGG	CTCTCTGGC	AAAGATGGCG	TCCGTGGTCT	GACCGGCCCC
3181	ATTGGTCTCTC	CTGGCCCTGC	TGGTCCCTCT	GGTGACAAGG	GTGAAAGTGG	TCCCAGCGGC
3241	CCTGCTGGTC	CCACTGGAGC	TCGTGGTGGC	CCCGGAGACC	GTGGTGAGCC	TGGTCCCCC
3301	GGCCCTGCTG	GCTTTGCTGG	CCCCCTGGT	GCTGACGGCC	AACCTGGTGC	TAAAGGCGAA
3361	CCTGGTGATG	CTGGTGCTAA	AGGCGATGCT	GGTCCCCTG	GCCCTGCGG	ACCCGTGGA
3421	CCCCCTGGCC	CCATTGGTAA	TGTTGGTGCT	CCTGGAGCCA	AAGGTGCTCG	CGGCAGCGCT
3481	GCTCCCCCTG	GTGCTACTGG	TTTCCCTGGT	GCTGCTGGCG	GAGTCGGTCC	TCCTGGCCCC
3541	TCTGGAATG	CTGGACCCCC	TGGCCCTCCT	GGTCCCTGCTG	GCAAAGAAGG	CGGCAAGGT
3601	CCCCGTGGTG	AGACTGGCCC	TGCTGGACGT	CCTGGTGAAG	TTGGTCCCCC	TGGTCCCCCT
3661	GGCCCTGCTG	GCGAGAAAGG	ATCCCCCTGGT	GCTGATGGTC	CTGCTGGTGC	TCCTGGTACT

3721 CCCGGGCTC AAGGTATGC TGGACAGCGT GGTGTGGTCG GCCTGCCTGG TCAGAGAGGA
3781 GAGAGAGGCT TCCCTGGTCT TCCTGGCCCC TCTGGTGAAC CTGGCAAACA AGGTCCCTCT
3841 GGAGCAAGTG GTGAACGTGG TCCCCCTGGT CCCATGGGCC CCCCTGGATT GGCTGGACCC
3901 CAGGTGAAT CTGGACGTGA GGGGGTCTCT GGTGCCGAAG GTTCCCCTGG ACGAGACGGT
3961 TCTCCTGGCG CCAAGGGTGA CCGTGGTGAG ACCGGCCCCG CTGGACCCOC TGGTGTCTCT
4021 GGTGCTCTCG GTGCCCCCTGG CCCCCTGGCC CCTGCTGGCA AGAGTGGTGA TCGTGGTGAG
4081 ACTGGTCTCG CTGGTCCC GCCTGCTGTC GGCCCTGTTG GCGCCCGTGG CCCC GCCGA
4141 CCCC AAGGCC CCCGTGGTGA CAAGGGTGAG ACAGGCGAAC AGGGCGACAG AGGCATAAAG
4201 GGTCACCGTG GCTTCTCTGG CCTCCAGGGT CCCCCTGGCC CTCCTGGCTC TCCTGGTGAA
4261 CAAGGTCCTT CTGGAGCCTC TGGTCTGCT GGTCCCCGAG GTCCCCCTGG CTCTGCTGGT
4321 GTCCTGGCA AAGATGGACT CAACGGTCTC CCTGGCCCCA TTGGGCCCCC TGGTCTCGC
4381 GGTCGCACTG GTGATGTGG TCCTGTTGGT CCCCCCGGCC CTCCTGGACC TCCTGGTCCC
4441 CCTGGTCTCT CCAGCGCTGG TTTCGACTTC AGCTTCTGTC CCCAGCCACC TCAAGAGAAG
4501 GCTCACGATG GTGGCCGCTA CTACCGGGCT GATGATGCCA ATGTGGTTCG TGACCGTGAC
4561 CTCGAGGTGG ACACCACCTT CAAGAGCCTG AGCCAGCAGA TCGAGAACAT CCGGAGCCCA
4621 GAGGGCAGCC GCAAGAACCC CGCCCGCACC TGCCGTGACC TCAAGATGTG CCACTCTGAC
4681 TGGAAGAGTG GAGAGTACTG GATTGACCC AACCAAGGCT GCAACCTGGA TGCCATCAA
4741 GTCTTCTGCA ACATGGAGAC TGGTGAGACC TGCGTGTACC CCACTCAGCC CAGTGTGGCC
4801 CAGAAGAACT GGTACATCAG CAAGAACCCC AAGGACAAGA GGCATGTCTG GTTCGGCGAG
4861 AGCATGACCG ATGGATTCCA GTTCGAGTAT GGCGGCCAGG GCTCCGACCC TGCCGATGTG
4921 GCCATCCAGC TGACCTTCTT GCGCCTGATG TCCACCGAGG CCTCCAGAA CATCACCTAC
4981 CACTGCAAGA ACAGCGTGGC CTACATGAC CAGCAGACTG GCAACCTCAA GAAGGCCCTG
5041 CTCCTCCAGG GCTCCAACGA GATCGAGATC CGCGCCGAGG GCAACAGCCG CTTACCTAC
5101 AGCGTCACTG TCGATGGCTG CACGAGTAC ACCGGAGCCT GGGGCAAGC AGTATTGAA
5161 TACAAAACCA CCAAGTCTCT CCGCCTGCCC ATCATCGATG TGGCCCCCTT GGACGTTGGT
5221 GCCCAGACC AGGAATTCCG CTTGACGTT GGCCCTGTCT GCTTCTGTGA A

