

엽록체계놈연구의 발전

리 영 철

경애하는 최고령도자 김정은동지께서는 다음과 같이 말씀하시였다.

《현대육종의 물질적기초로 되는 유용한 동식물유전자원을 더 많이 수집확보하고 특성평가와 보존을 잘하며 효율적인 리용방법을 적용하여 현대적인 육종연구의 튼튼한 토대를 마련하여야 합니다.》

유용한 유전자원을 효과적으로 리용하는데서 생물의 계놈해석에 기초하는것이 현대적인 육종연구에서 매우 중요하다.

엽록체계놈은 빛합성과 기타 대사과정에 관여하는 많은 주요단백질들을 암호화한다. 고성능배열분석기술의 출현은 엽록체유전학과 계놈학분야의 연구에서 빠른 전진을 가져오게 하였다. 최근까지 300여종의 작물과 키나무들을 포함하여 800여종이상의 엽록체계놈배열이 완전히 해석되어 세포내에서의 유전자이동, 보존, 다양성 그리고 분자유종에 대한 리해를 더 풍부히 하는데 이바지하였다.[1, 2, 54]

론문에서는 주요작물들의 기원과 재배기간의 변화 등을 리해하는데서 엽록체계놈배열에 대한 해석이 가지는 커다란 의의에 대하여 론의하였다.

1. 엽록체계놈배열분석기술의 발전

계놈학분야에서 비약적인 발전을 가져올수 있게 한 중요한 요인은 배열분석기술의 획기적인 개선에 있다. 물론 계놈분석에서 단위당 분석처리능력을 높이는것이 중요하다. 그러나 이에 앞서 중요한것은 우선 엽록체계놈을 분리하고 굴림식복제(Rolling Circle Amplification)로 엽록체의 웅근계놈을 통채로 증폭에 리용하는것[3, 4]이며 다음으로 엽록체계놈배열을 탐침으로 하여 세균인공염색체(BAC)나 포스미드(fosmid)서고를 제작하는것[5, 6]이다.

력사적으로 보면 엽록체계놈배열분석기술에서 제일먼저 쓰인 방법은 탐침법이였는데 정확도가 낮은 본질적인 결함을 비롯하여 여러가지 문제들이 제기되였다. 실례로 BAC나 포스미드서고의 제작에서 높은 질을 보장하는 문제, 엽록체계놈증폭을 위한 많은 PCR과정에 생기는 오류 그리고 다른 세포기관자 DNA에 의한 오염 등과 같은 여러가지 난문제들에 부딪치게 된다. 이밖에도 PCR기술을 재배열빈도가 높은 엽록체계놈이나 련관이 없는 식물종의 엽록체계놈들의 배열분석에 리용하기에는 어렵다는것이다. 그리하여 실천적으로 보다 개선된 분석기술이 요구되였다.

2세대배열분석법인 NGS(Next-Generation Sequencing)는 과학자들에게 있어서 더 빠르고 더 비용이 적게 드는 분석기술로 되였다. NGS는 2006년에 처음 개발되어 남천속(*Nandina*)과 방울나무속(*Platanus*)의 엽록체계놈들을 배열분석하는데 리용되였다.[8] 여기서 다중NGS조작기술이 엽록체계놈의 배열분석에 쓸모가 있게 되는데는 Illumina분석장치가 주요NGS분석장치에 리용되였기때문이다.[7, 9] 그것은 Illumina에서 회전주기증폭산물을 리용할수 있었기때문이다.[19, 37] 그리하여 표본계놈배열이 없이도 새로운 배열의 조립에 생명정보학적조작기술을 리용할수 있었고 또한 이러한 조립물로부터 엽록체계놈의 일치성을 동정할수 있

었다.[10]

3세대 배렬분석방법이 개발되었는데 이것은 단분자실시간(SMRT)배렬분석에 PacBio체계를 적용한 분석기술[13]로서 현재 가장 광범히 쓰이고있다.[11, 13] 이 방법의 우점은 긴 배렬을 읽을수 있고 새로운 계놈을 조립할수 있으며 특히 엽록체계놈에서 거꿀반복배렬(IR)과 단순-코피 영역사이에 위치한 4개의 접합영역을 결정하는데 편리한것이다.

