



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0072410
(43) 공개일자 2020년06월22일

- | | |
|--|--|
| <p>(51) 국제특허분류(Int. Cl.)
C12Q 1/6888 (2018.01)</p> <p>(52) CPC특허분류
C12Q 1/6888 (2018.05)
C12Q 2600/156 (2013.01)</p> <p>(21) 출원번호 10-2019-0161201</p> <p>(22) 출원일자 2019년12월06일
심사청구일자 2019년12월06일</p> <p>(30) 우선권주장
1020180159746 2018년12월12일 대한민국(KR)</p> | <p>(71) 출원인
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)</p> <p>(72) 발명자
이준현
대전광역시 유성구 어은로 57, 103동 606호(어은동, 한빛아파트)
이승환
대전광역시 유성구 어은로 57, 127동 803호(어은동, 한빛아파트)
(뒷면에 계속)</p> <p>(74) 대리인
최규환</p> |
|--|--|

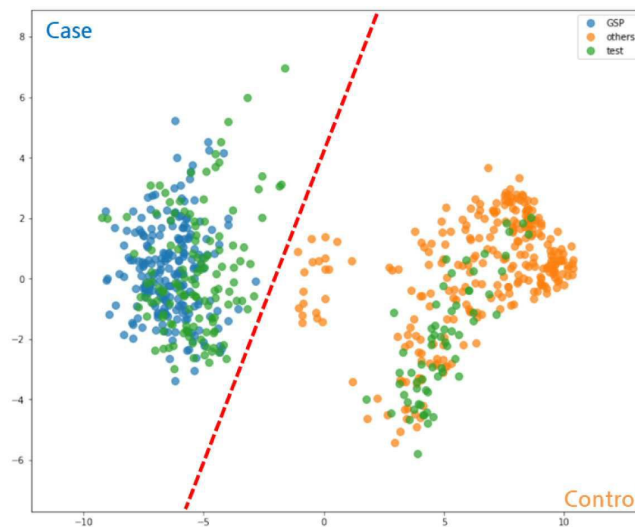
전체 청구항 수 : 총 9 항

(54) 발명의 명칭 토종닭의 유전적 배경 또는 품종을 판별하기 위한 SNP 마커 세트 및 이의 용도

(57) 요약

본 발명은 토종닭의 유전적 배경 또는 품종을 판별하기 위한 SNP 마커 조성물 및 이의 용도에 관한 것으로, 본 발명의 SNP 마커 조성물은 신품종 토종닭을 정확하게 식별하기 위한 품종 특이적이고 최적화된 SNP 마커의 조합으로, 토종닭의 유통과정에서 부정행위를 방지하는 수단으로 활용될 수 있으며, 토종닭에 대한 소비자의 신뢰도를 상승시켜 가금 생산 농가의 소득 증대에 도움을 줄 수 있을 것으로 기대된다.

대표도 - 도4



(72) 발명자

서동원

세종특별자치시 장군면 외암길 17

김형용

경기도 안양시 동안구 관악대로 135, 126동 1006호(비산동, 비산 삼성 래미안)

이 발명을 지원한 국가연구개발사업

과제고유번호 213010-05-2-SB250

부처명 농림축산식품부

연구관리전문기관 농림식품기술기획평가원

연구사업명 골든씨드프로젝트

연구과제명 토종닭 개체 및 집단식별 분자표지 개발과 검증체계 구축

기 여 율 1/1

주관기관 충남대학교 산학협력단

연구기간 2017.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

서열번호 14, 18, 25, 26, 30, 37, 38, 39, 41, 46, 52, 54, 56, 57, 62, 63, 64, 67, 68, 69, 71, 73, 74, 78, 79, 82, 85, 86, 92, 93, 95, 97, 98, 100, 103 및 105의 염기서열로 이루어진 폴리뉴클레오티드에 있어서, 각각의 염기서열 중 36번째에 위치한 SNP(single nucleotide polymorphism) 염기를 포함하는 8개 이상의 연속된 뉴클레오티드로 구성된 폴리뉴클레오티드 또는 이의 상보적인 폴리뉴클레오티드를 포함하는, 토종닭의 유전적 배경 또는 육계 신품종을 판별하기 위한 SNP 마커 조성물.

청구항 2

제1항에 있어서,

서열번호 1 내지 13, 15 내지 17, 19 내지 24, 27 내지 29, 31 내지 36, 40, 42 내지 45, 47 내지 51, 53, 55, 58 내지 61, 65, 66, 70, 72, 75 내지 77, 80, 81, 83, 84, 87 내지 91, 94, 96, 99, 101, 102, 104, 106 및 107의 염기서열로 이루어진 폴리뉴클레오티드에 있어서, 각각의 염기서열 중 36번째에 위치한 SNP(single nucleotide polymorphism) 염기를 포함하는 8개 이상의 연속된 뉴클레오티드로 구성된 폴리뉴클레오티드 및 이의 상보적인 폴리뉴클레오티드로 이루어진 군으로부터 선택되는 하나 이상의 폴리뉴클레오티드를 추가로 포함하는 것을 특징으로 하는, 토종닭의 유전적 배경 또는 육계 신품종을 판별하기 위한 SNP 마커 조성물.

청구항 3

제1항 또는 제2항에 있어서, 상기 연속된 뉴클레오티드는 8개 내지 100개의 연속된 뉴클레오티드인 것을 특징으로 하는 SNP 마커 조성물.

청구항 4

제1항 또는 제2항에 있어서, 상기 토종닭은 한협 품종 H, F 또는 Y인 것을 특징으로 하는 SNP 마커 조성물.

청구항 5

제1항 또는 제2항에 기재된 SNP(single nucleotide polymorphism) 염기를 포함하는 폴리뉴클레오티드 또는 이의 cDNA를 포함하는, 토종닭의 유전적 배경 또는 육계 신품종 판별용 마이크로어레이.

청구항 6

토종닭 의심 개체 또는 육계 신품종 개체에서 게놈 DNA를 분리하는 단계; 및

상기 분리된 게놈 DNA에서 제1항 또는 제2항에 기재된 폴리뉴클레오티드의 SNP(single nucleotide polymorphism) 위치 염기의 유전자형을 결정하는 단계;를 포함하는, 토종닭 또는 육계 신품종의 유전적 배경 또는 품종 판별 방법.

청구항 7

제6항에 있어서, 상기 토종닭은 한협 품종 H, F 또는 Y인 것을 특징으로 하는 토종닭 또는 육계 신품종의 유전적 배경 또는 품종 판별 방법.

청구항 8

제1항 또는 제2항에 기재된 SNP(single nucleotide polymorphism) 염기를 포함하는 폴리뉴클레오티드를 증폭하기 위한, 토종닭 또는 육계 신품종의 유전적 배경 또는 품종 판별용 프라이머 세트.

청구항 9

제8항의 프라이머 세트; 및 증폭 반응을 수행하기 위한 시약을 포함하는, 토종닭 또는 육계 신품종의 유전적 배

경 또는 품종을 판별하기 위한 키트.

발명의 설명

기술 분야

[0001] 본 발명은 토종닭의 유전적 배경 또는 품종을 판별하기 위한 SNP 마커 세트 및 이의 용도에 관한 것이다.

배경 기술

[0002] 가금육은 우리나라 전체 육류생산량의 20% 이상을 차지하여 농가 소득 및 농촌 경제에 중요한 역할을 담당하고 있다. 그 중에서 가장 높은 비율을 차지하는 닭은 그 종자를 대부분 축산선진국(미국, 영국, 프랑스, 독일, 덴마크 등)으로부터 수입에 의존하고 있는 실정이다. 또한, 소(2008년), 돼지(2014년)의 경우 생산이력제 사업이 추진되어 국내종과 수입종의 구분으로 신뢰성을 확보해 나가고 있는 반면, 재래닭의 경우 보존 및 개발이 많이 이루어져 있지 않은 상태에 있고 그 명칭이나 구분이 명확하지 않아 국내종으로서의 신뢰성 확보가 시급한 실정이다. 농림축산식품부는 닭, 오리고기, 계란 등 가금산물의 유통경로를 소비자들이 한눈에 확인할 수 있도록 2019년 하반기 '가금 및 가금산물 이력제' 도입을 목표로 2018년 11월 시범사업을 추진하는 계획을 발표하였다.

[0003] 한국 정부는 가금육의 수입 의존도를 낮추고 미래농업을 선도하는 종자강국 실현을 위해 토종닭 품종을 개발하는 골든씨드 프로젝트(Golden Seed Project)를 기획하였다. 상기 프로젝트를 통해 개발되는 토종닭 신품종은 국내시장 점유율 30% 이상, 해외시장 수출 100만 달러의 목표로 진행되고 있다. 따라서, 앞으로 개발될 토종닭 품종의 경우 유전자형을 이용한 개체식별 방법을 확립하여 신속하고 정확하게 식별할 수 있는 방법이 필요하다. 그러나, 닭의 유전자 상에서 개체 식별이 가능한 도메인(domain)은 닭의 품종에 따라 다양하게 나타나기 때문에, 닭의 개체 식별을 위해서는 닭의 특이적인 유전양상에 근거한 표지 유전자를 선정하고, 이들을 활용한 유전자 감식기법을 설정하는 것이 중요하다.

[0004] 한편, 한국등록특허 제1751932호에는 한국산 재래닭의 일당 증체량에 대한 유전능력을 조기에 예측하고 식별할 수 있는 '신규한 DNA 표지인자 및 이를 이용한 선별방법'이 개시되어 있고, 한국공개특허 제2018-0050470호에는 '재래닭 개체 식별용 초위성체 마커 및 이를 이용한 재래닭의 개체 식별 방법'이 개시되어 있으나, 본 발명의 토종닭의 유전적 배경 또는 품종을 판별하기 위한 SNP 마커 및 이의 용도에 대해서는 기재된 바가 없다.

발명의 내용

해결하려는 과제

[0005] 본 발명은 상기와 같은 요구에 의해 도출된 것으로서, 본 발명자들은 신품종 토종닭을 정확하게 식별하기 위한 품종 특이적이고 최적화된 SNP 마커의 조합을 개발하기 위해서, 한협의 토종닭 9계통, 국립축산과학원의 토종닭 6계통, 및 육용 및 산란 실용계 5계통의 총 283마리 닭의 유전자형 정보를 600K 고밀도 SNP 칩을 이용하여 획득하였고, 획득된 유전자형의 SNP들의 정보를 바탕으로 주성분 분석(Principal Coordinates Analysis, PcoA)을 실시한 결과, 일부 품종을 제외하고는 대부분의 품종이 클러스터를 형성하며 구분이 잘 이뤄지는 것을 확인하였다. 또한, 본 발명자는 골든씨드프로젝트에서 신품종 토종닭을 생산하기 위해 주로 사용되는 한협의 HH, HF 및 HY 계통을 명확하게 구분할 수 있는 SNP 조합을 선별하기 위해 총 107개의 SNP를 선별하였고, 선별된 107개의 SNP 조합을 이용하여 600K 칩으로 유전자형 분석을 수행한 동일한 샘플로 다시 주성분 분석을 수행한 결과, 각 집단의 구분이 잘 되는 것을 확인함으로써, 본 발명을 완성하였다.

과제의 해결 수단

[0006] 상기 과제를 해결하기 위해, 본 발명은 서열번호 14, 18, 25, 26, 30, 37, 38, 39, 41, 46, 52, 54, 56, 57, 62, 63, 64, 67, 68, 69, 71, 73, 74, 78, 79, 82, 85, 86, 92, 93, 95, 97, 98, 100, 103 및 105의 염기서열로 이루어진 폴리뉴클레오티드에 있어서, 각각의 염기서열 중 36번째에 위치한 SNP(single nucleotide polymorphism) 염기를 포함하는 8개 이상의 연속된 뉴클레오티드로 구성된 폴리뉴클레오티드 또는 이의 상보적인 폴리뉴클레오티드를 포함하는, 토종닭의 유전적 배경 또는 육계 신품종을 판별하기 위한 SNP 마커 조성물을 제공한다.

[0007] 또한, 본 발명은 상기 SNP(single nucleotide polymorphism) 염기를 포함하는 폴리뉴클레오티드 또는 이의 cDNA를 포함하는, 토종닭의 유전적 배경 또는 육계 신품종 판별용 마이크로어레이를 제공한다.

- [0008] 또한, 본 발명은 토종닭 의심 개체 또는 육계 신품종 개체에서 게놈 DNA를 분리하는 단계; 및 상기 분리된 게놈 DNA에서 본 발명에 따른 폴리뉴클레오티드의 SNP(single nucleotide polymorphism) 위치 염기의 유전자형을 결정하는 단계;를 포함하는, 토종닭 또는 육계 신품종의 유전적 배경 또는 품종 판별 방법을 제공한다.
- [0009] 또한, 본 발명은 상기 SNP(single nucleotide polymorphism) 염기를 포함하는 폴리뉴클레오티드를 증폭하기 위한, 토종닭 또는 육계 신품종의 유전적 배경 또는 품종 판별용 프라이머 세트를 제공한다.
- [0010] 또한, 본 발명은 상기 프라이머 세트; 및 증폭 반응을 수행하기 위한 시약을 포함하는, 토종닭 또는 육계 신품종의 유전적 배경 또는 품종을 판별하기 위한 키트를 제공한다.

발명의 효과

- [0011] 본 발명에 따른 SNP 마커 조성물은 토종닭의 유통과정에서 발생하는 부정행위를 방지하는 수단으로 활용될 수 있으며, 토종닭 또는 육계 신품종에 대한 유전적 배경을 정확하게 판별할 수 있으므로, 토종닭에 대한 소비자의 신뢰도를 상승시켜 가금 생산 농가의 소득 증대에 도움을 줄 수 있을 것으로 기대된다.

