초파리 상동 재조합 유전자 RECQE에 관한 연구

정상민/다카라코리아바이오메디칼(주)

우리는 초파리에서 새로운 형태의 대장균 *recQ* 상동유전자인 RECQE를 분리하였다. *Recqe*유전자는 4개의 intron과 5개의 exon으로 구성되며 분자량은 12만 Da으로 추정되는 1058 아미노산의 단백질을 코딩한다. RECQE 단백질은 7개의 helicase 효소 motive를 가지고 있다. Helicase 부위는 대장균의 RecQ DNA helicase와 42% 동일하며 사람의 RecQL5와 선충의 E03A3.2.과 가장 가깝다. RECQE 단백질의 C-말단은 특이한 구조를 가지며 RecQ superfamily중에서 가장 길다. 우리는 RECQE 단백질이 DNA helicase 효소 활성이 있으며 C-말단이 그 효소 기능의 유지에 불가결하다는 사실을 확인하였다. *RECQE* mRNA는 난모세포에 축적되며 초기배에서 고발현 되었다. 우리는 RecQ의 상동유전자의 발현이 배의 발생 과정에서 조절됨을 확인하였고, 이러한 결과는 RECQE의 DNA helicase 활성이 초기배의 DNA대사 과정에 관여함을 시사하는 것이다.

서론

RecQ 상동유전자는 다세포 생물에 있어서 중요한 역할을 한다. 사람에 있어서 Bloom's와 Werner's 증후군과 같은 유전 질병을 유발하는 원인 유전자가 대장균의 RecQ 상동유전자임이 알려졌다. Bloom's 증후군은 유전적 질환를 보이는데, 주된 임상적 징후는 일광에 민감한 작은 홍반이 얼굴에 나타나며 면역결핍증세를 보인다. Werner's 증후군은 상염색체 열성의 질환이며 드물게 암을 수반하는 조로증 등의 증세를 보인다. 사람의 또 다른 RecQ의 상동유전자인 *RecQL4*에 변이가 생기면 Rothmund-Thompson 증후군과 유사한 증세를 유발한다. Rothmund-Thompson 증후군은 희귀한 상염색체 이상의 유전성 질환이며 피부와 골격 이상, 소년 백내장, 조로증과 종양형성이 쉬워지는 경향을 보인다. RecQ 상동유전자에서 나타난 돌연변이는 여러 가지 표현형을 나타내며 그런 현상은 불완전한 상동재조합(recombination)과 수복(repair), 노화, 게놈의 불안정, 암의 고발병율 등을 포함한다. 최근까지 최소한 사람에 있어서 *RecQL, Blm, Wrn, RecQL4, RecQL5* 등 5개의 RecQ 상동유전자가 알려졌다. 또한 선충의 게놈 프로젝트의 성과로 적어도 4개의 가능한 RecQ 상동유전자가 밝혀졌다. 일반적으로 원핵생물과 단세포 진핵생물이 단 하나의 RecQ 상동유전자를 갖고, 다세포 생물이 복수의 RecQ 상동유전자를 갖는 것에서 알 수 있듯이 하나의 생물체에서 발현되는 RecQ 상동유전자의 수는 바로 생물체의 복잡성과 연관된다. 모든 RecQ 상동유전자들은 helicase 부위를 보유하고 있다. DNA helicase 활성은 대장균, 효모, 사람에서 다수 알려져 있다. 다수의 RecQ 상동체 들의 기능은 다양하지 않지만 그것이 발현되는 기관이나 발생과정의 시기 또는 세포의 형태 등에 따라 제한적인 기능을 갖는다. 사람의 RecQ 상동유전자의 superfamily 중에서 그 돌연변이가 질병에 관련된 Blm, Wrn, *RecQL4* 등 3개의 유전자 산물은 helicase 부위 상류에 긴 N-말단이 있으며 특정한 조직에서만 발현되는데 반하여 짧은 N-말단이 갖는 *RecQL*과 *RecQL5*는 보편적으로 발현한다. 아직까지 *RecQL*과 *RecQL5*의 변이에 있어서는 유전적 질환이 알려지지 않았으며, 짧은 N-말단의 기능에 대해서는 잘 알려지지 않았다.

