

Current Status of Cattle Genome Sequencing and Analysis using Next Generation Sequencing

Jung-Woo Choi, Han-Ha Chai, Dayeong Yu, Kyung-Tai Lee, Yong-Min Cho and Dajeong Lim*

Animal Genomics and Bioinformatics Division, National Institute of Animal Science, Rural Development Administration, Jeonju 565-851, Korea

Received March 2, 2015 / Revised March 23, 2015 / Accepted March 23, 2015

Thanks to recent advances in next-generation sequencing (NGS) technology, diverse livestock species have been dissected at the genome-wide sequence level. As for cattle, there are currently four Korean indigenous breeds registered with the Domestic Animal Diversity Information System of the Food and Agricultural Organization of the United Nations: Hanwoo, Chikso, Heugu, and Jeju Heugu. These native genetic resources were recently whole-genome resequenced using various NGS technologies, providing enormous single nucleotide polymorphism information across the genomes. The NGS application further provided biological such that Korean native cattle are genetically distant from some cattle breeds of European origins. In addition, the NGS technology was successfully applied to detect structural variations, particularly copy number variations that were usually difficult to identify at the genome-wide level with reasonable accuracy. Despite the success, those recent studies also showed an inherent limitation in sequencing only a representative individual of each breed. To elucidate the biological implications of the sequenced data, further confirmatory studies should be followed by sequencing or validating the population of each breed. Because NGS sequencing prices have consistently dropped, various population genomic theories can now be applied to the sequencing data obtained from the population of each breed of interest. There are still few such population studies available for the Korean native cattle breeds, but this situation will soon be improved with the recent initiative for NGS sequencing of diverse native livestock resources, including the Korean native cattle breeds.

Key words : Cattle, CNV, next generation sequencing (NGS), signature of selection, SNP

서 론

20세기 축산 특히 동물유전, 동물생명공학 관련 분야는 가축통계육종의 발전, 인공수정 기술, 체세포복제 등 연구결과에 힘입어 가축개량 및 제반 기술에 획기적인 발전이 있어왔다. 금세기 들어 이를 잇는 또 하나의 기념비적 연구가 수행되었는데 이는 미국, 캐나다, 유럽 등의 연구진들에 의해 주도된 주요 가축들에 대한 전장유전체 해독 및 표준유전체지도(reference genome sequence assembly) 작성이 그것이다[19, 21, 22]. 이러한 전장유전체해독에 대한 결과는 고품질의 단백질을 생산하는 가축의 독특한 학술적 가치와 더불어 주요 경제형질(economic traits)의 개량 가속화, 질병치료, 신약품 개발 등에 유용하게 사용될 것으로 기대되고 있다.

이러한 기존의 전장유전체 해독은 상당한 연구 인력과 자금을 요했었기에 일부 대형연구기관 혹은 국제 컨소시엄 형태로

수행되어 왔으나, 최근 차세대염기서열(Next Generation Sequencing, NGS)의 등장과 해독기술의 발전은 전통적인 sanger sequencing에 비해 급속히 단축된 시간과 비용으로 전장 유전체를 해독할 수 있다는 점에서 유용성이 각광받고 있다. 기본적으로 NGS 플랫폼은 수백만에서 수억의 시퀀스 조각(short reads)의 생산을 통해 전장유전체 해독을 수행하며, NGS 등장 초기 주요 문제점으로 지적되어 온 정확성의 향상과 지속적인 시퀀싱 가격 하락으로 인해 소위 시퀀싱의 민주화(the democratization of the sequencing)를 보다 가속화 시켜주고 있는 형국이다. 기본적으로 이들 NGS에서 생성된 데이터는 assembly 방식에 따라 크게 resequencing과 de novo sequencing으로 분류될 수 있다. Resequencing은 기존에 만들어진 표준염기서열(reference genome assembly)과 비교하여 대용량의 유전다형성 발굴 및 분석에 활발히 이용되고 있으며, de novo sequencing의 경우 de novo 단어 자체가 의미하듯 표준염기서열을 통한 mapping이 아닌 기본적으로 다양한 크기의 라이브러리(sequencing libraries)에서 생산된 reads의 자체적 assembly 수행을 지칭할 수 있을 것이다.

이러한 NGS를 이용한 전장유전체 해독은 최근 전세계 가축유전체 분야 특히 국제 공동연구를 통해 reference genome assembly가 既 생성된 중에서 대용량 유전다형성 발굴 등의 목적으로 resequencing 형태로 활발히 수행되고 있으며[1, 13,

*Corresponding author

Tel : +82-31-290-1606, Fax : +82-31-290-1602

E-mail : lim.dj@korea.kr

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

26, 29, 37], 반면 표준유전체 제작에 사용된 가축품종과 뚜렷한 특성이 기대되는 일부 품종의 경우 NGS 적용을 통해 자체적인 de novo assembly를 통한 표준유전체지도 작성이 이루어지고 있다[14]. 더 나아가 이들 NGS는 지속적인 가격 하락에 힘입어 기존 초위성체마커(Microsatellite marker) 혹은 SNP Chip-array 등으로 접근해왔던 집단유전체학(Population genomics)적 분석의 고도화를 통해 정밀한 선발신호(signature of selection) 및 가축화(domestication) 연관 유전자 지역 발굴에도 활용되기 시작하고 있다[33, 34].