도면의 간단한 설명

- [0012] 도 1은 600K 고밀도 SNP 칩을 이용하여 획득된 SNP들의 유전자형 정보를 바탕으로 닭 283마리의 주성분 분석(Principal Coordinates Analysis)을 실시한 결과이다.
- 도 2는 선발된 107개의 SNP 조합을 이용하여 600K 칩으로 유전자형 분석을 수행한 동일한 샘플로 다시 주성분 분석을 수행한 결과로, Case는 HH, HF, HY 품종을 의미하고, Control은 다른 계통의 품종을 의미한다.
- 도 3은 선발된 107개의 SNP들의 식별능력을 시뮬레이션하기 위해 확보된 유전자형을 이용하여 가상의 자손을 생성하여 주성분 분석을 수행한 결과로, 노란색은 Control 그룹을 파란색은 Case 그룹을 의미한다.
- 도 4는 가상의 집단과 실제 집단 182마리의 유전자형 정보를 이용하여 MDS(Multidimensional scaling) 분석을 실시한 결과이다.
- 도 5는 MDS 분석을 통해 얻은 분산 정보를 바탕으로 기존의 600K 데이터에서 획득한 283 수와 가상자손 488 수를 포함한 총 771 수의 데이터를 학습 데이터 세트로 사용하였고, 새롭게 획득한 192 수의 유전자형 정보를 테스트 데이터 세트로 사용하여 기계학습 모델별 ROC(Receiver operating characteristic) 커브를 도출한 결과이다.
- 도 6는 마커 조합의 최적화를 위해 전체 107개의 SNP 중 품종 구분에 실질적으로 영향을 주는 SNP를 각 기계학습의 모델별 특징 중요도(Feature Importance; FI)에 기초하여 선발한 결과이다.

발명을 실시하기 위한 구체적인 내용

- [0013] 본 발명의 목적을 달성하기 위하여, 본 발명은 서열번호 14, 18, 25, 26, 30, 37, 38, 39, 41, 46, 52, 54, 56, 57, 62, 63, 64, 67, 68, 69, 71, 73, 74, 78, 79, 82, 85, 86, 92, 93, 95, 97, 98, 100, 103 및 105의 염기서열로 이루어진 폴리뉴클레오티드에 있어서, 각각의 염기서열 중 36번째에 위치한 SNP(single nucleotide polymorphism) 염기를 포함하는 8개 이상의 연속된 뉴클레오티드로 구성된 폴리뉴클레오티드 또는 이의 상보적인 폴리뉴클레오티드를 포함하는, 토종닭의 유전적 배경 또는 육계 신품종을 판별하기 위한 SNP 마커 조성물을 제공한다.
- [0014] 또한, 본 발명은 상기 SNP 마커 조성물에 서열번호 1 내지 13, 15 내지 17, 19 내지 24, 27 내지 29, 31 내지 36, 40, 42 내지 45, 47 내지 51, 53, 55, 58 내지 61, 65, 66, 70, 72, 75 내지 77, 80, 81, 83, 84, 87 내지 91, 94, 96, 99, 101, 102, 104, 106 및 107의 염기서열로 이루어진 폴리뉴클레오티드에 있어서, 각각의 염기서열 중 36번째에 위치한 SNP 염기를 포함하는 8개 이상의 연속된 뉴클레오티드로 구성된 폴리뉴클레오티드 및 이의 상보적인 폴리뉴클레오티드로 이루어진 군으로부터 선택되는 하나 이상의 폴리뉴클레오티드를 추가로 포함하는, 토종닭의 유전적 배경 또는 육계 신품종을 판별하기 위한 SNP 마커 조성물을 제공한다.
- [0015] 본 발명의 SNP 마커 조성물에 있어서, 상기 연속된 뉴클레오티드는 8 내지 100개의 연속된 뉴클레오티드일 수 있으나, 이에 제한되지 않는다.
- [0016] 본 명세서에서 용어, "뉴클레오티드"는 단일 가닥 또는 이중 가닥 형태로 존재하는 디옥시리보뉴클레오티드 또는 리보뉴클레오티드이며, 특별하게 다르게 언급되어 있지 않은 한 자연의 뉴클레오티드의 유사체를 포함한다.

- [0017] 용어 "다형성(polymorphism)"이란 같은 종의 생물이라도 모습이나 고유한 특징이 다양하게 나타나는 것 또는 하나의 유전자 좌(locus)에 두 가지 이상의 대립 유전자(allele)가 존재하는 경우를 말하며, 다형성 부위 중에서 개체에 따라 단일 염기만 다른 것을 "단일염기다형성(single nucleotide polymorphism, SNP)"이라 한다. 우리나라 고유 유전자원에 대한 연구와 유전자원의 보전을 위해 단일염기다형성을 이용한 SNP DNA 분석법을 이용한 다양한 연구가 수행되고 있으나, 토종닭과 관련된 연구는 제한적으로 진행되고 있는 실정이다. 상기 다형성을 확인하기 위한 다형성 마커는 선택된 집단에서 1% 이상, 바람직하게는 5% 이상 또는 10% 이상의 발생빈도를 나타내는 두 가지 이상의 대립유전자를 가지는 것일 수 있다.
- [0018] 본 명세서에서 용어 "토종닭"은 순수 혈통의 재래닭과, 외국에서 유래하였으나 도입 경위가 명확하고 개량을 거쳐 최소 7세대 이상 우리나라의 기후와 풍토에 안정적으로 정착한 품종을 포함하는 의미이다.
- [0019] 본 발명의 일 구현 예에 따른 SNP 마커 조성물에 있어서, 상기 토종닭은 한협 품종 H, F 또는 Y 일 수 있으나, 이에 제한되지 않으며, 상기 육계 신품종은 한협 품종 H, F 또는 Y를 이용하여 육종한 품종 혹은 자손 개체일 수 있다. 상기 한협 품종 H, F 또는 Y는 GSP(golden seed project) 신품종 토종닭 개발의 핵심계통으로 활용되고 있는 집단으로, 본 발명에 따른 SNP 마커 조성물은 한협 품종 H, F 또는 Y와 다른 닭 품종간의 유전형을 높은 정확도로 구별할 수 있으므로, 토종닭 자원에 대한 권리확보 및 보호에 유용하게 사용될 수 있을 것이다.
- [0020] 본 발명의 일 구현 예에 따른 SNP 마커 조성물에 있어서, 상기 SNP 위치 염기는 서열번호 1 내지 107의 염기서열 모두 36번째로, 다형성 염기 정보는 표 2 내지 5의 SNP 염기서열 정보에 [/]로 표시하였다.
- [0021] 본 발명에 따른 SNP 조성물은 서열번호 14, 18, 25, 26, 30, 37, 38, 39, 41, 46, 52, 54, 56, 57, 62, 63, 64, 67, 68, 69, 71, 73, 74, 78, 79, 82, 85, 86, 92, 93, 95, 97, 98, 100, 103 및 105의 염기서열로 이루어진 36개의 폴리뉴클레오티드를 최소 마커 조합으로 포함할 수 있고, 상기 외에 서열번호 1 내지 13, 15 내지 17, 19 내지 24, 27 내지 29, 31 내지 36, 40, 42 내지 45, 47 내지 51, 53, 55, 58 내지 61, 65, 66, 70, 72, 75 내지 77, 80, 81, 83, 84, 87 내지 91, 94, 96, 99, 101, 102, 104, 106 및 107의 염기서열로 이루어진 폴리뉴클레오티드로 이루어진 군으로부터 선택되는 하나 이상의 폴리뉴클레오티드를 추가로 포함할 수 있다. 상기 36개의 폴리뉴클레오티드로 이루어진 SNP 마커 조합은 특징 중요도(Feature Importance; FI)를 기준으로 선발되었으며 AdaBoost 기계 학습 모델에서 토종닭과 그 외의 닭 품종을 높은 정확도로 구별할 수 있는 최소의 마커 조합이다.
- [0022] 본 발명은 또한, 상술한 SNP(single nucleotide polymorphism) 염기를 포함하는 폴리뉴클레오티드 또는 이의 cDNA를 포함하는, 토종닭의 유전적 배경 또는 육계 신품종 판별용 마이크로어레이를 제공한다. 상기 폴리뉴클레오티드는 바람직하게는, 아미노-실란(amino-silane), 폴리-L-라이신 또는 알데히드 활성기가 코팅된 기판에 고정될 수 있으나, 이에 제한되지는 않는다. 상기 폴리뉴클레오티드를 기판에 고정화시키는 방법으로는 피에조일렉트릭(piezoelectric) 방식을 이용한 마이크로피펫팅(micropipetting), 핀(pin) 형태의 썬터(spotter)를 이용한 방법 등을 사용할 수 있으나, 이에 한정되지 않으며, 당업계에 공지된 다양한 방법을 이용할 수 있다. 또한, 상기 기판은 실리콘 웨이퍼(silicon wafer), 유리, 석영(quartz), 금속 또는 플라스틱일 수 있으나, 이에 제한되지는 않는다.
- [0023] 본 발명에 따른 마이크로어레이는 전술한 것과 같이 염기서열 내에 SNP 염기를 포함하는 서열번호 14, 18, 25, 26, 30, 37, 38, 39, 41, 46, 52, 54, 56, 57, 62, 63, 64, 67, 68, 69, 71, 73, 74, 78, 79, 82, 85, 86, 92, 93, 95, 97, 98, 100, 103 및 105의 염기서열로 이루어진 36개의 폴리뉴클레오티드를 최소 조합으로 하여 구성될 수도 있고, 상기 36개의 폴리뉴클레오티드에 서열번호 1 내지 13, 15 내지 17, 19 내지 24, 27 내지 29, 31 내지 36, 40, 42 내지 45, 47 내지 51, 53, 55, 58 내지 61, 65, 66, 70, 72, 75 내지 77, 80, 81, 83, 84, 87 내지 91, 94, 96, 99, 101, 102, 104, 106 및 107의 염기서열로 이루어진 SNP 염기를 포함하는 폴리뉴클레오티드로 이루어진 군으로부터 선택되는 하나 이상의 폴리뉴클레오티드를 추가로 포함하여 구성될 수도 있다. 상기 폴리뉴클레오티드는 토종닭인 한협 품종 H, F 또는 Y와 다른 닭 품종간에 다형성을 나타내는 SNP 염기를 포함하고 있으므로, 본 발명에 따른 마이크로어레이는 토종닭 특히, 한협 품종 H, F 또는 Y를 다른 닭 품종들로부터 판별하는데 유용하게 활용될 수 있다.
- [0024] 본 발명은 또한, 토종닭 의심 개체 또는 육계 신품종 개체에서 게놈 DNA를 분리하는 단계; 및 상기 분리된 게놈 DNA에서 본 발명에 따른 폴리뉴클레오티드의 SNP(single nucleotide polymorphism) 위치 염기의 유전자형을 결정하는 단계;를 포함하는, 토종닭 또는 육계 신품종의 유전적 배경 또는 품종 판별 방법을 제공한다.
- [0025] 본 발명의 토종닭 또는 육계 신품종의 유전적 배경 또는 품종 방법에 있어서, 상기 SNP 위치의 유전자형은 표 2

내지 5에 개시된 것과 같다.

