

인간 유전체 ENCODE 프로젝트와 정크 DNA의 기능성 논쟁에 대한 기원과학적 고찰

김아람*
현창기**

논문초록

인간 유전체의 DNA 서열 중 98%가 비암호화 DNA(non-coding DNA), 즉 단백질 정보가 없는 비유전자 영역으로서, 진화론적으로는 오랜 진화의 과정에서 기능을 상실한 '정크 DNA(junk DNA)'로 해석되어왔다. 그러나 인간 유전체 기능 연구를 위한 ENCODE 프로젝트가 2012년에 인간 유전체의 80%가 고유의 생화학적 기능을 가지고 있다는 연구결과를 발표하면서 정크 DNA는 잘못된 개념임을 밝혔다. 본 논문에서는 정크 DNA 개념과 ENCODE 연구결과와의 갈등에 대해 고찰하고, 이를 통해 인간 유전체에 대한 기원과학적 관점을 비교 평가하였다. 인간 유전체를 무신론적 자연발생의 결과로 보는 기존의 진화론적 관점을 비판적으로 분석하고, 특정한 목적성과 복잡성을 가지는 설계의 산물로 보는 유신론적 관점의 해석을 시도함과 동시에 그 타당성을 논증하였다.

주제어 : 인간 유전체, ENCODE, 정크 DNA, 진화론, 지적설계론, 기원과학

* 한동대학교 생명과학부 교수

** 한동대학교 생명과학부 교수, 교신저자

2016년 6월 13일 접수, 7월 18일 최종수정, 7월 31일 게재확정

1. 서론

1990년에 시작된 인간 유전체 프로젝트가 인간 유전체의 염기서열 지도를 완성한 후, 학자들은 그 염기서열이 갖고 있는 유전자 정보를 분석한 결과 인간 유전체 염기서열 중 단백질 정보를 가지고 있는 부분은 단 2% 정도에 불과하며 98%의 영역에는 단백질 정보가 없다는 사실을 밝혀내었다(Elgar & Vavouri, 2008). 유전체의 DNA 서열 중에서 이렇게 무의미해 보이는 부분들을 가리켜 진화생물학자들은 1970년대부터 생명유지를 위한 특정한 기능이 없다는 의미로 ‘정크 DNA(junk DNA)’라 불려왔다(Ohno, 1972). 한편 인간 유전체가 가지는 기능들에 대한 연구도 활발하게 진행되었는데, 그 중 대표적인 연구는 2003년부터 국제 컨소시엄 형태로 진행된 초대형 연구 프로젝트로서 ENCODE(ENCyclopedia of DNA Elements) 프로젝트가 시작되었다. ENCODE 프로젝트는 이름 그대로 DNA 요소들로 이루어지는 백과사전을 만들어내고자 수행된 연구로서 단백질 정보를 가지지 않는 DNA 염기서열들이 과연 기능이 없는 정크 DNA가 맞는지 기능이 있다면 어떠한 기능을 담당하고 있는지를 탐구하고자 하였다(The ENCODE Project Consortium, 2012).

2012년에 발표된 ENCODE 프로젝트의 결론은 인간 유전체의 어떤 부분이 발현, 즉 전사되며 그 전사과정은 어떻게 조절되는지, 그리고 그 과정은 세포핵 내의 염색체 DNA의 구조에 의해 어떻게 영향을 받고 있는지에 대해 매우 흥미로운 사실을 보여주었다. 즉, 인간 유전체가 보유하고 있는 2만개 정도의 유전자들이 차지하는 유전체 상의 부분은 2%에 불과한데 유전체의 기능은 그 유전자들에만 국한된 것이 아니라 각각의 유전자의 활성을 조절하는 광범위한 영역에 부여되어 있으며, 실제로 인간 유전체의 최대 80%가 고유의 기능을 가지고 있다는 사실이었다(The ENCODE Project Consortium, 2012). 인간 유전체에는 400만개에 달하는 ‘유전자 스위치’가 존재하고 있으며 지금까지는 기능이 없다고 평가되었던 DNA 부분들의 98%가 유전자의 발현을 조절함에 있어 다양한 기능을 수행할 수 있다는 사실이 확인된 것이다. 또한 질병과 관련된 대부분의 변화는 유전자 자체가 아닌 스위치 부분에 있다는 새로운 사실을 확인함으로써 질병의 예방과 치료를 위한 매우 중요한 토대를 마련하기도 하였다.

인간 유전체 DNA의 최대 80%가 고유의 기능을 가지고 있다는 ENCODE 프로젝트의 결론은 인간 유전체의 대부분을 정크 DNA라고 믿어왔던 기존의 진화론적 해석에

커다란 숙제를 안겨주는 결과이기도 하였다. 최근 급속하게 발전해 온 최첨단의 DNA 분석기술과 후생유전학적 지식으로 가능했던 ENCODE 프로젝트의 결과는 이른바 ‘인간 DNA 백과사전’, 즉 인간의 유전체가 오랜 세월 동안 진화를 거쳐 형성된 것이 아니라 이미 각 부분에 고유의 기능이 부여되어 설계된 결과물임을 강력하게 시사해주고 있는 것이다. 본 고찰에서는 인간 유전체의 기원 문제에 주목하여, 자연발생적 진화과정에 의해 인간 유전체가 만들어졌다고 보는 무신론적 가설의 논리적 문제점을 분석함과 동시에, 인간 유전체를 지적설계의 산물로 보는 유신론적 관점을 제안하고 그 타당성을 논증하고자 한다.

II. 본론

1. ENCODE 프로젝트

인간 유전체를 구성하는 약 30억쌍의 DNA 염기서열이 인간 유전체 프로젝트(Human Genome Project)의 결과로 발표된 후, 그 후속 연구로서 인간 유전체가 갖는 기능적 요소들을 찾아내고자 2003년 9월부터 ENCODE(ENCyclopedia of DNA Elements) 프로젝트가 시작되었다. 미국 국립인간유전체연구소(National Human Genome Research Institute, NHGRI)가 주도하는 이 ENCODE 프로젝트에는 미국, 영국, 일본, 스페인, 싱가포르 등 여러 나라의 32개 연구기관, 442명의 연구자들이 참여하여 수행되었으며 그 결과들은 ‘네이처(Nature)’, ‘사이언스(Science)’, ‘유전체생물학(Genome Biology)’, ‘유전체 연구(Genome Research)’를 비롯한 6개 주요 과학저널에 30편의 논문으로 발표되었다 (Maher, 2012).

ENCODE 프로젝트에서 주목한 유전체 내의 기능요소로는 유전자와 프로모터(promoter), 유전자 발현을 촉진 또는 억제하는 조절인자(전사인자, transcription factors), 유전자 전사(transcription)를 조절하는 부위(인핸서, enhancers), 메틸화(methylation) 부위 및 다양한 생물종 사이에서 보존되는 부위 등이었다. ENCODE 프로젝트의 결과는 기존의 인간 유전체에 대한 상식을 크게 벗어나는 새롭고 놀라운 사실을 보여주었다. 특정한 단백질 정보를 가지고 있지 않은 서열로서 지금까지 ‘정크(junk) DNA’로

해석되어왔던 인간 유전체의 상당 부분이 실제로는 인체에서 결정적인 작용을 하는 기능을 보유하고 있다는 발견이었다.

지금까지 생물학계에서는 인간 유전체 DNA에서 유전자, 즉 단백질 정보를 담고 있는 단백질 암호화 영역(protein-coding region)은 2%에 불과하며 그 나머지 중 상당부분이 단백질 정보와 무관한 ‘쓸모없는(junk)’ 서열로 이해하고 그 이유로 인간의 진화 과정에서 기능을 잃어버린 결과로 해석해왔다. 그러나 ENCODE 프로젝트의 결과에 의하면 전체 인간 유전체의 80%가 생화학적 기능을 가지고 있는 중요한 서열들이 밝혀짐과 동시에, 이 기능적 부분들은 상호간에 또는 유전자 영역과 물리적으로 연관되어 있어서 유전자 발현 조절 메커니즘은 물론 유전체의 조직화에 대한 새로운 해석 방법이 필요함을 시사해주었다(The ENCODE Project Consortium, 2012). 다시 말해, 기존의 유전체학적 지식으로는 유전체의 기본적 구성단위가 당연히 단백질을 암호화하는 유전자라고 해야 하겠지만 ENCODE 프로젝트의 결과는 실제로는 유전체로부터 전사된 다양한 종류의 RNA들을 유전체의 기본 단위로 보는 새로운 관점이 필요함을 제시하고 있는 것이다. 즉 단백질 정보를 코딩하지 않는 정크 DNA로 불렀던 염기서열들이 실제로는 mRNA로 전사되어 유전자 발현 조절 등 특정한 역할들을 수행하고 있음이 밝혀진 것이다.

요컨대 유전체는 유전자들의 단순한 집합체가 아니라는 사실이 강조된 것인데, 주목할 만한 것은 유전체 DNA 서열중의 일부분은 유전자 발현을 조절하는 단백질에 대한 결합지점(landing spot)으로 작용하며 일부 서열은 RNA 사슬로 변환되어 유전자 발현을 조절하는 기능을 수행한다는 것이다. 지금까지 RNA는 DNA 서열이 전사되어 만들어짐으로써 단백질 합성 정보를 제공하는 정보 전령자 역할을 담당하는 것으로 보편적으로 이해되어 왔으나, ENCODE에서는 상당수의 RNA 분자가 단백질 합성에 관여하기 보다는 유전자 발현의 조절이라는 그 자체의 독자적인 기능을 가지고 있음을 밝혀내게 된 것이다. ENCODE 프로젝트의 결과를 보다 구체적으로 살펴보면 다음과 같이 요약된다. 인간 유전체는 80.4%에 달하는 영역이 기능을 보유하고 있는데 그 영역들은 다음 3가지로 분류되는 영역들, 즉 1) 다양한 RNA 분자들을 발현하는 부위, 2) 히스톤 수식(histone modification)이 강화된 영역, 3) 느슨한 염색질(open chromatin) 영역 또는 전사인자(transcription factor) 결합 부위 영역 중 하나이다. 그 중 가장 광범위하게 보유하고 있는 영역이 RNA 분자 발현 부위로서 전체 유전체의 62%에 달하

며, 그 다음으로는 히스톤 수식 강화 영역으로서 56.1%, 이어서 느슨한 염색질 영역 15.2% 및 전사인자 결합 영역 8.1%의 순으로 많다(Djebali et al., 2012; Thurman et al., 2012; Sanyal et al., 2012; Neph et al., 2012; Gerstein et al., 2012).