현재 이 PacBio체계로 엽록체계놈배렬을 분석하면 정확도가 최소한 약 85%이지만 여기에 계층계놈조립알고리즘과 함께 마지막단계의 화학처리를 결합하여 오류교정을 하면 최대 99.999%까지 배렬분석의 정확성을 높일수 있다.[12] 실제로 양지꽃과(*Potentilla*)식물의 엽록체계놈을 Illumina분석장치로 배렬분석하였을 때에는 7개의 클론콘티그가 조립된것이므로 정확도가 90.59%였지만 PacBio체계에서는 통채로 하나의 콘티그로 분석되므로 오류를 수정하여 100%의 정확도를 보장하였다.[39]

1986년에 처음으로 담배(*Nicotiana tabacum*)에서 첫 옹근엽록체계놈배렬이 분석[14]되었다. 이때는 PCR에 의한 유전자증폭도 시작에 불과한 때였고 더우기 세균인공염색체도 개발되지 않은 환경에서 고전적인 방법으로 배렬분석을 하였기때문에 해석기일도 오래 걸렸고 정확도도 높지 못하였다. 그리하여 1998년에 담배엽록체계놈배렬에 대한 정확한 수정판이 다시 제기되었다.[15] 담배엽록체계놈분석에 이어 1990년대초에 벼와 땅밥(*Marchantia polymorpha*)엽록체의 전계놈이 분석되었다. 2000년대 중엽까지만 하여도 불과 몇십종의 식물엽록체계놈분석자료밖에 없었지만 최근 10년사이에 폭발적으로 자료들이 축적됨으로써 현재 300여종의 작물과 키나무들의 엽록체계놈을 포함하여 800여종이상 식물들의 엽록체계놈배렬분석자료가 NCBI에 구축되어있다.[21, 54] 이렇게 많은 식물종들의 엽록체계놈배렬자료가 구축되게 된것은 계놈배렬분석기술의 비약적인 발전과 떼어놓고 생각할수 없다.

식물엽록체계놈배렬자료들이 수많이 축적됨으로써 식물의 다양성과 진화, 계통발생학적연구에서 일정한 전진이 이룩[51, 56]되었는데 이것들은 핵, 사립체, 엽록체계놈들사이에 진행되는 기능적인 유전정보의 협동을 이해하는데서와 농업실천에서 가치있는 여러가지 스트레스건달성농작물들을 육종하는데서 매우 중요한 자료로 되었다. 또한 식용작물들에서 왁찐이나 생물약제를 발현시키기 위한 식물생물반응기를 개발하는데서도 큰 도움으로 되었다.[21] 특히 사람혈액단백질생산을 위한 조작이 쉽고 실용성있는 기술공정(cGMP)을 확립하는데서 큰 의의를 가진다.[6, 16]

2. 엽록체계놈의 일반적구조

지상식물엽록체계놈은 고도로 보존된 구조로 조직화되어있다. 엽록체계놈은 큰 단순배렬영역(LSC)과 작은 단순배렬영역(SSC) 그리고 이 영역들을 구분해주는 2반복의 거꿀반복영역(IR)들로 구성되어있는 하나의 고리형분자이다. 그러나 강냉이에서 연구한데 의하면 세포내에 고리형과 함께 일정한 비율로 선형엽록체계놈이 존재한다는것도 알려졌다.[17, 18] 엽록체계놈은 120~130개의 유전자를 가지고있는데 이 유전자들은 대부분 초기 빛합성에 관여하는 효소, 단백질들을 암호화하는 유전자들이거나 DNA전사와 mRNA의 단백질제로의 번역과 관련된 유전자들이다. 최근에 연구한데 의하면 비암호삽입사이영역이 상당히 다양하며 흔히 이 영역에 중요한 조절배렬들이 포함되어있다는것이 밝혀졌다.