Primers (5' -> 3') for procollagen-SBP-mGFP-COL1a1 complete sequencing:

CMV (from MWG)	fwd	CGC AAA TGG GCG GTA GGC GTG	1 – 630bp
(Procollagen1a1_m1	fwd	TCGTGATGGGATAACCAGGGC	472 – 630bp)
mGFP_m1	rev	GGTGTGCGCCCTCGAACT	607 – 925bp
mGFP_COL1a1	rev	CCATGGGGCCAGGCAC	879 – 1350bp
mGFP_m1	fwd	CACTACCTGAGCACCCAGTC	1272 – 1731bp
EGFP_C1 (from MWG)	fwd	GATCACTCTCGGCATGGAC	1361 – 2306bp
COL1a1_m1	fwd	GGCAGCAAAGGAGACTG	2248 – 2765bp
COL1a1_m4	rev	GACCTTTGCCGCTTCT	2662 – 3542bp
COL1a1_m3	rev	CACCACGATCACCCTCTTG	3508 – 4036bp
COL1a1_m2	rev	GGGGCAGGAAGCTGAAGT	3689 – 4435bp
COL1a1_m1	rev	GATGGCCACATCGGCAG	4389 – 4876bp
COL1a1-C_m1	fwd	TTCTGCAACATGGAGACTGGTG	4790 – 5175bp
COL1a1-C_m2	fwd	CAGCAGACTGGCAACCTC	5057 – 5271bp + vector

Primers for amplifying inserts:

Gene	Primer sequence 5' -> 3'	Type
Pro-COL-SBP*	GATCTCGAGCTCAAGCTTCGATGTTTCAGCTTTGTGGACCTC	Fwd
	CCATGCAGCAGCAGCACATGGGCTCTCTCTGGCC	Rev
mGFP	AGAGCCCATGTGCTGCTGCTGCTGCATGGTGAGCAAGGGCGAG	Fwd
	GGGGCAGCAGCAGCACTTGTACAGCTCGTCCATGC	Rev