이들 영역을 통해 나타나는 기능을 보면, 1) 우선 비암호화 영역 중 일부가 유전자 발현 조절 단백질들의 결합지점으로 작용하여 발현 조절작용에 참여하거나, 2) 또한 일부 영역은 화학적 수식작용의 특이적 위치로 작용하여 위(僞)유전자(pseudogene) 부위를 침묵화(silencing)시키는 조절작용에 기여하고, 3) 한편으로는 전사작용의 결과로 생성된 RNA 산물 그 자체가 완성된 형태로서 단백질 합성과는 무관하게 조절작용에 참여한다는 사실이다. 다시 말해서 ENCODE 프로젝트는 특정 유전자가 얼마나 활발하게 발현할 것인가를 조절하는 메커니즘이 지금까지 알고 있던 것에 비해 훨씬 복잡하다는 사실을 보여준 것이다. ENCODE 프로젝트에서는 인간 유전체로부터 400만개에 달하는 유전자 스위치들이 존재함을 발견하였는데 이들은 조절 대상이 되는 유전자에 가까이 근접해서 존재할 수도 있고 멀리 떨어진 위치에서도 발견된다(Pennisi, 2012). 결국 서로 다른 세포들이 보여주는 각각의 독특한 유전자 발현특성, 즉 유전체적 정체성(genomic identity)은 이 스위치들이 서로 다른 패턴으로 조합됨으로 인해 나타나게 되는 것이다. 뿐만 아니라 유전자 발현 조절작용은 RNA 분자를 통해서도 이루어질 수 있는데, 유전체로부터 전사되어 생성된 특정 RNA 분자들이 특정 유전자의 단백질 발현 즉 단백질의 발현량을 조절하고 있음이 밝혀짐으로써, 유전자 발현 메커니즘이란 극도로 복잡한 전사조절 네트워크를 통해 이루어지는 현상임을 보여주었다(Djebali et al., 2012). 여기에 신호전달 네트워크와 단백질 간 상호작용까지 더해진다면 유전자 발현 현상의 복잡성은 훨씬 더 심화될 것임을 추정하는 것은 어렵지 않다.

또한 ENCODE 프로젝트 연구성과의 의의는 질병 연구에 있어서의 새로운 기초를 제공하고 있다는 점도 빼놓을 수 없다. 각종 질병이 유전체 염기서열 상의 돌연변이에 의해 발병 가능성이 높아지거나 낮아질 수 있는데, ENCODE 프로젝트의 연구결과들은 우선 각종 질환에 대한 발병 가능성이 관련 유전자 자체보다는 그 유전자의 발현 조절에 의해 결정될 수 있음을 제시하고 있다. 뿐만 아니라 비암호화 영역에서의 돌연변이들을 특정해줌으로써 그 변이가 어떤 유전자의 조절에 어느 정도로 영향을 미치는지를 가늠할 수 있도록 하여 질병의 예방과 치료에 활용할 수 있다는 것이다. 질병을 유발하는 대부분의 돌연변이는 유전자 상의 이상 유발이 아닌 유전자 조절에서의

이상을 유발한다는 사실은 염색체 DNA 돌연변이에 의한 질병 유발 메커니즘에 대해 새로운 시각을 제공하는 개혁적 발견이라 할 만하다.

2. 정크 DNA

유전체에서 특정한 단백질 정보를 가지고 있지 않은 DNA 염기서열을 비암호화 DNA(non-coding DNA)라고 부르는데, 유전체 내의 비암호화 DNA의 비율은 생물체마다 크게 달라서 인간 유전체의 경우에는 98%를 넘는 것으로 추정되지만 세균 유전체에서는 20% 수준에 그친다(Elgar & Vavouri, 2008; Costa, 2012). 이 비암호화 DNA 서열 대부분은 그 기능을 알 수 없기 때문에 ‘정크 DNA’라고 불려왔다. 정크 DNA는 1960년대에 만들어진 개념으로서(Ehret & De Haller, 1963), 1972년 Susumu Ohno에 의해 공식적인 용어로 사용되기 시작하였는데, 그는 생물체에서 일어나는 해로운 돌연변이들로 인해 유전체 상의 기능적 영역의 위치(로커스, locus)는 생물체마다 특정 수준을 넘을 수 없다고 주장하면서 포유동물 유전체의 경우에는 자연선택의 과정에서 30,000개 이상의 로커스를 가질 수 없고 그 이외의 영역은 특정 기능이 관찰되지 않는 영역이라고 추정하였다(Ohno, 1972). 한편 1980년대에 들어와 거대 유전체에서 나타나는 비암호화 DNA의 대부분이 전이인자(transposable element 또는 transposon, 유전체 내에서 다른 부위로 이동하는 능력이 있는 DNA 단위)들의 이기적 유전자 증복(selfish amplification)으로부터 기인한다는 관점이 지배적으로 대두되었다(Orgel & Crick, 1980). 그 이후 현재까지 정크 DNA라고 알려진 가장 많은 경우는 전이인자들로 규정되면서 사람 유전체의 40% 이상을 차지한다고 주장되어왔다. 그 외에도 유전체 내의 내인성 레트로바이러스(endogenous retrovirus) 서열이 8% 이상 포함되어 있다. 이러한 정크 DNA 서열들은 발현이 억제된 상태의 무의미한 반복 서열의 형태로 총 인간 유전체의 98%를 차지하고 그 사이에서 드물게 나타나는 유전자 서열들이 2% 이하의 빈도로 포함되어 전체 유전체를 구성하고 있다는 것이다.

결국 유전체의 대부분을 차지하고 있는 정크 DNA 서열들이 왜 이처럼 유전체의 상당 부분을 차지하고 있는가라는 핵심적 질문이 대두되었다. 진화론적 관점에서 시작된 정크 DNA의 개념은 자연스럽게 오랜 진화과정에 의한 기능 상실을 보여주는 증거로 귀결되었고 한편으로는 생물체가 복잡해져가는 과정에서의 산물이라고 해석되었다.

Ohno는 1970년에 유전자 중복을 진화와 종분화의 주된 원인이라고 주장하였다(Ohno, 1970). 이와 같은 맥락에서 W. Ford Doolittle과 Carmen Sapienza는 1980년에 발표하기를, 유전체 내의 어떤 DNA 영역이 그 표현형적 기능을 보여주지 않더라도 그 유전체의 생존(genomic survival)을 위해 진화한 것이라면 그것이 바로 그 존재의 이유가 된다고 주장하였다(Doolittle & Sapienza, 1980).

위와 같이 정크 DNA가 등장하고 그 의미를 부여한 역사를 살펴볼 때 정크 DNA는 결코 객관적인 과학적 사실에 근거한 개념이 아닌 무신론적 진화론에 근거하여 만들어진 기원과학 이론의 하나임을 확인할 수 있다. 즉 비암호화 DNA에 대한 해석을 내림에 있어 철저한 진화론적 관점 아래, 이기적 유전자의 중복이라는 논리를 배경으로 만들어낸 하나의 추론적 개념인 것이다. 그러나 이 추론은 잘못된 가정으로부터 시작되어 잘못된 결론으로 귀착된 오류 이론임이 ENCODE 프로젝트의 연구결과에 의해 여지없이 밝혀지고 만 것이다. 앞서 설명한 바와 같이, 무의미 반복서열로만 해석되었던 DNA 서열의 상당부분이 유전자 발현을 조절하는 중요한 역할을 하는가 하면 또는 실제 다양한 조절 RNA 분자들을 만들어내는 전사 발현부위라는 사실이 확인되었다. 다양한 조직을 대표하는 140여 종류의 세포에서 정크 DNA 부분의 활성화는 다양하게 나타났고 그러한 다양성은 곧 각 세포 즉 조직의 특징적 발현에 중요한 원인으로 분석된 것이다. ENCODE 연구결과가 발표되자 사이언스지는 “ENCODE 프로젝트는 정크 DNA의 명복을 비는 글을 썼다”고 평가하였다(Pennisi, 2012). ENCODE 프로젝트의 연구결과를 시작으로 앞으로 전개될 유전체학적 연구들로 인해 정크 DNA 개념의 종말이 예고되고 있는 것이다.

3. 정크 DNA의 생화학적 기능에 관한 논쟁

위에서 설명한 바와 같이, 유전체 내의 비암호화 DNA 영역들이 생물체의 진화과정에서 유전체의 생존을 위해 어떤 특정한 기능을 보여주지 않는 상태로 나타나고 존재할 수 있으며 그것이 종분화의 원인이 된다는 정크 DNA의 관점은 1970년대 이후 1980년대에 이르면서 탄탄하게 뿌리를 내리기 시작하였다. 그러나 ENCODE 프로젝트 연구결과가 발표되기 이전에도 정크 DNA를 제시하는 진화론적 관점의 예측과는 큰 차이를 보여주는 의외의 연구 결과들이 잇달아 발표되면서 정크 DNA 개념에 대한 의

문은 지속적으로 제기되어 왔다.