- [0026] 본 발명의 토종닭 또는 육계 신품종의 유전적 배경 또는 품종 판별 방법은 피검체(토종닭 의심 개체 또는 육계 신품종 개체)로부터 분리한 게놈 DNA에서 유전자형의 결정을 통해, 피검체가 한협 품종 H, F 또는 Y인지, 혹은 한협 품종 H, F 또는 Y를 이용하여 육성된 품종인지를 판별할 수 있게 되는 것이다.
- [0027] 본 발명의 방법에 있어서, 피검체(토종닭 의심 개체 또는 육계 신품종 개체)로부터 게놈 DNA를 분리하는 방법은 당업계에 알려진 통상적인 방법을 통하여 이루어질 수 있다. 예를 들면, 조직 또는 세포로부터 DNA를 직접적으로 정제하거나 PCR과 같은 증폭 방법을 사용하여 특정한 영역을 특이적으로 증폭하고 이를 분리함으로써 이루어질 수 있다. 본 발명에 있어서, DNA란 DNA 뿐만 아니라 mRNA로부터 합성되는 cDNA도 포함한다. 피검체로부터 핵산을 얻는 단계는 예를 들면, PCR 증폭법, 리가제 연쇄 반응(ligase chain reaction), 전사증폭(transcription amplification), 자가유지 서열복제(self-sustained sequence replication system; Guatelli 등, Proc. Natl. Acad. Sci. USA (1990) 87:1874-1878) 및 핵산 서열 기재 증폭(nucleic acid sequence-based amplification)이 사용될 수 있으나, 이에 제한되지는 않는다.
- [0028] 분리된 DNA의 염기서열의 분석은 당업계에 알려진 다양한 방법에 의하여 이루어질 수 있다. 예를 들면, 디데옥시법에 의한 직접적인 핵산의 뉴클레오타이드 서열의 결정을 통하여 이루어지거나, SNP 부위의 서열을 포함하는 프로브 또는 그에 상보적인 프로브를 상기 DNA와 혼성화시키고 그로부터 얻어지는 혼성화 정도를 측정함으로써 다형성 부위의 뉴클레오타이드 서열을 결정/분석하는 방법 등이 이용될 수 있으나, 이에 제한되지는 않는다. 상기 혼성화의 정도는 예를 들면, 검출가능한 표지를 표적 DNA에 표지하여, 혼성화된 표적 DNA 만을 특이적으로 검출함으로써 이루어질 수 있으며, 그외 전기적 신호 검출방법 등이 사용될 수 있으나, 이에 제한되지는 않는다.
- [0029] 상기 토종닭 의심 개체 또는 육계 신품종 개체로부터 분리한 핵산 시료를 본 발명에 따른 SNP를 포함하는 폴리뉴클레오타이드 또는 이의 상보적 폴리뉴클레오타이드, 또는 이와 혼성화하는 폴리뉴클레오타이드와 혼성화시킨 후 혼성화 결과를 검출하는 단계를 포함할 수 있다.
- [0030] 본 명세서에서 용어 "프로브(probe)"는 타겟 핵산 서열에 실질적으로 상보적인 부위 또는 부위들을 포함하는 단일-가닥 핵산 분자를 의미한다. 또한, 용어 "혼성화(hybridization)"는 상보적인 단일 가닥 핵산들이 이중-가닥 핵산을 형성하는 것을 의미한다. 혼성화는 완전히 매칭되거나 일부 미스매치로 실질적으로 매칭되는 2개의 핵산 가닥 사이에서 일어날 수 있다. 혼성화를 위한 상보성은 혼성화 조건, 특히 온도에 따라 달라질 수 있다.
- [0031] 또한, 본 발명에 따른 토종닭의 유전적 배경 또는 품종 판별 방법은 각 개체의 유전자형 분석 결과를 기계학습 모델을 통해 정확도 및 특이도를 검증할 수 있고, 상기 기계학습 모델은 AdaBoost, Decision Tree 또는 Random Forest 모델일 수 있으나, 이에 제한되지 않는다.
- [0032] 본 발명은 또한, 본 발명에 따른 SNP(single nucleotide polymorphism) 염기를 포함하는 폴리뉴클레오타이드를 증폭하기 위한, 토종닭 또는 육계 신품종의 유전적 배경 또는 품종 판별용 프라이머 세트를 제공한다.
- [0033] 본 발명의 프라이머 세트에 있어서, 상기 SNP 염기를 포함하는 폴리뉴클레오타이드는 서열번호 1 내지 107의 염기 서열로 이루어진 것일 수 있으며, 자세한 정보는 전술한 것과 같다.
- [0034] 본 명세서에서, 용어 "프라이머"는 카피하려는 핵산 가닥에 상보적인 단일 가닥 올리고뉴클레오타이드 서열을 말하며, 프라이머 연장 산물의 합성을 위한 개시점으로서 작용할 수 있다. 상기 프라이머의 길이 및 서열은 연장 산물의 합성을 시작하도록 허용해야 한다. 프라이머의 구체적인 길이 및 서열은 요구되는 DNA 또는 RNA 표적의 복잡도(complexity) 뿐만 아니라 온도 및 이온 강도와 같은 프라이머 이용 조건에 의존할 것이다. 본 발명에 따른 프라이머 세트는 정방향 및 역방향 프라이머가 하나의 프라이머 세트를 이룰 수 있고, 대립유전자(allele) 특이적 정방향 프라이머 2개와 역방향 프라이머 1개가 하나의 프라이머 세트를 이룰 수도 있으며, 또는 ASP(SNPtype assay allele specific primer)1, ASP2, LSP(SNPtype assay locus specific primer) 및 STA(SNPtype assay specific target amplification primer) 프라이머가 하나의 프라이머 세트를 이루는 Fluidigm SNP 유전형 분석용 프라이머 세트일 수 있으나, 이에 제한되지 않는다.
- [0035] 본 발명의 올리고뉴클레오타이드 프라이머 세트는 증폭 반응을 통해 표적 서열을 증폭할 수 있다. 표적 핵산을 증폭하는 방법은 중합효소연쇄반응(polymerase chain reaction, PCR), 리가아제 연쇄반응(ligase chain reaction), 핵산 서열 기재 증폭(nucleic acid sequence-based amplification), 전사 기재 증폭 시스템(transcription-based amplification system), 가닥 치환 증폭(strand displacement amplification) 또는 Q β -복제효소(replicase)를 통한 증폭 또는 당업계에 알려진 핵산 분자를 증폭하기 위한 임의의 기타 적당한 방법이 있다. 이 중에서, PCR이란 중합효소(polymerase)를 이용하여 표적 핵산에 특이적으로 결합하는 프라이머 쌍

으로부터 표적 핵산을 증폭하는 방법이다. 이러한 PCR 방법은 당업계에 잘 알려져 있으며, 상업적으로 이용가능한 키트를 이용할 수도 있다.

[0036] 본 발명은 또한, 상기 프라이머 세트; 및 증폭 반응을 수행하기 위한 시약을 포함하는, 토종닭 또는 육계 신종종의 유전적 배경 또는 품종을 판별하기 위한 키트를 제공한다. 본 발명의 키트에서, 상기 증폭 반응을 수행하기 위한 시약은 DNA 폴리머라제, dNTPs, 버퍼 등을 포함할 수 있다. 또한, 본 발명의 키트는 최적의 반응 수행 조건을 기재한 사용자 안내서를 추가로 포함할 수 있다. 안내서는 키트 사용법, 예를 들면, PCR 완충액 제조 방법, 제시되는 반응 조건 등을 설명하는 인쇄물이다. 안내서는 팜플렛 또는 전단지 형태의 안내 책자, 키트에 부착된 라벨, 및 키트를 포함하는 패키지의 표면에 설명을 포함한다. 또한, 안내서는 인터넷과 같이 전기 매체를 통해 공개되거나 제공되는 정보를 포함한다.

[0037] 이하, 본 발명을 실시예에 의해 상세히 설명한다. 단, 하기 실시예는 본 발명을 예시하는 것일 뿐, 본 발명의 내용이 하기 실시예에 한정되는 것은 아니다.

[0038] 1. SNP 마커 선별을 위한 재료 및 방법

[0039] 공시축은 15계통의 한국 토종닭 집단과 5계통의 실용계 집단으로 구성된 총 283마리의 닭 샘플을 실험에 이용하였다. 한국 토종닭 집단은 한협 종계회사가 보유한 순계 9계통(HH: 23마리, HF: 23마리, HG: 23마리, HS: 23마리, HV: 23마리, HW: 23마리, HA: 20마리, HY: 21마리, HZ: 15마리)과 국립축산과학원(The National Institute of Animal Science, NIAS)에서 제공한 순계 6계통(NC: 6마리, ND: 6마리, NH: 6마리, NS: 6마리, NR: 6마리, NY: 5마리)으로 구성되어 있으며, 실용계 집단은 3종의 Cobb, Arbor Acre, Ross 회사의 육용계(Cobb: 12마리, Ab: 10마리, Ross: 12마리)와 2종의 Hyline brown, Lohman brown 회사의 산란계(HL: 10마리, LO: 10마리)로 구성되어 있다.

[0040] 게노믹 DNA(gDNA)는 PrimePrep™ DNA Isolation 키트(GenetBio, Deajeon, Korea)를 이용하여 공시축의 혈액을 통해 추출하였다. 추출된 gDNA의 품질과 농도는 NanoDrop 분광광도계(Thermo Fisher Scientific, Waltham, MA, USA)를 이용하여 측정하였다. 추출된 DNA는 실험 전까지 -20℃에서 냉동 보관하였다. 상기 확보한 gDNA는 Axiom 600K chicken array SNP chip(Affymetrix, Santa Clara, CA, USA)을 이용하여 580,954개의 SNP(Single Nucleotide Polymorphism)의 유전자형 정보를 확인하였다. 확인된 유전자형 정보는 더욱 정확한 결과 도출을 위해 PLINK 소프트웨어(ver. 1.91, <http://zzz.bwh.harvard.edu/plink/>)를 활용하여 geno 옵션으로 유전자형 분석 오류가 10% 이상인 SNP들을 제거하는 QC(quality control)를 수행하였다. QC 과정을 수행한 SNP 유전자형 정보를 바탕으로 PLINK 소프트웨어를 이용하여 주성분 분석(Principal Coordinates Analysis, PcoA)을 실시하였다. 이를 통해 얻어진 데이터는 R software package를 이용하여 plot을 시각화 하였다.

[0041] 2. SNP 마커 후보 세트 구성

[0042] 600K SNP 칩으로 유전형을 확인한 집단 중에서 GSP(golden seed project) 신종종 토종닭 개발의 핵심계통으로 활용되고 있는 집단은 case 그룹(HH, HF, HY)으로 설정하고 HH, HF 및 HY를 제외한 나머지 집단은 모두 control 그룹으로 설정하여 case 그룹의 특이적인 SNP를 확보하기 위해 두 집단간의 GWAS(Genome-Wide Association Study)를 실시하였다. 분석결과에서 case 및 control 그룹을 유의적으로 구분할 수 있는 SNP들은 chi-square p-value 값에 따라 우선순위를 설정하였다. GWAS 방법을 이용하여 도출한 SNP 순위 중 상위 SNP들의 대부분이 거대 염색체에 포진되어 있는 단점과 집단의 연관불평형(Linkage Disequilibrium, LD) 다양성 정보를 활용하기 위해 LD 블록을 계산하였으며 계산 결과를 토대로 1, 50, 100 LD당 SNP를 선별하여 세트 당 107개의 SNP를 포함한 3종의 SNP 조합 세트(set 1, 2, 3)를 구성하였다.

[0043] 상기 3종의 SNP 조합 세트들의 품종 식별능을 평가하기 위하여 각각의 SNP 조합 세트를 이용하여 PLINK 소프트웨어를 이용하여 주성분 분석(PcoA)을 실시하였다. 그 결과, 도 2에 개시된 것과 같이 50 LD당 SNP를 선별한 set 2의 결과가 최적의 집합양상을 나타내었다. 비록 도 2에서 set 3이 set 2와 비슷한 패턴을 보이는 것처럼 확인되었으나, 이전의 분석에서 마커들 간의 LD 간격이 커질수록 set 2가 set 3에 비해 집단 클러스터링이 더 잘 되는 결과를 보인 바 있어, set 2를 최종적으로 선별하였다. 또한 50 LD당 SNP를 선별한 set 2의 SNP는, 107개 SNP의 대다수가 1번 염색체에 몰려있는 set 1의 결과와 달리 전반적으로 SNP들이 각각의 염색체에 고루 분포하였다(하기 표 1 참고).

[0044] 3. 최적의 SNP 마커 세트 구성

[0045] 조합이 구성된 SNP 세트 중에서 set 2의 효율이 좋은 것을 확인하여 검증실험을 수행하기에 앞서 품종 구분 가

능성을 평가하기 위해 가상의 자손 488마리(F_1 , F_2 및 F_3)를 생성하여 평가하였다. 가상의 자손은 아비(2n)와 어미(2n)로부터 유래한 반수체(n) 대립유전자(allele)들의 임의적 결합을 통해 생성하였다. 가상의 자손 생성시 부모 계통으로 한협 회사의 HH, HF 및 HY 계통이 이용된 경우 case 그룹으로 추가 지정하였고, 그 외의 가상의 자손들은 control 그룹으로 포함되었다. 세트들의 SNP 유전자형 정보를 PLINK 소프트웨어(ver. 1.91)를 이용하여 각각 PcoA를 실시하였다. 추가적으로 PcoA 결과를 바탕으로 집단들을 가장 잘 구분할 수 있는 최적의 분류 모델(classification model)을 확인하기 위해 9가지의 기계학습 알고리즘(Nearest Neighbors, Linear SVM, Radial-Basis-Function SVM, Random Forest, AdaBoost, Naive Bayes, Linear Discriminant Analysis, Quadratic Discriminant Analysis algorithm, Decision tree)을 적용하여 최적의 107개 SNP 조합 세트 및 분류 모델을 판별하였다.

[0046] 선발된 107개 SNP 조합 세트의 SNP들을 이용하여 실제집단을 구분할 수 있는 최적의 SNP 개수를 확인하기 위해 Fluidigm's Biomark 96.96 Dynamic Array(Fluidigm Corporation, CA, USA)를 제작하여 총 182마리의 닭 샘플에 대한 유전자형 분석을 실시하였다. 유전자형 분석은 case 집단으로 이용될 한협 종계회사가 보유한 순계 3계통(HF: 36마리, HH: 36마리, HY: 26마리)과 GSP 신품종 토종 실용계 집단 2종(GSP_CC(HFHY): 10마리, GSP_CC2(FHFY): 10마리) 및 control 집단으로 이용될 시중에 유통 중인 육용계(Ross: 20마리, Cobb: 8마리, Abor Acres: 11마리)와 산란계(Lohmann brown: 5마리), 한국 토종닭 계통의 실용 육용계 3종(WM2: 10마리, Yelim: 5마리, Hyunin: 5마리)을 이용하였다.

[0047] 이어서 SNP 마커 개수에 따른 품종 구분 가능 정도를 확률적으로 계산하였다. 민감도(sensitivity)는 본 발명의 SNP 마커에 의해 test 그룹으로 예상되는 개체가 실제로 test 그룹일 확률의 정확도를 의미한다.

$$\text{Sensitivity} = \frac{\text{true Test group chicken}}{\text{estimate Test group chicken}}$$

[0048]

[0049] 특이도(specificity)는 본 발명의 SNP 마커에 의해 non-test 그룹으로 예상되는 개체가 실제로 non-test 그룹일 확률의 정확도를 의미한다.

$$\text{Specificity} = \frac{\text{true non - Test group chicken}}{\text{estimate non - Test group chicken}}$$

[0050]

[0051] 실시예 1. 토종닭 구분을 위한 SNP 마커의 선별

[0052] 283마리의 SNP 유전자형 정보를 600K 고밀도 SNP 칩을 이용하여 각각 획득하였고, 획득된 SNP들의 유전자형 정보를 바탕으로 PcoA를 실시한 결과 일부 품종을 제외하고는 대부분의 품종이 클러스터를 형성하며 구분이 잘 이뤄지는 것을 확인할 수 있었다(도 1).

