

SEQUENCE LISTING

<110> DSM IP Assets BV
 Verwaal, René
 Wu, Liang
 Sagt, Cornelis Maria Jacobus
 Damveld, Robbertus Antonius

<120> Dicarboxylic acid production in recombinant yeast

<130> 26346

<140> 26346WO
 <141> 2008-11-14

<160> 6

<170> PatentIn version 3.2

<210> 1
 <211> 494
 <212> PRT
 <213> Rhizopus oryzae

<400> 1

Met Leu Arg Ala Ser Ala Thr Arg Phe Leu Ser Gln Ala Lys Asn Met
 1 5 10 15

Asn Asn Ser Pro Arg Leu Phe Ser Ser Ala Ser Ala Ala Leu Gln Lys
 20 25 30

Phe Arg Ala Glu Arg Asp Thr Phe Gly Asp Leu Gln Val Pro Ala Asp
 35 40 45

Arg Tyr Trp Gly Ala Gln Thr Gln Arg Ser Leu Gln Asn Phe Asp Ile
 50 55 60

Gly Gly Pro Thr Glu Arg Met Pro Glu Pro Leu Ile Arg Ala Phe Gly
 65 70 75 80

Val Leu Lys Lys Ala Ala Ala Thr Val Asn Met Thr Tyr Gly Leu Asp
 85 90 95

Pro Lys Val Gly Glu Ala Ile Gln Lys Ala Ala Asp Glu Val Ile Asp
 100 105 110

Gly Ser Leu Ile Asp His Phe Pro Leu Val Val Trp Gln Thr Gly Ser
 115 120 125

Gly Thr Gln Thr Lys Met Asn Val Asn Glu Val Ile Ser Asn Arg Ala
 130 135 140

Ile Glu Leu Leu Gly Gly Glu Leu Gly Ser Lys Ala Pro Val His Pro
 145 150 155 160

Asn Asp His Val Asn Met Ser Gln Ser Ser Asn Asp Thr Phe Pro Thr
 165 170 175

Ala Met His Val Ala Ala Val Val Glu Ile His Gly Arg Leu Ile Pro
 180 185 190

Ala Leu Thr Thr Leu Arg Asp Ala Leu Gln Ala Lys Ser Ala Glu Phe
 195 200 205

Glu His Ile Ile Lys Ile Gly Arg Thr His Leu Gln Asp Ala Thr Pro
 210 215 220

Leu Thr Leu Gly Gln Glu Phe Ser Gly Tyr Thr Gln Gln Leu Thr Tyr
 225 230 235 240

Gly Ile Ala Arg Val Gln Gly Thr Leu Glu Arg Leu Tyr Asn Leu Ala
 245 250 255

Gln Gly Gly Thr Ala Val Gly Thr Gly Leu Asn Thr Arg Lys Gly Phe
 260 265 270

Asp Ala Lys Val Ala Glu Ala Ile Ala Ser Ile Thr Gly Leu Pro Phe
 275 280 285

Lys Thr Ala Pro Asn Lys Phe Glu Ala Leu Ala Ala His Asp Ala Leu
 290 295 300

Val Glu Ala His Gly Ala Leu Asn Thr Val Ala Cys Ser Leu Met Lys
 305 310 315 320

Ile Ala Asn Asp Ile Arg Tyr Leu Gly Ser Gly Pro Arg Cys Gly Leu
 325 330 335

Gly Glu Leu Ser Leu Pro Glu Asn Glu Pro Gly Ser Ser Ile Met Pro
 340 345 350

Gly Lys Val Asn Pro Thr Gln Cys Glu Ala Met Thr Met Val Cys Ala
 355 360 365

Gln Val Met Gly Asn Asn Thr Ala Ile Ser Val Ala Gly Ser Asn Gly
 370 375 380

Gln Phe Glu Leu Asn Val Phe Lys Pro Val Met Ile Lys Asn Leu Ile
 385 390 395 400

Gln Ser Ile Arg Leu Ile Ser Asp Ala Ser Ile Ser Phe Thr Lys Asn
 405 410 415

Cys Val Val Gly Ile Glu Ala Asn Glu Lys Lys Ile Ser Ser Ile Met
 420 425 430

Asn Glu Ser Leu Met Leu Val Thr Ala Leu Asn Pro His Ile Gly Tyr
