

## ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> Автономная некоммерческая образовательная организация высшего образования «Сколковский институт науки и технологий»/ Autonomous Non-Profit Organization for Higher Education "Skolkovo Institute of Science and Technology"

<120> СРЕДСТВО РАЗРЕЗАНИЯ ДНК НА ОСНОВЕ CAS9 БЕЛКА ИЗ DEFLUVIIMONAS SP.

<130> 417254

<150> RU 2018141534

<151> 2018-11-26

<160> 2

<170> BiSSAP 1.3.6

<210> 1

<211> 1079

<212> PRT

<213> Defluviimonas

<220>

<223> белок, гомологичный Cas9

<400> 1

Met	Tyr	Arg	Phe	Ala	Phe	Asp	Leu	Gly	Thr	Asn	Ser	Leu	Gly	Trp	Ala	
1			5						10					15		
Val	Tyr	Arg	Leu	Thr	Asn	Asp	Thr	Arg	Pro	Ala	Ser	Ile	Glu	Arg	Thr	
			20					25					30			
Gly	Val	Arg	Ile	Phe	Pro	Asn	Gly	Arg	Asp	Pro	Gln	Ser	Lys	Glu	Ser	
		35					40					45				
Asn	Ala	Ala	Gly	Arg	Arg	Met	Pro	Arg	Gly	Ala	Arg	Arg	Arg	Gln	Asp	
	50					55					60					
Arg	Ser	Leu	Gly	Arg	Arg	Lys	Arg	Leu	Leu	Asp	Asp	Leu	Ile	Gly	Phe	
65					70					75					80	
Gly	Leu	Leu	Pro	Ser	Asp	Ala	Asp	Ala	Arg	Ser	Ala	Val	Phe	Ala	Ala	
				85					90					95		
Asn	Pro	Ile	Glu	Ala	Arg	Ala	Arg	Ala	Ala	Arg	Glu	Gln	Val	Ala	Leu	
			100					105					110			
Glu	Gln	Leu	Gly	Arg	Ala	Leu	Trp	His	Met	Ser	Lys	His	Arg	Gly	Phe	
		115					120					125				
Lys	Ser	Asn	Arg	Arg	Ala	Asp	Lys	Asp	Ala	Asp	Glu	Lys	Gly	Lys	Ile	
	130					135					140					
Ala	Ile	Ala	Ser	Ala	Ala	Leu	Leu	Glu	Arg	Leu	Arg	Ala	Asp	Gly	His	
145					150					155					160	
Pro	Thr	Tyr	Gly	Ala	Phe	Leu	His	Ala	Arg	Leu	Val	Arg	Gly	Glu	Gly	
				165					170					175		
Thr	Arg	Ile	Arg	Pro	Ala	Gly	Asp	Gly	Ala	Lys	Leu	Ser	Tyr	Glu	Phe	
			180					185						190		
Tyr	Pro	Thr	Arg	Ala	Leu	Leu	Glu	Ala	Glu	Phe	Asp	His	Ile	Trp	Ala	
		195					200					205				
Ile	Gln	Ala	Gly	Phe	His	Pro	Ser	Leu	Thr	Glu	Ala	Met	His	Glu	Arg	
	210					215					220					
Leu	Arg	Asp	Thr	Ile	Phe	Phe	Gln	Arg	Pro	Leu	Arg	Pro	Val	Arg	Pro	
225					230					235					240	
Gly	Lys	Cys	Thr	Phe	Phe	Pro	Asp	Gln	Asp	Arg	Leu	Pro	Arg	Trp	His	
				245					250					255		
Pro	Ala	Ala	Gln	Glu	Phe	Leu	Ile	Leu	Ser	Gln	Leu	Asn	His	Leu	Arg	
			260					265						270		