[0053] GSP에서 신품종 토종닭을 생산하기 위해 주로 사용되는 품종은 HH, HF, HY의 3개의 계통이 3원교잡의 방법으로 생산되게 되는데, HH, HF, HY계통의 경우는 한협에서 특이적으로 보유하고 있는 재래닭 품종에 해당하며, HH 및 HF 계통은 육용특성의 부계통으로, HY 계통은 산란특성의 모계통으로 활용되고 있어 교배 계획 전반에 활용되고 있을뿐만 아니라, 유전성분과 특성으로도 충분한 독립성을 띄고 있기 때문에 SNP 선발 주요 품종으로 선택하였다. HH, HF와 HY 계통을 명확하게 구분할 수 있는 SNP 조합을 선발하기 위해 유전자형 분석 에러가 10%이상 확인된 SNP를 QC한 유전자형 정보를 통계분석 도구인 PLINK를 이용하여 HH, HF 및 HY를 Case 그룹으로 설정하고, 나머지 토종닭 집단 및 실용계 집단을 모두 Control 그룹으로 설정한 후 Case, Control 집단에 대한 연관 분석을 수행하였다. 그 결과의 유의미한 차이인 X^2 (Chi-Square) p-값을 획득하여 낮은 순서로 정렬한 후 성염색체인 Z 염색체 SNP를 제거하고, Case 집단의 특이 동형접합 좌위를 선발하였다. 또한, 유전체 전체에서 SNP를 고르게 선발하고자 집단의 연관불평형(Linkage Disequilibrium; LD) 블록 정보를 분석하여 LD 블록 50개당 하나의 마커를 선발한 결과(set 2), 1번 염색체에서 37개, 2번 및 5번 염색체에서 각각 10개, 6번 염색체에서 8개, 3번 및 15번 염색체에서 각각 7개, 24번 염색체에서 6개, 10번 염색체에서 5개, 20번 및 26번 염색체에서 각각 4개, 7번 및 18번 염색체에서 각각 2개, 11번, 13번, 22번, 25번 및 28번 염색체에서 각각 1개씩 총 107개의 SNP 조합을 얻을 수 있었다. 선발된 107개의 SNP 조합을 이용하여 600K 칩으로 유전자형 분석을 수행한 동일한 샘플로 다시 주성분 분석을 수행한 결과 Case 집단 구분이 잘 되는 것을 확인할 수 있었다(도 2).

[0054]

Set2의 SNP 정보

염색체 번호	SNP 수	염색체 번호	SNP 수
1	37	26	4
2	10	7	2
5	10	18	2
6	8	11	1
3	7	13	1
15	7	22	1
24	6	25	1
10	5	28	1
20	4		

표 2

107개의 SNP 마커

CHR	SNP	Physical Position	Strand	Flank (서열번호)
1	AX-75508404	6881076	+	GAAAGCAACTTTGCTTTGGTGGTTGTTATTCTCA/C/TTACAAAGGATGGCTATTTCTTTCTGGTGTAGC (1)
1	AX-75428829	32560632	+	AGAAACTCTCACCAGTCATCAGCTACTTTTCTTC/C/TTACAGGTGCTGTGATATCCAGTTTGGCTTCTGTG (2)
1	AX-75429083	32697379	+	TACGTAAGAACCATTAAAGGAAGAAAGGAAGAA/C/GTGAAGAAAGACTTCGGCATGTTTGGCCCAATTCTCAAGA (3)
1	AX-75441489	38348135	+	TTGTACTGAGCAAGACTTTGGCCAAGACCTTCTTA/C/GTTCCTTACTGTTTGTAAAGACTGCAATGATGTGG (4)
1	AX-75441731	38464663	+	AAGACTCTACTTTGTTTATTAACCAAGACTTAAGCATTA/GTAGAGTGCAATACCTGCTGCATCACAATAACAGGGA (5)
1	AX-75441745	38471715	+	TAGTATTGTTGATGGCCATTATTCGCCAGTACTGGTA/G/CTGATTATTGTGCATAGCCGTCCACACAGAGGA (6)
1	AX-75443426	39248473	+	TGTTCTGTGCTCTCGGGCAACAGATGTTTCTCTG/C/TTCCATGATAGTCAAGTTTTTTAGTTTCAATTTTT (7)
1	AX-75445847	40430277	+	GCCTAAGTATCATCTTTTATAGTGTGAGGTTGCTTA/G/CTAGTGTGCAAGAACTGCCAATTAGTACTGTGCTT (8)
1	AX-75445919	40467557	+	CTCAAGTATAGGTCATGCTTTATATTACTTTAA/C/ATAGGTATGTTGAAGTTTCTCACTTTATTTCCGG (9)
1	AX-75446390	40710381	+	TTACGCAAGAGCTGTTGAACAGCTGTTTATTAATAA/C/GTTGACCACTCCTGTTACAGCACAGAAAAAAGGGT (10)
1	AX-75446448	40738206	+	CATTGCACAGCTCCCAATCAATGTGATGCTGA/C/TTGGCCCTCAAAAGCTTCAGGGTGCCCTCCCAACC (11)
1	AX-75446643	40838785	+	GCAGTTTACTTTTGGCAAGGAAGAACCCAGTAAGGA/C/G/ATGAGCACACTGTGTGGGAAGGACTGGTACAGTGGC (12)
1	AX-75446774	40897183	+	CACATCTTTCAGTTGATTTGTTGTTTGGAAAGGA/C/TTAGAGAAGCAAAATG6AAAGTCAATTTGGACAGTGT (13)
1	AX-75446836	40926139	+	ATGCTTTATAGGTTCTTCAATCTTATGCTTCAG/C/TTGATTGTTTCTCAACTTTTCTACTAGCTATAAG (14)
1	AX-75447593	41321969	+	TTTCAATAAAATCTCTGCTTGAATGTAGTGTCTTG/C/TTGAATCTAATAGTTGGGATTGTCTATAAGTGCAT (15)
1	AX-75450008	42616585	+	TACCTGAGAGTAAACATTTCTTACAGATAATGCA/C/TTTGTGTAAGAGAGTCTTCTAGAAACGTAATCTTA (16)
1	AX-75450200	42700221	+	TAATTAGGTAAAAATCATCTCCAGTAGGATACTTA/G/TAAGTGTGCTGCAATGCTGCAATAAAGGCTGACAC (17)
1	AX-75450269	42731066	+	AGGCTGAGGAGTGAATTTTCAAGTTTCCACAGAA/C/TTATTATTTTACATGCTACTGTCTATTGCAAGATTA (18)
1	AX-75450652	42909573	+	CTATTTTAGTACTTCAAGAGAGATGTTAGAGAA/C/GTACTGGATTAAATTTTCTGTGTAGAGTGTCA (19)
1	AX-75450664	42913912	+	TCATTTTAATATGCTGAGTGGCCACAATGTGTAA/C/GTTCCTTCTCTTGCAATCCACACCACTTGTCTCTCT (20)
1	AX-75451181	43155104	+	ATAAGTATTTCCCAAGCAGAGAAAGCAGTAAATTA/C/GTATGCTTGACCTATTCCAGGTAACAGTTAAACCTA (21)
1	AX-75451215	43171485	+	GGCAAGAGACCCAGCCATGCACTTTGTTTATTTAGTA/GTAGAGTGAGAGACTGAATGGCAAGCAAGGCTTATT (22)
1	AX-75451445	43263769	+	TTCTTTAGATGACTCATTTTCTAGGTTTATCAGCA/C/TACAGTGGCCACCACTTCAATGACTAGTAAACAACAA (23)
1	AX-75452996	43997714	+	TGTGGCTGCATCTGCTGGGATGCTGCTAAAGATTATG/TTATGAGCCTTTGGCAAGCTTCAATGCAATTAAT (24)
1	AX-75458405	46632353	+	TTATTTTGAATCAGGTTCTTATTTCCATCAGAGCTG/C/TTGTCAAGTGTGCTGCTGCTGCAAGCTAGCTGTGCTAC (25)
1	AX-75459486	47136098	+	GAAAGCAACTTTGCTTTTGGTGGTTGTTATTCTCA/C/TTACAAAGGATGTTGCTAATTTCTTTCTGGTTAGC (26)
1	AX-80870454	48341664	+	CAAGCTAGAACTCTGTAACCTCATCAGCAATCTCTTA/TTATTGTCAAGCCCATGGCAGAAATCAACCTTAAGA (27)

[0055]

표 3

107개의 SNP 마커 (continued)

CHR	SNP	Physical Position	Strand	Flank (서열번호)
1	AX-75474264	54082924	+	ACGGAAAAAGAACTGTTATTTTATCTGACAGAGC/A/GTACAGTTTTTACTGGGATCAAGCACACAGAATAGA (28)
1	AX-75541721	86899349	+	GTGTTGACCTTGCTTCTCTGCTTGAAGTCACTC/TTAGAACTTGCTGATGAGCTTCTGCAAGGACTCTT (29)
1	AX-75541725	86901036	+	CTTAGCATGAATTGAGAAAAATGAGATACTTTTTTG/C/TTTCCATTCTCAGTCAGTTTTACAGCTCTGCTCCCT (30)
1	AX-75207417	108589118	+	GAATTAGCAATTTCTATGTTACCAGATGTAAGTACC/A/GJAATCATGGAAAGTTACTTTCTTTAGATGTTAGCT (31)
1	AX-75207430	108595746	+	TGTATTCAACCCAGTTGATGCAACCATCTTGCCAG/C/TTGTCACTGATTACAAAAGTGGACACATGTTGCCCTC (32)
1	AX-75207482	108624993	+	CTTATCCACTACTGATGAGGACTGAAGCATACTCT/A/GTGTGTTTGCAGGATCTAATACGCTGATCTTATTTT (33)
1	AX-75226732	117702156	+	TGGTACTACCATTTATGTTGCTACAAAAAATAA/C/A/CACCTTCATAGCTGTAAGTACTGACCTTATAGTGCC (34)
1	AX-75377378	186339052	+	GGTATGTTTTCAAGCTTTTCTCAGACACTTGAAAAG/C/TTAAGATAGGGTGGAGTGGGGCTTCTGCAATTGTGTG (35)
1	AX-75377768	186516318	+	ATCCAAGGCTTCTGTAACATGATGAAGACGGATAGC/A/GTCTACATCAGTATAATCACACCATTTGAAGTAGTTAC (36)
1	AX-75379158	187002396	+	TTACAGATGTGTAATGGTTTCCAAAGGAATGGGA/A/GJAGTTGACACAAAATGCCATGCTGCTGCACTAC (37)
2	AX-76046791	16801291	+	TTTTCTGACAGAGAGCAGTTAGTAGTCCACTTTG/TTGTAAAACTTTCTGCAAGTCAATTTCTTAAGTATGCT (38)
2	AX-76060521	24681889	+	GAATCTGTGTTGAAGATGAGCAAAACTTGAGATG/A/GJCCAAACTACGGTCAATCTTGCTGGCTGTGG (39)
2	AX-76073577	31602596	+	CGAGCAGCAGCTGGAGTTTCAAAATGTTATGTAC/A/GJTAAAGCAATGACACTGAAGTTACAATGTAAGATAC (40)
2	AX-76076543	33597664	+	GGCACATCAGCATTTGATTTCTTGTTATATTTTCA/A/GJTAACCTTTTCAACTTCTCTGCAATTTATTTCTGCC (41)
2	AX-76103350	49121225	+	CAGAACACTTTATGATTAACCAAGAAATTTCTG/C/TTGATTTCTGGAAGTCTGAATTAACCAAGACTTA (42)
2	AX-80984170	63960038	+	AGATGATTCATGCAAGGCACTTCTTACCTCA/C/GJCATCTAGTGAATGGGGTTATGTTGCATGTTAGA (43)
2	AX-76160487	80330589	+	CAGCTGCAAGGGAATTCATCACTTCCAAATACTAT/A/GJCTGGCAAAATTCATGATGTTTCACTATGTTTTT (44)
2	AX-80901802	120833799	+	GCACACGTTATTGAAGCGCTCTTTAGGGTAAACAG/A/TTATCACTCTATAGGCTCTGTTGACTGTTTAGATAGA (45)
2	AX-76004707	129702668	+	AGATGTTCCACTCTCTGAAAGACAGGAGCTCCAC/A/GJGGGAGAGAAAAATAAAGTACCATTGTTATCCACA (46)
2	AX-76011876	133448941	+	TTACAAAACCATGACAGCTGGAGATGTCTGTCTAAAC/C/TTGATTTATCATCACTTCTATGACAAATAGCCATTGAGA (47)
3	AX-76424142	15742614	+	TGGAGACGGGTTGAACCTGGGACAACGGATGGAAC/C/TTGGGGTGGATGGAATGGGGATGGAATGGAACTGGG (48)
3	AX-76471524	38893502	+	TTCAATTAACCTTTATCAAGCAGATCATTTGCCACAC/A/GJGTACCTCAAACTTAACATCCATTATTTTATATT (49)
3	AX-76535428	68491734	+	TTATGCTGATTTTCACTCTATGAACATTAAGTCTA/C/TTGGGTGTGTTCTTATCTGATCTTTTGGAGGAGGT (50)
3	AX-80939801	70475559	+	AAGAAATGAACCTGGCATCTTAGCACCCCAATATCTC/GJAGACATTTGACCATAAGGTCGCAAACTGATCTTTAA (51)
3	AX-76404489	106465724	+	GAATTTGGGAGGAATGCGCAGATTTTAAAGGAAAGAT/C/TTATTTGTTCCAGTGTCTCATTTGTTCTTAATGTT (52)
3	AX-76405796	106989833	+	TGTGCTCATTTAAGAGCTAGGGGATTATTTTACAT/A/JAATCCAGATTTTCTATTTGGAAAAAGTTGTCTGT (53)
3	AX-76406048	107075180	+	TTATGTTCCCATGGCAGCTCTGTAGTGTGTCATG/A/GJAGTACAGCAGATAGACCTTGAACCAATGAATTAC (54)

표 4

107개의 SNP 마커 (continued)