우리는 짧은 N-말단을 가진 초파리의 RecQ 상동유전자 *Recqe*의 분리와 동정에 관하여 기술하고자 한다. RECQE helicase에는 하나의 짧은 N-말단, 긴 C-말단과 7개의 helicase 부위가 있다. 우리는 초기배에서 RecQ 상동유전자의 발현이 발생과정에서 조절되는 것을 처음으로 알아 내었다. 또한 다른 RecQ superfamily와의 염기서열 비교에서 사람의 RecQL5와 선충의 E03A3.2가 새로운 그룹을 형성함을 알았다.

결론

초파리 RecQ상동유전자 cDNA의 단리와 그 구조

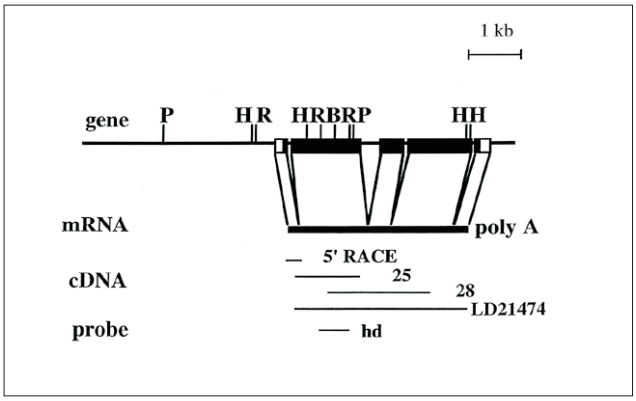


그림 1 초파리 *Recqe* 유전자의 게놈 구조
Recqe 유전자 와 cDNA 구조. 검은 선은 게놈 DNA, mRNA, cDNA clone, *Recqe* clone을 분리하기 위한 probe를 나타낸다. 검은 굵은 막대는 *Recqe* 유전자의 코딩(암호화)영역을 표시한다. 흰 부분은 5′ 또는 3′ non-coding 영역을 표시한다. 제한효소 부위: H, *Hind* II; R, *Eco*R I; P, *Pst* I; B, *Bcl* I. probe1은 library의 스크리닝에, probe2는 northern blot 및 *in situ* 해석에 쓰인 부위를 나타낸다.

우리는 처음에는 사람의 ATPaseQ1(RecQL) DNA단편을 이용하여 보합법으로 genomic DNA library를 검색하였다. 그러나 결국 찾을 수 없었고 다른 생물의 RecQ 상동유전자의 helicase 부위와 유사한 두 곳의 degenerated primer를 합성하여 PCR을 행한 후, 다행히도 증폭된 DNA단편은 cloning 후 염기서열을 분석하였다. 그중에 하나의 clone이 0.6 kb의 intron이 없는 단편임을 알았고, 다른 RecQ superfamily와

도 유의성있는 상동성을 보여 초파리의 게놈과 cDNA library를 그 단편을 probe로 하여 검색하였다 (그림 1). 2개의 큰cDNA clone을 분리하여 염기서열도 분석하였다. 5'-말단은 5'-RACE법으로 찾아낸 223염기의 5'-말단은(그림 1) 개시코돈을 포함하고 있었다. 이러한 염기서열을 근거로 우리는 초파리 국제 게놈해독 사업에서 나온 EST 데이터베이스에서 polyA tail을 갖는 동일한 overlapping clone LD21474를 발견하였다 (그림 1). 그 LD21474와 단리된 clone을 비교하여 접합이나 제한효소 부위 등을 밝혔다.

게놈 library에서 원하는 clone을 단리한 후 분석한 염기서열을 근거로 다른 상동유전자들과 염기 서열을 비교한 결과, 그 clone은 4개의 짧은 intron과 5개의 exon으로 구성되어 있었다. 3174염기의 ORF는 초파리의 번역개시에 필요한 consensus와 흡사한 서열과 개시코돈이 존재하였다. 또한 두개의 종결 코돈이 개시코돈의 상류에 존재하는 것으로 보아 가능한 개시코돈이 ORF끝 종결 코돈의 상류에 존재함을 유추할 수 있었다.