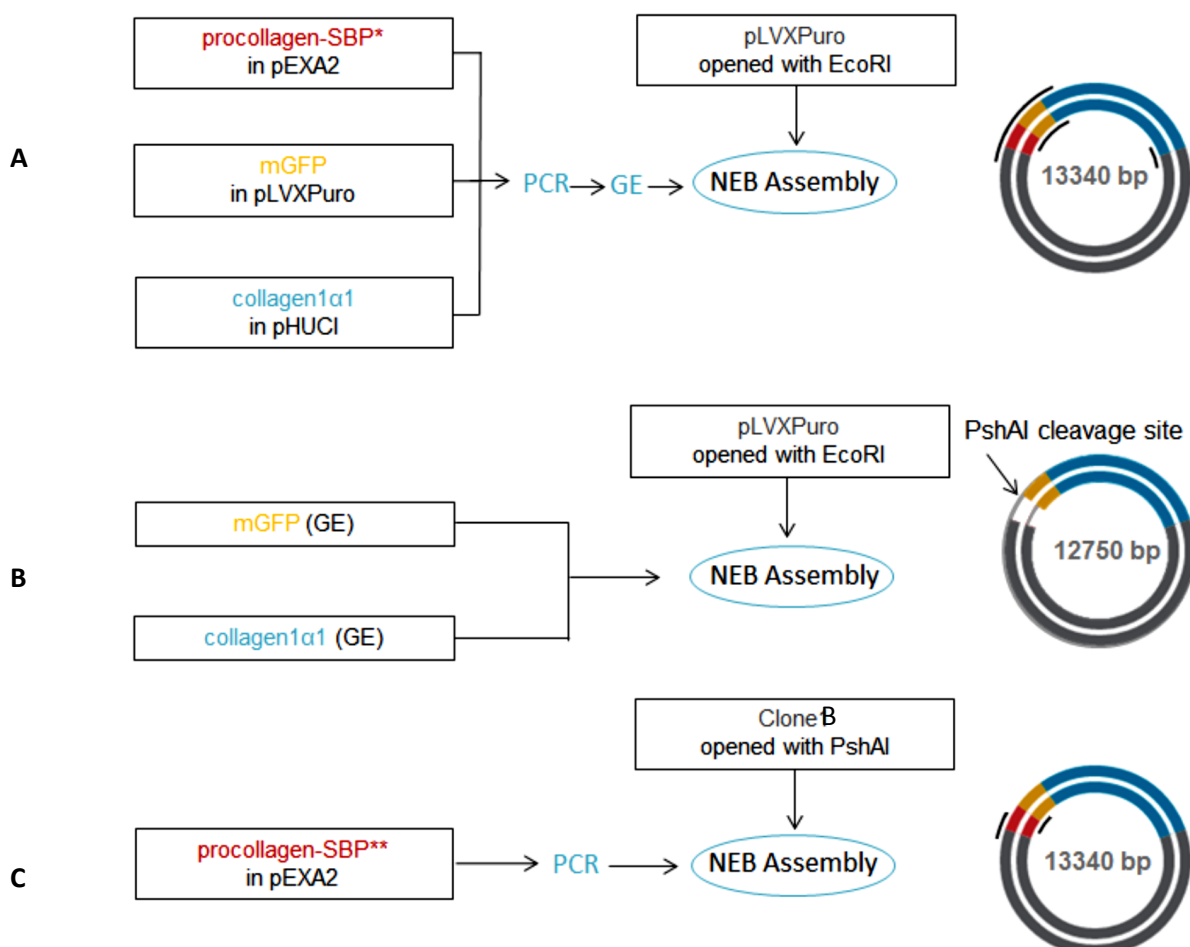
COL1a1	GCTGTACAAGTGCTGCTGCTGCCCCCAGCTGTCTTATGGC	Fwd
	CGCGGTACCGTCGACTGCAGTTACAGGAAGCAGACAGG	Rev
Pro-COL-SBP**	ACTCAGATCTCGACACCGGTCGCCACCATGTTTCAGCTTTG	Fwd
	CCATGGTGGCGACCGGTGTCTTCATGGGCTCTCTCTGG	Rev

Cloning strategy:

Procollagen-SBP (pro-COL-SBP; synthetic construct by MWG), mGFP and COL1a1 inserts were amplified via PCR and the last two were subsequently purified via gel extraction. Prior to assembly the remaining template vector of pro-COL-SBP-pEX-A2 was fragmented via DpnI digest. pLVXPuro was digested for 12 h (overnight) at 37 °C with EcoRI to open the vector backbone. EcoRI was heat inactivated at 65 °C for 20 min. For the NEB HiFi Assembly reaction (based on the Gibson Assembly), the compounds to be combined were added in equimolar proportions of 0.05 pmol except for procollagen-SBP* with 0.18 pmol and incubated with the assembly master mix for 3 h at 50 °C. The first assembly reaction resulted in **A**. As the generated construct was missing the pro-COL-SBP, but introduced a unique PshAI site (**B**), another reaction was performed (**C**).

B was linearized using a PshAI digest at 37 °C for 30 min with subsequent heat inactivation for 20 min at 65 °C. And the second assembly reaction was done using **B** (0.02 pmol) together with procollagenSBP** (PCR amplified; 0.04 pmol), which resulted in the final construct **C**.

We recommend growth in StbI3 and fast processing of transformed bacteria to prevent possible loss of the construct.



Complete Sequence with pLVXPuro:

gtgatgcggttttggcagtacatcaatgggctggatagcggtttgactcacggggatttccaagtctccacccattgacgtcaatgggagttt
ttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgcccattgacgcaaattgggctgtaggctgtacggtgggaggtcta
tataagcagagctcgtttagtgaacctgagatcgctggagacgcatccacgctgtttgacctcatagaagacaccgactctactagagga
tcgctagcgtaccggactcagatctcgaCACCGGTCGCCACCATGTTACGCTTTGTGGATCTGCGGCTCCTGCTGCTGC
TTGCCGCAACTGCGCTGCTCACACATGGCCAGGAAGAGGGACAGGTGGAAGGGCAGGATGAGGACATTCT
CCCATCACTTGCCTCAGAATGGGCTGAGGTACCACGATCGCGATGTCTGGAAACCCGAACCCTGTGGATT
GCGTGTGCGATAATGGCAAGGTGCTGTGTGACGACGTTATCTGCGACGAAACCAAGAAGTGTCCCGGAGCTG
AGGTCCCAGAGGGCGAATGCTGTCCGGTGTGTCCCGATGGCAGCGAGTCTCCACAGACCAGGAAACGACCG
GAGTAGAGGGCCCCGAAAGGGGACACAGGACCACGAGGACCCAGAGGTCCAGCCGGACCACCCGGTCGTGA
TGGGATACCAGGGCAACCAGGGCTGCCTGGCCCTCCTGGCCCACCTGGACCTCCTGGCCCGCCGGCTTGGG
TGGAACCTTTCAGCCGCTGCAGCGATGGACGAAAAGACCACCGGATGGAGAGGTGGGCACGTTGTGGAAG
GTCTTGCTGGCGAGTTGGAGCAGCTGCGCGCCAGACTCGAGCATCACCCACAAGGGCAGAGAGAGCCCATG
AAGACACCGGTCGCCACCATGGTGTGAGCAAGGGCGAGGAGCTGTTACCCGGGGTGGTGCCCATCCTGGTCGA
GCTGGACGGCGACGTAACCGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCA
AGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCCGTGCCCTGGCCCACCTCGTGACCACCTGAC
CTACGGCGTGAGTTCAGCCGCTACCCCGACCACATGAAGCAGCAGACTTCTTCAAGTCCGCCATGCC
GAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAG
TTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGCACAAGCTGGAGTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGGCATC
AAGGTGAAGTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAAC
ACCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCAAGCTGAGCAAAG
ACCCCAACGAGAAGCGCGATCACATGGTCTGCTGGAGTTCGTGACCGCCCGGGATCACTCTCGGCATGG
ACGAGCTGTACAAGTGTGACGCTGCCCCAGCTGTCTTATGGCTATGATGAGAAATCAACCGGAGGAATTC
CGTGCTGGCCCCATGGGTCCCTCTGGTCTCCTGGTCTCCCTGGCCCCCTGGTGCACCTGGTCCCCAAGGCT
TCCAAGGTCCCCTGGTGTGAGCCTGGCGAGCCTGGAGCTCAGGTCCCATGGGTCCCCGAGGTCCCCAGGTCC
CCCTGGAAAGAATGGAGATGATGGGGAAGCTGGAAAACCTGGTCTGCTGGTGTGAGCGTGGGCCTCCTGGGC
CTCAGGGTGTGCTGAGGATTGCCCGAACAGCTGGCCTCCCTGGAATGAAGGGACACAGAGGTTTCAGTGGTT
TGGATGGTGCCAAGGGAGATGCTGGTCTGCTGGTCTAAGGGTGTGAGCCTGGCAGCCCTGGTGAAGATGGA
GCTCCTGGTGTGAGTGGGCCCCGTGGCCTGCCTGGTGTGAGAGAGGTGCCCCCTGGAGCCCTGGCCCTGCTGGT
GCTCGTGGAAATGATGGTGTACTGGTGTGCCGGGCCCCCTGGTCCCACCGGCCCCGCTGGTCTCCTGGCT
TCCCTGGTGTGTTGGTGCTAAGGGTGAAGCTGGTCCCCAAGGGCCCCGAGGCTCTGAAGGTCCCCAGGGTG
TGCGTGGTGTGAGCCTGGCCCCCTGGCCCTGCTGGTGTGCTGGCCCTGCTGGAAACCTGGTGTGATGGAC
AGCCTGGTGTAAAGGTGCAATGGTGTCTCTGGTATTGCTGGTGTCTCTGGCTTCCCTGGTGGCCGAGGCC
CTCTGGACCCAGGGCCCCGGCGGCCCTCCTGGTCCCAAGGGTAACAGCGGTGAACCTGGTGTCTCTGGCAG
CAAAGGAGACTGTTGCTAAGGGAGAGCCTGGCCCTGTTGGTGTTCAGGACCCCTGGCCCTGCTGGAGA
GGAAGGAAAGCGAGGAGCTCGAGGTGAACCCGGACCCACTGGCCTGCCCGGACCCCTGGCGAGCGTGGTG
GACCTGGTAGCCGTGGTTTCCCTGGCGCAGATGGTGTGCTGGTCCCAAGGGTCCCCTGGTGAACGTGGTT
CTCCTGGCCCTGCTGGCCCCAAGGATCTCCTGGTGAAGCTGGTGTGCTCCCGGTGAAGCTGGTGTGCTGGTGC
CAAGGGTCTGACTGGAAGCCCTGGCAGCCCTGGTCTGATGGCAAACTGGCCCCCTGGTCCCAGCCGGTCA
AGATGGTGTCCCCGACCCCGAGGCCACCTGGTGGCCGTGGTGTGATGGGATTCCCTGGACC
TAAAGGTGCTGCTGGAGAGCCCGGCAAGGCTGGAGAGCGAGGTGTTCCCGGACCCCTGGCGCTGTGCGTC
CTGCTGGCAAAGATGGAGAGGCTGGAGCTCAGGGACCCCTGGCCCTGCTGGTCCCCTGGCGAGAGAGGT
GAACAAGGCCCTGCTGGCTCCCCGATTCCAGGGTCTCCTGGTCTGCTGGTCTCCAGGTGAAGCAGGCA
AACCTGGTGAACAGGGTGTTCCTGGAGACCTTGGCGCCCTGGCCCTCTGGAGCAAGAGGGCAGAGAGGTT
TCCCTGGCGAGCGTGGTGTGAAGGTCCCCCTGGTCTGCTGGTCCCCGAGGGGCCAACGGTGTCCCCGCA
ACGATGGTGTGAAGGGTGTGCTGGTGGCCCTGGAGCTCCCGGTAGCCAGGGCGCCCTGGCCTTCAGGGAA

