(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0120236
(43) 공개일자 2015년10월27일

(51) 국제특허분류(Int. Cl.)
C12N 15/70 (2006.01) *C12N 1/21* (2006.01)
C12P 7/44 (2006.01)
(21) 출원번호 10-2014-0046251
(22) 출원일자 2014년04월17일
심사청구일자 없음

(71) 출원인
(주)에스티알바이오텍
강원도 춘천시 소양강로 56, 바이오벤처프라자 3-8호 (후평동)
(72) 발명자
박선옥
서울특별시 마포구 토정로32길 11 한강삼성아파트 103동 1002호
김상용
충청남도 천안시 서북구 시청로 73 동일하이빌아파트 208동 804호
(뒷면에 계속)
(74) 대리인
김영철, 김민정, 김 순 영

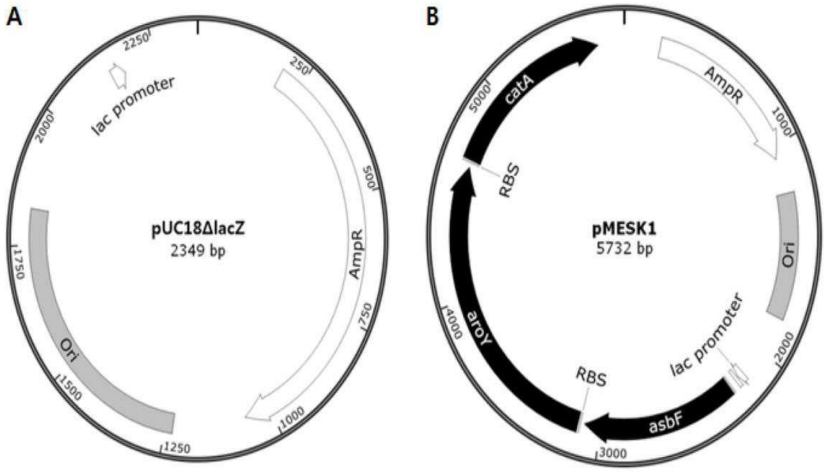
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 **뮤코닉산 생산용 미생물 및 그 제조방법**

(57) 요 약

본 발명에 의한 재조합 벡터로 형질전환된 대장균은 높은 뮤코닉산 생산능을 보유함으로써, 바이오 유래의 TPA 또는 아디프산의 전구체인 뮤코닉산을 효과적으로 생산할 수 있다. 구체적으로, 본 발명에 의한 미생물은 대장균이 뮤코닉산을 생산하는데 필요한 세 외래유전자인 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드를 각각 하나의 오페론(operon) 형태로 제조한 재조합 벡터를 대장균에 도입함으로써, 상기 외래유전자를 각각 동일한 비율로 전사할 수 있다.

대 표 도 - 도2



(72) 발명자

전계택

강원도 춘천시 공지로 361 공무원아파트 A동 606호

김응수

서울특별시 강남구 언주로30길 26 타워팰리스 G동 5006호

최시선

인천광역시 부평구 부흥로 246 동아아파트 36동 407호

송지수

서울특별시 도봉구 시루봉로6길 87 유송맨션 A동 401호

명 세 서

청구범위

청구항 1

뮤코닉산 생산용 재조합 벡터로서,

asbF^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드를 포함하고,

상기 세 폴리뉴클레오티드의 각 업스트림(upstream)에 대장균 유래의 리보솜 결합 부위(ribosome binding site, rbs)를 포함하고, 상기 세 폴리뉴클레오티드 중 첫 번째 전사되는 폴리뉴클레오티드의 리보솜 결합부위(rbs)의 업스트림(upstream)에 프로모터(promoter)를 포함하는 뮤코닉산 생산용 재조합 벡터.

청구항 2

제 1 항에 있어서, 상기 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드는 서열번호 1을 포함하는 뮤코닉산 생산용 재조합 벡터.

청구항 3

제 1 항에 있어서, 상기 *aroY*^{opt}를 코딩하는 폴리뉴클레오티드는 서열번호 2를 포함하는 뮤코닉산 생산용 재조합 벡터.

청구항 4

제 1 항에 있어서, 상기 *catA*^{opt}를 코딩하는 폴리뉴클레오티드는 서열번호 3을 포함하는 뮤코닉산 생산용 재조합 벡터.

청구항 5

제 1 항에 있어서, 상기 리보솜 결합 부위는 서열번호 4를 포함하는 뮤코닉산 생산용 재조합 벡터.

청구항 6

제 1 항 내지 제 5 항 중 어느 하나의 항에 따른 재조합 벡터로 형질전환된 숙주 세포.

청구항 7

제 6 항에 따른 숙주 세포를 포함하는 대장균.

청구항 8

제 7 항에 있어서, 상기 대장균은 *aroE* 를 코딩하는 폴리뉴클레오티드(서열번호 5)가 결손된 대장균.

청구항 9

제 7 항에 있어서, 상기 대장균은 상기 $asbF^{opt}$, $aroY^{opt}$ 및 $catA^{opt}$ 을 코딩하는 각 폴리뉴클레오티드의 전사량이 0.9~1.1 : 0.9~1.1 : 0.9~1.1인 대장균.

청구항 10

뮤코닉산 생산용 대장균을 제조하는 방법으로,

벡터에 $asbF^{opt}$를 코딩하는 폴리뉴클레오티드, $aroY^{opt}$를 코딩하는 폴리뉴클레오티드 및 $catA^{opt}$를 코딩하는 폴리뉴클레오티드를 삽입하고, 상기 세 폴리뉴클레오티드의 업스트림(upstream) 방향으로 대장균 유래의 리보솜 결합 부위(ribosome binding site, rbs)를 각각 삽입하여 재조합 벡터를 제조하는 단계; 및

상기 재조합 벡터를 대장균 모미생물에 도입하여 형질전환하는 단계;

를 포함하는 뮤코닉산 생산용 대장균의 제조방법.

청구항 11

제 10 항에 있어서, 상기 대장균 모미생물은 상기 대장균은 *aroE* 를 코딩하는 폴리뉴클레오티드(서열번호 5)가 결손된 대장균인 뮤코닉산 생산용 대장균의 제조방법.

청구항 12

제 10 항에 있어서, 상기 벡터는 lac 프로모터 유래 lacZα 유전자가 제거된 벡터인 뮤코닉산 생산용 대장균의 제조방법.

청구항 13

제 12 항에 있어서, 상기 벡터에서 제거된 lac 프로모터 유래 lacZα 유전자는 서열번호 8을 포함하는 뮤코닉산 생산용 대장균의 제조방법.

청구항 14

제 10 항에 있어서, 상기 재조합 벡터를 제조하는 단계는 상기 세 폴리뉴클레오티드 중 첫 번째로 전사되는 폴리뉴클레오티드의 리보솜 결합부위(rbs)의 업스트림(upstream)에 프로모터(promoter)를 삽입하는 것을 더 포함하는 뮤코닉산 생산용 대장균의 제조방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 뮤코닉산 생산능을 갖는 신규한 재조합 벡터, 이를 포함하는 미생물 및 그 제조방법에 관한 것이다.

배 경 기 술

[0002] 에너지 및 화학물질의 원재료로 사용되는 화석연료는 지속적인 유가상승 및 자원고갈로 인하여 지구온난화를 포함한 각종 환경-경제적 문제에 직면하고 있으며, 이를 보완하기 위한 친환경적이고 재생 가능한 대체물질을 찾는 노력이 계속되고 있다. 그 중에서도 미생물을 이용한 생합성 전략이 활발하게 연구되고 있으며, 특히 특정 대사경로가 유전공학적으로 재설계된 재조합미생물을 이용한 다양한 화학물질, 바이오 연료, 아미노산 및 식물체의 2차 대사산물 등을 생합성 하는 것이 가능해졌다.

[0003] 현재 세계 고분자시장은 바이오 유래의 환경친화적 원료의 사용을 요구받고 있으며 이에 따라 코카콜라, 펩시와 같은 거대 음료회사들이 100% 식물 유래의 PET(폴리에틸렌테레프탈레이트) 개발에 총력을 기울이고 있다. 한편,

PET와 유사한 구조를 가지면서도 독특한 특성을 갖는 PTT(폴리트리메틸렌테레프탈레이트) 역시 잠재력 높은 시장규모를 가질 뿐만 아니라, 연평균 17%의 고성장을 진행 중에 있다. 따라서 PET와 같이 PTT 역시 바이오 유래의 환경친화적 고분자합성이 필요한 시점이다. 현재 PTT는 석유 유래의 TPA(terephthalic acid)와 바이오 유래의 1,3-PDO(1,3-프로판디올)와의 축합반응에 의한 합성을 통해 생산이 이루어지고 있다. 한편, 바이오 유래의 1,3-PDO에 대한 연구가 활발히 진행된 것에 비해 바이오 유래의 TPA 연구는 전 세계적으로 매우 미비한 형편이다. 특히 대부분의 TPA(terephthalic acid)에 관한 연구가 석유화학물질을 출발점으로 하고 있기 때문에 바이오 유래의 친환경적인 TPA 생산공정의 개발 필요성이 대두되고 있다. 그러나 생물공정을 이용하여 바이오 유래의 탄소원으로부터 직접적으로 TPA를 생합성하는 공정은 아직까지 개발되지 못하고 있다. 따라서 환경친화적인 TPA 생산을 위해 먼저 생물공정을 통해 바이오 유래의 TPA 전구체를 생산하고, 생물공정을 통해 생산된 TPA 전구체를 화학합성공정을 통하여 TPA로 전환시키는 융합생산공정의 개발이 필요하다.