한편, *in situ* hybridization법을 이용하여 3번 염색체의 70D6-E2영역에 존재함을 밝혔다. 70E1-E2영역을 포함하는 P1 clone을 이용하여 PCR 해석 후, 실제로 PCR로 증폭한 단편의 제한효소 부위와 probe로 쓰인 0.6 kb단편과 비교함으로써 염색체상에 하나의 유전자로 존재함을 알아냈다.

예상된 아미노산 서열에 7개의 helicase motif

예상된 ORF는 1058개의 아미노산의 polypeptide로 구성되었으며 분자량은 12만 Da이다. BLAST 검색에서 그 아미노산 서열은 신규임을 알았다. 사람의 RecQL5와 선충의 E03A3.2와 가장 흡사하였으나 다른 원핵 및 진핵생물의 RecQ family와도 유사하였다. 그리하여 *RecQ*의 elongated 또는 새로운 그룹 E라는 의미를 주어 Recqe로 명명하였다. 그 C-말단의 서열은 어떤 단백질과도 분명한 유사성은 없었다. N-말단쪽에 보이는 7개의 helicase motive는 보편적인 DNA/RNA의 그 것과 같았다. Motif II의 DEAH서열을 보면 DExH family에 속한다는 것을 알 수 있다. Motif I, Ia, II, III, IV, V, VI는 대장균의 RecQ 단백질과 각각 61, 50, 57, 53, 33, 53, 72%가 동일하였다. 전체적으로 helicase 부위는 RecQ와 42%의 상동성을 갖는다. 7개의 motif 외에 Ia-II, II-III, III-IV, IV-V사이의 4곳에서 RecQ와 그 superfamily와 같이 짧은 보존영역이 있었다.

RECQE는 새로운 RecQ superfamily

RECQE와 RecQ superfamily사이에 연속적인 7개의 motive가 발견되어, 계통발생학적 해석을 하였다. RECQE는 대장균의 RecQ, 효모의 RecQ, Wrn RecQ, Blm RecQ, RecQL과는 거리가 있으나 사람의 RecQL5, 선충의 동정되지는 않은 E03A3.2와 높은 연관성을 보였다. 이러한 결과는 RECQE, RecQL5, E03A3.2가 이미 동정된 다른 RecQ helicase들과는 독립적으로 진화되었음을 시사한다. 그들 helicase 부위 중에서 RECQE는 여타의 RecQ superfamily에 32-42% 상동적인 것에 반하여 RecQL5와 E03A3.2에 각각 54%, 48% 상동성이 있다 (그림 2). RECQE, RecQL5, E03A3.2는 motive III-IV 사이에 약 10 아미노산 정도 더 긴 간격이 있다. RECQE의 C-말단은 다른 어떤 RecQ superfamily보다도 길며, 짧은 N-말단과 더불어 다른 고등생물의 RecQ 상동유전자에서도 볼 수 없는 독특한 구조를 하고 있다 (그림 2).

초기배에서의 Recqe의 발현

우리는 northern blot을 통해 초파리 *Recqe*의 발현을 알아보았다. Recqe cDNA의 크기에 상응하는 3.4 kb의 하나의 전사 산물이 1.2 kb의 *Cla* I-*Xho* I DNA 단편으로 작성된 probe로 검출되었다 (그림 3). *RECQE* mRNA는 0-3h AEL (산란후 시간)의 배아에서 가장 발현이 높았고, 그보다는 다소 낮지만 3-6h, 6-9h의 AEL에서도 검출되었다. 아주 적은 양의 발현이 후기 배아, larvae, purpae와 성충에서 검출되었다. 배나 난모세포 발생과정에서 RECQE mRNA의 분포를 배아나 미성숙란에 antisense probe를 쓴 *in situ* hybridization으로 조사하였다 (그림 4). Antisense probe로 *RECQE* mRNA를 조사해보면 다핵 포배엽기의 배아에 균일하게 분포하며 세포배엽의 주변 (그림 4A, C)에 국재하는 것을 알 수 있으며 sense probe로는 검출되지 않았다 (그림 4B, D). *Recqe*의 전체적인 발현 수준은 오래된 배아일수록 감소되었다 (그림 4E). 난모세포 발생 중의 *RECQE* mRNA는 초기 10기의 egg chamber의 nurse cell에 주로 분포하며 소포세포에는 분포하지 않았다 (그림 4G). 또한 11-12기 egg chamber의 난모세포에서도 축적되었다 (그림 4H). 난소와 초기배의 *RECQE* mRNA는 northern blot과 같은 결과를 얻었다 (그림 3). 이러한 결과는 RecQ 상동유전자가 발생학적으로

