

고고유전학의 분석 원리와 최근 고유전체 연구 동향

김태호¹, 우은진², 박순영¹

¹서울대학교 사회과학대학 인류학과 생물인류학 실험실, ²세종대학교 인문과학대학 역사학과

(2018년 11월 1일 접수, 2018년 12월 3일 수정접수, 2018년 12월 5일 게재승인)

간추림 : 고고유전학은 고DNA에 대한 분석을 고고학 및 인류학적 증거와 교차검증함으로써 인류사에 대한 과학적인 사실을 정립하고자 하는 학문이다. 30여 년 전 시작된 고DNA 연구는 급격하게 발전하여 최근 10년간은 전장유전체의 염기서열을 분석하는 고유전체 연구로 그 범위를 확장하였다. 이를 통해 고인류 종들과 현생 인류 고대 집단들의 기원 및 이주 패턴들을 진화유전학적으로 엄밀하게 연구하는 것이 가능해졌다. 본 연구에서는 고고유전학의 전반적인 분석 원리와 최근의 고유전체 연구 성과 및 경향을 검토하였다. 시료 채취 기술 및 통계 분석 방법의 발전, 고인류 및 서유라시아 고대 집단의 고유전체 연구들을 통해 정립된 연구 방법들은 현재 다른 지역들에도 활발하게 적용되고 있는 추세다. 그러나 한반도를 포함한 동아시아 고유전체 연구는 아직 부진한 실정이다. 본 연구를 통해 아직 수행된 바 없는 한반도 고유전체 연구가 진행될 시 어떠한 사실들을 밝힐 수 있는지 그 가능성을 전망하고자 한다.

찾아보기 낱말 : 고고유전학, 고DNA, 고유전체, 고인류, 고대 집단

서 론

생물인류학은 고고학과 상호보완 하며 물질적 증거를 통해 인류의 기원과 진화 및 다양성 등을 밝혀 왔다. 고고학자들은 유물들을 주로 다루며 생물인류학자들은 현대인의 생물학적 특성 및 화석과 고인골의 해부학·형태학적 특징을 주로 다룬다. 그런데 30여 년 전부터 유전학적인 분석을 통해 인류학적 주제들을 다루는 연구가 시도되기 시작했다. 1985년 고고학자이자 분자생물학자였던 Pääbo가 고대 이집트 미라의 조직으로부터 DNA를 추출하여 보고한 논문이 그 시작이었다[1]. 이러한 고DNA(ancient DNA) 분자를 고인골로부터 화학적으로 추출하는 기술도 곧 확립되었다[2]. 결과적으로 생물 시료에서 가장 근원적인 정보를 담고 있는 생화학적 자료인 DNA 염기서열을 고DNA로부터 분석하는 연구가 시작되었다.

특히 지난 30년 동안 분자생물학 분야에서 일어난 두 가지 기술적 발전 덕분에 체계적인 고DNA 연구가 가능해졌다. 첫째로, 고DNA의 양을 급격하게 증폭시킬 수 있는 중합효소연쇄반응(polymerase chain reaction, 이하 PCR) 기술이 1987년에 개발되었다[3]. 이전에는 원핵생물의 DNA에 연구하고자 하는 DNA를 삽입하여 증식하게 한 뒤 균주로부터 다시 DNA를 분리하는 클로닝(cloning) 과정을 거쳐야 DNA의 양을 늘릴 수 있었다[1]. 그러나 고DNA는 그 양이 극히 제한된 만큼 분석을 위해서는 시료의 양을 크게 늘리는 것이 중요했다. 클로닝보다 훨씬 효율적으로 DNA를 증가시키는 PCR의 등장은 이를 쉽게 해결하여 고DNA 연구가 활발해지는 데 크게 기여했다. 그런데 PCR은 최소 60~70 bp 크기의 DNA만을 증폭시킬 수 있다. 이 때문에 DNA의 사후 손상(postmortem damage) 과정 중 하나인 파편화(fragmentation)에 따라 생성된 60~70 bp보다 작은 크기의 고DNA는 PCR을 적용하기 어렵다는 한계점이 있었다[3].

두 번째 발전은 DNA에 담겨 있는 실질적인 생체 정보인 염기서열을 단시간에 대용량으로 분석할 수 있는 일명

저자(들)는 '의학논문 출판윤리 가이드라인'을 준수합니다.
저자(들)는 이 연구와 관련하여 이해관계가 없음을 밝힙니다.
교신저자 : 박순영(서울대학교 사회과학대학 인류학과)
전자우편 : suny@snu.ac.kr

차세대염기서열분석법(Next-Generation Sequencing, 이하 NGS)의 등장이다. NGS 이전에는 PCR로 증폭한 짧은 DNA 조각의 염기서열을 분석해야만 했는데 Sanger 염기서열분석법(Sanger sequencing)이 대표적인 염기서열 분석 방법 중 하나였다[3]. 이러한 방식으로 긴 DNA 염기서열을 분석하기 위해서는 PCR을 통해 얻은 염기서열들을 각각 분석한 뒤 연결하는 작업이 필요했으므로 장시간이 소요되었다. 즉, 유전체(genome) 전체를 분석하는 데에는 적합하지 않았다. 그런데 2000년대 중후반부터 대용량의 염기서열을 단시간에 분석하기 위한 방법들이 개발되었다. 이를 통틀어 NGS라고 부른다. NGS를 통해 시료의 유전체 전체를 단시간에 모두 분석하는 전장유전체 염기서열분석(whole genome sequencing)이 가능해졌다. 일부의 유전자만을 분석할 수 있던 기존의 유전학과 달리 유전체 전체를 다루는 유전학은 유전체학(genomics)이라는 새로운 생물학 분과 학문이 되었다. 고DNA에 대해서도 이전에는 유전체의 일부분인 미토콘드리아 DNA (mitochondrial DNA, 이하 mtDNA)나 Y염색체의 염기서열만을 분석했지만 NGS의 발전 및 상용화가 이뤄지면서 고유전체(ancient genome)의 염기서열도 분석할 수 있게 되었다. 특히 NGS는 분석하고자 하는 DNA를 인위적으로 최소 20~30 bp의 짧은 DNA 조각들로 나눈 후 동시에 염기서열을 분석하여 단시간에 긴 염기서열을 얻어내는 방식이다. 따라서 이미 파편화된 조각들로 존재하는 고DNA를 분석하는 데 적합하며, 더 이상 PCR을 거칠 필요도 없게 되어 상술했던 PCR 시료의 최소 DNA 크기 문제도 해결되었다[3].

본 연구에서는 이렇게 분자생물학 기술의 발전에 따라 활발하게 수행되고 있는 고DNA 연구의 기본적인 분석 원리와 최신 동향을 검토하여 파악하고자 한다. 고DNA는 미생물을 포함한 여러 생물종들에 대해 광범위하게 연구되고 있지만 이 연구에서는 고인류 및 현생 인류 집단들에 대한 연구로 한정하여 기존의 고고학·생물인류학적 증거를 어떻게 보완할 수 있는지에 초점을 맞추었다. 이를 통해 고고유전학 최신 연구들의 경향을 파악하고 앞으로 국내 고고유전학 연구가 나아갈 방향을 제시하는 데 기여할 수 있기를 기대한다.

한편, 고DNA를 연구하는 분과들은 여러 명칭을 가지고 있다. 고생물학처럼 멸종한 종을 대상으로 하는 경우 고유전학(paleogenetics)이 일반적인 명칭이며, 상술했듯이 인간 외에도 다양한 생물종들에 대한 고유전학 연구들이 수행되고 있다. 하지만 기존의 고고학 증거나 생물인류학 증거와 교차검증 가능한 주제를 다루는 경우는 주로 고고유전학(archaeogenetics)이라고 부른다. 예를 들면 최근의 생물고고학이나 고병리학 연구들은 유물이나 고인골로부터

발견된 식물, 병원균 등의 고DNA를 분석하여 다른 증거들과 비교검증하기도 하는데 이 또한 고고유전학 연구에 해당한다. 따라서 고인류나 현생 인류의 고DNA를 분석하는 연구에 한해서는 인류유전학(anthropological genetics)이라는 말도 드물게 혼용되고 있다. 본 연구에서는 한국 고고학·생물인류학계에서 관용적으로 가장 많이 통용되어 온 고고유전학으로 명칭을 통일한다.

재료 및 방법

고고유전학의 전반적인 분석 방법과 최근의 연구 경향을 검토하기 위해 먼저 고고유전학의 근간이 되는 진화유전학(evolutionary genetics)적 분석의 일반적인 전제들과 분석 대상들을 정리하였다. 특히 유전학으로부터 유전체학이 새롭게 발전함에 따라 진화유전학의 연구 대상도 일부 유전자들로부터 전장유전체로 범위가 확장되었다. 이러한 흐름이 최근 고고유전학 연구들이 고유전체 연구로 발전하는 데 어떤 영향을 미쳤는지를 함께 다루었다. 또한, 고유전체 연구의 가장 큰 문제는 오래된 뼈에서 시료를 채취하는 데에서 발생한다. 최근 연구들에서 밝힌 시료 채취 문제 해결 방안들에 대해서도 검토하였다.

최신 연구 경향에 대해서는 최근 10년간 출판된 고DNA 논문들 중에서도 전장유전체를 분석한 연구들에 한해 검토하였다. 고유전체 연구가 불가능했던 이전에는 고DNA에서 어떤 부분을 분석하는지에 따라 결과가 달라질 수 있었기에 아프리카 기원론(Out of Africa theory) [4] 외에는 고대 인류 집단들의 기원과 이동에 대해 명확히 지지되는 이론이 드물었다. 하지만 고유전체 연구를 통해 과학적인 이론의 정립이 가능해지면서 미국의 사이언스(Science)와 영국의 네이처(Nature) 두 저널에서 많은 고DNA 논문들이 출판되고 있다. 또한, 고DNA가 아닌 현존하는 현대인 집단들의 유전체(modern genome)만을 다루는 최근 연구들에서도 이들의 기원이 되는 현생 인류(*Homo sapiens*)의 고대 집단(ancient populations)들을 추정하는 분석이 함께 수행되고 있는데 이는 고유전체 연구의 핵심 주제이다. 본고에서는 이러한 연구들도 포함하여 최근 10년간 출판된 유전체 분석 연구 중 고인류, 고대 집단을 다룬 논문들을 사이언스, 네이처 두 저널을 중심으로 검토하였다.

검토 결과 현재까지의 고유전체 연구는 주로 유럽을 포함한 서유라시아(Western Eurasia) 지역의 고인골들을 대상으로 이루어졌다는 사실을 확인할 수 있었다. 네안데르탈인(Neanderthals 또는 Neandertals)과 데니소바인(Denisovans)의 뼈로 최초의 고유전체 연구가 시작된 뒤 서유라시

아 지역에서 발견된 현생 인류 고인골들로 연구 범위가 확장되었다. 이외 지역의 고유전체 연구들은 아직 태동하고 있는 단계이기에 고대 집단들의 기원과 이동에 대해 확립된 사실들이 상대적으로 적은 실정이다.