 435 440 445

Asp Lys Ala Ala Lys Cys Ala Lys Lys Ala His Lys Glu Gly Thr Thr
 450 455 460

Leu Lys Glu Ala Ala Leu Ser Leu Gly Tyr Leu Thr Ser Glu Glu Phe
 465 470 475 480

Asp Gln Trp Val Arg Pro Glu Asp Met Ile Ser Ala Lys Asp
 485 490

<210> 2
 <211> 1419
 <212> DNA
 <213> *Rhizopus oryzae*

<400> 2
 atgagtagtg cctctgctgc ttgcaaaaa ttccgtgctg agcgcgatac ttttggtgat 60
 ctccaagttc ctgctgatag atattggggg gctcaaacc aaaggtctct tcaaaatttt 120
 gacattggtg gccctactga acgtatgcc gaacctttga tccgtgcctt tgggtgcctc 180
 aaaaaggctg ctgctactgt caacatgact tatggcttgg atcctaaagt tggatgaagct 240
 attcaaaagg ctgctgacga ggtcattgat ggaagcttga ttgatcattt ccctcttgtt 300
 gtctggcaaa ctggttccgg tactcaaacc aagatgaacg ttaacgaagt tatctccaac 360
 cgtgctattg aactttttggg tggtagactt ggtagtaagg ctctgttca tcccaacgat 420
 catgtcaaca tgagtcaatc atccaatgac acgttcccta ctgccatgca cgttgctgct 480

```

gttgttgaaa ttcacggtcg acttattcct gctttgacca ctttgcgtaga tgcccttcaa 540
gccaaatccg ctgagtttga acacatcatc aagatcggtc gtactcactt gcaagatgca 600
actcctttga ctctcgggtca agaattctct gggtatactc aacaattgac ttacgggtatt 660
gctcgtgtac aagggtacctt ggagcgcctc tataaccttg ctcaagggtg tactgctgtt 720
ggtactggtc ttaacaccag aaaagggttc gatgccaagg tagctgaagc tattgcttct 780
attaccggtc ttcctttcaa gaccgcccct aataagtttg aagcccttgc tgctcacgat 840
gctctcgttg aagctcacgg agctctcaat accgttgctt gttctcttat gaagatcgcc 900
aacgatatcc gttatcttgg ttctggacct cgctgtggtc ttggtgaact ttccttgcct 960
gaaaacgaac ccggtatctt tatcatgccc ggtaagggtta atcctactca atgtgaagct 1020
atgaccatgg tctgtgctca agtcatgggt aacaacactg ctatttctgt tgctggttcc 1080
aatggtcaat tcgagcttaa tgtcttcaaa cccgtcatga tcaagaactt gatccaatcc 1140
attcgtctta tttctgatgc ctctatttca ttcacaaaaa actgtgttgt tggtattgaa 1200
gccaatgaaa agaagattag cagcattatg aatgagtcac tgatgttggt cactgctctt 1260
aaccctcata ttggttacga taaagctgct aaatgtgcca agaaggcca caaggaaggc 1320
accaccttga aggaagctgc cctttctctt gggtacttga cttctgaaga attcgaccag 1380
tggttagac ccgaagatat gatctctgcc aaggattaa 1419

```

<210> 3

<211> 472

<212> PRT

<213> Artificial sequence

<220>

<223> Fumarase R. oryzae minus putative targeting signal

<400> 3

```

Met Ser Ser Ala Ser Ala Ala Leu Gln Lys Phe Arg Ala Glu Arg Asp
1           5           10           15

```

```

Thr Phe Gly Asp Leu Gln Val Pro Ala Asp Arg Tyr Trp Gly Ala Gln
          20           25           30

```

```

Thr Gln Arg Ser Leu Gln Asn Phe Asp Ile Gly Gly Pro Thr Glu Arg
          35           40           45