[0004] 뮤코닉산은 TPA 합성을 위한 바이오 유래의 전구체로서의 용도 외에도 아디프산의 전구체로도 잘 알려져 있다. 아디프산(adipic acid)은 나일론, 윤활유, 플라스틱, 가소제 등의 다양한 물질의 전구체 역할을 함으로써 전 세계적으로 연간 2 X 10^9 kg정도의 많은 양이 소비되고 있다. 하지만 현재 아디프산이 합성될 때 사용되는 벤젠 유래의 사이클로헥산은 산화되는 과정에서 지구온난화를 야기하는 N_2O를 발생시키는 문제를 가지고 있다. 또한 각 단계별 다양한 중간화학물질은 인체에 유해할 뿐만 아니라 발암물질로도 알려져 있다. 이러한 문제를 해결하기 위해 식물 유래의 재생 가능한 원료를 이용하여 아디프산을 만드는 기술, 즉 바이오 유래의 뮤코닉산의 수소화 반응을 통한 아디프산 생산공정의 개발이 필요하다.

선행기술문헌

비특허문헌

[0005] (비특허문헌 0001) Alper H, Miyaoku K, Stephanopoulos G. (2005) Construction of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets. Nature Biotechnology, 23, 612-616.

(비특허문헌 0002) Atsumi S, Hanai T, Liao JC. (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. Nature, 451, 86-89.

(비특허문헌 0003) Azizkhan JC, Jensen DE, Pierce AJ, Wade M. (1993) Transcription from TATA-less promoters: Dihydrofolatereductase as a model. Critical Reviews in Eukaryotic Gene Expression, 3, 229-54.

(비특허문헌 0004) Berry, A. (1996) Improving production of aromatic compounds in Escherichia coli by metabolic engineering. Trends in Biotechnology, 14(7), 250-256.

(비특허문헌 0005) Branlant C, Oster T, Branlant G. (1989) Nucleotide sequence determination of the DNA region coding for Bacillus stearothermophilus glyceraldehyde-3-phosphate dehydrogenase and of the flanking DNA regions required for its expression in Escherichia coil. Gene, 75, 145-155.

(비특허문헌 0006) Conway T, Ingram LO. (1988) Phosphoglycerate kinase gene from Zymomonasmobilis: Cloning, sequencing and localization within the gap operon. Journal of Bacteriology, 170, 1926-1933.

(비특허문헌 0007) Draths KM, Frost JW. (1994) Environmentally compatible synthesis of adipic acid from D-glucose. Journal of the American Chemical Society, 116, 399-400.

(비특허문헌 0008) Fox DT, Hotta K, Kim CY, Koppisch AT. (2008) The missing link in petrobactin biosynthesis: AsbF encodes a ()-3-dehydroshikimate dehydratase. Biochemistry, 47(47), 12251-12253.

(비특허문헌 0009) Frost JW, Draths. KM. (1995) Biocatalytic syntheses of aromatics from D-glucose: Renewable microbial sources of aromatic compounds. Annual Review of Microbiology, 49, 557-579.

(비특허문헌 0010) Harris JI,Waters M. (1976) Glyceraldehyde-3-phosphate dehydrogenase. The Enzyme, 13, 1-49.

(비특허문헌 0011) Lin Y, Yan Y. (2012) Biosynthesis of caffeic acid inEscherichia coli using its

endogenous hydroxylase complex. Microbial Cell Factories, 11, 42.

(비특허문헌 0012) Lin YL, Blaschek HP. (1983) Butanol production by a butanol-tolerant strain of Clostridium acetobutylicum in extruded corn broth. Applied and Environmental Microbiology, 45, 966-973.

(비특허문헌 0013) Miller JH. (1972) Experiments in molecular genetics; Cold spring harbor laboratory. Planview.

(비특허문헌 0014) NiuW, DrathsKM,Frost JW. (2002) Benzene-free synthesis of adipic acid. Biotechnology Progress, 18, 201-211.

(비특허문헌 0015) Pfleger BF, Kim Y, Nusca TD, Maltseva N, Lee JY, Rath CM, Scaglione JB, Janes BK, Anderson EC, Bergman NH, Hanna PC, Joachimiak A, Sherman DH.(2008) Structural and functional analysis of AsbF: Origin of the stealth 3,4-dihydroxybenzoic acid subunit for petrobactin biosynthesis.

(비특허문헌 0016) Proceedings of the National Academy of Sciences, 105(44), 17133-17138.

(비특허문헌 0017) Pittard AJ. (1996) Biosynthesis of the aromatic amino acids.in Escherichia coli and Salmonella. InNeidhart FC, Curtiss R, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE (ed.). vol. 1.American Society for Microbiology.Washington, D.C.pp. 458-484.

(비특허문헌 0018) Sprenger GA. (2007) From scratch to value: Engineering Escherichia coli wild type cells to the production of L-phenylalanine and other fine chemicals derived from chorismate. Applied Microbiology and Biotechnology, 75 (4), 739-749.

(비특허문헌 0019) Thiemens MH, Trogler WC. (1991) Nylon production: An unknown source of atmospheric nitrous oxide. Science, 251, 932-934.

(비특허문헌 0020) Ulrich H. (1988) Raw materials for industrial polymers. Germany: Hanser Publishers, Munich.

(비특허문헌 0021) Warhurst AM, Clarke KF, Hill RA,Holt RA, Fewson CA (1994) Production of catechols and muconic acids from various aromatics by the styrene-degrader Rhodococcusrhodochrous NCIMB 13259.Biotechnology Letters, 16, 513-516.

(비특허문헌 0022) Weber C, BrC, Weinreb S, Lehr C, Essl C, Boles E.(2012) Biosynthesis of cis,cis-muconic acid and its aromatic precursors, catechol and protocatechuic acid, from renewable feedstocks by Saccharomyces cerevisiae. Applied and Environmental Microbiology, 78(23), 8421.

(비특허문헌 0023) Wu CM, Lee TH, Lee SN Lee YA, Wu JY (2004) Microbial synthesis of cis,cis-muconic acid by Sphingobacterium sp. GCG generated from effluent of a styrene monomer (SM) production plant. Enzyme and Microbial Technology, 35, 598-604.

(비특허문헌 0024) Yan Y, Chemler J, Huang L, Martens S, Koffas MA. (2005) Metabolic engineering of anthocyanin biosynthesis inEscherichia coli. Applied and Environmental Microbiology, 71, 3617-3623.

발명의 내용

해결하려는 과제

[0006] 본 발명은 높은 뮤코닉산 생산능을 보유한 대장균 및 그 제조방법을 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0007] 상기 목적을 달성하기 위하여, 본 발명의 일 구현예는 뮤코닉산 생산용 재조합 벡터로서, *asbF*^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드를 포함하고, 상

기 세 폴리뉴클레오티드의 각 업스트림(upstream)에 대장균 유래의 리보솜 결합 부위(ribosome binding site, rbs)를 포함하고, 상기 세 폴리뉴클레오티드 중 첫 번째 전사되는 폴리뉴클레오티드의 리보솜 결합부위(rbs)의 업스트림(upstream)에 프로모터(promoter)를 포함하는 뮤코닉산 생산용 재조합 벡터를 제공한다.

[0008] 또한, 본 발명의 일 구현예는 뮤코닉산 생산용 대장균을 제조하는 방법으로, 벡터에 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드를 삽입하고, 상기 세 폴리뉴클레오티드의 업스트림(upstream) 방향으로 대장균 유래의 리보솜 결합 부위(ribosome binding site, rbs)를 각각 삽입하여 재조합 벡터를 제조하는 단계, 및 상기 재조합 벡터를 대장균 모미생물에 도입하여 형질전환하는 단계를 포함하는 뮤코닉산 생산용 대장균의 제조방법을 제공한다.

발명의 효과

[0009] 본 발명에 의한 재조합 벡터로 형질전환된 대장균은 높은 뮤코닉산 생산능을 보유함으로써, 바이오 유래의 TPA 또는 아디프산의 전구체인 뮤코닉산을 효과적으로 생산할 수 있다. 또한, 본 발명에 의한 미생물은 대장균이 뮤코닉산을 생산하는데 필요한 외래유전자를 각각 하나의 오페론(operon) 형태로 제작하여 대장균에 도입함으로써, 상기 외래 유전자가 각각 동일한 비율로 전사할 수 있다. 따라서, 제조된 미생물 자체의 구현여부를 하나의 유전자에 대한 구현여부로 판단할 수 있으므로 제조한 미생물의 성공여부를 용이하게 확인할 수 있다.

도면의 간단한 설명

[0010] 도 1은 대장균내 방향족 아미노산 및 시스,시스-뮤코닉산의 생합성 경로 및 조절경로를 나타낸 것으로, 파선은 피드백 억압(repression)을 의미한다. 도 1에 기재된 물질명 E4P(erythrose-4-phosphate), PEP(phosphoenolpyruvate), DAHP(3-deoxy-d-arabinoheptulosonate-7-phosphate), DHQ(3-dehydroquinic acid), DHS(3-dehydroshikimic acid), SA(shikimic acid), PCA(protocatechuate)와 그리고 유전자명 *aroF*(DAHP synthase, l-tyr), *aroG*(DAHP synthase, l-phe), *aroH*(DAHP synthase, l-trp), *aroB*(DHQ synthase), *aroD*(DHQ dehydratase), *aroE*(shikimate dehydrogenase), *asbF*^{opt}(dehydroshikimate dehydratase), *aroY*^{opt}(protocatechuate decarboxylase) 및 *catA*^{opt}(catechol 1,2-dioxygenase)은 *E. coli* K-12 연결지도에 따른 것이다.

도 2는 재조합 플라스미드 지도로서 (A) pUC18△lacZ 및 (B) pMESK1를 나타낸 것이다.

도 3은 (A) LB medium (B) M9 medium의 서로 다른 배지에서, AB2834/pMESK1(실시예) 및 AB2834/pUC18△lacZ(비교예)의 배양시간에 대한 시스,시스-뮤코닉산(Muconic acid)과 세포성장(Cell growth)을 측정한 결과를 나타낸 그래프이다. (◆: AB2834/pMESK1의 시스,시스-뮤코닉산 생산량, ○: AB2834/pUC18△lacZ의 세포성장, ■: AB2834/pMESK1의 세포성장)

도 4는 AB2834/pMESK1(실시예)의 시스,시스-뮤코닉산(◆), 3-DHS(□), cathecol(△) 및 PCA(○) 생산량을 나타낸 그래프이다.