따라서 본 연구에서는 먼저 고인류 및 서유라시아 지역을 대상으로 한 대표적인 논문들을 검토하여 고유전체 연구들의 전형적인 특징 및 주제들을 정리하였다. 그리고 이러한 특징들이 타 지역들에 대해서는 어떻게 적용되고 있는지 남아시아와 중앙아시아, 아메리카, 오세아니아, 아프리카, 동아시아의 순서로 검토하였다. 이렇게 연구 논문들을 검토한 결과는 고고유전학자 Reich가 2018년에 출간한 책에서 제시한 틀을 대체로 따랐다[5].

특히 동아시아 지역에 대해서는 다른 지역들과 달리 아직까지도 고유전체 연구들이 많지 않은 상황이다. 한반도 고대 집단들의 기원과 이동을 논하기 이전에 필요한 동아시아 고대 집단들의 기원과 이동에 대해 매우 제한적인 사실들이 수립되어 있다. 본고에서는 먼저 한반도 이외 동아시아 지역들의 현대인 유전체 및 고유전체 연구들을 검토하여 이러한 사실들을 정리하였다. 마지막으로 한반도 지역에 한해서만 현대인 DNA(modern DNA)와 고DNA 양쪽에 대해 모두 유전체 수준의 연구가 없었던 상황을 감안하여 mtDNA, Y염색체 등 부분적으로나마 DNA를 다른 집단유전학 및 고고유전학 연구들을 모두 검토하였다. 이를 통해 국내 고고유전학 연구들의 성과 및 한계를 정리한 후 타 지역에 대한 최근 연구들처럼 전장유전체 수준의 분석이 접목될 때 어떤 연구가 가능할지를 고찰했다.

결 과

1. 고고유전학의 분석 원리

1) 연구의 기본 전제

고인류 및 현생 인류를 대상으로 하는 고고유전학은 인간의 소진화(microevolution)를 다루는 인간진화유전학(human evolutionary genetics)이라고 할 수 있다. 고인류 종들 및 현대인 고대 집단들의 고DNA 염기서열을 현존하는 현대인 집단들의 DNA 염기서열과 비교분석함으로써 고대 인류 집단들의 기원, 유래 및 이주 양상 등을 소진화적으로 재구성하기 때문이다. 이러한 연구 방식은 유전정보를 통계적으로 분석함으로써 집단(population, 지리적으로 함께 분포하는 한 생물종의 무리를 뜻하며 ‘개체군’으로도 번역 가능)들이나 다른 생물종들 간의 지리적 관계 및 진화사를 재구성하는 진화유전학의 방법을 응용한 것이다. 집

단유전학(population genetics), 계통유전학(phylogenetics), 계통지리학(phylogeography) 등이 이러한 분석을 사용하는 진화유전학 분과들이다.

이러한 진화유전학적 분석의 기본적인 전제들이 어떻게 고고유전학에 적용되는지를 크게 3가지로 정리할 수 있다. 첫째, 한 종이지만 지리적으로 격리되어 있던 집단들이 인접한 지역으로 이주하면 집단 구성원 간의 상호교배가 자연적으로 일어난다. 이를 통해 발생한 자손들은 재조합된 DNA에 의해 양쪽 집단의 평균적인 유전적 특징을 절반씩 가진다. 따라서 이렇게 세대가 계속될수록 자손들이 가지는 양쪽 집단의 유전적 특징들은 지속적으로 섞이게 된다. 이러한 현상을 두 집단 간 유전자 흐름(gene flow)이 존재한다고 표현하는데 고고유전학을 포함한 인간유전학 분야에서는 집단 혼합(admixture)이라는 표현도 널리 쓰이고 있다. 요컨대 서로 다른 집단들이 인접한 지역에서 함께 분포할 시 교배에 의한 집단 혼합이 발생하여 하나의 집단이 될 가능성이 높다. 집단 혼합이 유지된 시간이 길수록 자손들의 DNA에서 양쪽 집단의 DNA가 더 많이 혼재되므로 이를 통해 조상 집단들의 유전적 특징을 유추하기는 어려워진다. 반대로 집단 혼합 후 적은 시간이 지났다면 상대적으로 추정이 더 쉽다.

둘째, 집단들이 오랫동안 격리될수록 각 집단의 평균적인 유전적 특징은 서로 달라지기에 지역별 분포집단들을 유전적으로 다르게 규정할 수 있으며, 이렇게 집단마다 가진 고유한 유전적 특징을 통해 한 집단으로부터 여러 집단으로 갈라진 시기도 추정 가능하다. 이전에 같은 지역에 있던 집단이 두 집단으로 나뉘어 서로 다른 지역에 정착할 경우 두 집단은 유전적으로 다른 집단으로 분기한다. 유전적 표류(genetic drift) 및 돌연변이(mutation)라는 두 가지 진화 메커니즘에 의해 이를 설명할 수 있다. 유전적 표류는 집단이 분리되거나 세대가 바뀌면서 평균적인 유전적 변이의 빈도가 무작위적으로 달라지는 현상을 의미한다. 이 원리에 의해 한 집단이 여러 집단으로 분리될 때 각 집단은 평균적으로 다른 유전적 특징을 가진다. 즉, 분리 시점부터 이미 각 집단이 가지는 유전적 변이는 서로 다르다. 이후 각 집단에서 세대가 지속되면서 집단 간 유전적 차이는 점점 커지는데 이는 주로 돌연변이에 기인한다. 유전물질인 DNA의 모든 부위들은 일정한 돌연변이율을 가진다. 이러한 무작위 돌연변이에 의해 새로운 유전적 변이가 지속적으로 생성되어 자손에게 전달된다. 이때 집단이 개별적으로 놓인 환경에 따라 생존에 이점을 제공하는 특정한 유전자가 자연선택(natural selection)될 수도 있지만 생존이나 기능에 지대한 영향을 주지 않는 ‘중립적인’ 돌연변이는 무작위적으로 자손에게 전달된다. 결과적으로 다음 세대의

유전적 변이 빈도가 무작위적으로 바뀌는데 이 과정도 유전적 표류의 일종이다. 정리하면 원래 하나였던 집단이 두 집단으로 나뉘어 서로 다른 지역에 정착할 경우 1차적인 유전적 표류에 의해 초기부터 유전적 차이가 발생하며, 시간이 경과하면서 중립적인 돌연변이에 의한 2차적인 유전적 표류에 의해 두 집단의 유전적 특징은 지속적으로 달라진다. 중립진화 이론(neutral theory of molecular evolution)에 따르면 집단 간에 관찰되는 대부분의 유전적 차이는 이렇게 자연선택되지 않은 중립적인 돌연변이 및 유전적 표류에 의해 결정된다. 최근의 인류 진화에서 발생한 유전적 다양성에 자연선택은 매우 드물게 기여했음이 유전학적으로 검증되기도 했다[6,7]. 결론적으로 집단 간 유전적 차이는 대부분 중립진화에 의해 축적된 결과이다. 따라서 한 세대의 돌연변이율을 분자시계(molecular clock) 지표로 사용하면 집단 간 유전적 차이를 양적으로 분석하여 과거에 한 집단으로부터 여러 집단으로 분기한 시기를 추정하는 것이 가능하다.

셋째, 유전적 차이가 커지면 종 분화가 일어날 수 있다. 일단 종 분화가 발생하면 유사한 특징을 가진 종들이더라도 상호교배 시 생식상의 문제가 발생할 확률이 높아진다. 이종 간 교배가 아예 불가능하거나 가능하더라도 자손은 생기지 않을 수 있으며, 자손이 생기더라도 생존력이나 생식력에 유전적 문제가 생기는 경우가 많다. 따라서 종 분화 이후엔 지리적으로 격리되지 않아도 종 분화가 유지되는 경향이 있다. 일례로 네안데르탈인은 화석 증거를 통해 여러 지역에서 현생 인류와 함께 분포한 것으로 추정된다. 이 종은 형태학적인 분석에 의해 먼저 현생 인류와 다른 종으로 규정되었으나 이후 고유전체 분석을 통해 유전적으로도 현생 인류와 명확히 구분됨이 확인되었다[8,9].

2) 유전체상의 분석 대상

진화유전학자들은 상술한 전제하에 유전정보인 DNA 염기서열을 수학적 모델링 및 추론통계학 등의 방법으로 분석하여 집단이나 종들 간의 진화적인 거리를 추론한 뒤 계통수(phylogeny)를 재구성한다. 이때 분석하는 DNA 염기서열의 범위는 집단 또는 종 간의 변이가 존재하는 부분들이다. 이렇게 변이가 있는 DNA 부위들의 대표적인 예는 단일염기다형성(single nucleotide polymorphism, 이하 SNP)이다. 전 세계 현대인 집단들의 유전적 다양성을 밝히는 연구들에서는 집단들의 유전정보를 광범위하게 비교하여 이러한 SNP들을 판별해낸다[7]. 전장유전체 염기서열분석이 가능해진 후에는 유전체상에 존재하는 거의 모든 SNP들을 비교분석할 수 있게 되었으나 이전에는 일부 SNP들만을 일종의 표지(marker)로 사용해 왔다.

대표적인 표지들이 mtDNA와 Y염색체상에 존재하는 SNP들이다. mtDNA는 한 세포 내에 한 쌍으로만 존재하는 핵 DNA와 달리 몇 천 개가 존재하며 유전자 재조합이 안 되는 등의 여러 특성상 SNP 분석이 용이하다. 따라서 진화유전학 분야 전반에서 폭넓게 활용된 분석 대상이었다. 특히 한 세포 내 몇 천 개가 존재한다는 점은 시료에 소량 존재하는 고DNA를 추출하는 데도 큰 이점으로 작용하였다. 아울러 mtDNA는 모계로만 유전되기에 전 세계 현대인 집단들에서 이 부분들만 분석해도 진화적인 모계 계통수를 구성할 수 있다. 이를 통해 모든 현대인의 모계 조상은 약 200,000년 전에 아프리카에 존재했다는 사실이 밝혀짐으로써 현생 인류의 다지역 기원설은 기각되고 아프리카 기원설이 지지되었다[4]. mtDNA에 비하면 Y염색체는 분석의 용이성은 부족하지만 부계로만 유전되므로 같은 원리에 의해 부계 계통수를 구성 가능하며 널리 활용되었다.

또한, 계통수에서 한 공통조상을 공유한 모든 자손들의 군집(monophyletic group)을 하플로그룹(haplogroup)이라고 한다. 현존하는 현대인 집단들의 mtDNA 및 Y염색체 변이가 대부분 분석되어 모든 mtDNA 하플로그룹과 Y염색체 하플로그룹을 분류할 수 있다. 따라서 고DNA로부터 mtDNA나 Y염색체의 하플로그룹을 분석한 후 이렇게 현존하는 하플로그룹 정보와 비교함으로써 고대 집단들의 ‘모계 진화사’ 및 ‘부계 진화사’를 구성하는 것이 가능하다.

