

## Sox 유전자의 다양성

- 총설 -

홍경원 · 김희수\*

부산대학교 생명과학부

Received July 1, 2004 / Accepted August 4, 2004

**Multiple Facets of Sox Gene.** Kyung-Won Hong and Heui-Soo Kim\*. *Division of Biological Sciences, College of Natural Sciences, Pusan National University, Busan 609-735, Korea* – Sox protein family, a transcription factor, has been found in whole animal kingdom, and contains a sequence-specific DNA binding domain called high mobility group domain (HMG). The Sox protein family based on the amino acid sequence of HMG domain was classified into 10 groups. Each group of Sox family shows significant conservation from nematode to human. The HMG domain affect to various developmental cell differentiation through binding to enhancer and regulating other transcription factors. Recently, many molecular biologists focus their research on the illustration of Sox-related disease, evolution and phylogeny. Especially, stem cell research with Sox gene family is indispensable field for understanding of their biological functions. The understanding of Sox genes may contribute to understand their role in human genetic disease and whole animal evolution.

**Key words** – Sox family, HMG, transcription factor, enhancer, evolution and phylogeny

당신의 머리에서 발끝까지 세포, 조직, 기관 들은 왜 지금의 모습인가? 계놈내의 어떤 인자들이 인간을 지금의 모습으로 만드는가? 최근 인간 발생에 대한 기작은 분자 생물학의 발달과 더불어 조금씩 밝혀지고 있다. 인간과 동물의 발생 과정에 중요한 영향을 미치는 인자들 중의 하나가 바로 이 총설에서 소개하고자 하는 Sox 패밀리에 속하는 단백질이다 [26]. Sox 패밀리는 동물계 전체에서 발견되고, 이들은 동물의 생식선 형성, 기관 발달 그리고 세포 형태 특수화와 같은 다양한 발달 과정의 조절에 관여한다 [16,28].

Sox 패밀리가 가지는 가장 보존적인 영역은 high-mobility group (HMG) 도메인이다. HMG 도메인은 79 아미노산 길이의 DNA 결합모티프로써 두 개의 서브패밀리로 나눌 수 있다. 하나는 TCF/SOX/MATA 서브패밀리며 서열 특이적 HMG 도메인 하나를 가진다. 또 다른 서브패밀리인 HMG/UBF는 서열 특이성이 조금 낮은 다중 HMG 도메인들을 가진다. TCF/SOX/MATA 서브패밀리는 DNA의 minor groove에 (A/T)(A/T)CAA(A/T)G 라는 특이적인 서열에 결합하는 경향이 있다 [23].

Sox 패밀리 중 SRY가 처음으로 포유류에서 결정되었고, 나머지 Sox패밀리에 속하는 단백질들은 Sry 단백질의 HMG 도메인과 비교하여 상대적 보존성에 기초하여 동정되었다 [21]. Sox라는 이름은 Sry HMG box의 약자에서 왔으며, 각 Sox 단백질들의 HMG 도메인과 Sry의 HMG 도메인의 아미노산 서열이 약 50% 이상의 상동성을 가지는 것으로 정의되어 왔다. 그러나 최근 상동성이 50% 이하로 낮더라도 HMG

도메인내의 특이적인 아미노산 서열인 'PMNAFMVW'를 보존하고 있는 단백질들을 Sox패밀리로 정의하는 것이 일반적이다 [4].

Sox 패밀리의 분류는 마우스, 초파리, 선충 등의 모델 생물들의 계놈이 밝혀지면서, Sox패밀리의 유전자, 유전자 단편 그리고 종내 유전자 파라로그, 종들 사이의 유전자 오소로그들이 빠르게 분류 되었다. 현재까지 약 30종의 척추동물과 12종의 무척추 동물에서 Sox 유전자와 그 단편들이 동정되었다. 인간과 마우스 계놈에서 약 20쌍의 오소로그 Sox 유전자들이 동정되었고 Sox6, Sox13을 제외하고 모두 동일한 계놈 구조를 가지고 있었다 [19].

이 Sox 유전자에 돌연변이가 발생했을 때 동물의 발생단계에서 비정상적인 표현형을 보여주었다. 예를 들어, 인간에서 Sox9 돌연변이는 골격형성 기형과 성 역전 신드롬을 야기시킨다. 또, Sry의 돌연변이는 성역전과 비정상 고환 형성의 결과를 가져왔다 [8].

이처럼 발생단계에서 중요한 기능을 하는 Sox 유전자는 생물학을 연구하는 많은 연구자에게 흥미롭고, 도전해볼 만한 연구거리를 제공한다. 본 총설에서 현재까지 보고된 Sox 유전자의 분류, 진화 및 기능 등을 고찰하여, 인간의 유전적인 질병들을 이해하고, Sox 유전자를 통한 전체 동물계의 진화 과정을 추정할 수 있는 유전적인 지표를 개발할 수 있는 기초 자료를 제공할 수 있을 것으로 생각된다.