도 5는 AB2834/pMESK1(실시예)의 RT-PCR 분석결과를 나타낸 것이다(M: 100bp ladder, 1: *aroB*, 2: *aroD*, 3: *aroF*, 4: *aroH*, 5: *aroG*, 6: *asbF*^{opt}, 7: *aroY*^{opt}, 8: *catA*^{opt}, 9: GAPDH).

발명을 실시하기 위한 구체적인 내용

[0011] 이하, 본 발명을 상세히 설명한다.

[0012] 뮤코닉산(muconic acid)은 바이오 유래의 TPA(terephthalic acid) 또는 아디프산의 전구체로서, 방향족아미노산 생합성 경로의 중간물질을 통해 합성될 수 있다. 이에 본 발명은 뮤코닉산을 생산할 수 있는 미생물을 제공하기 위하여, 미생물 중 대장균(*Escherichia coli*)에 뮤코닉산을 합성할 수 있는 유전자들을 도입하였다.

[0013] 이에 본 발명의 구현예들은 뮤코닉산 생산용 재조합 벡터로서, *asbF*^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드를 포함하고, 상기 세 폴리뉴클레오티드의 각 업스트림(upstream)에 대장균 유래의 리보솜 결합 부위(ribosome binding site, rbs)를 포함하고, 상기 세 폴리뉴클레오티드 중 첫 번째 전사되는 폴리뉴클레오티드의 리보솜 결합부위(rbs)의 업스트림(upstream)에 프로

모터(promoter)를 포함하는 뮤코닉산 생산용 재조합 벡터를 제공한다. 일 구현예로서 상기 세 폴리뉴클레오티드 중 첫 번째 전사되는 폴리뉴클레오티드는 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드 중에서 선택되는 어느 하나이며, 상기 세 유전자의 배열 순서는 상기 유전자가 기재된 순서로 한정되지 않는다. 예를 들면, 세 유전자는 뮤코닉산 생산용 재조합 벡터내에 *aroY*^{opt}를 코딩하는 폴리뉴클레오티드, *asbF*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드 순으로 배열될 수 있다. 또는, *catA*^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드 순으로 배열될 수 있다.

[0014] 구체적으로, 본 발명의 일 구현예로서 사용되는 대장균(*Escherichia coli*)은 그람 음성 호기성 세균으로 독성이 없으며, 성장이 매우 빠르고 실험실과 산업현장에서 쉽고 저렴하게 키울 수 있다는 장점을 가지고 있다. 그러나 대장균 내에는 뮤코닉산을 합성할 수 있는 유전자들이 존재하지 않기 때문에, 뮤코닉산을 생산하기 위하여는 외래유전자의 도입이 필요하다. 이에 본 발명은 대장균 대사경로 중 방향족아미노산 생합성 경로에서 3-DHS(3-dehydroshikimate)가 합성되는 것에 주목하여, 3-DHS로부터 뮤코닉산을 생합성 하는데 있어서 필요한 유전자로서, *asbF, aroY* 및 *catA*를 각각 코딩하는 유전자를 도입하였다. 일 구현예로서, 상기 뮤코닉산은 시스-시스 뮤코닉산(cis,cis-muconic acid, CCM)이며, 상기 3-DHS에서 시스-시스 뮤코닉산으로의 대사경로는 도 1에 나타내었다. 도 1에 나타난 바와 같이, 대장균의 탄수화물 대사경로 중에는 방향족아미노산인 페닐알라닌(Phenylalanine,Phe), 타이로신(Tyrosine,Tyr), 트립토판(Tryptophan,Try)을 생합성하는 경로가 있다. 이 때 중간물질인 3-DHS (3-dehydroshikimate)가 앞서 언급한 바와 같이 뮤코닉산 전구체로서 사용될 수 있는데, 뮤코닉산 생산을 위해서는 대장균에 원래는 존재하지 않는 세 종류의 외래유전자를 도입해야만 가능하다. 즉, 3-DHS 탈수효소(3-DHS dehydratase)를 코딩하는 *asbF*, PCA 탈카르복실화효소(protocatechuic acid decarboxylase)를 코딩하는 *aroY*, 카테콜 1,2-산소화효소(catechol 1,2-dioxygenase)를 코딩하는 *catA*이다. 이러한 3종의 외래유전자의 도입으로 3-DHS로부터 PCA, 카테콜을 거쳐 뮤코닉산을 합성할 수 있게 된다.

[0015] 이에 본 발명의 구현예들은 상기 세 종류의 외래유전자를 대장균에 도입함에 있어서 대장균에서의 발현율을 증가시킴으로써 뮤코닉산 생산능을 향상시키기 위하여, 상기 *asbF, aroY* 및 *catA*를 대장균 코돈으로 변형한 염기서열인 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드를 합성한다. 그리고 상기 세 폴리뉴클레오티드와, 상기 세 폴리뉴클레오티드의 각 업스트림(upstream)에 대장균 유래의 리보솜 결합 부위(ribosome binding site)를 포함하고, 상기 세 폴리뉴클레오티드 중 첫 번째 전사되는 폴리뉴클레오티드의 리보솜 결합부위(rbs)의 업스트림에 프로모터(promoter)를 포함하는 재조합 벡터로 제조할 수 있다. 일 구현예로서, 상기 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드는 서열번호 1을 포함하고, 상기 *aroY*^{opt}를 코딩하는 폴리뉴클레오티드는 서열번호 2를 포함하며, 상기 *catA*^{opt}를 코딩하는 폴리뉴클레오티드는 서열번호 3을 포함할 수 있다. 상기 리보솜 결합 부위는 서열번호 4를 포함할 수 있다. 일 구현예는 상기와 같이 제조한 재조합 벡터로 형질전환된 숙주세포를 제공하며, 상기 숙주 세포를 포함하는 대장균을 제공한다.

[0016] 이로써 상기와 같은 *asbF*^{opt}, *aroY*^{opt} 및 *catA*^{opt}를 포함하는 본 발명의 일 구현예에 따른 대장균은 상기 세 외래유전자 각각을 단일 오페론 구조로 포함함으로써, 상기 *asbF*^{opt}, *aroY*^{opt} 및 *catA*^{opt} 을 코딩하는 각 폴리뉴클레오티드의 전사량이 0.9~1.1 : 0.9~1.1 : 0.9~1.1로 구현된다.

[0017] 상기 세 종류의 폴리뉴클레오티드는 각각의 해당 효소인 3-DHS 탈수효소(3-DHS dehydratase), PCA 탈카르복실화효소(protocatechuic acid decarboxylase), 카테콜 1,2-산소화효소(catechol 1,2-dioxygenase)를 효율적으로 발현시켜준다. 본 발명의 다른 일 구현예로서 상기 세 유전자 중 *asbF* 유전자는 *Klebsiella pneumonia, Acinetobacter sp, Podospore anserine* 또는 *Bacillus thuringiensis*유래의 *asbF*, 상기 *aroY* 유전자는 *Klebsiella pneumonia, Sedimentibacter hydroxybenzoicus* 또는 *Enterobacter cloacae*유래의 *aroY, catA* 유전자는 *Burkholderia xenovorans, Acinetobacter calcoaceticus* 또는 *Acinetobacter calcoaceticus* 유래의 *catA*를 사용할 수 있다.

[0018] 또한, 본 발명의 일 구현예로서 상기 대장균은 3-DHS의 축적을 위해 *aro E* 유전자(서열번호 5)가 결손되어, *aro E* 효소활성이 상실된 대장균이며, 예를 들면 *E.coli* AB2834(구입기관: *E. coli* Genetic Stock Center, 기탁번호: AB2834)를 사용할 수 있다. 상기 areE유전자는 DHS(3-dehydroshikimic acid)를 SA(shikimic acid)로 전환

시키는 3-DHS 탈수효소를 코딩하는 유전자로 aroE 유전자가 결손되면 DHS가 SA로 전환되는 것을 차단하여 DHS가 방향족아미노산 생합성에 이용되는 것을 차단함으로써 대사 흐름을 뮤코닉산 생합성 쪽으로 바꾸어 줄 수 있어 상기 areE유전자가 결손된 대장균을 사용하면 뮤코닉산의 생합성 효율을 더욱 증가시킬 수 있다. 또한 방향족 아미노산의 생합성이 차단되면 대사경로 초기물질인 PEP(phosphoenolpyruvate)와 E4P(erythrose-4-phosphate)가 DAHP로의 전환에 대한 feedback inhibition의 해제(저해)가 일어나 뮤코닉산 생합성으로의 대사흐름이 원활해진다. 다만 aroE 유전자의 결손으로 방향족아미노산이 합성되지 않으므로 배양배지에 방향족아미노산을 첨가해 주어야 하는데 이때 과량의 방향족아미노산을 첨가해 주면 다시 피드백 억제(feedback inhibition)가 일어나므로 세심한 주의가 필요하다.

[0019] 본 발명의 다른 일 구현예는 상기와 같은 뮤코닉산 생산용 대장균을 제조하는 방법으로, 벡터에 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드를 삽입하고, 상기 세 폴리뉴클레오티드의 업스트림(upstream) 방향으로 대장균 유래의 리보솜 결합 부위(ribosome binding site, RBS)를 각각 삽입하여 재조합 벡터를 제조하는 단계, 및 상기 재조합 벡터를 대장균 모균주에 도입하여 형질전환하는 단계를 포함하는 뮤코닉산 생산용 재조합 대장균의 제조방법을 제공한다. 또한, 일 구현예로서 상기 재조합 벡터를 제조하는 단계는 상기 세 폴리뉴클레오티드 중 첫 번째로 전사되는 폴리뉴클레오티드의 리보솜 결합부위(rbs)의 업스트림(upstream)에 프로모터(promoter)를 삽입하는 것을 더 포함할 수 있다.