Ile	Val	Asp	Asp	Gln	Gly	Glu	Gln	Pro	Leu	Asp	Ile	Thr	Ala	Arg	Asp
		275					280						285		
Leu	Val	Ala	Arg	Thr	Leu	Met	Ala	Gly	Thr	Lys	Leu	Ser	Trp	Ser	Gly
		290				295					300				
Leu	Lys	Lys	Thr	Leu	Lys	Leu	Pro	Ser	Gln	Ala	Glu	Phe	Asn	Leu	Glu
305					310					315					320
Lys	Gly	Gly	Leu	Lys	Glu	Leu	Ala	Arg	Asn	Asp	Val	Ala	Ala	Arg	Leu
				325					330					335	
Leu	Gly	Asp	Thr	Lys	Lys	Pro	Gly	Pro	Leu	Ala	Thr	Leu	Trp	Pro	Thr
			340					345						350	
Leu	Asp	Ser	Ala	Thr	Gln	Glu	Glu	Ile	Leu	Trp	Gln	Ile	Ser	Glu	Val
		355					360						365		
Ala	Asp	Pro	Glu	Glu	Leu	Ile	Cys	Trp	Leu	Met	Glu	Arg	Leu	Gly	Leu
		370				375					380				
Ser	Arg	Glu	Val	Ala	Glu	Arg	Val	Glu	Lys	Ile	Pro	Leu	Pro	Asp	Gly
385					390					395					400
His	Leu	Arg	Phe	Cys	Lys	Thr	Ala	Thr	Gln	Ala	Ile	Val	Glu	Lys	Leu
				405					410					415	
Arg	Cys	Asp	Val	Ile	Pro	Tyr	Asp	Glu	Ala	Val	Arg	Arg	Ala	Pro	Leu
			420					425					430		
Leu	Gly	Gly	Ala	Gly	Leu	Asp	His	Ser	Asp	Phe	Gln	Ala	Glu	Glu	Gly
		435					440					445			
Val	Asp	Thr	Leu	Pro	Pro	Tyr	Asn	Arg	Leu	Pro	Val	Leu	Gln	Arg	Met
	450					455					460				
Ile	Gly	Asn	Gly	Thr	Ser	Asp	Pro	Lys	Asp	Thr	Asp	Leu	Leu	Arg	Tyr
465					470					475					480
Gly	Arg	Ile	Thr	Asn	Pro	Thr	Val	His	Ile	Ala	Leu	Gly	Gln	Phe	Arg
				485					490					495	
Arg	Val	Met	Asn	Ala	Leu	Ile	Ala	Glu	Tyr	Gly	Lys	Pro	Ala	Gln	Val
			500					505					510		
Val	Ile	Glu	Ala	Thr	Arg	Asp	Met	Ala	Lys	Ser	Ala	Glu	Glu	Leu	Asn
		515					520					525			
Lys	Ile	Glu	Lys	Thr	Ile	Arg	Asp	Asn	Glu	Lys	Arg	Asn	Asp	Arg	Trp
	530					535					540				
Arg	Ala	Glu	Leu	Gln	Lys	Ala	Gly	Leu	Leu	Val	Glu	Gly	Ala	Arg	Ile
545					550					555					560
Gly	Asp	Arg	Phe	Leu	Arg	Met	Arg	Leu	Trp	Glu	Glu	Ile	Gly	Arg	Gly
				565					570					575	
Pro	Ala	Asp	Arg	Leu	Cys	Pro	Tyr	Thr	Gly	Arg	Pro	Ile	Ala	Leu	His
			580					585					590		
Gln	Leu	His	Ser	Asp	Glu	Ile	Glu	Ile	Asp	His	Ile	Leu	Pro	Phe	Glu
		595					600					605			
Gln	Thr	Phe	Asp	Asp	Ser	Pro	Ala	Asn	Lys	Ser	Leu	Cys	Phe	Arg	Asp
	610					615					620				
Ala	Asn	Arg	Arg	Lys	Gly	Lys	Leu	Ser	Pro	Ala	Ala	Ala	Ala	Asp	Arg
625					630					635					640
Gln	Pro	Asp	Phe	Phe	Asp	Leu	Ala	Ala	Ile	Ile	Gly	Arg	Thr	Lys	His
				645					650					655	
Leu	Pro	Ala	Asn	Lys	Ala	Trp	Arg	Phe	Leu	Pro	Gly	Ala	Met	Glu	Lys
			660					665					670		
Trp	Glu	Glu	Thr	Arg	Gly	Phe	Glu	Asp	Arg	Gln	Leu	Asn	Ala	Thr	Gly
		675					680					685			
Tyr	Leu	Ala	Arg	Val	Val	Arg	Ala	Tyr	Thr	Glu	Ala	Leu	Phe	Pro	Lys
	690					695					700				
Thr	Asp	Ala	Glu	Gly	Lys	Arg	Arg	Ser	His	Val	Trp	Val	Leu	Pro	Gly
705					710					715					720
Arg	Met	Thr	Ala	Met	Leu	Arg	His	Arg	Trp	Gly	Leu	Asn	Leu	Gly	Asp
				725					730					735	
His	Asn	Arg	Lys	Ser	Arg	Asp	Asp	His	Arg	His	His	Ala	Ile	Asp	Ala
			740					745					750		