국내재래가축은 지난 세기 주요경제형질의 개량, 외모특성의 고정정책 및 해외개량품종의 전 세계적 확산 등을 이유로 상대적으로 생산성이 떨어졌던 고유도착종들이 한때 멸종위기에 처하며 한반도 내 재래가축품종의 유전적 다양성이 심각하게 훼손 되었었다. 이에 재래고유의 생물유전자원에 대한 중요성이 최근 인식되며 재래가축 유전자원에 대한 본격적인 보존 및 관리 등이 농촌진흥청 국립축산과학원을 책임기관으로 지정, 운영되어 오고 있다[5]. 더 나아가 재래가축의 중요성은 생물다양성협약(convention on biological diversity)과 더불어 생물자원에 대한 주권적 권리인정 및 이들 유전자원에 대한 지식재산권과 이익 분배에 대한 나고야 의정서가 채택됨으로써 한반도 내 유전자원의 다양성 보존 및 이의 유전체학적 분석이 시급해지고 있는 실정이다.

NGS를 이용한 국내재래가축에 대한 전장유전체 해독은 소, 돼지, 닭 등 일부 주요 경제가축에 대해 이루어졌으며(completed sequencing reaction 기준), 이에 대한 본격적 분석 결과는 최근 약 2년간 집중적으로 출판이 이루어 졌다[7-9, 10, 25]. 이들 대부분의 연구는 초기 NGS 도입 당시 집단수준에서 분석하기에는 고가의 분석이었기에 각 품종당 단위 개체 및 각 해당품종의 표준유전체지도를 이용하여 대용량의 유전다형성 발굴 등 및 유전적 구성의 기본적 탐색 등의 연구목적으로 주로 수행되었다. 이 중 주요 경제동물인 소는 현재 한우, 칩소, 흑우, 제주흑우의 4개 품종이 국제연합식량농업기구 가축다양성정보시스템(domestic animal diversity information system in Food and Agriculture Organization of the United Nations, 약칭 DAD-IS in FAO)에 등록돼 있으며, 이들 4개 품종은 Applied Biosystems by Life Technologies社의 SOLiD 3, Illumina社의 Genome analyzer 및 HiSeq 2000 플랫폼을 통한 high coverage depth의 시퀀싱 데이터를 생산하여 정확한 유전다형성 발굴에 성공하였다. 더 나아가 지속적으로 하락하고 있는 NGS 시퀀싱 가격 및 다부처유전체사업 출범으로 인해 이들 재래소 품종에 대한 NGS의 적용을 최근 집단수준에서 확장하고자 하는 연구가 개시, 현재 진행 중에 있다. 본 총설에서는 최근 주목 받고 있는 이 NGS 기술을 이용한 소 유전체학 분야에서 수행된 다양한 연구결과와 현황에 대해 고찰하고, 이러한 기술의 향후 연구 및 적용방향에 대해 논의해보고자 한다.

본 론

소 품종에 대한 전장유전체 연구

소의 표준유전체 지도는 미국이 주도한 국제 소 유전체 및 반수체 국제 컨소시엄(The Bovine Genome Sequencing and Analysis & The Bovine HapMap Consortium)에서 6년간 약 700억원의 연구비를 투입하여 작성되었으며[19, 20], 사상 최초의 반수동물에 대한 연구 결과물은 2009년 4월 Science의 표지논문으로서 두 편이 동시에 발표 되었다. 이의 주요 결과물로서 두 종류의 표준유전체지도(Bovine genome sequence assembly version Btau4 and UMD3)가 당시 배포됐으며[28, 40], 이들 genome assembly는 현재 소 유전체학 분야에서 광범위하게 쓰이고 있는 single nucleotide polymorphism (SNP: 단일염기다형성) chip 제작[30] 및 활용[2, 17, 23, 36]과 더불어 NGS의 적용분야 특히 resequencing 형태로 전장유전체에 대한 대용량의 유전다형성 발굴(e.g. SNP, Insertion and Deletion (InDel))과 Copy Number Variation (CNV)등 전장유전체에 걸친 광범위한 유전다형성 발굴에 활발히 사용돼오고 있다. 이러한 연구들은 대부분 European *Bos taurus* 및 *Bos indicus* 품종에서 활발히 진행돼오고 있으며[1, 13, 37, 4, 11, 27] 아시아 소 품종에 대한 연구는 극히 일부 품종[24]을 제외하고 최근까지 이루어지지 않는 상황이다.

국내 재래소는 현재 4품종이 국제연합식량농업기구(food and agricultural organization of the United Nations: FAO)의 가축다양성 정보시스템(domestic animal diversity information system: DAD-IS)에 등재되어 있으며(Fig. 1), 대표적 乳牛 품종인 홀스타인에 기반한 한국형 젖소 역시 상기 시스템에 등록돼있는 상태이다. 이들 4개의 재래소품종은 널리 알려진 한우와 더불어 칩소, 흑우, 및 제주흑우이며, 특히 한우를 제외한 3개 품종은 FAO가 정한 멸종위기 품종으로 분류되어 있다. 이들의 중요성은 유전자원 보존, 종 다양성 유지 및 우리 고유의 품종에 대한 새로운 축산물의 수요를 창출한다는 측면에서 최근 중요성이 부각되고 있지만 아직까지 이들에 대한 체계적인 전장유전체 연구와 대용량의 유전다형성 발굴 체계가 미약한 실정이다. 이에 이들 4품종(한국형젖소 포함 時 5품종)에 대해 NGS를 이용하여 한우에 대한 전장유전체 해독을 필두로 칩소, 내륙흑우, 제주흑우 및 한국형젖소의 전장유전체 해독결과가 최근 발표되었다.