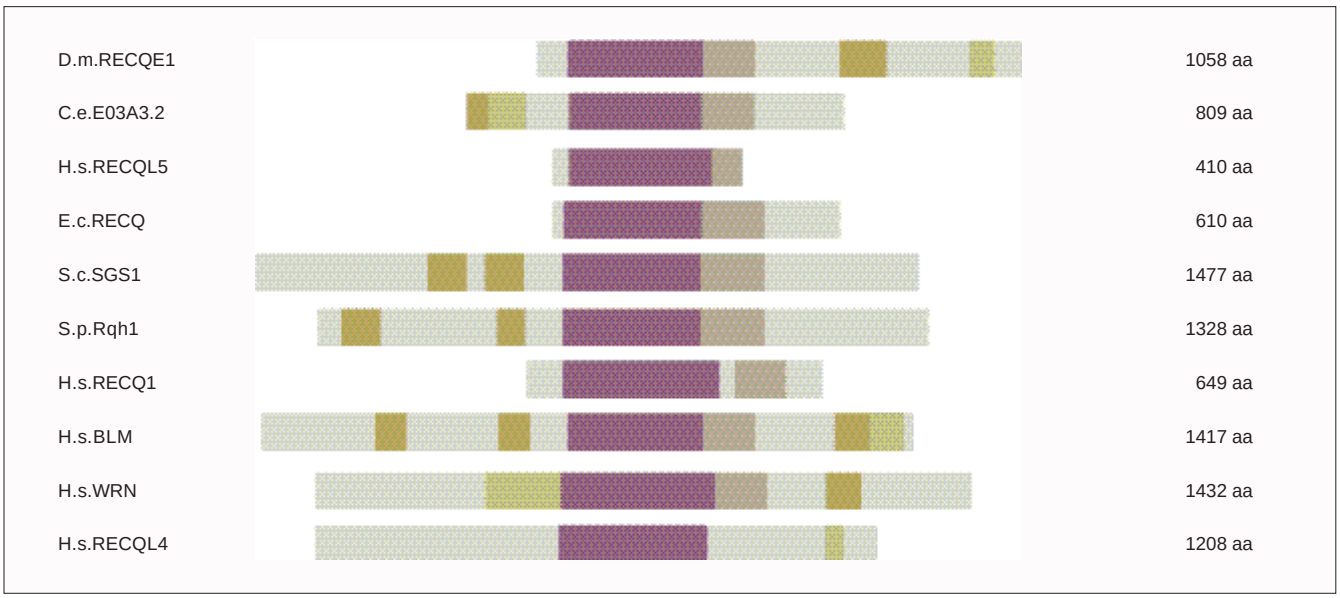


그림 2 RecQ superfamily 1차 구조의 개략
자주색 부분이 비교에 쓰인 helicase 영역이다. 초파리 RECQE와 비교된 각 helicase 부위의 상동성을 나타내는 수치(%)는 막대 안에 표시하였다. 유전자 산물과 생물의 이름은 왼쪽에 아미노산의 길이는 오른쪽에 표시하였다. 산성 영역은 회색, 염기성 영역은 사선으로 표시하였다.

초기배에서 발현 조절됨을 보여주는 것이다.