Ala Val Val Gly Val Ile Asp Arg Arg Met Val Gln Val Leu Gln Thr  
755 760 765  
His Ala Arg Asn Leu Gly Val Glu Arg Leu Asp Arg Val Leu Pro Ala  
770 775 780  
Pro Pro Glu Pro Phe Glu Gly Phe Arg Asp Ala Val Leu Ala Ala Val  
785 790 795 800  
Glu Lys Val Asn Val Ser His Arg Ala Gln His Gly Ser Ile Asp Pro  
805 810 815  
Thr Asp Pro Ser Gln Thr Ser Gly Arg Leu His Glu Asp Thr Val Phe  
820 825 830  
Gly Leu Val Arg Asp Val Pro Glu Asn Gln Ala Glu Arg Thr Ile Gly  
835 840 845  
Asn Val Val Val Arg Lys Pro Ile Thr Gly Leu Thr Glu Lys Glu Ile  
850 855 860  
Gly Gln Val Arg Asp Val Lys Met Arg Leu Ser Leu Gln Asp Ala Thr  
865 870 875 880  
Ser Leu Ala Arg Asp Lys Ser Leu Pro Glu Ala Glu Arg Lys Lys Arg  
885 890 895  
Leu Pro Glu Ala Leu Ala Lys Trp Val Gln Glu Thr Gly His Arg Lys  
900 905 910  
Leu Arg Ile Leu Lys Pro Glu Ala Gly Val Arg Pro Val His Asp Arg  
915 920 925  
Asp Gly Lys Pro Tyr Lys Trp Leu Val Pro Gly Glu Ile Ser Trp Leu  
930 935 940  
Asp Ile Leu Glu Ala Pro Asp Gly Thr Trp Phe Gln His Ala Thr Asp  
945 950 955 960  
Ile Trp Ala Ala Asn Ser Gly Gly Ala Glu Pro Trp Asn Val Ala His  
965 970 975  
Pro Ser Ala Arg Phe Ile Met Arg Val His Lys Asn Asp Thr Ile Gln  
980 985 990  
Val Phe Asp Trp Asp Asp Arg Asn Lys Cys Val Val Glu Gly Ser Asn  
995 1000 1005  
Gln Ile Lys Arg Ile Val Arg Leu Glu Pro Ser Ala Asn Arg Met Arg  
1010 1015 1020  
Leu Val Gly Ile Asn Asp Ala Gly Lys Phe Asp Glu Arg His Asn Glu  
1025 1030 1035 1040  
Pro Asn Asp Pro Phe Arg Trp Asp Leu Ala Thr Ile Ser Lys Leu Lys  
1045 1050 1055  
Leu Arg Arg Ala Arg Arg Val Arg Ile Asp Glu Leu Gly Arg Val His  
1060 1065 1070  
Thr Ile Pro His Gly Thr Val  
1075

<210> SEQ ID NO 2  
<211> 3240  
<212> DNA  
<213> Defluviimonas

<400> 2  
atgtatcgtt tcgctttcga cctcggaacc aactcccttg gctgggccgt ctatcgcttg 60  
accaatgaca cgcgccccgc ttccattgaa aggacagggg tgaggatctt cccaaacggt 120  
cgggatccac agtcgaagga gtcgaatgcc gctggccgca gaatgccgcg cggcgcaagg 180  
cggaggcagg atcgaagcct gggacgaagg aagcggctgc tcgatgacct gatcgggttt 240  
ggcctgctgc catcagatgc agatgccgc agcgcggtct ttgccgcaaa tcccatcgaa 300