## 본 론

## Sox 단백질의 명명

Sox 패밀리의 구성 단백질들의 이름은 다음과 같은 일반적인 규칙으로 명명되어진다. 모든 Sox 단백질들은 Sry를 제외

\*Corresponding author

Tel : +82-51-510-2259, Fax : +82-51-581-2962

E-mail : khs307@pusan.ac.kr

하고 그들의 이름 내에 Sox라는 용어를 가지고, 접두어나 접미어가 첨가된다. 인간의 단백질들은 전부 'SOX' 세 글자를 대문자로 하고, 다른 종들은 'Sox'처럼 첫 글자만 대문자로 한다. Sox 뒤의 숫자들은 Sox 단백질들이 동정된 순서에 따라서 결정된 것이다. 초기에 이 숫자들이 33까지 정해졌었지만, 최근 인간 게놈 프로젝트가 완성되면서 PCR 증폭이나 염기서열 결정의 실수로 인하여 같은 그룹이지만 다른 단백질로 동정된 것들이 제거되어 실제로는 그 숫자가 33보다 적다. 예를 들어, 원래 SOX29로 명명되었던 유전자는 SOX5와 상당한 서열 유사성을 보였다. 그러나, HMG 도메인 영역에서 결실(microdeletion)을 가지는 것이 발견되었고, 발현되는 mRNA가 존재하지 않았다. 그래서 SOX29는 SOX5의 pseudogene인 것으로 판명되어 SOX5p로 명명되었다. 다른 SOX 유전자들도 검증을 통해 최종적으로 인간에 20가지의 SOX 유전자들이 존재한다고 보고되고 있다. 한편, Sox 단백질의 이름 내에 접두어가 존재한다면, 그것은 동정된 종을 가리킨다. 이 접두어는 포유류에서는 잘 사용되지 않고, 포유류 이외의 척추동물과, 무척추동물들의 Sox이름에 많이 이용된다. 예를 들어, 접두어가 'xSox17 또는 xeSox3'인 것은 *Xenopus laevis*, 'cSox11 또는 chSox9'는 chicken 그리고 'rtSox24'는 rainbow trout을 가리킨다[20,26].

### Sox 단백질의 분류

초기 Sox 패밀리의 연구자인 Wright 등은[29] 마우스 Sox 유전자의 HMG 도메인들이 암호화하는 아미노산 서열에 기초하여 다음과 같이 6개의 그룹으로 분류하였다. 먼저 그룹 A에는 Sry; B에는 Sox-1,-2,-3 그리고 -14; C에는 Sox4, -11 그리고 -12; D에는 Sox5, -6 그리고 13; E에는 Sox8, -9 그리고 -10; 마지막으로 F에는 Sox7로 분류하였다. 이것은 이후에 Bowles 등에 [4] 의하여 그룹 A-H까지 8개의 그룹으로 재분류 되었고, 추가로 I, J 두 개의 새로운 그룹을 첨가 하였다.

Bowles 등의 분류는 현재 일반적으로 받아들여지는 Sox 패밀리의 분류이므로 여기서 좀더 자세히 설명 하고자 한다. 그들은 기존에 부분적인 HMG 도메인 서열만 사용했던 것을 전체 길이의 HMG 도메인 서열로 확장시켰다. 그 전체 서열 정렬의 결과는 Fig. 1에서 보여진다. Fig. 1에서 서열들은 이전에 정의된 그룹들과 유사하게 배열되는 것을 볼 수 있다. 또한, 서로 다른 Sox 그룹은 각 그룹의 특이적인 서열이 존재하는 것을 볼 수 있다. 예를 들어, 14-19 위치에서 그룹D는 'A(K/D)(D/E)ERR'이고 그룹 E는 'AQAARR'이며 그룹 F는 'AK(I/D)ERK'임을 보여준다. 즉, 이 결과는 척추동물 Sox 오소로그들의 HMG 도메인 서열이 매우 보존되어있다는 증거가 된다 [20]. 5-10 위치에 있는 서열 모티프인 'RPMNAF'는 Sry를 포함한 모든 Sox 서열들에서 보존되어 있지만, 아웃그룹 서열인 fu-MATA1, mo-LEF1 또는 mo-TCF1과는 많이 다른 것을 볼 수 있다. 이 보존된 서열을 좀더 확장시킨

'RPMNAFMVW'는 Sry이외의 다른 모든 Sox 단백질들에서 보존되어 있음을 알 수 있다. 이러한 특이적인 서열 보존성은 Sox 패밀리 분류에 단지 HMG 도메인의 상동성을 이용할 뿐만 아니라, 이 서열의 보존성을 분류에 이용할 수 있음을 가리킨다[4].

또한 Bowles 등은 Sox 패밀리의 각 그룹마다 HMG 도메인 내에 인트론의 위치가 다르고, 그 것이 그룹 내에서 보존되어 있음을 발견하였다. 그래서 그들은 인트론의 보존성 에 대해서 또한 점수를 부여하였다. 왜냐하면, 인트론이 서로 다른 Sox 그룹에서 같은 서열 위치에 발생했을 가능성은 극히 희박하기 때문에, 특수한 위치에서 인트론의 존재 또는 부재에 대한 점수화는 Sox의 분류와 그 Sox를 가지는 종들 간의 관련성을 유추할 수 있는 계통분류학적 유전적 지표로 사용될 수 있다[4].

Fig. 1에서 세로의 짧은 막대가 있는 부위가 게놈 서열에서 인트론이 있는 부위이다. HMG 도메인의 인트론은 척추동물의 Sox 그룹에서는 찾아지지 않았지만, 초파리와 선충의 그룹 B와 C에 속하는 Sox 단백질들의 HMG 도메인 내에 인트론을 가지는 것이 찾아졌다. 그룹 B의 인트론들은 후구동물에서는 없어졌다고 생각된다. 또한, 서로 다른 그룹에 속하는 HMG 도메인의 인트론이 그룹 D, E, F의 그룹들 사이에서 보존되어 있음을 알 수 있다. 더욱이, 그룹 D와 F에서 HMG box 인트론은 그 위치까지 두 그룹 사이에 보존되어 있었고, 그것은 두 그룹이 선조를 공유했을 가능성을 제안한다[4].

Fig. 1의 아미노산 서열 정렬의 결과를 Sox 서열들 사이의 유전적 거리에 기초하여 계통도를 분석하면 Fig. 2와 같은 결과가 보여진다. 계통도에서 대부분의 무척추동물의 서열들은 같은 그룹의 척추동물들의 서열과 같은 그룹을 형성함으로써 공통 조상을 가지는 것으로 보여진다. 그러나, 그룹 A의 Sry는 단계통 분류를 형성하지 않는 것을 보인다. Bowles 등은[3] 여기에 대해 다음과 같이 설명하고 있다. Sry는 다른 Sox 그룹들과 달리, 오소로그한 포유류의 Sry영역은 HMG 영역 바깥부분에서 아주 작은 상동성을 가지고 있는데 이는 Sry의 존재위치가 진화속도가 빠른 Y 염색체 상에 위치한 결과일 것이다. Sry의 HMG 도메인 내에서 진화속도는 다른 Sox 유전자와 비교에서도 높다. 즉, 다양한 포유류의 Sry 단백질들에서 관찰된 차이들은 HMG 도메인 이외의 영역에만 관련된 것이 아니고, 전체 단백질의 비정상적인 빠른 분기를 반영한다. 이처럼 Sry 서열들은 몇몇 상황에서 HMG 도메인 서열들 만을 사용한 Sox패밀리의 그룹화가 가지는 한계를 반영하는 것이라 여겨진다[4].