[0020] 일 구현예로서 상기 세 폴리뉴클레오티드 중 첫 번째 전사되는 폴리뉴클레오티드는 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드 중에서 선택되는 어느 하나이며, 상기 세 폴리뉴클레오티드의 삽입 및 배열 순서는 상기 유전자가 기재된 순서로 한정되지 않는다. 예를 들면, 세 유전자는 뮤코닉산 생산용 재조합 벡터 내에 *aroY*^{opt}를 코딩하는 폴리뉴클레오티드, *asbF*^{opt}를 코딩하는 폴리뉴클레오티드 및 *catA*^{opt}를 코딩하는 폴리뉴클레오티드 순으로 배열될 수 있다. 또는, *catA*^{opt}를 코딩하는 폴리뉴클레오티드, *aroY*^{opt}를 코딩하는 폴리뉴클레오티드 및 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드 순으로 배열될 수 있다.

[0021] 상기와 같이 *asbF*^{opt}, *aroY*^{opt} 및 *catA*^{opt} 를 각각 코딩하는 폴리뉴클레오티드의 업스트림 방향으로 대장균 유래의 리보솜 결합 부위(RBS)를 추가하며, 상기 외래유전자 각각이 단일 오페론 형태로 제조되어 각 유전자의 전사량을 0.9~1.1 : 0.9~1.1 : 0.9~1.1로 동일하게 구현할 수 있다.

[0022] 일 구현예로서 상기 재조합 대장균 미생물의 제조방법의 상기 대장균 재조합 벡터를 합성하는 단계는 *asbF*^{opt}를 코딩하는 폴리뉴클레오티드에 추가된 상기 리보솜 결합 부위 앞에 *Xba*I 제한효소(서열번호 6), 상기 *catA*^{opt}를 코딩하는 폴리뉴클레오티드 뒤에 *Hind*III 제한효소(서열번호 7) 자리를 각각 추가하는 단계를 더 포함할 수 있다. 또한, 일 구현예로서 상기 벡터는 플라스미드를 사용할 수 있으며, 구체적으로 *Xba*I와 *Hind*III 제한효소 자리를 가지고 있으며 프로모터(promoter) 방향과 일치하는 플라스미드라면 그 종류에 제한없이 모두 사용할 수 있다.

[0023] 다른 일 구현예로서, 상기 벡터는 lac 프로모터 유래 lacZα 유전자가 제거된 벡터를 사용할 수 있다. 일 구현예로서 상기 lac 프로모터 유래 lacZα 유전자는 서열번호 8을 포함할 수 있다. 상기 lacZα 유전자가 제거된 벡터는 lacZα 유전자만 제거하고 프로모터 부분은 그대로 이용하고 있어 lac 리프레서(repressor)가 lac 오퍼레이터(operator)에 결합하는 것이 유지된다. 그러므로, *asbF*^{opt}, *aroY*^{opt} 및 *catA*^{opt} 가 발현되도록 하기 위해 IPTG(Isopropyl β-D-1-thiogalactopyranoside)을 도입하여 유전자 발현을 유도하는 단계가 필요로 하며, 공정 진행 중 글루코스가 고갈되거나 낮은 농도로 유지되어야 생산성이 높게 유지된다.

[0024] 본 발명은 상기와 같은 제조방법을 통해 뮤코닉산 생산용 대장균을 제조함으로써, 궁극적으로 대장균에서 합성생물학적 뮤코닉산 생산을 위한 최적의 외래유전자군 발현-조절오페론 시스템을 구축할 수 있다.

[0025] 이하, 본 발명의 제조예 및 시험예를 참조하여 본 발명을 상세히 설명한다. 이들은 오로지 본 발명을 보다 구체적으로 설명하기 위해 예시적으로 제시한 것일 뿐, 본 발명의 범위가 이 제조예 및 시험예에 의해 제한되지 않는다는 것은 당업계에서 통상의 지식을 가지는 자에 있어서 자명할 것이다.

[0026] [제조예]

[0027] 본 발명의 일 구현예에 따른 재조합대장균 미생물을 하기의 방법에 따라 제조하였다.

[0028] 사용 미생물, 배지 및 배양조건

[0029] 뮤코닉산 합성을 위한 대장균 미생물은 *E. coli* AB2834로 예일대학교의 *E. coli* Genetic Stock Center에서 구입하였다 (하기 표 1 및 표 2 참조).

표 1

[0030]

미생물(*E. coli*)	특성
AB2834	*aroE* 결손 *E. coli*
AB2834/pUC18△lacZ (비교예)	pUC18△lacZ 함유 AB2834
AB2834/pMESK1(실시예)	pMESK1 함유 AB2834

표 2

[0031]

벡터(플라스미드)	특성
pUC18△lacZ	*lacZ*α결손 pUC18
pMESK1	$asbF^{opt}$, $aroY^{opt}$, $catA^{opt}$ 포함 pUC18△lac

[0032] 먼저, 대장균 *E. coli* DH5α(구입처: TaKaRa, 구입번호: 9057)를 재조합 플라스미드 제작 시 모미생물로 사용하였다. 모든 대장균 미생물은 37℃, 220rpm에서 진탕배양 하였으며, 1L의 LB(Luria-Bertani)액체배지(박토트립톤 10.0g, 박토효모추출물 5.0g, NaCl 10.0g 함유)와 1L의 M9 액체배지(Na_2HPO_4 6.8g, KH_2PO_4 3.0g, NaCl 0.5g, NH_4Cl 1.0g를 980㎖의 1차수에 멸균 후, 멸균된 20% 글루코스 20㎖, 1M $CaCl_2$ 100㎕, 1M $MgSO_4$ 2㎖와 함께 L-try, L-phe, L-tyr 및 시킴산(shikimate)을 각각 0.04g씩 첨가)를 사용하였다. 항생제로 사용한 앰피실린(Ap)은 최종농도 50mg/L에 맞춰 배지에 넣어주었다. IPTG(Isopropyl-β-D-thiogalactopyranoside)는 0.5M 스톡(stock) 상태로 준비하였다. 프로모터를 조절하기 위해서 인덕션(induction)시에 IPTG를 사용하였으며, IPTG는 접종 후 36시간 뒤에 0.5mM로 6시간마다 인덕션해주었다. 3차 증류수에 녹인 아미노산, 시킴산, 앰피실린, IPTG, 20% 글루코스는 모두 0.25㎛ 멤브레인을 사용하여 멸균하였다.

[0033] pUC18△lacZ 플라스미드의 제조

[0034] *lacZ*α유전자(서열번호 8)가 제거된 pUC18 플라스미드의 제조를 위하여 *lacZ*α유전자를 기준으로 반대방향의 프라이머 한 쌍(서열번호 9 및 10)을 제작하여 PCR에 사용하였다.

[0035] - pUC18_lac_forward: 5'-agctgtttcctgtctagaaattgttatc-3' (서열번호 9),

[0036] - pUC18_lac_reverse: 5'-ttaagcttgccccgacacccgccaac-3' (서열번호 10)

[0037] 상기 두 염기서열 중 밑줄 친 서열은 각각 *Xba*Ⅰ, *Hind*Ⅲ 제한효소(서열번호 6 및 7) 자리이다.

[0038] PCR은 최종부피를 50㎕로 맞춰 5X 퓨전(Phusion) HF 버퍼 10㎕, 0.2mM dNTPs, 0.5μM의 상기 pUC18△lacZ forward 및 pUC18△lacZ reverse 두 프라이머, 3% DMSO, 0.25㎕의 pUC18 플라스미드가 포함된 혼합물을 제작한 후, C1000 Thermal Cycler (BIO-RAD)에서 수행하였다. 이때, PCR 방법은 다음과 같다: 1단계: 98℃ 30초, 2단계: 98℃ 10초, 60℃ 30초, 72℃ 1분 30초 (30회 반복), 3단계: 72℃ 10분.

[0039] PCR 산물은 1% (w/v) 아가로스겔에 전기영동을 통해 확인 후 PCR 정제(purification) 키트를 사용하여 정제 후 T4 라가아제(Takara)를 사용하여 자가-라이게이션(self-ligation)하였다. 라이게이션된 벡터는 pUC18△lacZ sequencingprimer (5'-ttggccgattcattaatgcag-3', 서열번호 11)를 제작 후 그 서열을 확인하였다 (Macrogen, Korea).

[0040] AB2834/pMESK1 대장균(실시예)의 제조

[0041] AB2834/pMESK1 대장균(실시예)을 제조하기 위하여, 먼저 데이터베이스 검색을 통해 *B. thuringiensis*유래의 *asbF*, *K. pneumonia* 유래의 *aroY*, *Acinetobacter calcoaceticus* 유래의 *catA*를 선별하여 대장균 코돈으로 최적화하여, $asbF^{opt}$(서열번호 1), $aroY^{opt}$(서열번호 2) 및 $catA^{opt}$(서열번호 3)를 코딩하는 폴리뉴클레오티드를 합성하였다. 그 다음, 폴리뉴클레오티드의 각 업스트림(upstream)에 대장균 유래 리보솜 결합 부위(RBS, 서열번호 4)를 추가하여 합성 후 pUC57 플라스미드에 *Xba*Ⅰ, *Hind*Ⅲ 제한효소 자리를 이용하여 클로닝하였다 (Cosmo genetech, Korea). 최종적으로 pUC18△lacZ 플라스미드에 상기 합성한 $asbF^{opt}$, $aroY^{opt}$ 및 $catA^{opt}$를 각각 코딩하는 상기 폴리뉴클레오티드 및 리보솜 결합 부위를 포함하는 유전자 단편을 *Xba*Ⅰ, *Hind*Ⅲ 제한효소 자리에 클로닝하여(도 2 참조), pMESK1 플라스미드를 제작하였다.

[0042] 하기의 표 3은 상기 제조된 pMESK1 플라스미드의 염기서열을 나타낸 것(서열번호 32)이다.

표 3

[0043]

[0044] 그리고 이 플라스미드를 *E. coli* AB2834에 형질전환시켜, AB2834/pMESK1(실시예, 기탁번호: KCCM11505P), 즉 본 발명에 따른 뮤코닉산 생산용 대장균을 제조하였다.