CHR	SNP	Physical Position	Strand	Flank (서열번호)
5	AX-80770611	8975995	+	TGACCTTTCAGAGCGTGTAGCGTGGTTATTGCTGTATATCTGCAGAGTCCGAAACAGTTTGTACAGCTCATG (55)
5	AX-76772443	10525709	+	AGAGCAAAATTCAGGTGCTGAAGTCAATCTCATAC/AGTCTTGTCTATTGGAAAGGAGTGTATGTAAGAAAT (56)
5	AX-76783359	15422952	+	CAACAAGAGGCCACGAGAGCGCTGGCTCACAACTTTC/TTGTGAAGCACAAAGCTGAGACAGGAAGAAGAACATGG (57)
5	AX-76795927	21044319	+	TAATCAATTAAATCAAGAACCTCAATTAAGTATCTGA/AGTCACTGGAAAGTTCTCAAGTATAGTTAGTTTACA (58)
5	AX-76798646	22318499	+	GCCATAAACCTTGTGAACCTCCAGGAATGCTGAAGATG/AGATTACAAATCTTACAAATTTGATATATCCCTATA (59)
5	AX-76813402	28944704	+	CTTCTCTCCCTTAGCTCCAGGAATGCTAAGAAG/CTTTTGTACACCAATGTTTCTGCTTATCAACAGGACC (60)
5	AX-76814216	29316596	+	GGGATACGTTTCAATAGTAAATAGGAGATGAGCACTG/TTCAAAGCTCTATGAAAACAAATTTCCATGTTCCATC (61)
5	AX-76816988	30557101	+	CTCTGCTTATATTTGGAAATTACAACAGTACCAG/AGTGAAGATCTGAAGATCTGAAGATGAATGGTTGGAATGT (62)
5	AX-80900040	31919073	+	TCGGAAGCTACCATGTGTAAAGTGCACAAACATG/C/GJAAATGGATTTTACATATACCTACAAACTTTTCAAT (63)
5	AX-76826610	35241221	+	AAATAGTGCACGAAGTATTTGTAAAGCAGTGCCTAA/AGTCAATGATCAAGGAATGAAAACTTGACTATTTATG (64)
6	AX-76897968	123160	+	TCCCTTCCCATTTCTGAAGTTCACCGCAATCTTCA/AGTGAATCTTACACTGGCTTCACTCAGTAACCCAGC (65)
6	AX-76905366	152308	+	GACTGCTTATAGCAATAGCAATTTTCTCAGCAGAC/AGTTGAATGCAGTATCAAGTTAAACAGATTAATTAGT (66)
6	AX-76956620	354242	+	CCTTGTGTGCTTGCACACACCTTTCACCTCTA/C/TTGGATTTACAGGTGACATCAAGGTACCAAGAAC (67)
6	AX-76929414	2483135	+	AAACTATTTTGTCTTAAATGACAAGTAGGTGGCTT/C/TTGTGAGCAATCTACTCAAAACAGATGTATATCA (68)
6	AX-76929922	2502241	+	ATCAGAAGAACTGAAGGACGGCACAATTCCTTCTG/AGTGAATGATGTCCTTCTTGTGGTCAGTAGCAGAGCT (69)
6	AX-76958587	4349782	+	TTCTCACTAGCAAAATGAATCCAGCTTGTGAC/AG/AAATAGGCATTAAATCAGCAAGTTGCTAAATATATGAT (70)
6	AX-76936945	27629403	+	AGCCCTTCATGAGTTAGCTGATTTCTTGTGTCA/AGTGAATCTGATTTAAATATGCTGTGATACTGTTTCAAG (71)
6	AX-76953936	33890827	+	AGTGAATCAGTGTAAACCGGGATCCACCTGGGGCA/C/GTGTGACACTGTAGCACTTATTTGACCCCATG (72)
7	AX-77013891	25566642	+	CAGATACCAAGATTCAGATCAACCCCCGAA/AG/CAATTTCACTCACCCTACCAAGAGCGTTGGCTACACAGAAA (73)
7	AX-77016453	26550071	+	AGGATTTGCATACAGATTTTCTCATATTTGCCCA/AG/TTTAAACAAAACATCCACCAAGTTCTTCTATGTCT (74)
10	AX-75606960	3013930	+	ATTTCCAAAGGAAGATTAATTCATGAGACGGATG/C/TAATTTCTGTGCTTGACTTGAGACTTCAGGCATTCC (75)
10	AX-75608022	3364198	+	TGTCATGTGTCTTAAATGACCCCTACTTGTG/C/TTGTCTTGTAAAGGGTAGGTGTTCTTTTCTTA (76)
10	AX-75612981	4896692	+	TTTAAAAATGATGATATACATATACAGTAGAC/AG/AGATTCCATAGGAGCTTTTGTCTATGGATGTGTTAA (77)
10	AX-75576766	12309949	+	ATCAGCATAGCTGTAAATCTCTTGTCTTAAATCA/TTAGTCTTGCAAAAAGAAATCTCTAACAGACTTCAGGCCT (78)
10	AX-75584784	14613017	+	TAGCTGGGCTGTGGGCACTCCCACTCAGGGTAA/TTGCG/TTAGTGTGGGGCTTTGAAAGTAGTTTAAATTAGAAAAGC (79)
11	AX-75645919	15968804	+	TTACCCGTGTTGATGATGCTACAGAGCTGCAAAAC/AG/TTCTTTGTGTTGAGGAGTTGAATTTTCTTCTGGTCA (80)
13	AX-75730899	10336153	+	CTATGATCCCATAGAGATGCCATGAATTAATTCAGG/C/TTTGGACGAAGAAATCCCTCACACTTTCATACAAAGT (81)

표 5

107개의 SNP 마커 (continued)

CHR	SNP	Physical Position	Strand	Flank (서열번호)
15	AX-75830104	3155963	+	CCAAATAGGACCTACCAACAGTAAGTCACAGCAAA/CJATCTATGAATCATTTCCATTAGAGCTTCTCTAT (82)
15	AX-75838335	5837883	+	AAAAAATGATTCACAGTAGAGAACTGCAGCATTTTC/TTCCCTTGATATTGTACTTTAGGAGTTTAAAGA (83)
15	AX-80941187	6059528	+	TAGCAGGTATGTGCTTTGTTAGGGGTAGCCAGAA/C/GTGGGGCTGGTGTGGTTTCTGCACTGGTATGTAGG (84)
15	AX-75817170	10442003	+	ATCAAACTGAACTGGGAAGACTGCCAATGTCTGGGA/GJATGGGATCTCCGTTCAAAACCAAGTGTGGGGAG (85)
15	AX-75817181	10446252	+	GGGGCAACCACTGACATTAAGGCTTTAATACCAAGG/C/TTGGGAGGAGTGAAGCTGTGCATCACACTTTTGA (86)
15	AX-75817263	10470240	+	TCAAGGACAGTAAAGAGAGGGCTCAGACTGGAAAAA/A/CCTTGCTCAATTGGGCTTTGGCACAGTATCATCT (87)
15	AX-75818576	10861056	+	TCTGGCAGTGGGCAACCCCTCACTAATATAC/A/GJTGACAGTGGTTTCTGTATTATTGAAGCCATCAAA (88)
18	AX-75887276	1073313	+	TTCAATTTTCTTGAGAGAACATCAGCTTGTAC/C/TTTGGTGGAAATTTTTCATCACCATGAACCCCGTTA (89)
18	AX-75893870	3095141	+	GCTAAACTGTCCAGAACTGTGAACCTCTAGTATG/A/GJCTTAATGAACCTCCAGAAATAAAGCTGAAAAAC (90)
20	AX-80804652	2981370	+	GCCCTGGGCTGGGTGCTGACTTATGGAGACTGT/C/GJCTTAGGGCAGGGCTTCCACGTCCAGGGCTGCT (91)
20	AX-76200839	11501023	+	CAAGTGGTTCTACAGGGTAAACATTAGTTGTAAT/G/TTACATACTCAGCACTCTCAGGACATTCAGCATTTT (92)
20	AX-76200845	11504796	+	TAAGAAAGTCTTAAGCATGTACTATATGGTAA/C/A/GJTAGGGTACTTGTTCTGCTTATCTGGGATCTT (93)
20	AX-76202441	12151057	+	CAGTGTGCTCCCTCAACCTCTGAGAGTGAATG/C/TTCTCTGATTCTGCTTCTTCTGACTGAATATAAAGTAACTAG (94)
22	AX-76272972	3936204	+	ATTCCCGACCATTAATCCAAAGGGAGGGAGTGT/C/TTCTGCTTGTCTACTGAATATAAAGTAACTAG (95)
24	AX-76302050	1921841	+	TTCTGTTTCAGGAGTGTGCTTGTGCTTTTCTCC/C/TTGAGATCTTTCGCTTCCGTTGCAATTTGGGAGGAAG (96)
24	AX-76304444	2532351	+	TAATCAGCAGAAACAGAAATCATGCAAGAAAC/C/TTGATACCTGGAGAGCAGATTTCTTCACAGTGGTTTC (97)
24	AX-76305675	2804520	+	ACTCCAGGTAGGAAAAAGCATCTTGTGCTGCA/A/GJATTGCTTGTTCATAGTACTGTCTTCAGTTTGC (98)
24	AX-80996298	3197370	+	TTCTTAAGTCAATTTCTCCGCATCTTAAAGAA/C/GJTAATTTGCTGACCAATGACCCATGCTAAGGATGGCA (99)
24	AX-76307382	3198302	+	TCTGGCAATACCGTAAGTCAATCAAGGGGAGGAGTGT/C/TTCTGACAAAGAGACCTTCAATTACCCTCCATA (100)
24	AX-76314648	4922109	+	TGAACACAGGAACCTCTTCTTCAACAGATGTGCT/C/TTTGCAAGCTAACAATAGCCCGGCTATAAATA (101)
25	AX-76328766	540367	+	TGCAATGGAGAGAGGGGGAGATCTGCTCATTT/C/TTATCCAGTGTGTCAAGGTGCAAACTGATTGCAAA (102)
26	AX-76333998	1849162	+	GTCTGCTCACTACTCTAACCTTTCTCTAGAT/A/GJATCCACCTTCTACAGTGTCTTGTTTCAAAATGG (103)
26	AX-76334482	1956241	+	AAAAAAGAGCATTTTAAAAAATTCATGTGCTCA/TACJATTTGTCTTTTACATACACCCGTTATTCATTTCA (104)
26	AX-76336047	2304029	+	TTGGATGGGAGCCCAAGATGGAAGGTGTGGAG/C/TTGGGAGTTAAAGACGAAAAAGCAAAACCTC (105)
26	AX-76338943	2962024	+	TTGAGCAAGGAGTTAAATCCGAGTCAAGCTG/A/GJAGAGATGAAGATGTAGGATGTAGTGTAGGG (106)
28	AX-76388478	737078	+	AGAGCAGAGTGTGGCACAATACCTCAAGGAGCAAA/CJTGTAATTAATGAACCATATTAACACACACACTGAAC (107)

실시예 2. 선별된 SNP 마커의 집단 구별능 평가

선발된 107개 SNP들의 식별능력을 시뮬레이션하기 위해 확보된 유전자형을 이용하여 가상의 자손을 생성하여 주 성분 분석을 수행한 결과에서도 case와 control 집단이 분리 가능한 것을 확인할 수 있었고, case 집단에서 유래된 가상의 자손의 경우 case와 control 사이에 위치한 결과까지 확인할 수 있어, 집단간의 명확한 분리를 유도할 수 있어 신품종 토종닭의 품종식별과 부정유통 방지용 검증수단으로 활용 가능할 것으로 예측되었다(도 3).

실시예 3. 선별된 SNP 마커의 토종닭 집단 구별 분석

앞서 확증된 107개 SNP 마커의 식별력을 검증하기 위해 Fluidigm genotyping chip을 제작하여 시중에 유통 중인

실용 육용계(Ross, Cobb, Abor Acres)와 난용계(Lohmann Brown), 국산 토종닭 실용 육용계(WM) 및 한협의 원종계(GPS) 교잡종을 포함한 총 182마리의 추가 유전자형 분석을 실시하였다. 유전자형 분석 결과 전체 call rate가 99.85%로 전반적으로 유전자형 분석이 잘 이루어졌음을 확인하였다. 일부 call이 되지 않은 데이터는 분석 결과에 영향을 주지 않는다고 판단되어 전체 평균값으로 대체하여 분석을 진행하였다. 선발된 107개 SNP 마커의 식별능을 검증하기 위해 앞서, 생성한 가상의 집단과 식별능 검증을 위해 새로이 유전자형 정보를 획득한 실제 집단 182마리의 유전자형 정보를 이용하여 MDS(Multidimensional scaling) 분석을 실시하였다. MDS plot을 확인한 결과, 기존 결과와 유사한 결과를 확인할 수 있었다. 즉 case와 control이 잘 분리되는 것이 확인되어, 본 발명의 107개의 SNP 칩을 이용한 분석 결과의 신뢰성이 높음을 확인할 수 있었다(도 4).

[0063]

107개 SNP 마커의 식별능 검증을 보다 정확하게 수행하기 위해 MDS 분석을 통해 얻은 분산 정보를 바탕으로 기계학습을 실시하여 식별능을 수치화하였다. 기계학습은 기존의 600K 데이터에서 획득한 283 수와 가상자손 488 수를 포함한 총 771 수의 데이터를 학습 데이터 세트로 사용하였고, 새롭게 획득한 182 수의 유전자형 정보를 테스트 데이터 세트로 이용하여 ROC(Receiver operating characteristic) curve를 도출하였다. 이때 정확도(Area Under Curve)는 기계학습의 모델에 따라 변하기 때문에, 본 분산 정보에 가장 최적의 기계학습 모델을 정확도 수치에 기반하여 선택하였다(도 5, 표 6).