최근 포유류 Y 염색체 Sry 유전자의 진화에 대한 의문을 X에 존재하는 Sox3을 통하여 특수하게 해명했는데, maximum likelihood의 통계적인 방법을 통한 계통도를 작성하여, Sry의 출현이 그것의 포유류 기원과 일치하고, 다양한



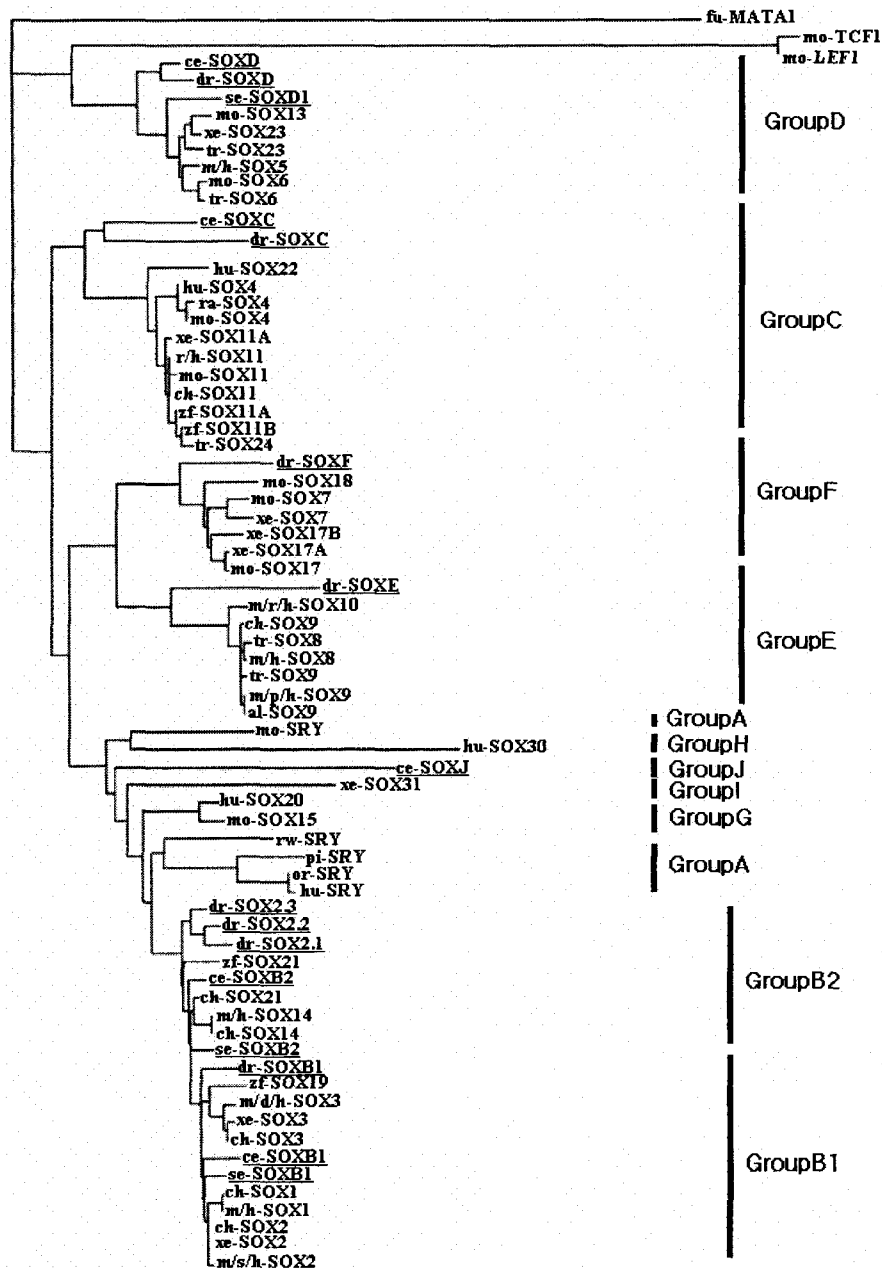


Fig. 2. Neighbor-Joining tree for the Sox HMG domain sequences. The tree was originated from Bowles *et al.* (2000). fu-MATA, mo-LEF1 and moTCF1 were used as the outgroup. Previously defined SOX family groups A-G are supported. hu-SOX30, xe-SOX31 and ce-SOXJ do not fall into any of the previously defined groups and so each designed the sole member of group H, I and J, respectively. Invertebrate sequences are underlined.

Sry 유전자들이 일정하게 단계통의 계통도를 지지하는 것을 보여주었다[13]. 그러므로 Sry가 Fig. 2에서처럼 비정상적인 분기는 상당히 높은 진화 속도와 관련이 있는 것 같다.

Bowles 등은 또한 HMG 도메인을 포함한 Sox 단백질 전체 길이에 내재하는 도메인들을 도식화하고 그룹별로 분류하였다. 여기서 보이는 포유류의 도메인 구조는 Sry를 제외하고 포유류들 사이에 상당히 잘 보존되어 있으므로, 그림에서는 마우스의 것을 대표적으로 보여 주었다. Fig. 3에서 구

조와 기능에 관련된 도메인과 보존성이 높은 영역 그리고 인트론 위치가 정확하게 표시되어있다. Fig. 3에서 Sry를 제외하고 각 그룹간의 관계가 계통도와 잘 조화되는 것을 보여준다. 이 그림에서 가장 분명한 특징은 HMG 도메인의 크기와 전체 단백질의 길이가 그룹별로 잘 보존되어 있다는 것이다. 예를 들어, leucine zipper 와 proline/glutamine rich 도메인은 그룹 D 내에 보존되어있다[4].