보라콩, 팔, 완두와 같은 일부 콩과식물과 소나무류식물들의 엽록체계놈들에서는 거꿀반복영역(IR)이 완전히 결실되었거나 해당 유전자족들이 결실되어 흔적으로만 남아있다. 엽록체계놈에서 이 IR영역은 구조적으로 재배렬이 자주 일어나는 영역이다.

엽록체게놈은 일반적으로 구조적조직화와 유전자구성에서 보존적이지만 그 크기는 사 이배렬들간의 차이로 하여 식물종에 따라 각이하다. 보통 은삼나무(*Cathaya argyrophylla*)의 엽록체게놈의 크기는 107kb로서 고등식물엽록체게놈가운데서 가장 작고 제라니움(*Pelargonium*)의 엽록체게놈은 218kb로서 제일 크다.[54] 엽록체게놈의 크기는 식물종들의 핵 게놈크기에는 관계없다.

지상식물의 엽록체게놈에서 인트론도 유전자들과 같이 일반적으로 보존적인 영역이다. 그러나 여러 식물종들, 실크로 보리(*Hordeum vulgare*), 참대(*Bambusa* sp.), 목감자(*Manihot esculenta*), 커피콩(*Cicer arietinum*) 등에서는 단백질을 암호화하고있는 유전자들에 인트론이 없다. 이 인트론이 없는 유전자들에 해당한 단백질들은 서로 다른 기능을 수행한다. 이러한 단백질들로는 ATP합성효소아단위(atpF), Clp프로테아제(clpP), RNA폴리메라제(rpoC2), 리보솜단백질(rpl2, rps12, rps16)들을 들수 있다. 인트론이 없는 유전자들의 대부분은 비록 다양한 식물종들에 널리 존재하지만 일정한 특이적인 식물집단 혹은 종들에 국한되어있다. 실크로 *clpP*유전자를 보면 한쌍잎식물에서 벼과(*Poaceae*), 두쌍잎식물에서 바늘꽃과(*Onagraceae*)와 물푸레나무과(*Oleaceae*) 그리고 겉씨식물식물에서는 소나무류(*Pinus*)에만 대체로 존재한다.[19]

3. 엽록체게놈배열해석에 의한 식물종들의 다양성과 계통발생연구

보다 높은 분류학적수준(과수준)에서 엽록체게놈의 단백질암호영역과 보존성배열들이 계통발생학적연구와 식물재배연구에 리용될수 있다.[19]

초기의 계통발생학적연구에 부분적인 엽록체DNA배열들이 리용되었다. 엽록체게놈의 다양한 영역들이나 혹은 반복(다중)DNA단편들을 리용하여 계통발생학적분석에서 일정하게 전진이 있었으나 배열정보가 불충분한것으로 하여 서로 연이 가까운 종들을 구별하기는 어려웠다. 그리하여 완전한 엽록체게놈배열이 필요되었다. 또한 일정한 류형의 지상식물엽록체게놈들에서는 구조적으로 현저한 재배열이 일어나므로 식물종내와 종사이에 엽록체게놈배열들을 서로 비교하면 식물종들의 진화관계를 밝힐수 있다.

엽록체게놈에 대한 자료가 농업실천에서 절실히 요구되는것은 경제적인의가 있는 재배종들을 확정하는것과 함께 그것들의 순종들을 결정하기 위해서이다. 엽록체게놈기원의 DNA암호(barcode)는 변이체들의 확정과 육종원천의 보존에 리용할수 있다. 육종의 성과는 유전적적합성에 의하여 결정되는데 여기서 엽록체게놈자료는 유전적으로 편편이 있는 식물들을 확정하는데서 귀중한 자료로 된다.

수수(*Sorghum bicolor*), 보리(*Hordeum*), 강냉이(*Zea mays*), 밀(*Triticum aestivum*), 호밀(*Secale cereale*) 그리고 벼(*Oryza sativa*)들의 엽록체게놈배열들이 발표되었다.[23, 25] 벼는 세계적으로 가장 중요한 작물중의 하나이며 세계적범위에서 사람들의 1차적인 당질원천이다.