TGCTGGTGAACGTGGTGCAGCTGGTCTTCCAGGGCCTAAGGGTGACAGAGGTGATGCTGGTCCCAAAGGTG
CTGATGGCTCTCCTGGCAAAGATGGCGTCCGTGGTCTGACCGGCCCATTTGGTCTCCTGGCCCTGCTGGTGC
CCCTGGTGACAAGGGTGAAAGTGGTCCCAGCGGCCCTGCTGGTCCCCTGGAGCTCGTGGTGGCCCCGAGA
CCGTGGTGAGCCTGGTCCCCCGGCCCTGCTGGCTTTGCTGGCCCCCTGGTGGTGCACGGCCAACCTGGTGT
AAAGGCGAACCTGGTGTGCTGGTGTAAAGGCGATGCTGGTCCCCTGGCCCTGCCGGACCCGCTGGACCC
CCTGGCCCCATTGGTAATGTTGGTGTCTCTGGAGCCAAAGGTGCTCGCGGCAGCGCTGGTCCCCTGGTGTCTA
CTGGTTTCCCTGGTGTGCTGGCCGAGTCGGTCTCCTGGCCCCTCTGGAAATGCTGGACCCCTGGCCCTCCT
GGTCTGCTGGCAAAGAAGGCGCAAAGGTCCCCTGGTGGAGACTGGCCCTGCTGGACGTCTGGTGAAGTT
GGTCCCCTGGTCCCCTGGCCCTGCTGGCGAGAAAGGATCCCCTGGTGTGATGGTCTGCTGGTGTCTCTG
GTACTCCCGGCCCTCAAGGTATTGCTGGACAGCGTGGTGTGGTGGCCTGCCTGGTCAAGAGAGGAGAGAGA
GGCTTCCCTGGTCTTCTGGCCCCTCTGGTGAACCTGGCAAACAAGGTCCCTCTGGAGCAAGTGGTGAACGTG
GTCCCCCTGGTCCCATGGGCCCCCTGGATTGGTGGACCCCTGGTGAATCTGGACGTGAGGGGGCTCCTG
GTGCCGAAGGTTCCCCTGGACGAGACGTTCTCCTGGCGCAAAGGGTGACCGTGGTGGAGACCGGCCCGCTG
GACCCCTGGTGTCTCTGGTGTCTCTGGTGGCCCTGGCCCCGTTGGCCCTGCTGGCAAGAGTGGTGTGCTGG
TGAGACTGGTCTGCTGGTCCCGCCGTCTGTGGCCCTGTTGGCGCCCGTGGCCCCGCGGACCCCAAGGC
CCCCGTGGTGACAAGGGTGAGACAGGCGAACAGGGCGACAGAGGCATAAAGGGTCACCGTGGCTTCTCTGG
CCTCCAGGGTCCCCTGGCCCTCTGGTCTCCTGGTGAACAAGGTCCCTCTGGAGCCTCTGGTCTGCTGGTGC
CCCGAGGTCCCCTGGTCTGCTGGTGTCTCTGGCAAAGATGGACTCAACGGTCTCCCTGGCCCCATTGGGCC
CCCTGGTCTCGCGTGCAGTGGTGTGCTGGTCTGTTGGTCCCCCGGCCCTCTGGACCTCCTGGTCCCC
CTGGTCTCCAGCGCTGGTTTCGACTTCAGTTCCTGCCCCAGCCACCTCAAGAGAAGGCTCACGATGGTGG
CCGCTACTACGGGCTGATGATGCCAATGTGGTTCGTGACCGTGCAGCTCGAGGTGGACACCACCTCAAGAGC
CTGAGCCAGCAGATCGAGAACATCCGAGCCAGAGGGCAGCCGCAAGAACCCCGCCGCACCTGCCGTGA
CCTCAAGATGTGCCACTCTGACTGGAAGAGTGGAGAGTACTGGATTGACCCCAACCAAGGCTGCAACCTGGA
TGCCATCAAAGTCTTCTGCAACATGGAGACTGGTGGAGACCTGCGTGTACCCCACTCAGCCCAGTGTGGCCAG
AAGAAGTGGTACATCAGCAAGAACCCCAAGGACAAGAGGCATGTCTGGTTCGGCGAGAGCATGACCGATGG
ATTCCAGTTCAGATATGGCGGCCAGGGCTCCGACCCTGCCGATGTGGCCATCCAGCTGACCTTCTGCGCCTG