그러나 mtDNA, Y염색체는 전체 유전체에서 아주 일부분의 정보에 해당하므로 이들만으로는 집단들의 진화사를 완전하게 밝힐 수 없다. 특히 mtDNA, Y염색체는 자손이 부모로부터 단 하나만 받기에 이들을 통해서 두 집단이 아닌 한 집단의 정보만 다음 세대로 전달된다. 반면, 유전체의 다른 부분들은 모두 부모로부터 한 쌍을 받으므로 두 집단의 정보를 모두 포함한다. 특히 여러 집단들이 혼합할 시 세대가 계속될수록 유전정보는 각 집단에서 온 정보들로 혼재되기 때문에 유전체 전체에서 한 집단으로부터 기원한 정보의 양은 지속적으로 감소하게 된다. 유전체의 일부가 아닌 전체를 분석할 경우에는 이러한 잠재적인 조상 집단들의 모든 정보를 포함할 수 있다. 따라서 유전체 수준에서 집단들을 비교분석해야만 모든 조상 집단들의 완전한 진화사를 밝히는 것이 가능하다.

이러한 이유로 전장유전체 염기서열분석이 가능해진 후에는 현존하는 집단들의 유전적 다양성뿐만 아니라 고인류 및 현생 인류 고대 집단들의 고DNA를 연구할 때도 유전체 수준의 분석이 일반적인 분석 방식이 되었다[7-9]. 특기할 만한 점은 일부 유전자를 비교할 때에 비해 유전체를 분석할 때는 대상이 되는 SNP의 개수가 무수히 많아지게 된다. 따라서 적은 SNP를 비교분석할 때보다 더 복잡한 통

계적 방법이 요구된다. 현재 고유전체 연구에 널리 쓰이는 통계 분석 방법으로는 주성분분석, 유전적 군집(genetic cluster) 분석, D-통계량 및 F-통계량 분석 등이 있다.

3) 고DNA의 시료 채취 문제

고DNA는 고인골의 특정 부위를 치과드릴 등으로 뚫은 뒤 나온 뺏가루(bone powder)에 화학적 처리를 하여 DNA를 분리하는 방식으로 얻는다[2,10]. 이때 외부 DNA로 인한 오염을 방지하기 위해 멸균시설 및 특정 장비들을 이용하는 엄밀한 방법으로 시료를 채취한다. 그럼에도 불구하고 고DNA 시료에는 두 가지 중대한 문제점이 있었다. Reich는 이 문제들이 해결되고 난 후에야 고유전체 연구가 훨씬 용이해졌음을 주목하고 있다[5]. 첫째, 시료 채취 방식과는 별개로 외부 DNA의 오염을 근본적으로 방지할 수 없다. 오래된 뼈나 조직일수록 미생물 등의 DNA가 섞일 가능성이 크기에 분석하고자 하는 고DNA와 외부 DNA를 명확히 구분하는 기술이 필요하다. 2013년에 화학적 탐침(probe)의 혼성화(hybridization)를 이용하여 고DNA만 선택적으로 농축시키는 새로운 방법이 개발되어 이 문제가 크게 개선되었다[11]. 또한, 고DNA의 화학적 특성을 지표로 하여 정량적인 평가를 통해 고DNA를 외부 DNA와 구분하는 통계적 모델들도 개발되고 있다[12].

둘째, DNA는 생분자로 영구히 보존되지 않으므로 고DNA는 추정연대가 오래되거나 따뜻한 기후의 지역일수록 화학 반응이 가속화되어 더 쉽게 분해된다는 큰 문제점이 있다. 그러나 최근 새로운 시료 채취 방식이 등장했다. 2015년 출판된 연구에 따르면 관자뼈에서 고DNA를 추출할 경우 다른 뼈에 비해 최대 약 400배나 높은 고밀도의 고DNA를 추출할 수 있다[10]. 이후 열대 지역에서 발굴된 고인골에 대한 고유전체 연구들도 빠른 진척을 보이기 시작했다.

2. 최근 고유전체 연구 동향

1) 고인류 연구: 현생 인류와 다른 호미닌 종들과의 관계

현생 인류 및 가까운 고인류 종들을 포함하는 인간 속(Homo)은 영장류 분류 체계상 인간 아과(Homininae)의 인간 족(Hominini, 이하 호미닌)에 속한다. 호미닌 종들 중 현재까지 대용량의 고유전체 해독에 성공한 종은 현대인 외에 네안데르탈인과 데니소바인, 단 두 종이다. 그 이상 오래된 호미닌 표본들의 경우 인골이 오랜 기간 화석화되면서 고DNA가 충분히 보존되지 않기 때문이다. 비교적 온전한 표본들이 유라시아 각지에서 발굴된 네안데르탈인의 경우 완전한 고유전체 분석이 끝난 상태다[8,9]. 이와 달리 데니소바인은 러시아 시베리아 지역의 데니소바 동굴(Deniso-

va Cave)에서 발견된 큰어금니 1개와 손가락뼈 1개가 발견된 표본의 전부였다. 하지만 이 표본의 DNA 보존 상태가 매우 양호했기에 대용량의 고유전체 분석이 진행될 수 있었다[13,14]. 네안데르탈인의 유전체가 현생 인류의 것과 확연히 구별되듯이 데니소바인의 유전체도 현생 인류, 네안데르탈인의 것과 명확하게 구별되었다. 뼈가 거의 남아 있지 않기에 형태학적으로는 기존의 호미닌 종들과 다른 종으로 규정하는 것이 불가능했으나 유전학적인 분석을 통해 다른 종으로 규정한 첫 사례이기도 하다. 분석 결과에 따르면 데니소바인은 현대인보다 네안데르탈인에 더 유전적으로 가까웠다. 약 804,000년 전에 현대인, 네안데르탈인과 데니소바인의 공통조상이 먼저 존재했고, 약 640,000년 전에 네안데르탈인과 데니소바인의 공통조상이 존재한 것으로 추정되었다[13].

이렇게 유전적으로 구별되는 호미닌 종들의 고유전체 분석 연구는 현생 인류 유전체 연구에 대해 두 가지 중요한 점을 시사하고 있다. 첫째, 10% 미만의 적은 양이지만 현대인들의 유전체에서 두 종의 유전체와 공유하는 부분이 존재한다[8,9,13,14]. 이러한 공유 DNA의 존재를 통해 현생 인류를 포함한 세 호미닌 종들이 공존하고 있던 당시 이들의 이종 간 교배를 통해 자손이 태어날 수 있었고, 살아남은 자손들 중에서도 생존력과 생식력을 갖춘 개체들이 존재했다는 사실이 명백해졌다. 이런 점에서 유전적으로도 두 종을 현생 인류와 완전히 다른 종으로 규정하기는 어렵다.

둘째, 현대인 집단에 따라 두 호미닌 종과 공유하는 DNA의 양이 다르며, 이를 통해 호미닌 종들과의 이종 간 혼합이 일어난 위치를 추론할 수 있다. 유라시아인 집단들의 유전체에서 1~4%가 네안데르탈인 DNA로부터 기원했는데 이는 사하라 이남 아프리카인(Sub-Saharan Africans) 집단들이 평균적으로 가지고 있는 네안데르탈인 DNA보다 더 많은 양이다[8]. 즉, 현생 인류가 아프리카 대륙 밖으로 나온 후에야 네안데르탈인과의 교배가 발생했다고 할 수 있다. 네안데르탈인의 화석이 유럽에서부터 시베리아까지 광범위하게 유라시아에서 발견된 걸 고려하면 현생 인류와 네안데르탈인의 혼합 집단은 유라시아 전역에 분포했을 가능성이 크다. 2014년에는 서시베리아(western Siberia)에서 발굴된 약 45,000년 전의 것으로 추정되는 현생 인류 표본에 대한 고유전체 연구가 출판되었는데 이 연구에 따르면 네안데르탈인과 현생 인류 간의 혼합은 약 52,000~58,000년 전에 발생했다[15]. 이후 새로운 연대 추정 방법을 제안한 최근 연구에서는 그 시기를 약 40,000~54,000년 전으로 추론하였다[16].

이렇게 비교적 분명한 사실이 수립된 네안데르탈인과 달리 데니소바인 DNA의 경우는 분석 결과가 복잡하다. 현대

인 집단들의 유전체와 비교한 결과 멜라네시아인 집단이 다른 어떤 유라시아인 집단보다 더 많은 데니소바인 DNA를 가지고 있었고, 이는 유전체의 4~6% 정도였다[13]. 그런데 멜라네시아인 유전체에 남아 있는 네안데르탈인 DNA의 양은 이 양보다 적기에 과거 멜라네시아인의 조상 집단이 네안데르탈인과 먼저 혼합된 후 데니소바인과 혼합되었음을 유추할 수 있다. 실제 혼합 시기를 추정된 결과, 네안데르탈인과의 혼합은 약 50,000~60,000년 전, 데니소바인과의 혼합은 약 44,000~54,000년 전이었다[17]. 하지만 상술했듯이 데니소바인은 멜라네시아와 지리적으로 떨어져 있는 시베리아 지역의 데니소바 동굴 한 군데에서만 발견되었고 아직까지 비교분석할 만한 동아시아 및 오세아니아 지역의 고유전체 자료도 부족한 실정이다. 또한, 고유전체 자료로부터 추정된 데니소바인의 유전적 다양성은 현대인과 비교했을 때 매우 낮은 편이었다[14]. 데니소바인이 시베리아에서 멜라네시아까지 넓게 지리적으로 분포했다고 가정하면 이는 이례적인 결과다. 따라서 데니소바인과 현생 인류 집단 간 혼합의 구체적인 과정에 대해서는 많은 후속연구가 요구된다.

나아가 최근까지도 호미닌 종들 간의 상호작용을 예측하는 증거가 지속적으로 출판되고 있다[18,19]. 2016년 연구에서는 스페인에서 발굴된 약 430,000년 전의 것으로 추정되는 호미닌 화석으로부터 불완전하게나마 고DNA를 분석했다[18]. 네안데르탈인과는 다른 종으로 추정되는 이 화석은 이미 온전한 DNA가 많이 남아 있지 않았지만 mtDNA는 완전히 복원할 수 있었고, 핵 DNA의 일부분도 얻어낼 수 있었다. 분석 결과, mtDNA는 데니소바인의 것에 더 가깝지만 핵 DNA는 네안데르탈인의 것과 더 유사한 결과가 나왔다. 스페인과 데니소바 동굴 사이 지리적 거리를 감안하면 이는 매우 특이한 결과다. 그러나 분석 수준의 한계를 감안하면 호미닌 화석들에 대한 후속 고유전체 연구가 지속적으로 시도되어야 할 것으로 보인다. 그럼에도 불구하고 이 결과를 통해 아직 예측하기 어려운 호미닌 종들 간의 진화사가 많을 수 있다는 사실이 더 명백해졌다. 이는 호미닌 종들의 유전체를 비교분석한 이전 연구에서 제시했던 알려지지 않은 호미닌 종의 존재 가능성[9]을 암시하는 결과이기도 하다.