이러한 계통분석 결과 가장 최근에 정리된 Sox 패밀리의

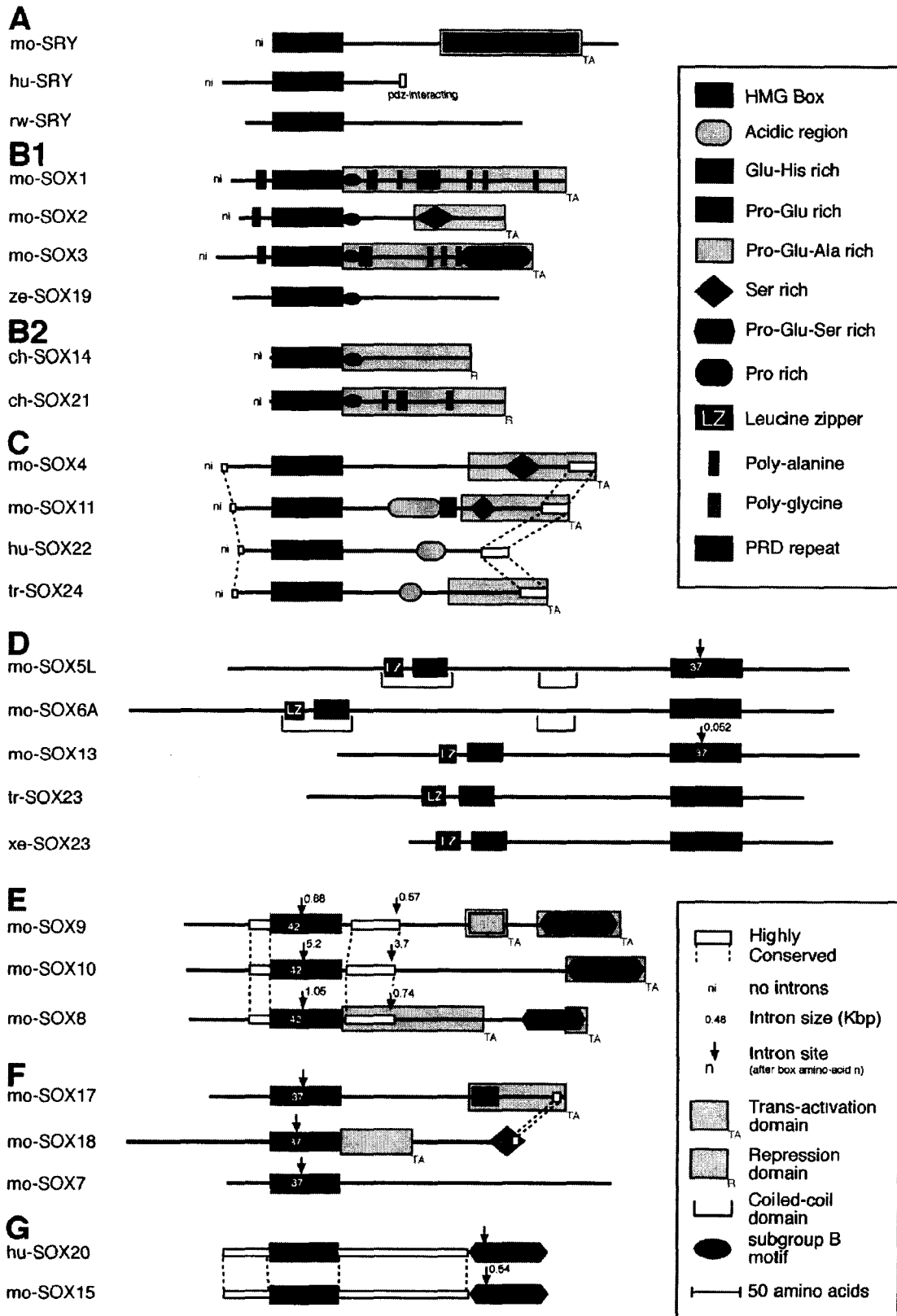


Fig. 3. Schematic representation of Sox protein domains.

분류와 명명은 Table 1에서 보여진다. 표에서 인간과 마우스 간에 오소로그한 Sox 단백질들을 분류하고, 그들의 기능과 그 유전자가 가지는 엑손의 개수들을 한눈에 알 수 있다. 여기서 특이한 것은 마우스와 인간의 Sox6과 Sox13에서 엑손의 수가 다르다는 것이다[19]. 우리는 이러한 엑손의 차이는 인간과 마우스 뿐 아니라 침팬지나 원숭이들의 엑손을 비교함으로써, 언제 새로운 엑손의 삽입이 있었는지, 또는 결실이 있었는지 연구해 볼 필요가 있다고 생각한다.

### Sox HMG의 특징

Sox 단백질의 HMG 도메인은 target 유전자의 DNA에 결합하는 도메인이다. 다른 HMG 도메인 단백질들처럼, Sox 단백질들은 특수한 DNA 구조를 인지하는 능력을 가지고, 그 서열에 결합한다. HMG 단백질들 중, 서열 특이적 DNA 인지 능력은 Sox 단백질에만 유일하다. Sox 단백질의 인식 모티프는 5'-(A/T)(A/T)CAA(A/T)G-3'이다[9].

Sox 단백질의 HMG 도메인 구조는 다른 HMG 도메인들처럼 L-모양으로 꼬여서 배열된 세 개의 알파 나선으로 구성되어 있다. 1번과 2번 나선은 역평행 구조를 길게 형성하고 나선 3과 N-말단은 L-모양의 짧은 팔 부분을 형성한다. 또한, 그 전체 구조는 소수성 아미노산으로 구성된 core에 의해 L-모양이 유지된다. 이 core를 구성하는 아미노산은 염기 특이적인 DNA 결합을 제공하는 것으로서 Sox 단백질들 사이에서 매우 보존되어 있다. HMG의 전체적인 구조는 DNA와 결합 해도 변하지 않지만, DNA는 큰 구조적인 변화를 일으킨다. 즉 DNA의 'minor groove'가 HMG 도메인의 표면에 결합하여 오목하게 된다. 결과적으로, Sry나 다른 Sox 단백질들에 의해 결합된 DNA는 전체적으로 70-85도정도 굽어지게 된다(Fig. 4.) [27].