```

```

Met Pro Glu Pro Leu Ile Arg Ala Phe Gly Val Leu Lys Lys Ala Ala

```

50						55						60			
Ala	Thr	Val	Asn	Met	Thr	Tyr	Gly	Leu	Asp	Pro	Lys	Val	Gly	Glu	Ala
65					70					75					80
Ile	Gln	Lys	Ala	Ala	Asp	Glu	Val	Ile	Asp	Gly	Ser	Leu	Ile	Asp	His
				85					90					95	
Phe	Pro	Leu	Val	Val	Trp	Gln	Thr	Gly	Ser	Gly	Thr	Gln	Thr	Lys	Met
			100					105					110		
Asn	Val	Asn	Glu	Val	Ile	Ser	Asn	Arg	Ala	Ile	Glu	Leu	Leu	Gly	Gly
		115					120					125			
Glu	Leu	Gly	Ser	Lys	Ala	Pro	Val	His	Pro	Asn	Asp	His	Val	Asn	Met
	130					135					140				
Ser	Gln	Ser	Ser	Asn	Asp	Thr	Phe	Pro	Thr	Ala	Met	His	Val	Ala	Ala
145					150					155					160
Val	Val	Glu	Ile	His	Gly	Arg	Leu	Ile	Pro	Ala	Leu	Thr	Thr	Leu	Arg
				165					170					175	
Asp	Ala	Leu	Gln	Ala	Lys	Ser	Ala	Glu	Phe	Glu	His	Ile	Ile	Lys	Ile
			180					185					190		
Gly	Arg	Thr	His	Leu	Gln	Asp	Ala	Thr	Pro	Leu	Thr	Leu	Gly	Gln	Glu
		195					200					205			
Phe	Ser	Gly	Tyr	Thr	Gln	Gln	Leu	Thr	Tyr	Gly	Ile	Ala	Arg	Val	Gln
	210					215					220				
Gly	Thr	Leu	Glu	Arg	Leu	Tyr	Asn	Leu	Ala	Gln	Gly	Gly	Thr	Ala	Val
225					230					235					240
Gly	Thr	Gly	Leu	Asn	Thr	Arg	Lys	Gly	Phe	Asp	Ala	Lys	Val	Ala	Glu
				245					250					255	
Ala	Ile	Ala	Ser	Ile	Thr	Gly	Leu	Pro	Phe	Lys	Thr	Ala	Pro	Asn	Lys
			260					265					270		
Phe	Glu	Ala	Leu	Ala	Ala	His	Asp	Ala	Leu	Val	Glu	Ala	His	Gly	Ala
		275					280					285			