가장 최근에는 네안데르탈인과 데니소바인의 이종 교배로부터 자손이 태어날 수 있었음을 증명하는 연구가 출판되었다[19]. 이 연구에서는 데니소바 동굴에서 발견된 수천 개의 뼈조각들(bone fragments) 중 호미닌 종의 것으로 추정되는 한 조각의 유전체를 분석했다. 분석 결과, 이 뼈를 가진 개체는 네안데르탈인(모)-데니소바인(부) 부모의 자식으로 밝혀졌다. 이는 두 호미닌 종 간의 교배가 일어났음

을 가장 직접적으로 보여주는 증거다. 이러한 호미닌들의 고유전체 연구들을 종합하면 과거에 현생 인류를 포함한 호미닌 종들 간의 교배 및 혼합이 드물지 않게 발생했을 것으로 추측할 수 있다.

2) 서유라시아 연구: 고유전체 연구의 전형적 특징들

서유라시아 지역은 고대 근동(Ancient Near East)이라고 지칭되는 서아시아와 유럽 지역을 통틀어 일컫는다. 특히 근동 지역은 약 50,000~100,000년 전 아프리카 대륙 밖으로 이주한 현생 인류 집단[7]이 처음으로 마주한 지역이기도 하다. 고인류 연구를 통해 시작된 고유전체 연구는 현생 인류의 고대 집단들로 범위를 확장했고, 서유라시아 지역에서 발견된 표본들을 대상으로 처음 수행되기 시작했다. 연구 성과가 축적됨에 따라 근동 지역 진출 이후 서아시아로부터 서유럽에 걸쳐 발생한 현생 인류 고대 집단들의 다양한 이주 패턴들이 지역과 시간대에 따라 세밀하게 규명되고 있다[20,21]. 이러한 서유라시아 고유전체 연구들은 현생 인류의 과거사를 밝히는 고고유전학 연구의 전형적인 4가지 특징들을 잘 보여준다. 첫째, 집단의 표본인 고인골이 발굴된 고고학적 맥락에 따라 분포 지역, 추정 시기, 생활 방식, 도구명 등으로 집단을 지칭한다. 예를 들면, 유럽 신석기 수렵채집민(European Neolithic hunter-gatherers)이나 아나톨리아 농경인(Anatolian farmers), 얄나야 스텝 목축인(Yamnaya steppe pastoralists) 등이다. 이러한 명칭은 현대인 집단들을 주로 국가, 인종, 민족 등으로 부르는 것과 구별된다.

둘째, 고고유전학 연구는 고고학적 증거로부터 예측되는 패턴을 상호보완 할 수 있다. 특히 상술한 명칭을 통해 고고학적 맥락에서 추론한 문화의 확산과 고DNA 증거로부터 도출된 집단들의 이동 패턴을 쉽게 비교할 수 있다. 고유전체를 분석해 보면 일반적으로 고대 집단들의 지리적 위치 및 문화적 차이에 따라 집단 간 유전적 차이가 명확하게 나타나는 경향이 있다. 인간 집단들이 각자 거주 지역의 환경에 적응하면서 서로 다른 문화를 발전시킬 수 있다는 점과 앞서 살펴본 중립진화 이론을 종합하면 이는 자명한 사실이다. 그런데 유사한 문화가 여러 지역에서 발견될 경우에는 여러 가능성이 존재한다. 특정 집단의 이주에 의한 것일 수도 있지만 집단 간 모방에 따른 문화만의 확산에 의한 것일 수도 있다. 고고학적 증거만으로는 풀 수 없었던 이 문제는 문화를 공유했던 고대 집단들의 유전적 특징을 분석함으로써 비로소 추정할 수 있게 되었다. 실제로 최근 고고유전학 연구들은 문화와 집단이 함께 움직였음을 보여주는 사례들을 축적하고 있다. 예를 들어, 근동 스텝(steppe) 지역에서 발견된 얄나야(Yamnaya) 문화와 독일 지역에서

발견된 매듭무늬토기(Corded Ware) 문화는 유사한 점을 상당 부분 공유하고 있으며, 둘 다 신석기 시대로 추정된다. 2015년 연구에서는 다양한 서유라시아 지역의 고유전체를 비교분석함으로써 매듭무늬토기 집단 유전체의 약 75%가 얀나야 집단으로부터 기원했다는 사실을 입증했다[20]. 이를 통해 신석기 후기에 유럽 지역으로 얀나야 집단의 대규모 이주가 있었으며, 이 패턴이 두 문화의 유사성과 관련 있음을 밝힐 수 있었다.

셋째, 위와 동일하게 문화의 확산이라는 맥락에서 고고유전학 연구는 언어학적 증거와도 상호보완 가능하다. 물질문화가 유물과 유적을 통해 드러난다면 이에 대비되는 정신문화의 대표적인 사례는 언어다. 최근 연구들은 같은 계통의 언어를 사용하는 집단들이 공유하는 DNA의 양을 통해 언어의 변이도 고대 집단들의 이동 및 혼합과 밀접한 관련을 가지고 있음을 보여주고 있다. 예를 들어, 유럽에 널리 퍼져 있는 인도유럽어족 언어들의 확산 및 기원에 대해서 몇 가지 고고학·언어학적 가설들이 공존하고 있다. 고유전체 연구는 유전체 자료에 따라 가장 지지되는 이주 패턴을 도출하므로 특정 가설을 더 지지할 수 있다. 상술한 2015년 연구에서 밝힌 얀나야 집단의 대규모 이주 패턴은 인도유럽어족 언어가 유목민 집단으로부터 기원했고 바퀴의 발명 후 유럽 전역으로 확산되었다는 가설을 지지하였다[20]. 그러나 2018년 출판된 오세아니아 지역을 대상으로 한 연구에서는 한 지역에 거주하는 집단들의 조성이 바뀌더라도 사용 언어 계통은 유지되는 이례적인 사례가 제시되기도 했다[22]. 이는 고대 집단들 간 상호작용이 아주 복잡할 수 있으며, 고고학·언어학 증거 및 고유전체 증거에 대한 종합적인 분석이 필요하다는 점을 시사한다.

넷째, 현존하는 집단들의 유전체만을 비교분석함으로써 고대 집단들에 대한 가설을 수립할 수 있으며, 이 가설은 해당되는 고유전체 자료가 발견될 때 검증 가능하다. Reich는 이에 대한 대표적인 사례로서 다음의 세 연구들을 제시하고 있다[5]. 먼저 Patterson 등은 현재 고유전체 연구에서 널리 쓰이고 있는 통계 분석방법을 처음 소개하면서 프랑스인 집단의 유전체가 브라질의 원주민 카리티아나족(Karitiana) 집단과 지중해의 사르데냐(Sardinian) 지역 집단의 DNA를 모두 일정 수준 가지고 있음을 보였다[23]. 이 결과로부터 유럽에 분포했던 프랑스, 사르데냐 집단의 조상 집단과 아메리카에 분포했던 브라질 원주민 집단의 조상 집단에 모두 기여한 고대 집단이 있었으리라 추정하였다. 즉, 이 고대 집단이 양쪽으로 분기하여 한쪽은 유럽인 집단의 조상이 되었고, 다른 한쪽은 아메리카 대륙으로 넘어가 아메리카 원주민 집단의 조상이 되었다는 것이 유력한 통계적 가설이었다. 현존하는 현대인 집단들의 유전체만

을 진화유전학적으로 비교분석하여 과거 집단들의 분기를 예측한 것이다. 이 집단은 당시까지 고DNA로 발견된 적이 없는 집단이었기에 고대 북유라시아(ancient northern Eurasia) 집단으로 지칭되었다. 이후 시베리아 말타(Mal'ta) 지역에서 발굴된 약 24,000년 전으로 추정되는 현대인의 고유전체 연구에서 실제로 이 조건을 만족하는 고대 집단의 존재가 입증되었다[24]. 같은 원리로 고유전체 간의 비교 분석을 통해서도 새로운 고대 집단에 대한 가설을 수립할 수 있다. 2016년 연구에서는 근동 지역의 고유전체 분석을 통해 아직 고유전체 증거가 발견되지 않은 가상의 고대 집단인 기저 유라시아 집단(Basal Eurasians)의 DNA가 네안데르탈인 DNA를 거의 포함하지 않을 것으로 예측했다[21]. 이는 유라시아에서 현생 인류의 고대 집단들과 네안데르탈인 간의 혼합이 균일하게 발생하지 않았다는 것을 뜻한다. 앞서 살펴본 데니소바인과 현생 인류 간의 밝혀지지 않은 집단 역학[13,14]과 유사한 사례다. 따라서 현생 인류 고대 집단들의 고유전체 증거를 통해 호미닌 종들과 현생 인류의 이종 간 혼합 집단이 존재했던 시기와 위치도 더 명확하게 추론하는 것이 가능해진다.

정리하면 서유라시아 고유전체 연구들은 고고학적 증거, 언어학적 증거, 현존하는 현대인 집단들의 유전체 증거 등의 기존 연구들을 통합하면서 고대 집단들의 과거사를 점점 세밀하게 정립해 나가는 전형적인 과학적 연구 과정을 잘 보여준다. 새로운 고유전체 증거를 통해 아직 발견되지 않은 고대 집단이나 이동 패턴에 대해 통계적으로 추론하여 가설을 수립할 수 있으며, 이후의 증거를 통해 기존 가설을 지지하거나 기각 가능하다.

3) 타 지역 연구: 전형적 방법들의 적용

서유라시아 지역은 지역별로 현존하는 현대인 집단들의 유전체 정보가 이미 많이 축적되어 이를 토대로 가설을 세울 수 있었고[23], 이후 전장유전체 분석이 끝난 고유전체와의 비교분석도 용이했다. 그러나 타 지역들은 인도인(Indian) 집단들이나 원주민 집단들처럼 현존하는 집단들의 유전체에 대해서도 분석된 자료가 부족했다. 따라서 이러한 집단들의 유전체를 먼저 분석하면서 고대 집단들에 대한 가설들을 세운 후 점차적으로 고유전체 연구들에서 이를 검증해 나가는 과정을 거치고 있다. 이러한 타 지역 연구들은 서유라시아 연구들이 정립한 고고유전학 연구의 전형적 특징들을 잘 따르고 있다.