ATGTCCACCGAGGCCTCCAGAACATCACCTACCACTGCAAGAACAGCGTGGCCTACATGGACCAGCAGACTG
GCAACCTCAAGAAGGCCCTGCTCCTCCAGGGCTCCAACGAGATCGAGATCCGCGCCGAGGGCAACAGCCGCT
TCACCTACAGCGTCACTGTGATGGTGCACGAGTACACCCGAGCCTGGGGCAAGACAGTGAATGATAACA
AAACCACCAAGTCTCCCGCCTGCCATCATGATGTGGCCCCCTGGACGTTGGTGGCCCCAGACCAGGAATT
CGGCTTCGACGTTGGCCCTGTCTGCTTCTGTAActgcagtcgacggtagccgaggccgggatcccgcactctagataattct
accggtaggggaggcgctttccaaggcagtcgagcatgccttagcagccccgctgggactggcgctacacaagtggcctctggcct
cgcacacattccacatccaccgtaggcgccaaccggctcgttctttggtagcccttcgcccactctactctcccctagtcaggaagtcc
ccccgccccgagctcgcgctgagcagcgtgacaaatggaagtagcacgctctcactagtctcgtcagatggacagcaccgctgagcaatg
gaagcggttaggcctttggggcagcggccaatagcagctttgctccttctgctttctgggctcagaggctgggaaggggtgggtccggggcggg
ctcaggggcgggctcagggcgggcgggcgcccgaagtcctccggaggcccgattctgcagcttcaaagcgacgctgcccgcgctg
ttcctcttctctatctccggcctttcagctgcagccaagcttaccatgaccgagtaagccacggctgcctcgcaccccgagcagc
tccccaggccgtacgcaccctcgcgcccgttcgcccactccccgccacgccaaccctcgtacccgaccgccaacatcgagcgggtcac
cgagctgcaagaactcttctcagcgcgctgggctcgacatcggcaaggtgtgggtcggcagcagcgccgctggcggtctggaccag
cgggagagcgtcgaagcggggcggtgttcgcccagatcggcccgcgcatggccgagttgagcgggtccggctggccgcgagcaacagatg
gaaggcctctggcggcaccggccaaggagcccgcgtggttctggccaccctcggcgtctcggccaccaccaggggcaagggtctgggc
agcggctcgtctcccggagtgaggcgggcagcgcgcccggggtgcccgccttctggagacctccgccccgcaacctccccttctacg
agcggctcggcttaccgtcaccgcccagcgtcagaggtcccgaaggaccgcacctggtgcatgaccgcaagcccggtgctgaccgctc
ggaacaatcaacctgattacaaatgtgaaagattgactggtattcttaactatggtgctcttttacgctatgtgatacgtgctttaatg
cctttgatcatgctattgcttcccgtatggcttctcttctcctctgtataaatctggtgctgctctttatgaggagttggtggccgtgtgag
gcaacgtggcgtggtgtgactgtgttctgacgcaacccccactggttggggcattgcccaccctgtcagctctttccgggacttctgcttc
cccctccattgccaaggggcaactcatcgccctgcttcccgcgtgctggacaggggctggctgttgggactgacaattccgtggtggtg
cggggaagctgacgtccttccatggctgctcgcctgtgttccacctggattctgcggggacgtccttctgctacgtcccttcggccctcaatcca