한우(Hanwoo) 전장유전체해독의 경우 기본적으로 대량의 SNP 정보(~4.7백만개)가 생산됐으며, 캐나다와의 국제공동연구를 통해 북미 대표적 肉牛, 乳牛 품종인 블랙 앵거스, 홀스타인과의 유전자 비교분석을 수행하여 한우에서 16개의 염색체 고정영역을 발견했으며 ACTR3, ARPC2, VIL1, DSTN, AOX1 등과 같이 마블링, 연도와 관련된 유전자들이 다수 존재하는 것으로 조사됐다[25]. 또한 상기 데이터를 이용하여 유전자 복제수 변이(Copy Number Variation, CNV)를 탐색 하였으며,



Fig. 1. Morphological characteristics of Hanwoo, Jeju Heugu, Heugu, and Chikso (presented as clockwise).

한우와 블랙 앵거스 사이 CNV는 1,173개, 한우와 홀스타인 사이 CNV는 총 963개가 발굴 되었다. 이들 CNV의 부분적 결과는 qRT-PCR을 통한 추가적 실험적 검증을 통해 NGS를 적용한 CNV 발굴이 기존 SNP BeadChip array 와 array genomic comparative hybridization (aCGH) 실험방법과 거의 동일한 정확도를 가지고 정밀한 CNV 발굴이 가능함을 보여주었다[8]. 생물학적 의의로서 이러한 전체 발굴된 CNV중 상당수의 증가 CNV(CNV-gain)가 상대적으로 블랙 앵거스와 홀스타인에서 나타났는데 이는 오랜 선발과정에서 CNV가 축적된 결과로 유추할 수 있으며 이는 한우가 앞으로 개량 가능성이 많다는 점을 시사해주고 있다. 이 분석에 사용된 sequencing 플랫폼은 Applied Biosystems by Life Technologies社의 SOLiD 3 이며, color space에 기반한 read 파일(e.g. csfasta)처리는 생물정보학적으로 다소 까다로운 것으로 알려져 왔으나 상기 연구를 통해 이러한 종류의 대용량의 데이터 처리 기반을 구축했다는 점 역시 본 연구의 또 하나의 의의이기도 하다 [25, 37].

최초의 한우 유전체 해독연구를 제외한 재래소 품종에 대한 resequencing 연구는 일루미나(Illumina)社의 sequencing 플랫폼 HiSeq 2000을 통해 최근 이루어졌다. 칩소(Chikso)의 경우 UMD 3.1 reference assembly를 이용한 전장유전체 해독을 통해 약 5.9백만 개의 SNPs 및 약 오십 오만 개의 Insertions and Deletions (InDels)이 발굴 되었다[9]. 발굴된 유전다형성은 Bovine 50K BeadChip 및 sanger sequencing 등을 활용하여 SNP 및 InDels에 대한 실험적 검증이 수행됐으며, 이들

전장유전체에 걸쳐 발굴된 SNP이 상당한 정확도(~99%)를 가진 것으로 추정되었다. 그러나 InDel의 validation 수치는 SNP validation 수치보다 낮았으며(~70%), 이는 기존 Human에서의 연구와 유사한 결과로서 이는 생물정보학 파이프라인에의 InDel calling의 어려움이 아직 완벽히 해소되지 않았음을 의미한다. 따라서, 추후 InDel 분석결과의 해석 및 실제 육종현장에 유전자마커로 사용시 상당한 주의가 요망됨을 시사해주고 있다. 흑우(Heugu)의 전장유전체 해독 역시 동일 시퀀싱 플랫폼으로 UMD 3.1 reference assembly와 비교하여 약 6백만개 이상의 SNP이 발굴 되었다[7]. 또한 제주흑우(Jeju Heugu), 한국형젓소(Korean Holstein) 등이 추가로 해독되어[10] FAO에 등록된 전 대한민국 재래소품종의 전장유전체 해독이 이루어졌으며, 이들 추가된 품종들을 이용 한우의 추가 유전체해독을 통해 총 ~1천 4십만 개의 SNP 발굴 및 1.1백만 개의 Indels이 발굴되었다(54.1% and 78.9% for Novel SNP and InDels, respectively). 한우, 칩소, 제주흑우, 및 한국형 젓소 시퀀싱 데이터를 이용한 CNV 발굴 역시 동일 논문에 통합 발표 되었으며, 한국형젓소를 비교 컨트롤로 잡아 각 유전체와의 총체적 비교분석을 통해 992, 284, 그리고 1,881 개의 CNV가 한우vs乳牛, 제주흑우vs乳牛, 및 칩소vs乳牛에 각각 발굴되었다[10]. 특히 발굴된 대용량의 유전다형성 중 몇 가지 흥미로운 유전자 지역이 즉각적으로 주목됐으며 그 한 예로서 inducible nitric oxide synthase 2 (NOS2)는 재래소품종간 비교 분석 시 한국형젓소에서 지속적으로 적은 뚜렷한 복제수 변이를 보여주었다. 이는 지난 세기를 거쳐 특히 우유와 연관된

형질들에 있어 강력하게 선발된 乳牛의 주요 경제형질과 NOS2와의 관련성을 간접적으로 제시해주고 있으며, 본 연구를 통해 생산된 다양한 유전정보들은 추후 기능성 연구 및 유전체정보의 육종에 활용에 있어 유용한 자료가 되리라 사료된다.

소의 집단유전체 자료 활용 형질 연관 진화분석

집단수준에서의 연구 즉 집단유전체학(Population genomics) 이론의 소 유전체학 분야의 적용은 상대적으로 높은 NGS의 가격으로 인해 아직까지 주로 대용량 고밀도 칩을 활용한 집단기반의 분석 및 유전체 기능해석을 통해 이루어지고 있다. 기본적으로 2009년 소만수체분석 국제공동연구(International Bovine HapMap project)가 그 대표적 예가 될 수 있으며[20], 이후 소전장유전체 해독 및 분석 프로젝트 이후 지속적으로 개발 보완된 commercial Bovine SNP-Chip array에 힘입어 다양한 소 품종에 대해 집단 수준에서 기후변화, 특정 경제형질 등과 연관되어 진화적으로 품종이 개량되었는지에 대한 유전체학적 연구가 가속화 되고 있는 형국이다.