(1) 남아시아 및 중앙아시아 연구

인도를 포함하는 남아시아 및 중앙아시아 지역은 서유라시아와 인접하고 있음에도 아직까지 고유전체 연구가 많이 이뤄지지 않았다. 인도 내 다양한 현대인 집단들에 대해 처

음으로 광범위한 유전체 분석을 수행한 2009년 연구에서는 인도인 집단들과 서유라시아인 집단들의 유전적 조성을 비교분석하였다[25]. 분석 결과, 서유라시아인과 유전적으로 가까운 집단들과 그렇지 않은 집단들이 구분되었다. 이에 따라 서유라시아인과 유전적으로 가까운 ‘북인도 조상 집단(Ancestral North Indians, 이하 ANI)’과 그렇지 않은 ‘남인도 조상 집단(Ancestral South Indians, 이하 ASI)’이라는 두 가상의 고대 집단이 존재했을 것으로 예측했다. 또한, 연구자들은 ANI, ASI의 고유전체 증거가 없음에도 통계적 분석을 통해 현존하는 인도인 집단들이 ANI, ASI 각각에 얼마나 가까운지를 추정하였다. 특기할 만한 점은 ANI로부터 100% 기원한 집단은 없었던 반면, 육지로부터 격리된 안다만 제도(Andaman Islands)의 원주민 집단인 온게족(Onge)은 거의 100% ASI로부터 유래했다는 점이다. 나아가 후속연구에서는 추가적인 유전체 분석을 수행하여 이 두 고대 집단 간의 혼합이 1900~4200년 전에 발생한 것으로 추론했다[26]. 그러나 아직까지 그에 상응하는 고유전체 증거는 없는 상황이다.

한편, 2009년 연구에서는 인도 현대인 집단들 간의 유전적 차이가 유럽인 집단들 간의 유전적 차이에 비해 더 크다는 점도 제시했다[25]. 이 차이는 집단들 간의 유전자 흐름이 제한되면서 집단 혼합이 자연적으로 일어나지 않을 때 발생할 수 있다. 즉, 집단들이 인접한 지역에서 조우했음에도 서로 상호작용하지 않고 유전적인 창시자 집단들(founder populations)로서 초기의 유전적 특징을 유지할 때 나타나는 현상이다. 이러한 창시자 효과(founder effect)는 앞서 개관한 유전적 표류의 일종이다. 연구자들은 분석 집단들의 유전체에 약 30~100세대 이전의 창시자 효과를 암시하는 유전적 신호가 현재까지도 잘 남아 있다는 사실을 통해 이를 실증하였다. 이러한 집단 간 유전적 차이는 인도 내 엄격한 족내혼 관습으로부터 기인한 것으로 추측된다[25]. 이 사례는 집단들이 인접한 지역에 분포할 때 집단 간 상호교배가 자연적으로 발생하여 자손을 생산하고, 결과적으로 하나의 집단으로 혼합된다는 진화유전학적 전제가 인간을 대상으로 할 때는 항상 유효하지는 않음을 잘 보여준다.

최근 2018년 연구에서는 중앙아시아부터 남아시아에 이르는 넓은 지역에서 발굴된 고인골들로부터 광범위하게 고유전체를 분석하였다[27]. 연구자들은 분석 대상 중 한 고대 집단이 고고학적 증거에 따라 고대 인더스 문명과 문화적으로 접촉한 것으로 추정되어 인더스 인접(Indus Periphery) 집단이라고 명명하였다. 이 집단은 아직까지 발견된 바 없는 고대 인더스 문명인들의 유전적 특징을 가지고 있는 집단일 수 있으며, ANI 집단의 형성에 기여한 조상 집

단으로 추정된다. 향후 더 많은 고유전체 연구가 진행될수록 중앙아시아와 남아시아에 분포했던 다양한 고대 집단들의 이주 및 혼합 양상이 기존에 수립된 서유라시아의 고대 집단 역학과 점차 연결될 것으로 기대된다.

(2) 아메리카, 오세아니아 및 아프리카 연구

아메리카, 오세아니아 및 사하라 이남 아프리카 지역의 공통점은 현재까지도 수렵채집 양식을 유지하기도 하는 원주민 집단들과 대항해시대 이후 이 지역들로 진출한 유럽인들로부터 기원한 집단들이 기본적으로 함께 분포하며, 원주민들의 역사는 기록으로 잘 남아 있지 않다는 점이다. 문명의 등장을 기점으로 지역 집단들 간의 이주, 혼합, 전쟁 등의 역사가 사료로써 비교적 잘 정립되어 있는 유라시아 지역과 크게 다른 점이다. 따라서 이 지역들은 유럽인들의 진출 이전의 역사를 재구성하기 위해서 고고학 및 인류학적 증거가 필수적이었다. 고유전체 연구의 발전 이후 이러한 기존의 증거들을 획기적으로 보완할 수 있게 되었다. 즉, 이 지역들은 타 지역에 비해 고유전체 분석 결과로 밝힐 수 있는 역사의 범위가 상대적으로 더 넓은 편이다. 먼저 살펴본 남아시아 사례와 유사하게 지역 원주민들의 유전체 분석을 통해 고대 집단 역학을 추정한 후 고유전체 분석으로 검증하는 순서로 연구들이 진행되었다.

먼저 아메리카 지역에 대해서는 2012년에 출판된 현대 원주민 집단들의 유전체 연구를 통해 최소 15,000년 전에 시베리아로부터 베링 육교(Bering Land Bridge)를 건너 아메리카로 넘어온 ‘첫 아메리카인(First Americans)’ 집단으로부터 대부분의 아메리카 원주민 집단들이 유래했다는 가설이 수립되었다[28]. 연구자들은 이외에도 동아시아로부터 아메리카로 두 번의 이주가 더 있었던 것으로 추정했으나 이 두 고대 집단의 유전적 흔적이 남아 있는 집단은 분석 대상들 중 에스키모알류트어족(Eskimo-Aleut languages) 언어를 사용하는 집단들과 나데네어(Na-Dene)를 사용하는 치퍼와이언족(Chipewyan) 집단에 국한되었다. 이후 시베리아 말타 지역의 고유전체 연구에서 ‘첫 아메리카인’ 집단의 조상 집단이 가질 만한 유전적 특징을 제시하였고[24], 미국 몬태나 주에서 발굴된 약 12,600년 전의 것으로 추정되는 고유전체 증거도 이 가설을 지지하였다[29].

그러나 이 지역에 현생 인류가 최초로 정착한 과정에 대해 첨예한 논쟁이 지속되었다. ‘첫 아메리카인’ 가설과는 달리 두 아메리카 대륙에서 발굴된 고인골에 대한 이전의 생물인류학 연구는 형태학적으로 명확히 구별되는 두 고대 집단으로부터 아메리카 원주민 집단들이 기원했다고 추측했다[30]. 이는 현대 아메리카 원주민들의 머리뼈와 북아메리카에서 발굴된 머리뼈들이 대부분 동아시아인의 것에 가

까운 반면, 남아메리카에서 발굴된 많은 머리뼈들은 오세아니아 지역 원주민의 것과 더 유사하다는 사실로부터 추론한 것이었다. 즉, 현대 동아시아인에 가까운 고대 집단과 현대 오세아니아인에 가까운 고대 집단이 모두 조상 집단에 해당한다는 것이 가설이다. 2015년에 출판된 두 연구는 각자 독자적인 분석을 통해 이 가설을 검증하였으나 서로 상반된 결론을 도출하였다[31,32]. 한쪽은 조상 집단이 두 집단이라는 가설을 지지했다[31]. 그러나 다른 한쪽은 이 가설뿐만 아니라 상술했던 2012년 연구에서 밝힌 동아시아로부터 세 번의 이주가 발생했다는 가설[28]조차 기각하고 단 한 번의 이주로부터 모든 집단이 파생됐다고 추정했다[32]. 최근 연구에서는 미국 알래스카 주에서 발굴된 약 11,500년 전의 것으로 추정되는 고인골의 고유전체를 분석하였는데 이러한 단일집단 기원설이 더 지지되었다[33]. 이후 같은 연구팀의 후속연구에서는 현대 오세아니아인과 형태학적으로 가깝다고 보고된 고인골[30]의 고유전체를 분석 대상에 포함시켰다. 분석 결과, 이 표본들 중 일부로부터 실제 오세아니아인의 유전적 신호를 감지할 수 있었으나 머리뼈의 형태학적 특징과의 상관성은 기각되었다[34]. 이러한 일련의 연구들을 통해 서유라시아 연구들이 주로 시간대 및 지역별로 세부적인 이주 패턴을 지속적으로 축적해온 것과 달리 대륙 내 모든 집단들의 기원과 유래라는 하나의 사실에 대해서도 여러 고고유전학 가설들이 첨예하게 대립할 수 있다는 것을 확인할 수 있다. 이는 새로운 고유전체 증거들의 필요성을 잘 보여준다.

다음으로 오세아니아 지역의 경우는 앞서 살펴본 사례들과는 달리 박물관에 보관되어 있던 약 100년 전의 호주 원주민 머리카락 표본으로부터 먼저 유전체 분석이 이루어졌다[35]. 분석 결과, 약 62,000~75,000년 전에 현생 인류가 아시아로 처음 이주하기 시작했고, 이로부터 분기하여 오세아니아 지역으로 이주한 고대 집단이 현대 호주 원주민들의 조상 집단인 것으로 추정되었다. 또한, 이보다 늦은 약 25,000~38,000년 전에 유럽인과 아시아인이 갈라진 것으로 추정됐다. 연구자들은 이 결과들을 종합하여 현생 인류가 동아시아 이주 후 인접한 오세아니아로 연이어 이주한 것이 아니라 독립적으로 두 지역에 이동하여 정착했다는 가설을 지지했다. 이후 2016년에 출판된 연구에서는 현존하는 호주 원주민 집단들과 멜라네시아인 집단들에 대해 광범위하게 유전체를 분석하여 두 집단들의 조상 집단이 하나였던 것으로 추정했다[36]. 이 가상의 오세아니아 고대 집단은 상술했던 연구[35]에서 지지한 오세아니아 정착 가설에 잘 상응한다.

그런데 비스마르크 제도와 솔로몬 제도를 포함하는 ‘근접 오세아니아(near Oceania)’ 지역보다 더 동쪽에 위치한

남태평양 섬들을 지칭하는 ‘원격 오세아니아(remote Oceania)’ 지역에서는 이보다 복잡한 양상이 나타난다. 2010년 이 지역의 현대인 집단들을 유전체 수준에서 분석한 연구에서는 폴리네시아인의 DNA에 동아시아인 DNA가 상당 부분 존재한다는 사실로부터 약 3,000년 전 멜라네시아인과 동아시아인 간의 집단 혼합이 선행된 후 이 혼합 집단이 폴리네시아로 이주하기 시작한 것으로 추론했다[37]. 그러나 2016년 폴리네시아의 바누아투(Vanuatu)와 통가(Tonga) 섬의 고유전체들을 분석한 연구에서는 멜라네시아인 집단의 DNA를 거의 가지지 않은 고유전체 증거를 제시했다[38]. 이로부터 새롭게 도출된 가설은 동아시아에 분포하던 고대 집단이 ‘근접 오세아니아’에 분포하던 멜라네시아인 조상 집단과의 혼합 없이 먼저 ‘원격 오세아니아’로 최소 3,000년 전에 이동하여 정착했다는 것이다. 나아가 최근 연구에서는 새로운 고유전체 및 현대 폴리네시아인 유전체 증거를 통해 약 2,500년 전부터 멜라네시아인 집단이 꾸준히 이 지역으로 유입되면서 점차 동아시아 집단을 대체한 것으로 추정했다[22].