우선, 기존에는 아무런 기능이 없는 정크 DNA로 여겨지던 DNA 서열들이 실제로는 특정한 기능을 가지고 있음이 새롭게 발견되는 경우들이다. 예를 들면, 인간 유전체에서 반복하여 나타나는 전이인자들이 기존에는 특별한 기능이 없는 것으로 규정하였던 진화론적 주장과는 달리 배아의 발달과정에서 유전자 발현을 조절하는 중요한 역할을 담당하고 있다는 사실이 밝혀진 것이 대표적이다(Hardman, 2004). 또한 무엇보다도 세포 내에서 단백질 정보를 암호화하고 있는 유전자들도 중요하지만 그들의 발현을 조절하는 작용이 유전자 자체 못지않게 중요할 뿐만 아니라, 생물체의 복잡성은 그 조절작용의 복잡성으로 대변할 수 있음이 밝혀져 왔다. 유전체의 비암호화 영역인 인트론(intron)이 조절작용을 가지고 있음을 보여주는 다양한 연구결과들이 보고되고(Stern et al., 1996), 유전자 발현 스위치 작용을 하는 피크논(pyknon)이라 불리는 유전자 조절작용 모티프(motif)들이 발견되는 등(Rigoutsos et al., 2006), 비암호화 DNA 영역에는 더 이상 ‘정크 DNA’라 규정할 수 없는 고유의 기능들이 존재한다는 사실이 확인되어 왔다. 이렇게 다양한 조절작용을 담당하는 비암호화 DNA 서열들이 발견됨으로 인해 정크 DNA 개념의 오류 가능성은 이미 충분히 예견되어 왔음을 알 수 있다.

한편, 정크 DNA 개념을 무력하게 만드는 각종 생물체의 유전체를 발견한 것도 주목할 만하다. 열대에서 서식하는 식충 식물인 열대통발 기바(*Utricularia gibba*)의 유전체를 서열 분석한 결과가 2013년 Nature에 발표되었다(Ibarra-Laclette et al., 2013). 열대통발 기바의 유전체는 현재까지 알려진 식물체의 유전체 중 가장 크기가 작다. 결과에 따르면 열대통발 유전체는 단백질을 암호화하는 유전자와 그들을 제어하는 짧은 DNA 서열들이 유전체의 약 97%로서 대부분을 차지하고 있어서 정크 DNA라 판단할 만한 서열을 거의 가지고 있지 않았다. 연구자들은 열대통발 기바의 유전체가 작고 정크 DNA 부분이 매우 적은 이유를 정크 DNA가 사라지는 방향으로 진화과정이 있어났기 때문이라는 해석을 내리면서 여전히 진화론적 입장을 고수하고 있으나, 이는 정크 DNA가 진화과정에서 발생된다는 기존의 개념과의 정면 충돌을 피하기 위한 변명적 해석임을 쉽게 알 수 있다.

이렇듯 여러 연구결과들이 정크 DNA 개념의 오류 가능성을 강력하게 제기하여 왔음에도 불구하고 그 개념 자체에 대한 근본적인 검토와 수정은 이루어지지 않은 채 지금에 이르면서 학생이나 비전문적인 대중에게 무차별적으로 가르쳐지고 제시되어

왔다. 결국 정크 DNA 가설에 대해 결정적인 반론을 제기한 것은 ENCODE 프로젝트의 방대한 연구결과들이었다. 인간 유전체의 98%를 차지하는 비암호화 영역 중 80%가 다양하고 복잡한 기능을 수행하고 있다는 분석결과는 40여 년간 지속되어온 진화론적 정크 DNA 가설이 심각하게 잘못된 가정 하에 설정된 이론임을 강하게 시사해준 것이다. 하지만 정크 DNA 가설을 지지하는 학자들은 단순히 포기하지 않고 있다.

우선 Graur 등(2013)은 ENCODE 프로젝트의 결과를 ‘진화론 없는 복음(evolution-free gospel)’이라고 빗대면서 반박하였다. ENCODE에서는 기능(function)과 기능성(functionality)의 의미를 구별하지 못하였다고 지적하면서, 진화의 역사를 보여주는 기능을 선택된 효과(selected effect)로서의 기능으로, 진화의 역사와는 무관한 기능을 원인 역할(causal role)로서의 기능으로 규정한 후, ENCODE 프로젝트에서는 선택된 효과로서의 기능을 지나치게 강조하는 오류를 범하고 있다고 비판하였다. 이들은 2009년 Griffiths가 설명한 예(Griffiths, 2009)를 바탕으로 논리적 근거를 제시하였다. 유전체 내에서 전사인자와 결합하는 유사한 두 서열이 있는데 하나는 전사가 일어나지만 다른 하나는 일어나지 않는다고 할 때, 첫 번째 서열(TATAAA)은 우연한 돌연변이로 염기가 조금 바뀌었지만 자연선택에 의해 전사인자와의 결합과 전사작용이 유지되었고, 두 번째 서열에는 전사인자는 결합되지만 실제로 전사가 일어나지 않는다고 가정한다면 이 경우에는 진화를 위한 적응을 초래할 수는 없으므로 결국 이 두 번째 서열은 ‘선택된 효과’로서의 기능은 가지지 않지만 ‘원인 역할’로서의 기능이 관찰될 수 있다는 것이다. 즉, ENCODE 프로젝트는 이러한 진화와는 무관한 ‘기능성’을 지나치게 강조하여 해석에 적용함으로써, 유전체 내 기능적 요소를 단백질 또는 RNA를 만들어 내거나 단백질과 결합 하는 등의 특정한 성질을 나타내는 부분 영역으로 단정적인 규정을 내리고 있다는 것이 문제라고 지적하였다. 이들은 또한 ENCODE 프로젝트의 해석 논리, 즉 1) 특별한 생물학적 과정에서의 기능을 가지고 있는 DNA 서열들은 특정한 “특성(property)”을 나타내는 경향이 있는데, 2) 어떤 DNA 서열이 그와 동일한 특성을 가지면, 3) 그 DNA 서열은 “기능을 갖는다(functional)”라는 논리 전개에 대해서도 문제를 제기한다. 예를 들어, 어떤 무작위적인 서열이 전사인자와 결합할 수도 있지만 그 결과 반드시 전사가 일어나지 않을 수도 있는 것처럼, 어떤 DNA 서열은 반드시 그 추정되는 기능을 나타내지 않은 채로 어떤 특성을 보여줄 수 있다는 것이다.

한편 ENCODE 프로젝트에서 ‘기능을 갖는’ 기준으로 삼고 있는 관찰 대상들, 즉 전

사, 히스톤 수식, open chromatin, 전사인자 결합, DNA 메틸화(DNA methylation) 등에 대해서도, 그것들을 특정한 ‘기능’으로 볼 수 없다는 전혀 다른 관점에 기초한 비판도 제기되었다. 예를 들면, ENCODE에서는 RNA로 전사된 모든 DNA 조각 서열들을 기능 서열로 목록화하고 있지만, 실제 RNA polymerase II에 의해 생성되는 전사체들의 90%가 전사 노이즈(transcriptional noise)로 여겨지고 있는 점과(Struhl, 2007) 그 전사체들이 리보솜에 결합할 수도 있으며 어떤 경우는 번역될 수도 있다는 점에서(Wilson & Masel, 2011), 전사작용 자체를 ‘기능’과 동일한 의미로 볼 수 없다는 주장이다. 더불어, 히스톤 수식 중 2% 이하만이 ‘기능’과 관련된 작용을 나타낸다는 연구결과(Karlic et al., 2010)에 근거하여 히스톤 수식 현상의 일부를 관찰하여 그 모두에 기능적 유의성을 부여한 것도 잘못된 것이라고 지적한다. 이와 더불어, 대부분의 전사 개시 위치(transcription start site)가 open chromatin 영역에 존재하는 것, 전사인자가 결합하는 DNA 서열, DNA 메틸화 등의 현상에 대해서도 다소 차이는 있으나 유사한 논리에 근거하여 ‘기능’에 연결된 해석에 대하여 잘못된 접근임을 문제제기하고 있다.

하지만 이러한 ‘기능’이라는 용어의 정의에 대한 논쟁에 대하여, ENCODE 프로젝트는 기능에 대한 분명한 정의를 제시하였으며 의과학적으로 관련된 기능요소들을 찾고자 하였기 때문에 프로젝트가 내린 결론은 그 자체로 받아들여져야 한다고 반박하는 입장도 제기되었다(Germain et al., 2014). 또한 ‘기능’에 대한 서로 다른 정의를 내림으로써 소모적인 논쟁을 벌일 것이 아니라 ENCODE가 미치는 과학적 영향력, 즉 유전체의 비암호화 영역이 갖는 생화학적 활성화에 대한 후속 연구를 위해 필수적인 데이터들을 제공해주고 있다는 점을 평가하여야 한다는 주장도 주의를 환기시켜주고 있다(Nature Editorial, 2013). 개념에 대한 정의도 중요하고 과학 역시 언어의 한계와 연결되어 있는 것은 맞지만, ENCODE 데이터들을 이용한 연구논문들이 ‘기능’의 정의를 다루는 논문들에 비해 훨씬 많다는 점에서 ENCODE는 그 목적에 맞게 과학계에 잘 받아들여지고 있다는 점을 강조하고 있는 것이다. ENCODE 프로젝트의 리더 연구자 중 한 사람이었던 Ewan Birney는 ‘기능’이라는 개념에 대하여, RNA, “넓은” 또는 “좁은” 의미의 히스톤 수식, DNase I 과민성 부위(DNase I hypersensitive sites), ChIP-seq, DNaseI Footprint, 전사인자 결합 모티프 및 엑손 등을 분석하는 다양한 부류의 분석 방법들에 의해 확인되는 “특정한 생화학적 활성화”를 의미하는 실용적인 개념이라고 설명한 바 있다(Birney, 2012).