벼속 종들은 10개의 계놈형으로 분류되는데 그것은 6개의 2배체(AA, BB, CC, EE, FF, GG)와 4개의 겹2배체(BBCC, CCDD, HHJJ, HHKK)들이다. 재배벼와 야생벼들간의 진화적관계를 밝히기 위한 논쟁은 아직도 계속 진행되고있다.[20] 실크로 AA계놈형은 오스트랄리아의 *Oryza meridionalis*(1년생)와 *Oryza rufipogon*(다년생)의 2개의 종에도 있다.[23] 그럼에도 불구하고 오스트랄리아와 아시아의 야생벼엽록체게놈을 분석한데 의하면 오스트랄리아의 *O. rufipogon*엽록체게놈들은 아시아의 *O. rufipogon*보다 오스트랄리아의 *O. meridionalis*에 보다 유사하다는것이 밝혀졌다.[24] 벼 *Oryza* AA종의 19개의 엽록체게놈을 리용하여 계통수를 작성하였는데 이 자료는 벼작물들의 개량과 보존전략을 세우는데 도움이 될것

이다.[25]

세계적으로 가장 중요한 섬유작물인 목화(*Gossypium hirsutum*)의 첫 엽록체계놈은 2006년에 분석되었다.[22] 2배체목화에는 8개의 계놈집단(A, B, C, D, E, F, G와 K)들이 있다. *Gossypium hirsutum*(고지대에서 재배하는 목화)은 세계적으로 광범히 재배되고있는데 A와 D 계놈집단으로부터 만들어진 겹2배체이다. 현재 22개의 목화종들에 대한 엽록체계놈배열들이 해석되어 목화의 진화와 재배에 대한 정보를 얻는데 이용되고있다.[11, 26, 27] 단순반복배열프라임들을 이용하여 2배체목화의 8개의 계놈집단과 겹2배체들을 포함하여 41종의 목화들을 분석[28]한 결과 4배체(AD)종들은 2개의 새로운 A계놈종들과 *Gossypium herbaceum*, *Gossypium arboreum*으로부터가 아니라 새로운 A계놈을 가지고있던 사멸한 조상종으로부터 기원되었다는것이 밝혀졌다.

란과(Orchidaceae)는 전체 속씨식물의 6~11%를 차지하는 큰 과인데 그 대부분의 식물들은 관상용화초들이다. 화초재배에 이용되는 많은 란초종들은 Epidendroideae아과에 속하며 이 아과의 여러 종들에 대한 엽록체계놈배열들이 분석되었다. 란과식물들에서는 일반적으로 속간교잡이 쉽게 일어나고 품종계열이 명백치 않기때문에 화초로서 관상적가치가 큰 변종(품종)들의 어미아비기원을 정확히 밝히기가 힘들다.[29] 어미아비정보는 육종과 다양성결정에서 중요하다. 란과식물 특히 Epidendroideae아과들에 대한 PCR분석을 진행하고 종수준에서의 계통발생학적관계를 15개의 변종들에서 밝히는데 의하면 보존영역이 8개나 동정되었다. 이것은 어미아비의 기원을 정확히 해명하는데 도움으로 되었다. 어미아비목록에 의하면 치설란속(*Odontoglossum*)의 품종들인 《Violetta von Holm》, 《Margarete Holm》, 《Golden Gate》들은 다 같은 어미로부터 기원된것으로 알려져있었다. 그러나 《Violetta von Holm》의 계통발생학적분석은 《Golden Gate》 또는 《Margarete Holm》들과는 련관이 없다는것을 보여주었다.[29] *Odontoglossum bictoniense*엽록체DNA배열분석에 기초하여 작성한 계통수와 어미아비목록사이에 생긴 불일치는 엽록체에 해당 유전자들이 포획되었기때문이라고 보고있다. 한 종의 엽록체계놈에 다른 종의 엽록체계놈이 삽입되면 속간 또는 속내에서 잡종화가 일어난다. 비록 엽록체계놈배열분석이 밀접히 련관된 집단들에 대한 계통발생학적정보를 제공한다 하여도 삽입된 엽록체계놈이나 혹은 유전자들이 계놈배열에 있다면 잡종화결과 계통발생학적관계를 모호하게 할수 있다.[30] 이 경우 핵과 엽록체계놈을 둘다 분석에 이용하면 보다 완전한 계통발생학적관계를 알수 있다.