Leu Asn Thr Val Ala Cys Ser Leu Met Lys Ile Ala Asn Asp Ile Arg
 290 295 300

Tyr Leu Gly Ser Gly Pro Arg Cys Gly Leu Gly Glu Leu Ser Leu Pro
 305 310 315 320

Glu Asn Glu Pro Gly Ser Ser Ile Met Pro Gly Lys Val Asn Pro Thr
 325 330 335

Gln Cys Glu Ala Met Thr Met Val Cys Ala Gln Val Met Gly Asn Asn
 340 345 350

Thr Ala Ile Ser Val Ala Gly Ser Asn Gly Gln Phe Glu Leu Asn Val
 355 360 365

Phe Lys Pro Val Met Ile Lys Asn Leu Ile Gln Ser Ile Arg Leu Ile
 370 375 380

Ser Asp Ala Ser Ile Ser Phe Thr Lys Asn Cys Val Val Gly Ile Glu
 385 390 395 400

Ala Asn Glu Lys Lys Ile Ser Ser Ile Met Asn Glu Ser Leu Met Leu
 405 410 415

Val Thr Ala Leu Asn Pro His Ile Gly Tyr Asp Lys Ala Ala Lys Cys
 420 425 430

Ala Lys Lys Ala His Lys Glu Gly Thr Thr Leu Lys Glu Ala Ala Leu
 435 440 445

Ser Leu Gly Tyr Leu Thr Ser Glu Glu Phe Asp Gln Trp Val Arg Pro
 450 455 460

Glu Asp Met Ile Ser Ala Lys Asp
 465 470

<210> 4

<211> 1419

<212> DNA

<213> Artificial sequence

<220>

<223> nt FumR minus putative targeting codon optimised for *S.cerevisae*

<400> 4
atgtcctctg cttctgctgc tttgcaaaaa ttcagagctg aaagagatac cttcggtgac 60
ttgcaagttc cagctgaccg ttactggggg gctcaaactc aaagatcttt gcaaaacttt 120
gacattggtg gtccaactga aagaatgcc gaaccattaa tcagagcttt cgggtgttttg 180
aagaaggctg ctgccaccgt caacatgacc tacggtttgg acccaaagg tggatgaagcc 240
atccaaaagg ctgctgacga agttatcgat ggttctttga ttgaccattt cccattggtt 300
gtctggcaaa ccggttctgg tactcaaacc aagatgaacg tcaatgaagt catctccaac 360
agagccattg aattgttggg tggatgaatta ggttccaagg ctccagtcca cccaaacgat 420
catgtcaaca tgtctcaatc ttccaacgac actttcccaa ctgccatgca cgttgctgcc 480
gttgttgaaa ttcacggtag attgattcca gctttgacca ctttgagaga tgctttgcaa 540
gccaaatctg ctgaattcga acacatcatc aagattggta gaaccactt gcaagatgct 600
acccattga ctttaggtca agaattctcc ggttacactc aacaattgac ctacggtatt 660
gctcgtgttc aaggactttt ggaaagatta tacaacttgg ctcaagggtg tactgctgtc 720
ggtactggtt tgaacaccag aaagggtttc gatgccaaagg ttgctgaagc cattgcttcc 780
atcactggtt taccattcaa gaccgctcca aacaaattcg aagctttggc tgctcacgac 840
gctttggttg aagctcacgg tgctttgaac accggttgctt gttctttgat gaagattgcc 900
aacgatatcc gttacttggg ttctggtcca agatgtgggt taggtgaatt gtctctacca 960
gaaaacgaac caggttcttc catcatgcca ggtaagggtca acccaactca atgtgaagct 1020
atgaccatgg tttgtgctca agtcatgggt aacaacactg ccatctctgt tgctggttcc 1080
aacggtcaat tcgaattgaa tgtcttttaa ccagtcatga tcaagaactt gatccaatcc 1140
atcagattaa tctctgacgc ttccatctct ttcaccaaga actgtgttgt cggtattgaa 1200
gctaacgaaa agaagatctc ctccatcatg aacgaatctt tgatgttggc cactgctttg 1260
aacctcaca ttggttacga caaggctgcc aagtgtgcca agaaggctca caaggaagg 1320
accactttga aagaagctgc tctatctttg ggttacttga cctctgaaga attcgaccaa 1380
tggttagac ctgaggacat gatctctgcc aaggattaa 1419

<210> 5
<211> 1000
<212> DNA
<213> Artificial sequence

<220>

<223> TDH1 promotor

<400> 5

```

cttccctttt acagtgttc ggaaaagcac agcgttgtcc aagggaacaa tttttcttca      60
agttaatgca taagaaatat ctttttttat gtttagctaa gtaaaagcag cttggagtaa     120
aaaaaaaaat gagtaaattt ctcgatggat tagttttctca caggtaacat aacaaaaacc     180
aagaaaagcc cgcttctgaa aactacagtt gacttgtatg ctaaagggcc agactaatgg     240
gaggagaaaa agaaacgaat gtatatgctc atttacactc tatatcacca tatggaggat     300
aagttgggct gagcttctga tccaatttat tctatccatt agttgctgat atgtcccacc     360
agccaacact tgatagtatc tactcgccat tcacttccag cagcgccagt agggttgttg     420
agcttagtaa aaatgtgctc accacaagcc tacatgactc cacgtcacat gaaaccacac     480
cgtgggggct tgttgcgcta ggaataggat atgcgacgaa gacgcttctg cttagtaacc     540
acaccacatt ttcagggggg cgatctgctt gcttccttta ctgtcacgag cggcccataa     600
tcgcgctttt tttttaaaag gcgcgagaca gcaaacagga agctcggggt tcaaccttcg     660
gagtggctgc agatctggag actggatctt tacaatacag taaggcaagc caccatctgc     720
ttcttaggtg catgcgacgg tatccacgtg cagaacaaca tagtctgaag aaggggggga     780
ggagcatgtt cattctctgt agcagtaaga gcttggtgat aatgaccaa actggagtct     840
cgaaatcata taaatagaca atatattttc acacaatgag atttgtagta cagttctatt     900
ctctctcttg cataaataag aaattcatca agaacttggg ttgatatttc accaacacac     960
acaaaaaaca gtacttcact aaatttacac acaaaacaaa                        1000

```

<210> 6

<211> 500

<212> DNA

<213> Artificial sequence

<220>

<223> TDH1 terminator

<400> 6

```

ataaagcaat cttgatgagg ataatgattt ttttttgaat atacataaat actaccgttt      60
ttctgctaga ttttgtgaag acgtaaataa gtacatatta ctttttaagc caagacaaga     120
ttaagcatta actttaccct tttctcttct aagtttcaat actagttatc actgtttaaa     180
agttatggcg agaacgtcgg cggttaaaat atattaccct gaacgtgggtg aattgaagtt     240
ctaggatggg ttaaagattt ttcttttttg ggaaataagt aaacaatata ttgctgcctt     300

```


tgcaaaacgc acataccac aatatgtgac tattggcaaa gaacgcatta tcctttgaag	360
aggtggatac tgatactaag agagtctcta ttccggtcc acttttagtc cagagattac	420
ttgtcttctt acgtatcaga acaagaaagc atttccaaag taattgcatt tgcccttgag	480
cagtatatat atactaagaa	500