위와 같은 ‘기능’ 개념에 대한 논쟁 외에도, ENCODE 프로젝트에서 사용된 세포주 (cell line)들과 타겟으로 설정된 전사인자 뿐만 아니라 실험방법조차 적절하지 못하였다는 문제 제기도 있었다(White et al., 2013). 이러한 비판에 대해서 Mattick 등(2013)은 세포 또는 조직 특이적인 전사작용과 스플라이싱(유전자로부터 전사된 mRNA에서, 비발현부위를 제거하고 발현부위만을 이어 붙여 단일 폴리펩타이드 사슬에 번역하기 위한 mRNA로 개조하는 과정) 작용이 넓게 분포하면서 정밀하면서도 역동적으로 조절되고 있다는 사실은 인간 유전체의 유전적 기능을 보여줌에 있어서 보존된 DNA 염기서열을 추정하는 것(conservation estimates)에 비해 보다 더 정확한 증거가 되는 것이라고 반박하였다. 왜냐하면 보존 서열 추정은 상대적이기도 하고 유전체의 큰 다양성으로 인해 알아내기 어려운 작업이기 때문이라는 것이다. 또한 논란의 중심에 있는 유전체 정보들은 주로 후생유전학적 조절(epigenetic regulation, 즉 DNA 메틸화 또는 히스톤 번역 후 수식 등과 같이 염기서열의 변화 없이 유전자의 활성이 조절되는 것)과 관련된 것들이며, ENCODE 프로젝트 이전에 이미 많은 앞선 연구들로부터 가능성이 제기되어 왔기 때문에 ENCODE의 결과가 전혀 예상치 못할 일은 아니었다(Carey, 2015). 한편으로는 ENCODE 프로젝트가 애초부터 질병 치료와 관련된 유전체 내 기능요소를 찾고자 했던 것이지 진화적 기능요소를 찾는 것이 아니었고, 진화를 위한 자연선택은 유전적 기능을 확립하는데 충분하지도 않고 반드시 필요한 것도 아니기 때문에 서로가 관련은 되어있되 동일한 것은 아니라고 하는 중재적 입장이 주장되기도 하였다(Germain et al., 2014).

ENCODE 연구자들은 2014년에 다음과 같이 정리하여 발표하였다(Kellis 등, 2014): 지금까지 유전체의 기능에 대해서는 세 가지 접근방법, 즉 표현형의 변화를 조사하는 유전학적 방법, 보존 서열을 조사하는 진화론적 방법, 생화학적 활성을 분석하는 생화학적인 방법으로 연구되어왔다. 그러나 이 세 가지 접근방법 모두가 각각의 한계점을 가지고 있다. 유전학적 방법은 생물체에 실제 나타나지 않는 기능요소들을 놓칠 수 있고, 진화론적 방법은 진화적으로 가까운 생물종들의 유전체라 하더라도 서로 상당히 다르기 때문에 여러 종간의 서열 정렬(sequence alignment) 분석은 정확도가 낮을 수밖에 없으며, ENCODE 프로젝트에서 사용한 생화학적 접근방법 역시 높은 재현성에도 불구하고 어떤 기능을 항상 자동적으로 나타내주는 것은 아니라는 단점이 있다. 하지만 결국 DNA 요소들의 분자적 기능을 알 수 있는 단초를 제공해주는 것은 생화학적 데

이터들이기 때문에 ENCODE에서 제공되는 생화학적 활성 지도들(biochemical maps)은 가장 가치 있는 결과물이며, 기능요소들이 실제로 어떻게 분자, 세포, 생물체 차원에서의 기능과 연관되는지를 조사하는 출발점이 될 것이라고 결론지었다.

지금까지 설명한 과학적 측면에서의 토론들과 함께 ENCODE 프로젝트가 과학계와 사회에 미치고 있는 영향에 대해서도 논쟁이 있다. ENCODE 프로젝트가 거대한 연구비를 사용하고 있다는 점을 지적하면서 총 4억불에 달하는 고비용의 거대 프로젝트로 인해 생산성에 대한 비판도 있었다(Timpson, 2013). 442명의 연구자들이 2억 8천8백만 달러의 연구비를 들여서 밝혀낸 것이 무엇인지를 반문하며 거대 과학프로젝트가 생산성이 높은 연구자들이 사용해야 할 연구비를 낭비한 셈이라고 비난하는 것이다. ENCODE 결과가 프로젝트에 투입된 시간을 정당화하지 못하는 수준이라는 비판도 있다. 인간 유전체 프로젝트(Human Genome Project, HGP)와 비교하면서 HGP는 분명한 마무리가 있었는데 ENCODE는 사실상 언제 끝날지 알 수 없다는 점에서 부정적인 평가를 받고 있는 것이다. ENCODE 연구 성과를 과학계뿐만 아니라 일반 대중매체에 설명하는 방식에 대해서도 문제 삼는 목소리도 있었다. ENCODE 프로젝트의 리더 과학자가 ENCODE 연구결과로 인해서 유전체학 교과서를 다시 써야 할 것이라고 평가한 것(Pennisi, 2012; Stamatoyannopoulos, 2012)에 대하여 Graur 등은 “다시 써야 할 책은 유전체학 교과서가 아니라 마케팅, 대중매체를 동원한 광고, 대중 홍보 등에 관련된 책일 것”이라고 비난하였다(Graur et al., 2013). 그러나 이와 같은 과학 안에서의 이슈 이외의 비판들은 뒷받침할만한 타당한 근거가 빈약하다는 점에서 바람직한 비판이 아니라는 지적도 나오면서 ENCODE를 둘러싼 논쟁은 점입가경의 상황 중에 있다.

4. 인간 유전체의 기원

지금까지 인간 유전체에 대한 진화생물학적 해석으로부터 나온 정크 DNA 개념에 대해 살펴보고 최근 발표된 ENCODE 프로젝트 연구결과와의 갈등과 논쟁에 대해 살펴해보았다. 본 장에서는 인간 유전체의 기원에 대한 진화론적 관점을 종합적으로 분석하고, ENCODE 프로젝트를 비롯한 최근 연구결과들을 바탕으로 이에 대한 비평 및 유전체 기원에 대한 새로운 시각을 제시하고자 한다. 인간유전체의 기원을 다루기 위해서는 다음 세 가지에 대한 분석이 선행되어야 한다: 1) 자가복제가 가능한 최초의

유전체와 원시세포(proto cell)가 어떻게 탄생하였는가에 대한 기존의 진화론적인 이론들, 2) 어떻게 단순 원시세포 또는 미생물의 유전체가 복잡하고 정교한 동물의 유전체로 발전할 수 있는지에 대한 가설들, 3) 인간 유전체의 기능에 대한 최신의 연구성과들. 마지막으로 이러한 분석을 기초로 하여 인간 유전체의 기원에 대한 새로운 추론을 제안하고 그 타당성에 대해 논증하고자 한다.

(1) 자가복제 가능한 유전체 및 원시세포의 형성에 대한 진화가설과 비판적 고찰

최초의 생명체 탄생에 대한 부분은 지금까지 여러 가지 가설들이 제시되어 왔으나 (Crick, 1968; Orgel, 1968; Woese, 1967; Robertson & Joyce, 2012; Cech, 2012; Patel, 2015) 가장 단순한 형태의 생명체 조차도 극도의 복잡성을 띄고 있어 진화론 학자들 사이에서도 이에 대해 확고한 결론을 내지 못한 채 치열한 토론이 이루어지고 있는 분야 중의 하나이다. 생명체의 가장 중요한 특징인 생식(reproduction)은 세포 수준에서는 세포의 분열, 분자 수준에서는 유전정보 물질의 자가복제(self-replication)를 기반으로 한다. 따라서 최초 유전체의 생성 및 진화를 설명하기 위해서는 우선 유전체의 자가복제 기능에 대해 살펴볼 필요가 있다. 또한 생명체의 진화가 일어나기 위해서는 새로운 정보들이 유전체 안에 지속적으로 축적되어야 하는데 이러한 유전정보의 자연 발생적 축적이라는 가정이 논리적으로 타당한지에 대해 살펴보아야 하며, 아울러 새롭게 생성된 유전정보들이 다음 세대에 전달될 수 있도록 담보하는 유전물질의 안정성을 고찰해야 한다.

현존 생명체들의 유전체 복제과정은 복제 대상인 DNA와 이를 복제하는 정교한 단백질 복합체들이 동시에 존재할 때 가능한데 이러한 시스템이 우연히 만들어지는 것은 보편적인 물리화학적 법칙 상 매우 확률이 낮을 수밖에 없다. 따라서 진화론 학자들은 현존하는 DNA 복제 시스템보다 훨씬 단순하면서도 자가복제가 가능한 물질들을 찾기 시작했고 그 대안으로 제시된 것이 바로 RNA 효소(RNA enzyme), 즉 라이보자임(ribozyme)이었다(Crick, 1968; Orgel, 1968; Woese, 1967). RNA 분자는 어떤 염기서열을 갖느냐에 따라 종종 효소의 기능도 갖게 되는 것으로 알려져 있는데(Guerrier-Takada et al., 1983; Gilbert, 1986; Cech, 1987), 리보솜(ribosome) 안에 들어있는 리보솜 RNA(rRNA)가 그 예이다. 하지만 이러한 RNA는 그 특성상 오랜 시간 그 기능적

구조를 유지하지 못하는 불안정성 때문에 함께 결합된 단백질들의 보호와 도움을 받아 리보솜 안에서 그 기능을 유지하고 있다. 하지만 만약 자가복제가 가능한 형태의 RNA가 우연히 만들어질 수 있다면, 이러한 RNA와 이를 보호하고 둘러싸는 지질막만으로 가장 원시적인 세포의 기능을 수행할 수 있다고 보는 모델이다.

그러나 이러한 가설에는 입증하기 어려운 논리적 허점이 있다. 최초의 생명체가 자기 복제 가능한 RNA에서 우연히 시작되었다면 이러한 기능을 가진 RNA는 그 기능을 잃어버리기 전에 계속해서 빠르게 복제되어야 한다. 뿐만 아니라 이러한 신속한 복제 기능이 매우 오랜 시간 동안 유지된 상태에서 그 분자 자체에 지속적인 돌연변이가 일어나 새로운 유전정보를 생성, 저장해야 한다. 또한 DNA보다 훨씬 불안정한 RNA 효소가 자연상태에서 보존되기 위해서는 오랜 시간 동안 기능하는 일종의 특별한 보호장치가 반드시 필요하다. 그러나 이 모든 필요성들에도 불구하고 진화론적 관점의 논리는 RNA 효소에 의한 자가복제 생명체의 발생 가능성만을 단순하게 제시하는 수준에 머물러 있다.