엽록체계놈에 대한 정보는 여러 작물들 특히 콩과작물의 재배과정을 리해하는데 크게 도움을 준다.[31] 콩과작물의 엽록체계놈구조는 긴 단편이 회전하여 반대로 결합되던가 역반복배열이 결여되기도 한 다중재배열로 되어있다.[32] 실제로 51kb크기의 긴 영역이 회전되어 반대로 결합되었는데 이러한 구조적변화는 콩(*Glycine max*)엽록체계놈에서 처음 동정되었고 이에 이어 78kb 크기의 회전배열영역이 *Phaseolus*와 *Vigna*엽록체계놈들에서 련이어 확정되었다.[33-35] 보다 최근에 51kb크기의 회전영역에 36kb크기와 5.6kb크기단편이 역회전되었다는것이 동정되었다.[34] 이 역배열들에 대부분 중요한 유전자들이 있지만 어느 한 유전자도 파괴되지 않았을뿐만아니라 식물체의 생존과 기능에도 아무런 영향을 주지 않았다. 이러한 특성들은 계통발생학적연구에서뿐만아니라 콩과식물의 엽록체형질전환에서 중요한 기초자료로 된다. 또한 엽록체계놈구조에 대한 자료는 유전자배열증폭에 필요한 프라이머설계와 더 나아가서 재배와 형태발생학적분석에서 매우 가치가 있다.

귤나무속(*Citrus*)은 경제적가치가 큰 열매나무들중의 하나이다. 2006년에 처음으로 감귤나무(*Citrus sinensis*)엽록체계놈이 분석[12]되었고 이것을 표본계놈으로 하여 련달아 다른 귤나무속 식물종들의 엽록체계놈분석에 성공하였다.[36] *Citrus*(28종)와 *Citrus*관련속(6종)

을 포함하여 총 34종 식물들의 엽록체게놈에 대한 형태발생학적분석에 의하면 굴나무는 다 같은 공통선조를 가지었다.[37] 분석한데 의하면 4개의 유전자들(*matK*, *ndhF*, *ycf1*, *ccsA*)에서 일누클레오티드변이와 삽입/탈락빈도는 다른 유전자들에 비하여 높다는것을 보여주었다. 일반적으로 II형스플라이싱관련인트론과 마투라제유전자인 *matK*는 계통발생학과 진화 연구에 흔히 리용되고있다.[38] *matK*유전자들에 대하여 양성선발한데 의하면 굴나무뿐만 아니라 몇가지 다른 식물종들에서도 공통적으로 나타난다는것을 보여주었다. 실지 30개 이상의 식물집단들을 조사한데 의하면 일반적으로 유전자들이 여러가지 각이한 생태적선발 압을 받게 되므로 *matK*유전자들도 예외없이 양성선발지표로 될수 있다. *ndhF*유전자는 엽록체NAD(P)H데히드로게나제(NDH)복합체의 아단위를 암호화한다. 스트레스에 대한 적응이 *ndhF*유전자들에 관련될수 있다고 보면 엽록체에서 합성되는 NDH단량체들은 높은 빛스트레스에 대한 감수성으로 나타난다.[38, 47] 이러한 연구들을 통하여 오스트랄리아종들에서 *matK*와 *ndhF*는 고온과 건조한 기후에 대한 적응성으로 표현된다는것을 알수 있다.[37, 39]