마지막으로 아프리카 기원론[4] 외에는 분명히 입증된 고대 집단 역할이 많지 않았던 사하라 이남 아프리카 지역도 앞서 살펴본 사례들처럼 현존하는 아프리카인 집단들과 아프리카계 미국인 집단들을 유전적으로 분석한 연구가 선행되었다[39]. 그런데 이 2009년 연구는 유전체 수준의 변이를 분석한 연구임에도 이후의 유전체 연구들과 비교할 때 분석 수준이 충분하지 않은 편이다. 그러나 100개 이상의 다양한 사하라 이남 아프리카 집단들을 광범위하게 분석함으로써 아프리카 대륙의 민족, 문화적 특징, 언어적 특징에 상응하는 14개의 아프리카 조상 집단 군집(ancestral population clusters)을 도출하였다. 이후 2017년 출판된 고유전체 연구는 사하라 이남 아프리카에 분포했던 고대 수렵채집민, 농경민, 목축민 집단들의 상호작용을 재구성함으로써 아프리카 고대 집단들의 이주 및 혼합 과정에 대한 새로운 사실들을 밝히기 시작했다[40]. 최근 2018년 연구에서는 약 15,000년 전의 것으로 추정되는 북아프리카 지역의 고유전체를 분석하였는데 이 집단은 이전 서유라시아 연구에서 밝힌 근동 지역의 고대 집단[21]과 유전적으로 가장 가까웠지만 사하라 이남 아프리카 계통 집단의 DNA도 약 37%나 가지고 있었다[41]. 이 고유전체 증거를 통해 아프리카의 현생 인류 집단이 처음 외부로 진출한 이후 홍적세(Pleistocene) 시기에 서유라시아 고대 집단들과 아프리카 고대 집단들 간 상호작용이 존재했음이 규명되었다.

상술한 최근 연구[41]처럼 아메리카, 오세아니아 지역도 후속 고유전체 연구에 따라 기존에 밝혀진 서유라시아 지

역의 고대 집단 분포 양상과 통합될 가능성을 내포하고 있다. 단 이 두 지역은 동아시아를 거쳐야만 서유라시아 지역과 연결된다. 따라서 동아시아 지역에 대한 규명이 선행되어야만 전 세계 고대 집단 역학이 하나로 이어질 수 있다.

(3) 동아시아 연구

2009년에 현대 동아시아인 집단 73개의 유전체를 비교 분석한 연구가 출판되었다[42]. 분석 결과, 집단의 언어나 민족, 지리적 위치 등의 특성들과 유전적 특징이 잘 상응하였다. 즉, 동아시아인 집단들은 이 특성들을 기준으로 할 때 유전적으로 서로 다른 집단들로 대체로 잘 구분되었다. 집단들의 기원에 대해서는 서유라시아 지역으로부터 여러 번이 아닌 단 한 번의 이주를 통해 모든 동아시아 집단들이 파생되었다는 가설이 수립됐다. 즉, 현존하는 동북아시아와 동남아시아 집단들이 단일 조상 집단으로부터 기원했다는 가설인데 이를 직접적으로 증명하는 고유전체 증거는 밝혀진 바가 없다.

동아시아 지역을 대상으로 하는 고유전체 연구는 앞서 검토한 모든 지역들에 비해 더디게 진행되고 있다. 먼저 동남아시아의 경우 최초의 고유전체 연구 2편이 2018년에 동시에 출판되었다[43,44]. 한 연구는 신석기 후기부터 철기 시대까지 추정 시기가 다양한 25개 동남아시아 고유전체를 분석하였다[43]. 분석 결과, 현존 동남아시아 집단들은 토착 고대 집단인 호아비니언(Hòabinhian) 집단과 나중에 이주한 동북아시아 농경인 고대 집단을 포함한 최소 4개의 고대 집단으로부터 기원한 것으로 추정되었다. 다른 연구에서는 신석기 시대로 추정되는 18개의 동남아시아 고유전체를 분석하여 중국 남부의 농경인 고대 집단, 오스트로아시아어족(Austroasiatic languages) 언어를 사용하는 유라시아 수렵채집민 고대 집단 등이 기존에 분포하던 고대 집단에 유입한 것으로 추정했다[44]. 결과적으로 두 연구 모두 동남아시아에 새로운 고대 집단이 여러 번 이주했다는 사실을 밝혔다.

이와 비교하면 동북아시아의 지역은 현존하는 중국, 일본 집단들을 대상으로 하는 유전체 연구가 여러 차례 수행되었으나[45-48] 고유전체 연구는 매우 부족한 실정이다. 먼저 중국의 경우는 2009년에 출판된 유전체 연구에서 현대 한족 집단들이 유전적 특징에 따라 남부, 중부, 북부의 3개 군집으로 구분 가능하다고 밝혔다[45]. 이는 상술한 2009년 동아시아인 유전체 연구에서 한족 집단의 소집단들을 분석한 결과와 유사했다[42]. 최근 출판된 대규모 연구에서는 약 12,000명에 이르는 다양한 지역의 한족 집단 유전체를 분석하여 중국과 인접한 여러 지역 고대 집단들

과의 혼합을 검증하였다[46]. 분석 결과, 중국 북서부 지역에서 유럽 집단과의 집단 혼합이 발생했을 가능성이 제시되었다. 그러나 이를 직접적으로 증명할 만한 중국 고대 집단 고DNA 연구는 중국 티안유안 동굴(Tianyuan Cave)에서 발굴된 약 40,000년 전의 것으로 추정되는 표본을 고인골을 분석한 2013년 연구가 유일하다[11]. 이마저도 표본이 오래되어 mtDNA와 21번 염색체 등 유전체의 일부만을 분석할 수 있었기에 고유전체 증거로써는 유효성이 불충분하다.

일본의 경우도 상황은 유사하다. 현존하는 일본인 집단들은 크게 아이누인(Ainu), 류큐인(Ryukuans), 본토 일본인(Mainland Japanese) 세 집단으로 구분되는데 2012년 연구에서는 이 세 집단을 유전체 수준으로 분석하여 고대 집단들을 추론하였다[47]. 연구자들은 세 집단이 두 고대 집단으로부터 파생됐다고 추정했고, 이들이 고대 일본의 조몬인과 야요이인일 가능성을 제시했다. 2015년 후속연구에서는 아이누인을 조몬인의 자손 집단으로, 한국인과 중국 한족 집단을 야요이인의 자손 집단으로 가정하여 고대 집단들에 대한 추가 분석을 수행하기도 했다[48]. 이 연구들을 뒷받침할 만한 약 3,000년 전의 것으로 추정되는 조몬인 고DNA 연구가 2017년에 출판되었으나 상술한 중국 고DNA 연구[11]처럼 이 연구도 유전체의 일부만을 분석했다는 한계를 지니고 있다[49].

하지만 최근에 동북아시아 집단들의 기원을 밝히는 데 도움이 될 만한 고유전체 연구들이 2편 출판되었다[43,50]. 먼저 2017년에 출판된 연구에서는 러시아 연해주 지역의 동굴 ‘악마의 문(Devil’s Gate)’에서 발굴된 고인골의 고유전체를 분석하였다[50]. 분석 결과, 이 고대 집단은 동굴이 위치한 아무르강 유역(Amur Basin)에 분포하는 현대 울치족(Ulchi) 집단과 가장 유전적으로 가까웠으며, 현존하는 한국인, 일본인 집단들과도 가까웠다. 또한, 앞서 살펴본 동남아시아 연구에서는 25개의 동남아시아 고유전체에 고대 일본 조몬인 고유전체 1개를 포함시켜 비교분석하였다[43]. 이 연구에 따르면 조몬인 집단은 동남아시아 고대 집단과 동북아시아 고대 집단의 혼합에 가장 가까운 것으로 나타났으며, 현대 일본인 집단은 조몬인과 다른 동아시아 집단의 혼합으로 추정되었다.

연해주와 일본은 한반도를 사이에 두고 있으므로 위 연구들[43,50]을 한반도 고대 집단들에 대한 연구와 종합한다면 동북아시아 지역의 집단 역학을 상당 부분 재구성할 수 있을 가능성이 크다. 그러나 현재까지 현대 한국인 집단들의 유전체를 비교분석하거나 한반도 고대 집단 고유전체를 분석한 연구는 수행된 적이 없다.

고찰

본 연구에서는 고고유전학 연구의 기본적인 분석 원리와 최근 고유전체 연구 동향을 검토하였다. 검토 결과, 호미닌 종들의 유전적 특징 및 현생 인류와의 관계, 서유라시아 지역에 분포하던 현생 인류 고대 집단들의 기원 및 상호작용 등을 밝힌 고유전체 연구들을 토대로 타 지역에 고유전체 연구들이 활발하게 수행되고 있음을 확인할 수 있었다. 서유라시아 지역과 달리 현존하는 집단들의 유전체 정보가 조사되어 있지 않던 타 지역들은 현대 인도인, 아메리카인, 오세아니아인, 아프리카인 등의 집단들에 대해 먼저 전장 유전체 분석이 수행되었다. 이를 통해 고대 집단들에 대한 가설을 먼저 수립한 후 고유전체 연구의 검증을 통해 기존 가설을 지지하거나 새로운 가설을 세우는 과정을 거치고 있다. 그러나 동아시아 고유전체 연구는 전반적으로 지지 부진하다.

특히 한반도 고유전체 연구는 아직 시도조차 없는 상황이다. 타 동아시아 집단들과 달리 현대 한국인 집단을 분석 대상에 포함시킨 유전체 연구들도 부족하다. 앞서 살펴본 현대 동아시아인 집단들의 유전체 분석 연구[42]와 최근 수행된 전 세계 현대인 집단들의 유전체 분석 연구[7]에서는 현존하는 한국인 집단과 일본인 집단의 유전체가 유사하다는 사실을 밝히고 있다. 그러나 이 연구들 외에 세계 전 지역의 현존하는 집단들을 폭넓게 다룬 타 유전체 연구들에서는 한국인 유전체가 분석에서 제외되어 있다. 또한, 먼저 살펴보았듯이 중국, 일본 지역에 대해서는 국내 집단들의 유전체만을 비교분석하여 그 기원이 되는 고대 집단들의 특징을 추정하는 연구들이 수행되었다[45-48]. 그러나 현대 한국인 집단들만을 분석한 유전체 연구는 출판된 적이 없다.