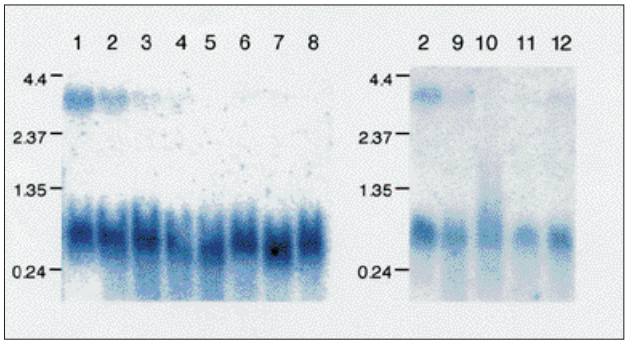


그림 3 초파리의 발생 과정에서의 *RECQE* mRNA 변화
Total RNA는 각 발생과정에서 준비되었다. 20 ug의 RNA를 1.3%의 formaldehyde 변성 아가로스 겔로 분획한 후 나일론막에 전사하여 ^{32}P 표식된 probe로 보합을 실시하였다. RNA marker는 그림의 각 왼쪽에 표시하였다. RNA blot은 1.2 kb의 DNA 단편으로 보합을 실시하였다. 그림 윗부분이 *RECQE*이며, 아래는 *rp49*이다. Lane 1, 난소; Lane 2-8, 배아: 2, 0-3h AEL; 3, 3-6h AEL; 4, 6-9h AEL; 5, 9-12h AEL; 6, 12-15h AEL; 7, 15-18h AEL; 8, 18-21h AEL; Lane 9, 구더기; Lane 10, 번데기; Lane 11, 성충 수컷; Lane 12, 성충 암컷.

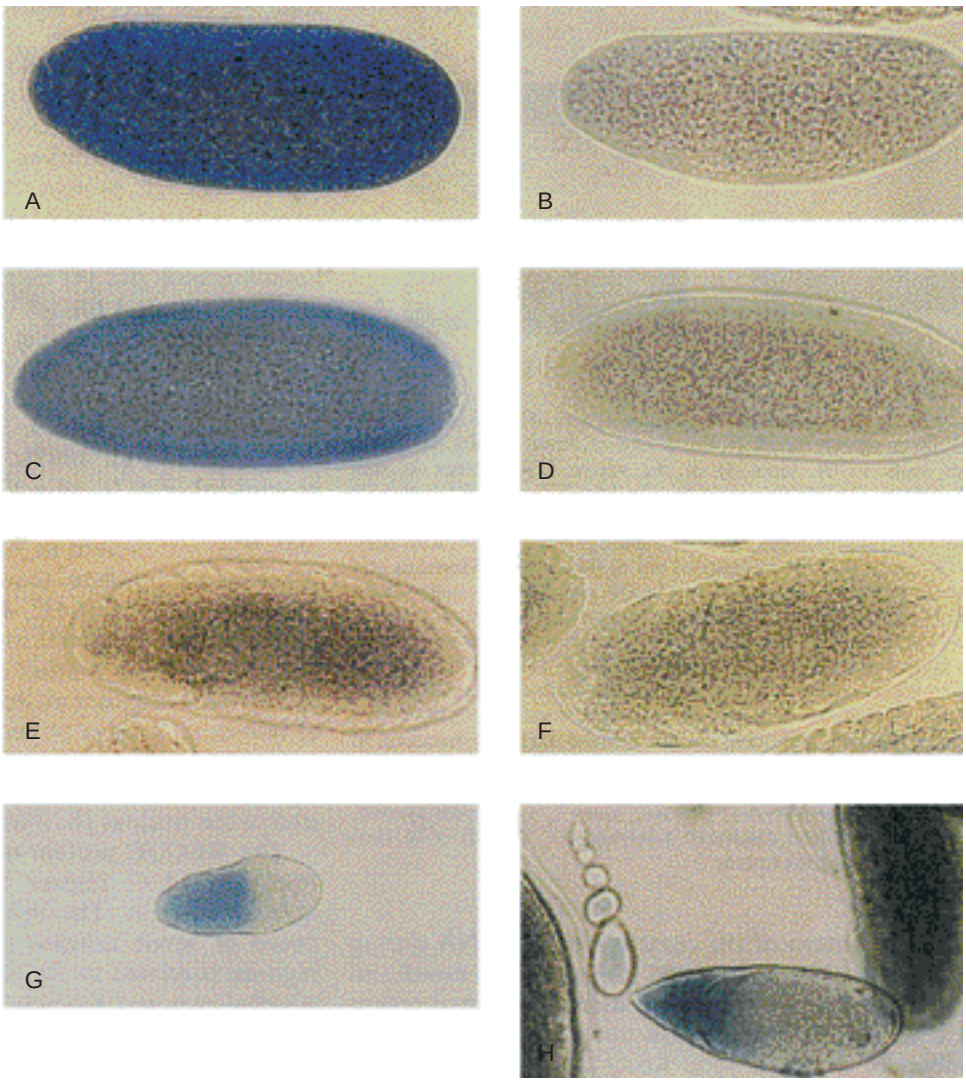
RECQE 단백질은 DNA helicase이다.