표 6

토종닭 집단 구별에 대한 기계학습 모델별 정확도

	AdaBoost	Decision Tree	Linear Discriminant Analysis	Naive Bayes	Nearest Neighbors	Quadratic Discriminant Analysis	Random Forest
Accuracy	0.99	0.93	0.99	0.92	0.97	0.92	0.91
Sensitivity	0.99	0.89	0.99	0.88	0.95	0.87	0.86
Specificity	1.00	1.00	1.00	1.00	1.00	1.00	1.00

[0064]

총 7가지의 기계학습 모델(AdaBoost, Decision Tree, Linear Discriminant Analysis, Naive Bayes, Nearest Neighbors, Quadratic Discriminant Analysis, Random Forest)을 적용하여 107개 SNP 마커의 품종 식별능을 측정하였다. 기계학습 모델별 정확도를 도출한 결과, 대부분의 모델에서 전반적으로 높은 수준의 정확도로 품종 구분이 가능함을 확인하였다. 특히, AdaBoost 및 Linear Discriminant Analysis 모델의 경우 GSP 품종(HF, HH, HY, YH, FH, FY, CH, CF)과 그 외 품종을 99.45% 확률로 구분가능한 것으로 확인되었다. 이 외에도 Nearest Neighbors, Decision Tree 모델에서도 각각 96.7%와 92.85%로 높은 수준의 정확도를 확인하였다. 더불어, 실제 거짓값을 거짓값으로 인지할 확률의 지표인 Specificity는 모든 모델에서 100%의 정확도를 나타내었다. 이는 임의의 개체가 GSP 품종이 아닐 경우 이를 GSP 품종이 아니라고 판단할 확률이 100%임을 의미한다. 위와 같이 전반적으로 높은 수준의 품종 구분력을 보인 이유는 앞서 선발한 107개 SNP 마커들이 본 발명의 주요 타겟인 HF, HH, HY의 특이적인 SNP를 잘 반영하였기 때문인 것으로 사료되었다.

[0066]

마커 조합의 최소화 및 최적화를 위해 전체 107개의 SNP 중 품종 구분에 실질적으로 영향을 주는 SNP를 특징 중요도(Feature Importance; FI)를 기준으로 최소조합 마커를 선발하였다. 선발 결과 Decision Tree 모델에서 8개, Random Forest 모델에서 44개, AdaBoost 모델에서 36개의 SNP 마커가 각각 선별되었다. 모델별 특징 중요도로 선별된 SNP들은 일부 중복된 값을 가지고 있었으며, 상기 3개의 학습모델에서 선별된 총 61개의 SNP로 구성되어 있었다. 따라서, 품종 구분에 실질적으로 영향을 준다고 판단되는 61개의 SNP를 최적의 SNP 마커 조합으로 선발하였다(도 6, 표 7, 8 및 9).

표 7

AdaBoost 기계학습 모델에서 최적 SNP 마커

CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	AX-75446836	40926139	T	0.9457	0.2521	C	160.9	7.35E-37	51.63
1	AX-75450269	42731066	T	0.9348	0.2438	C	161.4	5.69E-37	44.46
1	AX-75458405	46632353	T	0.04348	0.5661	G	84.57	3.72E-20	0.03484
1	AX-75459486	47136098	C	0.2826	0.09574	T	24.34	8.07E-07	3.721
1	AX-75541725	86901036	C	0.8222	0.3347	T	74.5	6.07E-18	9.193
1	AX-75379158	1.87E+08	G	0.01087	0.5269	A	83.36	6.85E-20	0.009869
2	AX-76046791	16801291	T	0.7174	0.1062	G	177.2	1.96E-40	21.35
2	AX-76060521	24681889	G	0.7841	0.2062	A	119.2	9.41E-28	13.98
2	AX-76076543	33597664	A	0.5326	0.1116	G	93.33	4.44E-22	9.074
2	AX-76004707	1.3E+08	G	0.8804	0.2913	A	113.3	1.89E-26	17.91
3	AX-76404489	1.06E+08	T	0.3043	0.1467	C	13.5	0.000239	2.545
3	AX-76406048	1.07E+08	A	0.8696	0.3616	G	80.86	2.42E-19	11.77
5	AX-76772443	10525709	A	0.9783	0.3388	G	128.2	1.01E-29	87.8
5	AX-76783359	15422952	T	0.8152	0.186	C	149.8	1.96E-34	19.31
5	AX-76816988	30557101	G	0.3152	0.07025	A	47.62	5.16E-12	6.092
5	AX-80900040	31919073	C	0.6739	0.1188	G	144.4	2.88E-33	15.34
5	AX-76826610	35241221	G	0.4022	0.1736	A	24.35	8.04E-07	3.203
6	AX-76956620	354242	C	0.7826	0.2707	T	88.78	4.42E-21	9.701
6	AX-76929414	2483135	T	0.7717	0.2273	C	106.3	6.19E-25	11.5
6	AX-76929922	2502241	A	0.7889	0.1983	G	128.3	9.7E-30	15.1
6	AX-76936945	27629403	G	0.8261	0.249	A	114.4	1.05E-26	14.33
7	AX-77013891	25566642	A	0.6413	0.1777	C	88.21	5.88E-21	8.274
7	AX-77016453	26550071	A	0.6304	0.2107	C	67.87	1.74E-16	6.389
10	AX-75576766	12309949	A	0.3696	0.1095	C	40.77	1.71E-10	4.767
10	AX-75584784	14613017	G	0.6957	0.1909	T	99.46	2E-23	9.689
15	AX-75830104	3155963	A	0.03333	0.5847	C	91.94	8.96E-22	0.02449
15	AX-75817170	10442003	A	0.6222	0.04752	G	211.2	7.59E-48	33.01
15	AX-75817181	10446252	C	0.6196	0.04959	T	207.8	4.11E-47	31.21
20	AX-76200839	11501023	T	0.8889	0.2996	G	110.6	7.32E-26	18.7
20	AX-76200845	11504796	G	0.8913	0.2718	A	127	1.81E-29	21.97
22	AX-76272972	3936204	C	0.8913	0.166	T	200.6	1.52E-45	41.2
24	AX-76304444	2532351	C	0.8587	0.1446	T	205.6	1.28E-46	35.94
24	AX-76305675	2804520	A	0.913	0.3361	G	105	1.22E-24	20.74
24	AX-76307382	3198302	A	0.9239	0.3802	G	91.82	9.47E-22	19.8
26	AX-76333998	1849162	G	0.9783	0.3678	A	115.8	5.24E-27	77.36
26	AX-76336047	2304029	T	0.4884	0.1799	C	39.53	3.24E-10	4.351

[0067]

표 8

Random forest 기계학습 모델에서 최적 SNP 마커

CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	AX-75443426	39248473	T	0	0.5475	C	93.29	4.509E-22	0
1	AX-75446390	40710381	A	0.7065	0.1736	G	114.5	1.008E-26	11.46
1	AX-75446643	40838785	G	0.9457	0.3216	A	123.4	1.148E-28	36.71
1	AX-75446774	40897183	T	0.9565	0.2562	C	163	2.488E-37	63.87
1	AX-75446836	40926139	T	0.9457	0.2521	C	160.9	7.346E-37	51.63
1	AX-75450200	42700221	G	0.6196	0.157	A	93.13	4.892E-22	8.743
1	AX-75450269	42731066	T	0.9348	0.2438	C	161.4	5.688E-37	44.46
1	AX-75451215	43171485	G	0.02174	0.5768	A	95.25	1.678E-22	0.01631
1	AX-75451445	43263769	A	0.4457	0.1632	C	37.39	9.683E-10	4.121
1	AX-75452996	43997714	G	0.8043	0.2115	T	126.6	2.319E-29	15.32
1	AX-75458405	46632353	T	0.04348	0.5661	G	84.57	3.716E-20	0.03484
1	AX-75459486	47136098	C	0.2826	0.09574	T	24.34	8.071E-07	3.721
1	AX-75474264	54082924	G	0.9239	0.3306	A	111.3	4.978E-26	24.59
1	AX-75377378	186339052	T	0.5217	0.1384	C	71.08	3.428E-17	6.79
1	AX-75379158	187002396	G	0.01087	0.5269	A	83.36	6.846E-20	0.009869
2	AX-76073577	31602596	G	0.3478	0.04752	A	80.72	2.601E-19	10.69
2	AX-76076543	33597664	A	0.5326	0.1116	G	93.33	4.436E-22	9.074
2	AX-80984170	63960038	C	0.7391	0.1343	G	159.2	1.653E-36	18.26
2	AX-76004707	129702668	G	0.8804	0.2913	A	113.3	1.888E-26	17.91
3	AX-80939801	70475559	C	0.6196	0.1467	G	100	1.515E-23	9.473
3	AX-76404489	106465724	T	0.3043	0.1467	C	13.5	0.0002387	2.545
3	AX-76405796	106989833	A	0.5652	0.2149	C	48.05	4.162E-12	4.75
5	AX-76772443	10525709	A	0.9783	0.3388	G	128.2	1.008E-29	87.8
5	AX-76795927	21044319	G	0.7609	0.1245	A	178.6	9.831E-41	22.38
5	AX-76813402	28944704	T	0.3889	0.1021	C	49.62	1.87E-12	5.597
5	AX-76816988	30557101	G	0.3152	0.07025	A	47.62	5.164E-12	6.092
5	AX-80900040	31919073	C	0.6739	0.1188	G	144.4	2.879E-33	15.34
5	AX-76826610	35241221	G	0.4022	0.1736	A	24.35	8.035E-07	3.203
6	AX-76897968	123160	G	0.9783	0.4021	A	102.5	4.248E-24	66.92
6	AX-76956620	354242	C	0.7826	0.2707	T	88.78	4.419E-21	9.701
6	AX-76958587	4349782	A	0.8913	0.1941	G	175.9	3.8E-40	34.05
10	AX-75606960	3013930	T	0.8261	0.3058	C	88.06	6.352E-21	10.78
10	AX-75584784	14613017	G	0.6957	0.1909	T	99.46	2.001E-23	9.689
15	AX-75838335	5837883	C	0.6957	0.04792	T	251.2	1.447E-56	45.42
15	AX-75817170	10442003	A	0.6222	0.04752	G	211.2	7.594E-48	33.01
15	AX-75817263	10470240	A	0.6304	0.08058	C	166.9	3.524E-38	19.46
20	AX-76200839	11501023	T	0.8889	0.2996	G	110.6	7.322E-26	18.7
24	AX-76302050	1921841	T	0.4565	0.1095	C	67.59	2.012E-16	6.831
24	AX-76305675	2804520	A	0.913	0.3361	G	105	1.219E-24	20.74
24	AX-76314648	4922109	T	0.4783	0.1715	C	42.33	7.713E-11	4.429
26	AX-76333998	1849162	G	0.9783	0.3678	A	115.8	5.242E-27	77.36
26	AX-76334482	1956241	A	0.4239	0.1245	G	48.53	3.252E-12	5.175
26	AX-76336047	2304029	T	0.4884	0.1799	C	39.53	3.235E-10	4.351
28	AX-76388478	737078	C	0.7065	0.06405	A	229.7	6.765E-52	35.18

[0068]

표 9

Decision Tree 기계학습 모델에서 최적 SNP 마커

CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	AX-75450269	42731066	T	0.9348	0.2438	C	161.4	5.688E-37	44.46
1	AX-75379158	187002396	G	0.01087	0.5269	A	83.36	6.846E-20	0.009869
3	AX-76405796	106989833	A	0.5652	0.2149	C	48.05	4.162E-12	4.75
10	AX-75576766	12309949	A	0.3696	0.1095	C	40.77	1.709E-10	4.767
10	AX-75584784	14613017	G	0.6957	0.1909	T	99.46	2.001E-23	9.689
15	AX-75817263	10470240	A	0.6304	0.08058	C	166.9	3.524E-38	19.46
26	AX-76333998	1849162	G	0.9783	0.3678	A	115.8	5.242E-27	77.36
26	AX-76336047	2304029	T	0.4884	0.1799	C	39.53	3.235E-10	4.351

[0069]

[0070] 선발된 최적의 마커 조합 61개를 바탕으로 검증 시험(validation test)을 진행하였다. 검증 시험은 앞선 방법과 동일하게 기존의 600K 데이터에서 획득한 283 수와 가상자손 488 수를 포함한 총 771 수의 데이터를 학습 데이터 세트로 사용하였고, 새롭게 획득한 182 수의 유전자형 정보를 테스트 데이터 세트로 이용하여 마커 조합의 정확도를 평가하였다(표 10).

표 10

토종닭 집단 구별에 대한 기계학습 모델별 정확도

	AdaBoost	Decision Tree	Linear Discriminant Analysis	Naive Bayes	Nearest Neighbors	Quadratic Discriminant Analysis	Random Forest
Accuracy	0.995	0.93	1.00	0.918	0.984	0.918	0.934
Sensitivity	0.992	0.890	1.00	0.873	0.975	0.873	0.898
Specificity	1.00	1.00	1.00	1.00	1.00	1.00	1.00

[0071]

[0072] 최적의 SNP 마커 조합 61개에 대한 정확도 평가 결과 Linear Discriminant Analysis 모델에서 100%의 정확도로 가장 높은 정확도를 보였다. AdaBoost 모델 또한 99.45%로 높은 수준의 정확도를 나타냈다(표 11).

표 11

최적의 SNP 마커 조합을 이용한 토종닭 집단 구별능 결과

Breed	Experiment samples	AdaBoost	Decision Tree	Linear Discriminant Analysis	Naïve Bayes	Nearest Neighbors	Quadratic Discriminant Analysis	Random Forest
Hanhyup F	36	1.0	0.917	1.0	0.972	1.0	0.972	0.972
Hanhyup H	36	1.0	0.944	1.0	0.972	1.0	0.972	0.94
Hanhyup Y	26	0.961	0.769	1.0	0.731	0.962	0.73	0.769
WM_2	10	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Yelim K	5	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Hyunin	5	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Arbor Acre	11	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Cobb	8	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Ross	20	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Lohmann brown	5	1.0	1.0	1.0	1.0	1.0	1.0	1.0
GSP_cc (HFHY)	10	1.0	1.0	1.0	1.0	1.0	1.0	0.9
GSP_cc (PHFY)	10	1.0	0.9	1.0	0.4	0.8	0.4	0.6

[0073]

[0074] 더불어, AdaBoost 모델에서 선발된 36개의 최소의 SNP 마커 조합으로도 99.45%의 높은 수준의 품종 구분이 가능한 것을 확인하였다(표 12). 상기 최소 마커 조합을 활용하면 일부 주요 PL 계통인 HY를 제외하고 모든 집단에 서 정확한 구분이 가능하였다. PL 집단의 특성상 시중에서 유통되지 않는 점과 기계학습의 학습데이터가 축적될

수록 정확도가 개선되는 점을 감안한다면 실제 현장에서 신종품 토종닭을 구분 가능한 마커로 활용 가능 할 것으로 사료된다.