한편 원시세포의 형성과정에 대한 진화이론도 가능성 낮은 가정을 전제하고 있는 동일한 문제를 안고 있다. 대표적인 원시세포의 자연발생 이론을 요약하면, 원시 지구에서 단순한 형태의 지방산들이 수중의 특정 조건 하에서 자연적으로 생성되고, 그 지방산들이 밀집될 수 있는 특정한 환경 - 지방산을 포함한 물방울의 기화 등 - 에서는 지질 소낭(lipid vesicle)이 형성될 수 있을 것으로 본다. 그 소낭에 만약 자가복제가 가능하도록 자연발생된 단일가닥 RNA 분자가 포함된다면 이러한 조합이 원시세포의 원형이 될 수 있다고 보는 것이다(Szostak et al., 2001). 즉 지질 소낭 안에서 자가복제가 가능한 RNA가 소낭 안팎의 농도차에 의해 소낭 내부로 확산되어 들어온 RNA 단량체(RNA monomer)들을 이용해서 자기 자신을 계속 복제할 수 있을 것이라는 가설이다. 하지만 이를 위해 필요한 전제 조건은 첫째, 최초의 RNA는 소낭 안에서 순수 화학반응에 의해 자신을 주형(template)으로 하여 이중나선 형태의 중합체(polymer)를 형성해야 한다. 둘째, 이렇게 형성된 이중나선 중합체는 RNA 자가 복제를 위해 단일 가닥들로 분리되어야 한다. DNA나 RNA의 이중나선은 높은 온도에서 분리(denaturation)되고 낮은 온도에서 합체(renaturation)되는데 지속적인 복제과정이 일어나기 위해서는 이러한 반복적인 온도 변화가 선행되어야 한다. 셋째, 이중나선 RNA 분자가 분리된 상태에서 온도가 내려가면 그 중 한 가닥은 3차원 구조를 형성하면서 RNA 중합효소

(RNA polymerase)의 기능을 갖고, 그 상보적 가닥(complementary strand)을 이용하여 자신과 동일한 서열을 가진 RNA 분자를 복제한 뒤 재차 단일가닥들로 분리되어 위의 과정을 반복한다. 이러한 가정이 제안된 이유는 처음 한번은 우연히 화학적인 방법으로 이중 나선형태의 중합체를 형성한다고 가정하지만 매번 이와 같은 방식으로 진행된다면 그 더딘 속도로 인하여 그 전에 RNA 중합효소로 작용하는 기능성 RNA 분자가 분해될 가능성이 높아지고 결국 지속적인 복제가 불가능해지기 때문이다. 넷째, 이런 방식으로 형성된 원시세포들이 지속적으로 진화하기 위해서는 서로 경쟁하는 군집이 형성되어야 한다. 즉, 다양한 돌연변이를 가지고 있는 수많은 자가복제 원시세포들이 함께 존재해야 한다. 이 집단에서 특정 환경에 우세한 특성을 가진 원시세포가 경쟁을 통해 자연선택되고 그 유전정보가 집단 내에 보편화될 수 있다고 보기 때문이다.

그러나 이렇게 요약되는 원시세포 형성에 대한 진화론적 시나리오를 분석해볼 때, 각 단계에 대한 타당한 증거의 제시 없이 단순한 가정과 추론들의 연속적 구성으로 이루어져 있다는 문제점이 발견된다. 자가복제가 가능한 RNA 분자와 원시세포가 특정 환경에서, 자연발생적으로, 그리고 안정적으로 다량으로 존재했을 것이라는 가설은 말 그대로 가설일 뿐, 그를 뒷받침할 다양한 실험적 또는 이론적 증거들은 물론 그 가설을 실제로 검증할만한 과학적 방법론조차도 제대로 제시되지 않았다. 이러한 입장에 동조하는 진화론 학자들까지도 이러한 시나리오는 최초의 생명체 발생과 관련하여 ‘개연성이 있는’ 순서와 과정에 대한 가설이라고 평가한다(Szostak, 2012). 뿐만 아니라 최근의 생명체 발생 또는 생명체 합성에 대한 연구들도 이 가설에 대한 충분한 증거를 제시하지 못하기는 마찬가지이다. 오히려 지금까지 보고되고 있는 연구결과들에 따르면, 자가복제가 가능한 단순 소낭을 형성하는 것조차도 고도화된 기술이 적용된 실험실 조건 하에서 극도로 통제되고 정밀하게 조절되는 다수의 인위적인 단계들에 의해서만 일부 구현이 가능하며, 더욱이 이를 위해서는 실험자의 의지적 개입과 미리 고안된 DNA 프라이머(DNA primer)는 물론, DNA 중합효소(DNA polymerase) 등과 같은 단백질 기능체가 필요하다는 것을 확인하였다(Kurihara et al., 2015). 이는 최신 생명공학 기술을 동원한 인위적 실험에서조차 RNA 분자와 지질 소낭으로부터 자가복제가 가능한 시스템을 구현해내지 못하고 있으며, 이는 RNA와 지질 소낭이라는 물질 자체만으로 스스로 자기복제 시스템을 만들어낸다는 진화이론이 단순 가설 수준에 머물러 있음을 보여준다.

(2) 단순 원시세포 유전체로부터 복잡한 동물 유전체로의 진화가설과 비판적 고찰

또한 가상의 원시세포가 만들어진다 해도, 그것이 현존하는 생명체의 조상이 되는 최초의 생명체, 즉 LUCA(Last Universal Common Ancestor)로 발전하기 위해서는 앞서 살펴본 자가복제 시스템의 형성과는 비교할 수 없는 복잡한 진화의 단계들을 거쳐야 하지만, 이에 대해서는 “개연성이 높은” 가설 수준의 해석조차 아직까지 제시되지 못하고 있다. 이를 위한 핵심과정, 즉 오랜 시간에 걸쳐 원시세포들의 진화를 가능케 하는 플랫폼인 원시 유전체가 어떻게 지속적으로 유전정보를 축적하여 현존하는 가장 단순한 박테리아와 비슷할 것으로 추정되는 LUCA의 유전체로 발전할 수 있을지, 그리고 그 과정 중에 어떻게 RNA 효소의 기능이 자연스럽게 DNA와 단백질 효소로 분산되어 전이될 수 있는지 등 핵심적 진화단계들에 대해서 설득력 있는 가설이 제시되지 못하고 있을 뿐 아니라 오히려 진화학자들 사이에서도 의견이 나뉘어 있다. 더 나아가 가장 단순한 형태의 박테리아가 이와는 비교할 수 없을 정도의 복잡성과 체계성을 가진 다세포 동물의 유전체로, 그 안에서도 무척추동물을 거쳐 어류, 양서류, 파충류로 이어지는 척추동물의 유전체로 발전하는 과정을 설명하기 위한 설득력 있는 유전체학적 해석들은 요원한 상황이다.

현재까지 발견된 박테리아 중 가장 단순한 경우는 최소 1,000개 정도의 유전자를 포함하고 있는 것으로 알려져 있다(Hutchison et al., 2016). 인간을 포함한 대부분의 척추동물은 최소 19,000~20,000개 정도의 유전자를 가지고 있는 것으로 보아 가상의 최초 생명체에서 척추동물까지 진화가 이루어지기 위해서는 오랜 시간을 거친다고 가정하더라도 급격한 유전정보의 증가가 이루어져야 한다. 뿐만 아니라 생명 현상의 유지를 위해 단순한 양적 증가뿐 아니라 유전자들의 발현을 조절하는 조절 DNA의 진화가 함께 이루어져야 한다. 지금까지 유전체의 진화를 설명하기 위해 제시되어 온 가설로는 전체 유전체 중복(whole genome duplication), 부분 유전체 중복(partial genome duplication), 유전자 이동(gene transfer) 등을 들 수 있다. 하지만 이들 가설들은 이미 기능적으로 완성된 상태의 자가복제 가능 유전체가 존재한다는 가정 하에서만 설명이 가능하다. 또한 자가복제가 가능한 유전체가 우연에 의해 만들어졌다 하더라도 전혀 다른 기능을 수행하는 수많은 독립 유전자들이 어떻게 하나 또는 소수의 유전자로부터 파생되어 만들어질 수 있는지를 설명하는 것은 매우 어려운 주제이다. 뿐만 아니라,

지금까지 실험적으로 발견된 대부분의 돌연변이들은 정보의 증가를 가져오지 못한 것으로 보고되고 있다(Bergman, 2005). 유전적 다양성의 감소는 현재 관찰되는 가장 뚜렷하고 지속적인 현상이며, 현재까지 관찰된 돌연변이들이 초래한 현상은 중복(duplication)을 통한 동일한 유전자 숫자의 증가, 그리고 유전자의 변형이었다(Liu & Moran, 2006). 또한 이른바 소진화(microevolution)적 적응과 동반하는 유전적 변화가 축적된다 하더라도 그것이 대진화(macroevolution)로 연계되어 특정 생물종에서 다른 생물종으로 종분화(speciation)가 일어나는 현상은 실험적으로 전혀 관찰된 바 없다(Dembski & Wells, 2008)

미국의 비영리 유전체 연구소인 J. Craig Venter Institute(JCVI)는 현존하는 가장 단순한 박테리아 중 하나인 *Mycoplasma mycoides*의 유전체로부터 생명현상을 유지하는 최소한의 유전체를 만들고자 시도한 내용을 2016년 사이언스(Science)지에 발표하였다(Hutchison et al., 2016). 즉, 정크 DNA를 제거하여 생명현상에 필수적인 유전자들만을 가지는 유전체를 만들고자 한 것인데, 흥미로운 점은 이 단순한 박테리아조차 그 기능이 이미 최적화되어 있다는 사실이었다. 특히 대사와 관련된 유전자들은 거의 대부분 핵심적인 기능을 수행하고 있었으며, 반복적인 실험을 통해 유전체의 크기를 최소한으로 줄여보려 하였으나, 현존하는 가장 작은 유전체보다 작게 만드는 것은 실패하였다. 뿐만 아니라 이렇게 인위적으로 축소된 유전체를 주입한 박테리아는 오직 실험실의 특정한 조건에서만 증식이 가능하며 여러 미생물생리학적 특징이 현저하게 약화되는 경향을 보여주었다. 이러한 박테리아들은 본래 보호된 실험실과 같은 조건이 아닌 척박하고 다양한 자연환경에서 살아남을 수 있음을 감안할 때, 정크 DNA라고 불리는 많은 부분이 사실은 생존과 관련된 핵심적인 기능을 가지고 있으며 이를 담고 있는 세포 내 유전체는 현재의 과학 수준으로는 아직 다 이해하기 어려운 매우 정밀한 생명 정보 시스템으로 설계되어 있음을 미루어 알 수 있다.