gcacgggctc gggccgcgcg cgaacaggtg gctctggaac aacttggccg ggcactctgg 360  
cacatgtcga aacatcgcgg tttcaagagc aaccgccgcg ccgacaagga tgcggatgaa 420  
aagggcaaga tcgcgatcgc ctcggccgcc cttctagagc gactgagggc agacggggcac 480  
ccgacatacg gcgcctttct gcacgcgcgc ctcgctccgag gtgaaggaac gcgcatacgc 540  
ccggcaggcg acggcgcaaa gctgtcctac gaattctacc cgacacgcgc tcttctggag 600  
gccgagtctg atcacatctg ggccatccaa gcgggcttcc acccgtccct gaccgaagcg 660  
atgcacgagc ggetgcgcga caccatcttc tttcagcgcc ccctgcgccc cgtccgtccg 720  
gggaaatgca cattcttccc cgatcaggac cggctgccgc gctggcaccg gcccgcacag 780  
gagtttctga tcctgtcaca gctcaacct ttgcgcatcg tcgacgatca aggcgaacag 840  
ccactggaca tcaccgcgcg ggatcttgtc gcgcgtacgt tgatggctgg aacaaaactg 900  
tcatggctcg gctgaaaaa gaccctgaag ctgccgtcac aggccgagtt caacctcgaa 960  
aagggcggcc tgaaggagct tgcacggaat gacgtggcgg cacgactgct gggcgatacg 1020  
aagaaacccg gaccgctggc gacgctctgg ccgaccctcg acagcgccac gcaagaggaa 1080  
atcctgtggc agatatcgga ggtcgcagac ccggaagagc tcactctgctg gcttatggaa 1140  
aggctcgggc tttctcgcga ggtggcggaa agggctgaga agatcccact gccggacggt 1200  
cacctgcggt tctgcaagac agcgcgcgag gcgattgtcg aaaaactgcg ctgcgacgtc 1260  
attccctatg acgaagccgt ccggcgcgcg cttttgcttg gcggggcggg gctggaccac 1320  
tccgattttc aggccgaaga aggcgtcgat accctgccgc cctacaaccg gctgcccggt 1380  
ctacagcgca tgatcggcaa cggaaaccagc gatccgaaag acacggacct gctgcgctat 1440  
ggccggatca ccaaccggac cgtgcacatc gccttgggcc agttccggcg cgtgatgaat 1500  
gcgctgatcg cagaatacgg taagcctgcg caggtcgtca tcgaggccac gcgagacatg 1560  
gccaagttag ccgaggaact gaacaagatc gagaaaacca tccgcgacaa cgaaaaacgc 1620  
aatgatcgtt ggcgtgccga gcttcagaaa gcggggctgc tgggtggaagg cgcgcggatc 1680  
ggcgaccggt tcttgaggat gcgcctatgg gaggagattg gcagaggccc ggcagatcgg 1740  
ctctgtccct acaccggtcg ccccatcgcg ctgcatcagc tgcatccga cgagatcgag 1800  
atcgatcata tcctgccctt cgagcaaacc ttcgacgaca gtccagccaa caagagcctg 1860  
tgtttccgcg atgccaaccg ccgcaagggc aagctgtcgc cggctgcagc ggccgaccgg 1920  
cagccagact tcttcgatct ggctgccatc atcgggcgca cgaaacacct gcccgccaac 1980  
aaggcctggc ggtttctgcc cggtgccatg gaaaaatggg aagagacgcg cggtttctgag 2040  
gatcgccagt tgaacggcac cggctatctg gcccgagttg tccgcgccta taccgaggcg 2100  
ctgttcccga aaaccgatgc cgagggcaag cgacgcagcc atgtctgggt cctgcccggg 2160  
cgcatgaccg cgatgttgcg gcaccgctgg ggtctgaacc tgggcgatca caaccgtaag 2220

tcacgcgacg atcaccgccca ccatgcgatc gacgccgccg tggtagggcgt gatcgaccgg	2280
cggatggtgc aggtggttga gaccatgcg cggaacctgg gggtaggagcg gctggaccgc	2340
gtgttaccgg cgccgccaga acccttcgaa ggtttccgcg acgccgtgct ggctgccgtc	2400
gaaaagggtca acgtcagcca tcgtgcgcaa cacggcagca tcgaccaac cgatccatcg	2460
cagaccagcg gcaggctgca tgaggataca gtctttggtc ttgtgcggga tgtgccagag	2520
aaccaggccg agcgcacat cggcaatgtc gttgtgcgca aaccataac cggcttgacc	2580
gaaaaggaaa tcggacaggt gcgcgacgtc aagatgcgcc tcagtctgca ggatgccacc	2640
agtttagccc gcgacaaatc cttgccggag gcagagcgca agaagcgctt gcccgaggcg	2700
ctggcgaaat gggttcagga aaccggccac cgcaaactgc gtatcctgaa gcccgaaact	2760
ggcgttcgcc ccgttcatga ccgtgatggc aagccctaca aatggctggt gcccggtgaa	2820
atctcctggc tcgacatcct agaagcgctt gacggcacat ggtttcagca cgccacagac	2880
atctggggcg caaactcggg tggggccgag ccctggaacg tcgcgcaccc cagcgcacgt	2940
ttcatcatgc gcgtccacaa gaatgacacc attcaagtgt tcgattggga cgacagaaac	3000
aagtgtgtcg tggaaggaag caaccagatc aagaggatcg tacgcttggga accgagcgcc	3060
aacagaatgc ggctcgtcgg cataaacgac gcaggcaagt ttgacgaacg gcacaacgag	3120
ccaaacgatc ccttccgctg ggatcttgcc acgatcagca aactgaagct ccgccgcgcc	3180
cggcgcgtgc ggatcgacga attggggcgg gttcacacga tcccgcattg gacagtttag	3240