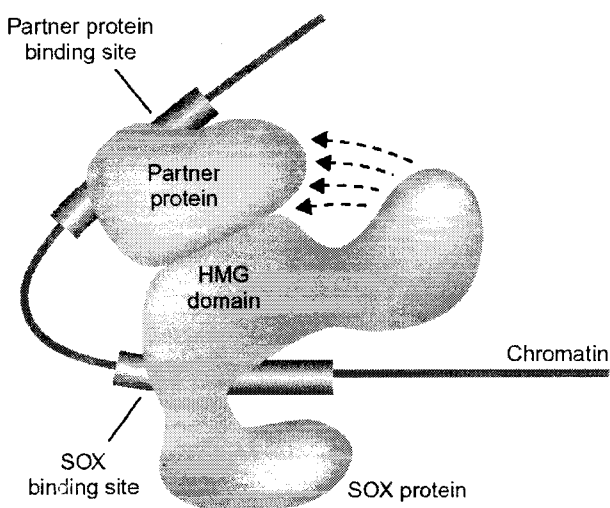


Fig. 4. A model for the specificity of Sox transcription factor action. Dotted arrows indicates that the non-HMG Sox domain may stabilize protein-protein interaction.

Sox 단백질들의 'minor groove' 결합 특징은 대부분의 다른 전사 인자들의 DNA 결합과 다르지만, 이러한 Sox 단백질의 결합은 다른 전사인자들이 DNA에 결합하기에 적합한 공간을 제공한다. DNA를 굽히는 능력은 Sox 단백질들이 크로마틴 구조를 조율하고, 다른 DNA 결합 전사 인자들이 생물학적인 활성을 가지게 조립하는 기능을 수행한다[25].

Sox 단백질들은 또한 그들 스스로 DNA에 결합할 수 없는 단백질들을 이 복합체 내로 끌어들이는데, 그것은 신호 전달 기작과 관련되어 있다. 그런 기작은 TCF/LEF (HMG 슈퍼패밀리)에 속하는 또 다른 HMG 도메인 패밀리의 단백질들과 Sox 단백질간에 연관된 유사성으로 제안된다. TCF/LEF 단백질들은 신호전달물질인 윈트 단백질(Wnt) 신호에 의해  $\beta$ -catenin에 결합하고, 그럼으로써 이 세포 외 신호를 번역하여 유전자 발현을 변화시킨다. 윈트 단백질 신호 기작에 대한 개입은 TCF/LEF/SOX 단백질들에서 공통적인 특징이지만,  $\beta$ -catenin과의 상호작용 도메인은 TCF/LEF에서만 보존되어 있다[1].

Sox 단백질들은 척추동물 발생과정에서 목표 유전자들의 발현 촉진과 억제 등 상당히 다양한 역할들을 다른 파트너 단백질들과 상호작용을 통하여 수행한다. 이 상호작용들은 HMG 도메인에 의해 종종 매개 된다는 것이 생화학적인 연구에서 검증되었다[28].

HMG 도메인은 모든 Sox 단백질들에서 유사하다. 그러므로 만일 Sox 인자들 사이에서 기능의 차이가 있다면, HMG 도메인 이외의 부분들에서 차이가 날 것이라 생각되었다 [18]. 그러나 Sox에 대한 생화학적인 연구에서 HMG 도메인이 DNA-binding 영역 절반과 protein-protein 상호작용을 위한 영역 절반으로 중첩되어 있다는 것이 밝혀졌다. 최근 연구들은 혈관 형성인자인 Sox18과 근세포에서 중요한 전사 활성 인자인 MEF2C 사이[10], 배아 신경축 인자 Sox10과 Sp1/3 사이[17], 그리고 SRY와 남성 호르몬 수용체의 Zn finger 사이에서 HMG 도메인과 파트너 단백질들 사이의 상호작용이 보고되었다[33].

HMG 도메인과 파트너 단백질 사이의 상호작용은 Sox 단백질들 외의 단백질들에서도 발견되었다. 예를 들어, Alx4라는 단백질은 전사인자인 LEF1의 HMG 도메인에 결합하고, 결합한 두 단백질 모두 모낭 발생 동안 전사활성을 위한 N-CAM 전사인자에 결합한다[2]. Sox 단백질들과는 달리, HMG 도메인 단백질들 HMG1, HMG2는 DNA 결합에 있어서 특이성은 보이지 않는다. 그러나 이 단백질들은 그들의 목표유전자들에 HMG 도메인을 매개한 상호작용을 통해 POU와 HOX 단백질들을 포함하는 다른 서열 특이적 전사인자들의 결합을 조절하는 것이 발견되었다[34]. 이러한 결과들은 protein-protein 상호작용을 위한 HMG 도메인의 기능이 보존된 특징이라는 것을 제안한다[21].