gcggaccttccctcccgccgctgctgccggctctcgccgcttccgcttccgcttccgcttccgctcagacgagtcggatctcccttgggcccctc
cccgcctggaattaattctgcagtcgagacctagaaaaacatggagcaatcacaagtagcaatacagcagctaccaatgctgattgtcctggc
tagaagcacaagaggaggagggtgggtttccagtcacacctcaggtaccttaagaccaatgacttaacaggcagctgtagatcttagcca
cttttaaaagaaaagaggggactggaaggctaattcactccaacgaagacaagatacttgatctgtggatctaccacacacaaggctac
ttcctgattagcagaactacacaccaggccagggtcagatatccactgaccttggatgggtgctacaagctagtagcagttgagccagataa
ggtagaagaggccaataaaggagagaacaccagctgtttacacctgtgagcctgcatgggatggatgacccggagagagaagtgttagagt
ggaggttgacagccgctagcatttcatcacgtggccgagagctgcatccggagtactcaagaactgctgatatcgagcttctacaaggga
ctttccgctggggactttccagggaggcgtggcctgggaggactggggagtgccgagccctcagatcctgcatataagcagctgcttttgcctg
tactgggtctctggttagaccagatctgagcctgggagctctctggtaactagggaaaccactgcttaagcctcaataaagcttgccttgagt
gcttcaagtagtgtgcccgtctgtgtgactctggtaactagagatccctcagacccttttagtcagtgaggaaatctctagcagtagtagt
catgtcatcttattttagtatttataactgcaagaaatgaatatcagagagtgagaggccttgacattgctagcgttttaccgtcgacctta
gctagagcttggcgtaatcatggatcatagctgttctgtgtgaaattgttatccgctcacaattccacacacatacagccggaagcataaagt
gtaaagcctgggggtcctaatgagtgagtaactacattaattgcttgcgctcactgccgcttccagtcgggaaacctgctgccaagctgc
attaatgaatcgccaacgcgaggagaggcggttgcgtattgggctcttccgcttccgctcactgactcgcgctcggctgcttccgct
gaggcgagcggatcagctcactcaaaggcgtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggcc
agcaaaaggccaggaaccgtaaaaaggccgcttggcgttttccataggtcggccccctgacgagcatcaaaaaatcgacgctcaa
gtcagaggtggcgaacccgacaggactataaagataccaggcgttccccctggaagctcctcgtgctcctctgttccgacctgcccgtt
accggatacctgtccgcttctcccttgggaaagcgtggcgttctctatagctcagctgtaggtatctcagttcgggttaggtcgttccgctcaa
gctgggctgtgtgcagcaacccccgttccagccgacctgctgccttaccgtaactatcgtcttgagccaaccggtaagacacgacttacc
gccactggcagcagccactggtaacaggattagcagagcaggtatgtagggcgtgctacagagttctgaagtgggtggcctaactacggctac
actagaagaacagtatgtgctgctgctgaagccagttaccttggaaaaagagttgtagctcttgatccggcaaaacaaaccaccgc
tgtagcgggtgtttttgttgaagcagcagattacgcgcaaaaaagagatcacaagaagatccttgatcttttctacggggtctgacgct
cagtggaacgaaaactcacgttaagggtttgttcatgagattatcaaaaaggatcttccactagatcctttaaatataaagttttaa
tcaatcaaaagtataatagtaaaacttggctgacagttaccaatgcttaatcagtgaggcaactatctcagcagatctgtctatttctgctcatccat
agttgctgactccccgtgctgtagataactacgatacgggagggttaccatctggccccagtgctgcaatgataccgcgagaccacgctcac
cggctccagattatcagcaataaaccagccagccggaaggccgagcgcagaagtggctcctgcaactttatccgctccatccagcttattaat
tgttccgggaagctagagtaagtagttcgccagttatagtttgcgcaacgttgttccattgctacaggcatcgtggtgtcacgctcgtctgtt
gtatggcttattcagctccggttcccaacgatcaaggcaggttacatgatccccatgttggcaaaaaagcgggttagctcctcggctcctccgat
cgttgcagaagtaagttggccgaggttatcactcatggtatggcagcactgcataattcttactgtcatgccatccgtaagatgctttctg
tgactggtgagtactcaaccaagtcattctgagaatagtgatgcccgcagcaggttgccttcccggcgtcaatacgggataataccgcgcca
catagcagaacttaaaagtgtcatcattggaacgttcttggggcgaaaactctcaaggatcttaccgctgttgagatccagttcagtgtaa
cccactgctcacccaactgatcttcagatctttactttaccagcgttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagg
aataaggcgacacggaatgttgaatactcatacttcttttcaatatttgaagcattatcagggtattgtctcatgagcggatacatat
ttgaatgtatttagaaaaatacaaatagggttccgcgacattccccgaaaagtgccacctgacgtcgacggatcgggagatcaactgtt
tattgcagcttataatggttacaataaagcaatagcatcaaaatccacaaataaagcatttttactgactttagttgtgtgttgcacaa
tcatcaatgtatcttcatgtctggatcaactggataactcaagtaacaaaatcatccaaactcccaccatacctattaccactgcaa
ttacctgtggttcttactctaaactgtgattcctctgaatttttcttttaagaaattgtatttgttaaatagtactacaaacttagtagtt
ttaagaaattgtattgttaaatagtactacaaacttagtagttggaaggctaattcactccaaagaagacaagatacttgatctgtgga
tctaccacacacaaggctacttccctgattagcagaactacacaccaggccagggtcagatatccactgaccttggatgggtgctacaagcta
gtaccagttgagccagataaggtagaagaggccaataaaggagagaacaccagctgtttacacctgtgagcctgcatgggatggatgacccg
gagagagaagtgttagagtgagggttgacagccgctagcatttcatcacgtggccgagagctgcatccggagtacttcaagaactgctgat
atcgagcttctacaaggactttccgctggggactttccagggaggcgtggcctgggaggactggggagtgccgagccctcagatcctgcat
ataagcagctgcttttgcctgtactgggtctctggttagaccagatctgagcctgggagctctctggtaactagggaaaccactgcttaagcc
tcaataaagcttgccttgagtcttcaagtagtgtgtgcccgtctgtgtgactctggtaactagagatccctcagacccttttagtcagtgtaga
aaatctctagcagtgccgcccgaacagggacttgaagcgaagggaaccagaggagctctctcagcagcaggactcggcttctgtaagcgc
gcacggcaagaggcgaggggcggcgactggtgagtagccaaaaattttagctagcggaggctagaaggagagagatgggtgagagcgt
cagtttaagcgggggagaattagatcgcgatgggaaaaatcggtaaggccagggggaaagaaaaataaataaataatagtagt
gggcaagcaggagctagaacgattcaggttaactcggcctgttagaaacatcagaaggctgtagacaaatactgggacagctacaacct