표 12

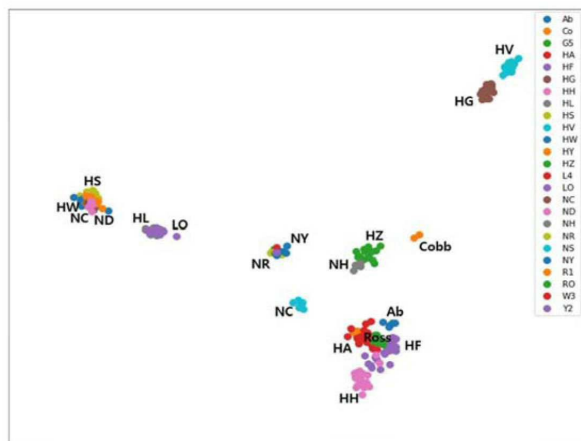
최소의 SNP 마커 조합(36 SNP)을 이용한 토종닭 집단 구별능 결과

Breed	Experiment samples	AdaBoost	Decision Tree	Linear Discriminant Analysis	Naïve Bayes	Nearest Neighbors	Quadratic Discriminant Analysis	Random Forest
Hanhyup F	36	1.0	0.972	1.0	1.0	1.0	1.0	1.0
Hanhyup H	36	1.0	0.972	1.0	1.0	0.972	0.972	1.0
Hanhyup Y	26	0.962	0.769	0.962	1.0	0.962	0.923	0.885
WM_2	10	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Yelim K	5	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Hyunin	5	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Arbor Acre	11	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Cobb	8	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Ross	20	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Lohmann brown	5	1.0	1.0	1.0	1.0	1.0	1.0	1.0
GSP_cc (HFHY)	10	1.0	0.9	1.0	1.0	1.0	1.0	0.7
GSP_cc (FHFY)	10	1.0	0.8	1.0	1.0	1.0	0.9	1.0

[0075]

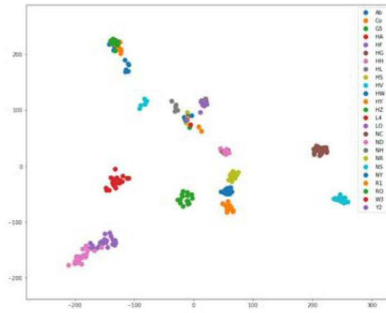
도면

도면1

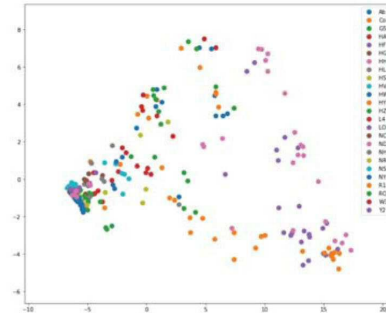


도면2

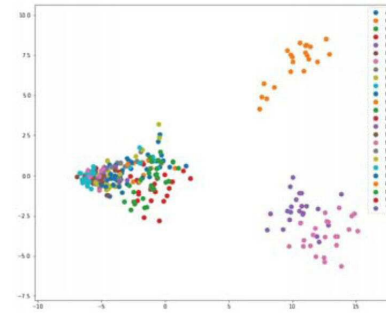
107 SNP 선발 전



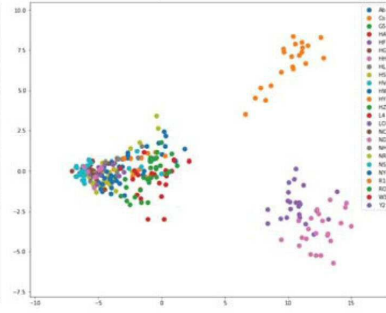
Set 1. 1 LD당 SNP를 선발



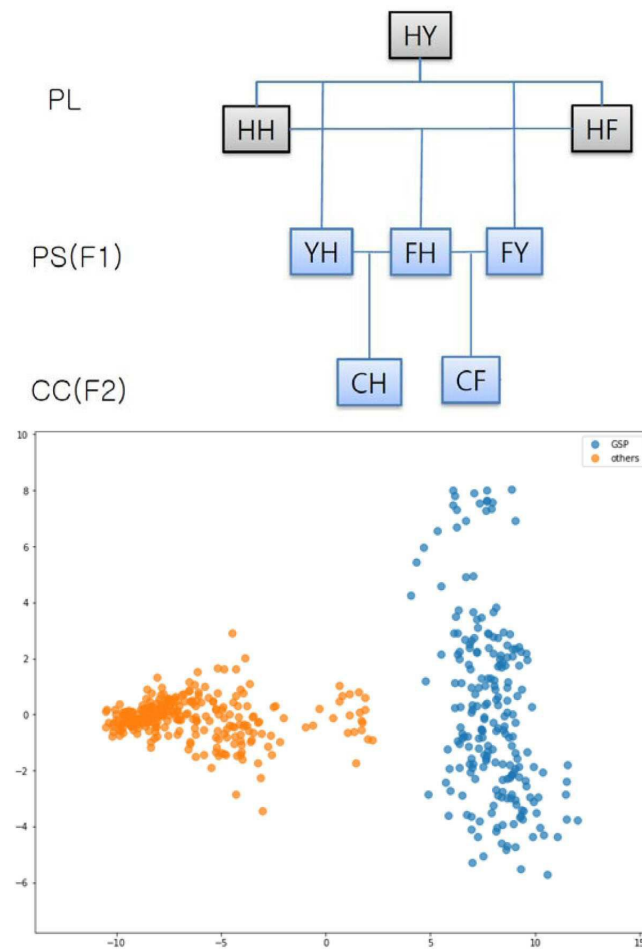
Set 2. 50 LD당 SNP를 선발



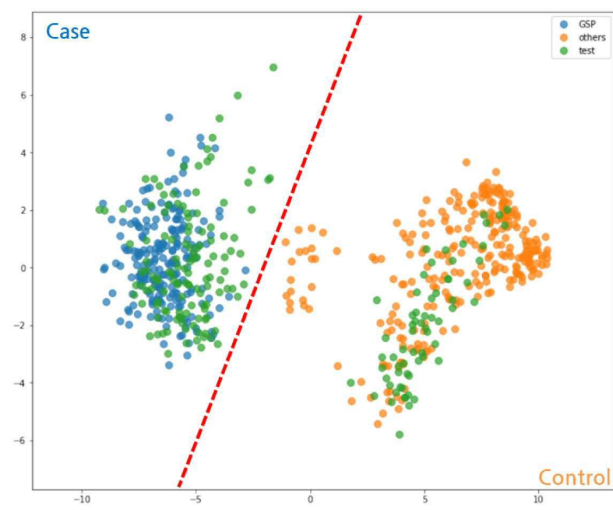
Set 3. 100 LD당 SNP를 선발



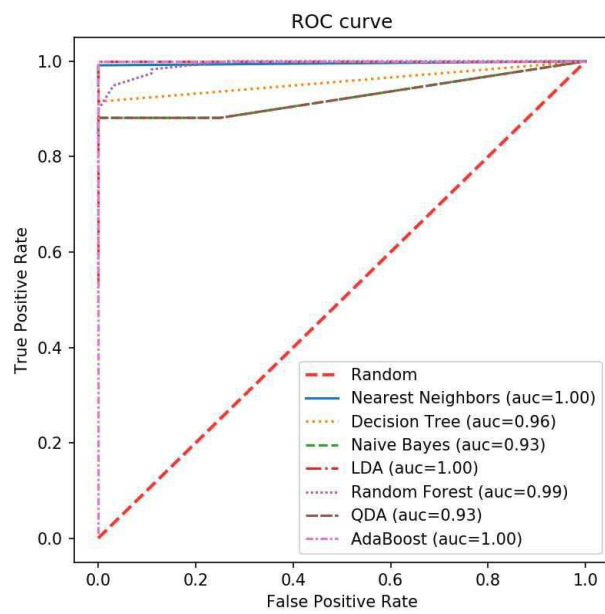
도면3



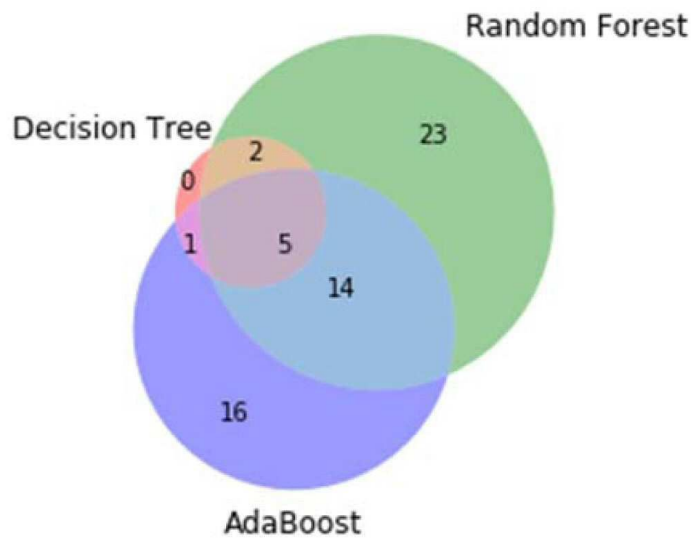
도면4



도면5



도면6



서열 목록

<110> The Industry & Academic Cooperation in Chungnam National University (IAC)

<120> SNP marker set for discriminating genetic background and cultivar
of Korean native chicken and uses thereof

<130> PN19418

<160> 107

<170> KoPatentIn 3.0

<210> 1

<211> 71

<212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 1

gaagcaactt tgcttttggg ggttggttat tctcacacaa aaggtgtggc tatttctttt 60

ctggtgttag c 71

<210> 2

<211> 71

<212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 2

agaaactctc accagtcacg agctactttt cttccacacag gttgctgtgc atatccagtt 60

tgcttcctgt g	71
<210> 3	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 3	
tacgtatgaa ccattaaagg aagaaaagga agagaagaaa agacttcgc atgtttgcc	60
aattctcaag a	71
<210> 4	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 4	
ttgtactgag caagactttg ccaagaccta cttctacttc acttactgtt tgtaaagac	60
tgcatgtatg g	71
<210> 5	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 5	
aagactctac ttgtttatag ccagaactta agcataagag tgcatactg ctgcatcaca	60
ataacagggt a	71
<210> 6	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 6	
tagtattgtt gatggccatt atccccagta ctggactga ttatttgtca catagcgcc	60
accacagagg a	71

<210> 7
 <211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 7
 tgttctgtgc tcigcggcaa cagattgttt tctgtcccat gtatgctcag gtttttagt 60
 tttcattttt t 71

<210> 8
 <211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 8
 gcctaagtat catcttttat agtgctgagg ttgctacata gtgcagaaac tgccaattag 60
 tacctgtgct t 71

<210> 9
 <211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 9
 ctcaagtgat aggtcatgcc tttatattat acttaaatag gtatgttgaa gtttctcact 60
 ttattttccg g 71

<210> 10
 <211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 10
 ttcagcaaga gctgttgaa acgttgcttt aataaatgac cactcctgtt cagcacagaa 60
 aaaaaaagg t 71
 <210> 11

<211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 11
 cattgcacag ctcccaaac aatgtgatct gctaacggcc cctcaaagct ctcagggtgc 60
 cctccccacc c 71

<210> 12
 <211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 12
 gcagtttact ttggcaagg aagaaccag taaggagatg agcactgtgt gggaaggact 60
 ggtacagtgg c 71

<210> 13
 <211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 13
 cacatctttc agttgatttg ttggttttgg aaggacagag aagcaaatgg aaagtcaatt 60
 tggacagtgt t 71

<210> 14
 <211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 14
 atgcctttat aggttcttca atccttatgc ttcagcttga ttgttttct caactttca 60
 ctagctataa g 71

<210> 15
 <211> 71

<212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 15
 tttcaataaa atctctgctt gcaatgtagt gtcttgtgaa tctcattagt tgggattgtc 60
 tataagtgca t 71

<210> 16
 <211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 16
 tacctgagag taaacatttc ttcacagata atgcactttg ctgatgcaga gtcttctaga 60
 aacgtatect a 71

<210> 17
 <211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 17
 taattaggta aaaatcattc caagtaggat actctaaact tgctgcagta actggcataa 60
 aaggctgaca c 71

<210> 18
 <211> 71
 <212> DNA
 <213> Unknown
 <220><223> Gallus gallus domesticus
 <400> 18
 aggcctgagg catgatattc aagtttccac cagaacatta attttacatg ctactgtcta 60
 ttgcaagatt a 71

<210> 19
 <211> 71
 <212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 19

ctatTTtaga tacttcaaag aggagatggT agagaagact ggattaaatt tttctgtgtg 60

ctaagctgtc a 71

<210> 20

<211> 71

<212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 20

tcactTTaat atgtctgagtG ccacaatgtg cttaaactct tctccttgca tccacaccac 60

ttgctctctc t 71

<210> 21

<211> 71

<212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 21

ataagtattc ccaaagcaga gaagcagcgt aattcatatg cttgacctat tccaggtaac 60

agttaaacct a 71

<210> 22

<211> 71

<212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 22

ggcaaggacc cagcccatgc actttgttca tttagagact ggaggactga atggcaaagc 60

aaggccttat t 71

<210> 23

<211> 71

<212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 23

ttccttagat gactcatttt ctaggtttca tcagcataca gtgccaccac ttcattcagct 60

agtaacaaca a 71

<210> 24

<211> 71

<212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 24

tgtggctgca tctgctggga tgctgctaag attatgatga gcctttgcaa gccttcaaat 60

gcaataataa t 71

<210> 25

<211> 71

<212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 25

ttattttgaa tcaggttctt attcccatca gagctgctgt cagtcgcctg ctgccagact 60

agctgtgcta c 71

<210> 26

<211> 71

<212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 26

gaagcaactt tgcttttggg ggttggttat tctcacacaa aaggtgtggc tatttctttt 60

ctggtgttag c 71

<210> 27

<211> 71

<212> DNA

<213> Unknown

<220><223> Gallus gallus domesticus

<400> 27
cagctcagaa tctgttaacc tcatcagcaa tcctaattg tcaagcccca tggcagaaat 60
caaccctaag a 71