DNA에 결합 능력과 전사 인자 단백질들과 상호작용 매

Table 1. Pairing of Mouse and Human Genes

Gene	Sox Group	Major known (or Deduced) Functions	Species	Accession no.	No. of exon	Chromosomal Location
Sry	A	Testis determination	Mouse	NM_011564	1	Y(cM)
			Human	NM_003140		Yp11.3
Sox1	B1	Lens development, (neural determination)	Mouse	NM_009233	1	8(4cM)
			Human	NM_005986		13q34
Sox2	B1	Neural induction, (lens induction, pluripotency)	Mouse	NM_011443	1	3(15cM)
			Human	BC13923		3q26.3
Sox3	B1	(neural determination, lens induction)	Mouse	NM_009237	1	X(24.3cM)
			Human	NM_005634		Xq27
Sox4	C	Heart, lymphocyte, thymocyte development	Mouse	NM_009238	1	13(20cM)
			Human	NM_003107		6q22.3
Sox5	D	Chondrogenesis	Mouse	NM_011444	15	6(69.5cM)
			Human	NM_006940		12p11.1
Sox6	D	Chondrogenesis (Cardiac myogenesis)	Mouse	NM_011445	17	7(55cM)
			Human	NM_033326	16	11p15.3
Sox 7	F	(Development of vascular and many other tissues)	Mouse	NM_011446	2	14(28cM)
			Human	NM_031439		8p22
Sox8	E	(Development of many tissues)	Mouse	AF191325	3	17(8cM)
			Human	NM_014587		16p13.3
Sox9	E	Chondrogenesis, sex determination	Mouse	BC024958	3	11(69.5cM)
			Human	NM_000346		17q25
Sox10	E	Neural crest specification	Mouse	AF047043	3	15(46.5cM)
			Human	NM_006941		22q13
Sox11	C	(Neuronal, glial maturation)	Mouse	NM_009234	1	12(11cM)
			Human	NM_003108		2p25
Sox12	C	(Development of many tissues)	Mouse	BF714412	1	2(86cM)
			Human	NM_006943		20p13
Sox13	D	(Development of arterial walls, pancreatic islets)	Mouse	AB006329	13	1(70cM)
			Human	NM_005686	14	1q31
Sox14	B2	(Interneuron specification, limb development)	Mouse	AF193437	1	9(53cM)
			Human	NM_004189		3q22
Sox15	G	(Myogenesis)	Mouse	AB014474	2	11(39cM)
			Human	NM_006942		17p13
Sox17	F	Endoderm specification	Mouse	NM_011441	3	1(7cM)
			Human	NM_022454		8q11.2
Sox18	F	Vascular and hair follicle development	Mouse	NM_009236	2	2(96cM)
			Human	NM_018419		20p13.3
Sox21	B2	(CNS patterning)	Mouse	BE647677	1	14(50cM)
			Human	NM_007084		13q32
Sox30	H	(Male germ cell maturation)	Mouse	AV255326	5	11(20cM)
			Human	NM_007017		5q35

개 기능에 추가로 HMG 도메인은 핵내 물질 수송을 위한 신호들을 포함한다. 유전자 발현을 조절하기 위하여, 전사인자들은 세포질에서 핵 내로 수송 되어야 한다. 이것의 촉진 과정은 임포틴(impotin) 이라는 전사인자 단백질들에 의해 조절된다. 이 임포틴의 역할과 관련된 핵내 위치 결정 신호(Nuclear localization signal; NLS)서열들이 Sox 단백질의 HMG 도메인의 C-말단에서 동정되었다[18]. Forwood와 그의 동료들은[7] SRY의 핵내 수송기능이 HMG 도메인 C-말단의 NLS를 요구하는 임포틴 에 의해 매개된다는 것을 보였다. 이 NLS에서 돌연변이가 발생하면 성염색체를 일으키고, DNA결합이나 굽어짐의 현상이 보이지 않았다. 한 예로, 핵수송의 기작이 SOX9의 기능을 위해 필수 적이란 것을 보여주었다. 성분화 이전의 SOX9 단백질은 미분화 성선의 세포질에서 발견되지만, SRY가 발현된 이후에는 SOX9는 정소의 핵내에 위치하게 된다[James et al., 2004]. 이것은 Sry의 핵내 위치결정이 정소 분화를 위해 필수적이란 것을 보여주는 것이다. 그리고, 임포틴 는 SOX9 단백질의 HMG 도메인에 결합하는 것을 알 수 있었다. SOX 단백질의 NLS는 SOX 패밀리들 사이에서 보존되어 있고, 이것은 임포틴 에 의한 핵내 수송이 모든 Sox 단백질들에서 공통적인 특징임을 암시한다[28].

**Sox 단백질의 기능**

척추동물 종들에서 Sox 단백질들의 수는 20개 이상일 것이라고 가정되는데, 이런 많은 수가 주어진 것은 대부분의 조직과 세포 발생과정 중 최소한 한 단계 이상에서 발현되기 때문이다. 최근에 Sox 단백질의 중요한 몇 가지의 생물학적인 역할들이 규명되었다.

포유류의 성 결정에 관여하는 Sry 유전자는 Sox 패밀리의 원형이며, Y 염색체 상에 위치하는 남성 결정의 결정적인 인자이다. Sry는 생식능(genital ridge)의 세포 들에서 발현되어 Sertoli 세포와 정소의 분화를 시작하게 한다[5]. Sry의 분자수준에서 기작은 아직 분명하지는 않지만, Dax1과 SF1 같은 성결정에 관계하는 다른 단백질들과 상호작용 한다는 것이 알려져 있다[24]. SOX9는 SRY와 함께 정소발달에 중요한 기능을 하고, 뼈 발생과정에 연골 형성에 관계한다[14,31].

Sox1, Sox2, Sox3은 안구 유도와 분화의 초기 과정에 관여하고, 최근에 Kamachi 등은 [11] Sox2와 다른 전사 인자인 PAX6이 안구 특이적인  $\delta$ -crystallin 유전자의 촉진좌위(enhancer)에서 촉진의 파트너 단백질로써 함께 역할을 하는 것을 보고했다. 그들은 Sox2와 PAX6이 함께 발현되면 둘 다 촉진좌위에 결합하지만, 각각 따로 발현이 되면 SOX2나 PAX6은 혼자서는 비록 그들의 인식 서열에서 각각 결합할 수는 있지만 안정된 protein-DNA 복합체를 형성하지 못한다는 것을 발견하였다. 그러므로 안정한 Sox-promoter 결합을 위해서 이 두 단백질이 결합하는 것이 중요하다[12].