[0045] AB2834/pUC18△lacZ (비교예)의 제조

[0046] 상기에서 제조된 pUC18△lacZ 플라스미드를 *E. coli* AB2834에 형질전환시켜, AB2834/pUC18△lacZ (비교예)를 제조하였다.

[0047] [시험예 1]

[0048] 상기 제조예에서 제조한 재조합대장균의 생산곡선 및 대사산물(뮤코닉산 및 뮤코닉산 전구체)을 하기와 같이 분석하였다.

[0049] 먼저, 상기 비교예(AB2834/pUC18△lacZ) 및 실시예(AB2834/pMESK1)를 250㎖ 플라스크 배양한 후 50㎖ M9배지 및 LB배지 모두에서 각각 배양하고 일정시간 간격으로 샘플링을 하여 각 미생물의 생장곡선 및 대사산물의 HPLC 분석을 수행하였다.

[0050] 구체적으로, 채취한 배양액 1㎖을 13000rpm에 5분 원심분리하여 세포를 분리하여 상등액을 0.25㎛ 멤브레인 필터에 여과하여 분석에 사용되었다. 생합성된 시스,시스-뮤코닉산(CCM) 및 전구체들은HPLC(고속액체크로마토그래피)로 분리-분석되었으며, 사용한 컬럼은 Aminex HPX-87H, 이동상은 5mM H_2SO_4,, 유속은 0.6㎖/L, 분석온도는 65℃에서 수행되었다 (FRC-10A, SHIMADZU). UV 검출은 PCA (259nm), 3-DHS (237nm), 카테콜 (275nm), CCM (262 nm) 에서 수행하였다. 정량분석을 위해 사용한 표준(standard) (PCA, 3-DHS, Catechol, CCM: Sigma)는 각각 증류수에 녹여 0.25㎛ 멤브레인에 여과하여 사용하였다.

[0051] OD_{600} 값을 바탕으로 생장곡선을 분석한 결과, 두 종류의 배지에서 모두 공벡터만 넣은 대조군 미생물인 비교예 (AB2834/pUC18△lacZ)가 실시예(AB2834/pMESK1)보다 더 빠르게 지수기에 도달하는 것을 확인하였다. 최고 OD_{600} 값은 두 미생물 모두 LB배지에서 높은 값을 보였으며 정지기에서 수행하는 0.5 mM IPTG 인덕션 시간 또한 LB배지가 더 빠른 것으로 나타났다.

[0052] 그 다음, 뮤코닉산이 생산되는지 확인한 결과 실시예(AB2834/pMESK1)의 경우 LB배지에서 3.5mg/L, M9배지에서 9.5mg/L의 뮤코닉산이 생산되는 것을 확인하였다(도 3 참조). 또한 3L 발효조에서 뮤코닉산의 생산을 확인한 결과, 실시예(AB2834/pMESK1)의 경우 배양 72시간 후에 뮤코닉산 7.7 g/L, 3-DHS 0.3 g/L, PCA 0.2 g/L이 각각 생산되었으며 이때까지 뮤코닉산의 생성속도가 일정하게 유지되고 있어 배양을 계속할 경우 뮤코닉산의 생산량도 계속 늘어갈 것으로 예상된다 (도 4 참조).

[0053] [시험예 2]

[0054] 본 발명의 일 구현예에 따른 상기 제조예에서 제조한 실시예 (AB2834/pMESK1)에 있어서, 도입한 3종의 외래유전자와 기존의 방향족아미노산 생합성 경로 유전자들의 발현양상, 즉 뮤코닉산 생합성 경로 유전자들의 전사 확인을 위한 RT-PCR을 하기와 같이 실시하였다.

[0055] 먼저, pMESK1 플라스미드를 형질전환 시킨 AB2834/pMESK1(실시예)를 5㎖ LB 배지에 전배양 하고50㎖ M9배지에 1/100로 접종하여 배양 23시간 뒤 0.5mM IPTG 인덕션을 하고, 3시간 후 배양액을 1300rpm에 5분 원심분리 후 세포를 RNase가 처리된 멸균된 막자사발로 파쇄하였다. RNA는 RNeasy Mini Kit (Qiagen)을 사용하여 분리하였고, cDNA 합성은 PrimeScript 1^{ST} strand cDNA Synthesis Kit (Takara)를 사용하여 합성하였다. RT-PCR은 최종부피 20㎕로 맞춰 10X 버퍼 2㎕, 0.25mM dNTPs, 0.4μM의 각 두 쌍의 프라이머 (하기 표 4 참조), 10% DMSO, 1㎕의 합성한 cDNA를 혼합물로 제작 후 C1000 Thermal Cycler (BIO-RAD)에서 수행되었다. RT-PCR 조건은 다음과 같다: 1단계: 95℃ 15분, 2단계: 95℃ 45초, 60℃ 45초, 72℃ 40초 (30회 반복), 3단계: 72℃ 10분. PCR 산물은 2% (w/v) 아가로스겔에 전기영동을 통해 확인하였다(도 5 참조).

표 4

[0056]

서열번호	프라이머	염기서열(5' -> 3')	타겟[a]
11	pUC18_lac_sequencing	ttggccgattcattaatgcag	Plasmid pUC18_lac
12	pUC18_lac_F	agctgtttcctgtctagaaattgttatc	Plasmid pUC18
13	pUC18_lac_R	ttaagcttgccccgacacccgccaac	Plasmid pUC18
14	aroB_RT-PCR_F	TGTCGTTACTCTCGGGGAAC	
15	aroB_RT-PCR_R	CCGCGGACCTTATCGAGATA	
16	aroD_RT-PCR_F	CCGAAGAAATCATTGCCCGT	
17	aroD_RT-PCR_R	TCACCAGCCAGACGAGAAAT	
18	aroF_RT-PCR_F	CTGAAGGCCGCTTTTCCATT	
19	aroF_RT-PCR_R	TTCCAGAGCAGTTTCCGGAT	

20	aroG_RT-PCR_F	CTGACGTTTGCCAGCAGATT	
21	aroG_RT-PCR_R	TGACGTAACAGAGCATCGGT	
22	aroH_RT-PCR_F	AAACCACGAACTGTTGTCGG	Strain AB2834/pMESK1 cDNA
23	aroH_RT-PCR_R	CCGGTCACCATATCGAGGAA	Strain AB2834/pMESK1 cDNA
24	asbF_RT-PCR_F	TCCTGCATATTTGGGAAAGC	Strain AB2834/pMESK1 cDNA
25	asbF_RT-PCR_R	ATGCCCTCAAACAGTGGAAC	
26	aroY_RT-PCR_F	ATCCTGTGGGCTATGACGAC	
27	aroY_RT-PCR_R	GGTGCAGTCGAAAATGGTTT	
28	catA_RT-PCR_F	TATTTCACCGATGCTGGTCA	
29	catA_RT-PCR_R	ATACAGCGGACCTTCAATGG	
30	GAPDH_RT-PCR_F	TTTCCGTGCTGCTCAGAAAC	
31	GAPDH_RT-PCR_R	GTCAACACCAACTTCGTCCC	

[0057] 전사 발현량 비교 시 대조군으로 사용한 하우스키핑(housekeeping) 유전자는 GAPDH(D-glyceraldehyde-3-phosphate dehydrogenase)로 D-글리세랄데히드-3-포스페이트(D-glyceraldehyde-3-phosphate)를 1,3-포스포글리세레이트(1,3-phosphoglycerate)로 전환시키는 효소로서, 해당과정과 글루코오스 신생합성에 관여하는 중요한 효소이다. GAPDH와 같은 하우스키핑 유전자들은 조절과정 없이 모든 세포에서 항상 발현되는 유전자이기 때문에 정량적 RT-PCR를 수행할 때 대조군 유전자로 사용된다. RT-PCR 분석 결과, 도 5에 나타난 바와 같이 고 사본수(high copy number) 벡터에 클로닝하여 도입해준 *asbF*^{opt}, *aroY*^{opt}, *catA*^{opt} 유전자들은 GAPDH와 비교 했을 때 높은 발현량을 보이는 것을 확인할 수 있었다. 반면 방향족아미노산 생합성 경로의 유전자들 (*aroB, aroD, aroF, aroG, aroH*)은 상대적으로 매우 낮은 발현량을 보였다. 특히 *aroD, aroG* 유전자들은 2% (w/v) 아가로스겔 전기영동을 통한 발현량 관찰에서 육안으로 확인하기에 힘들 정도의 수준이었다. 이는 사용한 미생물인 *E.coli* AB2834이 *aroE* 유전자의 결손으로 인해 배양 시 배지에 넣어준 방향족아미노산인 페닐알라닌, 타이로신, 트립토판에 의해 세 종류의 DAHP 생성효소인 *AroF, AroG, AroH*가 피드백 억제(feedback inhibition)을 받아 유전자들이 전사수준에서 제한을 받기 때문으로 예상된다. 따라서 뮤코닉산 합성 초기단계인 DAHP synthase 유전자의 전사에 대한 제한에 의해 다음 단계인 DHQ synthase인 *AroB*, DHQ dehydratase인 *AroD*의 합성 또한 낮은 수준으로 이루어져 뮤코닉산 및 중간단계물질의 생산량에 부정적으로 작용할 것으로 예상된다.