<210> 28
<211> 71
<212> DNA
<213> Unknown
<220><223> Gallus gallus domesticus

<400> 28
agggaaaagg aactgttatt ttatctgaca ggagcataca gtttttactg ggatcaggca 60
cacagaatag a 71

<210> 29
<211> 71
<212> DNA
<213> Unknown
<220><223> Gallus gallus domesticus

<400> 29
gtgttgacct tgcttctctc tgcttgaagg tcatccagaa ctgtctgatg agcttctctgc 60
aaggacttct t 71

<210> 30
<211> 71
<212> DNA
<213> Unknown
<220><223> Gallus gallus domesticus

<400> 30
cttagcatga attgagaaaa tgagatactt ttttgcttcc attctcagtc agttttacag 60
ctctgctccc t 71

<210> 31
<211> 71
<212> DNA
<213> Unknown
<220><223> Gallus gallus domesticus

<400> 31

gattagcaat ttctatgtta ccagatgtaa gtaccaaadc atggaaagtt actttcttta	60
gatgttttagc t	71
<210> 32	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 32	
tgtattcacc cagttgatgc aaccatcttg gccagctgtc actgattaca aagtgtgaca	60
catgttgcct c	71
<210> 33	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 33	
cttatccacc tactcatgag gactgaagca tactcagtgt ttgcaggat ctaatacgt	60
gatcttattt t	71
<210> 34	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 34	
tggtactacc atttatgttg ctacaaaaa ataacaacct tcatagctgt agctactgac	60
cttatagtgc c	71
<210> 35	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 35	
ggtatgtttt caagcttttc tcagacactt gaaagcaaga tagggtgagt ggggcttcct	60

gcaatttgtgt g	71
<210> 36	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 36	
atccaggcct tctgtaactg ataagaacgg atagcactac atcagtataa tcacaccatt	60
gaagtagtta c	71
<210> 37	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 37	
ttaacagatg tgtaatggtt tccaaaggaa tgggaaagtt gacacaaaaa tgccatgcat	60
gcctgcacta c	71
<210> 38	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 38	
ttttctgaca gagagcagtt agtgagtccc acttttggtga aagctttctg caagtcattc	60
ttaagtatgc t	71
<210> 39	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 39	
gattcctgtg ttgaagatga gcaaaacttg agatgaccaa actacggtca ttccaatctt	60

gctggctgtg g	71
<210> 40	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 40	
cgagcagcac tggagtttca aatgtgttat gtacaattag gcaatgacac tgaagttaca	60
atgtaagata c	71
<210> 41	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 41	
ggcacatcag cattgattcc ttgttatatt ttccaataac ttttcaactt ctctgcatt	60
tattttctgc c	71
<210> 42	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 42	
cagaacactt tattgataac cagacaaaat ttctgcgatt tcctggaagt ctgaattaaa	60
accaagactt a	71
<210> 43	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 43	
agatgattcc atgcaaggca tcttccttta cctcaccatc tagtgaaatg ggggttatgt	60

tgcatgttag a	71
<210> 44	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 44	
cagctgcaag ggaaatcatc acttccaaat actatactgg caaatccaat gatagtttca	60
ctatgttttt t	71
<210> 45	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 45	
gacacgttat tgagcgcctc ttttagggta aacagaatca ctctataggc tctgttgact	60
gtttagatag a	71
<210> 46	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 46	
agatgttcca ctctctgaa agacagggac tccacagga gagaaaataa gctaccatt	60
ggtatccac a	71
<210> 47	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 47	
ttacaaaacc atgacagtgg agatgtctgt ctaaactat tcatacttc tatgacaaat	60

agccattcag a	71
<210> 48	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 48	
tggagacggg ttgaactggg gacaacggat ggaaccgggg gtggatggaa tgggggatgg	60
atggaactgg g	71
<210> 49	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 49	
tcatattact ttatcaagca gatcattgcc cacacagtac cctcaaacta acatccatt	60
attttaatat t	71
<210> 50	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 50	
ttatgctgat ttcagtctat gaacattcaa gtctacgggt gtgttcttat ctgatctttt	60
gcgaggagg t	71
<210> 51	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 51	
aagaaatgaa actggcatct tagcacccaa atcatcagac attgtacat acggtcaaac	60

tgtatcttta a	71
<210> 52	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 52	
gatttgggag gaaatgccag atttaaagga aagatccatt tgttcccagt gtctcatttg	60
ttcctaattgt t	71
<210> 53	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 53	
tgtgctcatt aagagctagg ggattatttt acataaatcc agcatttctc attgttggaa	60
aagttgtctg t	71
<210> 54	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 54	
tcatgtccca tggcagcctc tgtagtggtg ccatgaagta cagcagatag accctgaacc	60
aaatgaatta c	71
<210> 55	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 55	
tgactttcag agcgtgctag cgtgggttatt gctgtaatct gcagagtcg aacagtttgt	60

cacagctcat g	71
<210> 56	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 56	
agagcaaatt caggtgctga agtcatatct catacactct tgctattgga aggaagtgct	60
atggtaggaa t	71
<210> 57	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 57	
caacaaggag cccacgagag cctggctcac aacttcgtga agcacaagct gagacaggaa	60
gaagaacatg g	71
<210> 58	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 58	
taatcaatta atcagaacct catttaagta tctgaacact ggaagttctt caagtatagt	60
taggttttac a	71
<210> 59	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 59	
gccataaacc ttgtgaactc catgaactgg agatcaattc acaaatcttc acaatttgat	60

atataccctat a	71
<210> 60	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 60	
ctttctctcc ccttagctcc acgaatgcta agaagctttg tacaccatgt ttctgcctta	60
tcaacaggac c	71
<210> 61	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 61	
gggatacgtt tcatagtaaa taggagatga gcactgcaaa gctctatgaa aacaatttc	60
atgttcccat c	71
<210> 62	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 62	
ctctgtctta tatttggaat tacaacagta ccggaagatg aaagatctga ggatgaattg	60
gttggaatg t	71
<210> 63	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 63	
tctgaagcta ccatgtgtaa gtgcacaaac acatgcaaat ggattttcac atatacctac	60

aaaactttca t	71
<210> 64	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 64	
aatagtgcc a gaagtattg taaagcagtg cctaaatcat gatcaggaat gaaaaacttg	60
cactatttat g	71
<210> 65	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 65	
tccctttccc attctgagtt ccaccgcaaa tcttactga ttctacact ggcttcactc	60
agtaaccag c	71
<210> 66	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 66	
gactgcttta tagcaatagc attttctcag cagacattga atgcagtatc aagttaaaca	60
ggataattag t	71
<210> 67	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 67	
ccttggtgtgc ctigccacac cactttcacc ctctacggat tctacaggtg acatcaaggt	60

acccaagaac c	71
<210> 68	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 68	
aaactatattt gccttaaatg acaagtaggt ggcttcgtgt gagcaatcta ctcaaaacag	60
atgttatatc a	71
<210> 69	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 69	
atcagaagaa ctgaaggacg gcacattccc ttctgagatg gtccttcctt ttgtgggtca	60
gtaggacagc t	71
<210> 70	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 70	
ttctcactag acaaatgaat ccagctttgc tgacaaaaat aggcattaat cagcagttgc	60
taaaatatga t	71
<210> 71	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 71	
agcccttcca tgagttagcc tgatttcttt gtgtcagagt ctgatttaaa tatgctgtga	60

tactgtttca g	71
<210> 72	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 72	
agtgagtcag tgctaaaacg ggatccacct ggggcagtgc tgcactgtag cactcttatt	60
tgaccccat g	71
<210> 73	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 73	
cagataccac agattcagat caaccccccg tgaaaaattt caccicacca gaagcgttcg	60
ctacacagaa a	71
<210> 74	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 74	
aggatttgca tacagatttc tcatattgcc caatgattta acaaaaacat ccaccagttc	60
ttctatgtc t	71
<210> 75	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 75	
attttccaaa ggaagattat tccatgagac gcatgcaatt ctctgttgct gcttggactt	60

caggcacttc c	71
<210> 76	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 76	
tgtcatgtgt cttctaaaat gaccctacc ttgtcctgtc tcttgtaagg gtaggtgtt	60
cttttttcct a	71
<210> 77	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 77	
tttaaaaatg catgatacat atactacagt agcacagatt ccataggagc ttttgctatg	60
ggatgtgtta a	71
<210> 78	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 78	
atcagcatag ctgtaatcct cttgcttaat caattatctt gcaaaacgaa tctctaacag	60
acttcaggcc t	71
<210> 79	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 79	
tagctgggtc tgggcatccc actcagggtt aatgcgagtg gtgggctttg aagtagttta	60

aattagaaag c	71
<210> 80	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 80	
ttaccctgtt tgatagtgt cacagctgca aacacatctt tgtgttcagg agttgaattt	60
tcttctgggc a	71
<210> 81	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 81	
ctatgatccc atagagatgc catgaataat tcaggcttgc agcaagaaat ccctcacact	60
tcatacaagt c	71
<210> 82	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 82	
ccaataagga cctaccaaac agtaagtcac agcaaaatct atgaatcatt tccatttaga	60
gcttctccta t	71
<210> 83	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 83	
aaaaaatgat tcacagtaga gaactgcagc atttccttcc ctttgatatt gtactttagg	60

agttttaag a	71
<210> 84	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 84	
tagcaggtat gtgccttggt cagggtagc cagaactggg gctggtgtgg ttctgcact	60
ggtatggtag g	71
<210> 85	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 85	
atcaaactg aactgggaag actgccaatg tctggaatgg gatctccgtt caaaaccagt	60
gtgctgggga g	71
<210> 86	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 86	
ggggcaacca ctgacataag gctttaatac cagggctggg aggagtgatg agctgtgcat	60
cacacttttg a	71
<210> 87	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 87	
tcaaggacag taagagaggg ctcagactgg aaaaaacttg ctcacaattg ggtcttgac	60

agtgatcatc t	71
<210> 88	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 88	
tctggcagtg ggcaccgacc ccccataact tatacatgac agtggttcct gtatttattg	60
aagccatcaa a	71
<210> 89	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 89	
ttcaattttc ttggagaaga catcagcttg gtaccctcgg tgggaatttt catcaccatg	60
aaccccggtt a	71
<210> 90	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 90	
gctaaactgt ccagaactgt gaacttctag taatgatctt aatgaaactc ccagaaataa	60
aagctgaaaa c	71
<210> 91	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 91	
gcccttgggc tgggtgcctg acttatggag actgtcctta gggcagggt tcccacgtcc	60

caggcgctgc t	71
<210> 92	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 92	
caagtgttc tacagggtac acattagttg taaatgcaca tactcagcac tctcaggaca	60
ttcagcattt t	71
<210> 93	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 93	
taagaaagtc tctaagcatg gtactatatg gtaacatagg gttacttgtc tgttcctta	60
tctgggatct t	71
<210> 94	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 94	
cagtgtgtct cctcaaccc tctcgtgagt gacatgtcc tgtattctgc cttcttctcg	60
acttcagacg c	71
<210> 95	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 95	
attccccagc cataattcca aaggggaggg atgtgctctg ctttgctatc actgaaatat	60

aaagtaatca g	71
<210> 96	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 96	
ttctgtttca ggaggtgctt tgcgcttttc ttccccgaga tctttcgctt tccgttgcat	60
ttgggaggaa g	71
<210> 97	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 97	
taatcacgca gaacaagaaa tcatgcaaag aacaccgtac ctggagagca gatttcctca	60
cagtgtgttt c	71
<210> 98	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 98	
actccaggta ggaaaagcat cttgttcct gcacaaattg cttgtttca tagtactgtc	60
ttcagttttg c	71
<210> 99	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 99	
ttcttaagtc atttcctccg catctttaga aagcactaat ttgctgacaa tgacctatgc	60

taaggatggc a	71
<210> 100	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 100	
tctggcaata ccgtaagtcc aatcaagggg cagcaactct ctgacaagag acactccaaa	60
taccctccat a	71
<210> 101	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 101	
tgaacacagg aatcctcttt cacaacgatg ctgctcggtc gcagctaaca aatagcccgg	60
gctataaaat a	71
<210> 102	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 102	
tgcattggag agagggggga gatcttgctt catttcatac cagtgtgtca aggtgcaaac	60
ctgattgcaa a	71
<210> 103	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 103	
gtctgcctca ctactctaac ctttctctct agataaatcc acctttctac agtgtcttgg	60

ttttcaaatg g	71
<210> 104	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 104	
aaaaaagagc atttttaaaa attcatgggc catacattg tcttttacia tcacaccgtt	60
attcatttca t	71
<210> 105	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 105	
ttgggatggg accccagaat ggagggtgtg gagcgcgggg agttaagaac gaaaagcaaa	60
agcaaaacct c	71
<210> 106	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 106	
ttcagcacia ggagttaaca tccgcagtca gactgaagag atgaagatgt agggatgtac	60
gtgatgtagg g	71
<210> 107	
<211> 71	
<212> DNA	
<213> Unknown	
<220><223> Gallus gallus domesticus	
<400> 107	
agagcagagt tgtggcacia tacctcaagg agcaaatggt aattaatgaa ccataataca	60
ccacactgaa c	71