이러한 분석은 기본적으로 전통적으로 사용되던 염기서열의 유전변이를 기반으로 한 Tajima test, 집단간 고정지수(F_{ST}), 이형접합성을 분석하는 연구 등이 있다. 각 분석의 특징을 살펴보자면, Tajima test(1989)는 염기서열 변이가 중립적인 특성을 가지고 있고 집단의 돌연변이와 유전표류(genetic drift)에

대해 평형상태(equilibrium)에 대한 비교로써 선발 여부(natural, purifying selection)를 결정한다. 어떤 집단이 중립평형(neutrality)특정 염기서열에 대한 진화적 역사를 살펴볼 수 있지만 선발(selection)의 기여 정도를 구분하기에 어려움이 있다. F_{ST} 분석은 종(species)에서 지리적, 생태학적 요소에 따라 더 작은 단위의 소집단으로 나뉘어지게 되는데 집단 간의 유전적 차이를 평가하기 위해 가장 대표적인 방법으로 현재까지 활용되고 있다. F_{ST} 는 0에서 1사이의 값을 가지며, 1에 가까울수록 집단의 분화가 많이 일어났음을 의미한다. 최근, 전통적인 집단의 유전적 다양성 분석을 보완하여 integrated Haplotype Score (iHS), Cross Population Extended Haplotype Homozygosity에 기반한 방법인 Rsb 와 XP-EHH 등 연관불평형(Linkage Disequilibrium) 기반의 분석법이 개발되어 좀 더 정밀한 분석이 진행되고 있다. iHS 분석은 집단 간의 일배체형 동형접합율(haplotype homozygosity)를 비교하여 neutral expectation에 대하여 상대적으로 일배체형 동형접합율의 정도가 높은 영역을 진화적으로 선발(selection)된 영역이라 간주한다[39]. iHS 분석은 조상의 대립유전자형(Ancentral allele)을 추정하고, 조상에서 유래된 대립유전자형(derived allele)을 비교하여 집단의 고정이 되지 않은 상태인 최근의 선발신호를 검출할 수 있는 특징을 가지고 있다. Rsb 분석은 iHS 분석과 기본적으로 유사한 방법을 사용하나, 조상의 대립유전자형을 추정하지 않고, 집단 간에 일배체형의 동형접합율

Table 1. Recent studies for selection signatures analyses that were associated with economic traits based on bovine population data

Breed information	Economic traits	Genomic regions (Mbp)	Candidate genes	Methods	References
Holstein, Brown swiss, Simmental, North American Angus, Piedmontese, Australian Angus, Brahman, Belmont Red, Hereford, Murray Gray, Santa Gertrudis, Shorthorns	Milk fat, Muscle formation	BTA4:12.5, BTA8:40.5, BTA10:30.0, BTA10:43.5, BTA18:58, BTA22:26.0	<i>SPATA17, MGAT1, PGRMC2, ACTC1, COL23A1, MATN2</i>	F_{ST} , iHS	Qanbari 등 (2011)
Holstein,	Milk fat, Milk protein	BTA6:12	<i>DGAT1, Casein cluster, GRH</i>	F_{ST}	Hayes 등 (2009), Flori et al (2009)
African taurine, European taurine, African hybrid, West African zebu, East African zebu	Trypanosomosis	BTA3:6.0	<i>CXCR4</i>	F_{ST}	Dayo 등(2009)
Piedmontese	Calving ease	BTA6:37.8 - 38.7	<i>LAP3, LCORL</i>	LD(Linkage Disequilibrium)	Bongiorni 등(2012)
Anatolian Black, Illyrian Mountain Buša, German Fleckvieh, Murnau-Werdenfelser, Franken Gelbvieh, Braunvieh, Red Holstein, Blanc-Bleu Belge, Galloway	Growth, milk-clotting, milk yield, red coat color, fertility	BTA4:4.6, BTA6:12, BAT18:12.2	<i>NPY, Casein cluster, MC1R, SALL1, NKD1</i>	XP-EHH	Rothammer 등(2013)
Hanwoo, Black angus, Holstein	Fat metabolism, disease resistance	BTA2: 95.3-96.4, BTA2: 100.9-101.4	<i>CREB1, CD28</i>	ROH(Runs of Homozygosity)	Lee 등(2013)

을 비교한다는 차이점이 있다. 특히 Rsb 와 XP-EHH의 방법은 intermediate allele만을 잡아내는 것이 아닌 집단 고정(fixation)에 근접한 selective sweep을 전장 유전체 수준에서 추출해 낼 수 있다는 장점이 있다[35, 38].