참대는 경제적으로나 생태적으로 중요한 아시아의 산림식물이다. 참대는 빨리 자라고 가을한 후에 뿌리줄기로부터 재생되므로 경작하기에 유리한 작물이다. 처음으로 2종의 참대엽록체게놈들만이 해석[40]되었지만 현재는 많은 참대엽록체게놈들이 분석되어 여러가지 목적에 리용할수 있게 되었다. 참대는 미성숙때에 길게 자라는 식물인것만큼 분류학적 연구에 필요한 꽃을 얻자면 어렵다. 이것은 전통적인 재생기관형태학적관찰에 근거하여 참대의 분류관계를 해석하는데서 하나의 장애로 되었다. 더우기 배럴분산빈도가 극히 낮고 균일(분류수준이 낮은)한 수림을 이루고있는 참대들의 분류학적, 계통발생학적관계는 해명하기 어렵다.[41] 그러나 이러한 유연관계는 25종의 참대엽록체게놈해석자료에 근거하여 계통발생수를 작성[42, 43]함으로써 명백해졌다. 참대와 함께 초본참대의 엽록체게놈에 대한 해석자료들도 발표되었다.[41] 초본참대의 엽록체게놈해석에서 흥미있는것은 동정된 유전자가 사립체게놈으로부터 엽록체게놈으로 옮겨진 유전자라는것이다. 이것은 엽록체게놈으로는 핵과 사립체게놈으로부터 DNA가 거의나 옮겨지지 않는다는 종전의 가설과는 다른 것이었다.[43] 일반적으로 효과적인 DNA흡수체계가 기관자에 없기때문에 DNA의 엽록체게놈으로의 이동에 대하여 쉽게 인정하려고 하지 않고있다. 초본참대에서 관찰되기 전에 이와 같은 현상은 한종의 쌍자엽(eudicot)식물과 한종의 단자엽식물(monocot) 엽록체게놈에서만 관찰되었다.[41, 43]

4. 핵게놈이나 사립체게놈으로의 엽록체유전자의 이동

식물세포에는 3개의 독자적인 게놈이 있다. 즉 식물은 핵, 엽록체, 사립체들에 독자적인 유전계들을 가지고있다. 사립체는 단일내공생으로 프로테오박테리아가 흡수되어 기원되었다고 보고있다. 이와 반면에 엽록체는 시안세균이 내공생한 후에 엽록체로부터 핵에 많은 유전자들이 이동하였다고 본다.[44] 엽록체나 사립체는 독자적인 번역계를 가지고있다. 핵에 암호화된 유전자들은 세포질에서 번역되고 번역된 단백질들은 엽록체를 포함한 해당 장소에 수송되어 기능을 수행하게 된다. 반면에 엽록체게놈에 암호화된 단백질들은 직접 엽록체안에서 자체로 합성되며 또한 빛합성 또는 단백질합성과 관련되는 여러 아단위로 이루어진 기능단백질복합체들도 엽록체안에서 조립된다.

엽록체게놈배열에서 유전자의 구성, 수, 구조는 자양성지상식물에서 대체로 보존되어 있다.[45, 46] 그러나 일부 단백질유전자들은 특이한 종에서 결핍되었다.[19] 엽록체게놈에서 *infA*, *rpl22*, *ndh*와 같은 유전자들이 없어진것은 그것들이 세포내 핵과 사립체으로 이동

하였기 때문이다.[53, 54] 이것은 계통발생학적분석과 진화적연구에 가치있는 정보로 된다. 식물세포내에서 사립체나 핵계놈으로 이동한 엽록체기원의 유전자들을 매우 쉽게 동정할수 있다. 그러나 유전자의 크기가 너무 작은 배열들에 대한 유연관계를 해석하는데 옹근엽록체계놈을 리용하면 오히려 오류를 범할수 있다.[54]

infA 유전자는 엽록체에서 번역개시인자 1로 알려져있는데 이 유전자는 *Escherichia coli*의 *infA* 유전자와 본질적으로 상동이다.[48, 49] 이 유전자는 mRNA, 리보솜과 tRNA-Met 개시인자들사이에 호상작용을 일으키도록 핵계놈에 암호화된 두종의 개시인자들과 협동하여 번역과정을 시작한다.[49] 엽록체계놈에 암호화되었던 *infA* 유전자는 속씨식물인 경우에 진화과정에 결실되었다.[49] 핵계놈에 암호화된 *infA* 유전자들은 애기장대(*Arabidopsis thaliana*), 콩, 도마도(*Solanum lycopersicum*) 그리고 나무채송화속의 식물(*Mesembryanthemum crystallinum*)들에서 동정되었다.[49] 이 4종의 식물들에서 *infA*의 발현산물인 해당 단백질사슬들에는 엽록체으로의 운반을 담당한 수송펩티드가 결합되어있다. 콩과 *A. thaliana*의 *infA*-GFP단백질을 리용한 연구를 통하여 핵계놈의 *infA* 유전자들이 세포줄에서 번역되어 그산물이 엽록체으로 수송된다는것을 보여주었다.[49] 보다 많은 식물체들에서 엽록체계놈의 *infA* 유전자들이 결실되었다는것을 최근에 동정하였다.