[0058] 대장균 코돈에 최적화된 염기서열로 합성한 3종의 외래유전자(폴리뉴클레오티드) *asbF*^{opt}, *aroY*^{opt}, *catA*^{opt} 의 도입으로, *aroE*가 제거된 재조합 대장균에서 뮤코닉산이 플라스크 배양에서는 9.5mg/L과 3L 발효기 배양에서는 7.7g/L이 각각 생성되었다. 이는 3종의 외래유전자 *asbF, aroY, catA*를 대장균 코돈에 최적화시킨 폴리뉴틀레오티드로 합성하고 각각의 유전자 앞에 리보솜 결합 부위(RBS)를 도입하여 고사본수(high copy number) 벡터의 단일 오페론 구조로 도입함으로써, 3종의 외래유전자들을 동시에 발현시킬 수 있는 효율적인 발현시스템이 구축된 것을 의미한다. 또한 상기 3종의 외래유전자 발현 수준을 동일하게 유지할 수 있을 뿐만 아니라 3종의 외래유전자 발현을 동시에 동일하게 극대화 시킬 수 있게 되었다. 도입된 3종의 외래유전자는 각각의 해당 효소인 3-DHS 탈수효소, PCA 탈카복실화효소(protocatechuic acid decarboxylase), 카테콜 1,2-산소화효소(catechol 1,2-dioxygenase)를 효율적으로 발현시켰으며, 궁극적으로 3-DHS 축적을 위해 *aroE*가 제거된 재조합된 대장균에서 뮤코닉산을 효율적으로 생합성 할 수 있었다.

수탁번호

[0059]

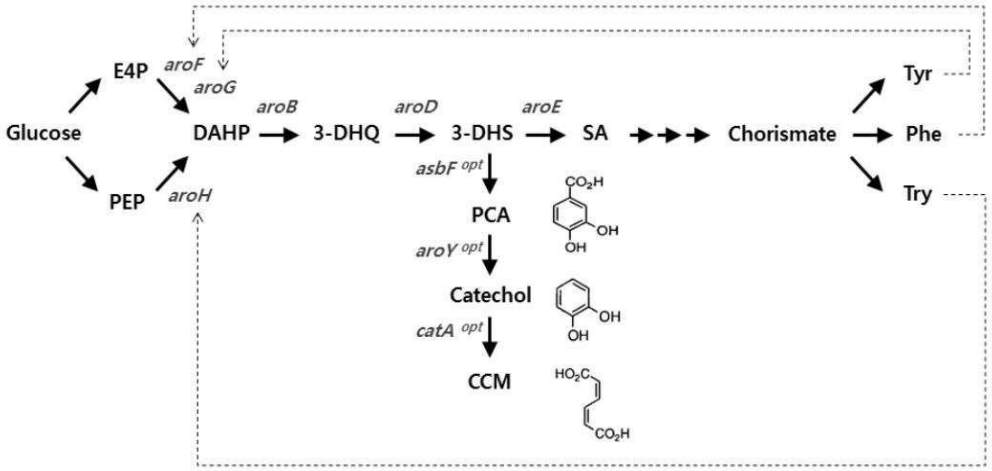
기탁기관명 : 한국미생물보존센터(국외)

수탁번호 : KCCM11505P

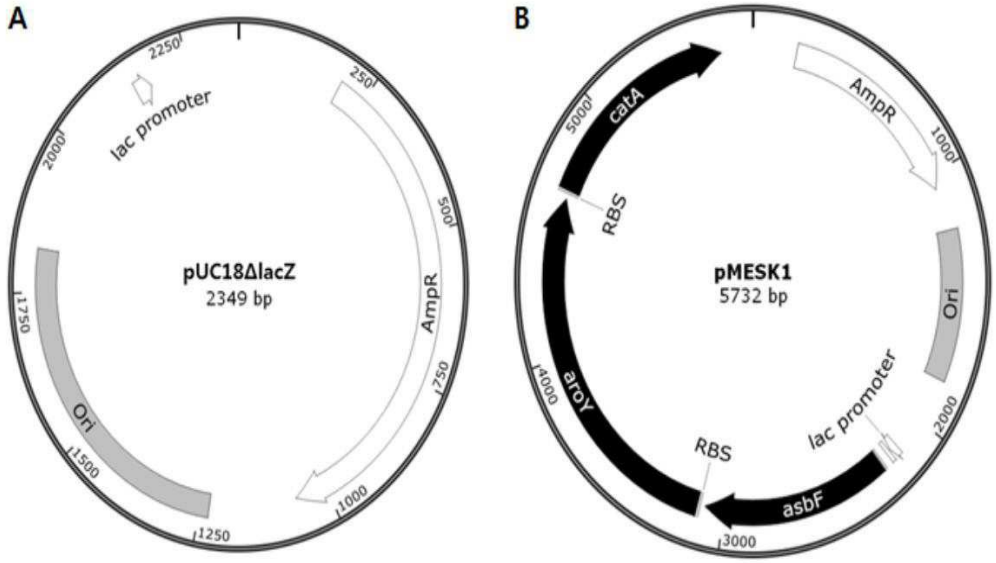
수탁일자 : 20140110

도면

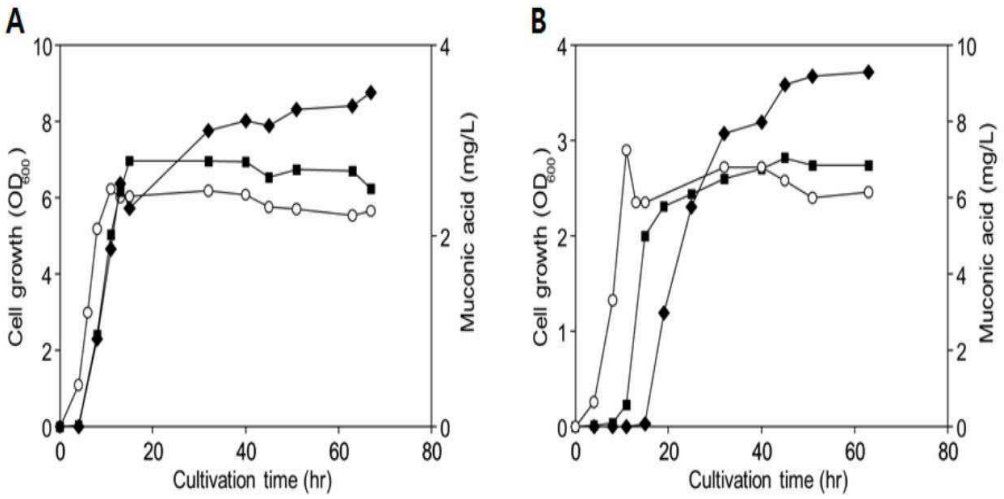
도면1


Glucose
E4P
PEP
aroF
aroG
aroH
DAHP
aroB
3-DHQ
aroD
3-DHS
aroE
SA
Chorismate
Tyr
Phe
Try
asbF opt
PCA
aroY opt
Catechol
catA opt
CCM

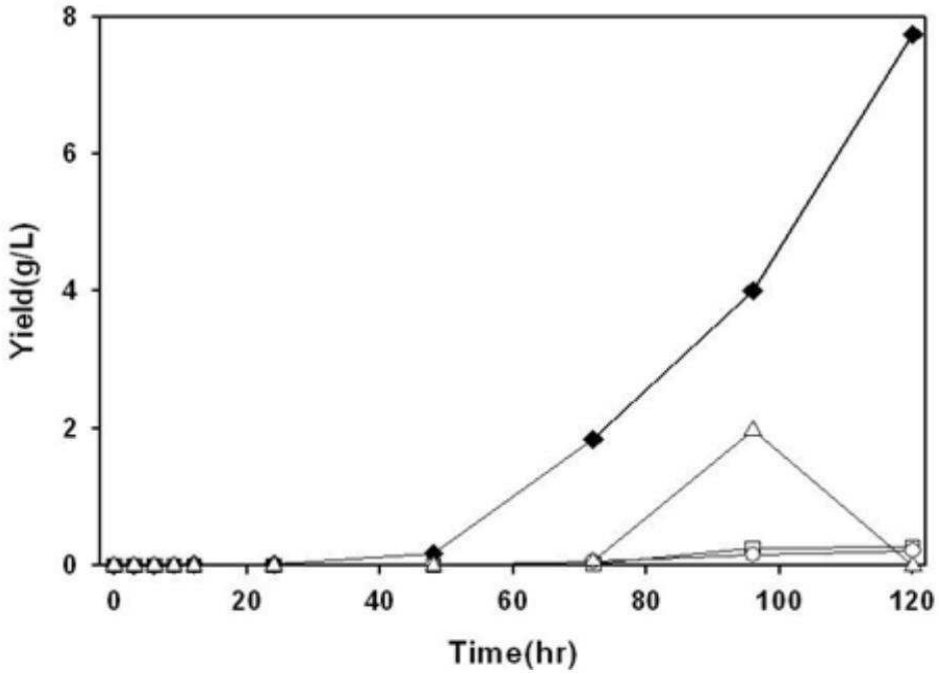
도면2


A
pUC18ΔlacZ
2349 bp
lac promoter
AmpR
Ori
B
pMESK1
5732 bp
catA
RBS
AmpR
Ori
aroY
RBS
lac promoter
asbF

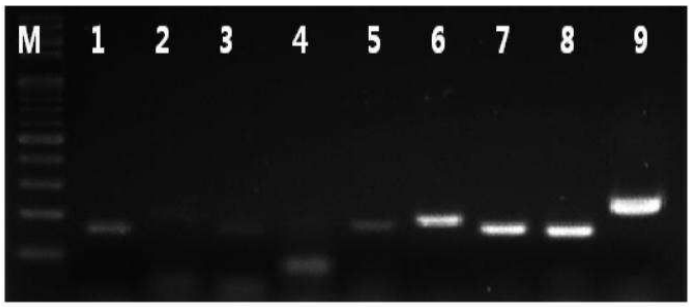
도면3


A
B
Cell growth (OD_{600})
Muconic acid (mg/L)
Cultivation time (hr)

도면4


Yield(g/L)
Time(hr)

도면5


M
1
2
3
4
5
6
7
8
9

서 열 목 록

<110> STR Biotech Co.,Ltd.