현대 한국인들의 유전자료를 통해 한국인의 기원을 추정하는 연구들은 현재까지 mtDNA와 Y염색체를 대상으로만 수행되었다[51-55]. 한국인을 포함한 현대 동아시아인 집단들의 Y염색체를 비교분석한 초기 연구들에서는 한국인이 한반도의 남북 양쪽 방향으로부터 모두 기원했다는 ‘이중기원설’을 제시했다[51,52]. 이후 mtDNA와 Y염색체를 모두 분석한 연구에서는 한국인의 mtDNA가 동북아시아 집단들의 것과 유사하지만 Y염색체는 동북·동남아시아 양쪽의 것들을 모두 포함한다는 분석 결과를 통해 과거에 한반도로 성 편향적인 이주(sex-biased migration)가 있었다고 해석하기도 했다[53]. 그러나 이후 분석 표지의 수를 늘린 Y염색체 연구에서는 한국인이 동북아 지역에서만 기원했다는 상반된 추정 결과가 나왔다[54]. 한편, 한국 내 지역 집단들에 대해서도 mtDNA만을 비교분석하는 연구

가 출판되었는데 다른 지역의 집단들은 유전적으로 서로 유사했으나 제주도 집단만 그렇지 않은 결과가 도출되었다[55]. 그러나 전장유전체 염기서열분석을 할 경우 이러한 모든 결과들과 전혀 다른 결과가 나올 수 있다.

한반도 고대 집단들에 대해서도 고유전체가 아닌 일부 DNA를 대상으로 한 고DNA 연구들은 다수 수행되었다[56-70]. 그런데 국내 대부분의 고DNA 연구들은 법의학적인 분석과 유사한 방법으로 단연쇄반복(short tandem repeats) 부위, 성 결정에 관여하는 유전자 부위, mtDNA 등을 분석하였다. 이러한 DNA 부위들의 분석을 통해 고인골 개체들의 성과 혈연관계를 추정하고 주피장자·순장자와 같은 고고학적 맥락 및 뼈에 대한 형태학적 분석과 비교하는 것이 연구의 목적이었다. 나주 복암리 옹관묘[56], 사천 늑도 유적 및 민통선 민묘[57], 아산 명암리 유적[58], 시흥 목감동[59], 완주 은하리 고분[60], 나주 대안리 고분 및 서천 옥남리 유적[61], 경산 임당 유적[62,63] 등에서 출토된 고인골들을 대상으로 이러한 연구들이 수행되었다. 이후 연구들에서는 이러한 분석 외에 고대 집단의 하플로그룹을 현존하는 동아시아 집단들의 하플로그룹과 비교하는 것이 추가 분석에 포함되었다. 아산 명암리 유적[64], 서천 옥남리 유적[65], 김포 장기동 유적[66], 경주 교동 유적[67] 등의 고DNA 연구에서 이러한 추가분석을 했다.

최근 고고유전학 연구 추세처럼 현대 한국인 집단의 기원을 밝히고자 했던 국내 고DNA 연구들도 있었다[68-70]. 먼저 2004년 경산 임당동과 사천 늑도 유적에서 출토된 고인골의 고DNA를 분석한 연구에서는 현대 동아시아인 및 코카시안(Caucasian) 집단들과 mtDNA를 비교분석하는 집단유전학적 분석을 국내 연구로서 처음 시도하였다[68]. 분석 결과, 늑도 유적의 시료들이 유전적으로 동북아시아와 가까운 것과 동남아시아에 가까운 것으로 나뉘었다. 국내 11개 유적에서 발굴된 고인골 표본들의 mtDNA를 비교분석한 2008년 연구에서도 유사하게 초기 한반도 집단이 북쪽에서 기원한 뒤 남쪽으로부터 다른 고대 집단들의 유입이 발생한 것으로 추정하였다[69]. 비교적 최근인 2015년 출판된 김포 양촌읍 유적의 고DNA 연구에서는 현대 동아시아인 집단들과 mtDNA를 비교분석한 결과, 시료가 동북아시아 지역과 전반적으로 유사하지만 동남아시아 지역과도 유전적인 관련이 크다는 결과가 나왔다[70]. 연구자들은 이를 통해 이중기원설[51,52]을 지지하는 한편, Y염색체와 mtDNA의 하플로그룹을 비교분석함으로써 모계와 부계에 따라 차이가 있는지도 확인했다. Y염색체의 하플로그룹 개수가 mtDNA의 것보다 적다는 분석 결과가 조선시대 부거(patrilocal) 혼인제를 반영할 가능성이 제시되었다. 결론적으로 이러한 일련의 연구들이 모두 한반도 집단이 남북 양

쪽에서 기원했을 가능성을 암시하고 있다.

그러나 이러한 연구들은 모두 분석 수준이 mtDNA, Y염색체에 한정되어 있다는 점에서 현재까지 축적된 고유전체 연구들과 종합하여 동북아시아와 한반도의 고대 집단 역사를 수립하기에는 유효하지 않다. 한반도 고대 집단들의 기원과 이동을 밝히기 위해서는 최근의 고고유전학 연구 추세처럼 유전체 수준의 고DNA 분석이 필요하다. 지리적 위치상 과거부터 지금까지 여러 국가들이 상호작용했던 곳이 한반도임을 감안할 때, 한반도의 고대 집단이 명확히 밝혀질 경우 동북아시아 고대 집단들의 유래를 밝히는 데 크게 기여할 수 있다. 이를 통해 고유전체 수준에서 동아시아 전 지역 고대 집단들의 기원과 이주에 대한 조감이 가능해진다면 본 연구에서 검토한 모든 타 지역의 연구들과 통합될 여지 또한 충분하다. 동아시아는 지리적으로 서유라시아, 오세아니아 지역과 인접해 있으며, 과거에는 아메리카 대륙과도 연결되어 있었기 때문이다. 즉, 동아시아 고대 집단들의 특징이 밝혀질 때 비로소 현생 인류 고대 집단 역학의 전체 그림이 짜여질 수 있다.

한편, 한반도의 고유전체 연구를 통해 새롭게 한반도 고대 집단들의 특징을 정립함으로써 기존의 고고학적 증거들과 비교하는 연구도 가능하다. 아울러 한국, 중국, 일본 등의 동북아시아 국가들이 많은 언어적 특징들을 공유하고 있듯이 교차검증할 만한 사료나 언어학적 자료들도 아주 다양할 것으로 예상된다. 또한, 앞서 살펴 보았듯이 고고유전학의 분석은 중립진화 이론을 전제로 하고 있지만 최근에는 고유전체 연구를 통해 특정 집단에서 자연선택된 유전자들을 검증하는 연구들도 점차 시도되고 있다[40]. 고대부터 자연선택되어 현대의 한국인들이 가지고 있는 유전자들을 탐색하는 연구들도 가능할 수 있다.

결론적으로 한반도 고유전체 연구는 매우 큰 가능성을 지니고 있다. 향후 국내 고유전체 연구를 통해 한반도의 인류사를 대상으로 하는 생물인류학·고고학 연구의 큰 진전이 있기를 기대한다.

REFERENCES

1. Pääbo S. Molecular cloning of ancient Egyptian mummy DNA. *Nature*. 1985; 314:644-5.
2. Höss M, Pääbo S. DNA extraction from Pleistocene bones by a silica-based purification method. *Nucleic Acids Res*. 1993; 21:3913-4.
3. Jobling M, Hollox E, Hurler M, Kivisild T, Tyler-Smith C. Finding and assaying genome diversity. In: Jobling M, Hollox E, Hurler M, Kivisild T, Tyler-Smith C. *Human evolutionary genetics*. New York: Garland Science; 2014. p. 95-6, 98-101, 123-8.
4. Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. *Nature*. 1987; 325:31-6.
5. Reich D. *Humanity's ghosts*. In: Reich D. *Who we are and how we got here*. New York: Pantheon Books; 2018. p. 77-86, 94.
6. Hernandez RD, Kelley JL, Elyashiv E, Melton SC, Auton A, McVean G, et al. Classic selective sweeps were rare in recent human evolution. *Science*. 2011; 331:920-4.
7. Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, et al. The Simons genome diversity project: 300 genomes from 142 diverse populations. *Nature*. 2016; 538:201-6.
8. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, et al. A draft sequence of the Neandertal genome. *Science*. 2010; 328:710-22.
9. Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, et al. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*. 2014; 505:43-9.
10. Pinhasi R, Fernandes D, Sirak K, Novak M, Connell S, Alpaslan-Roodenberg S, et al. Optimal ancient DNA yields from the inner ear part of the human petrous bone. *PLoS One*. 2015; 10:e0129102. DOI:10.1371/journal.pone.0129102.
11. Fu Q, Meyer M, Gao X, Stenzel U, Burbano HA, Kelso J, et al. DNA analysis of an early modern human from Tianyuan Cave, China. *Proc Natl Acad Sci U S A*. 2013; 110:2223-7.
12. Skoglund P, Northoff BH, Shunkov MV, Derevianko AP, Pääbo S, Krause J, et al. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. *Proc Natl Acad Sci U S A*. 2014; 111:2229-34.
13. Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, et al. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*. 2010; 468:1053-60.
14. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S, et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science*. 2012; 338:222-6.
15. Fu Q, Li H, Moorjani P, Jay F, Slepchenko SM, Bondarev AA, et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature*. 2014; 514:445-9.
16. Moorjani P, Sankararaman S, Fu Q, Przeworski M, Patterson N, Reich D. A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. *Proc Natl Acad Sci U S A*. 2016; 113:5652-7.
17. Sankararaman S, Mallick S, Patterson N, Reich D. The combined landscape of Denisovan and Neandertal ancestry in present-day humans. *Curr Biol*. 2016; 26:1241-7.
18. Meyer M, Arsuaga JL, de Filippo C, Nagel S, Aximu-Petri A, Nickel B, et al. Nuclear DNA sequences from the