26개 속에 달하는 57개 엽록체계놈에서 *rpl22* 유전자는 엽록체에서 결실되어 핵계놈으로 이동되었다.[50] 핵계놈의 *rpl22* 유전자는 세포줄로부터 엽록체으로 해당 단백질의 운반과 관련된 펩티드도 암호화한다. 이 펩티드들은 콩과나 참나무과에서 2개의 독자적인 *rpl22* 옹김인자가 있다는것을 시사하였는데 펩티드들은 서로 다르다. 핵으로의 이와 류사한 이동은 엽록체계놈으로부터 *rpl32*의 결실을 통하여서도 알수 있다.[52]

11개의 엽록체유전자가 빛합성에 관여하는 NDH의 아단위들을 암호화하는데 NDH단백질복합체들은 순환식전자전달계에 관여하는 빛체계 I 복합체를 조립[55, 56]하며 엽록체 숨쉬기를 촉진한다. 일부 제형양식물들의 엽록체계놈들에서 *ndh*가 결실되었으며 현재까지 밝혀진 수십개의 유전자들 즉 *accD*, *ycf1*, *ycf2*, *ycf4*, *psaI*, *rpoA*, *rpl20*, *rpl23*, *rpl33* 등과 같은 엽록체유전자들이 결실되어 핵계놈이나 사립체계놈으로 이동하였다. *psbJ*, *rps2*, *rps14*, *rps19*들이 1개 또는 몇종의 식물들에서 결실되었다. *ycf1*과 *ycf2* 유전자를 제외하고 엽록체계놈에서 결실된 많은 필수적인 유전자들이 핵으로 전이되어 식물의 빛합성능력을 유지하였다.

총체적으로 엽록체계놈배열들은 식물의 진화와 계통발생에 대한 리해에서 매우 가치가 있다. 또한 식물계놈뿐만 아니라 식물의 전사단위에 대한 자료기지들은 어떻게 혹은 언제 엽록체계놈으로부터 다른 기관자계놈으로 유전자들이 이동하는가에 대한 연구에 결실히 필요할것이다.

참 고 문 헌

- [1] P. Wambugu et al.; Sci. Rep., 5, 13957, 2015.
- [2] M. Brozynska et al.; Plant Biotechnol. J., 14, 1070, 2016.
- [3] R. K. Jansen et al.; Mol. Phylogenet. Evol., 48, 1204, 2008.
- [4] M. G. Bausher et al.; BMC Plant Biol., 6, 21, 2006.
- [5] H. Daniell et al.; Theor. Appl. Genet., 112, 1503, 2006.
- [6] H. Daniell et al.; Theor. Appl. Genet., 116, 723, 2008.
- [7] J. Wu et al.; Frontiers Plant Sci., 3, 234, 2012.
- [8] M. J. Moore et al.; BMC Plant Biol., 6, 17, 2006.