<120> Microorganism for producing muconic acid and method for manufacturing the strain

<130> 13P555IND

<160> 32

<170> KopatentIn 2.0

<210> 1

<211> 843

<212> DNA

<213> Artificial Sequence

<220><223> optimized asbF(3-dehydroshikimate dehydratase)

<400> 1

atgaaatact ccctgtgcac tattagcttc cgtcatcaac tgatttcttt cactgacatc 60

gttcagttcg cgtacgaaaa cggttttgaa ggcatcgagc tgtggggtac tcatgcccag 120

aacctgtaca tgcaggagcg tgaaaccacc gagcgtgagc tgaacttcct gaaagataag 180

aacctggaaa tcaccatgat ctctgactac ctggatattt ccctgtccgc cgacttcgag 240

aaaaccattg aaaaatccga acagctggtg gtgctggcaa actggttcaa caccaacaag 300

atccgtacct ttgcgggcca gaaaggtagc aaggactttt ctgaacagga acgtaaagag 360

tacgtaaagc gcatccgtaa aatctgcgac gtttttgcac agcataacat gtacgtcctg 420

ctggaaactc atccgaacac cctgaccgac actctgccta gcaccattga actgctggaa 480

gaagtgaacc atccaaacct gaaaatcaac ctggatttcc tgcatatttg ggaaagcggc 540

gcgaatccaa tcgactcttt ccatcgcctg aaaccgtgga ctctgcacta ccatttcaaa 600

aacatctcct ccgcggatta cctgcacgtg ttcgaaccga ataacgtcta cgccgctgca 660

ggctcccgta ttggtatggt tccactgttt gagggcatcg ttaactacga cgaaatcatt 720

caagaagttc gtggcaccga cctgtttgct tctctggaat ggttcggcca caactctaaa 780

gagatcctga aggaagaaat gaaagtgctg atcaaccgta aactggaagt ggtgacgagc 840

tga 843

<210> 2

<211> 1509

<212> DNA

<213> Artificial Sequence

<220><223> optimized aroY(protocatechuate decarboxylase)

<400> 2

atgactgcac caatccagga cctgcgcgac gcaatcgctc tgctgcaaca gcacgataac 60

cagtatctgg aaaccgatca cccagttgac ccgaatgcgg agctggcagg cgtctaccgt 120

catattggtg ccggtggcac tgttaaacgc ccgacccgca tcggcccagc aatgatgttt 180

aacaacatca aaggctaccc gcacagccgc atcctggtag gcatgcacgc ttctcgtcaa 240

cgtgcggcac tgctgctggg ctgtgaagca tctcagctgg ctctggaggt gggcaaggcc 300

gtcaaaaagc cggtggcgcc ggttgttgtt ccggcatctt ccgctccttg ccaggaacag 360

attttcctgg ccgatgatcc ggacttcgac ctgcgtacgc tgctgccggc tcacacgaac 420

actccgatcg acgcgggtcc gtttttctgc ctgggtctgg ctctggcgtc tgacccggtg 480

gatgcgagcc tgaccgacgt gaccatccac cgcctgtgcg ttcagggtcg tgacgaactg 540

agcatgttcc tggcagcagg tcgtcacatt gaagtattcc gccaaaaagc ggaagccgcc 600

ggtaaaccgc tgcctattac catcaacatg ggtctggacc cggcgatcta catcggcgct 660

tgctttgaag cacctacgac tccgttcggt tacaacgaac tgggtgtagc cggtgctctg 720

cgccagcgtc cggtggaact ggtacagggc gtgagcgtcc cagaaaaagc catcgcacgc 780

gctgaaatcg taatcgaagg cgaactgctg ccgggcgtcc gcgttcgtga agaccagcac 840

accaacagcg gtcacgcaat gcctgaattc ccgggctact gtggcggtgc caacccgtcc 900

ctgccggtta ttaaggttaa agctgttact atgcgtaaca acgcgattct gcagactctg 960

gtcggtccgg gtgaggaaca cacgaccctg gcgggcctgc cgactgaagc ctctatttgg 1020

aatgcggttg aagcagccat cccgggcttc ctgcagaatg tttacgctca taccgcgggt 1080

ggcggtaaat tcctgggtat cctgcaggtt aaaaagcgcc agccggcaga tgaaggtcgc 1140

caaggtcagg ctgccctgct ggcgctggct acctactctg aactgaaaaa catcattctg 1200

gttgacgaag atgtcgatat ctttgattcc gacgacatcc tgtgggctat gacgactcgt 1260

atgcagggcg atgtttctat caccaccatc ccgggcatcc gtggtcacca gctggacccg 1320

tcccagactc cggaatattc cccgagcatc cgcggcaacg gcatctcctg taaaaccatt 1380

ttcgactgca ccgttccgtg ggctctgaaa tcccacttcg agcgtgcgcc gtttgcggac 1440

gttgatccgc gtccgttcgc accggagtat ttcgctcgtc tggagaaaaa tcagggttcc 1500

gcgaaataa 1509

<210> 3

<211> 923

<212> DNA

<213> Artificial Sequence

<220><223> optimized catA(catechol 1,2-dioxygenase)

<400> 3

atgaaccgtc agcagatcga cgcgctggtt aaacagatga acgtggatac cgctaagggc 60

gaagttgacg ctcgcgtaca gcagattgta gtacgtctgc tgggtgacct gttccaggca 120

atcgaagatc tggatattca gccgtctgaa gtgtggaaag gtctggagta tttcaccgat 180

gctggtcagg cgaacgaact gggtctgctg gcggccggtc tgggcctgga gcactatctg 240

gacctgcgtg cggatgaagc agatgcaaaa gcaggtgtga ccggtggtac tccgcgtacc 300

attgaaggtc cgctgtatgt tgcaggtgct ccggaaagcg ttggtttcgc gcgtatggat 360

gacggtactg aatctggtaa aatcgatact ctgatcattg aaggcaccgt caccgacacc 420

gacggtaata tcatcgaaaa cgctaaagta gaggtttggc acgcgaactc tctgggcaac 480

tattctttct ttgataaatc ccagtccgac ttcaacctgc gtcgtaccat tctgaccgat 540

gcggacggta aatatgttgc gctgacgacg atgccagtag gctatggctg tccgccggaa 600

ggtaccaccc aagcgctgct gaacaaactg ggtcgccacg gtaaccgtcc ttctcacgta 660

cactatttcg tgtctgctcc gggctaccgt aaactgacga cccaattcaa tatcgagggt 720

gacgaatacc tgtgggatga tttcgctttt gctacccgtg atggtctggt ggcgaccgcg 780

gtagacgtga ccgatccagc tgaaatccag cgtcgcggcc tggatcacgc ttttaaacac 840

atcaccttca acattgaact ggttaaagat gcagccgcgg cacctagcac tgaggtagaa 900

cgccgtcgtg cgtccgctta att 923

<210> 4

<211> 6

<212> DNA

<213> Artificial Sequence

<220><223> ribosome binding site

<400> 4

gaagga 6

<210> 5

<211> 819

<212> DNA

<213> aroE(shikimate dehydrogenase)

<400> 5

atggaaacct atgctgtttt tggtaatccg atagcccaca gcaaatcgcc attcattcat 60

cagcaatttg ctcagcaact gaatattgaa catccctatg ggcgcgtgtt ggcacccatc 120

aatgatttca tcaacacact gaacgctttc tttagtgctg gtggtaaagg tgcgaatgtg 180

acggtgcctt ttaaagaaga ggcttttgcc agagcggatg agcttactga acgggcagcg 240

ttggctggtg ctgttaatac cctcatgcgg ttagaagatg gacgcctgct gggtgacaat 300

accgatggtg taggcttgtt aagcgatctg gaacgtctgt cttttatccg ccctggttta 360

cgtattctgc ttatcggcgc tggtggagca tctcgcggcg tactactgcc actcctttcc 420

ctggactgtg cggtgacaat aactaatcgg acggtatccc gcgcggaaga gttggctaaa 480

ttgtttgcgc acactggcag tattcaggcg ttgagtatgg acgaactgga aggtcatgag 540

tttgatctca ttattaatgc aacatccagt ggcatcagtg gtgatattcc ggcgatcccg 600

tcatcgctca ttcatccagg catttattgc tatgacatgt tctatcagaa aggaaaaact 660

ccttttctgg catggtgtga gcagcgaggc tcaaagcgta atgctgatgg tttaggaatg 720

ctggtggcac aggcggctca tgcctttctt ctctggcacg gtgttctgcc tgacgtagaa 780

ccagttataa agcaattgca ggaggaattg tccgcgtga 819

<210> 6

<211> 6

<212> DNA

<213> XbaI restrction enzyme

<400> 6

tctaga 6

<210> 7

<211> 6

<212> DNA

<213> HindIII restriction enzyme

<400> 7

aagctt 6

<210> 8

<211> 324

<212> DNA

<213> Artificial Sequence

<220><223> lacZ-alpha which codes for the N-terminus fragment of beta-galactosidase

<400> 8

atgaccatga ttacgaattc gagctcggta cccggggatc ctctagagtc gacctgcagg 60

catgcaagct tggcactggc cgtcgtttta caacgtcgtg actgggaaaa ccctggcgtt 120

acccaactta atcgccttgc agcacatccc cctttcgcca gctggcgtaa tagcgaagag 180

gcccgcaccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg gcgcctgatg 240

cggtattttc tccttacgca tctgtgcggt atttcacacc gcatatggtg cactctcagt 300

acaatctgct ctgatgccgc atag 324

<210> 9

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> pUC18_lacZ forward primer

<400

> 9

agctgtttcc tgtctagaaa ttgttatc 28

<210> 10

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> pUC18_lacZ reverse primer

<400> 10

ttaagcttgc cccgacaccc gccaac 26

<210> 11

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> pUC18_lacZ sequencing primer