Recqe cDNA로부터 *in vitro*에서 단백질을 발현시키기 위해 토끼의 망상적혈구 세포안에서 T7 promoter 하류에 그 cDNA를 놓고 T7 RNA 중합효소를 사용하여 전사하였다. *In vitro* 번역 시스템안의 표식된 단백질 산물은 SDS-PAGE후에 autoradiography 법으로 검출하였다. 단백질은 약 120 kDa정도의 위치에서 확인할 수 있었으며 Recqe 유전자에서 예상된 ORF 크기와도 일치하였다 (그림 5A). 또한 약 120 kDa 정도의 위치에서, helicase 부위 특이적인 항체를 affinity-정제를 한 후 사용하여 RECQE 단백질을 검출할 수 있었다 (그림 5B). 이러한 결과는 Recqe cDNA가 결국 기대되는 크기의 아미노산을 코드하고 있음을 시사하는 것이다. Recqe cDNA에서 유래한 *in vitro* 번역 단백질이 ATP 존재 하에서 표식된 올리고머와 그에 상보적인 환상의 단쇄 DNA의 복합체를 분리해내는 활성이 있음을 확인하였다 (그림 5C). Recqe cDNA 전체는 물론이고 N-말단의 helicase 부위만을 갖는 부분 cDNA로도 활성이 있음이 확인되었다 (그림 5C). Baculovirus system에서 발현 후 정제된 RECQE 단백질도 또한 DNA helicase 활성을 나타내었다. 이들 결과는 RECQE 단백질과 그 helicase 부위가 실제로 DNA helicase 활성을 갖고 있음을 보여준 것이다.

고찰

우리는 초파리로부터 대장균의 RecQ 상동체를 코드하는 cDNA clone을 분리하였다. 그리고 발생기의 mRNA 변화의 양상을 시험하였다. 아미노산 서열은 신규의 것이었고, helicase 부위를 갖고 있으며 다른 RecQ 상동체가 갖지 않는 짧은 N-말단, 긴 C-말단을 갖는 것이었다. 사람에 있어서 RecQ 상동체가 다수 발견되었고 초파리와 선충에 있

그림 4 배아와 난형성기의 RECQE 전사산물의 국재
모아진 배아는 고정한 후, dig-oxigenin 표식된 antisense (A, C, E) probe와 sense probe (B, D, F)로 보합을 하였다. A, B 초기 (2기) 배아. C, D 초기 (5기) 배아. E, F 후기 (12-14기) 배아. G, 10기의 egg chamber. H 11-12기의 egg chamber.