소의 집단 수준에서 대용량 유전체 자료의 생산이 가능해짐에 따라 *Bos taurus* 내 肉牛(beef cattle)와 乳牛(dairy cattle) 품종 간에 진화적으로 유전체 구조의 차이가 존재함을 밝혀졌다[18]. 또한, 홀스타인 품종에서 선발신호(Selection signatures) 분석을 진행한 결과, 유지방 성분에 영향을 주는 DGAT1(Diacylglycerol O-acyltransferase), Casein 유전자군, 성장호르몬(growth hormone receptor, GRH)등의 유전자가 존재하는 영역이 진화적으로 통계적으로 유의하게 선발되었

음을 알 수 있었다[31]. 뿐만 아니라, 유럽, 아프리카, zebu 소를 중심으로 집단을 통합하여 기후변화와 관련된 내서성 형질 및 면역에 영향을 미치는 유전자를 밝히기도 하였다[16]. 상기에에서 보여지듯 진화적 관점에서의 집단 간 혹은 집단 내 유전적 다양성 비교는 그 집단이 어떻게 관리되어 왔는지를 설명해 줄 뿐만 아니라 가축 각 품종의 특징을 규명할 수 있는 기초자료로 활용되어 분석할 수 있다(Table 1) [3, 12, 15, 25, 32]. 현재까지 연구된 진화분석은 대용량 SNP chip 기반의 분석이지만, NGS 자료가 소의 집단 수준에서 진화 분석이 진행된다면 좀 더 정밀한 마커 정보를 가지고 형질과 연관된 유전체 영역을 탐색할 수 있을 것이라 사료된다[11, 20].

국내 재래소에 대한 NGS를 이용한 집단수준에서의 분석은

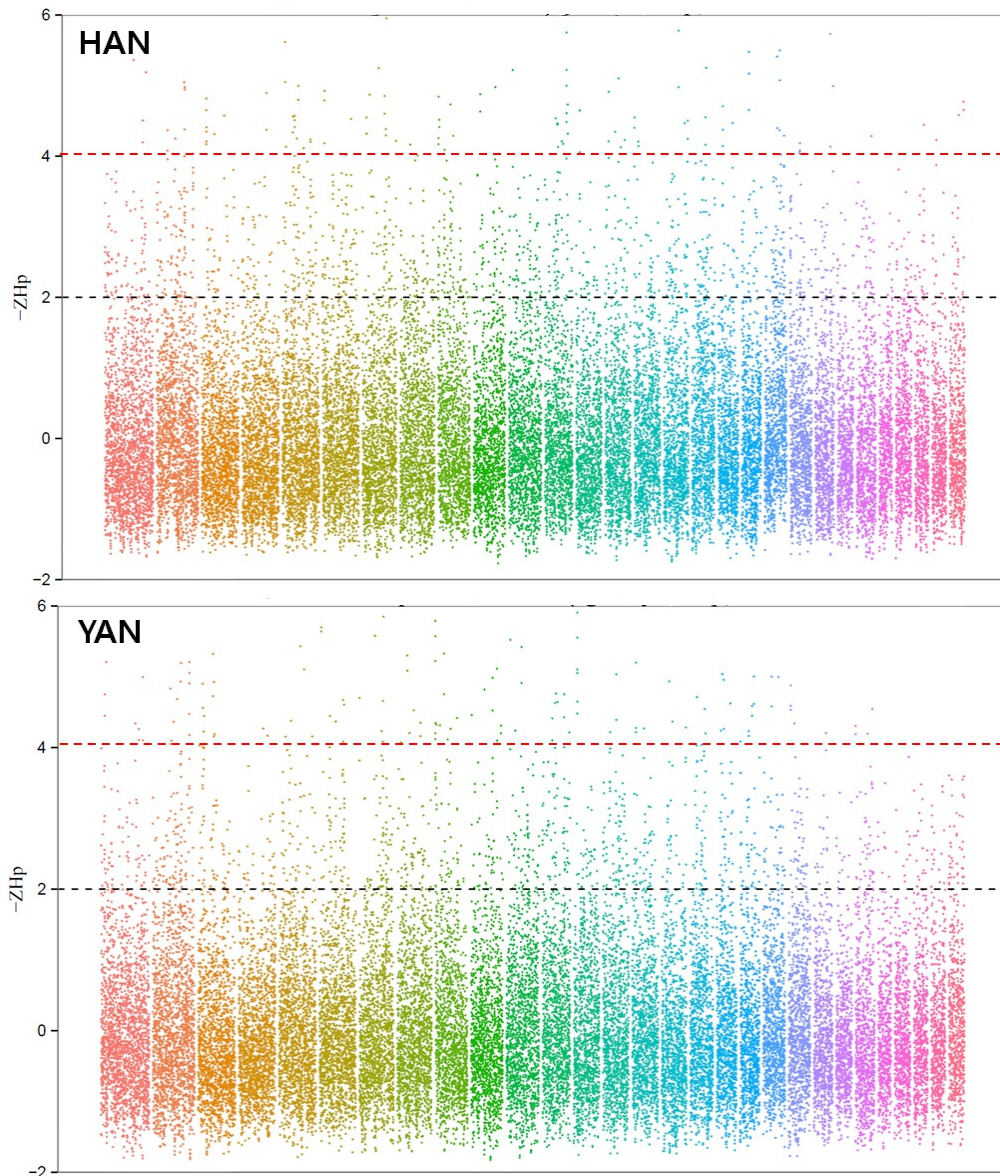


Fig. 2. Manhattan plot of ZHp scores across 29 autosomes for HAN and YAN. The black and red horizontal dotted-lines indicate ZHp thresholds of -2 and -4 that could be putative selective sweep candidates in this study.

최근 한우와 연변우 각 10개체에 대하여 Illumina HiSeq 2000 플랫폼을 통해 전장유전체해독 및 분석이 수행 되었으며, 이들 연구결과는 최근 국제학술지에 투고된 상태이다[6]. 본 연구를 통해 방대한 분량의 SNP 발굴 및 기존 각 품종 단수개체에서 수행키 어려운 선발신호(signature of selection) 발굴이 수행되었으며 전장유전체에 걸쳐 다수의 후보 유전자 지역이 한우의 개량프로그램에 의한 선발과 관련이 있을 것임이 밝혀졌다(Fig. 2). 더 나아가 최근 다부처 유전체 사업출범과 함께 다양한 종의 재래가축에 대한 집단수준에서의 연구가 진행 혹은 예정돼 있으며, 이들 연구는 기존 단수 혹은 소수개체의 전장유전체해독 연구의 한계인 집단유전체학 이론의 NGS에 대한 적용을 보다 가속화 시켜줄 것이라 기대한다.