또 다른 예로써, Sox2와 Sox3이 낭배형성선의 배아에서

Fgf4 (Fibroblast growth factor 4), UTF1 (배아 줄기세포에서 발현되는 전사인자의 조인자) 그리고 뼈 단백질(osteopontin) 유전자 발현을 조절하기 위해 Oct3/4 (Octamer-binding transcription factor 3/4)와 상호작용하는 것이 보고되었다 [32]. 또 다른 유전자가 그들의 발현이 SOX/OCT 복합체에 조절된다는 것이 규명되었다. 그것은 호메오도메인 단백질 Hoxb1 (Homeodomain box b1)이며, Hoxb1의 발현이 후뇌의 제4영역에만 제한 되고, 후뇌의 분화에 중요하다. Hoxb1 과 Pbx1(Hox 조인자) 단백질들은 둘 다 동일한 촉진좌위(b1-ARE enhancer)에 결합하는 것으로 알려져 있지만, 이 단백질들은 b1-ARE가 메틸레이션에 의해 활성화 되지 않은 다른 조직에서도 발현된다. 이것을 이해하기 위한 가능한 설명은, 추가적인 인자가 Hoxb1의 후뇌에서 영역 특이적 기능을 위해 필요하다는 것이다. Di Rocco는 [5] Sox와 Oct의 DNA 결합이 b1-ARE 촉진좌위의 활성을 위해 요구된다는 것을 시험관내와 생체내 실험에서 발견하였고, SOX2/Oct1 복합체가 후뇌에서 Hoxb1/Pbx1 복합체의 전사활성을 증가시킬 수 있음을 보고하였다[6].

신경계에서 강한 발현을 보이는 또 다른 Sox 단백질은 그룹 E의 Sox10이다. Sox10은 neural crest가 생겨나는 그때에 세포들 내에서 넓게 발현된다. Sox10 발현은 neural crest 세포들이 말단 신경계 형성에 관여하고 감각 수용 ganglia 뿐만 아니라 슈반 세포 계통의 세포들에서 일반적으로 신경을 따라서 검출된다[15].

**결론 및 전망**

Sox에 대한 연구는 지난 몇 년 동안 빠르게 진전되었고, 많은 패밀리들이 동정되었다. 또한, 동물 발생과정의 조절에 복잡하게 관여하는 것으로 알려지고 있다. Sox 단백질은 HMG 도메인 영역 이외의 부분에서 돌연변이가 빠르게 일어나는 것을 보여주었고, 이러한 돌연변이의 증가는 이 그룹의 전사인자와 인간의 유전적 질병에 대한 연구를 가속화시켰다. Sox 단백질들은 마우스에서 인위적인 유전자 결실 실험을 이용해 그들의 기능이 많이 밝혀지게 되었다. 그러나, 그들의 분자적인 기작의 이해는 그들이 전사활성 인자, 조절 인자, 혹은 구조적인 구성성분으로서 어떻게 기능을 하는지를 파트너 단백질들, 신호 기작들 그리고 목표 유전자들과의 상호작용을 통해 밝힘으로써 완전해 질 것이다. 최근 많은 분자 생물학자들이 이 Sox 유전자와 관련된 질병, 진화, 그리고 계통분류 등에 많은 관심을 보이고 있지만, 아직 국내에는 많은 연구가 이루어지지 못하고 있다. 이 총설에서 현재까지 국외에서 이루어지는 Sox에 관련된 논문들을 정리하여, 국내의 연구자들이 이해하기 쉽도록 설명하였다. 아마도 이 Sox 유전자들을 이해함으로써, 인간의 유전적 질병과 인간을 포함한 전체 동물계의 진화를 이해할 수 있는 열쇠가 될 것이



라 생각한다. 또한 현재 국내에 줄기 세포(stem cell)의 연구가 왕성하게 진행되고 있는데, 바로 이 Sox 유전자야 말로 그러한 줄기 세포 연구에 꼭 필요한 유전적인 지표가 될 것이라 생각된다.

## 요 약

Sox 패밀리는 동물계 전체에서 찾아지는 전사인자이고, HMG라는 특이적인 DNA결합 도메인을 가진다. 이 Sox 패밀리는 HMG 도메인의 아미노산 서열을 바탕으로 현재 10개의 그룹으로 분류된다. 각 그룹의 오소로그한 Sox 단백질들은 선충에서 인간까지 상당한 보존성을 보인다. HMG 도메인은 전사 촉진 좌위에 결합하고 다른 전사인자들의 결합을 조절함으로써 동물 발생과정의 다양한 세포에서 발현되어 그들의 분화에 결정적인 영향을 미친다. 최근 많은 분자생물학자들이 Sox 유전자와 관련된 질병, 진화, 그리고 계통분류 등에 많은 관심을 보이고 있다. 특히, 줄기세포에서 Sox 유전자의 연구는 그들의 생물학적인 기능을 이해하기 위해 꼭 필요한 분야이다. 아마도 이 Sox 유전자들을 이해함으로써, 인간의 유전적 질병과 인간을 포함한 전체 동물계의 진화를 이해할 수 있는 열쇠가 될 것이라 생각한다.

