



메타유전체 서열 조립의 문제점에 대한 연구

Issues on haplotype assembly of metagenomes

김우철(Woo-Cheol Kim)¹, 여윤구(Yun-Ku Yeo)², 김종현(Jong-Hyun Kim)³, 박상현(Sang-Hyun Park)⁴

요약

미생물의 경우 전체 생물종 중에서 많은 비율을 차지함에도 불구하고, 대부분의 미생물은 자연상태에서 분리시켜 배양하여 연구하기 힘들다. 최근의 미생물학은 자연상태에 있는 미생물을 직접 연구하려는 시도를 하고 있는데, 이런 연구를 가능하게 하기 위해서는 메타유전체의 염기서열을 밝혀내야 한다. 메타유전체의 염기서열을 밝혀내면, 현재 유전체학(Genomics)에 이용되는 분석방법을 적용할 수 있게 된다. 그러나, 단일 유전체가 아닌 특정 환경에서 살고 있는 다양한 미생물 유전체들의 집합인 메타유전체(Metagenome)의 경우에는 현재 확립된 서열조립 방법론이 없는 실정이다. 유전체의 염기 서열을 밝히는데 일반적으로 널리 쓰이고 있는 방법은 Whole Genome Shotgun Sequencing(WGSS)이다. 기존의 WGSS방식은 염기변이가 낮은 단일 유전체를 서열화(sequencing)하는데 초점을 맞춰서 개발되었기 때문에 염기변이율이 높은 유전체(polymorphic genome)의 서열화에 적용시킬 경우에는 염기 서열의 연속성(continuity)이 심각하게 손상된다. 염기변이로 인한 염기서열 연속성의 손상은 여러 미생물들의 유전체서열이 동시에 조립되는 메타유전체의 경우에는 더욱 심각해 진다. 이와 같이 메타유전체를 대상으로 하는 서열 조립에 대한 후속 연구들이 필요하다. 따라서 본 논문은 이러한 연구들의 기초 연구가 될 수 있는 메타유전체 조립 알고리즘에 대한 연구의 중요성과 메타유전체 서열화 과정에서 발생하는 문제점들의 원인을 분석한다. 이를 바탕으로 후속 연구들이 파생될 수 있는 기초 연구를 제공한다.

주제어 : 서열화, 유전체 조립, 메타유전체, 메타유전체학

이 논문은 2008년도 교육과학기술부의 재원으로 한국과학재단의 지원을 받아 수행된 연구임(2008-2004103)

1 연세대학교 컴퓨터학과 박사과정

2 연세대학교 컴퓨터학과 석사과정

3 연세대학교 컴퓨터학과 연구교수

4 연세대학교 컴퓨터학과 부교수, 교신저자

+ 논문접수 : 2008년 12월 30일, 심사완료 : 2009년 3월 25일