- [9] I. C. Pan et al.; PLoS One, 7, e34738, 2012.
- [10] R. A. Atherton et al.; Plant Methods, 6, 22, 2010.
- [11] X. Chen et al.; Frontiers Plant Sci., 6, 42, 2015.
- [12] C. S. Chin et al.; Nat. Methods, 10, 563, 2013.
- [13] M. Ferrarini et al.; BMC Genomics, 14, 670, 2013.
- [14] K. Shinozaki et al.; EMBO J., 5, 2043, 1986.
- [15] T. Wakasugi et al.; Plant Molecular Biology Reporter, 16, 231, 1998.
- [16] J. Su et al.; Biomaterials, 70, 84, 2015.
- [17] D. J. Oldenburg et al.; Frontiers Plant Sci., 6, 883, 2015.
- [18] D. J. Oldenburg et al.; Curr. Genet., 62, 431, 2016.
- [19] R. K. Jansen et al.; Proc. Natl. Acad. Sci. USA, 104, 19369, 2007.
- [20] P. Wambugu et al.; Sci. Rep., 5, 13957, 2015.
- [21] M. Adem et al.; Plant Methods, 13, 30, 2017.
- [22] S. B. Lee et al.; BMC Genomics, 7, 61, 2006.
- [23] M. Brozynska et al.; Tropical Plant Biol., 7, 111, 2014.
- [24] D. L. Waters et al.; Ecol. Evol., 2, 211, 2012.
- [25] M. Sotowa et al.; Rice, 6, 26, 2013.
- [26] Q. Xu et al.; PLoS One, 7, e37128, 2012.
- [27] R. I. Ibrahim et al.; Genes Genet. Syst., 81, 311, 2006.
- [28] P. Li et al.; Genet. Resour. Crop Evol., 61, 107, 2014.
- [29] F. H. Wu et al.; BMC Plant Biol., 10, 68, 2010.
- [30] D. E. Soltis et al.; Evolution, 49, 727, 1995.
- [31] B. Pickersgill et al.; Theor. Appl. Genet., 110, 432, 2005.
- [32] J. D. Palmer et al.; Curr. Genet., 14, 65, 1988.
- [33] X. Guo et al.; BMC Genomics, 8, 228, 2007.
- [34] S. Tangphatsornruang et al.; DNA Res., 17, 11, 2010.
- [35] M. G. Bausher et al.; BMC Plant Biol., 6, 21, 2006.
- [36] H. J. Su et al.; PLoS One, 9, e113049, 2014.
- [37] J. Caspermeier et al.; Mol. Biol. Evol., 32, 2217, 2015.
- [38] S. L. Chen et al.; J. Plant Res., 123, 241, 2010.
- [39] J. Carbonell-Caballero et al.; Mol. Biol. Evol., 32, 2015, 2015.
- [40] F. H. Wu et al.; Tree Physiol., 29, 847, 2009.
- [41] P. F. Ma et al.; Sci. Rep., 5, 11608, 2015.
- [42] P. F. Ma et al.; Syst. Biol., 63, 933, 2014.
- [43] D. R. Smith et al.; New Phytol., 202, 736, 2014.
- [44] J. N. Timmis et al.; Nat. Rev. Genet., 5, 123, 2004.
- [45] C. F. Barrett et al.; Mol. Biol. Evol., 31, 3095, 2014.
- [46] N. Tiller et al.; Mol. Plant., 7, 1105, 2014.
- [47] C. S. Lin et al.; Sci. Rep., 5, 9040, 2015.
- [48] H. S. Cummings et al.; J. Bacteriol., 176, 198, 1994.
- [49] R. S. Millen et al.; Plant Cell, 13, 645, 2001.
- [50] J. S. Gantt et al.; EMBO J., 10, 3073, 1991.

- [51] G. Peltier et al.; Annu. Rev. Plant Biol., 53, 523, 2002.
- [52] B. P. Cusack et al.; Trends Genet., 23, 270, 2007.
- [53] R. K. Jansen et al.; Mol. Biol. Evol., 28, 835, 2011.
- [54] H. Daniell et al.; Genome Biology, 17, 134, 2016.
- [55] M. Ueda et al.; Plant J., 72, 683, 2012.
- [56] Y. Munekage et al.; Nature, 429, 579, 2004.

주제 109(2020)년 1월 5일 원고접수

Progress of Chloroplast Genome Research

Ri Yong Chol

The chloroplast genome encodes many key proteins that are involved in photosynthesis and other metabolic processes. The advent of high-throughput sequencing technologies has facilitated a rapid progress in the field of chloroplast genetics and genomics. The chloroplast genome was sequenced on over 800 species of complete chloroplast genome sequences, including over 300 species of crop and tree genomes.

We discuss the impact of chloroplast genome sequences in understanding the origins of cultivated species economically important and changes that have taken place during domestication.

Keywords: chloroplast genome, high-throughput sequencing technologies, diversity, conservation