## 참 고 문 헌

1. Bienz, M. 1998. TCF: transcriptional activator or repressor? *Curr. Opin. Cell Biol.* **10**, 366-372.
2. Boras, K., and P. A. Hamel. 2002. Alx4 Binding to LEF-1 regulates N-CAM promoter activity. *J. Biol. Chem.* **277**, 1120-1127.
3. Bowles, J., L. Cooper, J. Berkman and P. Koopman. 1999. Sry requires a CAG repeat domain for male sex determination in *Mus musculus*. *Nat. Genet.* **22**, 405-408.
4. Bowlws, J., G. Schepers and P. Koopman. 2000. Phylogeny of the Sox family of developmental transcription factors based on sequence and structural indicators. *Dev. Biol.* **227**, 239-255.
5. Collignon, J., S. Sockanathan, A. Hacker, M. Cohentannoudji, D. Norris, S. Rastan, M. Stevanovic, P. N. Goodfellow and R. Lovell-Badge. 1996. A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2. *Development* **122**, 509-520.
6. Di Rocco, G., A. Gavalas, H. Popperl, R. Krumlauf, F. Mavilio and V. Zappavigna. 2001. The recruitment of SOX/OCT complexes and the differential activity of HOXA1 and HOXB1 modulate the Hoxb1 auto-regulatory enhancer function. *J. Biol. Chem.* **276**, 20506-20515.
7. Forwood, J., V. Harley and D. Jans. 2001. The C-terminal nuclear localization signal of the sex determining region Y (SRY) high mobility group domains mediates nuclear import through importin beta 1. *J. Biol. Chem.* **276**, 46574-46582.
8. Foster, J. W., M. A. Dominguez-Steglich, S. Guioli, C. Kwok, P. A. Weller, M. Stevanovic, J. Weissenbach, S. Mansour, I. D. Young, P. N. Goodfellow, J. D. Brook and A. J. Schafer. 1994. Campomelic dysplasia and autosomal sex reversal caused by mutation in an SRY-related gene. *Nature* **372**, 525-530.
9. Harley, V. R., R. Lovell-Badge and P. N. Goodfellow. 1994. Definition of a consensus DNA binding site for SRY. *Nucleic Acids Res.* **22**, 1500-1501.
10. Hosking, B. M., S. C. M. Wang, S. L. Chen, S. Penning, P. Koopman and G. E. O. Muscat. 2001. Sox 18 directly interacts with MEF2C in endothelial cells. *Biochem. Biophys. Res. Commun.* **287**, 493-500.
11. Kamachi, Y., M. Uchikawa, J. Collignon, R. Lovell-Badge and H. Kondoh. 1998. Involvement of Sox 1, 2, and 3 in the early and subsequent molecular events of lens induction. *Development* **125**, 2521-2532.
12. Kamachi, Y., M. Uchikawa, A. Tanouchi, R. Sekido, and H. Kondoh. 2001. Pax6 and Sox2 form a co-DNA-binding partner complex that regulates initiation of lens development. *Genes. Dev.* **15**, 1272-1282.
13. Katoh, K. and T. Miyata. 1999. A heuristic approach of maximum likelihood method for inferring phylogenetic tree and an application to the mammalian Sox-3 origin of the testis-determining gene SRY. *FEBS Lett.* **464**, 129-132.
14. Kent, J., S. C. Wheatley, J. E. Andrews, A. H. Sinclair and P. Koopman. 1996. A male-specific role for SOX9 in vertebrate sex determination. *Development* **122**, 2813-2822.
15. Kuhlbrodt, K., B. Herbarth, E. Sock, I. Hermans-Borgmeyer and M. Wegner. 1998. Sox10, a novel transcriptional modulator in glial cells. *J. Neurosci.* **18**, 237-250.
16. Lefebvre, V., P. Li and B. de Crombrughe. 1998. A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. *EMBO J.* **17**, 5718-5733.
17. Melnikova, I., H. Lin, A. Blanchette and P. Gardner. 2000. Synergistic transcriptional activation by Sox 10 and Sp1 family members. *Neuropharmacology* **39**, 2615-2623.
18. Poulat, F., F. Girard, M. P. Chevron, C. Goze, X. Rebillard, B. Calas, N. Lamb, and P. Berta. 1995. Nuclear localization of the testis determining gene product SRY. *J. Cell Biol.* **272**, 27848-27852.
19. Schepers, G. E., R. D. Teasdale and P. Koopman. 2002. Twenty pairs of Sox: Extent, Homology, and Nomenclature of the mouse and human Sox transcription factor gene families. *Dev. Cell* **3**, 167-170.
20. Schepers, G., M. Bullejos, B. Hoskings and P. Koopman. 2000. Cloning and characterization of the Sry-related transcription factor gene, Sox8. *Nucleic Acid Res.* **28**, 1473-1480.
21. Sinclair, A. H., P. Berta, M. S. Palmer, J. R. Hawkins, B. L. Griffiths, M. J. Smith, J. W. Foster, A. M. Frischauf, R. Lovell-Badge and P. N. Goodfellow. 1990. A gene from human sex determining region encode a protein with homology to a conserved DNA-binding motif. *Nature* **346**, 240-244.
22. Smith, J. M. and P. A. Koopman. 2004. The ins and outs

- of transcriptional control: nucleocytoplasmic shuttling in development and disease. *TRENDS genet.* **20**, 4-8.
23. Soullier, S., P. Jay, F. Poulat, J.-M. Vanacker, P. Berta, and V. Laudet. 1999. Diversification pattern of the HMG and Sox family members during evolution. *J. Mol. Evol.* **48**, 517-527.
  24. Swain, A., V. Narvaez, P. Burgoyne, G. Camerino and R. Lovell-Badge. 1998. Dax1 antagonizes Sry action in mammalian sex determination. *Nature* **391**, 761-767.
  25. Uwanogho, D., M. Rex, E. J. Cartwright, G. Pearl, C. Healy, P. J. Scotting, and P. T. Sharpe. 1995. Embryonic expression of the chicken Sox2, Sox3 and Sox11 gene suggests an inactivative role in neuronal development. *Mech. Dev.* **49**, 23-36.
  26. Werner, M. H. and S. K. Burley. 1997. architectural transcription factors: proteins that remodel DNA. *Cell* **88**, 733-736.
  27. Wegner, M. H. 1999. From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res.* **27**, 1409-1420.
  28. Werner, M. H., J. R. Huth, A. M. Gronenborn and G. M. Clore. 1995. Molecular basis of human 46X, Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. *Cell* **81**, 705-714.
  29. Wilson, M., and P. Koopman. 2002. Matching Sox: partner proteins and co-factors of the SOX family of transcriptional regulators. *Curr. Opin. Genet. Dev.* **12**, 441-446.
  30. Wright, E. M., B. Snopek, and P. Koopman. 1993. Seven new members of the Sox gene family expressed during mouse development. *Nucleic Acid Res.* **21**, 744.
  31. Wright, E., M. R. Hargrave, J. Christiansen, L. Cooper, J. Kun, T. Evans, U. Gangadharan, A. Greenfield and P. Koopman. 1995. The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. *Nature Genet.* **9**, 15-20.
  32. Yuan, H., N. Corbi, C. Basilico and L. Dailey. 1995. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Gene Dev.* **9**, 2635-2645.
  33. Yuan, X., M. Lu, T. Li, and S. Balk. 2001. SRY interacts with and negatively regulates androgen receptor transcriptional activity. *J. Biol. Chem.* **276**, 46647-46654.
  34. Zwilling, S., H. Konig and T. Wirth. 1995. High mobility group protein 2 functionally interacts with the POU domains of octamer transcription factors. *EMBO J.* **14**, 1198-1208.