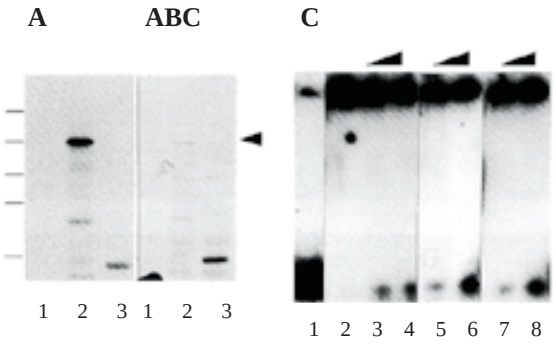


그림 5 Recqe cDNA의 *in vitro* 전사 및 번역
T7 전사/번역계를 ^{35}S 메티오닌 (A) 또는 비방사성 메티오닌 (B)으로 보온하였다. 전 Recqe cDNA (lane2), 5' 부분 Recqe cDNA (1-584 아미노산, lane 3), lane 1은 vector DNA. (A) SDS-PAGE 후의 autoradiography, (B)는 항체 반응 해석. 검은 화살표는 RECQE 단백질을 분자량 marker는 왼쪽에 각각 158, 116, 97, 66.4 kDa을 표시한다. C; Recqe cDNA로 *in vitro* 전사 및 번역을 한 산물의 helicase 활성. *in vitro* 번역물 2 ㎕를 lane 3, 5, 7에, 4 ㎕ 를 lane 4, 6, 8에 사용하였다. 그 *in vitro* 번역 산물의 분석 결과는 vector 유래가 lane 3,4이고 전체 Recqe cDNA가 lane 5와 6, 1-584 aa의 Recqe cDNA는 lane 7과 8이다. Lane 1은 열변성한 기질, 산물 없는 *in vitro* 번역물이 부가된 것은 lane 2이다.

어서는 적어도 두 종류(RECQE, DmBlm)가 알려진 셈이다. DmBlm은 긴 N-말단을 갖고 있는데 반하여 RECQE는 짧은 것을 갖고 있다. RecQ 상동체 중에서 N-말단의 길이는 서로 다르며 이는 각자의 특성을 결정짓는 요소라고 생각된다. Blm, Wrn, RecQL4, 효모의 RecQ 등은 모두 긴 N-말단을 가지고 있다. 긴 N-말단은 특이적인 기능을 갖고 있다. 효모의 SGS1과 사람 Wrn의 N-말단은 각각 TopIII 결합 영역과 exonuclease 부위를 갖고 있다. 사람의 RecQ 상동체의 긴 N-말단은 조직특이적인 발현 양상을 보인다. 그 영역에 변이가 일어날 경우, Bloom, Werner, Rothmund-Thomson 증후군과 같은 병의 유발과 관련이 있다. 그와 달리, 짧은 N-말단은 어떠한 유전 질병과도 관계없으며 모든 조직에서 보편적으로 발현한다. 또한 RecQL5의 발현은 세포주기 동안 일정하게 발현이 된다. RECQE는 RecQL5와 가장 유사하며 짧은 N-말단을 가지고 있다. RecQL5와 같이 RECQE의 발현은 분화된 세포에 특이적이지는 않으며 초기 배아에 제한되어 있다. RECQE 단백질은 가장 긴 C-말단을 갖고 있다. 일반적으로 C-말단의 기능은 아직 확실하지 않다. 효모 SGS1의 C-말단은 top3를 배경으로 한 조건하에서 특정하지 않은 기능을 하는 데 필수적이다. Wrn의 경우에는 C-말단의 변이가 질병으로 고생하는 환자들에게서 공통적으로 발견되며 일부이지만 핵으로의 이행에 손상을 주는 것으로 알려져 있다. RECQE 단백질은 중간 부분에 산성을 띠는 부위가 있으며 C-말단에 염기성 영역이 존재한다. 단백질의 핵이행에 필요한 후보 서열은 비교적 C-말단에 존재한다. 실제로 baculovirus를 쓴 발현계에서는 RECQE 단백질이 핵으로 이행되는 것을 확인하였다.

RECQE 단백질의 helicase와 N-말단, 또는 C-말단 부위를 다른 RecQ family와 계통적으로 비교해 보면 결론적으로 RECQE, RecQL5, E03A3.2와 함께 새로운 family로 분류되어야 함을 알 수 있었다. 최근에 알려진 초파리의 또 다른 RecQ 상동체인 Dmblm과 함께 RECQE는 사람의 BLM과 *RecQL5*에 다른 어떤 상동체보다 유사하다는 것을 알 수 잇었다. 아마도 *RecQL5*의 전사산물 중에서 RECQE와 같은 긴 C-말단을 갖는 것이 존재할지 흥미로운 일이다. RecQ family 전체적으로 잘 보존되어 있는 사실로 보아 같은 특이적 조상으로부터 진화된 것이 시사되지만 기능적 특성에 근거한 RECQE를 분류할 필요가 있다.