cccttcagacaggatcagaagaacttagatcattataataacagtagcaaccctctattgtgtgcatcaaaggatagagataaaagacaccaa
ggaagctttagacaagatagaggaagagcaaaacaaaagtaagaccaccgcacagcaagcggccggcctgatcttcagacctggaggag
gagatatgagggacaattggagaagtgaattatataataataaagtagtaaaaattgaaccattaggagtagcaccaccaaggcaaagaga
agagtggcagagagaaaaagagcagtggggaataggagctttgttccttgggttcttgggagcagcaggaagcactatggcgagcgtca
atgacgctgacggtagcaggccagacaattattgtctggtatagtcagcagcagaacaatttgctgagggctattgagggcgaacagcatctgtt
gcaactcacagctctggggcatcaagcagctccaggcaagaatcctggctgtggaaagatacctaaaggatcaacagctcctggggatttggggt
tgctctggaaaactcattgaccactgctgtgccttggatgctagtggagtaataaatctctggaacagatttggatcacacgacctggatg
gagtgggacagagaaattaacaattacacaagcttaatacactccttaattgaagaatcgaaaaccagcaagaaaagaatgaacaagaatt
attggaattagataaatgggcaagtttgggaattggttaacataacaaattggctgtggtatataaaattattcataatgatagtaggagctt
gtaggttaagaatagttttgctgactttctatagtagaatagagttaggcagggatattcaccattatcgtttcagaccacctccaacccg
aggggacccgacaggccgaaggaatagaagaagaaggtggagagagagacagagacagatccattcgattagtgaaacggatctcgacggt
atgcctttaaagaaaaggggggattgggggtacagtcaggggaaagaatagtagacataatagcaacagacatacaactaaagaatt
acaaaaacaaattacaaaaattcaaaatttcgggtttattacagggacagcagagatccagtttctgataagcttgggagttccggttacat
aacttacggtaaatggcccgctggctgaccgccaacgacccccgccattgacgtcaataatgacgtatgtcccatagtaacgccaataggg
actttcattgacgtcaatgggtggagatttacggtaaactgccacttggcagtacatcaagtgtatcatatgccaagtagccccctattgac
gtcaatgacggtaaatggcccgctggcattatgccagtagatgaccttatgggactttcctacttggcagtagatctacgtattagtcacgcta
ttacatg