결 론

최근 가축유전(체)학 분야는 chip technology 및 NGS의 기술적 발전에 힘입어 다양한 (品)種에 대한 유전적 구조, 주요 경제형질과의 연관성 및 실질적 개량프로그램 적용을 위한 광범위한 연구가 전 세계적으로 진행 중에 있다. 본 총설에서는 주요 경제동물인 소에 한정하여 NGS를 이용한 소 전장유전체 연구현황 및 국내 재래소품종에 대한 최근 연구성과 결과를 간략히 소개하였다. 이들 연구들은 기술적, 가격적 어려움으로 인해 대부분 resequencing 형태로 각 품종당 소수의 개체에 대해 수행 되었으며 이를 통해 전장유전체에 걸친 대용량의 SNP 발굴에 성공하였다. 더 나아가 기술적으로 발굴이 까다로운 유전체의 구조적 변이에 대한 특히 CNV 연구 역시 재래소품종에서 수행되었으며 이를 통해 전장유전체에 널리 분포한 대용량의 CNV 발굴 및 주요 경제형질과 관련된 후보 CNV의 발굴에 성공하였다. 그러나 최근 이루어진 NGS의 적용은 각 품종의 단수개체에 대한 유전체 해독이 주를 이뤄 집단수준에서의 연구가 어렵다는 한계가 있었으나, 최근 지속적인 시퀀싱 가격의 하락 및 다부처유전체 사업과 같은 사업추진으로 인해 집단유전체 연구와 NGS의 접목이 가시화되고 있는 형국이다. 추후 이러한 정보는 실제 개량 프로그램에 적용하기 위한 유용한 유전정보 자원으로 사용될 것이라 사료된다.

References

- Bickhart, D. M., Hou, Y., Schroeder, S. G., Alkan, C., Cardone, M. F., Matukumalli, L. K., Song, J., Schnabel, R. D., Ventura, M., Taylor, J. F., Garcia, J. F., Van Tassell, C. P., Sonstegard, T. S., Eichler, E. E. and Liu, G. E. 2012. Copy number variation of individual cattle genomes using next-generation sequencing. *Genome Res.* **22**, 778-90.
- Bolormaa, S., Hayes, B. J., Savin, K., Hawken, R., Barendse, W., Arthur, P. F., Herd, R. M. and Goddard, M. E. 2011. Genome-wide association studies for feedlot and growth traits in cattle. *J. Anim. Sci.* **89**, 1684-1697.
- Bongiorni, S., Mancini, G., Chillemi, G., Pariset, L. and Valentini, A. 2012. Identification of a short region on chromosome 6 affecting direct calving ease in Piedmontese cattle breed. *PLoS One* **7**, e50137.
- Canavez, F. C., Luche, D. D., Stothard, P., Leite, K. R., Sousa-Canavez, J. M., Plastow, G., Meidanis, J., Souza, M. A., Feijao, P., Moore, S. S. and Camara-Lopes, L. H. 2012. Genome sequence and assembly of *Bos indicus*. *J. Hered.* **103**, 342-348.
- Choi, S. B., Byun, M. J., Kim, Y. S., Kim, M. J., Choy, Y. H., Kim, D. H., Jeong, E. G., Kang, K. S., Kim, K. H. and Kim, J. H. 2012. National management system for conservation of livestock genetic resources: an overview. *Ann. Anim. Resour. Sci.* **23**, 142-148.
- Choi, J. W., Choi, B. H., Lee, S. H., Lee, S. S., Yu, D., Chung, W. H., Lee, K. T., Chai, H. H., Cho, Y. M. and Lim, D. 2015. Whole-genome resequencing analysis of Hanwoo and Yanbian cattle to identify genome-wide SNPs and signatures of selection. *Mol. Cells* Submitted.
- Choi, J. W., Chung, W. H., Lee, K. T., Lee, J. W., Jung, K. S., Cho, Y., Kim, N. and Kim, T. H. 2013. Whole Genome Resequencing of Heugu (Korean Black Cattle) for the Genome-Wide SNP Discovery. *Kor. J. Food Sci. An.* **33**, 715-722.
- Choi, J. W., Lee, K. T., Liao, X., Stothard, P., An, H. S., Ahn, S., Lee, S., Lee, S. Y., Moore, S. S. and Kim, T. H. 2013. Genome-wide copy number variation in Hanwoo, Black Angus, and Holstein cattle. *Mamm. Genome* **24**, 151-163.
- Choi, J. W., Liao, X., Park, S., Jeon, H. J., Chung, W. H., Stothard, P., Park, Y. S., Lee, J. K., Lee, K. T., Kim, S. H., Oh, J. D., Kim, N., Kim, T. H., Lee, H. K. and Lee, S. J. 2013. Massively parallel sequencing of Chikso (Korean brindle cattle) to discover genome-wide SNPs and InDels. *Mol. Cells* **36**, 203-211.
- Choi, J. W., Liao, X., Stothard, P., Chung, W. H., Jeon, H. J., Miller, S. P., Choi, S. Y., Lee, J. K., Yang, B., Lee, K. T., Han, K. J., Kim, H. C., Jeong, D., Oh, J. D., Kim, N., Kim, T. H., Lee, H. K. and Lee, S. J. 2014. Whole-genome analyses of Korean native and Holstein cattle breeds by massively parallel sequencing. *PLoS One* **9**, e101127.
- Daetwyler, H. D., Capitan, A., Pausch, H., Stothard, P., van Binsbergen, R., Brondum, R. F., Liao, X., Djari, A., Rodriguez, S. C., Grohs, C., Esquerre, D., Bouchez, O., Rossignol, M. N., Klopp, C., Rocha, D., Fritz, S., Eggen, A., Bowman, P. J., Coote, D., Chamberlain, A. J., Anderson, C., Van-Tassell, C. P., Hulsege, I., Goddard, M. E., Guldbbrandtsen, B., Lund, M. S., Veerkamp, R. F., Boichard, D. A., Fries, R. and Hayes, B. J. 2014. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat. Genet.* **46**, 858-865.
- Dayo, G. K., Thevenon, S., Berthier, D., Moazami-Gouzarzi, K., Denis, C., Cuny, G., Eggen, A. and Gautier, M. 2009. Detection of selection signatures within candidate regions underlying trypanotolerance in outbred cattle populations. *Mol. Ecol.* **18**, 1801-1813.
- Eck, S. H., Benet-Pages, A., Flisikowski, K., Meitinger, T.,