<400> 11

ttggccgatt cattaatgca g 21

<210

> 12

<211> 28

<212> DNA

<213> Artificial Sequence

<220><223> pUC18_lacZ_Forward primer tageting plasmid pUC18

<400> 12

agctgtttcc tgtctagaaa ttgttatc 28

<210> 13

<211> 26

<212> DNA

<213> Artificial Sequence

<220><223> pUC18_lacZ_Reverse primer tageting plasmid pUC18

<400> 13

ttaagcttgc cccgacaccc gccaac 26

<210> 14

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroB_RT-PCR_Forward primer

<400> 14

tgtcgttact ctcggggaac 20

<210> 15

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroB_RT-PCR_Reverse primer

<400> 15

ccgcggacct tatcgagata 20

<210> 16

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroD_RT-PCR_Forward primer

<400> 16

ccgaagaaat cattgcccgt 20

<210> 17

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroD_RT-PCR_Reverse primer

<400> 17

tcaccagcca gacgagaaat 20

<210> 18

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroF_RT-PCR_Forward primer

<400> 18

ctgaaggccg cttttccatt 20

<210> 19

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroF_RT-PCR_Reverse primer

<400> 19

ttccagagca gtttccggat 20

<210> 20

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroG_RT-PCR_Forward primer

<400> 20

ctgacgtttg ccagcagatt 20

<210> 21

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroG_RT-PCR_Reverse primer

<400> 21

tgacgtaaca gagcatcggt 20

<210> 22

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroH_RT-PCR_Forward primer

<400> 22

aaaccacgaa ctgttgtcgg 20

<210> 23

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroH_RT-PCR_Reverse primer

<400> 23

ccggtcacca tatcgaggaa 20

<210> 24

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> asbF_RT-PCR_Forward primer

<400> 24

tcctgcatat ttgggaaagc 20

<210> 25

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> asbF_RT-PCR_Reverse primer

<400> 25

atgccctcaa acagtggaac 20

<210> 26

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroY_RT-PCR_Forward primer

<400> 26

atcctgtggg ctatgacgac 20

<210> 27

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> aroY_RT-PCR_Reverse primer

<400> 27

ggtgcagtcg aaaatggttt 20

<210> 28

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> catA_RT-PCR_Forward primer

<400> 28

tatttcaccg atgctggtca 20

<210> 29

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> catA_RT-PCR_Reverse primer

<400> 29

atacagcgga ccttcaatgg 20

<210> 30

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> GAPDH_RT-PCR_Forward primer

<400> 30

tttccgtgct gctcagaaac 20

<210> 31

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> GAPDH_RT-PCR_Reverse primer

<400> 31

gtcaacacca acttcgtccc 20

<210> 32

<211> 5732

<212> DNA

<213> Artificial Sequence

<220><223> pMESK1_plasmid sequences

<400> 32

gacgaaaggg cctcgtgata cgcctatttt tataggttaa tgtcatgata ataatggttt 60

cttagacgtc aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt 120

tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat 180

aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt 240

ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg 300

ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga 360

tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc 420

tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac 480

actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg 540

gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca 600

acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg 660

gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg 720

acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg 780

gcgaactact tactctagct tcccggcaac aattaataga ctggatggag gcggataaag 840

ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg 900

gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct 960

cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac 1020

agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac caagtttact 1080

catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga 1140

tccttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt 1200

cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct 1260

gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc 1320

taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc 1380

ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc 1440

tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg 1500

ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggttt 1560

cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg 1620

agctttgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg 1680

gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt 1740

atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag 1800

gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt 1860

gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta 1920

ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt 1980

cagtgagcga ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc 2040

cgattcatta atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca 2100

acgcaattaa tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc 2160

cggctcgtat gttgtgtgga attgtgagcg gataacaatt tctagaaata attttgttta 2220

actttaagaa ggagatatac atatgaaata ctccctgtgc actattagct tccgtcatca 2280

actgatttct ttcactgaca tcgttcagtt cgcgtacgaa aacggttttg aaggcatcga 2340

gctgtggggt actcatgccc agaacctgta catgcaggag cgtgaaacca ccgagcgtga 2400

gctgaacttc ctgaaagata agaacctgga aatcaccatg atctctgact acctggatat 2460

ttccctgtcc gccgacttcg agaaaaccat tgaaaaatcc gaacagctgg tggtgctggc 2520

aaactggttc aacaccaaca agatccgtac ctttgcgggc cagaaaggta gcaaggactt 2580

ttctgaacag gaacgtaaag agtacgtaaa gcgcatccgt aaaatctgcg acgtttttgc 2640

acagcataac atgtacgtcc tgctggaaac tcatccgaac accctgaccg acactctgcc 2700

tagcaccatt gaactgctgg aagaagtgaa ccatccaaac ctgaaaatca acctggattt 2760

cctgcatatt tgggaaagcg gcgcgaatcc aatcgactct ttccatcgcc tgaaaccgtg 2820

gactctgcac taccatttca aaaacatctc ctccgcggat tacctgcacg tgttcgaacc 2880

gaataacgtc tacgccgctg caggctcccg tattggtatg gttccactgt ttgagggcat 2940

cgttaactac gacgaaatca ttcaagaagt tcgtggcacc gacctgtttg cttctctgga 3000

atggttcggc cacaactcta aagagatcct gaaggaagaa atgaaagtgc tgatcaaccg 3060

taaactggaa gtggtgacga gctgaaataa ttttgtttaa ctttaagaag gagatataca 3120

tatgactgca ccaatccagg acctgcgcga cgcaatcgct ctgctgcaac agcacgataa 3180

ccagtatctg gaaaccgatc acccagttga cccgaatgcg gagctggcag gcgtctaccg 3240

tcatattggt gccggtggca ctgttaaacg cccgacccgc atcggcccag caatgatgtt 3300

taacaacatc aaaggctacc cgcacagccg catcctggta ggcatgcacg cttctcgtca 3360

acgtgcggca ctgctgctgg gctgtgaagc atctcagctg gctctggagg tgggcaaggc 3420

cgtcaaaaag ccggtggcgc cggttgttgt tccggcatct tccgctcctt gccaggaaca 3480

gattttcctg gccgatgatc cggacttcga cctgcgtacg ctgctgccgg ctcacacgaa 3540

cactccgatc gacgcgggtc cgtttttctg cctgggtctg gctctggcgt ctgacccggt 3600

ggatgcgagc ctgaccgacg tgaccatcca ccgcctgtgc gttcagggtc gtgacgaact 3660

gagcatgttc ctggcagcag gtcgtcacat tgaagtattc cgccaaaaag cggaagccgc 3720

cggtaaaccg ctgcctatta ccatcaacat gggtctggac ccggcgatct acatcggcgc 3780

ttgctttgaa gcacctacga ctccgttcgg ttacaacgaa ctgggtgtag ccggtgctct 3840

gcgccagcgt ccggtggaac tggtacaggg cgtgagcgtc ccagaaaaag ccatcgcacg 3900

cgctgaaatc gtaatcgaag gcgaactgct gccgggcgtc cgcgttcgtg aagaccagca 3960

caccaacagc ggtcacgcaa tgcctgaatt cccgggctac tgtggcggtg ccaacccgtc 4020

cctgccggtt attaaggtta aagctgttac tatgcgtaac aacgcgattc tgcagactct 4080

ggtcggtccg ggtgaggaac acacgaccct ggcgggcctg ccgactgaag cctctatttg 4140

gaatgcggtt gaagcagcca tcccgggctt cctgcagaat gtttacgctc ataccgcggg 4200

tggcggtaaa ttcctgggta tcctgcaggt taaaaagcgc cagccggcag atgaaggtcg 4260

ccaaggtcag gctgccctgc tggcgctggc tacctactct gaactgaaaa acatcattct 4320

ggttgacgaa gatgtcgata tctttgattc cgacgacatc ctgtgggcta tgacgactcg 4380

tatgcagggc gatgtttcta tcaccaccat cccgggcatc cgtggtcacc agctggaccc 4440

gtcccagact ccggaatatt ccccgagcat ccgcggcaac ggcatctcct gtaaaaccat 4500

tttcgactgc accgttccgt gggctctgaa atcccacttc gagcgtgcgc cgtttgcgga 4560

cgttgatccg cgtccgttcg caccggagta tttcgctcgt ctggagaaaa atcagggttc 4620

cgcgaaataa aataattttg tttaacttta agaaggagat atacatatga accgtcagca 4680

gatcgacgcg ctggttaaac agatgaacgt ggataccgct aagggcgaag ttgacgctcg 4740

cgtacagcag attgtagtac gtctgctggg tgacctgttc caggcaatcg aagatctgga 4800

tattcagccg tctgaagtgt ggaaaggtct ggagtatttc accgatgctg gtcaggcgaa 4860

cgaactgggt ctgctggcgg ccggtctggg cctggagcac tatctggacc tgcgtgcgga 4920

tgaagcagat gcaaaagcag gtgtgaccgg tggtactccg cgtaccattg aaggtccgct 4980

gtatgttgca ggtgctccgg aaagcgttgg tttcgcgcgt atggatgacg gtactgaatc 5040

tggtaaaatc gatactctga tcattgaagg caccgtcacc gacaccgacg gtaatatcat 5100

cgaaaacgct aaagtagagg tttggcacgc gaactctctg ggcaactatt ctttctttga 5160

taaatcccag tccgacttca acctgcgtcg taccattctg accgatgcgg acggtaaata 5220

tgttgcgctg acgacgatgc cagtaggcta tggctgtccg ccggaaggta ccacccaagc 5280

gctgctgaac aaactgggtc gccacggtaa ccgtccttct cacgtacact atttcgtgtc 5340

tgctccgggc taccgtaaac tgacgaccca attcaatatc gagggtgacg aatacctgtg 5400

ggatgatttc gcttttgcta cccgtgatgg tctggtggcg accgcggtag acgtgaccga 5460

tccagctgaa atccagcgtc gcggcctgga tcacgctttt aaacacatca ccttcaacat 5520

tgaactggtt aaagatgcag ccgcggcacc tagcactgag gtagaacgcc gtcgtgcgtc 5580

cgcttaatta agcttgcccc gacacccgcc aacacccgct gacgcgccct gacgggcttg 5640

tctgctcccg gcatccgctt acagacaagc tgtgaccgtc tccgggagct gcatgtgtca 5700

gaggttttca ccgtcatcac cgaaacgcgc ga 5732