(3) 인간 유전체의 기능에 대한 최신의 연구성과들에 대한 분석

인간유전체는 2003년 이후 지금까지 38번 이상의 염기서열 개정과 재분석 작업을 거치면서 비교적 소상하게 그 실체가 드러나고 있다. 특별히 ENCODE 프로젝트와 관련 후속 연구들을 통해 인간 유전체의 구조적 기능들이 구체적으로 밝혀져 가고 있다.

현재 시점에서 인간 유전체 중 단백질 정보 암호화 서열, 즉 단백질 코딩 유전자는 대략 20,000개 정도인 것으로 추정되고 있다(Maher, 2012). 이는 기존에 예상했던 100,000개보다 훨씬 적은 수인 만큼 과학계에 큰 반향을 일으키는 발견이었으며, 그와 더불어 유전자 자체보다 유전자 발현을 조절하는 조절 DNA(regulatory DNA)의 중요성이 새롭게 밝혀지기 시작한 것도(Thurman et al., 2012; Neph et al., 2012; Djebali et al., 2012) 변혁적인 발견이 아닐 수 없다. 우리 몸은 대략 수십조개에 이르는 세포가 존재할 뿐만 아니라 세포의 종류도 최소 400여 가지가 있는 것으로 알려지고 있다(Vickaryous et al., 2006). 이러한 다양한 세포들의 고유한 기능을 유지하기 위해 유전체 내 유전자들은 각 세포 특성에 따라 각기 다양한 조합으로 발현되는데 이때 세포 특이적 발현을 조절하는 것이 바로 조절 DNA이다. 조절 DNA는 보통 유전자 코딩서열의 앞부분(upstream region) 혹은 뒷부분(downstream reason)에 복수로 위치하여 특정 세포 유형과 기능을 결정하는 유전자들의 발현을 정밀하게 조절한다. ENCODE 프로젝트 연구를 통해 140종의 인간 세포들의 고유한 활성에 관여할 것으로 예상되는 조절 DNA에 대해 확인한 결과 인간 유전체에서 차지하는 비율은 기존의 예상을 훨씬 뛰어넘는 수준(2012년 기준으로 약 40%)인 것으로 확인되었다(Stamatoyannopoulos, 2012). 하지만 120종의 세포들을 차례대로 분석하여 각 세포 특이적 조절 DNA 들을 찾아본 결과 각각의 세포종에서 발견된 조절 DNA들의 누적 속도가 감소하지 않고 지속적으로 유지되는 것으로 보아 400여 종류에 이르는 사람 세포내 유전체 내 조절 DNA의 양을 모두 계산한다면 그 값은 크게 늘어날 가능성이 높다. 뿐만 아니라 같은 세포 종이라 할지라도 다양한 생리학적 자극과 환경에 따라 기존에 발견되지 않은 조절 DNA 가 존재할 수 있다는 점도 주목해야 한다.

또한 본 논문의 전반부에서 고찰한 바와 같이, 지금까지 인간 유전체의 대부분이 전이인자(transposable element)등 또는 여러 반복적인 서열들로 인해 생성된 쓸모 없는 영역으로 해석되어 왔으나, ENCODE 프로젝트를 통해 이들 중 상당 부분이 실제로는 세포 특이적인 조절 인핸서(regulatory enhancer)의 역할을 담당하고 있는 것으로 확인되었다(The ENCODE Project Consortium, 2012; Thurman et al., 2012; Djebali et al., 2012; Gerstein et al., 2012; Sanyal et al., 2012). 즉, 정크 DNA 개념을 기반으로 하는 기존 진화론적 학설을 정면으로 부정하는 관찰 결과들이 계속해서 보고되고 있는 사실에 주목할 필요가 있다. 뿐만 아니라 진화 과정에서 축적된 DNA 회석이라 불

리는 위유전자(pseudogene)가 특정 세포 또는 조직에서 매우 중요한 기능을 수행하는 것이 밝혀지고 있다. 위유전자는 유전자로서의 특징들을 가지고 있으나 프로모터나 암호화 서열 내에서 발생한 돌연변이로 인해 단백질 발현 자체가 안 되거나 발현되더라도 프레임 이동(frameshift) 등에 의해 무의미한 비기능성 단백질(nonfunctional protein)을 합성하는 것으로 알려져 있다. 진화론에서는 이를 두고 중복 유전자(duplicated gene)가 진화 과정에서 그 기능을 상실한 후 화석처럼 유전체 안에 축적된 것이라 해석해 왔다. 일반적으로 위유전자는 프로세스드 슈도진(processed pseudogene)과 언프로세스드 슈도진(unprocessed pseudogene)으로 나누어지는데, 이 둘의 차이점은 인트론(intron) 혹은 이와 유사한 서열의 존재 여부이다. 언프로세스드 슈도진은 인트론 서열 등 유전자의 정상적인 요소들을 모두 가지고 있는 위유전자로서, 유전자 중복(gene duplication)에 의해 복사된 유전자 중 하나가 기능을 잃게 된 것이라는 설명이 진화론의 보편적인 해석이었다. 그러나 언프로세스드 슈도진 중의 한 예인 베타-글로빈(β -globin) 슈도진 이 몇 가지의 조절 RNA를 발현하는 기능성을 가지고 있으며, 이 부분에 일어난 돌연변이가 여러 가지 혈액 관련 질병과 관련이 있음이 밝혀졌다(Nuinoon, 2010; Giannopoulou et al., 2012; Roy et al., 2012).

한편 프로세스드 슈도진은 인트론 서열이 관찰되지 않는 위유전자 유형으로서, 진화론 학자들은 mRNA가 역전사(reverse transcription)된 후 cDNA 형태로 유전체에 재삽입되는 방식으로 발생된 것이라고 해석해왔으며, 그런 의미로 이러한 유형의 위유전자를 레트로진(retrogene)이라 불렀다. 하지만 이렇게 프로세스드 슈도진을 기능이 없는 DNA 화석이라고 보는 진화론적 해석과는 달리, 여러 레트로진들이 매우 중요한 기능들을 수행하고 있음이 계속하여 보고되고 있다(Soarse et al., 1985; Ciomborowska et al., 2013). 또한 더욱 흥미로운 것은 레트로진 위유전자 중 상당수가 실제 활발하게 전사가 일어날 뿐 아니라 일부는 다른 유전자를 조절하는 인핸서로 기능하고, 또 일부는 다른 유전자 발현을 억제하는 siRNA와 같은 조절 RNA로서의 기능을 수행하는 것으로 밝혀지고 있다는 사실이다. 예를 들면, 레트로진 중 하나인 PPM1K 위유전자는 여러 세포 내에서 활발하게 전사가 일어나며, 암환자의 경우 PPM1K의 RNA 발현량이 정상인에 비해 현저히 낮은 것으로 확인되었고, 이 PPM1K에서 발현하는 siRNA는 간세포암(hepatocellular carcinoma)의 발병과 관련 있는 유전자인 NEK8의 발현을 억제하는 역할을 하는 것으로 밝혀졌다(Chan et al., 2013). 진화론적으로는 PPM1K 위유

전자가 인간의 진화 과정에서 무작위적인 변이로 인해 우연히 만들어진 기능 없는 DNA 화석으로 이해하고 있었지만, 관찰된 현상은 전혀 그렇지 않았다. 세포의 생명현상을 유지해주는 중요한 기능을 가지고 있음이 확인됨과 동시에 원래부터 그러한 기능을 가지고 유전체의 한 요소로 존재하고 있었음을 강하게 시사해주고 있는 결과인 것이다.

(4) 인간 유전체에 대한 새로운 추론과 그 타당성

이상 정리한 내용으로부터 알 수 있는 바와 같이, ENCODE 및 후속 연구들로 인해 그 동안 정크 DNA로 불려왔던 인간 유전체의 대략 40%에 해당하는 다양한 반복서열들은 물론, 여러 위유전자들이 중요한 생물학적 기능을 수행하고 있는 것이 새롭게 밝혀지고 있다. 인간 유전체 기능연구가 진행될수록 인간 유전체의 다양한 부분들이 쓸모없는 진화과정의 화석들이라는 진화론적 추론은 그 설득력을 잃어가고 있다. 반면 최근의 인간 유전체 기능 연구 결과로부터 인간 유전체의 다양한 기능들이 새롭게 발견되고 있음을 주목할 만하다. 즉 겉으로 보기에는 무의미한 반복서열 내지는 돌연변이로 인해 유발된 변형 서열로 인간 유전체의 상당 부분이 다양하고 중요한 기능들을 수행하고 있다는 사실이 계속해서 확인되고 있다.