- Fries, R. and Strom, T. M. 2009. Whole genome sequencing of a single *Bos taurus* animal for single nucleotide polymorphism discovery. *Genome Biol.* **10**, R82.
14. Fang, X., Mou, Y., Huang, Z., Li, Y., Han, L., Zhang, Y., Feng, Y., Chen, Y., Jiang, X., Zhao, W., Sun, X., Xiong, Z., Yang, L., Liu, H., Fan, D., Mao, L., Ren, L., Liu, C., Wang, J., Li, K., Wang, G., Yang, S., Lai, L., Zhang, G., Li, Y., Wang, J., Bolund, L., Yang, H., Wang, J., Feng, S., Li, S. and Du, Y. 2012. The sequence and analysis of a Chinese pig genome. *Gigascience* **1**, 16.
 15. Flori, L., Fritz, S., Jaffrezic, F., Boussaha, M., Gut, I., Heath, S., Foulley, J. L. and Gautier, M. 2009. The genome response to artificial selection: a case study in dairy cattle. *PLoS One* **4**, e6595.
 16. Gautier, M. and Naves, M. 2011. Footprints of selection in the ancestral admixture of a New World Creole cattle breed. *Mol. Ecol.* **20**, 3128-43.
 17. Hayes, B. J., Bowman, P. J., Chamberlain, A. J., Savin, K., van Tassell, C. P., Sonstegard, T. S. and Goddard, M. E. 2009. A validated genome wide association study to breed cattle adapted to an environment altered by climate change. *PLoS One* **4**, e6676.
 18. Hayes, B. J., Chamberlain, A. J., Maceachern, S., Savin, K., McPartlan, H., MacLeod, I., Sethuraman, L. and Goddard, M. E. 2009. *Anim. Genet.* **40**, 176-184.
 19. International Bovine Genome Sequencing Consortium. 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science* **324**, 522-528.
 20. International Bovine HapMap Consortium. 2009. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science* **324**, 528-32.
 21. International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**, 695-716.
 22. International Swine Genome Sequencing Consortium. 2012. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **491**, 393-398.
 23. Jiang, L., Liu, J., Sun, D., Ma, P., Ding, X., Yu, Y. and Zhang, Q. 2010. Genome wide association studies for milk production traits in Chinese Holstein population. *PLoS One* **5**, e13661.
 24. Kawahara-Miki, R., Tsuda, K., Shiwa, Y., Arai-Kichise, Y., Matsumoto, T., Kanasaki, Y., Oda, S., Ebihara, S., Yajima, S., Yoshikawa, H. and Kono, T. 2011. Whole-genome resequencing shows numerous genes with nonsynonymous SNPs in the Japanese native cattle Kuchinoshima-Ushi. *BMC Genomics* **12**, 103.
 25. Lee, K. T., Chung, W. H., Lee, S. Y., Choi, J. W., Kim, J., Lim, D., Lee, S., Jang, G. W., Kim, B., Choy, Y. H., Liao, X., Stothard, P., Moore, S. S., Lee, S. H., Ahn, S., Kim, N. and Kim, T. H. 2013. Whole-genome resequencing of Hanwoo (Korean cattle) and insight into regions of homozygosity. *BMC Genomics* **14**, 519.
 26. Li, M., Tian, S., Yeung, C. K., Meng, X., Tang, Q., Niu, L., Wang, X., Jin, L., Ma, J., Long, K., Zhou, C., Cao, Y., Zhu, L., Bai, L., Tang, G., Gu, Y., Jiang, A., Li, X. and Li, R. 2014. Whole-genome sequencing of Berkshire (European native pig) provides insights into its origin and domestication. *Sci. Rep.* **4**, 4678.
 27. Liao, X., Peng, F., Forni, S., McLaren, D., Plastow, G. and Stothard, P. 2013. Whole genome sequencing of Gir cattle for identifying polymorphisms and loci under selection. *Genome* **56**, 592-598.
 28. Liu, Y., Qin, X., Song, X. Z., Jiang, H., Shen, Y., Durbin, K. J., Lien, S., Kent, M. P., Sodeland, M., Ren, Y., Zhang, L., Sodergren, E., Havlak, P., Worley, K. C., Weinstock, G. M. and Gibbs, R. A. 2009. *Bos taurus* genome assembly. *BMC Genomics* **10**, 180.
 29. Marklund, S. and Carlborg, O. 2010. SNP detection and prediction of variability between chicken lines using genome resequencing of DNA pools. *BMC Genomics* **11**, 665.
 30. Matukumalli, L. K., Lawley, C. T., Schnabel, R. D., Taylor, J. F., Allan, M. F., Heaton, M. P., O'Connell, J., Moore, S. S., Smith, T. P., Sonstegard, T. S. and Van Tassell, C. P. 2009. Development and characterization of a high density SNP genotyping assay for cattle. *PLoS One* **4**, e5350.
 31. Qanbari, S., Gianola, D., Hayes, B., Schenkel, F., Miller, S., Moore, S., Thaller, G. and Simianer, H. 2011. Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. *BMC Genomics* **12**, 318.
 32. Rothammer, S., Seichter, D., Forster, M. and Medugorac, I. 2013. A genome-wide scan for signatures of differential artificial selection in ten cattle breeds. *BMC Genomics* **14**, 908.
 33. Rubin, C. J., Megens, H. J., Martinez Barrio, A., Maqbool, K., Sayyab, S., Schwochow, D., Wang, C., Carlborg, O., Jern, P., Jorgensen, C. B., Archibald, A. L., Fredholm, M., Groenen, M. A. and Andersson, L. 2012. Strong signatures of selection in the domestic pig genome. *Proc. Natl. Acad. Sci. USA* **109**, 19529-19536.
 34. Rubin, C. J., Zody, M. C., Eriksson, J., Meadows, J. R., Sherwood, E., Webster, M. T., Jiang, L., Ingman, M., Sharpe, T., Ka, S., Hallbook, F., Besnier, F., Carlborg, O., Bed'hom, B., Tixier-Boichard, M., Jensen, P., Siegel, P., Lindblad-Toh, K. and Andersson, L. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* **464**, 587-591.
 35. Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E. H., McCarroll, S. A., Gaudet, R. and et al. 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913-918.
 36. Snelling, W. M., Allan, M. F., Keele, J. W., Kuehn, L. A., McDanel, T., Smith, T. P., Sonstegard, T. S., Thallman, R. M. and Bennett, G. L. 2010. Genome-wide association study of growth in crossbred beef cattle. *J. Anim. Sci.* **88**, 837-848.
 37. Stothard, P., Choi, J. W., Basu, U., Sumner-Thomson, J. M., Meng, Y., Liao, X. and Moore, S. S. 2011. Whole genome resequencing of black Angus and Holstein cattle for SNP and CNV discovery. *BMC Genomics* **12**, 559.
 38. Tang, K., Thornton, K. R. and Stoneking, M. 2007. A new approach for using genome scans to detect recent positive