- Middle Pleistocene Sima de los Huesos hominins. *Nature*. 2016; 531:504-7.
19. Slon V, Mafessoni F, Vernot B, de Filippo C, Grote S, Viola B, et al. The genome of the offspring of a Neanderthal mother and a Denisovan father. *Nature*. 2018; 561:113-6.
 20. Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*. 2015; 522:207-11.
 21. Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, et al. Genomic insights into the origin of farming in the ancient Near East. *Nature*. 2016; 536:419-24.
 22. Posth C, Nägele K, Collieran H, Valentin F, Bedford S, Kami KW, et al. Language continuity despite population replacement in Remote Oceania. *Nat Ecol Evol*. 2018; 2:731-40.
 23. Patterson NJ, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, et al. Ancient admixture in human history. *Genetics*. 2012; 192:1065-93.
 24. Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, Moltke I, et al. Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature*. 2014; 505:87-91.
 25. Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. *Nature*. 2009; 461:489-94.
 26. Moorjani P, Thangaraj K, Patterson N, Lipson M, Loh PR, Govindaraj P, et al. Genetic evidence for recent population mixture in India. *Am J Hum Genet*. 2013; 93:422-38.
 27. Narasimhan VM, Patterson NJ, Moorjani P, Lazaridis I, Mark L, Mallick S, et al. The Genomic Formation of South and Central Asia. *bioRxiv*. 2018. DOI:10.1101/292581.
 28. Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, et al. Reconstructing native American population history. *Nature*. 2012; 488:370-4.
 29. Rasmussen M, Anzick SL, Waters MR, Skoglund P, DeGiorgio M, Stafford Jr TW, et al. The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature*. 2014; 506:225-9.
 30. Neves WA, Hubbe M. Cranial morphology of early Americans from Lagoa Santa, Brazil: Implications for the settlement of the New World. *Proc Natl Acad Sci U S A*. 2005; 102:18309-14.
 31. Skoglund P, Mallick S, Bortolini MC, Chennagiri N, Hünermeier T, Petzl-Erler ML, et al. Genetic evidence for two founding populations of the Americas. *Nature*. 2015; 525:104-8.
 32. Raghavan M, Steinrücken M, Harris K, Schiffels S, Rasmussen S, DeGiorgio M, et al. Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science*. 2015. DOI:10.1126/science.aab3884.
 33. Moreno-Mayar JV, Potter BA, Vinner L, Steinrücken M, Rasmussen S, Terhorst J, et al. Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature*. 2018; 553:203-7.
 34. Moreno-Mayar JV, Vinner L, de Barros Damgaard P, de la Fuente C, Chan J, Spence JP, et al. Early human dispersals within the Americas. *Science*. 2018. DOI:10.1126/science.aav2621.
 35. Rasmussen M, Guo X, Wang Y, Lohmueller KE, Rasmussen S, Albrechtsen A, et al. An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science*. 2011; 334:94-8.
 36. Malaspina AS, Westaway MC, Muller C, Sousa VC, Lao O, Alves I, et al. A genomic history of Aboriginal Australia. *Nature*. 2016; 538:207-14.
 37. Wollstein A, Lao O, Becker C, Brauer S, Trent RJ, Nürnberg P, et al. Demographic history of Oceania inferred from genome-wide data. *Curr Biol*. 2010; 20:1983-92.
 38. Skoglund P, Posth C, Sirak K, Spriggs M, Valentin F, Bedford S, et al. Genomic insights into the peopling of the Southwest Pacific. *Nature*. 2016; 538:510-3.
 39. Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, et al. The genetic structure and history of Africans and African Americans. *Science*. 2009; 324:1035-44.
 40. Skoglund P, Thompson JC, Prendergast ME, Mittnik A, Sirak K, Hajdinjak M, et al. Reconstructing prehistoric African population structure. *Cell*. 2017; 171:59-71.
 41. van de Loosdrecht M, Bouzouggar A, Humphrey L, Posth C, Barton N, Aximu-Petri A, et al. Pleistocene North African genomes link Near Eastern and sub-saharan African human populations. *Science*. 2018; 360:548-52.
 42. HUGO Pan-Asian SNP Consortium. Mapping human genetic diversity in Asia. *Science*. 2009; 326:1541-5.
 43. McColl H, Racimo F, Vinner L, Demeter F, Gakuhari T, Moreno-Mayar JV, et al. The prehistoric peopling of Southeast Asia. *Science*. 2018; 361:88-92.
 44. Lipson M, Cheronet O, Mallick S, Rohland N, Oxenham M, Pietrusewsky M, et al. Ancient genomes document multiple waves of migration in Southeast Asian prehistory. *Science*. 2018; 361:92-5.
 45. Xu S, Yin X, Li S, Jin W, Lou H, Yang L, et al. Genomic dissection of population substructure of Han Chinese and its implication in association studies. *Am J Hum Genet*. 2009; 85:762-74.
 46. Chiang CW, Mangul S, Robles C, Sankararaman S. A comprehensive map of genetic variation in the world's largest ethnic group—Han Chinese. *Mol Biol Evol*. 2018; 35:2736-50.

47. Jinam T, Nishida N, Hirai M, Kawamura S, Oota H, Umet-su K, et al. The history of human populations in the Japa-nese Archipelago inferred from genome-wide SNP data with a special reference to the Ainu and the Ryukyuan popu-lations. *J Hum Genet.* 2012; 57:787-95.
48. Jinam TA, Kanzawa-Kiriyama H, Inoue I, Tokunaga K, Omoto K, Saitou N. Unique characteristics of the Ainu popu-lation in Northern Japan. *J Hum Genet.* 2015; 60:565-71.
49. Kanzawa-Kiriyama H, Kryukov K, Jinam TA, Hosomichi K, Saso A, Suwa G, et al. A partial nuclear genome of the Jomons who lived 3000 years ago in Fukushima, Japan. *J Hum Genet.* 2017; 62:213-21.
50. Siska V, Jones ER, Jeon S, Bhak Y, Kim HM, Cho YS, et al. Genome-wide data from two early Neolithic East Asian indi-viduals dating to 7700 years ago. *Sci Adv.* 2017; 3:e1601877. DOI:10.1126/sciadv.1601877.
51. Kim W, Shin DJ, Harihara S, Kim YJ. Y chromosomal DNA variation in east Asian populations and its potential for in-ferring the peopling of Korea. *J Hum Genet.* 2000; 45:76-83.
52. Jin HJ, Kwak KD, Hammer MF, Nakahori Y, Shinka T, Lee JW, et al. Y-chromosomal DNA haplogroups and their im-plications for the dual origins of the Koreans. *Hum Genet.* 2003; 114:27-35.
53. Jin HJ, Tyler-Smith C, Kim W. The peopling of Korea re-vealed by analyses of mitochondrial DNA and Y-chromo-somal markers. *PloS One.* 2009; 4:e4210. DOI:10.1371/jour-nal.pone.0004210.
54. Kim SH, Kim KC, Shin DJ, Jin HJ, Kwak KD, Han MS, et al. High frequencies of Y-chromosome haplogroup O2b-SRY465 lineages in Korea: a genetic perspective on the peopling of Korea. *Invest Genet.* 2011; 2:10. DOI:10.1186/2041-2223-2-10.
55. Hong SB, Kim KC, Kim W. Population and forensic genet-ic analyses of mitochondrial DNA control region variation from six major provinces in the Korean population. *Foren-sic Sci Int Genet.* 2015; 17:99-103.
56. Lee KS, Chung YJ, Han SH, Lee MH, Han MS, Choi DH. The genetic analysis on ancient human bone in jar coffin excavated from Naju Bokam-ni 3rd tumulus. *Conserv Stud.* 1999; 20:5-19. Korean.
57. Seo MS, Lee KS, Chung YJ, Lee MH. Personal identifica-tion of the excavated ancient human bone through molecu-lar-biological methods. *Conserv Stud.* 2001; 22:27-40. Ko-rean.
58. Seo MS, Lee KS, Chung YJ, Kim KK, Pak YJ. A genetic analysis of human remains from the Myeongam-ri site, Asan City. *Conserv Stud.* 2002; 23:59-75. Korean.
59. Seo MS, Chung YJ, Lee KS, Park KW. Genetic character-istics of mtDNA and STR marker in human bone excavated from Mokgam-dong, Siheung in Korea. *Conserv Stud.* 2003; 24:153-67. Korean.
60. Lee J, Kim J, Choi C, Han Y, Lee S, Cho Y, et al. An inves-tigation of Baekje mortuary practices using DNA analysis. *J Korean Archaeol Soc.* 2006; 61:70-91. Korean.
61. Jee S, Park J, Seo M, Hong J, Chung Y. Molecular genetic and histological analysis for the excavated human bone. *Conserv Stud.* 2007; 28:75-90. Korean.
62. Lee J, Ha D, Pak S, Woo EJ, Lee C, Kim D, et al. The rela-tions of the dead: identifying the relationship of individuals buried at Imdang, Gyeongsan, through the analysis of mito-chondrial DNA from human skeletal remains interred in large mounded tombs. *J Korean Archaeol Soc.* 2008; 68:128-55. Korean.
63. Ha D. A study on the status of the victims of the Silla sacri-ficial burials at Imdang. *J Korean Archaeol Soc.* 2011; 79:175-204. Korean.
64. Jee SH, Kim YJ, Chung YJ, Seo MS, Pak Y. A paleogenetic analysis of human skeletal remains from the Myeongam-ri site, Asan in Korea. *J Conserv Sci.* 2008; 23:81-93. Korean.
65. Kang S, Kwon E, Moon E, Cho E, Seo M, Kim Y, et al. Usefulness of biochemical analysis for human skeletal re-mains assigned to the Joseon Dynasty in Oknam-ri site in Seocheon, Korea. *J Conserv Sci.* 2010; 26:95-107. Korean.
66. Seo MS, Cho EM, Kim YJ, Kim SH, Kang SY. The genetic analysis study of ancient human bones excavated at Jang-gi-dong site, Gimpo. *J Conserv Sci.* 2014; 30:409-16. Kore-an.
67. Lee WJ, Woo EJ, Oh CS, Yoo JA, Kim YS, Hong JH, et al. Bio-anthropological studies on human skeletons from the 6th century tomb of ancient Silla Kingdom in South Ko-rea. *PloS One.* 2016; 11:e0156632. DOI:10.1371/journal.pone.0156632.
68. Seo MS, Lee KS. Genetic analysis of ancient human bones excavated in Sacheon Nuk-do and Gyeongsan Yim-dang-dong, Korea. *Conserv Stud.* 2004; 25:47-74. Korean.
69. Lee HY, Yoo JE, Park MJ, Kim CY, Shin KJ. Genetic characterization and assessment of authenticity of ancient Korean skeletal remains. *Hum Biol.* 2008; 80:239-50.
70. Jin HJ, Kim KC, Kim W. Phylogenetic analysis of two hap-loid markers of 500-years-old human remains found in a central region of Korea. *Gene Genom.* 2015; 37:33-43.

Principles of Archaeogenetics and the Current Trends of Ancient Genome Studies

Taeho Kim¹, Eun Jin Woo², Sunyoung Pak¹

¹*Biological Anthropology Laboratory, Department of Anthropology, College of Social Sciences, Seoul National University*

²*Department of History, College of Liberal Arts, Sejong University*

Abstract : Archaeogenetics is an academic discipline that aims to establish scientific facts of human history by integrating ancient DNA analyses with archaeological and anthropological evidence. After ancient DNA research was initiated about 30 years ago, it has been innovated so rapidly that the range of analysis has been extended toward the whole genome sequence of ancient genomes in recent 10 years. By this development, researchers have been able to study in detail the origins and migration patterns of hominin species and ancient human populations by approaches of evolutionary genetics. This study has reviewed main principles of the archaeogenetic analysis and the current trends of ancient genome studies with recent achievements. While sampling techniques and statistical analyses have been improved, typical research methods have been established by the findings on hominins and ancient western Eurasia populations. Recently, archaeogenecists have been applying the methods to studying those in other geographical areas. Nonetheless, there is still the lack of ancient genome research about populations in Eastern Asia including the Korean peninsula. This review ultimately aims to predict possibilities and promise of future ancient genome studies of ancient Korean populations.

Keywords : Archaeogenetics, Ancient DNA, Ancient Genome, Hominins, Ancient populations