이러한 최근의 기능유전체학적 관찰결과들을 통해 확인되고 있는 인간 유전체에 대한 과학적 사실들은, 결국 인간 유전체의 기원에 대해 기존의 진화론적 관점이 아닌 새로운 관점으로부터 출발하는 새로운 추론의 필요성을 역설한다. 본 논문에서는 그 새로운 이론으로서 인간 유전체에 대한 설계론적 관점을 제안한다. 즉, 인간 유전체를 처음부터 정교하게 설계된 시스템으로 해석하는 설계 이론이다. ENCODE 연구결과로부터 알 수 있는 바와 같이 인간 유전체는 20,000개 정도의 이외로 적은 수의 유전자들을 가지고 있지만 각 세포마다 서로 다른 유전자들이 활성화 되어 각 세포들의 차별화된 특성들을 나타내고 있으며 그 특성들이 곧 인간이라는 생명체의 생명현상을 구현해내고 있다(Ecker et al., 2012). 즉 유전체는 발현되는 주체이자 조절 받는 대상이 되는 유전자들과 그 유전자들의 발현을 정밀하게 조절해주는 비유전자 즉 단백질 비암호화 영역들로 구성되어 있는 것이다. 또한 인간 유전체의 구성 특징을 보면 발현 주체인 유전자보다는 그것을 조절하는 요소들이 훨씬 더 많은 부분을 차지하고 있다

는 사실이 인간 유전체의 복잡성을 대변해주는 현상이라 할 수 있다. 이러한 복잡성은 발현 주체와 그 발현을 조절하는 조절자들로 구성된 하나의 시스템으로 특징지을 수 있는 설계 특성으로 해석하는 것이 가능하다.

조절 DNA가 없는 유전자는 전혀 생명현상에 필수적인 단백질을 적시적소에 만들어낼 수 없고, 유전자가 없는 조절 DNA는 조절 대상이 없는 무의미한 존재가 될 수밖에 없다. 유전자와 조절 DNA는 이렇게 상호 불가결의 필요충분조건으로 공존한다. 이는 함께 동시에 존재해야만 존재의 의미가 부여되는 당위성과 필연성이 있는 시스템 특성의 전형적 예이다(Demski, 2001; Dembski & Wells, 2008; Meyer, 2010). ENCODE 인간 유전체 연구를 통해 발견된 결과들을 종합해보면, 유전자와 조절 DNA가 협력하여 구현하는 생명현상은 세포의 활성을 높이거나 낮추는 수준이 아니라 그 세포의 기본 요소들을 만들어내고 세포의 정체성과 기능을 결정하는 핵심 현상들이다(The ENCODE Project Consortium, 2012). 자동차로 비유하면, 차의 색깔이나 부가적인 액세서리가 아닌 차의 구조와 구동이라는 근본적인 기능에 해당한다고 할 수 있다. 인간 유전체의 유전자와 조절 DNA들도 인간 세포의 기존 기능의 정도를 변화시키는 것이 아니라 기능 그 자체의 존재와 더 나아가서 그 세포의 유형이 어떤 것인지를 결정하는 것이다. 이러한 이유로 인간 유전체는 그보다 단순한 어떤 유전체로부터 오랜 시간 동안의 변이를 거치면서 새로운 기능들이 점진적으로 추가되어 만들어진 존재라고 단정하기 어렵다. 인간 유전체의 복잡성은 마치 자동차에 액세서리를 하나 둘씩 추가함으로써 부가적인 기능이 향상되는 방식으로 형성된 복잡성이 아니라, 자동차가 자동차 되도록 규정해주고 자동차가 굴러가도록 해주는 복잡성인 것이다. 인간 유전체는 인간이라는 생명체가 갖추어야 할 기본적인 세포 종류들과 생존 및 작동 기능들에 대한 정보의 집합체로 확인된 만큼, 그러한 시스템이 점진적으로 발생, 진화하였다고 가정할 경우에는 그 진화의 과정 중에 나타나는 중간형태의 시스템은 생존과 존재 자체가 담보되지 않는다는 모순적 논리가 되고 만다. 인간 유전체는 그 존재의 출발점에서부터 이미 모든 필수 구성요소들이 완벽하게 갖추어진 상태로 설계된 산물로 해석된다. 또한 구성방식에 있어서도 다양한 유전자와 조절 DNA의 필수불가결한 상호작용들이 생명현상 유지라는 특정 목적을 위해 체계적으로 공존하고 있는 특정화된 복잡성의 전형을 보여주고 있다는 점에도 주목하여야 할 것이다. 그러한 복잡성은 스스로 발생할 수가 없으며, 따라서 인간 유전체는 설계자에 의해 설계된 시스템일 수 있다는

결론에 도달할 수 있다. 무신론적 진화론이 인간 유전체에 대한 올바른 해석을 내리는데 실패하고 있고, 관찰된 현상들에 근거할 때 가장 타당한 해석이 설계이론이라면 현대 과학계는 이를 받아들일 수 있어야 한다. 그러나 현대 과학계가 설계자의 존재를 받아들이지 않는 무신론적 세계관에 의해 편파적으로 주도되고 있다는 것은 주지의 사실이다. 하지만 과학 이론은 비록 기원과학의 분야라 할지라도 그에 대한 종합적인 타당성을 기준으로 판단되어야 한다. 그 이론의 타당성이 확보되었다면 그것이 단지 유신론적이라는 이유로 거부되어서는 안 될 것이다.

III. 결론

인간 유전체 DNA의 최대 80%가 고유의 기능을 가지고 있다는 ENCODE 프로젝트 연구결과는, 인간 유전체의 대부분을 ‘정크 DNA’라고 해석해왔던 기존의 진화론적 추론이 잘못된 것임을 명백하게 보여주었다. ‘비암호화 DNA’라는 객관적 설명이 아닌 ‘정크 DNA’라는 진화론적 개념으로 잘못 이해되었던 인간 유전체의 상당부분에 대해 이제는 새로운 관점으로 설명해야 할 것이라는 선언이기도 하다. ENCODE 프로젝트 연구결과를 총괄하는 논문(The ENCODE Project Consortium, 2012)의 제목처럼, 인간 유전체는 이미 정해진 고유의 기능을 갖는 구성 요소들을 총망라하고 있는 ‘백과사전(encyclopedia)’과 같은 존재이다. 인간 유전체를 구성하는 각 DNA 영역에는 이미 인간 생명현상의 기본적인 고유 기능들이 부여되어 있고 그러한 수많은 영역 요소들이 서로 필수불가결의 상호관계로 연계된 고도의 복잡성을 가지고 있음도 확인되었다. 본 고찰에서는 인간 유전체의 기원 문제와 관련하여, 이러한 최근의 기능유전체학적 연구 결과들과 기존의 진화론적 기원 이론 사이의 불일치로 인해 일어나고 있는 갈등의 내용들을 구체적으로 확인하고 실상을 파악할 수 있었다. 또한 최근의 연구결과들에 근거할 때, 인간 유전체를 무신론적 자연발생의 결과로 보는 진화론적 관점이 심각한 논리적 오류를 가지고 있음을 알 수 있었다. 오히려 인간 유전체가 유전자와 조절 DNA로 구성되는 특정화된 복잡성을 가지고 있음을 미루어 볼 때, 초월적인 설계자에 의한 설계의 산물로 보는 유신론적 관점이 더 타당함을 논증할 수 있었다.

많은 사람들이 유신론적 창조와 설계 개념을 받아들이면 과학의 영역이 없어진다고

말한다. 지루하고 과학의 발전이 없다고 말한다. 하지만 진화론적 해석 일변도의 과학 이야말로 오류와 편견으로 인해 과학 발전을 해치는 결과를 초래할 수 있다. 진화이론 역시 유신론적 이론과 마찬가지로 하나의 기원과학 이론으로서 무신론적 세계관에 입각하여 증거를 제시하는 것이지 과학적 사실을 증명해내는 것이 아니기 때문이다. 따라서 기존의 진화이론에 합리적인 이의를 제기하는 과정을 통해서 기원에 대한 다양한 이론들의 타당성을 견줄 수 있고 그로부터 여러 타당한 결론들이 도출되면서 과학의 진정한 발전을 이룰 수 있는 것이다. 자연에 대한 보다 타당하고 정확한 이해를 위해서는, 관찰되는 현상에 대해 유신론적 해석을 내린다고 할 때, 그 이론이 합리적이려면 그 자연현상 자체가 왜곡되는 것이 아닌 이상 단지 세계관을 이유로 그 이론을 거부해서는 안 된다. 뿐만 아니라 유신론적인 창조론 또는 설계이론을 받아들일 때 자연의 실체에 대한 올바른 이해와 창의적인 활용이 가능해질 수 있음을 인정해야 한다. 이러한 접근방식으로, 인간 유전체를 구성하는 유전자와 그 조절 DNA들의 상호작용은 어떻게 형성된 것인지, 돌연변이는 이것을 어떻게 변형시키고 있는지, 그 결과 어떻게 환경에 따라 다양한 인종과 특성을 나타내게 되었는지 알아낼 수 있다. 그리고 질병과는 어떻게 연관되는지 연구할 수 있으며 이를 통해 질병을 치료하는 새로운 대안을 제시할 수도 있다. 인간 유전체를 비롯한 모든 생명물질과 생명체를 통해서 발견되고 있는 심오한 설계 특성 앞에, 일방적인 진화론적 해석을 고집할 것이 아니라 보다 겸허한 마음으로 다가서서 탐구하는 자세가 필요한 것이다.

“이 논문은 다른 학술지 또는 간행물에 게재되었거나 게재 신청되지 않았음을 확인함.”