selection in the human genome. *PLoS Biol.* **5**, e171.

39. Voight, B. F., Kudravalli, S., Wen, X. and Pritchard, J. K. 2006. A map of recent positive selection in the human genome. *PLoS Biol.* **4**, e72.

40. Zimin, A. V., Delcher, A. L., Florea, L., Kelley, D. R., Schatz,

M. C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C. P., Sonstegard, T. S., Marçais, G., Roberts, M., Subramanian, P., Yorke, J. A. and Salzberg, S. L. 2009. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol.* **10**, R42.

초록 : 차세대유전체해독 기법을 이용한 소 유전체 해독 연구현황

최정우 · 채한화 · 유다영 · 이경태 · 조용민 · 임다정*
 (국립축산과학원 동물유전체과)

최근 차세대염기서열해독법(Next Generation Sequencing, NGS)의 급속한 발전에 힘입어, 다양한 가축 중에 대한 전장유전체 수준의 해독 및 분석 연구수행이 가능하게 되었다. 소의 경우 현재 한우, 뿔소, 흑우, 제주흑우 4품종의 재래소가 국제연합식량농업기구 가축다양성 정보시스템에 등록돼 있는 상태이다. 이러한 재래유전자원은 최근 NGS 기술을 이용 전장유전체에 걸친 대용량의 단일염기다형성 정보를 얻는데 성공하였으며, 또한 한국 재래소품종이 유럽기원의 소 품종들과 유전학적으로 차이가 있다는 점이 밝혀졌다. 또한 소 유전체학 분야에서 이 NGS의 응용은 유전체의 구조적 변이 특히 중전 대용량으로 정확한 발굴이 어려웠던 전장유전체에 널리 퍼진 복제수변이의 발굴에 성공적으로 적용되었다. 이러한 일련의 성공에도 불구하고 최근 NGS를 이용한 연구는 내재적인 한계점이 있었는데, 이는 연구 당시 고가의 연구비용 및 분석의 난해함으로 인해 각 대표 소 품종의 단수 또는 소수 개체에 대해서만 적용되었다는 점이 그 대표적 예라 할 수 있을 것이다. 즉, NGS에서 파생된 데이터의 보다 정확한 생물학적 의미를 찾기 위해서는 추가 실험적 검증과 더불어 면밀한 해석이 필요하다는 점을 시사하는 것이다. 최근 차세대염기서열 해독 비용이 지속적으로 하락하고 있으며, 이는 단수개체가 아닌 집단수준에서의 NGS 적용이 가능해짐에 따라 다양한 집단유전체학적 이론이 접목된 연구가 가능해지고 있다. 현재 국내 재래소 품종에 대한 집단수준에서의 연구는 극히 미흡한 상태이나, 이러한 상황은 최근 고밀도 칩, 차세대염기서열 자료와 같은 대용량 유전정보를 생산, 분석 중에 있어 재래가축에 대한 집단수준에서의 연구가 일부 해소될 것으로 기대된다.