*RECQE*의 mRNA를 배아나 난소에서 많이 볼 수 있는데 반해 척추동물의 RecQ 상동체에서는 그러한 발현의 예가 없다. 그 RECQE의 mRNA는 배아에서 주로 사용되기 위해 난소에 축적되는 것으로 추측된다. 초파리에 있어서 배아 유사분열의 13기까지 접합발생에 있어서 모성제어(maternal control)를 받는다. 접합게놈의 전사가 간기14기까지 요구되지 않는 것은, 아마도 RECQE의 전사산물이 초기 배아의 빠른 핵분열에 대비하기 위함일 것이다. RECQE의 전사산물은 다핵포배엽의 어느 특정영역에 국재하지 않고 세포배엽기 동안에 주변에 머무르다가 세포형성기에 극적으로 발현이 줄어드는 것을 알 수 있었다. 이러한 결과는 Recqe 유전자가 제5기의 배아에서 접합체적 mRNA로서 일부이긴 하지만 발현되었다는 것을 시사한다.

초파리 게놈프로젝트에서 알려진 EST clone (GH01404)는 약 3.7 kb의 cDNA는 54 kDa의 단백질을 만드는데 우리가 제작한 54 kDa에 대한 항체로는 초파리 추출물에서 반응하는 단백질을 찾지 못하였으나 120 kDa 영역의 단백질은 검출되었다. Recqe 유전자의 mapping에서 70D6-E2, P1 clone에 의한 PCR 해석에서 70E1-E2의 위치를 알아냈지만 아직 감수분열이나 상동재조합 관련의 어떠한 변이체도 보고 되지 않았다. 단 몇 종류의 치사 돌연변이체는 보고 된 바가 있다. 초파리는 초기의 13기까지는 세포분열 없이 배의 분할이 이루어지고 다핵체의 배아를 형성한다. 13기까지의 빠른 동시다발적인 핵분열은 거의 G1, G2기가 없이 M기와 S기로 구성되어 있다. 그 13기 중에서 초기 10기는 8-9분에 한번 꼴로 분열이 일어나고 10기 이후에 세포주기가 감속 된다. 10기 이전의 초파리의 배아에 aphidicolin을 주입하면 DNA 합성과 핵의 복제가 방해된다. 이러한 상황 하에서 염색질의 응축/이완, 핵막의 파괴/형성, 중심체의 복제 등과 같은 세포주기의 여러 가지 양상을 관찰할 수 있다. 10-13기 사이에서 aphidicolin은 초기배의 복제 주기 제어 기능을 의미하는 핵막의 파괴를 지연시킨다. 효모의 RecQ 상동유전자의 돌연변이체에서는 hydroxyurea에 대단히 민감하게 반응하는데 그 것은 S기에서 정지된 후 세포주기가 재개될 때 RecQ 상동체가 기능을 해야 가능해짐을 의미한다. 초파리 초기배아에서 충실하게 S기를 완성하는 것은 핵 DNA를 자세포에게 정확하게 전달한다는 의미에서 중요할 것이다. 이런 점에서, RECQE가 정확하고 신속한 S기의 완성에 관여하는지를 동정하는 것은 흥미 있는 일일 것이다.

정 상민
다카라코리아바이오메디칼 (주)
연구개발센터 부소장

1981-1985 건국대학교 축산대학 축산가공학과 (학사)
1989-1990 동경대학 대학원 농학계연구과 연구생 (문부성 국비장학생)
1990-1992 동경대학 대학원 농예화학 전공 (석사)
1992-1995 동경대학 대학원 응용생명공학 전공 (박사)
1995-2000 일본 이화학연구소 연구원 (Post Doc.)
2000-2001 동경대학 의학계연구과 신경내과 (Post Doc.)
2001-현재 다카라코리아바이오메디칼 (주) 연구개발센터 부소장