참고문헌

- Bergman, J. (2005). "Darwinism and the Deterioration of the Genome." *Creation Research Society Quarterly* 42. 110-112.
- Birney, E. (2012). "ENCODE: My own thoughts." Ewan's Blog: *Bioinformatician at large*. <http://genomeinformatician.blogspot.com/2012/09/encode-my-own-thoughts.html>. (검색일 2016.6.9)
- Carey, N. (2015). *Junk DNA: A Journey Through the Dark Matter of the Genome*. New York: Columbia University Press.
- Cech, T. R. (1987). "The chemistry of self-splicing RNA and RNA enzymes." *Science* 236 1532-1539.
- _____ (2012). "The RNA worlds in context." *Cold Spring Harbor Perspectives in Biology* 4. a006742.
- Chan, W. L., Yuo, C. Y., Yang, W. K., Hung, S. Y., Chang, Y. S., Chiu, C. C., Yeh, K. T., Huang, H. D., & Chang, J. G. (2013). "Transcribed pseudogene ψ PPM1K generates endogenous siRNA to suppress oncogenic cell growth in hepatocellular carcinoma." *Nucleic Acids Research* 41. 3734-3747.
- Ciomborowska, J., Rosikiewicz, W., Szklarczyk, D., Makułowski, W., & Makułowska, I. (2013). "'Orphan' retrogenes in the human genome." *Molecular Biology Evolution* 30. 384-396.
- Costa, F. F. (2012). "Non-coding RNAs, Epigenomics, and Complexity in Human Cells." in Morris, K. V. (Ed.) (2012). *Non-coding RNAs and Epigenetic Regulation of Gene Expression: Drivers of Natural Selection*. Norfolk: Caister Academic Press.
- Crick, F. H. (1968). "The origin of the genetic code." *Journal of Molecular Biology* 38. 367-379.
- Dembski, W. A. (2001). *Signs of Intelligence: Understanding Intelligent Design*. Grand Rapids: Brazos Press.
- Dembski, W. A., & Wells, J. (2008). *The design of life: Discovering signs of intelligence in biological systems*. Dallas: Foundation for Thought and Ethics.
- Djebali, S. et al. (2012). "Landscape of transcription in human cells." *Nature* 489. 101-108.
- Doolittle, W. F., & Sapienza, C. (1980). "Selfish genes, the phenotype paradigm

- and genome evolution.” *Nature* 284. 601–603.
- Ecker, J. R., Bickmore, W. A., & Barroso, I. (2012). “ENCODE explained.” *Nature* 489, 52–54.
- Ehret, C. F., & De Haller, G. (1963). “Origin, development, and maturation of organelles and organelle systems of the cell surface in *Paramecium*.” *Journal of Ultrastructure Research* 6. 3–42.
- Elgar, G., & Vavouri T. (2008). “Tuning in to the signals: noncoding sequence conservation in vertebrate genomes.” *Trends in Genetics* 24. 344–352.
- Germain, P., Ratti, E., & Boem, F. (2014). “Junk or Functional DNA? ENCODE and the Function Controversy.” *Biology and Philosophy* 29. 807–831.
- Gerstein, M. B. et al. (2012). “Architecture of the human regulatory network derived from ENCODE data.” *Nature* 489. 91–100.
- Giannopoulou, E., Bartsakoulia, M., Tafrali, C., Kourakli, A., Poulas, K., Stavrou, E. F., Papachatzopoulou, A., Georgitsi, M., & Patrinos, G. P. (2012). “A Single Nucleotide Polymorphism in the HBBP1 Gene in the Human β -Globin Locus is Associated with a Mild β -Thalassemia Disease Phenotype.” *Hemoglobin* 36. 433–445.
- Gilbert, W. (1986). “Origin of life: the RNA world.” *Nature* 319. 6055.
- Graur, D., Zheng, Y., Price, N., Azevedo, R. B., Zufall, R. A., & Elhaik, E. (2013). “On the immortality of television sets: “function” in the human genome according to the evolution-free gospel of ENCODE.” *Genome Biology and Evolution* 5. 578–590.
- Griffiths, P. E. (2009). “In what sense does ‘nothing in biology make sense except in the light of evolution’?” *Acta Biotheoretica* 57. 11–32.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). “The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme.” *Cell* 35. 849–857.
- Hardman, H. (2004). “Junk’ DNA may be very valuable to embryos.” *Developmental Cell* 7. 597–606.
- Hutchison, C. A. 3rd. et al. (2016). “Design and synthesis of a minimal bacterial genome.” *Science* 351. 6280.
- Ibarra-Laclette, E. et al. (2013). “Architecture and evolution of a minute plant genome.” *Nature* 498. 94–98.

- Karlic', R., Chung, H. R., Lasserre, J., Vlahovicek, K., & Vingron, M. (2010). "Histone modification levels are predictive for gene expression." *Proceedings of the National Academy of Sciences U.S.A.* 107. 2926-2931.
- Kellis, M. et al. (2014). "Defining functional DNA elements in the human genome." *Proceedings of the National Academy of Sciences U.S.A.* 111. 6131-6138.
- Kurihara, K., Okura, Y., Matsuo, M., Toyota, T., Suzuki, K., & Sugawara, T. (2015). "A recursive vesicle-based model protocell with a primitive model cell cycle." *Nature Communications* 6. 8352.
- Liu, Y., & Moran, D. (2006). "Do new functions arise by gene duplication?" *Journal of Creation* 20. 82-89.
- Maher, B. (2012). "ENCODE: The human encyclopaedia." *Nature* 489. 46-48.
- Mattick, J. S., & Dinger, M. E. (2013). "The extent of functionality in the human genome." *The HUGO Journal* 7. 2.
- Mitchell, L. J., Jorge, G. G., Christopher, J. M., Alain, L., & Satya, P. (2014). "Emerging science of the human microbiome." *Gut Microbes* 5. 446-457.
- Meyer, S. C. (2010). *Signature in the Cell: DNA and the Evidence for Intelligent Design*. New York: HarperCollins Publishers.
- Nature Editorial (2013). "Form and function." *Nature* 495. 141-142.
- Neph, S. et al. (2012). "An expansive human regulatory lexicon encoded in transcription factor footprints." *Nature* 489. 83-90.
- Nuonon, M. (2010). "A genome-wide association identified the common genetic variants influence disease severity in beta0-thalassemia/hemoglobin E. Human." *Genetics* 127. 303-314.
- Ohno, S. (1972). "So much "junk" DNA in our genome." *Brookhaven Symposia in Biology* 23. 366-370.
- _____ (1970). *Evolution by gene duplication*. New York: Springer-Verlag.
- Orgel, L. E. (1968). "Evolution of the genetic apparatus." *Journal of Molecular Biology* 38. 381-393.
- Orgel, L. E., & Crick, F. H. (1980). "Selfish DNA: the ultimate parasite." *Nature* 284. 604 - 607.
- Patel, B. H., Percivalle, C., Ritson, D. J., Duffy, C. D., & Sutherland, J. D. (2015). "Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism." *Nature Chemistry* 7. 301 - 307.

- Pennisi, E. (2012). "ENCODE project writes eulogy for junk DNA." *Science* 337. 1160-1161.
- Rigoutsos, I., Huynh, T., Miranda, K., Tsirigos, A., McHardy, A., & Platt, D. (2006). "Short blocks from the noncoding parts of the human genome have instances within nearly all known genes and relate to biological processes." *Proceedings of the National Academy of Sciences U.S.A.* 103. 6605-6610.
- Robertson, M. P., & Joyce, G. F. (2012). "The origins of the RNA world." *Cold Spring Harbor Perspectives in Biology* 4. a003608.
- Roy, P., Bhattacharya, G., Mandal, A., Dasgupta, U. B., Banerjee, D., Chandra, S., & Das, M. (2012). "Influence of BCL11A, HBS1L-MYB, HBBP1 single nucleotide polymorphisms and the HBG2 XmnI polymorphism on Hb F levels." *Hemoglobin* 36. 592-599.
- Sanyal, A., Lajoie, B. R., Jain, G., & Dekker, J. (2012). "The long-range interaction landscape of gene promoters." *Nature* 489. 109-113.
- Soares, M. B., Schon, E., Henderson, A., Karathanasis, S. K., Cate, R., Zeitlin, S., Chirgwin, J., & Efstratiadis, A. (1985). "RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon." *Molecular and Cellular Biology* 5. 2090-2103.
- Stamatoyannopoulos, J. A. (2012). "What does our genome encode?" *Genome Research* 22. 1602-1611.
- Stern, D. A., Carlo, T., Berget, S. M., (1996). "Architectural limits on split genes." *Proceedings of the National Academy of Sciences U.S.A.* 93. 15081-15085.
- Struhl, K. (2007). "Transcriptional noise and the fidelity of initiation by RNA polymerase II." *Nature Structural and Molecular Biology* 14. 103-105.
- Szostak, J. W., Bartel, D. P., & Luisi, P. L. (2001). "Synthesizing life." *Nature* 409. 387-390.
- Szostak, J. (2012). "Part I: The Origin of Cellular Life on Earth." in *The Origin of Life on Earth*. <https://www.youtube.com/watch?v=PqPGOhXoprU>. (검색일 2016.6.10.)
- The ENCODE Project Consortium (2012). "An integrated encyclopedia of DNA elements in the human genome." *Nature* 489. 57-74.
- Timpson, T. (2013). "Debating ENCODE: Dan Graur, Michael Eisen." Mendelspod. <http://mendelspod.com/podcast/debating-encode-with-dan-graur-and-michael-ei>

sen. (검색일 2016.6.10)

- Thurman, R. E. et al. (2012). "The accessible chromatin landscape of the human genome." *Nature* 489. 75-82.
- Vickaryous, M. K. & Hall, B. K. (2006). "Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest." *Biological Reviews* 81. 425-455.
- White, M. A., Myers, C. A., Corbo, J. C., & Cohen, B. A. (2013). "Massively parallel in vivo enhancer assay reveals that highly local features determine the cis-regulatory function of ChIP-seq peaks." *Proceedings of the National Academy of Sciences U.S.A.* 110. 11952-11957.
- Wilson, B. A., & Masel, J. (2011). "Putatively noncoding transcripts show extensive association with ribosomes." *Genome Biology and Evolution* 3. 1245-1252.
- Woese, C. R. (1967). *The genetic code: The molecular basis for genetic expression*. New York: Harper & Row. 186.

Abstract

Review of the evolutionary and design perspectives on the origin of human genome: The conflict between the ‘junk DNA’ concept and the ENCODE results

Ah-Ram Kim (Handong Global University)

Chang-Kee Hyun (Handong Global University)

Approximately 98% of human genome is non-protein-coding DNA. Majority of them has been considered to be junk DNA from the evolutionary perspective. However, the ENCODE project and its related research articles provided evidence that the concept of junk DNA is no longer supported based on the fact that a large portion of them, up to about 80%, shows cell-type specific biochemical functions. In this article, we examined the conflict between the concept of junk DNA and the result of the ENCODE project, and reviewed both the evolutionary and design perspectives on the origin of human genome.

Key Words: human genome, junk DNA, ENCODE, evolution, intelligent design, origin science

